

**CROSS CULTURAL ANALYSIS OF EMOTIONS ON SOCIAL
MEDIA BRANDING COMMUNICATION WITH EVIDENCE FROM
BIG DATA**

by
ALTUĞ TANALTAY

Submitted to the Sabancı Graduate Business School
in partial fulfilment of
the requirements for the degree of Doctor of Philosophy

Sabancı University
July 2023



ALTUĞ TANALTAY 2023 ©

All Rights Reserved

ABSTRACT

CROSS CULTURAL ANALYSIS OF EMOTIONS ON SOCIAL MEDIA BRANDING COMMUNICATION WITH EVIDENCE FROM BIG DATA

ALTUĞ TANALTAY

Ph.D in MAMANEGEMT Ph.D DISSERTATION, July 2023

Dissertation Supervisor: Prof. Dr. NİHAT KASAP

Dissertation Co-Supervisor: Prof. Dr. SELCEN ÖZTÜRKCAN

Keywords: Emotion classification, consumer emotion, social media marketing,
Twitter, brand post popularity, big data, global, local, emerging market

Localization efforts by global brands are frequently reflected in their social media strategies, with differing degrees of adaptation on the tone, emotion, and symbols employed. Moreover, emotions play a vital role in the consumption experience, influencing word-of-mouth communications and brand loyalty. The similarities and variations in the usage of emotional content by parent brands and their local counterparts and understanding their effect on branding communication constitute an emerging field of research that can uncover the valuable insights for better business. Besides quantifying cross-cultural differences in the preferences and use of emojis as nonverbal cues regarding brand communication context among two communities at different levels of social collectivism: English and Turkish speaking Twitter users, the aim of this research is twofold. First, a language, context and time aware framework for automatically extracting basic emotions of happiness, surprise, sadness, anger, fear and disgust from multinational brands' social media messages is proposed along with the evaluation of statistical and state-of-the-art deep learning methodologies on emotion classification domain. Second, the impact of emotions, as well as the most prominent structural and textual features on the popularity of brands' social media messages in a cross-cultural manner is addressed using an extensive data sample of Twitter messages collected for 33 brands from FMCG, Fast Food, Technology, Automotive, Apparel, Retail, Finance, and Logistics industries over a time window of three years. In order to mitigate the issues associated with the violation of OLS

regression assumptions, a random subspaces regression approach utilizing bootstrap resampling, which enables the quantification of feature importance and the assessment of their effects. Contributing to marketing, machine learning, social media marketing, and emotion literature, the findings are valuable for internationalizing and born-global firms to engage emotionally with their global and emerging market consumers in their social media marketing campaigns.



ÖZET

BÜYÜK VERİDEN ELDE EDİLEN KANITLARLA SOSYAL MEDYA MARKA İLETİŞİMİNDE DUYGULARIN KÜLTÜRLER ARASI ANALIZI

ALTUĞ TANALTAY

İŞLETME DOKTORA PROGRAMI DOKTORA TEZİ, TEMMUZ 2023

Tez Danışmanı: Prof. Dr. NİHAT KASAP

Tez Eş Danışmanı: Prof. Dr. SELCEN ÖZTÜRKCAN

Anahtar Kelimeler: Duygu sınıflandırması, tüketici duygusu, sosyal medya pazarlaması, Twitter, marka gönderi popülerliği, büyük veri, küresel, yerel, gelişmekte olan pazar.

Küresel markaların yerelleştirme çabaları, kullanılan üslup, duygu ve semboller üzerinde farklı derecelerde uyum sağlayarak sıklıkla sosyal medya stratejilerine yansır. Dahası, duygular, ağızdan ağza iletişimi ve marka sadakatini etkileyerek tüketim deneyiminde hayati bir rol oynar. Ana markalar ve yerel muadilleri tarafından duygusal içerik kullanımındaki benzerlikler ve farklılıklar ve bunların marka iletişimi üzerindeki etkisinin anlaşılması, daha iyi iş pratikleri için değerli içgörüler ortaya çıkarabilen, gelişmekte olan bir araştırma alanı oluşturmaktadır. Bu araştırmanın amacı farklı sosyal kolektivizm düzeylerine sahip iki topluluk olan İngilizce ve Türkçe konuşan Twitter kullanıcıları arasındaki marka iletişimi bağlamında "emoji" karakterlerinin tercihleriyle ilgili kültürlerarası farklılıkları nicelendirmenin yanı sıra iki yönlüdür. İlk olarak, çok uluslu markaların sosyal medya mesajlarından mutluluk, şaşkınlık, üzüntü, öfke, korku ve iğrenme gibi temel duyguların otomatik olarak çıkarılması için dil, bağlam ve zaman farkındalığına sahip bir sistem önerilmektedir. Bu sistem ayrıca duygusal sınıflandırma alanında istatistiksel ve en son derin öğrenme metodolojilerinin değerlendirilmesiyle birlikte sunulmaktadır. İkinci olarak, duyguların yanı sıra markaların sosyal medya mesajlarının popülerliği üzerindeki önemli yapısal ve metinsel özelliklerin etkisi, FMCG, Hızlı Yiyecek, Teknoloji, Otomotiv, Giyim, Perakende, Finans ve Lojistik sektörlerinden 33 marka için üç yıllık bir zaman diliminde toplanan Twitter mesajlarından oluşan geniş bir veri örneği

kullanılarak k lt rlerarası karşılařtırılmalı olarak ele alınmıřtır. En k   k kareler y ntemi ile geliřtirilen dođrusal regresyon varsayımlarının ihlalinden kaynaklanan sorunları hafifletmek i in, deđiřkenlerin  nemini nicelendirmeye ve etkilerini deđerlendirmeye olanak tanıyan bootstrap  rnekleme kullanarak rastgele alt uzaylar regresyon yaklařımı benimsenmiřtir. Bu  alıřmadan elde edilen bulgular, pazarlama, makine  đrenimi, sosyal medya pazarlaması ve duygu literat r ne katkı sađlamaktadır. Ayrıca, uluslararasılařan ve dođuřtan k resel olan řirketler i in, bu bulgular, sosyal medya pazarlama kampanyalarında k resel ve geliřmekte olan pazar t keticileriyle duygusal olarak etkileřime girmek i in deđerli bir rehber niteliđi tařımaktadır.



ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my academic advisors, Prof. Nihat Kasap and Prof. Selcen Öztürkcan, for their invaluable guidance, mentorship, and expertise. Their profound knowledge, critical insights, and continuous encouragement have played a pivotal role in shaping my research and enhancing my academic growth. I am truly fortunate to have had the opportunity to work under their supervision, and I am truly appreciative of their unwavering support.

I would also like to extend my sincere appreciation to the members of my Dissertation Committee, Enes Eryarsoy and Kemal Kılıç. Their insightful feedback, constructive criticism, and scholarly contributions have greatly enriched the quality of my research. I am truly grateful for their time, effort, and dedication in guiding me throughout this process.

Lastly, I would like to acknowledge the support and encouragement I have received from my mom and dad. Their belief in me and constant support have been crucial in my academic journey. I am ever grateful for their love, support, and understanding.

To all those mentioned above and countless others who have been a part of my academic journey, I extend my heartfelt appreciation. Their contributions, guidance, and support have been invaluable, and I am forever grateful for the impact you have had on my personal and professional growth.



*To my dear wife,
for her unwavering support and patience.*

TABLE OF CONTENTS

LIST OF TABLES	xiii
LIST OF FIGURES	xiv
LIST OF EQUATIONS	xvi
LIST OF ALGORITHMS	xvii
LIST OF ABBREVIATIONS	xviii
1. INTRODUCTION	1
1.1. Brand Post Popularity	2
1.2. Emotional Appeal in Brand Social Media Messages	4
1.3. Challenges in Extracting Emotions from Social Media Data	5
1.4. Cultural Influence on Emotional Responses to Brand Posts	7
1.5. Research Objective	9
1.6. Outline of the Dissertation	14
2. ANALYSIS OF EMOJI UTILIZATION IN SOCIAL MEDIA BRANDING COMMUNICATION ACROSS ENGLISH AND TURKISH LANGUAGES	16
2.1. Literature Review	19
2.1.1. Research Themes on Emojis	19
2.1.2. Emojis in Sentiment Analysis	20
2.1.3. Emojis in Cultural Studies	22
2.1.4. Emojis in Business Research	23
2.2. Methodology	24
2.2.1. Twitter Data and Text Pre-Processing	25
2.2.2. Attitude Toward Using Emojis in Twitter Messages across Languages (RQ-1)	28
2.2.3. Preference of Emojis across Languages (RQ-2)	30
2.2.3.1. Diversity of emojis preferred across languages (RQ-2a)	31
2.2.3.2. Similarity of popular emojis across languages (RQ2-b)	33

2.2.4.	Emotional Representation of Emojis across Languages (RQ-3)	36
2.3.	Experimentation and Results	38
2.3.1.	Attitude Toward Using Emojis in Twitter Messages across Languages (RQ-1)	38
2.3.2.	Preference of emojis across languages (RQ-2)	41
2.3.2.1.	Diversity of emojis preferred across languages (RQ-2a)	41
2.3.2.2.	Similarity of popular emojis across languages (RQ-2b)	42
2.3.3.	Emotional Representation of Emojis across Languages (RQ-3)	44
2.4.	Conclusion	47
3.	EMOTION EXTRACTION FROM SOCIAL MEDIA MESSAGES IN BRANDING CONTEXT ACROSS TURKISH AND ENGLISH LANGUAGES	51
3.1.	Literature Review	53
3.1.1.	Emotion Theories	54
3.1.2.	Emotion Extraction	55
3.1.2.1.	Benchmark data sources	55
3.1.2.2.	Text pre-processing & feature extraction	57
3.1.2.3.	Emotion classification models	59
3.2.	Methodology	62
3.2.1.	Generating the Auto Annotated Data Set	63
3.2.2.	Text Pre-Processing	67
3.2.3.	Emotion Classification Models	68
3.2.4.	Performance Metrics	70
3.3.	Experimentation and Results	71
3.4.	Conclusion	76
4.	UNDERSTANDING EMOTIONS AND BRAND POST POPULARITY IN SOCIAL MEDIA COMMUNICATION ACROSS ENGLISH AND TURKISH LANGUAGES	79
4.1.	Literature Review	82
4.1.1.	Brand Post Popularity in Social Media	82
4.1.2.	Conceptual Framework	85
4.1.3.	Random Subspace Regression with Bootstrap for Feature Selection	88
4.2.	Methodology	91
4.2.1.	Data	91
4.2.2.	Operationalization of Variables	93
4.2.3.	Models	95

4.2.4. Random Subspace Regression Using Bootstrap Method	98
4.3. Experimentation and Results	101
4.4. Conclusion	107
5. CONCLUSION	110
5.1. Discussion on Key Findings	110
5.2. Theoretical and Methodological Implications	113
5.3. Managerial Implications	115
5.4. Limitations and Future Research Directions	116
5.5. Concluding Remarks	119
BIBLIOGRAPHY	121



LIST OF TABLES

Table 1.1.	Multinational Brands Selected for Twitter Data Collection . . .	11
Table 1.2.	Research objectives and corresponding chapters	14
Table 2.1.	Number of Twitter Messages Collected by Industry	27
Table 2.2.	Summary Statistics: Number of Emojis in a Message & Pro- portion of Messages with Emojis	28
Table 2.3.	Parameter Estimates of the Evaluated Models	39
Table 2.4.	Number of preferred emojis and similarity	41
Table 2.5.	Diversity statistics	42
Table 2.6.	Dissimilarity (JD) across popularity levels	43
Table 2.7.	Number of preferred emojis and similarity	46
Table 3.1.	The most important 10 emojis by each emotion category	65
Table 3.2.	Number of selected emojis for each emotion category	66
Table 3.3.	Selected emojis for auto annotation	66
Table 3.4.	Number of messages in auto annotated Turkish and English data sets.	67
Table 3.5.	Accuracy and F-Scores of evaluated models	72
Table 3.6.	Precision, Recall and F Scores of Models	74
Table 3.7.	Training epochs and performance scores of deep learning models	75
Table 4.1.	Total number of posts, followers and interactions by industry .	92
Table 4.2.	Quantitative representation of features by antecedent categories	94

LIST OF FIGURES

Figure 1.1. Conceptual Framework of Brand Post Popularity proposed by De Vries, Gensler & Leeflang, (De Vries et al., 2012), page 84	4
Figure 1.2. The Influence of Cultural Individualism-Collectivism on Com- munication Behavior by Gudykunst, Matsumoto, Ting-Toomey, Nishida, Kim & S., (Gudykunst et al., 1996), page 512	7
Figure 1.3. Cross cultural perspective to model brand post popularity by Lin, Swarna & Bruning, (Lin et al., 2017) page 625.	9
Figure 1.4. Hofstede’s National Scores of Cultural Dimensions, (Hofstede, 2011)	10
Figure 1.5. Research Design of the Dissertation by Objectives and Chapters	12
Figure 2.1. Steps of Analysis for Research Questions	25
Figure 2.2. Distribution of number of emojis in messages	29
Figure 2.3. Distribution of emoji appearance frequencies.	34
Figure 2.4. Log transformed emoji frequencies ordered by log transformed ranks.	34
Figure 2.5. Segmentation of Log transformed emoji frequencies by log transformed emoji ranks	35
Figure 2.6. Diversity of emojis preferred.	42
Figure 2.7. Emoji popularity phases by frequencies of emoji groups pre- ferred.	43
Figure 2.8. Dissimilarity of emoji popularity levels by time, 2016/01 – 2021/07	44
Figure 2.9. Emotional distribution of emojis and the distance of their distributions across languages.	45
Figure 2.10. Distribution of distances by sentiments.	46
Figure 3.1. Research design	62
Figure 3.2. Data accumulation steps	64
Figure 3.3. Deep neural net architectures used in experiments	70
Figure 3.4. Accuracy scores of evaluated models	73
Figure 3.5. Accuracy scores by epoch on training and test data sets	75

Figure 4.1. Conceptual Framework	88
Figure 4.2. Research Design to Evaluate the Importance and Effects of Post Features on Brand Post Popularity	91
Figure 4.3. Distribution of Interactions by Language	93
Figure 4.4. Comparison of Regression Models Evaluated on Alternative Distributions of Interactions	97
Figure 4.5. Adjusted R^2 and RMSE given added features in regression model	101
Figure 4.6. RMSE over bootstrap cycles with increasing number of iterations	102
Figure 4.7. Average Importance Scores by Feature/Variable Category . . .	103
Figure 4.8. Random Subspace Bootstrap Estimates	106



LIST OF EQUATIONS

Equation 2.1.	31
Equation 2.2.	32
Equation 2.5.	37
Equation 2.6.	37
Equation 3.1.	64
Equation 3.2.	70
Equation 3.3.	71
Equation 3.4.	71
Equation 3.5.	71
Equation 4.1.	95
Equation 4.2.	96
Equation 4.3.	96
Equation 4.4.	96
Equation 4.5.	99

LIST OF ALGORITHMS

1. Estimation of $TTRM_i$ by Monte Carlo Permutation Test 32
2. Feature Importance and Significance by Random Subspace Regression
and Bootstrap 100



LIST OF ABBREVIATIONS

Acc Accuracy.....	70
BERT Bidirectional Encoder Representations from Transformers.....	53, 58
BIC Bayesian Information Criterion.....	30
BOW Bag of Words.....	57
CNB Complement Naïve Bayes.....	68
CNN Convolutional Neural Net.....	56, 59
CVC Cultural Variability in Communication.....	7
F F-Score.....	70, 73
FF Fully Connected Feed Forward Networks.....	68
GRU Gated Recurrent Unit.....	58
HAN Hierarchical Attention Network.....	56
JD Jaccard Distance.....	36
JS Jensen-Shannon Divergence.....	37
KNN K-Nearest Neighbors.....	59
LR Linear Regression.....	29
LSTM Long Short-Term Memory Networks.....	58, 68
MBERT Multilingual BERT.....	69, 70
ML Machine Learning.....	20
MNB Multinomial Naïve Bayes.....	59, 61, 68
NB Negative Binomial.....	30, 95, 96, 97, 98

NLP Natural Language Processing	11, 16, 52
OLS Ordinary Least Squares	3, 29, 80, 83, 84, 87, 95, 96
P Precision	70, 73
POS Part of Speech Tagging	57
PR Poisson Regression	29
R Recall	70, 73
RF Random Forest	59, 68
RMSE Root Mean Squared Error	xv, 101, 102
RNN Recurrent Neural Net	53, 56
RQ Research Question	9, 10, 17
RR Rate Ratio	29, 30
SA Sentiment Analysis	16, 20, 52, 57
SL Sentiment Lexicon	20
SVM Support Vector Machines	59, 61, 68
TEL Turkish Emotion Lexicon	56, 61
TF-IDF Term Frequency – Inverse Document Frequency	57, 61, 67
TTR Type-Token Ratio	31
VIF Variance Inflation Factor	87, 88, 98
ZINB Zero Inflated Negative Binomial	30, 39
ZIP Zero Inflated Poisson	30, 39

1. INTRODUCTION

In the era of globalization, brands are increasingly expanding their presence in foreign markets and adapting their marketing strategies to cater to diverse cultural contexts. This shift is driven by the recognition that local audiences have unique preferences, values, and behaviors that must be considered for effective brand communication. As a result, social media platforms have emerged as prominent channels for global brands to engage with local audiences, allowing them to tailor their messaging and content to specific markets.

One key aspect of this adaptation process is reflected in the tone, emotive content, and symbolic representations used by brands on social media platforms. Different cultures may respond differently to various communication styles, and brands need to strike a balance between maintaining their global identity and appealing to the local sensibilities. By analyzing and understanding the varying degrees of adaptation employed by global brands, researchers can gain insights into how brands navigate cultural boundaries and connect with their target audiences effectively.

Moreover, social media platforms have evolved into influential marketing tools for brands to promote their products and services in shaping consumer attitudes (Lee & Kahle, 2016). Consumers not only passively consume brand content but actively engage with it, leading to the development of attitudes and relationships. These interactions on social media platforms can foster deep-rooted emotional connections between consumers and brands (Hudson, Roth, Madden & Hudson, 2015). Such emotional connections play a pivotal role in driving positive word-of-mouth and brand recommendations (Edelman, 2010).

Given the significance of consumer engagement and emotional connections on social media, it becomes crucial to explore the similarities and differences between parent brands and their local versions in their use of contentual and emotional features along with the level of consumer engagement in social media messages. This emerging field of research offers an opportunity to uncover valuable lessons for better business practices. By examining the strategies employed by successful global

brands in adapting their social media communication to local markets, researchers can provide insights that can benefit other businesses in their pursuit of effective global marketing.

This dissertation examines cross-cultural differences in the effect of emotional appeals on social media brand engagements. Comprehensive investigation of this topic necessitates the adoption an interdisciplinary approach from a wide range of research areas to address the multifaceted and complex aspects of the issue at hand. Thus, before delving into the research objectives, not only the concept of "Brand Post Popularity", referring to the level of engagement and attention a brand's social media posts receive from its target audience, but also the challenges in extracting the emotional appeal from social media messages must be discussed. Moreover, a cultural point of view is required as the goal is also evaluating the impact of emotions and other impacting features on regional markets

1.1 Brand Post Popularity

Assessed by various metrics, including likes, comments, shares, retweets, and views, depending on the platform, it is a study of various factors such as vividness, interactivity, the content of the brand social media posts. The popularity of a brand post indicates its resonance with the audience and the effectiveness of the brand's communication strategy. It signifies the extent to which the post captures the attention and interest of the target market, as well as the degree to which it evokes positive responses and engagement. A highly popular brand post tends to generate more user interactions, such as comments and shares, which in turn can amplify the post's reach and influence.

Measuring brand post popularity provides brands with valuable insights into the effectiveness of their social media communication efforts. It helps them understand which types of content, messaging styles, and creative approaches resonate most with their audience. By analyzing the factors contributing to post popularity, brands can refine their social media strategies, optimize content creation, and tailor their messages to maximize engagement and audience connection.

First proposed by De Vries et al., earlier work on the effectiveness of banner advertising was inspired to model the effects of features on the popularity of brand messages. Following De Vries et al., these features are commonly categorized as

vividness, interactivity, practicality, information, entertainment and emotional appeal:

Vividness refers to what extent a post stimulates users' visual senses (De Vries et al., 2012; Schultz, 2017) with content such as images or videos.

Interactivity of a post is related to what extent it makes the consumers take an action or be able to influence the communication with hashtags, links, polls or questions.

Practicality of a post refers to its accessibility and efficiency of delivering the brand message represented by length of the textual content, time and day of the message.

Information in posts is related to what extent the content gives details about facts about alternative topics about a product or service.

Entertainment is related with how the brand post reflects fun, humour or excitement with its content.

Emotional appeal represents which feelings are targeted to be stimulated by the post, and its effect was also investigated as antecedents to popularity of brand posts.

Beginning with De Vries et al. analyzing Facebook brand pages, a set of factors are postulated within this research area, offering alternative conceptual frameworks to explore. These factors are operationalized by considering the content embedded in messages, and their effects are commonly inferred through the use of Ordinary Least Squares (OLS) regression. The conceptual framework proposed by (De Vries et al., 2012), which is the base for following studies in literature is shown in Figure 1.1. For vividness images and videos, regarding interactivity hashtags, links and polls were used for operationalization. Information and entertaining content were manually extracted from the textual content. Position was the duration for a post at the top position of the brand page and valence of comments was the aggregate sentiment of the users towards the brand message. Brand post popularity was measured in number of likes and comments, along with day of the message, message length and product category controlled. The effect and the significance of the variables were hypothesized and tested using ordinary least squares linear regression. Studies following De Vries et al., used alternative features and categories such as integration of practicality and emotional features.

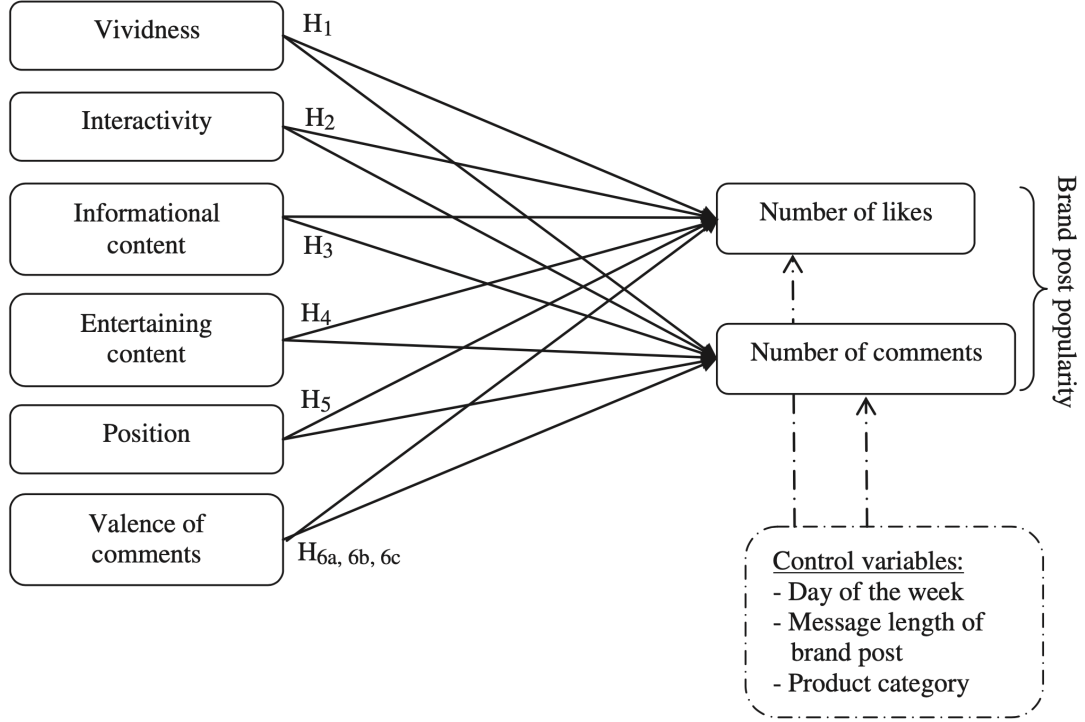


Figure 1.1 Conceptual Framework of Brand Post Popularity proposed by De Vries et al., (De Vries et al., 2012), page 84

While features related with vividness, interactivity, practicality, information or entertainment have been widely studied by previous work on Brand Post Popularity, as a result of certain linguistic and computational challenges regarding extracting the emotional appeal from textual content, only a limited number of studies have analyzed their effects (Karpinska-Krakowiak & Modlinski, 2020; Pinto & Yagnik, 2017; Swani, Milne & P. Brown, 2013).

1.2 Emotional Appeal in Brand Social Media Messages

The impact of emotional appeal on brand post popularity remains an area that requires further research attention within the realm of social media content. Emotions, defined as positive or negative experiences associated with specific patterns of physiological activities (Pinker, 1999; Schacter, Gilbert, Wegner & Hood, 2011), play a crucial role in the consumption experience (Laverie, Kleine & Kleine, 1993) and have the potential to influence word-of-mouth communications and brand loyalty (Crosby & Johnson, 2007). By analyzing customers' responses to products and

services, it becomes possible to identify areas for improvement or enhance communication by leveraging prior knowledge of their emotions, thereby increasing customer satisfaction. Extracting emotions from the target population aids in comprehending the impact of products, services, or policies and managing consumer reactions through emotion monitoring.

In the realm of brand social media interactions, only a handful of studies have focused on the incorporation of emotional sentiments in brand posts, suggesting that doing so leads to increased user attention and interactions (Swani et al., 2013). Conveying common emotional appeals such as happiness, love, motivation, and adventure in social media messages has been found to contribute to the popularity of brand posts (Pinto & Yagnik, 2017). Brands can therefore tailor the emotional content of their social media messages to elicit more positive responses from their audience. Additionally, monitoring the collective emotions of consumers can help prevent the spread of negative word-of-mouth about a product or service.

Understanding the role of emotions in brand post popularity not only provides insights into effective communication strategies but also offers opportunities for managing and influencing consumer perceptions and behaviors. By incorporating emotional appeals and monitoring the emotional responses of consumers, brands can better engage with their audience and foster positive brand experiences. Thus, investigating the relationship between emotional appeal and brand post popularity in the context of social media represents a vital area of research that can contribute to improving marketing practices and enhancing customer satisfaction. However, prior to measuring the effect of emotional appeal in post engagements one has to extract emotions from textual content which is a challenging task as a result of linguistic and computational challenges.

1.3 Challenges in Extracting Emotions from Social Media Data

Vivid or interactive characteristics of a brand post are represented by the structural content and extraction of these features is generally considered as a straightforward task as it simply requires text processing and counting the target elements such as hashtags, mentions, images, or videos, while on the other hand extracting emotions from social media data is more challenging. *Firstly*, due to the inherent complexity of human emotions and the diverse ways they are expressed through language, and

secondly, due to the vast amount of textual data generated and shared on social platforms. Regarding the first challenge, emotions are subjective experiences that involve a wide range of nuances, making it difficult to accurately capture and interpret them solely from textual content. The *second* challenge arise from the close relation of emotion extraction and big data. While the continuous stream of textual content generated and shared on social media platforms presents a rich source of information that can be leveraged for understanding human emotions at scale, challenges associated with the volume, velocity, and variety of data, as well as the need for advanced computational infrastructure, data storage, and analysis techniques make it harder to extract meaningful insights (Behera, Jena, Rath & Misra, 2021; Zucco, Calabrese, Agapito, Guzzi & Cannataro, 2020).

To tackle these challenges regarding emotion extraction, researchers have employed various approaches, such as lexicon-based analysis and machine learning algorithms. Lexicon-based approaches involve using predefined emotion lexicons or dictionaries that assign emotional labels to words based on their semantic and contextual associations. Machine learning techniques, on the other hand, rely on training models with labeled emotional data to automatically classify and extract emotions from text. These approaches leverage the vast amount of data available online, allowing for the development of more sophisticated models that can capture the subtleties and complexities of human emotions expressed through text. However, despite the advancements in these techniques, the inherent subjectivity and context-dependency of emotions continue to present ongoing challenges in accurately extracting and understanding emotions from text, thus, requiring further attention of research with contextual and cultural perspectives on the issue.

Besides its challenging nature, emotion extraction from text requires a significant amount of labeled data to train models effectively. Labeled data consists of text samples with annotations indicating the corresponding emotions expressed. However, acquiring large-scale labeled data specifically for emotion extraction is a challenging and time-consuming task. While there are abundant resources and labeled data available for emotion extraction in the English language, the same cannot be said for other languages. This language bias creates a limitation in the generalizability and applicability of emotion extraction techniques across different linguistic contexts and cultural settings as individualism or collectivism. Additionally, even in English, the available resources and labeled data for emotion extraction are still limited compared to the vast amount of text available on social media platforms and the internet as a whole. This scarcity of resources and limited research in emotion extraction hinders the development of robust and accurate models for understanding emotions in text across various languages and domains. Further efforts are needed to address

these limitations and promote research in the field of emotion extraction to enhance its applicability and effectiveness in diverse linguistic and cultural contexts.

1.4 Cultural Influence on Emotional Responses to Brand Posts

The effectiveness of aligning vividness, interactivity, and emotions in social media messages on the level of consumer engagement with brand posts is well-documented in existing literature (De Vries et al., 2012; Karpinska-Krakowiak & Modlinski, 2020; Schultz, 2017). However, it is important to recognize that the impact of these practices may vary across different regional markets, highlighting the need for a cultural perspective in business research (Lin et al., 2017).

National culture, as described by William Gudykunst's model of Cultural Variability in Communication (CVC), plays a significant role in shaping communication behaviors through cultural norms and the socialization processes of individuals within their respective cultures (Gudykunst, 1997; Gudykunst et al., 1996). The level of cultural individualism-collectivism within a nation directly influences communication styles and indirectly affects behavior through self-construals and values (Gudykunst, 2005) (see Figure 1.2).

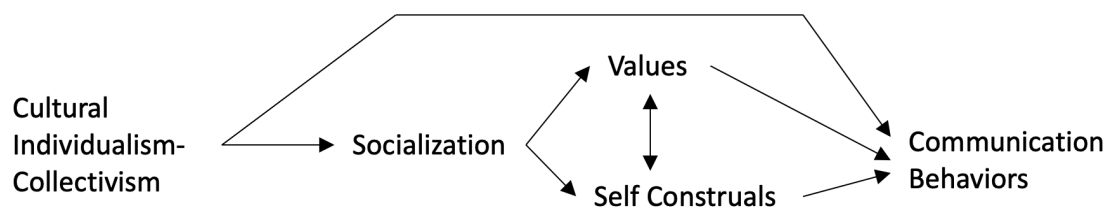


Figure 1.2 The Influence of Cultural Individualism-Collectivism on Communication Behavior by Gudykunst et al., (Gudykunst et al., 1996), page 512

Previous research has adapted Hofstede's national scores of cultural dimensions to examine the impact of individualism/collectivism characteristics on consumer behavior across different regional markets. For instance, studies have shown that Chinese (collectivist) consumers engage more in online information sharing activities compared to consumers from the United States (individualistic) (Chu & Choi, 2011). Similarly, consumers from individualistic cultures, such as Australia and the United Kingdom, are less likely to use social media for making purchasing de-

cisions compared to consumers from collectivist cultures (Goodrich & De Mooij, 2014). Considering these cultural characteristics and adapting social media practices accordingly can increase the effectiveness of brand campaigns by generating more positive interactions from targeted consumers. By understanding the cultural nuances and preferences of specific regional markets, brands can tailor their social media messages to resonate with local consumers, thereby enhancing engagement and fostering stronger connections.

Although numerous studies exist analyzing the vivid, interactive and emotional features effecting popularity of brands' social media in more individualistic - English - cultural market contexts, a research gap regarding their possible variation in culturally more collectivist markets such as Turkish still persists to exist. (Lin et al., 2017) proposed that social media practices might not yield the same effectiveness across diverse regional markets. Shown in Figure 1.3, Their model illustrates how social media practices presented to individuals with diverse cultural perspectives can be interpreted in accordance with those perspectives. Although the authors did not empirically evaluate this hypothesis, they drew insights from earlier research to suggest strategies for managing the popularity of brand posts from a cultural standpoint. Their discussion highlighted the relevance of Hofstede cultural dimensions theory, specifically individualism/collectivism, uncertainty avoidance, power distance, and long-term orientation (Hofstede, Hofstede & Minkov, 2010), in adapting social media practices for brands across different cultures. This adaptation entails tailoring the levels of vividness, interactivity, practicality, informativeness, and entertaining content to suit the preferences of the regional audience. Their recommendations were based on previous findings that revealed varying attitudes of consumers from different cultural backgrounds.

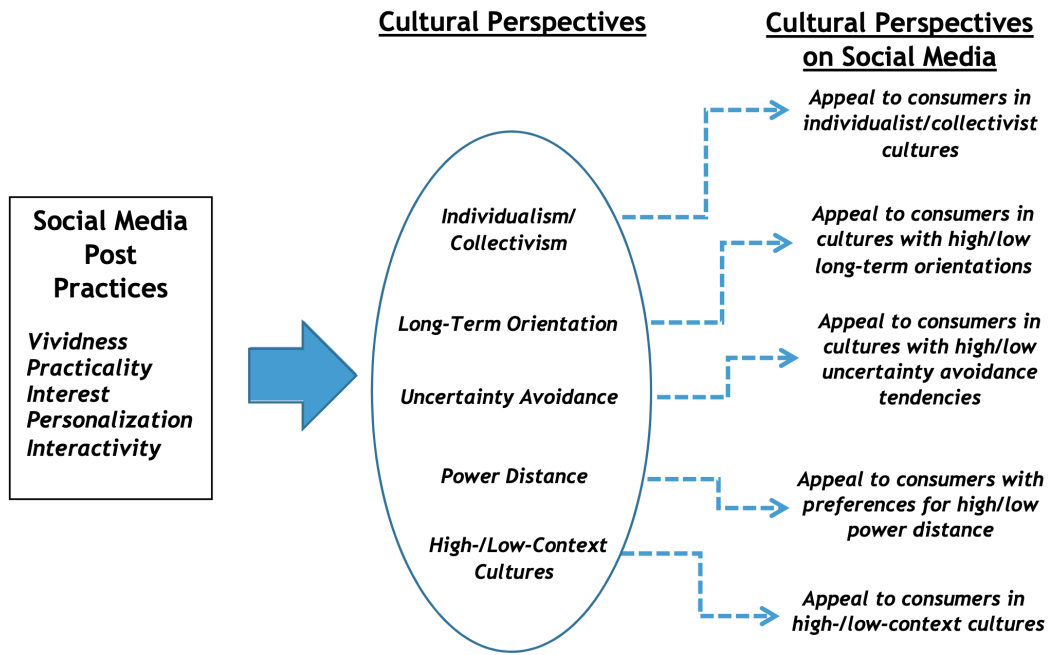


Figure 1.3 Cross cultural perspective to model brand post popularity by Lin et al., (Lin et al., 2017) page 625.

In summary, incorporating a cultural perspective in the adaptation of social media practices for brands across global and regional markets is crucial. By recognizing and leveraging the influence of cultural characteristics, brands can maximize the effectiveness of their campaigns, resulting in more positive interactions and engagement from their target consumers. This highlights the importance of considering cultural dimensions in research and practice to ensure the successful implementation of social media marketing strategies in diverse cultural contexts.

1.5 Research Objective

The focus of this dissertation is the examination of cross-cultural differences concerning the impact of emotions in social media branding communication. As mentioned, the issue must be approached from an interdisciplinary set of research fields related with brand post popularity, machine learning and information theory. Thus, approaching different aspects, the rest of this dissertation seeks answer to the following main Research Question:

RQ: What is the impact of emotions on brand communication in social media across

cultures/languages?

As shown in Figure 1.4, according to Hofstede’s national scores of cultural dimensions English speaking countries such as Australia, United Kingdom and United States are scored significantly higher than Turkish speaking Turkey regarding individualism dimension (*89-91 for English speaking countries, 37 for Turkey*). Thus, in order to address the empirical gap in cultural perspective in earlier studies in the context of branding communication, 8,101,034 English and 735,459 Turkish Twitter messages from 2016 to 2021, containing all posts and user replies of 33 multinational brand accounts from various industries are collected for analysis. These are Fortune 500 brands of 2021 actively operating separate Twitter accounts in English, representing markets with more individualistic cultures, and in Turkish, representing markets with less individualistic, more collectivist cultures, also inline with Hofstede’s national scores of cultural dimensions. The selected brands and their industries are shown in Table 1.1.

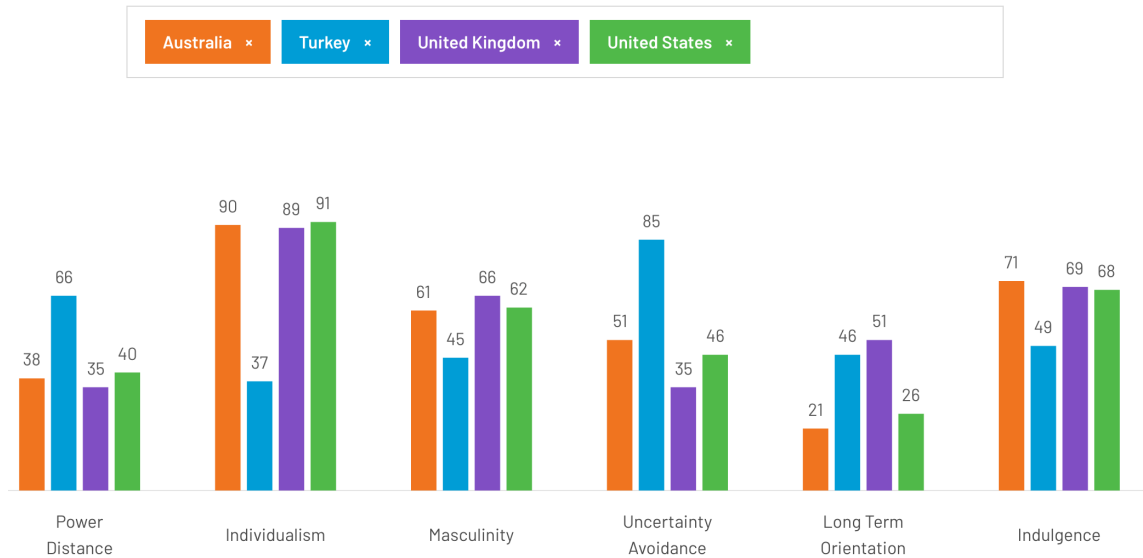


Figure 1.4 Hofstede’s National Scores of Cultural Dimensions, (Hofstede, 2011)

Visualization generated by "www.hofstede-insights.com"

Given the research focus, main RQ and culturally diverse data of multinational brands’ Twitter messages and consumer replies, the final research objectives of this dissertation are twofold. *Firstly*, the study aims to propose a flexible method for annotating textual data with six basic emotions (happiness, surprise, sadness, anger, disgust, fear) and neutrality by utilizing the emotional signals conveyed by emoji characters in text for automatically extracting emotions from social media messages. This method seeks to generate annotated data from social media messages in any language and context, specifically addressing the scarcity of annotated data sets in

Table 1.1 Multinational Brands Selected for Twitter Data Collection

FMCG	Fast Food	Technology	Apparel	Automotive	Finance	Logistics	Retail
Gillette	Burger King	Vodafone	Levi's	Hyundai	Allianz	UPS	Watsons
Coca Cola	KFC	Samsung	Marks. Spen.	Toyota	Axa	DHL	Carrefour
L'Oreal	McDonald's	Siemens	H&M	BMW	HSBC		
Algida/Walls		Sony					
Orkid/ Alldays		Microsoft					
Pepsi		HP					
Flormar		Huawei					
Oriflame		Netflix					
Yves Rocher							

languages other than English, such as Turkish. The study also evaluates the performance of alternative Natural Language Processing (NLP) models, including statistical and state-of-the-art deep learning methodologies, to extract emotions from the annotated data in Turkish and English. The goal is to collect contextual data sets for different languages and explore the most suitable classifier algorithm, ultimately proposing a framework for automatically extracting emotions from textual data. *Secondly*, the dissertation delves into the dynamics of brand post popularity by investigating the multifaceted elements that drive consumer engagement and shape the success of brand communication efforts by examining the effects of vividness, practicality, interactivity, information, entertainment features, as well as discrete emotions (happiness, surprise, anger, fear, disgust, and sadness) across English and Turkish data sets.

The overall process and research design for accomplishing the two final objectives, with regards to the subsequent chapters they are covered are shown in Figure 1.5. As the product of first research objective, an emotion classification model trained on English and Turkish Twitter data is used to extract discrete emotions from social media messages of brands in the Feature Extraction phase of the second objective. Processing meta information of Twitter posts, along with the predicted emotions, more features regarding the vividness, interactivity, practicality, informative and entertaining content of brand messages are extracted as factors. In the Predictions phase of Objective 2, following conceptual models proposed by previous literature on Brand Post Popularity, the importance and the effects of these factors are evaluated using bootstrap Monte Carlo simulations on alternative regression models predicting the sum of shares, likes and comments as the dependent variable.

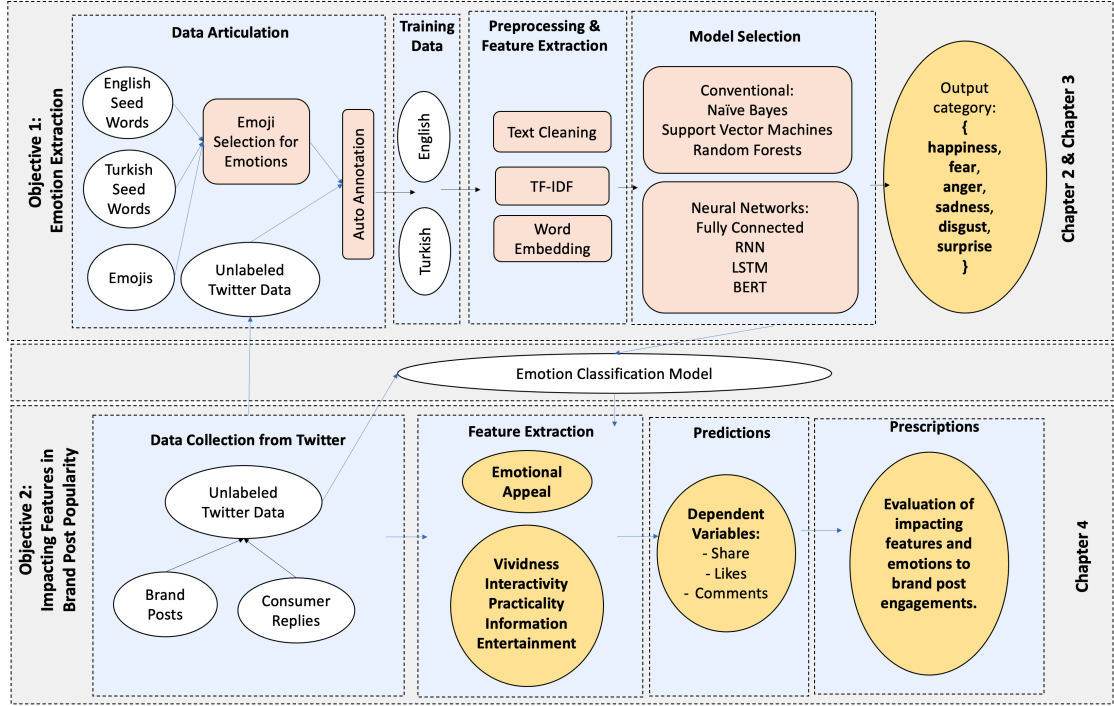


Figure 1.5 Research Design of the Dissertation by Objectives and Chapters
(Author's own representation)

The study provides valuable insights into understanding the significant variations in consumer preferences between local and global markets regarding emotions and branding communication in social media. It makes notable contributions to the fields of marketing communication and social media marketing, while also suggesting alternative approaches to methodological analysis that leverage big data, machine learning and information theory. The key findings and contributions of the study can be summarized as follows:

- 1) The study advances the understanding of brand and consumer behavior in the marketing context by examining the variations in emoji usage and their emotional representation across English speaking and Turkish speaking cultures using the most up-to-date textual big data collected over a wide time range. Findings show that besides being language and culture dependent, people's preferences and use of emojis are also dependent on the context and time they are used. Moreover, emojis representing negative emotions are more similarly whereas emojis representing positive emotions are less similarly used across cultures.
- 2) A flexible approach using emotional signals generated by emojis and emotional words is proposed to articulate textual data annotated with six basic emotions (happiness, surprise, sadness, anger, disgust, fear) and neutrality. The proposed approach is shown to be useful following the evaluation of state-

of-the-art emotion classification models on the generated data and achieving higher performance when compared to previous work on emotion extraction. While the focus is on brand social media communication context, the same approach can be applied to collect and annotate social media messages from various communication contexts such as sports, politics, or entertainment in any language.

- 3) In addition to assessing the importance of features related with vividness, practicality, interactivity, information, and entertainment aspects of brand Twitter posts across English and Turkish speaking cultures, this study incorporates the emotional content of messages as an additional determinant of brand post popularity. It represents the first research endeavor to examine the impact of these features and emotions conveyed by brands in their social media posts on the effectiveness of brand messages, particularly from a cross cultural perspective.

This study offers significant managerial contributions, particularly in the realm of branding communication. First, it introduces an end-to-end framework for automatic emotion extraction, which can be applied across various contexts including customer relationship management, brand management, and public opinion mining. Furthermore, proposed approach and the findings of cross cultural analysis on the impacting features and their effects to post engagements will allow brands to optimize their social media strategies and enhance the popularity of their brand posts across regional markets. By considering factors such as vividness, practicality, interactivity, entertainment, information, and emotional appeal managers can create more engaging and impactful brand posts that effectively resonate with their target audience. Understanding the dynamics and combined influence of these factors will enable social media campaigns to drive higher levels of engagement, increase brand visibility, and achieve their marketing objectives.

Following the overarching research purpose of investigating the relationship between emotional content in social media messages and brand post consumer engagements, while also considering the cultural contexts in which these interactions occur, the research aims to explore the impact of emotional appeal on brand post popularity, examine the challenges in extracting emotions from social media data, address the scarcity of labeled data and language bias in emotion extraction, and analyze the cultural influence of emotions on responses to brand posts. By studying these aspects, the research seeks to provide insights for global brands in adapting their marketing strategies and enhancing their engagement with local audiences on social media platforms.

1.6 Outline of the Dissertation

The research objectives along with the corresponding topic and the chapters covered are shown in Table 1.2. Given the research objectives described, the rest of this dissertation is organized as follows. In Chapter 2, using distributional measures and information-theory methodologies, cross-cultural differences in the preferences and use of emojis as nonverbal cues regarding brand communication context among English-speaking and Turkish-speaking Twitter users are analyzed. Emojis are shown to be useful in predicting sentiments and emotions in textual data, thus, the study described in Chapter 2 puts forward the need of approaching emoji characters differently across different cultural contexts to be able to further use them to understand the emotions conveyed in text.

Table 1.2 Research objectives and corresponding chapters

<i>Research Objective</i>	<i>Topic</i>	
<i>Automatically extracting emotions from social media messages</i>	Analysis of emojis utilization in social media	Chapter 2
	Emotion Extraction from social media messages	Chapter 3
<i>The effects of structural and emotional features on social media engagements</i>	Dynamics of brand post popularity.	Chapter 4

Following an extensive background of research on emotion classification from the perspectives of psychology and machine learning in Chapter 3, a framework for articulation of annotated textual data and emotion extraction on any language or context regardless of time is proposed. Focusing specifically on Turkish and English languages, using emoji characters to curate a data set of social media messages annotated by six discrete emotions of happiness, surprise, sadness, anger, fear and disgust in any language and context is described. Moreover, the performance state-of-the-art emotion classification models are evaluated on the articulated data for predicting the emotional appeal in brands' Twitter posts, further to be analyzed as a feature in the study detailed in Chapter 4.

While the studies in Chapters 2 and 3 serve the first objective of this dissertation, the study described in Chapter 4 accomplishes the second research objective. In

Chapter 4, underlying mechanisms of brand post popularity regarding factors related with vividness, practicality, interactivity, information, entertainment, and emotional appeal are explored together with the cultural differences across English and Turkish speaking markets.

Finally, in Chapter 5, a summary and discussion of findings from all chapters with concluding remarks and possible future research directions are given.



2. ANALYSIS OF EMOJI UTILIZATION IN SOCIAL MEDIA BRANDING COMMUNICATION ACROSS ENGLISH AND TURKISH LANGUAGES

The opportunity of being able to transmit one’s message to crowds and receiving immediate feedback in social media platforms attracted the attention of both brands and users. Communicating through micro-blogging platforms like Twitter consumers may interact with the brands directly, praising or complaining about products and services publicly, while brands are able to take immediate action. As messages are typically limited to a specific amount of characters in these platforms, it has become common to express emotions on these platforms using emoticons (Dresner & Herring, 2010). Abbreviated from “Emotional Icon”, emoticons’ usage got gradually replaced by emojis (Umashanthi & Jacob, 2015). As graphical symbols integrated into browser and mobile device keyboards, emojis have become non-verbal cues to convey facial expressions and a variety of other situations in textual communication. They are used to express basic emotions such as happiness or sadness as well as complex feelings and attitudes such as sarcasm, boredom, and friendship (Chang, 2016; Coyle & Carmichael, 2019). Their usage in text messaging positively affect consumers’ perceived enjoyment, usefulness, playfulness, and social connectedness, thereby promoting word-of-mouth (Hsieh & Tseng, 2017; Huang, Yen & Zhang, 2008). Analyzing their extracted emotional meaning is essential to comprehending the effective mechanisms involved in conveying the intended message’s delivery to result in perception in the desired manner.

The emergence of social media data accumulation and advancements in computational technology have led to the utility of Sentiment Analysis (SA) in gauging consumers’ emotional attitudes towards products and services. SA involves employing Natural Language Processing (NLP) techniques and algorithms to analyze and categorize text sentiment as positive, negative, or neutral. However, achieving reliable SA outcomes necessitates extensive annotated data, which is both expensive and complex to obtain. In response to this challenge, researchers in SA have focused on the emotional representations of emojis, leveraging theories from disciplines such

as Information Theory and Natural Language Processing. Consequently, there is a growing interest in exploring the influence of emojis on communication within the business context (Manganari, 2021). The selection and perception of emojis vary depending on the context in which they are used (Gawne & G., 2019). Therefore, given the ambiguity of their meaning in different contexts, it is imperative to understand how consumers interpret emojis in order to elucidate their application in consumer research (Jaeger, Roigard, Jin, Vidal & Ares, 2019). Moreover, emojis' potential to effectively represent consumers' emotions in social media advertising activities establishes them as valuable tools for measuring consumer emotions (Bai, Dan, Mu & Yang, 2019; Li, Lin & Chiu, 2014). Consequently, it becomes necessary to examine the cultural perspectives surrounding the usage and emotional portrayal of emojis in brand communication on Twitter.

From the cultural perspective, people learn how to express their feelings and values from their shared culture, so that culture influence both verbal and nonverbal communication (Gudykunst, 1997). Measured by the level of cultural individualism – collectivism, the ways people communicate in their societies vary across cultures (Gudykunst et al., 1996) and is considered to be as an influential factor on language and nonverbal cues (Hofstede et al., 2010). In individualistic cultures the emphasis is on the self and feelings are expressed more explicitly, whereas in collectivist cultures personal feelings are suppressed and people tend to use more subtle cues to forward their feelings to others (Nisbett & Masuda, 2003). There is a growing body of literature on cultural differences in the preference for emojis in text communication (Das, Wiener & Kareklas, 2019; Kavanagh, 2010; Lee & Hsieh, 2019; Takahashi, Oishi & Masaki, 2017; Yuki, Maddux & Masuda, 2007). However, mainly comparing patterns in cultures with obvious differences like USA and Japan to represent individualistic versus collectivist cultures majority of these studies use experimental design with limited external validity (Cook, Campbell & Shadish, 2002). There are also studies based on real life observational data (Kejriwal, Wang, Li & Wang, 2021; Li, Chng, Chong & See, 2019; Park, Barash, Fink & Cha, 2014), but their analysis are at aggregate level, not considering the effect of time on emoji usage patterns and the context of the communication. There is a gap in the literature for further research that focuses on how individuals use emojis, particularly in branding communication and marketing context across cultures those are in between regarding levels of cultural individualism - collectivism as Turkish community.

This study aims to quantify cross-cultural differences in the preferences and use of emojis as nonverbal cues regarding brand communication context among English-speaking and Turkish-speaking Twitter users. Using distributional measures and information-theory methodologies, following Research Questions (RQ's) are ad-

dressed:

RQ-1: *How different is the attitude toward using emojis in messages across cultures/languages?*

RQ-2: *How different are emojis preferred across cultures/languages?*

RQ-2a: *Does one culture use a more diverse set of emojis?*

RQ-2b: *How much different is the set of popular emojis across cultures/languages?*

RQ-3: *How different are emojis preferred semantically by basic emotions (Happiness, Surprise, Sadness, Anger, Fear, Disgust)?*

Insights from this study are useful for comprehending the significant differences in consumer preferences between local and global markets. Contributing to the fields of marketing communication and social media marketing, alternative approaches to methodological analysis that capture the inherent benefits of big data and information theory are highlighted:

- 1) As the first contribution, this study is the first one analyzing the cross-cultural differences of emoji usage and preferences processing most up-to-date textual big data collected over the widest time range, also considering their variation by time.
- 2) To compare variations in emoji usage and preferences, distributional measurements and information theory are integrated, as an interdisciplinary approach to marketing research.
- 3) As the final contribution, theory on brand and consumer behavior observing data specifically from marketing context, considering the differences in emojis' semantic emotional representation across cultures are intended to be advanced.

The rest of the chapter is organized as follows. In the Literature Review section, following presentation of research themes regarding the functional and perceptual sides of emojis, previous work related with their implications in sentiment analysis, cultural studies, and business research is discussed. In the Material & Methods section, following the details of data sample of this study, the methodology approaching each research question is discussed under a separate subsection. Findings regarding each research question are reported in "Experimentation and Results" section. In the "Discussion" section, the results are evaluated, and finally in "Conclusion" section

concluding remarks are made and future research directions are discussed.

2.1 Literature Review

Emojis were first introduced as emoticons in the 1990s by arranging several textual characters to resemble faces. It was proposed by American computer scientist Scott E. Fahlman that ":" and ":(" indicate humor and seriousness, respectively, in computer-mediated communication (Krohn, 2004). Later, at the start of the 2010s, emoticons were standardized as icons and dubbed "emoji characters" as part of Unicode 6.0 (Davis & Holbrook, 2021). Following the standardization of emoji characters, they were integrated into a vast array of mobile devices and browsers, which increased consumer interest in using them in social media messaging (Evans, 2017). Besides representing facial emotional expressions, emojis lately come in different flavors to indicate a variety of states and situations to symbolize travel, places, flags, animals, nature, or activities. The research community has also concentrated on emoji analytics in line with the widespread use of emojis by consumers, and interesting studies from a variety of scientific fields have come to light.

2.1.1 Research Themes on Emojis

Researchers' interest in emoji research has steadily grown since 2010 and especially after 2017 (Bai et al., 2019), following emojis' inclusion in mobile device keyboards and customers' desire in using them for digital communication (Evans, 2017). Since then, Emoji research was applied under various areas of computer science, marketing, behavioral science, linguistics, psychology, medicine and education (Bai et al., 2019). According to (Guntuku, Li, Tay & Ungar, 2019), prior research on emojis can be divided primarily under three themes: *“(1) studying emojis as a source of sentiment annotation, (2) analyzing differences in emoji perception based on rendering, and (3) understanding similarities and differences in emoji usage across different populations”*.

The first theme consists of studies from understanding the functional side of emojis to extracting emotions from various sources of text, like online blogs or messaging

platforms. They simply try to explore emojis’ emotional power as nonverbal cues to be further used in sentiment analysis. The second theme is majorly about how the emojis are perceived and viewed by users; this time trying to understand the emotional perception by the reader of the text. The final theme concentrates on cultural differences regarding the usage and perception of emojis under different contexts and cultures. While research on the second theme mainly depends on experiments and surveys with a group of participants being orthogonal to this work, the first and the third themes are the ones aimed to be contributed. Research on the unique effects of emojis on business-related communication is also gaining popularity (Manganari, 2021) . Therefore, it is equally crucial to discuss how this study pertains to the context of business research. More information on the contexts and themes, how they connect with this work, and contributions are provided in the following subsections.

2.1.2 Emojis in Sentiment Analysis

Research on emojis related with SA tasks majorly concentrated on their emotional representation using theory from various areas as Information Theory and Natural Language Processing. These research provided algorithms and models mainly based on two techniques, Machine Learning (ML) and Sentiment Lexicon (SL) (Bai et al., 2019).

Regarding ML technique, using emoticons for annotating text messages with their sentiments and then training supervised classifiers for sentiment analysis goes back to (Pang, Lee & Vaithyanathan, 2002; Read, 2005). Further on, (Liu, Li & Guo, 2012) as an example for Supervised Learning, proposed an emoticon smoothed language model trained on human annotated text data. More recently, (Redmond, Salesi & Cosma, Redmond et al.) showed that emoji outperform emoticons for training supervised text classifiers. (Li, Chng, Chong & See, 2018) proposed an effective sentiment model trained automatically on annotated Twitter data using an emoji approach. They showed that selected features performs well in classifying tweets into five different sentiment classes when compared to experimental results in related works.

Regarding SL technique, several lexicons have been proposed by researchers. Lexicons are simply dictionaries mapping words to sentiment scores, or to specific emotions. The emotion of texts is automatically determined by analysts and researchers utilizing the ML technique based on the cumulative polarity of the word content.

The first and the most well-known lexicon is (Novak, Smailović, Sluban & Mozetič, 2015b)’s “Emoji Sentiment Ranking”, where engaging 83 human annotators to labeled over 1.6 million tweets in 13 European languages. By having labeled the tweets with their sentiment polarity as positive, neutral and negative, they calculated the sentiment map of 751 most common emojis. Their results showed that most emojis contain positive sentiment, and there were no significant differences in the emoji rankings between the 13 languages and accessible publicly (Novak, Smailović, Sluban & Mozetič, 2015a).

Flexible approaches were proposed by (Tang, Qin, Liu & Li, 2013) and more recently by (Asghar, Khan, Bibi, Kundi & Ahmad, 2017), using a similar methodology to this work. Both of these works used public emotion lexicons of words to extract the emotional representation of emojis in vector space. In (Tang et al., 2013), authors first calculated the probability distribution of emojis according to their co-occurrences with emotional words acquired from public lexicons, and then ranked emojis according to their frequency and probability of representing an emotion. They selected top ranked emojis for each emotion category, and created a pseudo-annotated dataset to further train a belief network for sentiment classification. (Asghar et al., 2017) used a rule based model with extended lexicon support and an enhanced model of emotion signals like emotion words, slang and emojis. They achieved significantly better results in emotion classification task compared with baseline methods.

In more general terms, all of these examples may be regarded as applications of Domain Adaptation, considering a setting in which the training and testing data are sampled from different distributions (Glorot, Bordes & Bengio, 2011). Information in one domain (the training data), for this study word lexicons scored for emotion categories, is transferred to make predictions in another domain (testing data), for this study the emotional distribution of emoji characters in Twitter branding communication. By utilizing methods for modeling emojis in vector space and ranking emojis based on their co-occurrences with emotive terms gathered from previously created lexicons, this work complements earlier work. In contrast to the research described, the aim of this study is to identify the distinctions between eastern and western consumers’ attitude towards using emojis in branding rather than sentiment analysis. The adaption of a language other than English, which predominated in the earlier research, is another significant difference.

2.1.3 Emojis in Cultural Studies

The studies on cultural differences mainly focused on comparing geographic regions, languages, countries or groups in within a single country. Observing differences in frequency of usage, types and interpretation of emoji it was shown that emoji usage and their perception varied according to the identity of the users, thus, cultural background, language and user groups those also function as a micro cultural environment (Bai et al., 2019; Park et al., 2014). For example, it was shown that males and females perceived the function of emoji similarly (Herring & Dainas, 2018), whereas males used a more diverse set of emoji (Tossell, Kortum, Shepard, Barg-Walkow, Rahmati & Zhong, 2012), and females used higher number of emojis with positive sentiment (Prada, Rodrigues, Garrido, Lopes, Cavaleiro & Gaspar, 2018). Regarding age groups, overall frequency of emojis was not effected by age differences, but types of emojis preferred varied significantly (Prada et al., 2018; Xiang, Chen, Sun, Cheng & Bove, 2016).

Previous studies comparing collectivist and individualistic cultural traits between the east and west regarding emoji usage have produced interesting findings. Defined in numerous ways (Green, Deschamps & Paez, 2005), individualistic cultures may be referred as societies where people's decisions and actions are associated with independence and self-reliance. Whereas people in collectivist cultures depend more on social harmony, and interdependence with others (Oyserman, Coon & Kemmelmeier, 2002). Individualist features were typically used by scholars to define people from Western contexts (Western Europe, North America, Australia), whereas collectivist traits were typically used to describe non-Western (Asian, South American, and African) personalities (Green et al., 2005). (Park, Baek & Cha, 2014) analyzed 1.7 billion tweets from 2006 to 2019. By using Hofstede's national culture scores and indicators (Hofstede, 2011) across 78 countries, they concluded that people from individualistic cultures favored using horizontal and mouth oriented emoticons as ":", while people from more collectivist cultures favored vertical and eye-oriented emoticons as " ^ " more. Later, (Guntuku et al., 2019) following a similar study, comparing eastern cultures (represented by China and Japan) and western ones (USA, UK, Canada), revealed that recognizable normative and culture specific patterns existed for semantic differences in usage of emojis, however, frequency of usage was similar across east and west. Furthermore in (Li et al., 2019), researchers used Hofstede's culture index and emoji sentiment scores proposed by (Novak et al., 2015a,1), in order to examine the correlation between emoji usage and culture. Their results showed that preference of emojis to represent emotions as non-verbal cues in text communication complied with Hofstede's Cultural Dimensions Model.

This study is complementary to (Park et al., 2014,1) and (Li et al., 2019), in terms of how big data oriented approaches can be used to evaluate differences in cross-cultural communication, using Twitter as data source. In addition to earlier studies on cultural differences, there are examples that are similar to this study where researchers focused on variations in emoji usage across two or more languages or countries. An extensive study was done by (Kejriwal et al., 2021), analyzing 4 million tweets from 30 countries in a 29-day corpus. Their focus was how the prevalence, popularity and usage of emoticons varied across the selected countries. By observing the differences in the frequencies of unique emojis selected between users in different countries, and modelling the popularity of emojis as a 3-phase piecewise power-law distribution, they hypothesized that some emojis were both universal and globally popular across countries and languages. By comparing the similarities of usages using Kullback-Leibler (KL) Divergence, the researchers determined that countries, notably those in the Middle East, were more differentiated than others. They did this by calculating the probability distribution of emoji usage in each country in their sample. This work is inspired by theirs, in terms of analyzing popularity and usage of emojis. However, it differs being more focused on branding context as emojis do not usually represent a meaning on their own, but they are context dependent and their meanings must be considered in the scope of text and language they are used (Vandergriff, 2013). Moreover, emojis' usage patterns are analyzed in more detail considering the emotions they carry and the variation by time, focusing on two languages. Comparing Western usage patterns represented by English and Eastern usage patterns by Turkish, it is aimed to understand differences between a more individualistic culture and another relatively less individualistic but also less collectivists than countries like China or Japan.

2.1.4 Emojis in Business Research

Although the majority of studies on the use and perception of emojis concentrate on personal contexts, there is rising interest in how they affect communication in business context (Manganari, 2021). Emojis are selected and perceived differently according to the context they are used (Gawne & G., 2019). Thus, given the ambiguity in their meaning under variable contexts, understanding of how consumers interpret them is necessary to reveal their use regarding consumer research (Jaeger et al., 2019).

Emojis' potential to represent consumers' emotions in social media advertising ac-

tivities (Li et al., 2014) make them an effective tool to measure consumers' emotions (Bai et al., 2019). As shown by (Kim & Gupta, 2012), as a result of the two experiments on 129 and 138 student participants, negative emojis in online reviews lowered the perceived reviewer rationality and consequently decreased the informative value of the review, negatively impacting the product evaluations. In (Vidal, Ares & Jaeger, 2016), 12,260 tweets about breakfast, lunch, snack and dinner eating situations were analyzed and manually coded by researchers and it was found that emoticon and emoji use fit the emotional content of tweets. Moreover, their results suggested that emojis were an intuitive way to express emotions in food-related context. While age, gender or frequency of using them does not affect the emotional stimuli of the consumers (Jaeger, Xia, Lee, Hunter, Beresford & Ares, 2018), emoji positively effects purchase intentions (Das et al., 2019). In (Das et al., 2019), using a laboratory experiment and online study, authors aimed to understand when it was most effective to include an emoji in an advertisement communication. While their findings showed that emojis were effective at increasing positive affect and purchase intentions for hedonic products, primarily designed and consumed for the purpose of pleasure, enjoyment, or emotional satisfaction, other research found negative effects of using them in advertisement (Glikson, Cheshin & Kleef, 2018; Li et al., 2019). This work complements earlier research by examining the similarities and differences in usage and emotional depiction of emojis in brands' Twitter communication from a cultural viewpoint.

2.2 Methodology

Analysis of the emoji usage was carried out in three sections considering the differences and similarities across languages and detailed as the following sub-sections for both discussing the methodology and reporting the results of this work. The steps in the flow of analysis is shown in Figure 2.1, and the following subsections are organized for each research questions accordingly regarding "Material & Methods" and "Experimentation and Results" subsections. Following the description of the data collected and how it was processed, in sub-section 2.2.2, analysis of attitude of consumers toward using emojis in Twitter messages will be discussed. In section 2.2.3, the approach on analyzing the diversity and popularity of emojis preferred by consumers across cultures on a temporal basis will be summarized. Finally in sub-section 2.2.4, by extracting their semantic structural representation in text,

how different the emojis are used to represent emotions across languages will be discussed.

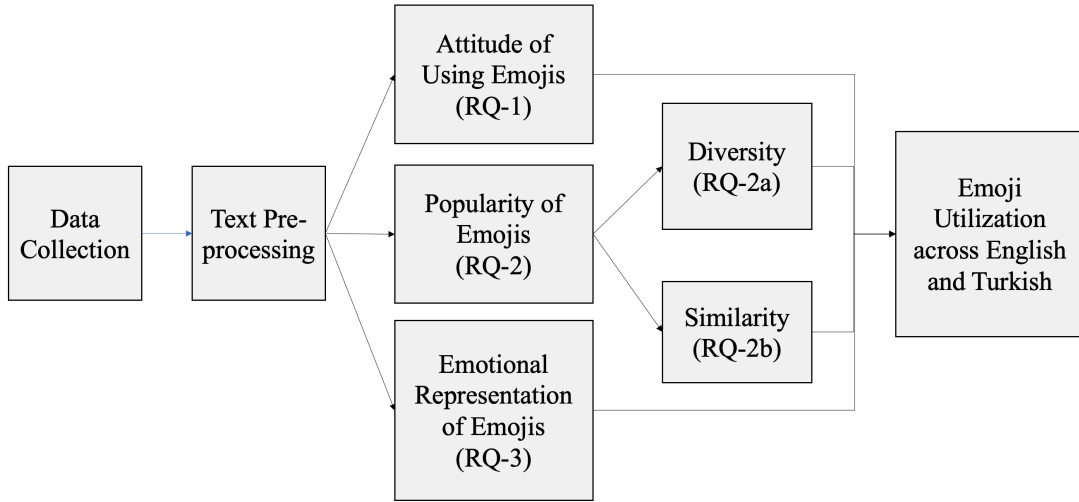


Figure 2.1 Steps of Analysis for Research Questions

2.2.1 Twitter Data and Text Pre-Processing

As shown by previous research, correlations exist between language and emoji usage both in terms of their frequency in Twitter messages (Kejriwal et al., 2021) and emotional representation (Park et al., 2014; Yuki et al., 2007). However, contradicting findings were observed when Hofstede’s cultural dimension of “individualism” was considered as the variable of interest effecting the emoji use as done by Li et al. (2019). Similarly, the effect of ‘third world’ and ‘second world’, former considered being more individualistic - latter more collectivist, as in Ljubešić & Fišer (2016) was also observed oppositely by (Kejriwal et al., 2021). For example, Western European countries used a less diverse set of emojis when compared to Middle East and USA. Differences were significant when USA and Asian countries like China, South Korea or Japan were compared (Yuki et al., 2007), but not very obvious when Europe and Middle East were considered. According to Hofstede’s cultural scores (Hofstede et al., 2010) English speaking countries like USA and UK have are substantially more individualistic cultures (*89 and 90 points respectively*) than China or South Korea (*20 and 18 points respectively*). Considering the future research directions of previous work suggesting analysis of larger scale data on specific languages and countries, English language was selected to represent messages from the more individualistic western cultures. As Turkey is at the center regarding its relatively

individualistic culture (*37 points*) and its geographic location when compared to China or South Korea, Turkish language was chosen to represent the countries in between. This way, a more detailed comparison was able to be made along with the opportunity of testing if the previous hypothesis of the individualism-collectivism or east-west factors effecting the usage of emojis in messages hold for the other countries in between.

A large scale of Twitter dataset was obtained using Twitter’s Search API. Twitter offers researchers specialized access options to use their API for running complex queries. Providing keywords and a time interval, all relevant messages can be retrieved from Twitter API in text format. Moreover, additional filters like language of the tweet, tweet content being sensitive, tweet status being ‘retweet’ or ‘reply’ may also be applied. As the context of this work is Branding Communication and the objective is cross-cultural analysis, all brands of Fortune 500 companies in 2021, having Twitter accounts both in English and Turkish languages were selected as sample. Finally, 33 brands from FMCG, Fast Food, Technology, Automotive, Apparel, Retail, Finance, and Logistics industries were found to be having accounts in both languages were chosen (see Table 1.1).

By using the selected brand names and accounts as keywords, all Twitter messages posted by brands and consumers’ replies to these messages between June 2016 and June 2021 were downloaded. A custom Java library was coded to download the messages and store them in Elasticsearch index for faster inquiries. The initial dataset included 8,101,034 English and 735,459 Turkish Twitter messages. English messages contained 2,708,874 brand posts and 5,392,160 consumer replies whereas Turkish messages contained 245,422 brand posts and 490,037 consumer replies. When compared to previous work in business context, the amount of data acquired was significantly higher. Number of Twitter messages, together with the number of messages containing any emoji by industries according to their types as brand posts (Post) and consumer replies (Reply) are shown in Table 2.1.

The messages’ plain textual content was kept separately in raw text format because the acquired Twitter data includes a lot of extra meta information that is not necessary for emoji analysis. Furthermore, using Python, distinct words and emojis were counted for each message, for each language. This helped to generate a numeric representation of the data. Word and emoji count data was stored in three dimensional sparse matrix format where first dimension may be referred as rows of messages, second dimension contain the non-empty indexes and the third contain the count value for each non-empty index. The indexes in the second dimension were the indexes of the dictionary of words and emojis previously counted. As a result, disk

Table 2.1 Number of Twitter Messages Collected by Industry

<i>Turkish</i>				
Industry	Messages		Contains Emoji	
	Post	Reply	Post	Reply
Automotive	8,924	16,447	575	2,395
Financial Services and Insurance	8,210	9,328	820	897
Food & Beverage	23,260	15,417	11,115	2,856
Food Service	18,027	18,424	2,780	2,307
Logistics	2,598	11,678	367	754
Personal Care	15,282	5,567	3,166	1,170
Retail	21,283	34,321	4,928	9,009
Technology	147,838	378,855	46,142	54,636
TOTAL	245,422	490,037	69,893	74,024

<i>English</i>				
	Messages		Contains Emoji	
	Post	Reply	Post	Reply
Automotive	71,046	189,724	10,751	33,974
Financial Services and Insurance	9,504	34,793	589	2,148
Food & Beverage	268,053	263,449	82,322	36,074
Food Service	1,359,714	1,372,728	172,436	239,418
Logistics	60,750	193,960	5,508	17,381
Personal Care	43,206	75,307	6,541	6,810
Retail	408,465	369,228	47,363	63,390
Technology	488,136	2,892,971	193,654	386,415
TOTAL	2,708,874	5,392,160	519,164	785,610

usage and processing time was fairly reduced when compared to running repeating queries on raw text data. The remaining analysis was carried out using the Python Numpy and SciPy modules following the production of the sparse matrices.

2.2.2 Attitude Toward Using Emojis in Twitter Messages across Languages (RQ-1)

For answering *RQ-1: How different is the attitude toward using emojis in messages across cultures/languages?*, mean number of emojis in Twitter messages was compared between English and Turkish datasets. Also, to differentiate between brands' and consumers' messages, the data was divided into conversational parts, such as brand posts and consumers' replies. As shown in Table 2.2 and Figure 2.2, despite the variances being slightly close, the distribution of number of emojis in posts and replies are highly right skewed, having first quartile, median and third quartile valued as zero. Majority of the tweets contain no emojis (Posts: 72% for Turkish, 80% for English – Replies: 84.8% for Turkish, 85.4% for English), and amongst the tweets with emojis majority contain a single emoji (Posts: 24% for Turkish, 14% for English – Replies: 9.8% for Turkish, 8.6% for English). Probability of observing a single emoji in Turkish messages is higher than in English, however, when there are more emojis in a message it is slightly higher for English. A common way to test the difference in means between two groups is applying t-test. However, the t-test implies that the data are normally distributed and that the groups have equal variances. Using t-test to compare the means of the skewed data such as the sample of this study could have led to unreliable results (Koplenig, 2019), as the confidence intervals would be calculated wrong. Moreover, when the data are discrete and display value inflation, statistical analysis of highly skewed distributions can be problematic (McElduff, Cortina-Borja, Chan & Wade, 2010).

Table 2.2 Summary Statistics: Number of Emojis in a Message & Proportion of Messages with Emojis

	English			Turkish		
	Posts	Replies	Overall	Posts	Replies	Overall
Mean	0.28	0.29	0.29	0.36	0.27	0.30
St. Dev.	0.69	1.25	1.09	0.73	1.24	1.10
Q1, Median, Q3	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Maximum	100	136	136	51	136	136
No Emoji*	80.8%	85.4%	84%	71.5%	84.9%	80%
Has Emoji**	19.2%	14.6%	16%	28.5%	15.1%	20%

* Proportion of messages containing no emoji.

** Proportion of messages containing at least one emoji.

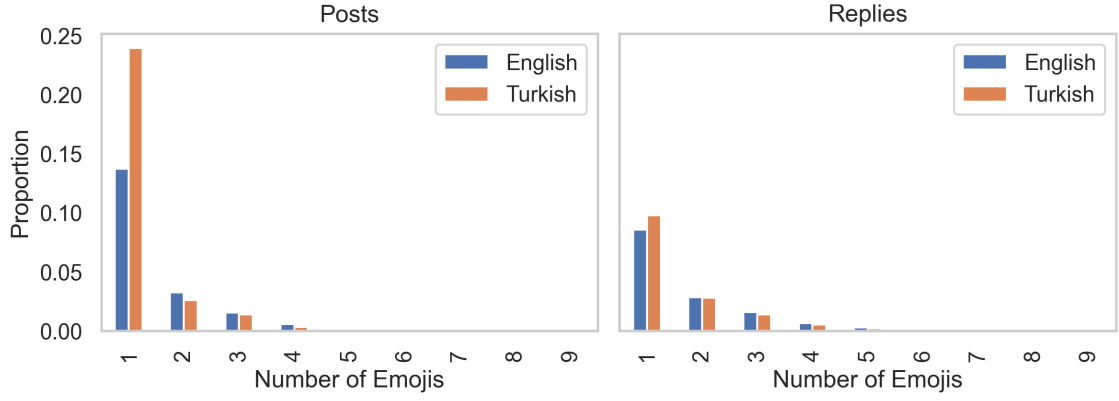


Figure 2.2 Distribution of number of emojis in messages

When count data with skewness is considered, an alternative approach to t-test to compare group means is applying a regression setting. Regarding this approach, one of discrete probability distributions is selected for the dependent variable, and its expected value is then modeled by a linear combination of the independent variables of concern. Being the only predictor regarding the case of this study, a binary language variable was created with English as the base value. Number of emojis in each message was considered as the dependent variable. A series of different regression models were built to describe the relationship between language and the number of emojis in a Twitter message. The p-values of the fit model's coefficients were used to determine if there was a significant impact of language on the amount of emojis used in messages or whether the observed coefficients were merely a result of chance.

Firstly, an Ordinary Least Squares (OLS) Linear Regression (LR) model was fit to compare the means between English and Turkish messages. When using a single predictor, this model's outcome is equivalent to a t-test. The assumptions of LR model's are exactly the same as t-test's: dependent variable is continuously numeric and normally distributed with equal variances in each group. The purpose of the linear model is to use it as a performance benchmark for comparison with other models. The parameter estimates in LR give the difference in means of emoji counts between English and Turkish messages.

Following, as the second model, Poisson Regression (PR) was fit which was more appropriate for count data. Regarding PR, the mean number of events and its variance are assumed to be equal in each of the compared groups (Afifi, Kotlerman, Ettner & Cowan, 2007), which was not valid for brands dataset. The Rate Ratio (RR), which is the estimate of the change in the relative mean number of occurrences between groups, is the only parameter of the PR. The binary variable for Turkish

language in this scenario and the model parameters, an intercept, combine linearly to generate RR. However, the variance of the distribution is significantly higher than the mean in emoji dataset (see Table 2.2). Negative Binomial (NB) regression, which incorporates an additional parameter to account for overdispersion, is an alternative approach (Afifi et al., 2007) as Poisson regression underestimates the standard errors of covariate parameters (Hilbe, 2011; Hinde & Demétrio, 1998).

As shown in Table 2.2, a fair amount of zero inflation was observed with majority of messages containing no emojis at all (English 84%, Turkish 80%). In order to allow for zero inflation, zero inflated versions of Poisson (Zero Inflated Poisson (ZIP)) and (Zero Inflated Negative Binomial (ZINB)) regression models were fit. These model types contain two sub models: a logit model for predicting excess zeros as a separate process, and a PR or NB model for predicting number of events occurring zero or more times. Regarding the emoji dataset, the logit part estimated the effect of language on the odds of probability of observing no emojis in a message. Dividing the number of events by the number of observations without the event, the odds of probability is determined. The secondary part (ZIP or ZINB) estimated the difference in emoji counts in messages between English and Turkish languages given that an emoji is used in a message. By nesting zero emoji count distribution and emoji count distribution in the same model, zero inflated models a provided a better view of the differences in preferring no emoji and using one or more emojis in branding messages across English and Turkish languages.

Ordinary Least Squares was used to estimate the LR model’s parameters. The Stats Models library in Python was used to estimate the parameters of the remaining models using Maximum Likelihood Estimation. Bayesian Information Criterion (BIC) as a measure of goodness of fit was used to compare the models’ performance. BIC is a common measure to assess how well a model fits a data set (Kuha, 2004), penalizing complexity of models and not being specific to any model type. Where the actual value of BIC is meaningless by itself, a lower BIC indicate better goodness of fit.

2.2.3 Preference of Emojis across Languages (RQ-2)

How differently emojis were preferred in English and Turkish speaking consumers (*RQ-2: How different are emoticons preferred across cultures/languages?*) was analyzed from two perspectives. Firstly, a comparison of richness of the emojis preferred was made by using tools for comparing lexical diversity amongst different corpus

(RQ-2a). Secondly, the mostly preferred emojis according to different popularity segments for each language were determined. By using similarity measures, English and Turkish Twitter messages were compared according to their emoji content, and how common and different were the set of emojis preferred was determined (RQ2-b).

2.2.3.1 Diversity of emojis preferred across languages (RQ-2a)

Aim of this subsection is describing the method employed for comparing the richness/diversity of emojis preferred between English and Turkish corpus answering (*RQ-2a: Does one culture/language use a more diverse set of emojis?*). A common method for quantifying the lexical richness of a corpus or text is calculating the Type-Token Ratio (TTR). TTR is defined as the number of various word types divided by the total number of word tokens (Baayen, 2001). As the aim was determining the richness/diversity of emojis used, not words, TTR was adapted as the number of different emoji types divided by the overall number of emoji counts:

$$(2.1) \quad TTR_i = \frac{V(C_i)}{C_i}$$

where i refers to language, $V(C_i)$ is the distinct count of emojis, and C_i is the total count of emojis in corpus i .

However when the richness of the words considered, the corpus size is known to have a significant impact on this measure, as shown by (Tweedie & Baayen, 1998). A remedy to this problem is described by (Johnson, 1944). They suggest a normalizing scheme by dividing the corpus into K segments ($K = C/N$), each segment containing N words. TTR is calculated for each segment and then averaged over the number of segments. As the English corpus contains around ten times more messages than Turkish corpus, this approach was also adapted to control the corpus size. Both datasets are divided into N message chunks ($N=1000$), and calculated TTR over each chunk. To omit zero counts, N is chosen to be large enough for the messages in chunks to contain some emoji. Finally, the mean chunk TTR is the average of the of all chunks:

$$(2.2) \quad TTRM_i = \frac{1}{K} \sum_{k=0}^K \frac{V(C_k)}{C_k}$$

where i refers to language, $V(C_k)$ denotes the number of distinct emojis in chunk k , C_k is the total number of emojis counted in chunk k .

$TTRM_i$ is in range from 0 to 1, higher values showing more richness/diversity. In order to test the significance of the difference of $TTRM_i$ across English and Turkish datasets, both a t-test and Monte Carlo permutation test was performed. As argued by (Koplenig, 2019), corpora cannot be considered as random samples when the complete known corpus of a specific context is used as the data set. The strong assumption behind statistical tests like t-tests are random sampling from the population, and finally be able to make inference. However, the randomness assumption is violated when corpus data are exhaustively collected for a particular context. A traditional test for statistical significance is becomes useless because there is no inference (Koplenig, 2019), or naturally it is not possible to make generalizations about the language as a whole. In this study, the goal was to determine if the observed variation in $TTRM_i$ was significant or the result of chance. Instead of drawing conclusions regarding emoji diversity for all English and Turkish speakers, this study’s goal was to investigate it in branding communication. Thus, for double checking the significance of the difference in $TTRM_i$ ’s, a modified version of Koplenig’s permutation test was run as first shuffling all messages into one dataset without regarding language.

Algorithm 1 Estimation of $TTRM_i$ by Monte Carlo Permutation Test

- 1: Shuffle all messages without regard to language into one dataset D
 - 2: Let $f_i \leftarrow$ sampling fraction for language i
 - 3: Let $s \leftarrow$ step size
 - 4: Let $TTRMS_i \leftarrow$ vector of TTRM for each step s and language i
 - 5: **for** $j = 1, 2, \dots, s$ **do**
 - 6: Generate two random datasets: D_1 for English and D_2 Turkish proportional to f_i
 - 7: Calculate $TTRM_i$ for D_i
 - 8: $TTRMS_i[j] = TTRM_i$
 - 9: **end for**
-

For the experimental setting 150.000 random observations were sampled from English and Turkish data sets to generate D . Algorithm was run for step size 1000 and f_i for each iteration was selected as 0.1 and 0.9 proportional to the original distribution by language.

The $TTRM_i$'s of 1000 iterations on the randomly generated data formed a bell-shaped distribution, allowing to make a comparison with the measure computed on non-random data. This procedure helped to compute a p-value indicating the probability of observing the measure as the proportion of the similar values extreme as the observed ones in the random data. If the p-value were less than a significance level, it might be considered as an evidence against the null hypothesis stating that the difference was observed by chance.

The procedure was also repeated by monthly basis from 2016 to 2021. Grouped by months, $TTRM_i$ was computed for English and Turkish, and plotted as a function of time, helping to analyze whether the variety of preferred emojis across languages are also changing through time.

2.2.3.2 Similarity of popular emojis across languages (RQ2-b)

In order to answer (*RQ-2b: How much different is the set of popular emojis across cultures/languages?*), the number of appearance of each emoji and how this differed between English and Turkish data sets was considered for analysis. Having rows for language, and columns for number of appearance and ranks respectively, in Figure 2.3 on the left column distribution of emoji appearances and on the right column frequencies of emojis ranked in decreasing order on the x axis are shown. It is clear that, regarding both English and Turkish data sets, a small proportion of emojis cover the majority all usage, similar to Pareto distribution. The distributions are heavily right skewed. For Turkish approximately 10% of emojis used contain the 90% proportion of the total count of all appearances of emojis. Similarly, for English around 9% of the emojis cover the 90% of all emoji appearances. Thus, the preference of emojis for both languages seem to follow a kind of power-law distribution. A common practice of identifying and quantifying power-law distributions is observing a straight-line behaviour where both ranks and frequencies are log (base 10) transformed and plotted (Clauset, Shalizi & Newman, 2009). However, as shown on Figure 2.4, a distribution with a constant power over the range of all ranks does not seem as the best fit, as there are variations of slopes along the x axis.

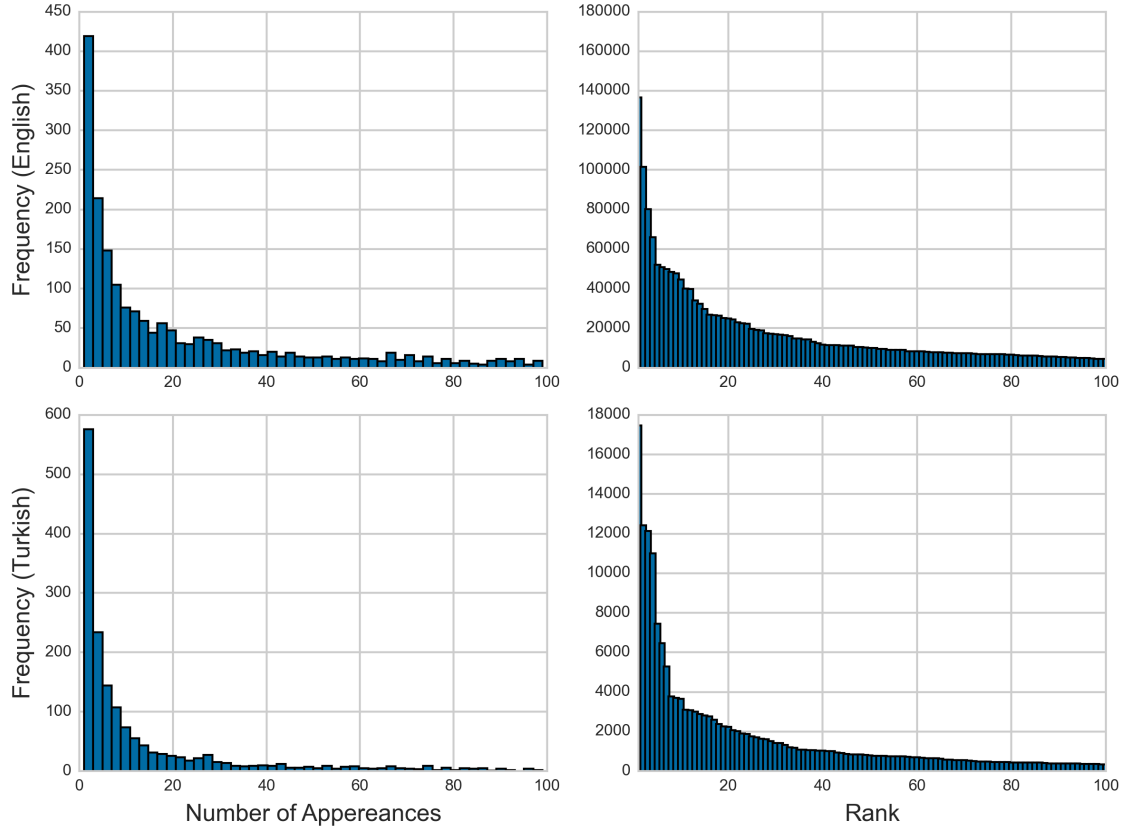


Figure 2.3 Distribution of emoji appearance frequencies.

Left column: Distribution of emoji appearances. Right column: Emoji's number of appearances by their ranks.

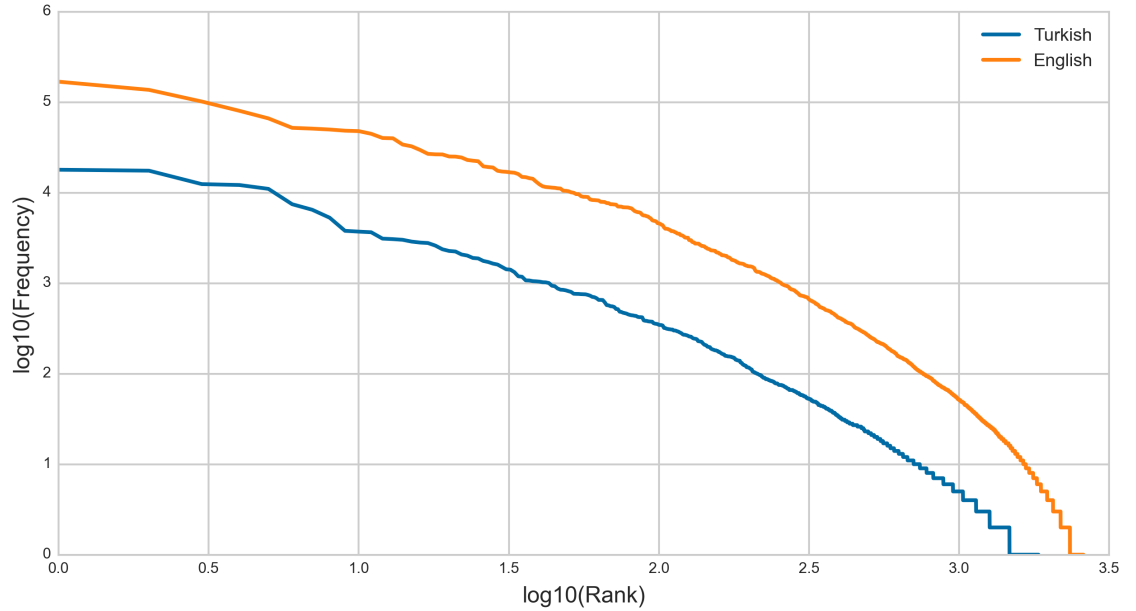


Figure 2.4 Log transformed emoji frequencies ordered by log transformed ranks.

According to (Kejriwal et al., 2021), while the overall log transformed frequencies

of emojis does not follow a straight line, when the rank axis is segmented into popularity phases it is possible to observe distributions close to power-law. Following the approach of (Kejriwal et al., 2021), the curve in Figure 2.4 was decomposed into three popularity phases (high, moderate, and low), which separately fit a power-law distribution with different parameters (see Figure 2.5). The high popularity phase contained 90% of all emoji counts in the message data set, followed by the moderate phase containing 90% of the rest of the emoji counts. In order to check the hypothesis that these three phases follow a discrete version of power-law distribution, or also named as Zipfian Distribution (Newman, 2005), a linear regression model for each of the phases for both languages was fit and R_2 values for each fit were computed. As seen in Figure 2.5, all R_2 values over 97% show strong evidence that these three phases were not generated randomly but through a Zipfian distribution similar to word frequency distributions. The slopes are correlated with the popularity category, and decrease as popularity of the emoticons decrease for both languages. For the first two segments, as the maximum x axis values are fixed, the slight difference between the slopes is caused by the data size difference between English and Turkish data sets.

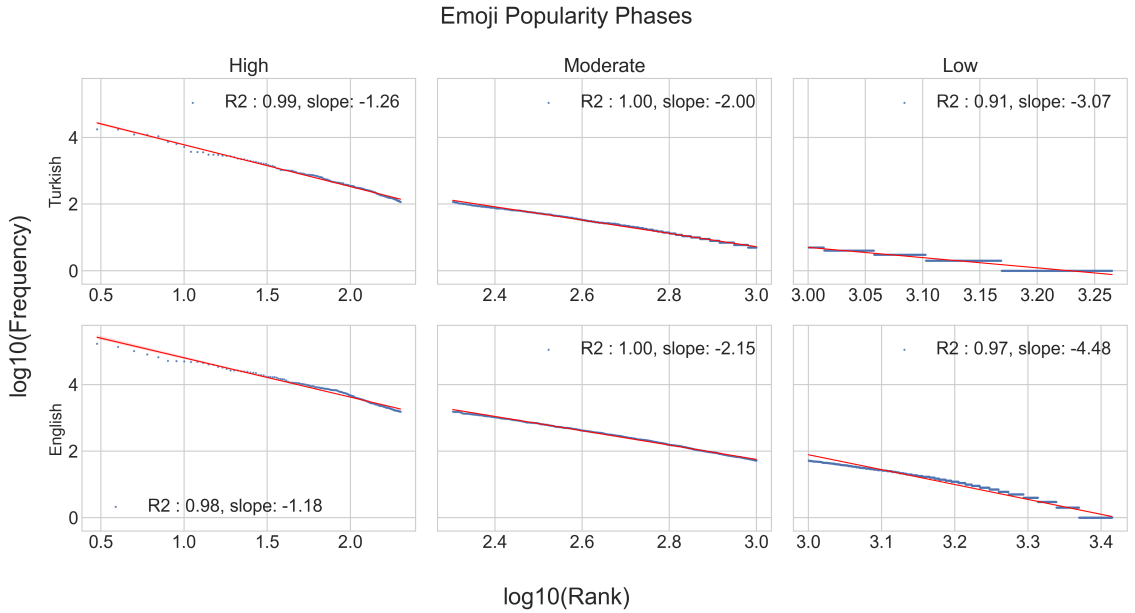


Figure 2.5 Segmentation of Log transformed emoji frequencies by log transformed emoji ranks

As number of distinct emojis were correlated with data size, it would not be correct to compare the slopes of the regression fits to evaluate the differences between the two groups. Furthermore, as popular emojis across English and Turkish were different, a comparison of the distributions in this sense could yield incorrect assessments. As a solution, a comparison of the items preferred in each popularity phase could be

made to assess how much similar or different the emojis are preferred, ranked by their importance among languages. In order to quantify the similarity of popular emojis between English and Turkish, Jaccard Distance (JD) for each emoji popularity phase was computed:

$$(2.3) \quad JaccardSimilarity_i(N_i^{en}, N_i^{tr}) = \frac{|N_i^{en} \cap N_i^{tr}|}{|N_i^{en} \cup N_i^{tr}|}$$

$$(2.4) \quad JD_i(N_i^{en}, N_i^{tr}) = 1 - JaccardSimilarity_i(N_i^{en}, N_i^{tr})$$

where i refers to popularity phase, N_i^{en} and N_i^{tr} denote the distinct set of emojis in phase i for English and Turkish languages respectively.

Jaccard Similarity is a measure of similarity between two sets of data, ranging from 0 to 100%, quantifying what proportion of distinct items are common in both sets. On the contrary, JD is a measure of dissimilarity between the two sets. It is the complement of the Jaccard similarity and can be found by subtracting the Jaccard similarity from 100%. In this case, JD was a proper statistic free of the unbalanced data size of each group, and did not get effected by the small differences in the ranking of the items in the sets. In order to check the significance of the JD statistic, whether it was substantial or by chance, a Monte Carlo permutation test was also applied as in Section 2.2.3.1, replacing the fourth step with JD_i statistic. The procedure was also repeated by monthly basis from 2016 to 2021. This time monthly similarity statistics (JD_i) were computed for English and Turkish and plotted as a function of time, helping to analyze whether the trend change over time.

2.2.4 Emotional Representation of Emojis across Languages (RQ-3)

A single emoji might represent a combination of emotions in different levels, it is hard to support an emoji to be used just for a single emotion by different consumers. Dating back to (Harris & Zellig, 1981), It is a common method to use co-occurrence analysis when underlying patterns in text data is to be explored. Regarding emotional representation of emojis, following (Riloff, 1996), (Tang et al., 2013) and (Song, Feng, Gao, Wang, Chen & Zhang, 2015) are some of the examples where emojis' co-occurrence with emotional seed words is used to compute the discrete probability distribution of emojis by 6 basic emotions of happiness, surprise, fear,

sadness, anger and disgust.

A similar approach was followed for (*RQ-3: How different are emojis preferred semantically by basic emotions?*), to understand how emojis represent basic emotions by analyzing their semantic usage. Top 90 words scored by each basic emotion were selected as emotional seed words from the Turkish Emotion Lexicon for Turkish (Tocoglu & Alpkocak, 2019a), and the NRC Emotion Lexicon (Mohammad & Turney, Mohammad & Turney, 2013) for English languages. The relevance of an emoji according to a set of emotional words was calculated using a co-occurrence frequencies of emojis with emotional words. Finally, a discrete probability distribution for each emoji was attained to show their relevance to the six basic emotions. For a specific emoji, the probability of representing one of the six emotions was calculated as:

$$(2.5) \quad P_i(e_j) = \frac{\sum_k co-frequency(e_j, ew_{ik})}{\sum_k \sum_i co-frequency(e_j, ew_{ik})}$$

where i refers to basic emotions as of happiness, surprise, fear, sadness, anger and disgust. e_j denotes emoji at index j . $co-frequency(e_j, ew_{ik})$ refers to frequency that that emoji e_j and the emotional word ew_{ik} at index k in the i 'th emotion category co-occur within a message in the data set. The denominator is the sum of all co-occurrences with all of the emotional words.

The emotion probability distributions of emojis for English and Turkish were compared against each other using Jensen-Shannon Divergence (JS), a symmetric distance index for discrete probability distributions (Endres & Schindelin, 2003). JS is the weighted sum of Kullback–Leibler Divergence (MacKay, 2003) by the probabilities of the events in question, which is the probabilities of observing six emotions for each emoji. Regarding discrete probability distributions, JS is defined as:

$$(2.6) \quad JS(P||Q) = \frac{1}{2} \sum_{x \in S} [P(x) \log(\frac{2Q(x)}{P(x) + Q(x)})]$$

where P and Q are discrete distributions on the sample space S .

In order to get an aggregate picture of the emotional similarities of semantic usage, from the most popular group, common emojis for each language are manually coded with the target sentiment they represent, as positive, negative and neutral using emoji sentiment scores by (Novak et al., 2015a). For example, 😊 represents a positive

sentiment, and is coded as "Positive", whereas 😞 represents a negative sentiment and is coded as "Negative". Following, two binary variables were created to represent negativity and positivity of an emoji, where neutrality is the baseline. As done in Section 2.2.2, regarding the fact that distance distribution was bell-shaped, an OLS regression model was fit in order to test the significance of differences by sentiment. By doing so, it is possible to distinguish between how similarly emojis were used by sentiments as well as identify the emotionally comparable and dis-similarly utilized emojis.

2.3 Experimentation and Results

2.3.1 Attitude Toward Using Emojis in Twitter Messages across Languages (RQ-1)

The outputs from the five models are shown in Table 2.3 in decreasing order by BIC values. As previously mentioned, the models examine the relationship between the language effect and the mean number of emojis used in a Twitter message. Categorized into posts and consumer replies, estimates of the model parameters, their standard errors as well as their p-values are reported from left to right. All models converged successfully, meaning that an optimum solution has been found for maximum likelihood estimation of the parameters.

Linear combination of the estimates of English (base variable, intercept) and Turkish (binary variable, 1 if Turkish) parameters of the LR model (*Model-1*) corresponds to the mean number of emojis used in a message (*see Table 2.3*). The results are exactly the same as t-test for the difference in means (*for posts t-stat: 57.32, p-value: 0.0 ; for replies t-stat: -11.52, p-value: 9.82*10⁻³¹*). LR parameters are significant for both posts and replies, showing strong evidence of the effect of language on the mean number of emojis observed in messages. Regarding posts, brands uses more emojis in Turkish messages on average than in their English messages (*difference in means: 0.0835*). However, regarding replies, Turkish-speaking users prefers fewer emojis in messages than English-speaking users (*difference in means: -0.0216*).

As the random variable is discrete count of emojis in a message and regarding the

fact that the distribution is not normal, PR model's (*Model-2*) BIC value drops significantly from $19 * 10^6$ to $9.0 * 10^6$, showing a better fit than LR. This time the exponential of the linear combination of parameters corresponds to the relative rate of observing emojis in messages. The events are observing certain number of emojis in a unit message, so the relative rate is actually corresponds to the mean number of emojis in messages same as in the LR model. PR produces similar results as LR, brands using more emojis in Turkish than in English and Turkish consumers tend to use less number of emojis. Being the last conventional model used, NB (*Model-4*) produces exactly same estimates with PR, but taking the variance into account through the dispersion parameter. The only difference of NB is the extended confidence intervals for estimates, so having a lower BIC value than PR.

Table 2.3 Parameter Estimates of the Evaluated Models

	Posts				Replies			
	Estimate	Std. Err.	p-val	BIC**	Estimate	Std. Err.	p-val	BIC**
<i>Model-1: Linear Regression</i>								
English (base)	0.2756	0.001	0.0*	6.20	0.2904	0.001	0.0*	19.0
Turkish	0.0835	0.000	0.0*		-0.0216	0.002	0.0*	
<i>Model 2: Poisson Regression (rate)</i>								
English (base)	-1.2887	0.001	0.0*	4.20	-1.2365	0.001	0.0*	9.0
Turkish	0.2648	0.004	0.0*		-0.0771	0.003	0.0*	
<i>Model 3: Zero Inflated Poisson Regression (rate, odds)</i>								
English (base, odds)	0.5977	0.002	0.0*	3.98	1.4943	0.001	0.0*	7.8
Turkish (odds)	-1.6585	0.013	0.0*		-0.1604	0.004	0.0*	
English (base, rate)	-0.2527	0.001	0.0*		0.4602	0.001	0.0*	
Turkish (rate)	-0.4740	0.004	0.0*		-0.2058	0.004	0.0*	
<i>Model 4: Negative Binomial Regression (rate, dispersion)</i>								
English (base)	-1.2887	0.001	0.0*	3.92	-1.2365	0.001	0.0*	7.18
Turkish	0.2648	0.005	0.0*		-0.0771	0.005	0.0*	
Dispersion	2.0897	0.007	0.0*		7.1159	0.013	0.0*	
<i>Model 5: Zero Inflated Negative Binomial Regression (rate, odds, dispersion)</i>								
English (base, odds)	-9.8977	2.224	0.0*	3.91	-8.0123	4.779	0.094	7.16
Turkish (odds)	-0.5118	0.014	0.0*		-1.9131	5.077	0.706	
English (base, rate)	-0.8191	0.005	0.0*		-1.2363	0.002	0.0*	
Turkish (rate)	-0.2049	0.007	0.0*		-0.0775	0.005	0.0*	
Dispersion	0.9860	0.011	0.0*		7.1146	0.018	0.0*	

* p<0.01
** All BIC values are normalized by 10^6 , i.e. $4.2 * 10^6$

To further examine the relationship between language and emoji usage, as also mentioned in Section 2.2.2, zero-inflated versions of the PR and NB models, which are ZIP and ZINB are also used for inference. These models allows us to analyze the likelihood of observing no emojis in a message and the likelihood of observing emojis in a message, given that the message contains at least one emoji. The results, presented in Table 2.3, shows that the ZINB model has a better fit to the data than

the ZIP model, as indicated by a lower BIC value. This is likely due to the higher variance of counts compared to the mean, which is not assumed by the Poisson part of the ZIP model.

The negative values of Turkish odds parameters in the ZIP model shows that the probability of observing no emojis in a message decrease when the message is Turkish, which is valid for brand as well as consumer messages. Estimate for Turkish rate of the Poisson Part in ZIP model of both brands posts and replies decrease the relative rate of observing emojis. This means that given a message contains an emoji, probability of observing more emojis increase when the language is English.

Considering the dispersion parameter, the standard errors of the parameters are higher and thus the confidence intervals are wider regarding the ZINB model. Also, this model has the lowest BIC, appearing as the best fit among the other models. Similar to PR, regarding brand posts, the logit part shows that being in Turkish decrease the probability of observing no emojis, and the NB part shows that being Turkish decrease the probability of observing one or more emojis in messages. However, regarding consumer replies, the odds parameters are not significant (*for base English(odds) p-value: 0.094, for Turkish(odds) p-value: 0.706*). This shows that the difference in probability of observing emojis between English and Turkish consumer replies (-0.02156) is not substantial. The effect of language in this case is observed by random chance. Finally, the NB part of the ZINB model, with a negative valued Turkish rate, shows that English consumer messages contain more emojis than Turkish messages, given that at least one emoji is observed.

When the results shown on Table 2.3 are considered, on average brands use more emojis in Turkish messages than in English. There is not enough evidence to state any significant difference in consumer messages between languages. However, when they use an emoji, English consumers prefer using more than one emoji in a message whereas Turkish consumers prefer using a single emoji.

2.3.2 Preference of emojis across languages (RQ-2)

2.3.2.1 Diversity of emojis preferred across languages (RQ-2a)

Emojis used in each language were counted and results are summarized in Table 2.4. As number of observations for English is much higher than Turkish, number of distinct emojis counted for English is higher than Turkish (*English: 2598, Turkish: 1840*). 1798 of the emojis preferred in English are also preferred in Turkish, which makes a Jaccard similarity of 69% and 98% respectively. Almost all emojis preferred in Turkish are also preferred in English.

Table 2.4 Number of preferred emojis and similarity

	English	Turkish
Distinct Emojis	2598	1840
Difference*	800	42
Jaccard Similarity	69%	98%

* Number of different emojis preferred compared to the other language.

Given the high similarity of Turkish speaking users' choices to English speaking users, richness/diversity of emojis preferred by language are also compared using the measure *TTRM*, normalized by corpus size. This helps to get a better insight of the similarity across languages. Table 2.5 summarizes the statistics regarding both the overall mean diversity (*TTRM*) and the monthly mean diversity by languages. The distributions of diversity for both languages are bell shaped, with close variances. For the overall comparison, the difference of 0.04 is substantial for both t-test and the permutation test ($p < 0.001$), showing that in Turkish richness of the emojis preferred is slightly less. However, when the monthly mean diversity is considered, this time Turkish is 1% above English, with near means and less variance. Moreover, the observed monthly difference is not substantial (p values 0.481 for *t-test*, 0.410 for *permutation test*), meaning that it is observed by random chance.

When diversity trend is analyzed over months from 2016 till mid 2021, similar findings are observed regarding the evidence for differences. In Figure 2.6, the two box plots on the left show the distribution of overall and monthly diversity statistics respectively, and the plot on the right show the monthly change over time. Trends are similar trends for English and Turkish, beginning from over 0.4's, converging to 0.25 in decreasing fashion.

Table 2.5 Diversity statistics

	English	Turkish	Difference	t-stat	p-val	*p-val perm.
Overall Diversity	0.33	0.29	0.04	4.982	0.001**	0.00**
Standard Deviation	0.16	0.18				
Monthly Mean Diversity	0.34	0.35	-0.01	-0.706	0.481	0.410
Standard Deviation	0.06	0.09				

*p-val perm: p-value of permutation test. No samples were observed with difference 0.04 following randomized experiments.

The results show that a more diverse set of emojis is preferred in English. However, the mean monthly diversity of emojis decreases for both languages from 2016 to 2021, meaning that diversity of the emojis preferred degrade similarly. A second question arise if the preferred set of emojis become more similar following the decrease in richness, which is summarized in the next section regarding RQ-2b.

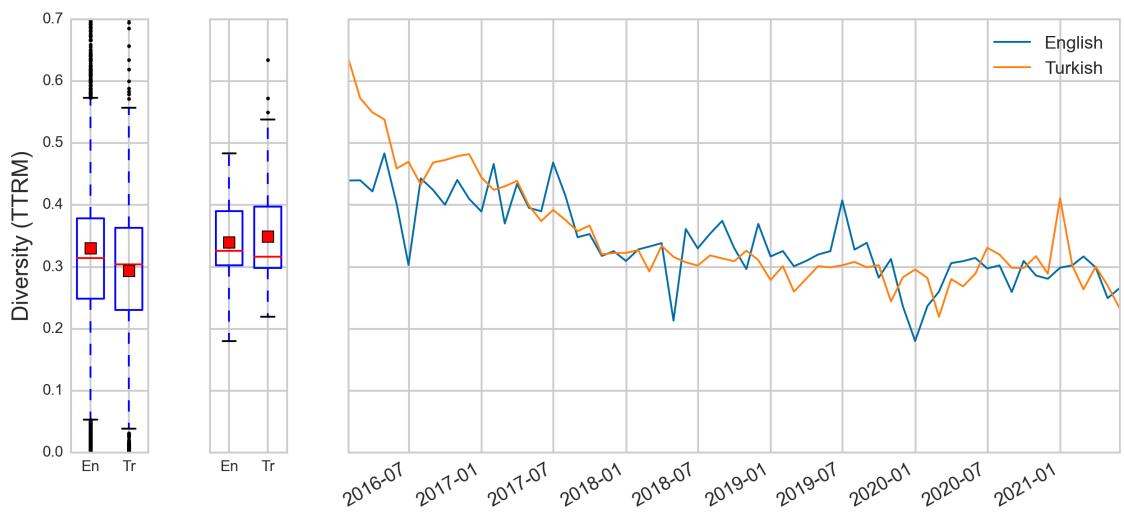


Figure 2.6 Diversity of emojis preferred.

From left to right: Overall diversity, monthly diversity, diversity by months

2.3.2.2 Similarity of popular emojis across languages (RQ-2b)

When the items in the three popularity sets are analyzed, different from the findings in previous work (Kejriwal et al., 2021), categories of emojis preferred have similar distributions for both languages. As shown in Figure 2.7, The top three ranked categories for high and moderate popularity levels are Smileys & Emotion, People & Body Parts and Travel emoticons. They are followed by less important emojis like symbols, activities or flags. The low popularity section contain the emojis with very few frequencies, thus, it is the section with the highest noise.

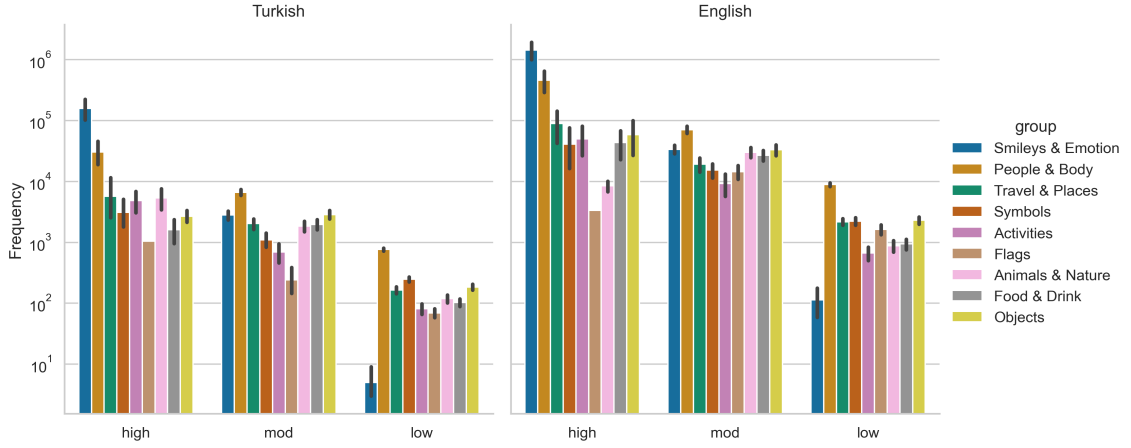


Figure 2.7 Emoji popularity phases by frequencies of emoji groups preferred. Popularity phases are grouped from left to right as high, moderate (mod) and low.

JD is a proper statistic free of the unbalanced data size of each group, and not effected by the small differences in the ranking of the items in the sets. Shown in Table 2.6, the statistics for the high, moderate and low popularity levels are 0.42, 0.49 and 0.65 respectively, in an increasing trend, showing that more popular emojis are more similarly preferred than less popular emojis.

For the three popularity levels from highest to lowest the p-values are 0.006, 0.028 and 0.0 respectively. The results show that the JD's for the first and last sections are substantial, but regarding the moderate level, JD is observable in the random data set more than 1% of the time, meaning that computed statistic is observed by random chance. Moreover, dissimilarity increase significantly to 0.65 in the low popularity level. While there is a substantial level of similarity in the most popular emojis between cultures, regarding the lower popularity levels dissimilarity increase. The reason is, the emojis in these two levels are counted less than 0.01 among the all emojis. Not following a specific statistical pattern, these less important emojis' usage may be regarded as noise.

Table 2.6 Dissimilarity (JD) across popularity levels

Popularity	Dissimilarity*	p-value**
High	0.42	0.006 <0.01
Moderate	0.49	0.028
Low	0.65	0.000 <0.01

*Jaccard Distance **Probability of observing measured JD statistic in Monte Carlo permutation test.

Dissimilarity is also measured as a function of time and some obvious patterns are observed. In Figure 2.8, monthly JD statistic between the English and Turkish

emoji preferences are plotted. For the two lower level popularity levels, emojis preferred are mostly dissimilar (70% on average). Following a similar pattern, both curves converge at around 74% and the slopes of the their trends are close to zero. However, the most popular 200 emoticons, beginning with 90% dissimilarity, follows a decreasing trend showing that the mostly preferred emojis in both languages are getting more common through time. When the decreasing diversity trends in the previous section is considered together with the increasing similarity of popularly preferred emojis, the preferences in both communities converge to a similar and smaller subset emojis.

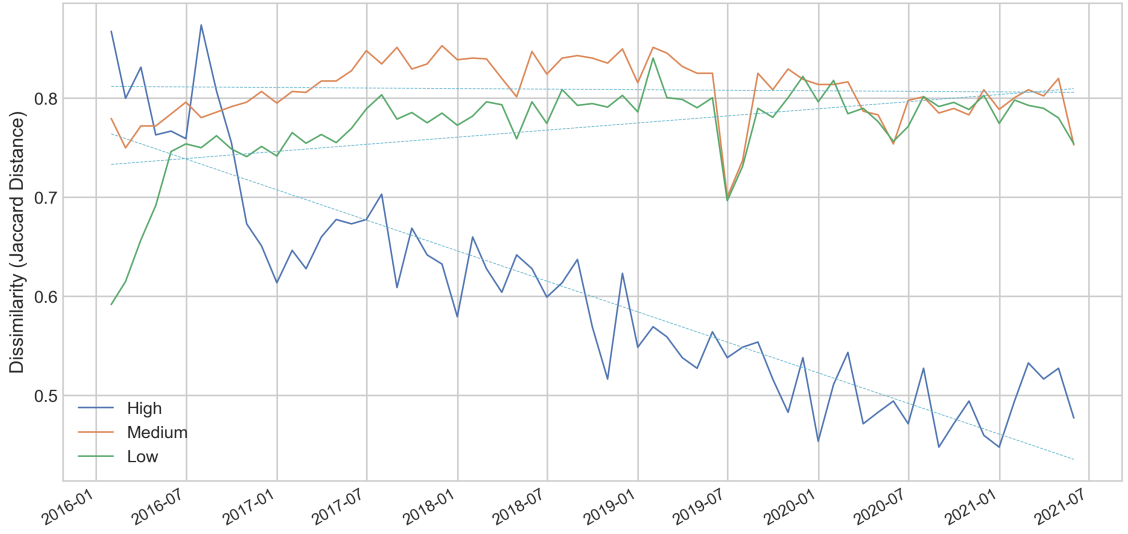


Figure 2.8 Dissimilarity of emoji popularity levels by time, 2016/01 – 2021/07 .

2.3.3 Emotional Representation of Emojis across Languages (RQ-3)

Among the 200 emojis in the most popular group selected for comparison, 128 of them were common in English and Turkish messages. The total frequency of these common ones cover more than 80% of the all emoji usages regarding both data sets. As described in 2.2.4, probability distribution of happiness, surprise, anger, disgust, sadness and fear for each of these emojis according to their co-occurrences with emotional words were computed. For comparing the differences in emojis' emotional usage, JS divergence between an emoji's emotional distribution in English and Turkish data sets is used as the distance measure. It is observed that emoticons, even they seem to represent a certain emotion, are used interchangeably with emotional words of different emotions.

The probability distribution by emotions is shown in the first two plots from the top of Figure 2.9, and the last plot shows the distance between the distributions for the English and Turkish messages, colored according to the sentiment of the emojis. Regarding English, emojis’ co-usage with happy words dominates the overall distribution. In Turkish, positive emotions of happiness and surprise are evenly distributed with negative emotions of anger, disgust, sadness and fear.

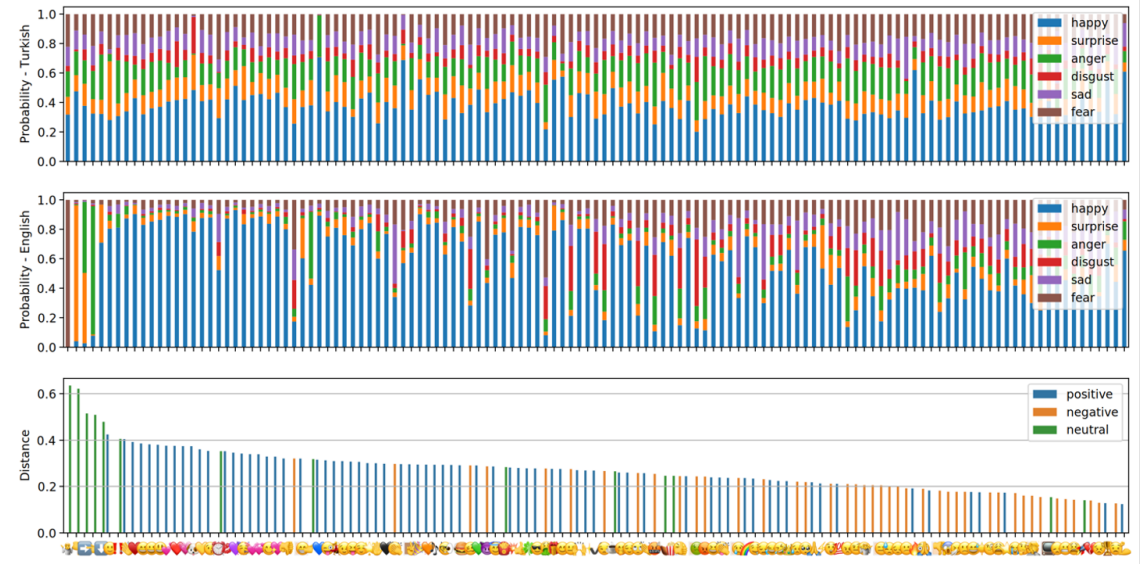


Figure 2.9 Emotional distribution of emojis and the distance of their distributions across languages.

Top: Emotional probability distribution of emojis in Turkish, center: Emotional probability distribution of emojis in English, bottom: distance in JS divergence colored by sentiments.

The bars in the distance plot are colored according to the manually coded sentiment of the emojis from Novak’s ranking scores (Novak et al., 2015a) as positive, neutral and negative. As shown, while distance decrease – so does similarity of emotional usage increase, emojis become more negative in their sentiments. In order to understand if negative emojis are more similarly used than positive emojis in the branding Twitter messages, mean distances by sentiments are compared. The distribution of distances by positive, neutral and negative sentiments are shown in Figure 2.10. Similar to the pattern in the distance plot (Figure 2.9), the most emotional diverse lexical usage is observed for neutral symbols (Mean: 0.360, Std: 0.15). On the other hand, emojis with negative sentiment are more similarly used than emojis with positive sentiment, having mean distances of 0.208 (Std:0.04) and 0.288 (Std: 0.06) respectively. To test if the observed differences in distances are substantial, an OLS regression model was fit as described earlier in sections 3.2 and 3.4. Prior to the model fitting process, two dummy variables are created to represent negativity and positivity of an emoji, where neutrality is the baseline. The output of the OLS

model is shown in Table 2.7.

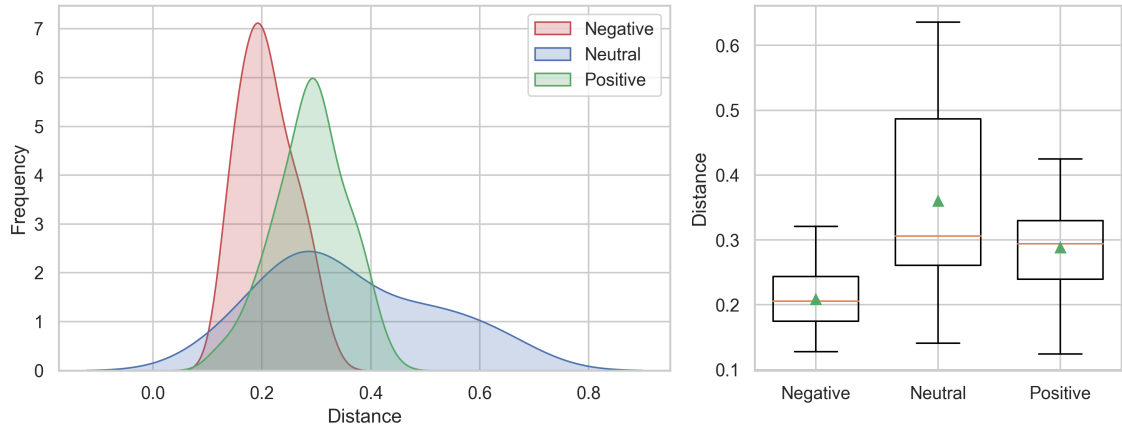


Figure 2.10 Distribution of distances by sentiments.

Left: Distance densities by sentiments, Right: Box plots of distances by sentiments.

Table 2.7 Number of preferred emojis and similarity

	Estimate	Std. Err.	t	p-val
Neutral (constant)	0.3699	0.021	18.005	0.00**
Negative	-0.0815	0.022	-3.626	0.00**
Positive	-0.1614	0.024	-6.806	0.00**

** $p < 0.01$

As the estimates are significant (Table 2.7), it may be argued that negative emojis are semantically more similarly used across cultures/languages regarding the combination of emotions they represent, where as positive ones are used more dissimilarly in brand posts and replies across languages.

2.4 Conclusion

This study aimed to measure cross-cultural differences in emoji choices and usage among English-speaking and Turkish-speaking Twitter users in relation to brand marketing setting. According to Hofstede’s Cultural Dimensions Model Turkish society is significantly more collectivist in culture than English speaking countries. However, more recent studies have also shown that Turkish people got more individualistic together with becoming more in-group collectivist since Hofstede’s model was proposed. Previous studies on the cultural variations in emoji preferences commonly reported that correlations exist among country and emoji use. Their data was collected for short periods of time, and their analysis were either at aggregate level or based on Hofstede’s model, which didn’t take the change in social behavior into account. By employing information theory-centric methods and considering the dimension of time, the suggestion of previous work was followed to further detail the analysis with alternative variables and countries for future research.

Alternative methods from information theory and distributional measures were used to answer three research questions related with emoji preferences and popularity across the two languages. In complement to previous work on emojis’ cultural usage, sampling from a more specific context of Twitter messages of a wider time range, the reported results shows that there are similarities as well as substantial differences among the preferences of both communities on aggregate level. However, when the preferences are analyzed over time, it is obvious that the usage pattern is getting more similar in terms of divergence and popularity.

Summary of the key findings regarding the three research questions are:

- *RQ-1: How different is the attitude toward using emojis in messages across cultures/languages?*

Brands used more emojis in Turkish messages than in English, whereas for consumers the difference is insignificant. Given that they use an emoji a message, Turkish speaking brands and consumers preferred using a single emoji whereas English speaking ones preferred multiple.

- *RQ-2: How different are emoticons preferred across cultures/languages?*

Majority of the emojis used in Turkish are also used in English, and when time factor is considered it is observed that the preferences in both communities is converging to a similar and smaller subset emojis.

- *RQ-3: How different are emojis preferred semantically by basic emotions (Hap-*

piness, Surprise, Sadness, Anger, Fear, Disgust)?

When their semantic emotional usage is considered, the findings shows that negative emojis are more similarly used across cultures whereas positive ones are used more dissimilarly in brand posts and replies across languages.

As an interesting finding, a certain group of emojis are becoming more common among the two languages. One reason might be social, where the English and Turkish users gets effected by the same popular culture in social media communication. Another reason might be technological, where the users may conveniently prefer the easily accessible emojis on device keyboards, which will also cause same emojis dominate the usage across other cultures. Being outside the scope of this work, the reasons behind the reported similarities and differences in emoji use across the two languages seek further attention from social theorists. Moreover, as also shown, besides being language and culture dependent, people’s preferences and use of emojis are also dependent on the context and time they are used.

While revealing interesting findings, this study has certain limitations. First, the only communication platform in concern is Twitter. Brands and consumers may act different in other social media platforms like Facebook or Instagram. However, it might also be considered as an advantage that messages in Twitter are limited to 280 characters (140 characters until 2017) and users may tend to use more emojis as non-verbal cues in such limited communication environments, helping the researchers to observe more emoji usages, thus explore a variety of emotional patterns. Second, Turkish language is selected to represent a culture which is in between in terms of individualism/collectivism among English and Asian speaking countries. Considered as more collectivist in societal culture than USA and UK according to previous studies (Hofstede et al., 2010), Turkey has been becoming less institutional collectivist and more in-group collectivist since 1995 (Pagda, Bayraktar & Jimenez, 2021). This means Turkish people have less value cohesiveness with society and more loyalty to sub-societal groups like family, becoming more individualistic in one sense and more collectivist in another. Similar to the fact that preferences might differ across platforms, they might also differ among countries like Turkey having in between collectivism. Moreover, models built in previous studies scoring the individualism-collectivism dimensions of countries are built on unrepresentative samples, with well-educated white collar participants. Naturally, the first and the second limitations makes it harder to generalize the findings of this work. By aggregation of messages from alternative platforms and languages, future research may justify if similar patterns exist through social media among other cultures. As long as data is publicly accessible, methodology applied in this study is applicable to other platforms and languages without any efforts other than data set expansion.

This study has a number of theoretical, methodological, and managerial implications and contributions besides the limitations mentioned. Theoretically, to the best of author's knowledge, this study is the first one analyzing the cross-cultural differences of emoji usage and preferences processing most up-to-date textual big data collected over the largest time span. In addition, sample data only consists of messages related with branding communication, making it context specific. A group of recent studies has already come up with findings of correlation among countries/language and emoticon/emoji preferences as (Park et al., 2014), (Park et al., 2014) or (Guntuku et al., 2019). However, their data is limited being collected over short periods of time as several months at most, and their analysis are at aggregate level not considering the effect of time, and contexts. This work has clearly shown that while similarities or differences are observed at aggregate level, preferences of brands and consumers regarding emojis are dynamic when time is considered. A less diverse and more similar set of emojis are preferred for both English and Turkish users over time converging to a common set of emojis. Still significant differences exist in their emotional semantic usage between the languages.

Methodologically, this study revealed interesting results employing methods from statistics and information theory on comparing cultural differences of emoji use in marketing context. Zero inflated models for semantic usage of words, word diversity, semantic information transfer from prior lexicons and computation of the emotional distribution of emojis were used as methods to understand the behavior of brands and consumers, in regards with social theory on differences of collectivism. As it was shown, combination of interdisciplinary approaches to such big data advance the theory on consumer behavior, observing the actual behavior of users instead of limiting the scope to a number of participants.

This study makes significant managerial contributions in furthermore to being the first to observe emoji data from specifically branding communication context. Word lexicons, or previously built sentiment models are commonly used by marketing researchers for analyzing sentiments towards a brand or a product from social media messages. Previously built word lexicons of emotions are not dynamic and it is an hard task to apply questionnaires on a certain number of participants continuously to update any word lexicon. Moreover, even there exists previous work on scoring emojis according to their sentiment, they are not dynamic either. It is shown that emojis represent a combination of emotions that varies according to culture and language. By computing their emotional distribution by the methods described, they can be used as sentiment or emotion indicators of consumer messages, as long as the computation is done continuously considering their potential for change. Focusing on branding communication and marketing context, emojis usage in other contexts like

sports, music or lifestyle comparing alternative cultures should be also be considered as topics to study for future research. Understanding if common global patterns of behavior exists, or if some cultures act more diverse than others for specific reasons are some of the questions seeking further answers.



3. EMOTION EXTRACTION FROM SOCIAL MEDIA

MESSAGES IN BRANDING CONTEXT ACROSS TURKISH AND ENGLISH LANGUAGES

One of the areas of research dealing with exploring the hidden information in information sources as social media platforms is Emotion Analytics, studying ways of extracting the true human emotions from unstructured data like text, images or videos. The motivation is, if emotions can be extracted from these data sources, then the attitude of the crowd towards a social policy or a certain product can be determined as well. Moreover, the signals of possible public events like mass protests can be foreseen, and certain precautions and actions may be taken. Emotions are one of the most crucial factors that affect how humans experiencing a stimuli event and take action. With a mechanistic perspective, “emotions can be defined as a positive or negative experience that is associated with a particular pattern of physiological activity (Pinker, 1999; Schacter et al., 2011). Psychologists have already been working on classifying and detecting human emotions since the last century, and as a result of the problem’s complex nature alternative theories have emerged. One of them, classifying emotions in a discrete way, is Paul Ekman’s six basic emotions as sadness, happiness, anger, fear, disgust, and surprise (Ekman, 2005). Another example is Robert Plutchik’s dimensional approach to emotions, describing emotions as a composition of primary emotions like joy, sadness, acceptance, disgust, fear, anger, surprise and anticipation (Plutchik, 1984). Naturally, the former research on emotions in Psychology constructed the basis of Emotion Analytics, and answered the first question of how emotions can be classified for computer science researchers working on this area.

Since late 90’s, one of the focuses of Computer Science research has been information extraction from textual content, and Twitter has been one of the most popular data sources in this context because of its availability though application programming interfaces. Based on the definitions made by psychology, extracting emotions from text is still a hot topic in areas of computer science as natural language processing or human-computer interaction. From the view of computer linguistics, emotion

detection or extraction is the process of identifying human emotions expressed in the text. The task can also be described as an extension to Sentiment Analysis (SA) in Natural Language Processing (NLP) context, and has been proven to be useful in several areas such as marketing, advertising (Jin, Li, Mah & Tong, 2007; Qiu, He, Zhang, Shi, Bu & Chen, 2010), question answering systems (Lita, Schlaikjer, Hong & Nyberg, 2005; Somasundaran, Wilson, Wiebe & Stoyanov, 2007), text summarizing (Seki, Eguchi, Kando & Aono, 2005), recommendation systems (Terveen, Hill, Amento, McDonald & Creter, 1997) or even improving information extraction (Riloff, Wiebe & Phillips, 2005). However, what Sentiment Analysis extracts from text is generally the polarity being negative, positive, or neutral. However, negative or positive feelings are not equal. As an example, if we consider fear and anger, they both have negative sentiment, but fear is represented by humans in a much more passive way, whereas anger usually leads to action such as not consuming certain products or supporting a political entity (Miller, Cronin, Garcia & Branscombe, 2009). Thus, being able to extract certain discrete emotions from text means acquiring more information, and creating new potential use cases, especially in political science, psychology, marketing, or human-computer interaction, which recently attracts emerging research.

Regarding marketing context, customers' reactions to products and services can be analyzed to decide which aspects of the products can change, or communications with the customers can be enhanced with prior knowledge about their emotions, to increase their satisfaction. In the decision-making process, being able to extract emotions of the target population will help to understand the impact of products, services, or policies and managing the consumer's reactions by monitoring the emotions. Brands may tailor the emotional content of their social media messages in order to increase the positive reactions. Moreover, monitoring of the aggregate consumer emotions can help to prevent a possible negative word of mouth about a product or service spread.

Aim of research described in this chapter is two-fold. Emotion classification requires extensive amount of annotated data. While such data sets are available in English, they are not very common in most of the languages like Turkish. So first, a flexible method to articulate textual data annotated by 6 basic emotions (*happiness, surprise, sadness, anger, disgust, fear*) (Ekman, 2005) and neutrality, by getting use of emotional signals of emoji characters in text is proposed. Using a scoring function and an algorithm to rank emojis according to the dominant emotion they represent, the method proposed may be applied to generate annotated data from social media messages in any context and language. As of our knowledge, this study is the first one to evaluate the emotional representation of emojis in branding communi-

cation and Turkish language. Second, performance of alternative NLP models to extract emotions from the articulated data using statistical and state-of-the-art deep learning methodologies are evaluated for Turkish and English languages. Previous research have already proposed complex models as Recurrent Neural Net (RNN) or Bidirectional Encoder Representations from Transformers (BERT) regarding emotion extraction from textual sources, and these models are proven to be useful for English language. However, there is still a research gap about the performance of these models for more local languages like Turkish and specific contexts like branding communication. In short terms, the goal of this work is to collect contextual data sets for different languages and then explore the best suited classifier algorithm prior to propose a framework for automatically extracting emotions from such textual data, answering the following research question:

RQ: *How can emotions be automatically extracted from brands' and consumers' social media messages in different languages and contexts?*

The rest of the chapter is organized as follows: Following a summary of previous work done on emotions and emotion extraction in the next section, technical details about the methodology used in generation of annotated data and model evaluation are detailed in section 3.2. Then, in section 3.3, performance of the evaluated models are discussed followed by a conclusion about the findings and projection for further studies.

3.1 Literature Review

Emotion Analytics research gets its roots from Psychology. In order for classifying emotions in text, images or any other kind of human generated content, one first has to be able to describe what an emotion is. According to (Hockenbury & Hockenbury, 2007), an emotion is a complex psychological state that involves three distinct components: a subjective experience, a physiological response, and a behavioral or expressive response. When making decisions in our daily lives, emotions, which are also bounded by our cultural context, play an important role. Human beings' cognition bounds the subjective experience; thus, everyone's different perception of events will create different psychological responses for each of them, finalizing with an action being the expressive response. For the case of Emotion Analytics, the

expressive response may be a message, an image or a video posted to the social media environment. But how are we going to differentiate emotions? In section 3.1.1 the theoretical basis of discrete emotion categories that this work is also based on will be discussed. Following, previous studies on emotion classification from textual data sources that this work aims to contribute will be summarized.

3.1.1 Emotion Theories

Research regarding extracting different emotions from human behaviors and actions goes back to the 19'th century, Darwin's work on animal and human emotions (Darwin & P., 1998), mostly influenced by Psychology and Cognitive Science. Even so, there is no accepted universal model for differentiating and detecting human emotions from their behaviors. The two existing viewpoints are Discrete and Dimensional approaches. Discrete Emotion Theory supports the idea that some emotions are distinctly distinguishable on the basis of neural, physiological, behavioral, and expressive features regardless of culture (Colombetti, 2009). Paul Ekman's studies are one of the bases of the Discrete Emotion Theory, following a cross-cultural research, defining six basic emotions as sadness, happiness, anger, fear, disgust, and surprise (Ekman, 2005). In the '80s, Robert Plutchik introduced the wheel of emotions. According to Plutchik, there exist eight primary emotions that serve as the foundation for all others: joy, trust, fear, surprise, sadness, disgust, anger and anticipation, which are also the opposites of each other in pairs, and a combination of these primary ones produce other complex emotions like boredom, disapproval or interest (Plutchik, 1984).

Most research related with Emotion Analytics gets use of the discrete approach to base their work on, while the dimensional approach introduces an alternative perspective. According to the dimensional emotion theorists, different than the discrete emotion theorists stating that varying emotions correspond to varying neurological processes in the brain, all emotions are the results of a common neuro-psychological process. This method defines emotions in two or more dimensions, as in Russel's Circumplex model (Russell, 1980) suggesting that emotions can be represented in a two dimensional circular space, arousal in one dimension and valance on the other. Although they are used scarcely in emotion detection literature, the dimensional approach has also been shown to be useful for textual data (Calvo & Mac Kim, 2013). This work tries to accomplish the same objective as prior work on extracting emotions from Turkish text, which mostly focused on Ekman's six basic emotions,

with the addition of neutrality for the purpose of comparison.

3.1.2 Emotion Extraction

As a sub-task of NLP and SA, research on emotion extraction from textual data is still developing, as the problem itself is context-sensitive and complex (Ben-Ze'ev, 2001). The problem is context sensitive, because the key terms in language used for alternating contexts like sports, healthcare, or politics differ, thus, flexible models for classifying emotions in different contexts is required. The complexity, firstly, comes from the problem's sensitivity related to human cognition and social context; secondly from expressions being mostly containing not a single emotion but a mixture of emotions; and lastly, the confusing linguistic use of emotional terms. Apart from being a more complex case of SA, the research steps are similar in emotion detection from textual resources: firstly data has to be collected, secondly how the data will be annotated must be determined, then data must be pre-processed to be suitable for models to be built on, and lastly alternating models' performances must be evaluated.

3.1.2.1 Benchmark data sources

Regarding the data collection and annotation phases, unlike SA, benchmark data sets are less common for emotion classification from text. However, for increasing the performance of any model, especially human annotated data is required to build statistical or machine learning models. Apart from the unsupervised approaches, most of the previous research gather teams of annotators, and make a start with creating data sets with certain reliability measures. Then, they evaluate their models' performances on the benchmark data sets. The most common one is the ISEAR data set, formed by responses of about 3000 people around the world, annotated with 7 basic emotions as joy, fear, anger, sadness, disgust, shame, and guilt (Scherer & Wallbott, 1994). (Asghar et al., 2017; Boynukalin & Karagoz, 2013) are examples getting use of ISEAR for model evaluation. Some of the other annotated data sources are SemEval, annotated with Ekman's 6 basic emotions from 1250 news headlines (Strapparava, Strapparava), updated frequently for the SemEval text classification contests (Mohammad, Bravo-Marquez, Salameh & Kiritchenko, 2018; Rosenthal,

Farra & Nakov, 2019). Alm’s annotated fairytale data set (Alm, Roth & Sproat, 2005) consisting of 1580 sentences composed from children’s fairy tales also annotated with Ekman’s 6 basic emotions. WordNet-Affect is another frequently used resource for lexical methodologies, containing English words labeled with their domain (sports, politics, medicine, etc.) and emotion category (Strapparava, Valitutti & others, 2004).

Regarding the Turkish case, public emotion annotated data for academic use are less common (Altmay & Turhan, 2022). Annotated with Ekman’s 6 basic emotions, Tocoglu and Alpkocak developed a Turkish text to emotion dataset called TREMO conducting a survey on 4709 individuals, and collecting 27,350 entries (Tocoglu & Alpkocak, 2018), with participants describing their experiences on certain emotions with text. By getting use of the TREMO dataset, they created a Turkish Emotion Lexicon (TEL) and further used their word scoring for classifying emotions from Turkish text (Tocoglu & Alpkocak, 2014,1). More recently, Tocoglu et al. expanded the size of TREMO, naming it TURTED, labeling more than 195K tweets by counting the emotional keywords from TEL appearing in twitter messages (Tocoglu & Alpkocak, 2019a). As training data is limited for the Turkish case, in (Shehu, Sharif, Sharif, Datta, Tokat, Uyaver, Kusetogullari & Ramadan, 2021) proposed three data augmentation techniques (shifting words of tweets, shuffling them and hybrid approach) and evaluate RNN, Convolutional Neural Net (CNN) and Hierarchical Attention Network (HAN) deep learning architectures on the twitter data set they generate. Their data augmentation method improved classification accuracy by the RNN model in all three cases. However, the target variable was the sentiment (negative, positive, neutral) of the tweets not discrete emotions.

As emotion expressions in text are complex to extract, the emotion to be extracted depends on language, context and time (Vosoughi, Zhou & Roy, 2016). Moreover, annotation of text data is time consuming and costly, thus, the researchers in the field began seeking for more flexible alternatives to automatically annotate the text data from additional information like hashtags or emoticons which are commonly present in micro blogging messages. Wang et al., in 2012, achieved to create a 2.5 million tweets data set annotated by hashtags and emoticons (Wang, Chen, Thirunarayan & Sheth, 2012). Similarly in (Asghar et al., 2017), a rule based approach to classify emotions from text is presented by annotating the twitter text with emoticons and slang words. Although in (Tocoglu & Alpkocak, 2019a) a lexicon based approach has been followed, more validation is required on the data accumulated on alternative contexts. One of this work’s purposes is understanding the usage of additional information in text as emotional words, emojis or hashtags, and explore their relation to emotion expressions of social media users.

3.1.2.2 Text pre-processing & feature extraction

Following the data collection and annotation phases, text data is pre-processed and transformed to its structured form. Similar to the approaches applied in SA, first, text is tokenized, then the common techniques of Part of Speech Tagging (POS), stop words removal, stemming and lemmatization are applied. POS is used for tagging each word in a sentence for its morphological category such as adjective, noun, verb or norm. For stop words removal, common terms as “the”, “a”, “or”, “and” are removed from text, as they don’t carry extra information about the content. Stemming converts the words into their root forms (i.e. “automatic” and “automate” become “automat”), and lemmatization replaces words with their base form (i.e. “caring” becomes “care”, “cars” become “car”). The cleaned and tokenized text is then transformed to its structured form by simply counting the words in the text, and representing the unstructured text data with words as features and words’ frequencies as their values. Named as Bag of Words (BOW) method, this approach does not take the sequence of words in account. As an addition, for sequence representation, n-grams are proposed. According to the n-gram representation, the words are counted in sequences of n, thus, text can be represented as frequency of single words (unigrams), two words in sequence (bi-grams) or more. Moreover, some words may be appearing whole data set very frequently, thus, not helping to differentiate between documents. The solution to such problem is Term Frequency – Inverse Document Frequency (TF-IDF) weighing. It is simply multiplying the word frequency to the logarithm of number of documents in the whole data set divided by the number of documents the word is appearing. This way, common words in the data set will have less weight, and the scarce words will have more importance.

Alternative methods for pre-processing of text data to evaluate the performance of text classifiers has been widely studied in SA research (Bao, Quan, Wang & Ren, 2014; Pak, Paroubek & others, 2010; Pang et al., 2002). As certain techniques are proven to be successful in SA, the researchers directly apply the text cleaning as described, and mostly represent the data with BOW model, especially if conventional statistical models are applied. As several examples, in (Hajar, 2016; Hancock, Landrigan & Silver, 2007; Jain, Kumar & Fernandes, 2017) unigram tokenization is applied for feature extraction. Differently in Refaee & Rieser (2014), when the authors incorporate morphological features upon the BOW model, they achieve a significant increase in model accuracy for Arabic Twitter feeds. In another study, the authors manually labeled around 8000 tweets and extracted 11 features: unigrams, bigrams, personal pronouns, adjectives, Word-net Affect lexicon, Word-net Affect lexicon with left/right context, Word-net Affect emotion POS,

POS-bigrams, Dependency-Parsing, and Emoticons, achieving a classification accuracy of 73.24% (Balabantaray, Mohammad & Sharma, 2012). Consequently, for the Turkish language case, there is a need to compare the contribution of morphological and syntactical features to the BOW model, an objective that this work also aims to achieve.

A more recent approach for feature representation of text documents is the incorporation of word embedding layers. The general idea is words co-occurring frequently in certain documents are also similar semantically. Unlike BOW model, word embedding models represent each word as a vector in n-dimensional space, and the distance between these vectors display the semantic similarity between them. Some of the examples of word embedding models in literature are GloVe (Pennington, Socher & Manning, 2014), ConceptNet (Speer, Chin & Havasi, 2017), Word2Vec (Strapparava et al., 2004), and BERT (Devlin, Chang, Lee & Toutanova, 2018b). BERT, developed in 2018, is designed to pre-train deep bidirectional representations from unlabeled text data by getting use of deep learning’s auto encoders. The pre-trained models can then be used on feature extraction on new data sets with fine tuning. In SemEval-2019 text processing competition, teams running for the task of “EmoContext: Contextual Emotion Detection in Text” getting use of BERT for feature representation achieved significantly higher precision among the other teams (Chatterjee, Narahari, Joshi & Agrawal, 2019).

In recent years several research papers comparing the performances of such vectorized representation of documents on Turkish language have been published. Using TREMO data set for Turkish and Twitter Emotion Corpus for English texts, (Ucan, Dörterler & Akçapınar Sezer, 2022) evaluated the accuracy of Gated Recurrent Unit (GRU), Long Short-Term Memory Networks (LSTM) and Attention (Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser & Polosukhin, 2017) deep learning networks alternatively on pre-trained embedding layers as multilingual BERT, DistilBERT and BERTurk. Their experimental studies showed that the proposed pre-trained emotion model has the highest success rate compared with the methods used in previous studies. (Uymaz & Metin, 2023) applied emotion enrichment models on Word2Vec, GloVe and BERT proposed by (Seyeditabari, Tabari, Gholizade & Zadrozny, 2019), for emotion classification on Turkish data sets predicting 6 basic emotions . The idea of (Seyeditabari et al., 2019) basically was words of opposite emotion must be represented by vectors far away from each other. According to their findings BERT outperforms Glove and Word2Vec for Turkish language. Moreover, emotionally enriched BERT embedding layer helped to increase classification accuracy by 22% over original BERT vectors.

3.1.2.3 Emotion classification models

There are generally three categories of methods for emotion classification and extraction from text in literature. These are learning-based, rule based and lexicon based methods. Learning based methods (or supervised methods) gets use of annotated data to train statistical classifiers as Multinomial Naïve Bayes (MNB), Support Vector Machines (SVM) (Cortes & Vapnik, 1995), or neural network architectures like Convolutional Neural Net (CNN)s and RNNs. Rule based methods are based on grammar and language structure. Finally, the lexicon based methods gets use of lexical resources such as Word Net Affect to calculate the emotion in text content from the emotional words.

Learning based approaches using statistical machine learning algorithms like Naïve Bayes, Decision Trees, Random Forest (RF) or SVM are the most common ones used for the English language-based analysis for modelling emotion classifiers. Moreover, in the last ten years, with their proven success in classifying sequences, neural networks have been a popular choice for such classification tasks. These supervised models are trained on human annotated data, and mostly evaluated on the open data sets as mentioned (i.e. ISEAR). (Hasan, Rundensteiner & Agu, 2014) created an emotion lexicon of 28 emotion words using the Circumplex model, and removed any non-emotional words from their data set. Following, they applied Naïve Bayes, SVM, Decision Trees and K-Nearest Neighbors (KNN) classifiers and achieved an accuracy close to 90% for the emotion categories in Circumplex model. (Rao, Pang, Xie, Liu, Wong, Li & Wang, 2017) created a combined model of two supervised layers. The first layer extracts the topics related with emotions, and the second layer, by using topics, establishes associations among bi-terms and emotions. In (Chaffar & Inkpen, 2011) and (Aman & Szpakowicz, 2007), authors compare the alternative supervised approaches, and similarly they conclude that SVM performs significantly better than the other techniques.

As their proven success on sequence classification, neural networks' performance has also been evaluated on the emotion classification domain regarding English language. As an example, (Zahiri & Choi, 2017) worked on emotions in TV show transcripts applying four types of sequence-based CNNs. In (Baziotis, Athanasiou, Chronopoulou, Kolovou, Paraskevopoulos, Ellinas, Narayanan & Potamianos, 2018) authors proposed a bi-directional LSTM architecture with Word2Vec embedding, and were ranked 1st in the 2018 SemEval competition. (Lazemi & Ebrahimpour-Komleh, 2018), with minimal feature engineering, developed a hybrid deep learning model that benefits both CNN and RNN deep models, and their result on a data

set with 629 sentences showed a about 0.12% higher performance compared with previous multi-label models.

Rule based approaches can be regarded as unsupervised methodologies, as they are applied by exploring the grammatical rules to classify emotions, or by extracting the specifically the emotional words from data getting use of lexical resources like Word-Net Effect. Rule based approaches define methods by using a dictionary with affective words and rules defined by linguists. They are more disadvantageous than the other methods as they need to consider complex grammatical rules of different languages, and these rules also vary according to context and time. Shaheen, El-Hajj, Hajj & Elbassuoni’s proposed method generates rules from input sentences using the ontology defined in Word Net and Concept Net (Shaheen et al., 2014). They, first, extract rules reference rules from the training set, and then compare the rules generated by their method using various ontology like Word Net for evaluation. A mixture of rule-based and learning-based method was applied by (Anusha & Sandhya, 2015), defining rules based on syntactic and semantic features of text and then using them as input features to learning-based classifiers including SVM and Naïve Bayes.

As another unsupervised group, regarding the lexicon approaches, the most simple method is using a predetermined set of words keywords (from lexical resources like WordNet Affect) that categorize text to emotion classes. An example is (Perikos & Hatzilygeroudis, 2013) proposing an algorithm that examines emotional words in text and calculates a rating from these words, and by comparing the strength of this rating towards a certain emotion category the classification is achieved. A more complex approach is proposed by (Gill, French, Gergle & Oberlander, 2008), getting use of co-occurrence semantic space techniques such as Latent Semantic Analysis and Hyperspace Analogue to Language in order to calculate the semantic similarity of text to the emotion keyword. Working on emotion classification of YouTube comments, (Hajar, 2016) compute a text entry’s similarity to each target emotions, using Pointwise Mutual information measure, in order to classify the text into a particular emotion category.

There are also studies on English emotion classification using word embeddings, where each word is represented with a vector of words and certain weights associated with the location of each word represented by its proximity to other words in the text corpus (Al-Omari, Abdullah & Shaikh, 2020; Huang, Trabelsi & Zaïane, 2019; Huang, Lee, Ma, Chen, Yu & Chen, 2019; Tomihira, Otsuka, Yamashita & Satoh, 2020). For example, in (Al-Omari et al., 2020) authors show that using pre-trained BERT transformers in deep learning classification architectures increase the F score

significantly on emotion classification domain for English text.

Although there is a growing literature on English text resources in recent years, the literature on emotion detection from Turkish text is still yet to emerge. For a similar analysis on the Turkish language-based content, literature involves a dozen of articles that use supervised or unsupervised approaches where evolving from evaluating statistical models to analysis with deep learning networks. As the first study on emotions in Turkish, in (Boynukalin & Karagoz, 2013), authors got ISEAR dataset’s joy, sadness, anger and fear categories translated to Turkish by 33 volunteers. After applying the conventional pre-processing and feature extraction phases, they apply MNB and SVM classifiers on the translated training data. In (Tocoglu & Alpkocak, 2014), authors develop a system for classifying the emotions of happiness, shame, guiltiness, disgust, sadness, anger and fear. They acquire their training data by using 500 university students’ descriptions of their most intense moments for the seven emotion categories. By weighing the text features with TF-IDF method, they create a vector space model to represent the text data, and apply Naïve Bayes for classification, where they achieve around 86% accuracy for the seven emotional classes. In (Tocoglu, Ozturkmenoglu & Alpkocak, 2019), authors choose 879 unique emotional keywords from the TEL data set (Tocoglu & Alpkocak, 2019b), and annotate around 205,000 twitter messages automatically if those messages contained the TEL keywords. Different than other studies in Turkish, they apply CNN and RNN algorithms for classification of the six basic emotion categories, in addition to conventional learning methods. With the highest accuracy of 74% attained from CNN classifier, their proposed approach performs better than state-of-the-art methodologies. However, they train their own embedding layer with a fully connected network, and do not experiment on the alternative word embedding techniques like GloVe, ConceptNet or BERT. As mentioned above, more recent studies as (Ucan et al., 2022) or (Uymaz & Metin, 2023) followed Tocoglu et al.’s approach in addition with integrating and evaluating pre-trained word embedding layers in deep language models.

Even though a lexicon approach is used in (Tocoglu et al., 2019) by the use of annotated text instances in (Tocoglu & Alpkocak, 2019b), since language changes according to culture, time, and context, a more generalized approach is required for extracting emotions from Turkish text. Moreover, previous work on Turkish depend mostly on annotated data, which is expensive and hard to acquire, and natural annotations like hashtags, emojis, or the images are neglected. However, by contributing these annotations, data sets can be enlarged to come up with better results in text classification. This study intends to advance past research by not only providing a benchmark data set of emotions from the context of branding communication, but

also by assessing the state-of-the-art methods for emotion extraction used in earlier studies for Turkish language.

3.2 Methodology

Detailed flow of the research design which is also the proposed framework for automatically extracting emotions from textual data is shown in Figure 3.1. Following data collection, similar steps for extracting basic emotions from textual content described in Section 3.1 have been implemented and evaluated regarding brand and consumer communication on Twitter. In simple terms these steps are data boosting & auto annotation of Twitter messages by co-usage of emotional seed words from English and Turkish lexicons and emojis, pre-processing and vectorizing text, and finally evaluation of supervised classifier methods on the accumulated data in order to extract the basic emotions of happiness, sadness, anger, surprise, disgust, fear including the neutral emotions.

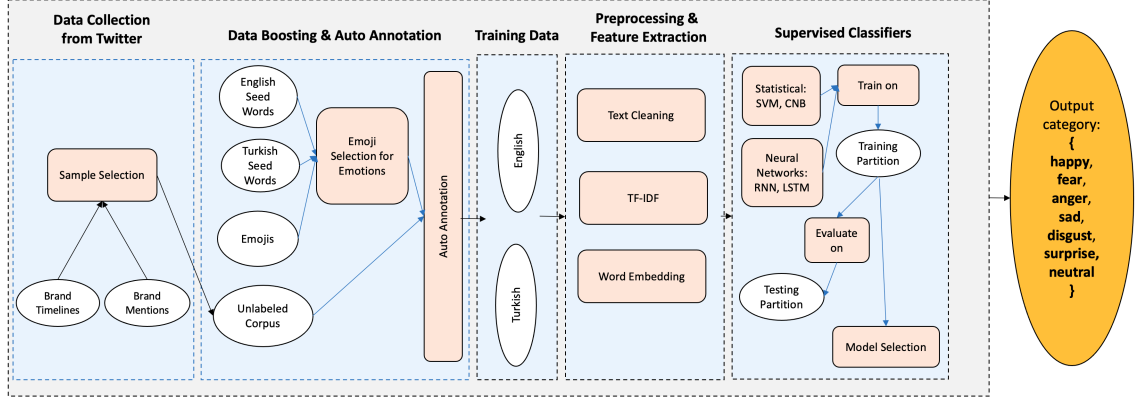


Figure 3.1 Research design

The majority of the state-of-the-art approaches proposed in the prior literature are English-based, and their effectiveness was measured using benchmark data sets. However, the same methodology may yield different findings in languages with notably different structural features, such as Turkish. Unlike English, Turkish is an agglutinative language having meaning of words changed when a suffix is added (Can, Kocberber, Balcik, Kaynak, Ocalan & Vursavas, 2008). Another distinctive character of the language is being a free word order language having a basic word order of subject-object-verb, where meaning does not get effected by changing the

order (Oflazer, 1994). Additionally, the emotional meaning of words may transmit different signals in various contexts such as sports and entertainment through time, where a word might represent opposite emotions. Described in more detail in the following sections, the aim of the research design is to suggest and discuss a flexible approach for extracting emotions from text in a variety of languages, contexts and time.

3.2.1 Generating the Auto Annotated Data Set

In order for proposing methods and curating an auto annotated data set for both Turkish and English languages using the Twitter Academic Search API 2.0, a massive amount of data from Twitter was collected. All brands of Fortune 500 firms in 2021 with Twitter accounts in both English and Turkish were chosen as a sample for specifically exploring the dynamics in branding communication. After a thorough search, 33 companies were selected (see Table 1.1) from the fast food, consumer packaged goods (FMCG), technology, automotive, apparel, retail, financial services (FinTech), and logistics (LogTech) sectors.

Using the chosen brand names and accounts as keywords, all tweets related from June 2016 to June 2021 were scraped. The messages were downloaded Java programming language and indexed in Elasticsearch for faster querying. A total of 8,101,034 tweets in English and 735,459 tweets in Turkish have been collected. Compared to the 245,422 brand posts and 490,037 consumer replies found in Turkish messages, 2,708,874 brand posts and 5,392,160 consumer replies were found in English messages.

To create supervised classifiers, textual data sources that have been labeled by humans based on their emotional content are required. These resources are both expensive and time demanding to create, and they are scarce and context specific for any language. Furthermore, studies have shown that emojis such as smileys, furious faces, and sobbing faces are used as strong indications of users' emotions and moods (Go, Bhayani & Huang, 2009; Liu et al., 2012; Read, 2005). In simple terms, by counting the co-usage of emojis and emotional words, first emojis representing the 6 basic emotions and neutrality are selected. Then, by filtering the twitter messages according to their emoji content, a pseudo annotated data set is generated for further training and evaluating the performances of emotion classification algorithms.

The detailed process of automatically annotated data accumulation is shown in

Figure 3.2. The process is applied for both languages separately. Inspired by (Riloff, 1996; Tang et al., 2013) and more recently by (Song et al., 2015) a co-frequency matrix of emotional seed words and emojis is generated. The co-frequency matrix is two dimensional where the first dimension contains the emotional words from specified lexicon, and the second dimension contains the emojis in the data set. Each value in the matrix displays the number of co-usage of an emotional word and an emoji in a Tweet. Then, by using the ranking function as shown in Equation 5 an importance score for each emoji according to the 6 basic emotions and neutrality is calculated.

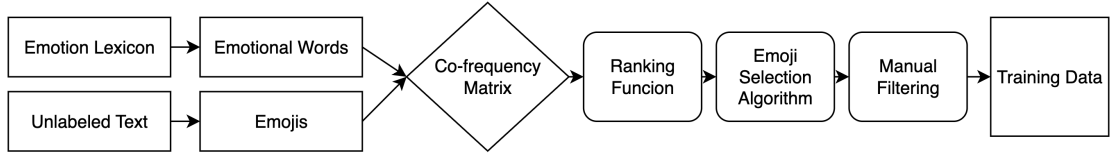


Figure 3.2 Data accumulation steps

$$(3.1) \quad Importance_i(e_j) = \frac{\sum_k co-frequency(e_j, ew_{ik})}{\sum_k \sum_i co-frequency(e_j, ew_{ik})} \log_{10} frequency(e_j)$$

In Equation 5, importance of an emoji j (e_j) for emotion category i is calculated. It is the $\log_{10}$ scaled version of Equation 3. ew_{ik} denotes the emotional word for emotion category i at index k . The first multiplier is the conditional probability observing an emoji with words in a certain emotion category given that the emoji is observed among all emotions, same as Equation 3 in Chapter 2. In brief, it shows how relevant an emoji is according to an emotion category. This probability is calculated by counting co-occurrences of a certain emoticon with emotional words for a certain emotion, and dividing by the count of co-occurrences of the emoticon with all emotional words. The second part of the multiplier is for decreasing the importance of less popular emoticons used rarely, and increasing the importance of frequently observed ones, by log scaling the frequency of the emoji in the data set.

This methodology is also called as “domain adaptation”, where transferring knowledge across domains where data distributions may be different (Li, Pan, Jin, Yang & Zhu, 2012). In our case, one domain is the emotional words developed in previous research as emotion lexicons, the other domain is the emoticons in the brands Twitter messages and replies. As in the examples cited above, this methodology has been applied vastly in English, whereas an application for Turkish language is missing.

Regarding the selection of the emotional seed words, TEL for Turkish (Tocoglu & Alpkocak, 2019a), and NRC Emotion Lexicon (Mohammad & Turney, Mohammad & Turney, 2013) for English are selected. Both lexicons are created by crowd sourcing, and in both emotional words are scored according to their importance for each emotion category. Top 90 scored words from each lexicon for each emotion category (happy, sad, anger, disgust, fear, surprise) are selected as seed words for importance calculation of emoticons. In Table 3.1, top 10 emoticons sorted according to their decreasing importance for each emotion category are listed.

Table 3.1 The most important 10 emojis by each emotion category

English						Turkish					
happy	surprise	sad	anger	fear	disgust	happy	surprise	sad	anger	fear	disgust
❤️	👉	⬆️	👈	👊	💩	💙	😊	😞	📅	😊	🚧
💜	📺	😭	👈	😬	👤	😬	📺	🕶️	💙	😊	🌫️
😊	📺	💤	👈	🚫	👤	🔥	👤	😞	😡	🐍	🐼
👏	📺	😞	👈	👊	👤	🕯️	🚚	📺	🇬🇷	👤	🐱
👤	🔗	😞	👤	👊	👤	💜	😬	🇹🇼	😬	🎬	❤️
👍	🔵	😞	👈	👤	👤	👤	👤	👤	😡	👤	🔔
👊	👉	😞	👉	?	👤	💪	🐼	👤	😞	👤	♣️
❤️	🚦	😊	👈	👤	👤	👤	👤	👤	😞	😞	👤
👤	👤	👤	👤	👤	👤	❤️	❤️	👤	😞	👤	👤
💙	🚚	👤	👤	👤	👤	👤	👤	👤	😞	👤	👤

When emojis are sorted by importance score in decreasing order, common emojis can be found among the sets of top n most important emojis of emotion categories. As an example, for Turkish ,right arrow emoji appears in the top ten most important emoji sets of both surprise and anger. Similarly for English, the smiling face appears in the top ten most important emoji sets of both surprise and fear. The reason is emojis being used interchangeably with words of different emotions. Words of negative emotions are less common in the branding messages resulting in probabilities for negative emotions to be relatively lower than positive ones. Also, multiplying the probabilities with $\log_{10}$ transformed frequencies, increase the importance of the most common emojis for all emotions. The smiling face emoji appears more than 17.000 times in the data set, naturally becoming important in terms of both positive and negative emotions.

For classifying the emojis more precisely, a secondary manual selection has been applied. First, emojis are assigned to the emotion with the highest relevance. Then, from the emotion category having the highest number of assigned emojis to the lowest, emojis are sorted according to their importance scores for each emotion and the most frequent and non-conflicting ones are re-assigned to their final emotion

category. Conflicting emojis, like a smiling face under fear emotion are not selected. Emojis never used together with an emotional word and not in "Smileys&Emotion" or "People&Body" categories are assigned to neutral emotion category. In Table 3.2, total number of selected emojis and in Table 3.3 the first 12 of the selected emojis for each emotion category are listed. Number of emotional words for each emotion variate for both languages, as in the case of happy words dominating the branding twitter data set. Thus, different number of non-conflicting emojis are selected for each emotion category.

Table 3.2 Number of selected emojis for each emotion category

	Happy	Surprise	Anger	Sad	Fear	Disgust	Neutral
English	22	10	9	13	19	10	29
Turkish	23	17	13	10	15	8	30

Table 3.3 Selected emojis for auto annotation

English							Turkish						
happy	surprise	sad	anger	fear	disgust	neutral	happy	surprise	sad	anger	fear	disgust	neutral
													®
													™

After the selection process, messages containing only one group the selected emojis are filtered and annotated with the emotion that the emojis in the message content represents. As an example, a message having 3 happy kind of emojis is annotated as 'happy', where as a message having 3 happy and 1 sad emoji is filtered out and not used in the final data set. As shown in Table 3.4, majority of the messages contain positive emotions for both languages. Total percentage of happy and surprise messages are 70% and 77% for English and Turkish respectively. Disgust and neutral messages are observed the least in Turkish, while it is fear for English. Amongst the negative emotions anger and sadness are the most observed ones in English with 7% and 9% proportions respectively. Anger dominates the negative emotional

observations for Turkish with 9% proportion. As the proportion of messages with disgust emotion is only 0.6%, to increase the accuracy of classifiers the Turkish data set is merged with TREMO data set prior to model evaluation.

Table 3.4 Number of messages in auto annotated Turkish and English data sets.

	Happy	Surprise	Sad	Anger	Fear	Disgust	Neutral	Total
English	483,339	36,333	63,459	5,079	13,769	22,504	21,475	691,671
	70%	5%	9%	7%	2%	3 %	3 %	
Turkish	64,561	5,369	2,173	7,849	2,930	524	919	84,325
	77%	6%	3%	9%	3%	1%	1%	
TREMO	4,416	1,980	5,942	15,051	7,005	2,798	0	19462
	23%	10%	31%	77%	36%	14%	0 %	

3.2.2 Text Pre-Processing

Both statistical and deep learning classifiers require textual data converted into a sort of numerical representation. First, words are converted to numbers according to the index of the dictionary of the text corpus, then sentences or messages are converted to a vector of these numbers. Prior to this operation, text is cleaned from meta information as hashtags, URL's and punctuation.

Before training any emotion classifiers, raw text of Twitter messages are cleaned from such meta information, while also keeping such information represented as short text in processed messages. 'RT' words in retweeted messages are removed. Mentions starting with "@" character are replaced with "REF" keyword. Hashtags starting with "#" character are replaced with "TAG" keyword. URLs, which are external links to web sites are replaced with "URL" keyword. Following removal of punctuation characters and unnecessary white space, text is converted to lower case letters. Moreover, as the data set has been annotated according to emoji content of message, emojis are removed from messages in order for the classifier models purely to be trained on the sequence of words.

Regarding statistical classifiers, stop words are removed and messages are converted to bag-of-words form where they are represented as vector of count of words of the dictionary created from corpus. Then, TF-IDF transformation is applied too all words in all messages, where common words in the corpus for all emotions are

weighed less while rare words are weighed more.

Regarding deep learning classifiers, a fixed sized dictionary of 15.000 words is created from the corpus and messages are vectorized similar to the application for statistical classifiers. Then, the vectors representing the messages are padded to size of 250 elements. The deep learning classifiers have an extra word embedding layer for representation of words in messages in vector form according to their distances in the corpus. In the experiments both a self trained and pre-trained versions of this word embedding layer has been used which will be described in the next section.

3.2.3 Emotion Classification Models

While earlier studies focused on evaluating more common and statistical classification models like MNB, CNB or SVM, following the advancement in parallel computing resources more recent studies analyze the contribution of complex models like RF or deep neural nets. As mentioned in Section 3.1 the most common neural net architectures are CNN, LSTM (Hochreiter & Schmidhuber, 1997) and lately pre-trained language models like BERT. In this study, deep neural net architectures such as Fully Connected Feed Forward Networks (FF) and LSTM have been evaluated alongside conventional classification methods such as MNB, CNB, SVM, and RF as baseline models. Furthermore, the impact of several BERT transformers on the resulting English and Turkish data sets is investigated. The deep neural network architectures are built as simple as possible for differentiating the impact of the method tested.

Figure 3.3 shows the network architectures used in experiments. FF is considered to be the baseline model for the other deep neural network architectures (*Figure 3.3i Left*). The network is composed of six layers. The input layer contains 250 input neurons followed by a trainable embedding layer converting each word in message to a 32 element fixed sized vector. The Flatten layer converts the the 3D vector of the previous step to a 1D vector. At this layer 20% drop out is applied before the dense layer consisting of 32 neurons and RELU activation function. The final layer consists of 7 neurons for each emotion category activated by Softmax (Bengio & others, 2009). All neurons in all layers of this architecture are fully connected to each other with weights to be calculated in training. Adam optimizer (Kingma & Ba, 2017) is used to fit the weights of the network with Categorical Cross Entropy loss function. The network has a total of 736,263 parameters.

In conventional methods as SVM, RF or even the more complex FF input text is represented as a set of words, then each text is represented by the count of these words. This is known as the bag-of-words strategy because the quantity of occurrences, rather than the sequence or order of words, is important in this approach. Even if TF-IDF transformation is applied, the context and the meaning of the words that might change in a sequence of words is neglected. RNN's are special network architectures that can process input sequences in variable lengths (Bengio et al. , 2009). They contain a layer consisting of feedback loops to maintain the memory over time. Each recurrent unit is only effected by its previous state which is basically a multiplication of the weights. Simple RNN architecture do not contain any memory units or gates. This causes the problem of vanishing gradients solved by GRU (Chung, Gulcehre, Cho & Bengio, 2014) or LSTM recurrent units (Hochreiter & Schmidhuber, 1997; LeCun, Bengio & Hinton, 2015). These units compromise gates that are utilised to sustain memory over long periods of time, making them an ideal classifier for textual data. The diagram in the center of Figure 3.3 shows the RNN architecture used in this study. It is composed of six layers having 250 neurons in its input layer. Similar to FF, input is followed by embedding layer converting each word in message to a 32 element fixed sized vector. LSTM recurrent layer consists 32 neurons activated by hyperbolic tangent activation function. Before and after the LSTM later 20% drop out has been applied. The final layer contains 7 neurons for each emotion category activated by Softmax. Adam optimizer has been used to fit the weights of the network with Categorical Cross Entropy loss function. The network has a total of 488,551 trainable parameters.

Regarding the experiments on deep networks with pre-trained transformers, Multilingual BERT (MBERT) (Devlin, Chang, Lee & Toutanova, 2018a), DistilBert and BERTurk have been fine-tuned on brands data set and evaluated. MBERT is a pre-trained on 102 languages with from Wikipedia corpus, so it is available for both Turkish and English. DistilBert, trained on the same Wikipedia corpus, is simplified distilled version of the base BERT model (Sanh, Debut, Chaumond & Wolf, 2019) with the objective of faster fine tuning times. Similarly trained on MBERT architecture on Wikipedia and translated texts from web, BERTurk (Schweter, 2020) is a community driven model specially for Turkish language. The pre-trained transformers are used to create the word vectors for the embedding layer of a five level network architecture. Shown in Figure 3.3 Two input layers containing 127 neurons each, one for word indexes and another for the attention mask are followed by a post BERT dropout layer to reduce over-fitting and improve generalization. A dense layer of 512 neurons with Relu (Bengio et al. , 2009) activation, and finally a dense layer of 7 neurons with Softmax activation function completes the

network component sequence. Including the pre-trained embedding layer, MBERT has 167,753,735, BERTurk has 111,014,663 and DistilBert has 109,879,559 parameters. Same as previously described networks, Adam optimizer has been used to fit the weights of the network with Categorical Cross Entropy loss function.

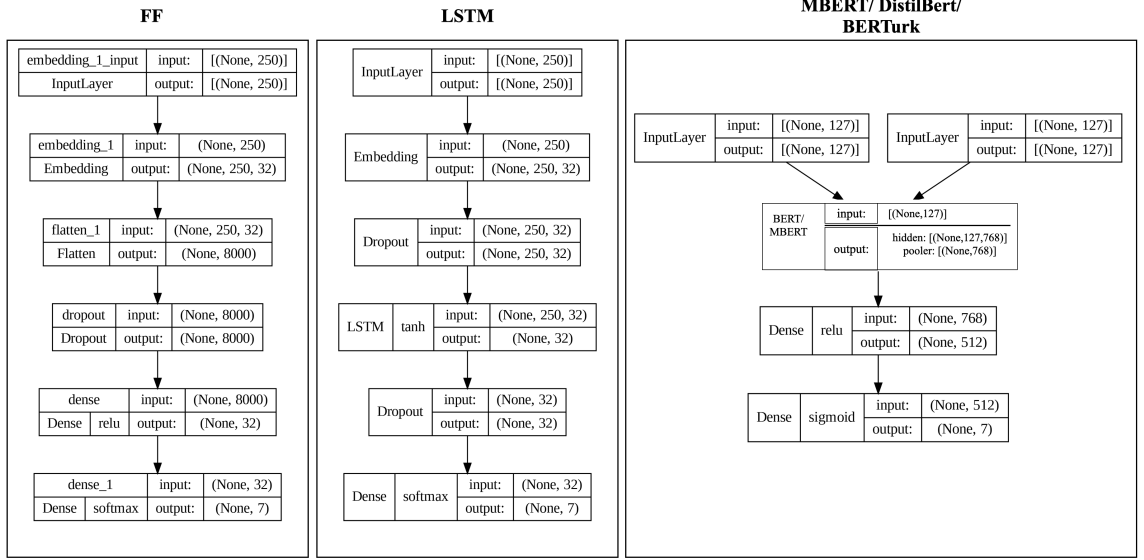


Figure 3.3 Deep neural net architectures used in experiments

Regarding the deep neural networks used in the experiments in this study, a subset of best performing architectures applied in (Tocoglu et al., 2019) and (Ucan et al., 2022) have been selected and modified.

3.2.4 Performance Metrics

The performance of the evaluated models will be reported as Precision (P), Recall (R), F-Score (F) and Accuracy (Acc). Precision reports the proportion of the relevant retrieved items, where Recall reports the proportion of relevant items retrieved among all relevant items. F-score is the harmonic mean Precision and Recall. Relevant items may be considered as true observations belonging to an emotion class. Accuracy is the proportion of correctly observed observations from all emotion classes. Acc, P, R and F are shown in equations 6-9.

$$(3.2) \quad Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

$$(3.3) \quad Precision(P) = \frac{TP}{TP + FP}$$

$$(3.4) \quad Recall(R) = \frac{TP}{TP + FN}$$

$$(3.5) \quad F-Score(F) = \frac{2TP}{2TP + FP + FN}$$

TP, TN, FP, FN denotes true positives, true negatives, false positives and false negatives. P, R and F are quantified per emotion class. As an example, precision or recall for happy emotion the positive class is considered to be happy messages, where the negative class is all the others.

3.3 Experimentation and Results

In order to validate the automatically annotated data set and complete the proposed framework for automatically extract emotions from textual data of any language or context, a set of experiments were conducted using both the conventional methods of MNB, CNB, SVM, RF and also the deep learning algorithms as FF, LSTM and BERT transformations. The accumulated data sets for both languages are partitioned into 80% training and 20% test sets. As the final training data is imbalanced, the class proportions are kept the same for training and test sets. The evaluation of the performances of the models are done and reported on the test partition separately for Turkish and English data sets.

Hyperparameters of all models are fine-tuned. MNB, CNB and RF classifiers are tested using Sklearn library in Python environment. For MNB and CNB, Laplace smoothing parameter is selected as 1.0. Regarding SVM classifier, a radial basis kernel function was used to fit the model with regularization parameter of 1.0 and cache size of 200. For RF, the model is build on 100 unpruned trees. Deep learning algorithms are run on Tensorflow backend using TPU environment provided by

Google Colab, and the algorithms are fed with batches of 256 observations.

Accuracy scores of evaluated models are shown in Figure 3.4. In Table 3.5 accuracy and F-scores of models are listed. For all models F scores are lower than accuracy, which shows that the data set is imbalanced and the models cannot predict every emotion class with the same accuracy. Thus, when evaluating the models, F-scores should also be considered as a performance indicator. MNB and CNB have the lowest accuracy and F scores compared to other models. While MNB is generally more accurate than CNB, it does not perform well on the rare classes as fear with 0.0 precision and recall on the test data. CNB is slightly better on the rare classes being able to predict several correct observations for fear and disgust. More complex models as SVM, RF or DL variations as FF, LSTM or MBERT perform significantly better on accuracy as well as F scores regarding both languages.

The performances of SVM and RF classifiers are also tested as baseline models similarly done in previous literature. Similar accuracy scores are observed for both SVM and RF (Turkish SVM: 0.81, RF: 0.81; English SVM: 0.80, RF: 0.79) higher than CNB and MNB. However on a single CPU configuration these models take too long to fit. While RF algorithm converges much faster than SVM in 8 minutes for Turkish and in around 45 minutes for English, SVM being problematic in large data sets converges in 48 minutes on Turkish data. It takes several hours for the SVM algorithm to complete execution for an optimized solution for the English data set. Even the accuracy scores of these models are higher than CNB and MNB, their F scores are significantly lower than 0.80, where SVM cannot predict any fear classes on the test data set.

Table 3.5 Accuracy and F-Scores of evaluated models

	Turkish		English	
	Acc	F	Acc	F
CNB	0.72	0.73	0.74	0.74
MNB	0.78	0.73	0.78	0.72
SVM	0.81	0.78	0.80	0.75
RF	0.81	0.78	0.79	0.74
FF	0.80	0.79	0.82	0.80
LSTM	0.80	0.79	0.83	0.81
MBERT	0.82	0.81	0.85	0.84
DistilBert	0.81	0.79	0.83	0.82
BERTurk	0.84	0.83	0.75	0.69

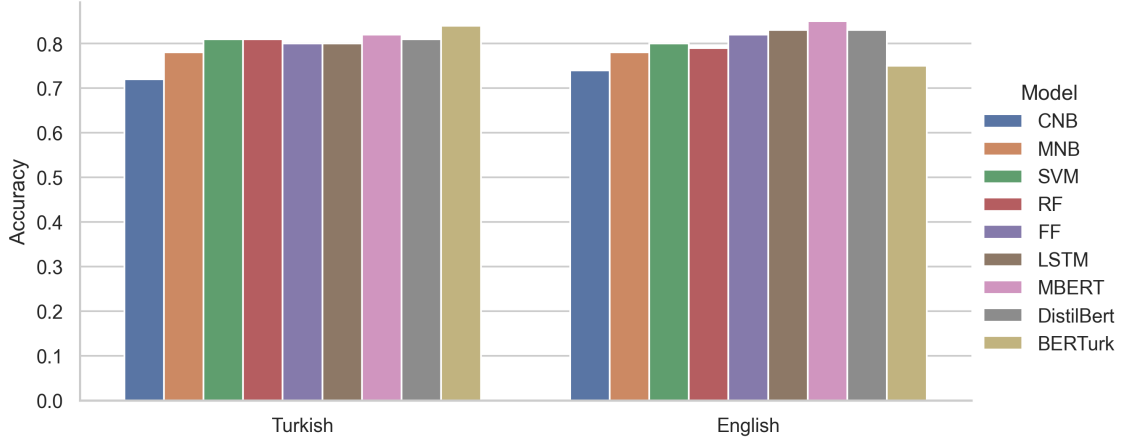


Figure 3.4 Accuracy scores of evaluated models

As also seen in Table 3.5 deep learning classifiers outperform the others when accuracy and F scores are considered together. Regarding English MBERT, trained on multi-language Wikipedia corpus, performs the best (Acc: 0.85, F: 0.84), while for Turkish, BERTurk, having the same architecture with MBERT but trained on Turkish Wikipedia corpus, has the highest scores (Acc: 0.84, F: 0.83). The detailed evaluation results of the deep learning classifiers and the RF classifier are listed in Table 3.6 for each emotion class. As noted, in 3.6 performance metrics reported are Precision (P), Recall (R) and F-Score (F). *Precision* is calculated as the sum of true predictions across all classes divided by the sum of true and false predictions across all classes. *Recall* is calculated as the number of correct positive predictions made out of all true predictions that could have been made across all classes. Finally, F-score is the harmonic mean of precision and recall. Support is the number of actual occurrences of the class in the dataset.

The performance of deep learning classifiers by each emotion depends on the algorithms to be able to train on sufficient amount of data. "Happy" messages dominate the training data. Thus, for both languages all algorithms produce over 0.85 F-scores for happy emotion category. For English, the second category best classified is surprise and for Turkish it is the neutral emotion category. Although, anger category seems to have enough observations in both languages, classifiers achieve F-scores from 0.50 to 0.65, which is relatively lower when compared to other emotion categories. This situation is more likely to be observed when the messages in a category contains noise, that is conflicting words from other emotions existing in sentences.

BERTurk, best performing classifier on Turkish data set, performs the worst on surprise category followed by fear (surprise F: 0.57, fear F: 0.62), which still outperforms the other classifiers. MBERT as the best performing classifier on English data set has the lowest F scores predicting fear and disgust categories (fear F: 0.16, disgust

F: 0.37). Overall prediction scores of the MBERT for English data set are observed to be lower than the BERTurk for Turkish data set. The reason for this is that the English data set is substantially more imbalanced than the Turkish data set, while having six times the amount of data. Fear, disgust, and surprise are appear less frequently in English data than in Turkish data, thus performance degrades when compared to Turkish.

Table 3.6 Precision, Recall and F Scores of Models

	Turkish																		
	RF			FF			LSTM			MBERT			BERTurk			DistilBert			support
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	
anger	0.76	0.37	0.50	0.57	0.55	0.56	0.54	0.62	0.58	0.72	0.50	0.59	0.67	0.64	0.65	0.56	0.59	0.58	2231
disgust	0.91	0.77	0.83	0.87	0.78	0.82	0.90	0.78	0.84	0.93	0.81	0.87	0.97	0.83	0.89	0.93	0.79	0.85	664
fear	0.82	0.48	0.60	0.88	0.43	0.57	0.80	0.48	0.60	0.80	0.50	0.62	0.68	0.56	0.62	0.94	0.45	0.61	1256
happy	0.80	0.98	0.88	0.83	0.95	0.88	0.85	0.92	0.89	0.83	0.96	0.89	0.86	0.95	0.91	0.83	0.94	0.88	13796
neutral	0.98	0.86	0.92	0.98	0.74	0.85	0.94	0.74	0.83	0.91	0.92	0.92	0.93	0.92	0.93	0.87	0.92	0.89	184
sad	0.86	0.54	0.66	0.83	0.51	0.64	0.82	0.54	0.65	0.88	0.59	0.70	0.87	0.62	0.72	0.92	0.52	0.67	1157
surprise	0.87	0.35	0.50	0.73	0.36	0.48	0.62	0.43	0.51	0.74	0.42	0.54	0.80	0.44	0.57	0.74	0.37	0.49	1470
weighted avg	0.81	0.81	0.78	0.80	0.80	0.79	0.80	0.80	0.79	0.81	0.82	0.81	0.83	0.84	0.83	0.81	0.81	0.79	
	English																		
	RF			FF			LSTM			MBERT			BERTurk			DistilBert			support
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	
anger	0.64	0.27	0.38	0.57	0.51	0.54	0.58	0.52	0.55	0.58	0.67	0.62	0.55	0.12	0.19	0.53	0.63	0.57	10158
disgust	0.60	0.06	0.12	0.44	0.29	0.35	0.46	0.30	0.37	0.49	0.30	0.37	0.28	0.01	0.01	0.42	0.30	0.35	4501
fear	0.43	0.02	0.04	1.00	0.00	0.00	0.50	0.04	0.08	0.40	0.10	0.16	0.00	0.00	0.00	0.35	0.08	0.13	2754
happy	0.79	0.99	0.88	0.86	0.97	0.91	0.88	0.96	0.92	0.92	0.95	0.94	0.76	0.99	0.86	0.90	0.95	0.92	96668
neutral	0.98	0.52	0.68	0.88	0.55	0.67	0.78	0.61	0.69	0.76	0.65	0.70	0.95	0.41	0.57	0.81	0.60	0.68	4295
sad	0.73	0.22	0.34	0.61	0.46	0.53	0.58	0.55	0.56	0.61	0.65	0.63	0.57	0.15	0.24	0.59	0.57	0.58	12692
surprise	0.99	0.74	0.85	0.97	0.74	0.84	0.94	0.76	0.84	0.94	0.78	0.85	0.96	0.55	0.70	0.98	0.74	0.84	7267
weighted avg	0.78	0.79	0.74	0.81	0.82	0.80	0.81	0.83	0.81	0.84	0.85	0.84	0.71	0.76	0.69	0.82	0.83	0.82	

As the accuracy and F scores of deep neural networks evaluated are significantly higher, more analysis has been made regarding their training process in order to propose the classification algorithm with the highest performance and lowest training efforts possible for the auto annotation framework. In addition to the requirement of huge amounts of data to train these networks, the training algorithm must run more than once on the training data. In each run, also called as epochs, parameters of the network are re-calculated, and the network learns the training data in a better sense. However, if the network is run more than a certain number of epochs it over-fits the training data, and this may cause stable or even decreasing accuracy score on the test data. The objective is training the network in least number of epochs achieving the highest accuracy on test data. Consequently FF, LSTM, MBERT, BERTurk and DistilBert classifiers are run for 10 epochs separately for English and Turkish data. Training accuracy scores, validation accuracy scores on test data for each epoch are shown in Figure 3.5. Regarding all models, it is clear that the training accuracy of the models tend to increase over 0.90, on the other hand the validation accuracy scores first tend to increase then get stable or start to decrease. Also, the sufficient number of epochs, training time in minutes and the accuracy scores on the test data are shown in Table 3.7, ranked by their complexity, with the RF model as

the baseline.

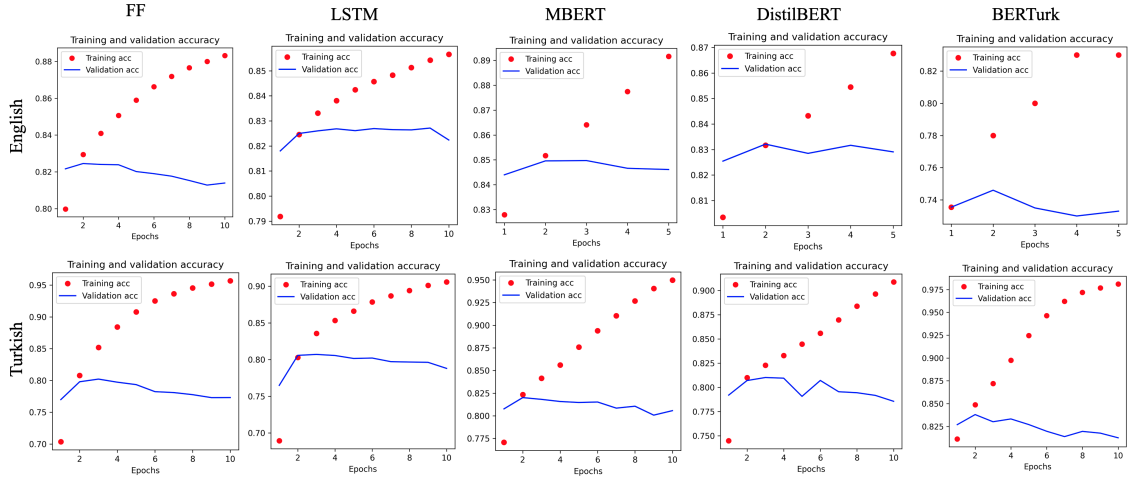


Figure 3.5 Accuracy scores by epoch on training and test data sets

Table 3.7 Training epochs and performance scores of deep learning models

	Turkish				English			
	Acc	F	Epochs	Time (mins)	Acc	F	Epochs	Time (mins)
RF	0.81	0.78	-	8.00	0.79	0.74	-	40.00
FF	0.80	0.79	3	0.35	0.82	0.80	2	1.00
LSTM	0.80	0.79	3	0.82	0.83	0.81	2	3.00
DistilBert	0.81	0.79	3	4.00	0.83	0.82	2	16.00
MBERT	0.82	0.81	2	5.00	0.85	0.84	2	20.50
BERTurk	0.84	0.83	2	5.00	0.76	0.69	3	8.70

As complexity of the networks represented by the number of trainable parameters increase, time required to run enough each epoch also tend to increase. While all networks converge in 2 to 3 epochs, more complex BERT networks complete in a minimum of 5 minutes for Turkish and 16 minutes for English. As an example LSTM has 488,551 trainable parameters and MBERT has 111,014,66 parameter. LSTM in English completes 2 epochs in 3 minutes, while MBERT completes the same number of epochs in 20.5 minutes. DistilBert, smaller than MBERT and larger than LSTM in parameter size, behaves similarly taking less time to complete than MBERT however much longer than LSTM. Moreover, along with complexity accuracy and F scores also tend to increase. For example LSTM can achieve up to accuracy of 0.80 in Turkish and 0.81 in English, MBERT outperform all of the simpler architectures having accuracy scores of 0.82 in Turkish and 0.84 in English. BERTurk, having same number of parameters with MBERT but trained on a Turkish corpus, is the most accurate model for Turkish (Acc: 0.84, F: 0.83) but naturally the worst performing one on the English data set.

The results about the performances of the models evaluated in experiments are compatible with previous literature. As also shown by (Tocoglu et al., 2019), deep neural networks outperform conventional classifiers as MNB or CNB and even RF on Turkish Twitter messages. Moreover, Recurrent Neural Nets with LSTM layers perform slightly better than FF or Convolutional networks. In (Ucan et al., 2022)’s study, higher accuracy scores are reported when BERT transformers are used in neural network architectures. Their results show that BERTurk on Turkish and MBERT on English data sets have the highest accuracy scores when compared to LSTM or SVM, similar to this study. However, their results are reported on more common data sets as TREMO or TURTED and has not yet been re-validated until this work. Moreover, trained on a similar data set consisting of Twitter messages as TURTED, the best performing architectures of BERTurk and MBERT of this study achieve significantly higher accuracy scores than the architectures proposed by (Tocoglu et al., 2019).

As noted, the objective of this study is not only evaluating classification algorithms for Turkish comparatively with English as in previous work, but also to propose a complete framework for automatically extracting emotions from text messages dependently on context, time and language. Thus, the accuracy scores reported over 0.80 also validates the proposed approach of generating labeled data for emotions. Also in line with earlier findings, the base BERT transformer (Devlin et al., 2018a) pre-trained on the target language, MBERT and BERTurk in our case, achieves substantially the best performance for English and Turkish. As the final layer of the proposed framework, fine-tuning deep neural nets with BERT transformers are suggested for context and language specific emotion extraction tasks.

3.4 Conclusion

Following an extensive background of research on emotion classification from the perspectives of psychology and computer science, this study aimed at building a framework for articulation of annotated textual data and emotion extraction on any language or context regardless of time. Turkish and English languages were the focus together with a curated data set of social media messages in branding communication. While English benchmark data sets required to build classifiers, annotated by emotions are common, they are sparse for more specific languages like Turkish. Moreover, the available data sets even in English or Turkish are not context

specific. Besides filling the gap in data articulation, state-of-the-art classifiers were evaluated not only in terms of their performance but also training time in order to complete the structure of the proposed framework.

Aiming to seek answer to research question *RQ: How can emotions be automatically extracted from brands' and consumers' social media messages in different languages and contexts?* key findings can be summarized as:

- Reported results shows that a language and context independent flexible method to articulate textual data annotated by 6 basic emotions (*happiness, surprise, sadness, anger, disgust, fear*) and neutrality is possible through emotional signals together produced by emojis and emotional words. While the concentration was on brand management, the same approach will be useful to collect and annotate social media messages from specific communication contexts as sports, politics or entertainment.
- Performance of conventional and deep learning classifiers were also evaluated and the findings were consistent with previous literature. Deep learning classifiers outperform conventional classifiers as MNB, CNB, SVM or RF in extracting emotions. Their performance is positively correlated with their complexity represented by the number of trainable parameters. However, performance comes with the cost of longer training time of the networks.
- Multilingual BERT variations, trained specifically on the target language achieve the highest accuracy for textual emotion classification, and consequently have the longest training time. The proposed architecture for Turkish (BERTurk) performed 10% higher than the LSTM architecture proposed by (Tocoglu et al., 2019) built similarly on generated data TURTED. The proposed architecture for English (MBERT) performed 7% higher than the BERT architecture proposed by (Ucan et al., 2022), however their data was not auto generated.

The theoretical contribution of this study is how information from other domains such as lexicons can be used to generate annotated data has been shown along with the validation of earlier algorithms proposed for emotion classification. The managerial contribution is the end-to-end framework to automatically extract emotions, applicable to any context such as customer relationship management, brand management or public opinion mining. Future studies should employ more features regarding the emotion classification task from textual resources, and apply analysis one more specific contexts in alternative languages like Spanish, Chinese or Arabic. Furthermore, future research should also examine the suggested framework's use

for various tasks such as question answering systems, topic extraction, or emotion analysis.



4. UNDERSTANDING EMOTIONS AND BRAND POST

POPULARITY IN SOCIAL MEDIA COMMUNICATION ACROSS ENGLISH AND TURKISH LANGUAGES

In today's hyper-connected digital landscape, brands are continuously striving to gain attention and establish meaningful connections with their target audience to promote their products and services (Lee & Kahle, 2016). Social media platforms have emerged as powerful tools for brand communication, offering a unique space where consumers can engage, share opinions, form communities, and exchange information (Delia & Armstrong, 2015). Consumers develop an attitude toward the brand and form relationships by interacting with it on social media (Hudson et al., 2015). Ultimately, a brand may become irreplaceable to consumers (Ekinci, Yoon & Oppewal, 2005) due to an established strong relationship. Successful online marketing efforts are known to influence consumer engagement both in spreading positive word-of-mouth and recommending the brand of concern (Edelman, 2010). With the unprecedented growth of user-generated content, understanding the factors that contribute to higher user engagement regarding a brand's social media post can uncover the applicable lessons for better business.

Targeting likes, comments, and shares as indicators of brand post popularity and engagement (Schultz, 2017), numerous studies have been conducted on measuring the effects of factors related with a brand's social media post's vivid, practical, interactive, informative and entertaining characteristics. Despite the growing research on social media posts' global attractiveness to consumers, three issues withstand to be further analyzed:

First, among these characteristics the impact of emotional appeal, representing the feelings to be simulated by the post, was not adequately investigated and, in most of the cases simplified as positive and negative appeal. Emotions play an important role in the consumption experience (Laverie et al., 1993) and impact word-of-mouth communications and brand loyalty (Crosby & Johnson, 2007). Thus, expanding the dimensions of emotions and understand the effects of more specific appeals such as happiness, sadness, anger or fear is required to

manage the tone in social media messages.

Second, studies on social media posts' global attractiveness to consumers, studies on the effect of cultural differences on consumer behavior regarding branding communication is still inadequate. Cultural characteristics of individualism/collectivism, uncertainty avoidance, power distance and long term orientation (Hofstede, 2011) are shown to be useful in understanding the impact of social media campaigns across cultures (Lin et al., 2017). Depending on the cultural perspective, consumers might interpret the brand messages differently. Thus, culture appears as an important factor to consider by managers to tailor their social media campaigns targeted to consumers belonging to specific cultural groups.

Third, regarding the effects of factors, conflicting findings were reported. Previous studies on Brand Post Popularity have employed dominantly Ordinary Least Squares (OLS) regression to model post interactions and derive insights. However, correct inference using OLS regression parameter estimates relies on certain assumptions. These assumptions include a linear relationship between the dependent and independent variables, no correlation among the independent variables, and normally distributed error terms with constant variance (Berry, 1993). While hard to validate, slight violations of these assumptions will lead to misleading estimates and statistics. Moreover, the proposed models may not have the same explanatory power on data other than the sample of concern. Thus, a more flexible approach to explore the impacting features and their effects is required for better generalization.

Considering these issues, this study delves into the dynamics of brand post popularity, investigating the multifaceted elements that drive consumer engagement and ultimately shape the success of brand communication efforts, answering the following research question:

RQ: *What are the most influential factors, including emotions, explaining brand post popularity across cultures?*

By exploring the underlying mechanisms and analyzing the effects of not only vividness, practicality, interactivity, information and entertainment features, but also discrete emotions of happiness, surprise, anger, fear, disgust and sadness addressing the first issue. In order to fulfil the gap of cultural perspective in earlier studies, regarding the *second* issue, Twitter posts of multinational brands' in English language representing markets from the more individualistic cultures, and in Turkish language representing markets from less individualistic, more collectivist cultures

are sampled to evaluate the cross cultural differences related with the features under concern. The third issue is addressed by evaluating multiple regression models under different target variable distributions and following the selection of the best fitting one, by running Monte Carlo simulations on random subset of features and data to eliminate the effects of violation of regression assumptions.

Consequently the main theoretical and practical contributions of this study are as follows:

- 1) Proposing a bootstrap approach for scoring and evaluating the effects of factors related with brand post popularity.
- 2) Understanding the effect of culture on these factors by building and comparing statistical models developed separately for English and Turkish datasets. This study is the first one concentrating on the cultural influence regarding brand post popularity.
- 3) As well as evaluating the cross cultural differences on vividness, practicality, interactivity, information and entertainment features of brand Twitter posts, the emotional content of a message as an additional factor of brand post popularity is integrated. This is the first study on analyzing how emotions transmitted by brands in their social media posts influence the success of brands' social media messages comparatively across cultures.

This study aims to shed light on the intricate relationship between brands, consumers, and the digital realm, providing valuable insights for marketers and researchers alike. The rest of the chapter is organized as follows: In Section 4.1 following a detailed coverage of previous work on "Brand Post Popularity in Social Media", the "Conceptual Model" that this study inspired from is discussed. Section 4.1 is finalized by a summary of related work regarding "Random Subspace Regression with Bootstrap" approaches applied on other domains and problems. In Section 4.2, "Materials & Methods" of this study under sub sections of "Data", "Operationalization of Variables", "Models", and "Method" is covered. Findings are reported and discussed in Section 4.3 followed by concluding remarks and future research directions in Section 4.4.

4.1 Literature Review

4.1.1 Brand Post Popularity in Social Media

The spread of brands' social media advertising efforts attracted the attention of research community and consequently numerous studies have been conducted on this topic. The first study was done by (De Vries et al., 2012) employing methods from earlier work on the effectiveness of banner advertising. Differences and similarities exist among these work in the context of data studied, operationalization of dependent/ independent variables and the conceptual model employed to explain brand post popularity. However, a cultural perspective on the dynamics of the variables is mostly neglected. Majority of the previous work has been done on English data sets. Moreover, contradicting findings were reported on the effect of chosen variables, which shows the need of validation using different approaches on alternative contexts.

Regarding the context of previous work, Facebook was the most studied platform, followed by Instagram (Geurin & Burch, 2017; Mazloom, Rietveld, Rudinac, Worring & Van Dolen, 2016; Yu & Sun, 2019), Twitter the least (Zadeh & Sharda, 2014). While some of the work focused on a broader range of companies from industries as FMCG, clothing, retail or telecommunication altogether (Banerjee & Chua, 2019; Chua & Banerjee, 2015; De Vries et al., 2012; Karpinska-Krakowiak & Modlinski, 2020; Thongmak, 2015; Zadeh & Sharda, 2014), some only focused on a certain industry as fast food (Mazloom et al., 2016; Yu & Sun, 2019), electronics (Rahman, Suberamanian, Zanuuddin, Moghavvemi & Nasir, 2017), mobile phone brands (Kujur & Singh, 2016), fitness (Pinto & Yagnik, 2017) or retail (Antoniadis, Paltsoglou & Patoulidis, 2019).

Brand post popularity was operationalized by number of likes, comments and shares. While some of the previous work created multiple models to explain likes, comments and shares separately (Antoniadis et al., 2019; Banerjee & Chua, 2019; De Vries et al., 2012; Schultz, 2016,1), in (Karpinska-Krakowiak & Modlinski, 2020) a weighed combination was preferred as dependent variable. The common antecedents analyzed were vividness, interactivity, informativeness and entertaining content of social media posts. Vividness refers to the visual appeal of a message represented by images or videos in content. Interactivity of a message increase if it has mentions, links or hashtags. If the content of the message contains any kind of information about a book, place or event the post is regarded as informative. If the message entertains the users through fun content the post is regarded as entertaining. Practicality of a message (Lin et al., 2017) is related to its accessibility and readability by the users, represented by the time of the message (day, night, evening, working hours,

non-working hours), day of the message(working / non-working), and the length. In some of the previous work emotions were also incorporated to represent message appeal as an antecedent to brand post popularity (Karpinska-Krakowiak & Modlinski, 2020; Swani et al., 2013; Swani & Milne, 2017; Thongmak, 2015). For modelling the effect of variables content analysis (Karpinska-Krakowiak & Modlinski, 2020; Pinto & Yagnik, 2017), ANOVA (Geurin & Burch, 2017; Pinto & Yagnik, 2017; Yu & Sun, 2019) and OLS regression were the most common methods employed to explain the effects.

As the first study inspiring the others, (De Vries et al., 2012) worked on a sample of 355 Facebook messages ranging from food, accessories, alcoholic beverages to mobile phone brands collected over a 9 month period. Log transformed number of likes and comments as dependent variable was explained by OLS regression having vividness, interactivity, informativeness, position of the message and valence of comments as determinants, controlling for week day, post length and product category. Their results showed that staying at the top position of brand pages, messages popularity increased. Vividness, interactivity and share of positive comments in brand posts enhanced number of likes, while entertaining messages had negative effect and informative messages had no significant effect on level of interaction by users. In (Sabate, Berbegal-Mirabent, Cañabate & Lebherz, 2014), using a sample of 164 messages of Spanish travel agencies effect of richness (represented by number of images, videos and links), week day and day time were measured over separate linear models to predict number of likes and comments. Post length and messages posted during business hours were supported to be attracting interactions. However, in later studies post length was found to be negatively related to number of likes(Schultz, 2016), and from late afternoon till late night were more likely to increase interactions (Pinto & Yagnik, 2017).

Schultz analyzed the effect of vividness, interactivity, content, and publication timing, while also controlling for post length, number of fans, and industry (Schultz, 2017). They worked on a sample of Facebook posts of apparel and food retail brands collected over a 45 day period, modelled brand post popularity using OLS regression similar to (De Vries et al., 2012). According to their findings post interactivity positively effected social interactions, whereas vividness had mixed effects. The informative and entertaining content had negative effect on interactions if they addressed only a certain part of the target consumers. Time at the top of the page had positive end day of the week had no effect on interactions. Working on a more culture specific sample of 18 Facebook brand pages from Greece, (Antoniadis et al., 2019) examined the effect of vividness, positive and negative reactions, message length and week days on number of likes, shares and comments. Dependent vari-

ables were normalized by number of fans of the brands. According to their OLS regression results images, videos, post length and positive comments had positive effect on the post popularity while posting time had no significant impact. Different than earlier examples, (Banerjee & Chua, 2019) explained likes, comments and shares by a conceptual framework of four antecedent categories named presentation, brand awareness, engagement and temporal. Hierarchical regression was used on a broader sample of 10,000 posts from 50 Facebook Fan Pages to measure the effect in terms of the dimensions of these four categories. Their findings showed that media richness had positive and post length had negative impact over post interactions. Reflecting brand centrality, competitor comparison and corporate social responsibility in brand posts increased popularity. In contrast contests or deal provisions in posts weren't found to be useful to get the attention of consumers.

Although not empirically evaluated, in (Lin et al., 2017) authors argued that social media practices would not be equally effective in different regional markets, and made suggestions for managing brand post popularity from a cultural perspective based on earlier studies. According to their discussion, Hofstede's cultural dimensions theory of individualism/collectivism, uncertainty avoidance, power distance and long-term orientation (Hofstede et al., 2010) would be useful to adapt social media practices of brands' across cultures. The adaptation may be considered as applying different levels of vividness, interactivity, practicality, informativeness and entertaining content targeted for the regional audience. They based their suggestions on the earlier findings about different attitudes of consumers from different societies of cultural characteristics. As an example, individualism/collectivism characteristics of the society the consumers are from would likely effect their reliance on online word-of-mouth opinions and purchase decisions. According to the study made by (Chu & Choi, 2011), Chinese (collectivist) consumers tends to engage more in online information sharing activities while U.S consumers less. (Goodrich & De Mooij, 2014), analyzing online and offline purchase decisions across markets of 50 different countries, found that consumers from individualistic cultures as Australia or U.K. are less likely to use social media for purchase decisions than consumers from individualistic cultures. This work aims to fulfil the empirical gap in earlier research on how the common antecedents/features of vividness, interactivity, practicality, information, entertaining content and the emotional appeal of brand posts impact consumers' interactions across Turkish speaking (collectivist) and English speaking (individualistic) consumers.

4.1.2 Conceptual Framework

Based on the findings of earlier research, this study aims to explore the most impacting features /factors of popularity of brands' Twitter posts from a cross cultural perspective. Similar to the first model proposed by (De Vries et al., 2012) and slightly modified later by (Antoniadis et al., 2019; Schultz, 2016), the conceptual framework shown in Figure 4.1 displays the relation of the features/factors with the brand post popularity used to evaluate the features in this study. The effects of the features are not only evaluated by their effect sizes and statistical significance, but also by their variation across culture along English speaking (individualistic) and Turkish speaking (collectivist) markets. Brand post popularity as the dependent variable is measured in total number of likes, shares and comments and examined over six categories of factors which are vividness, interactivity, practicality, informative content, entertaining content and emotional appeal.

Vividness refers to what extend a post stimulates users' visual senses (De Vries et al., 2012; Schultz, 2017). A post containing images or videos are considered to be more vivid than posts with plain textual content. According to (De Vries et al., 2012), posts containing images or videos attracted users to like brand posts, whereas in (Banerjee & Chua, 2019) it was found that vivid posts increased all kinds of interactions (likes, shares, comments).

Interactivity of a post is related to what extent it makes the consumers take an action or be able to influence the communication. Links, hashtags, mentions, asking a question or voting for a poll are features that define how much interactive a post is. According to (De Vries et al., 2012) more interactivity related to higher levels of post interactions, however, in (Chu & Choi, 2011) the same effect was only valid for low and medium level interactivity in posts, and less interactions by consumers were observed for high levels of interactivity. The positive effect was also observed by (Schultz, 2016; Swani & Milne, 2017), and more recently (Yu & Sun, 2019) studying Instagram posts argued that hashtags specifically are helpful to increase interactions.

Practicality of a post refers to its accessibility and efficiency of delivering the brand message. Length of the textual content, time and day of the message, the duration of the post at the top of the brand page are some of the factors related to how much practical a post is to the consumers for readability. For example, (Sabate et al., 2014) found that early morning and late night posts got 20% higher engagement than regular hours in Facebook. In contrast, according to (Kujur & Singh, 2016) non-working hours had no effect on shares and comments, but negative effect on

number of likes. Contradicting findings were found regarding post submission day and length. While (Kujur & Singh, 2016) argued that posts on weekdays increase the likelihood of comments, in (Antoniadis et al., 2019) a negative effect was observed. Regarding the length of posts, in (Schultz, 2016) a negative relation was reported for likes, while according (Sabate et al., 2014) longer posts increased number of likes. More recent research showed that post length decreased interactions in terms of likes and shares having no relation with comments (Banerjee & Chua, 2019). Regarding the duration of posts at the top position of brand pages, (Banerjee & Chua, 2019) argued that in order to increase user interactions a low submission frequency was required.

Information in posts is related to what extent the content gives details about facts about alternative topics. The information may be about an event, a social activity, books, places or people. Although a positive effect was measured in (Mazloom et al., 2016), other research found no relation between the informative content and popularity of brand posts (Antoniadis et al., 2019; De Vries et al., 2012; Karpinska-Krakowiak & Modlinski, 2020).

A post is described as *entertaining* when it reflects fun, humour or excitement with its content. According to (De Vries et al., 2012), the entertaining content in brand messages are unrelated with the focus of the brand, thus, consumers pay less attention on them. In contrast, in (Mazloom et al., 2016) a positive relation, and moreover in (Karpinska-Krakowiak & Modlinski, 2020) moderation of humor on narrative content increasing post interactions was observed.

Emotional appeal represents which feelings are targeted to be stimulated by the post, and its effect was also investigated as antecedents to popularity of brand posts. Analyzing posts of 193 Fortune 500 brands' Facebook accounts, (Swani et al., 2013)'s results suggested that including emotional sentiments in brand posts increased the attention and interactions of users. Using a similar approach, analyzing 421 posts of fitness brands, (Pinto & Yagnik, 2017) showed that happiness, love, motivation and adventure were the common emotional appeals to make a brand post popular. More recently, (Karpinska-Krakowiak & Modlinski, 2020) examined the popularity of 600 messages by 100 brands and analyzed the interaction between humour and product category over categorized messaged as narrative and non-narrative. They demonstrated that emotional appeal communicated through brand posts in a story-like structure works well for highly involving product categories having a long purchase cycle, but does not work for fast-moving consumer products. One of the objectives of this study is examining the impact of emotions on post interactions. Thus, different than previous studies, message appeal is represented by 7 categories derived

from the textual content of posts: 6 basic emotions defined by (Ekman, 2005) of happiness, surprise, anger, sadness, disgust and fear in addition to neutrality.

In line with previous research, number of followers (De Vries et al., 2012; Kujur & Singh, 2016; Schultz, 2016,1) and industry of the brand posts (Antoniadis et al., 2019; Schultz, 2017) are controlled as they directly effect level of post interactions (Sabate et al., 2014). Moreover, in order to be able to evaluate the differences in consumer behaviour across English and Turkish markets, in this study separate models on English and Turkish data sets based on the conceptual model (Figure 4.1) are employed.

In studies as (De Vries et al., 2012; Kujur & Singh, 2016; Schultz, 2016) or (Schultz, 2017) multiple models were used to explain the effects of factors separately on number of likes, shares and comments. As an alternative dependent variable to represent all interactions, (Antoniadis et al., 2019) uses Engagement Ratio, which is sum of likes, shares and comments divided by the number of followers. (Karpinska-Kraskowiak & Modlinski, 2020) uses a single dependent variable called *Relative Interactivity Index*, a weighted measure of brand post popularity, capturing all interactions such a way that more effortful activities like comments and shares have higher weights. In order not to put forward complex analysis results, this study focus on *Total Number of Interactions* as dependent variable, which is basically the sum of likes, shares and comments.

Majority of previous studies on Brand Post Popularity model the conceptual model with OLS regression and infer from the parameter estimates of the resulting models. In order to be able to make correct inference using OLS regression parameter estimates, linear relation among dependent and independent variables, no correlation among independent variables, normally distributed error terms with constant variance are the required assumptions (Berry, 1993). Multicollinearity, defined as high correlation or linear dependency among the independent variables in a regression model, can complicate coefficient interpretation, model stability, and appropriate inference (Lindner, Puck & Verbeke, 2020). Moreover, when one of collinear variables is omitted from the model the effect of the other variable remaining in the model may change significantly, given both of them are effect the dependent variable. The omitted variable may even cause the coefficient of collinear variables change signs, from positive to negative or vice versa regarding the direction of the effect. Together with non linearity of data, the omitted variables cause heteroscedasticity, where the errors variate according to the variation in the dependent variable.

Previous work employing OLS in their analysis commonly confirmed the absence of multicollinearity by reporting Variance Inflation Factor (VIF) among exploratory

variables below score 5.0 or 10.0, and continued their analysis by testing their hypothesis on the estimates, confidence intervals and p-values regarding their effects on likes, shares and comments (Antoniadis et al., 2019; Banerjee & Chua, 2019; Schultz, 2017). Sign of a coefficient estimate was interpreted as the direction of relation and p-value for the significance of the effect and contradictory findings were reported. Only considering VIF for testing these assumptions is not adequate. Even low levels of Pearson correlation exists among independent variables may cause multicollinearity. Heteroscedasticity, which is non-constant variance in error terms, will be observed when one of the assumptions of OLS regression are violated. Thus, evaluating the importance and the effect of the variables formulated by the conceptual model using the OLS approach will lead to unreliable results. As a remedy to the problem, the bootstrap approach of building multiple models on resampled observations and features, then aggregating the estimates on these models has been shown as a more reliable tool to estimate regression coefficients, confidence intervals and p-values for statistical inference (Ye, Du & Lu, 2023).

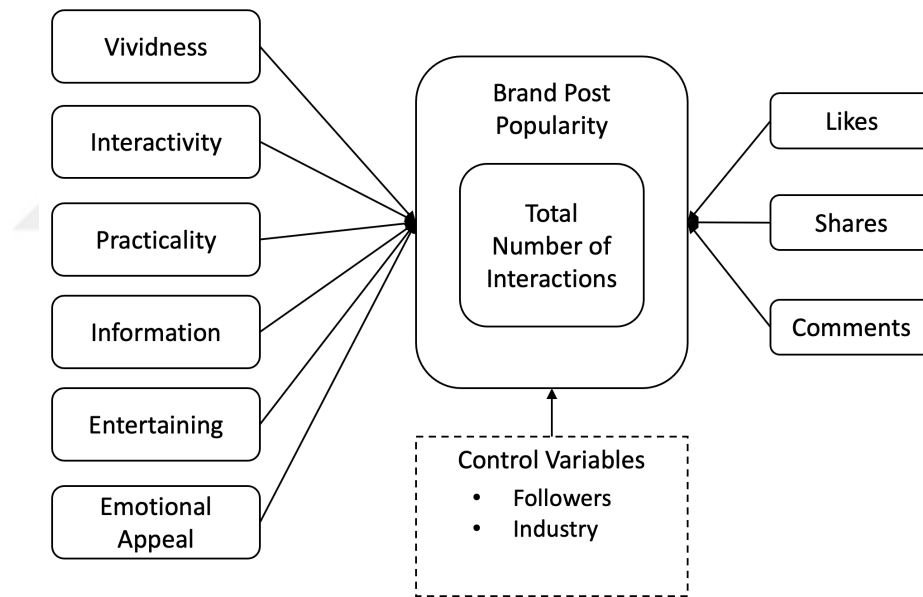


Figure 4.1 Conceptual Framework

4.1.3 Random Subspace Regression with Bootstrap for Feature Selection

Bootstrap is a resampling method to estimate the distribution of a test statistic or any parameter by sampling the data with replacement over a sufficient number of times. The original data to be sampled is considered as the population data. By gen-

erating multiple bootstrap samples and calculating the statistic of interest for each sample, sampling distribution can be estimated together with confidence intervals and standard errors (Efron & Tibshirani, 1994). Regarding linear regression, the average of a coefficient appearing in all bootstrap models is the estimate, .025 and .975 quantiles define the .95 confidence interval of the estimate. In the cases where the confidence interval contains zero, the null hypothesis stating that the value of the coefficient is zero is failed to be rejected. By resampling model parameters, their confidence intervals and rejection properties of hypothesis tests can more accurately be approximated regardless of the underlying probability distribution, under conditions that calculating the statistics is difficult. The difficulty may occur when data is large in volume, or certain statistical of the assumptions of the underlying model are not met (Horowitz, 2019).

When multiple models with full set of variables/features are built on B bootstrap samples and the final prediction is averaged over all models, this procedure is called bagging (Breiman, 1996). When a random subset of features are drawn B and fit on full data, it is called Random Subspaces (Ho, 1998). Finally, when both a subset of features and data are drawn randomly B times to fit B models, the procedure is referred as Random Patches (Louppe & Geurts, 2012). While some of the previous work use the keywords "Random Subspaces" and "Patches" interchangeably to refer to bootstrapping the features and the data altogether, building linear regression models with random features on bootstrapped samples is referred as Random Subspace Regression (Boot & Nibbering, 2019). Thus, Random Subspace Regression is actually a feature subset selection method where instead of applying an exhaustive search of models by subsets of features in many iterations, randomization speeds up analysis on big data, or when there is a large set of features (Ng, 2017).

As a feature selection / dimensionality reduction method, the performance of random subspace approach with bootstrap samples was shown to be useful by previous research. In (Diaz-Diaz, Aguilar-Ruiz, Nepomuceno & Garcia, 2005), random subset of features on bootstrap samples were used to measure an importance score for each feature according to its contribution to classification accuracy in each iteration. Then they used the scores to rank the features according to their importance. Their approach produced similar results to exhaustive one with very low number of iterations. In (Rabbi, Khalil, Khan, Almuqrin, Khalil & Anduallem, 2022), the authors used bootstrap with linear regression to ease out the effect of outliers in data and showed that consistent and satisfactory results were attained in the presence of response outliers and covariate outliers. More recently in (Wichitaksorn, Kang & Zhang, 2023), logistic regression model with random subspace of features was evaluated on simulated and empirical data in comparison with Lasso shrinkage, random

forest and neural networks. Their results showed that while random subspace regression performed generally well, computationally it was more efficient than the other methods. Similarly in (Hwangbo, Sharma, Arndt & TerMaath, 2023), randomized subset modelling was proposed to overcome the challenges in high-dimensional data analysis. Their results showed that the proposed model not only achieved the lowest prediction error when compared to common feature selection methods like Lasso, Random Forests or Neural Networks but also the important variables extracted were in alignment with the other alternatives verifying its feature selection capability.

Bootstrap resampling was shown to obtain more accurate generalization on high dimensional and non-normal data when residuals of OLS models revealed signs of heteroscedasticity in error terms. Bishara & Hittner compared Fisher z' , two Spearman rank-order methods, the Box–Cox transformation, rank-based inverse normal transformation, and various bootstrap confidence interval methods for correlations through Monte Carlo simulation (Bishara & Hittner, 2017). Their findings showed that the actual coverage of the 95% confidence interval on non-normal data had actual coverage as low as 68%, while the bootstrap method they used came closest to accurate coverage, though resulting often in an excessively long interval. Ye et al. considered one-sided hypothesis testing and interval estimation problems in the skew-normal unbalanced heteroscedastic one-way classification random effects model (Ye et al., 2023). They compared the test statistics and confidence intervals constructed by both bootstrap and generalized approaches and concluded that bootstrap approach performs better than the generalized approach in most cases.

Random Subspaces with bootstrap was also used for problems related with Business domain by previous work. In (Wang & Ma, 2011) and (Li, Lee, Zhou & Sun, 2011) the method is used to predict bankruptcy and credit risk. McCracken & Ng applied random subspace bootstrap to 130 macroeconomic time series and found that improve forecast accuracy was improved compared to well-known alternatives like principal component regression (McCracken & Ng, 2016). In (Boot & Nibbering, 2019) applied random subspace methods to empirical data and showed that the methods expected to work well in macroeconomic forecasting, where the factors that explained most of the variation in the predictors drove the dependent variable. However, the method has not yet been applied for determining the impacting features of brands' social media messages to increase brand post popularity, hence, random subspace regression with bootstrap is the method of this study.

4.2 Methodology

The research design of this study is shown in Figure 4.2. Following data collection, variables are operationalized with regards to reliability measured described by previous work on Brand Post Popularity. Then, by using the variables in the previous step, goodness-of-fit of OLS, Poisson and Negative Binomial regression models were evaluated on total number of interactions - sum of likes, shares and comments. Random subspace regression using bootstrap is further applied on the best fitting regression model for measuring importance score, coefficient estimate, confidence interval and p-value for the variables in the regression model.

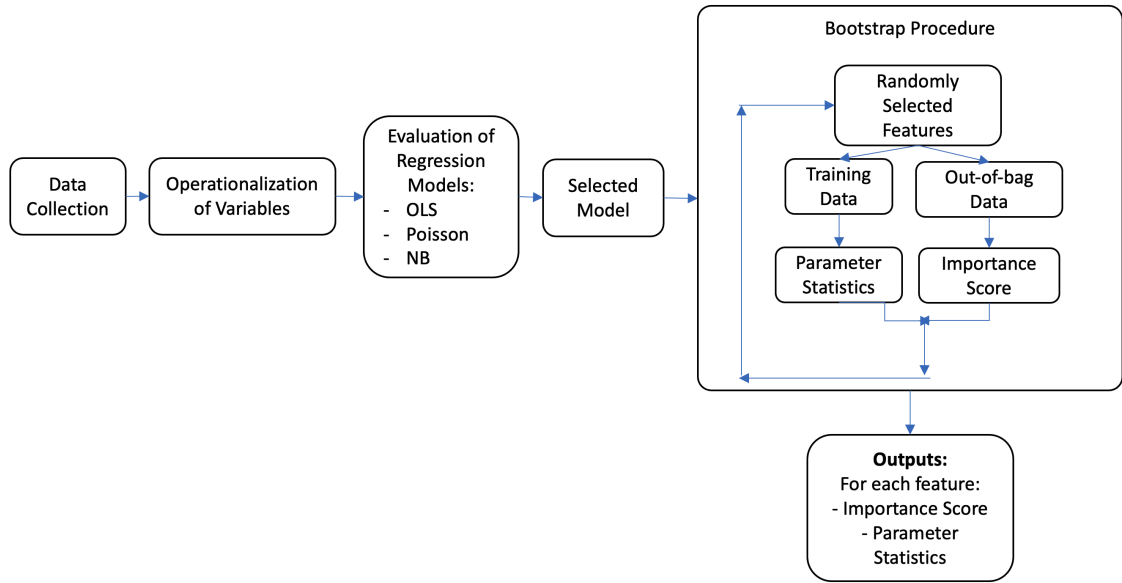


Figure 4.2 Research Design to Evaluate the Importance and Effects of Post Features on Brand Post Popularity

4.2.1 Data

As indicated by future directions of previous research on Brand Post Popularity, Twitter as a social medium demands further attention. Also reported to be among the four most popular social media platforms in North America and Europe, Twitter as the second most preferred platform has an increasing popularity trend in Turkey (GobalStats, 2023). Thus, Twitter posts of brands and their interactions are considered as the target data source for this study. Furthermore, there is no cross-cultural analysis of the effects of common features on post interactions on any social platform in earlier studies. According to Hofstede et al.'s cultural scores, English speaking countries like USA and UK have are substantially more individualistic cultures (89 and 90 points respectively) than Turkey (37 points) (Hofstede et al., 2010). Thus,

in order to fulfil this gap of cultural perspective in earlier studies, English language was selected to represent markets from the more individualistic cultures, whereas Turkish language was selected to represent markets from less individualistic, more collectivist cultures.

Twitter posts of 33 Fortune 500 brands from FMCG, Fast Food, Technology, Automotive, Apparel, Retail, Finance, and Logistics industries actively operating Twitter accounts in Turkish and English were selected as sample. Using Twitter's Academic API and Java programming language, all posts of the brands in sample from June 2016 to June 2021 were downloaded to be analyzed. Consequently for Turkish 37,825 and for English 571,561 brand posts were collected.

Total number of posts, followers, and interactions (sum of likes, comments and shares) are shown in Table 4.1. The highest volume of posts and interactions are from the Technology industry regarding both languages. For Turkish the least number of interactions are observed in Logistics, where for English it is observed in Financial Services industries. As also noted by previous research, a fair amount of positive correlation exists among Followers and Interactions (Pearson English: 0.60, Turkish: 0.99). Moreover, brand posts and interactions variate significantly by industry. As a result of their apparent effects number of followers and industry are required to be controlled in statistical models.

Table 4.1 Total number of posts, followers and interactions by industry

Industry	Turkish			English		
	Posts	Followers	Interactions	Posts	Followers	Interactions
Logistics	1,700	61,141	52,707	22,785	299,371	846,139
Personal Care	2,269	76,621	127,525	8,564	421,327	1,394,644
Food Service	3,286	39,342	136,144	4,785	8,015,842	19,113,026
Food & Beverage	856	135,096	232,423	5,676	6,427,611	2,782,856
Financial Services	3,364	49,412	287,831	5,348	329,809	433,440
Retail	9,678	426,035	494,603	14,763	9,772,636	2,401,469
Automotive	3,992	492,354	571,561	14,323	3,456,692	13,847,786
Technology	12,680	3,656,490	11,227,930	25,726	25,569,968	16,583,574

Total number of interactions together with the numeric independent variables have a highly right skewed non-symmetric distribution for both Turkish and English data, where proportionally more zeros are observed on interaction values regarding English (see Figure 4.3). All of the relationships of numeric variables and the target variable are non-linear regarding both languages. The distribution of dependent variable (total number of interactions) over binary variables are slightly right skewed for Turkish, bi-model as a result of the excess number of zeros in dependent variable. For English posts, highest values of Pearson correlation among the independent and

dependent variables are observed for followers (0.62), Logistics industry (-0.52) and informative content (0.40) followed by entertaining content (0.23) and Food Service industry (0.20). For Turkish, followers has the highest Pearson correlation (0.36) and other variables have relatively lower levels of correlation from -0.1 to 0.1 when compared to Turkish. Moreover, a fair amount of multicollinearity exists among the independent variables for both languages. Statistical models proposed for brand post popularity should consider remedies to skewness of distributions, non-linear relations together with the confounding effect of independent variables.

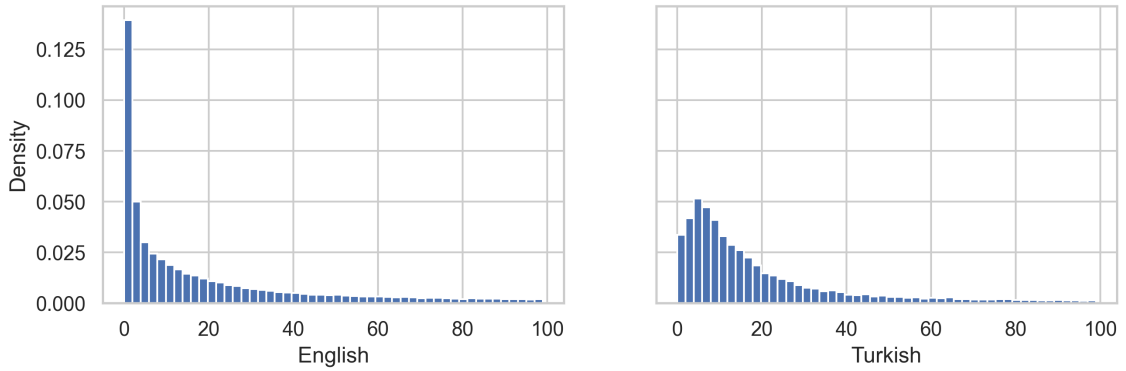


Figure 4.3 Distribution of Interactions by Language

x-axis: density of the observed count of interactions; y-axis: Total number of shares, comments and retweets less than value of 100.

4.2.2 Operationalization of Variables

The coding scheme and types of variables are shown in Table 4.2. While modelling is mostly similar, unlike previous work where *vividness* and *interactivity* were operationalized over ordinal levels as low, medium and high (De Vries et al., 2012; Schultz, 2017), in this study both categories are represented by the quantity of features existing in the content of brand posts. *Vividness* category is measured by the quantity of emojis, images and videos in a message. *Interactivity* of posts is measured by the quantity of hashtags, links, mentions, if there is a question and if a poll exists in their content.

Practicality category is constructed by daytime, weekday, position, text length and holiday variables. Daytime is a binary variable coded as 1 if the post submission is in early morning, morning or noon. Weekday is a binary variable coded as 1 if the message is posted on week days, not Saturdays or Sundays. Holiday variable is coded as 1 when the post mentions about a holiday concept or is submitted on

Table 4.2 Quantitative representation of features by antecedent categories

Features	Description	Type	Value
<i>Vividness</i>			
Emojis	Number of emojis in a post.	discrete, numerical	≥ 0
Images	Number of images in a post.	discrete, numerical	≥ 0
Videos	Number of videos in a post.	discrete, numerical	≥ 0
<i>Interactivity</i>			
Hashtags	Number of hashtags in a post.	discrete, numerical	≥ 0
Links	Number of links in a post.	discrete, numerical	≥ 0
Mentions	Number of user mentions in a post.	discrete, numerical	≥ 0
Poll	Post is a poll.	binary	1: if poll, 0: otherwise
Question	Post is a question.	binary	1: if question, 0: otherwise
<i>Practicality</i>			
Daytime	Post is submitted in daytime.	binary	1: if posted daytime, 0: otherwise
Weekday	Post is submitted on weekdays.	binary	1: if posted weekdays, 0: otherwise
Position	Duration at top position in minutes.	discrete, numerical	≥ 0
Text Length	Number of characters in post message.	discrete, numerical	≥ 0
Holiday	If the post is about a holiday.	binary	1: if about holiday, 0: otherwise
<i>Emotional Appeal</i>			
Happy	Post stimulates happy feelings.	binary	1: if happy, 0: otherwise
Surprise	Post stimulates surprise feelings.	binary	1: if surprise, 0: otherwise
Sad	Post stimulates sad feelings.	binary	1: if sad, 0: otherwise
Anger	Post stimulates angry feelings.	binary	1: if anger, 0: otherwise
Disgust	Post stimulates disgustful feelings.	binary	1: if disgust, 0: otherwise
Fear	Post stimulates fearful feelings.	binary	1: if fear, 0: otherwise
Neutral	No emotions are attached to the post.	Base Variable	
<i>Information</i>	Post is about a product.	binary	1: if about product, 0: otherwise
<i>Entertainment</i>	Post is about sports, TV, music, events.	binary	1: if entertaining, 0: otherwise
<i>Industry (Control)</i>			
Logistics	Brand in Logistics industry.	binary	1: if from Logistics, 0: otherwise
Food Service	Brand in Food Service industry.	binary	1: if from Food Service, 0: otherwise
Personal Care	Brand in Personal Care industry.	binary	1: if from Personal Care, 0: otherwise
Retail	Brand in Retail industry.	binary	1: if from Retail, 0: otherwise
Financial Services	Brand in Finance industry.	binary	1: if from Financial Services, 0: otherwise
FMCG	Brand in FMCG industry.	binary	1: if from FMCG, 0: otherwise
Automotive	Brand in Automotive industry.	binary	1: if from Automotive, 0: otherwise
Technology	Brand in Technology industry.	Base Variable	
<i>Followers (Control)</i>	Number of followers of the brand account.	discrete, numerical	≥ 0

holidays. Position is a discrete numeric variable of the duration of the post on the top position of a brand’s Twitter page until another Tweet is submitted. Text length is a discrete numeric variable quantified over the number of characters a post contains. Too large or too low values are observed for these two variables, and their probability distributions are not normal. Thus, text length and position variables were min-max normalized prior to analysis.

Information and *entertainment* are binary variables. When the post content is about sports, computer gaming, music, movies, TV, hobbies, or events they are coded as entertaining. Posts with messages related specifically to products or product features are coded as informative.

Emotional appeal is represented over 6 binary variables as *happy*, *surprise*, *sad*, *anger*, *disgust* and *fear*. *Neutrality* is the base variable for emotional appeal. Dominant emotion conveyed by brands is first extracted by Multilingual BERT (Devlin et al., 2018a) regarding English posts, and by BERTurk (Schweter, 2020) for Turkish posts using pre-trained deep learning classifiers as described in Chapter 3. Then, the discrete emotion extracted as text is encoded as a binary variable as described.

Regarding control variables, industry is represented over 7 binary variables as *Logistics*, *Food Service*, *Personal Care*, *Retail*, *Financial Services*, *FMCG* and *Automotive*. *Technology* is the base category for industry. As the second control variable *Followers Count* is a discrete numerical variable having the value of number of followers of brands at the time of post was downloaded.

4.2.3 Models

The most common method to explain the relationship between the features/variable described in Section 4.2.2 and brand post popularity measured in likes, shares, comments along with their transformations was the linear models. In (Antoniadis et al., 2019; De Vries et al., 2012; Schultz, 2017) and more recently in (Banerjee & Chua, 2019) natural log transformed likes, shares and comments were modelled using Ordinary Least Squares regression. As likes, shares and comments are discrete count variables, different than others Swani et al. and more recently Negative Binomial regression in (Swani & Milne, 2017).

In order to discover the most impacting features and their direction of effect, also in line with previous research, the linear relation in Equation 10 is used to model brand post popularity for English and Turkish languages.

$$(4.1) \quad y_i = X_i \beta_i + \varepsilon_i$$

Where,

i denotes language where $i \in \{English, Turkish\}$, and $y_i = likes_i + shares_i + comments_i$;

β_i represents the vector of regression coefficients for language i ;

X_i is the matrix of independent variables for language i ;

ε_i is the vector of error terms for language i

Then, the linear relation described in Equation 10 is used to fit OLS, Poisson and Negative Binomial (NB) regression models in order to find the best distribution to explain the variability of brand post popularity measured as y_i . OLS, Poisson

and NB versions of the regression models are described in Equation ??, ?? and ?? respectively.

OLS Regression:

$$(4.2) \quad f(\ln y_i | X_i, \beta_i) = X_i \beta_i$$

Where,

$\ln y_i$ is natural log transformed sum of likes, shares and comments for language i ;

X_i is the matrix of independent variables;

β_i represents the vector of regression coefficients;

Regarding OLS regression, the goal is to find the values of β_i 's that minimize the sum of squared residuals between the observed dependent variable y_i and the predicted values $X_i \beta_i$. The estimators are calculated by the normal equation is given by: $\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$, where $\hat{\beta}$ represents the estimates of the coefficients.

Poisson Regression:

$$(4.3) \quad f(y_i | X_i, \beta_i) = \frac{e^{-\lambda} \lambda^{y_i}}{y_i!}$$

Where,

y_i is the sum of counts of likes, shares and comments for language j ;

X_i is the matrix of independent variables;

β_i represents the vector of regression coefficients;

λ is the expected value or mean of the Poisson distribution, $\lambda = e^{X \beta_i}$;

NB Regression:

$$(4.4) \quad f(y_i | X_i, \beta_i, r_i) = \frac{\Gamma(y_i + 1) \Gamma(r_i)}{\Gamma(y_i + r_i)} \left(\frac{r_i + \mu}{r_i} \right)^{r_i} \left(\frac{r_i}{r_i + \mu} \right)^{y_i}$$

Where,

y_i is the sum of counts of likes, shares and comments for language j ;

X_i is the matrix of independent variables;

β_i represents the vector of regression coefficients;

r_i is the dispersion parameter of the Negative Binomial distribution;

$\Gamma(.)$ denotes the gamma function;

μ_i represents the mean parameter, which is typically modeled as $\mu_i = e^{X_i\beta_i}$;

Poisson regression variance is constant and equal to mean λ , where in NB regression accounts for an extra dispersion parameter r to model the over-dispersed count data. Regarding both models a closed form solution to find the best β 's doesn't exist, thus, maximum likelihood estimation (MLE) is used to find the values that maximize the likelihood of the observed data given the model.

In order to explore the most impacting features OLS, Poisson and NB regression models are fit separately on English and Turkish data sets. Goodness-of-fit of the models are measured by Akaike information criterion (AIC). AIC measures the relative quality of statistical models for a given set of data (Stoica & Selen, 2004) where lower scores are more favorable in for model selection. Thus, following the fitting procedure, the best regression model having the lowest AIC score is selected for further analysis of the features.

The real distribution of total interactions and the predictions of the models fit along with their AIC scores are shown in Figure 4.4.

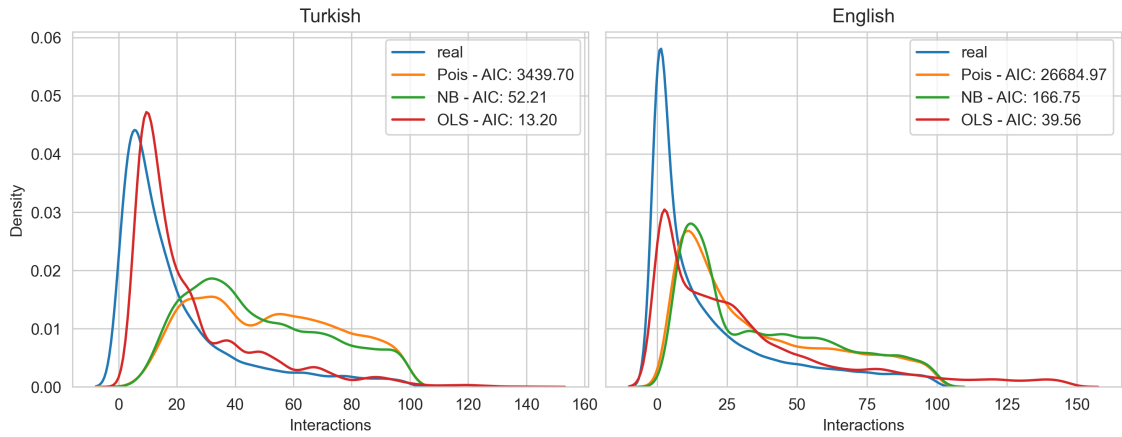


Figure 4.4 Comparison of Regression Models Evaluated on Alternative Distributions of Interactions

The lowest AIC scores for both languages are measured in OLS regression (*English: 39.56, Turkish: 13.20*) modeled for log transformed total number of interactions. OLS is followed by NB and Poisson regression models. Besides the AIC score, the distribution of OLS predictions are closest to the real distribution in English and Turkish data sets when compared to NB and Poisson as clearly seen on the blue and red lines in Figure 4.4. In both languages, the OLS model performs an acceptable performance of accounting for the skewness of distributions following the right tails. Consequently, OLS model is selected for further evaluation of the effect of variables on English and Turkish brand post consumer interactions in this study. Given the selected model, how the features are evaluated by random subspaces on bootstrap samples is discussed in the next section.

4.2.4 Random Subspace Regression Using Bootstrap Method

As discussed in Section 4.1.3, the results of OLS regression would be misleading when the assumptions are violated, and in such cases bootstrap resampling was shown to obtain more accurate generalization. Regarding the brands data set, the interactions are skewed and the relation of independent - dependent variables are non-linear. Moreover, a fair amount of multicollinearity exists although mean VIF for English and Turkish is 2.77 and 2.80 respectively. While individual VIF scores are mostly less than 10 and mean scores are low, there are natural correlations among independent variables. As some of the examples, images or videos are posted together with links; links, mentions, hashtags or emojis increase the post length. It is likely that residuals of an OLS model would show signs of heteroscedasticity, reducing confidence in generalizability of interpreting coefficient estimates.

As a remedy to the problems related with inference encountered when the assumptions of OLS regression are violated, random subspace regression using bootstrap method, similar to the previous work by Diaz-Diaz et al. and later by (Wichitakorn et al., 2023), is adapted to estimate coefficients, their confidence intervals and p-values along with an importance score for each coefficient given features and the conceptual model. The procedure defined in Algorithm 2 is applied using OLS regression as it predicts the closest form of the distribution of interactions on brand posts data. Adjusted R_{adj}^2 , shown in Equation 14 as performance metric for feature importance score calculation.

$$(4.5) \quad R_{adj}^2 = 1 - \left(1 - R^2\right) \frac{n-1}{n-p-1}$$

where n represents the sample size and p represents the number of predictors in the model.

14, scores the explained variance by the model, given the features. Unlike R^2 , R_{adj}^2 is resistant to increasing number of variables in the regression model by penalizing the score down when the number of parameters increase. Although OLS regression for modelling and Adjusted R^2 is used for importance scoring in this study, the procedure defined may also be applied on any regression or classification model together with other performance metrics such as negative mean squared error for regression or accuracy for classification.

Monte Carlo simulations are run on English and Turkish data sets to generate linear regression models on randomly selected features and observations. Given inputs data D , a set of features P , each feature indexed as p , sample size n , number of bootstrap iterations b , and maximum number of random features j , in each iteration a random sample of n observations D_{boot} are drawn from D . The left over observations D_{oob} are out-of-bag observations to be used for importance score calculation. A random sample of minimum 1 and maximum j features P_{boot} are selected to fit an OLS regression model on D_{boot} . So, each regression model each i 'th bootstrap iteration has a different number of randomly selected predictors. At the end of each iteration, the estimate of each selected coefficient is stored in vector β_p , and the adjusted R^2 of the model on D_{oob} is stored for each coefficient in vector R_p^2 .

Following the completion of i bootstrap iterations, for each feature in P the average of values in β_p is the coefficient estimate. Sorted in increasing order, .025 and .975 quantiles of β_p define the 95% confidence interval of the estimate. By using the standard deviation of β_p under normality assumption, a p-value for each coefficient estimate is calculated to determine whether the estimate is statistically significant.

The importance score for each feature is calculated by the regression models' explained variances they contribute to, and it is simply the average of values in vector R_p^2 . For example, if the feature 1 contributes to the models in bootstrap iterations 2,5 and 6 then its importance score is given by $Importance_1 = \frac{R_2^2 + R_5^2 + R_6^2}{3}$, where R_i^2 denotes the variance explained or goodness-of-fit of iteration i .

Finally, features are sorted according to decreasing importance and the associated confidence intervals and p-values indicate the significance of the coefficient estimates of features. For a given *alpha* level as 0.01 or 0.05, the features can further be elim-

inated according to their significance. sample size n , number of bootstrap iterations b and maximum number of random features j can be adjusted according to the problem on hand and computational capacity.

Algorithm 2 Feature Importance and Significance by Random Subspace Regression and Bootstrap

```

1: Inputs: data matrix  $D$ , feature set  $P$ , feature index  $p$ , sample size  $n$ , number
   of bootstrap iterations  $b$ , and maximum number of random features  $j$ 
2: Outputs: means of coefficient estimates  $\mu$ , lower bound 0.95 CI  $cli_l$ , upper
   bound 0.95 CI  $cli_u$ , importance score  $imp$ , standard deviation  $s$ , p-values  $pval$ 
3: Let coefficient matrix  $\beta$ , goodness-of-fit matrix  $R^2$ 
4: for  $i = 1, 2, \dots, b$  do
5:   Sample  $n$  observations from  $D$  to  $D_{boot}$  with replacement;
6:   Generate  $D_{oob}$  from  $D$  of observations not in  $D_{boot}$ ;
7:   Let  $r \leftarrow$  randomly generated number, where  $r \leq j$ ;
8:   Sample  $r$  features from  $P$  to  $P_{boot}$ ;
9:   Given features in  $P_{boot}$ , Fit OLS regression model  $Model_i$ , having coefficient
   estimates  $coef_{ikp}$  indexed as  $k$ 
10:  Calculate adjusted  $R_i^2$  on  $D_{oob}$ 
11:  for  $0..k$  do
12:    Store  $coef_{ikp}$  in  $\beta_p$ 
13:    Store  $R_i^2$  in  $R_p^2$  for  $coef_{ikp}$ 
14:  end for
15: end for
16: for  $0..p$  do
17:   Calculate mean of  $\beta_p$  and store in  $\mu_p$ 
18:   Store 0.025'th quantile of  $\beta_p$  in  $cli_{lp}$ 
19:   Store 0.975'th quantile of  $\beta_p$  in  $cli_{up}$ 
20:   Store standard deviation of  $\beta_p$  in  $s_p$ 
21:   Calculate P-value of observing data given  $\mu_p = 0$  using  $s_p$  and store in  $pval_p$ 
22:   Calculate  $\frac{1}{|R_p^2|} \sum R_p^2$  and store in  $imp_p$ 
23: end for

```

The Random Subspace Regression using bootstrap samples is applied on English and Turkish data sets described in Section 4.2.1. The method is consistent with asymptotic theory, or large sample theory, which states that as the sample size n grows larger, the sampling distribution of an estimator converges to a known distribution, also known as the limiting distribution. One can make statistical inferences and draw conclusions about the population from which the sample is obtained by studying the attributes of this limiting distribution (DasGupta, 2008; Höpfner, 2014). Regarding this study, the estimators are regression coefficients, their confidence intervals, p-values and importance scores. The results of experiments and comparison of these estimates across languages are discussed in the next section.

4.3 Experimentation and Results

As a preliminary analysis, starting with the null model having only the intercept term, features are selected in groups and by adding the groups one by one into the feature set an OLS regression model is fit. Following each fit performance metrics of adjusted R^2 and Root Mean Squared Error (RMSE) are recorded. The steps and the performance metrics obtained are shown in Figure 4.5. Control variables (Industry and Followers) are significantly the most contributing group to R^2 . Practicality, interactivity, vividness, entertainment, and emotions have slightly higher contribution to the variance explained by the model in English than in Turkish. Information and Entertainment variable contribute to explained variability more in English data set than in Turkish.

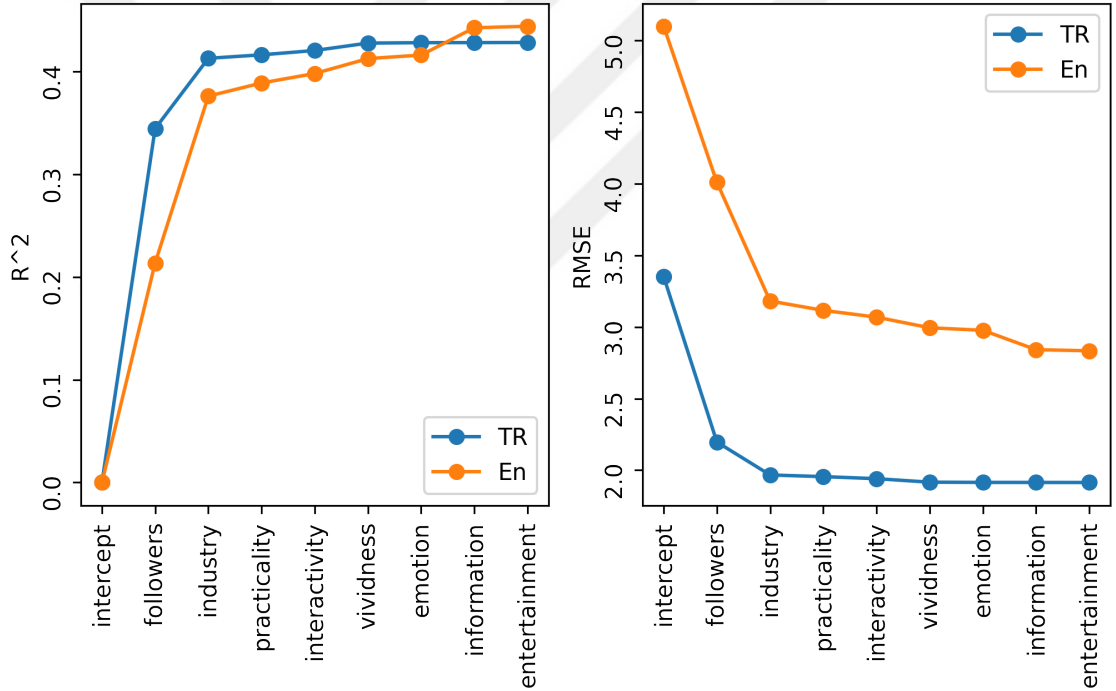


Figure 4.5 Adjusted R^2 and RMSE given added features in regression model

All OLS regression models are fit using StatsModels library, and Monte Carlo Simulations (bootstrap) described in Algorithm 2 are coded in Python development environment. In order to determine the minimum number of steps of bootstrap iterations adequate to detect the set of features best explaining the variance, bootstrap procedure is applied with 50, 100, 500, 1000 and 5000 iterations. Sample size n for each iteration is 10,000 observations, and in each iteration maximum number of 10 features are drawn from the full set of features. RMSE for each bootstrap cycle

with increasing number of iterations are shown in Figure 4.6. In each cycle, features with the highest importance scores and p-values less than α level of 0.01 are selected and cross validated RMSE of the OLS model is measured on data. Also, the cross validated RMSE of the full model with all features is calculated on all data for comparison (See blue line in Figure 4.6). The bootstrap approach shows lower RMSE for both languages when compared to the full model. Lowest RMSE is achieved with the bootstrap cycle running for 1000 iterations, which is significantly less than feature subset selection using exhaustive search, where the parameter space to search for is the subsets of all features making 2^{29} alternative combinations, given that 29 features are considered for the linear model.

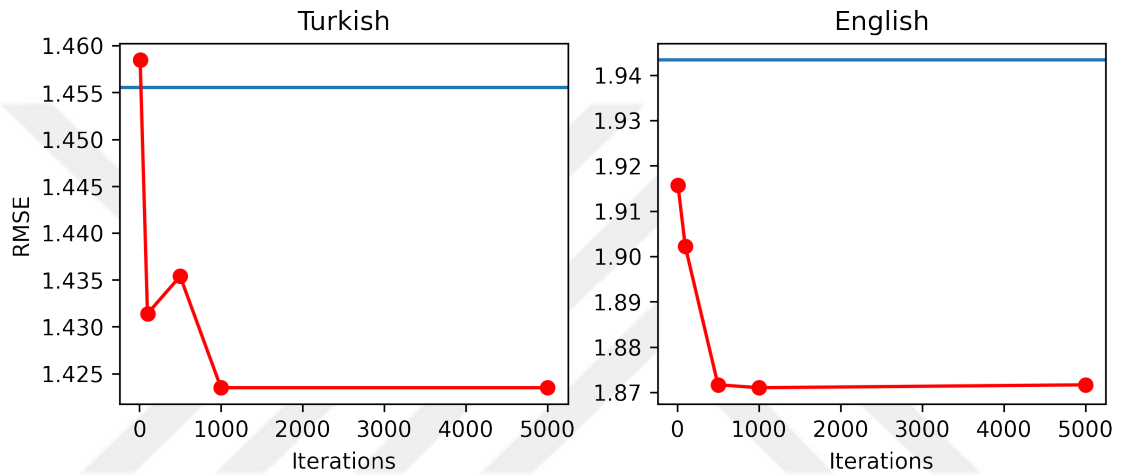


Figure 4.6 RMSE over bootstrap cycles with increasing number of iterations
Red line: cross validated RMSE of selected model after bootstrap iterations; Blue line: cross validated RMSE of the full model

Average importance scores by groups of variables of running 1000 random subspace bootstrap iterations on Turkish and English data sets are shown in Figure 4.7. While the control variables including number of followers and the industry of each brand account has the highest average importance for Turkish (0.159), it is the second important group for English (0.197) following the information variable (See Figure 4.7). The average scores suggest that industry and number of followers contribute to explaining engagement of interactions on Twitter regarding both languages. Regarding the Turkish post engagements, all categories other than control variables fluctuate around 0.12, having same level of importance. However regarding English data, Information and Entertainment are scored 0.243 and 0.184 respectively, suggesting that within the information and entertainment categories, the variables hold relatively more importance in driving Twitter post interactions, particularly in the English context. This result is consistent with the observed pattern of full model with all features shown in Figure 4.5. Vividness, emotional appeal, inter-

activity and practicality features are ranked according to their average importance scores similarly for both languages, fluctuating around 0.12 and 0.16 for Turkish and English data respectively, while vividness having relatively higher importance for English interactions. The Emotion category has importance scores of 0.120 and 0.160 in the Turkish and English data, respectively. This indicates that within the emotion group, the variable holds relatively more importance in driving Twitter post interactions, particularly in the English context. It suggests that evoking emotional responses through the content may be influential in driving engagement and interactions on Twitter among English-speaking users.

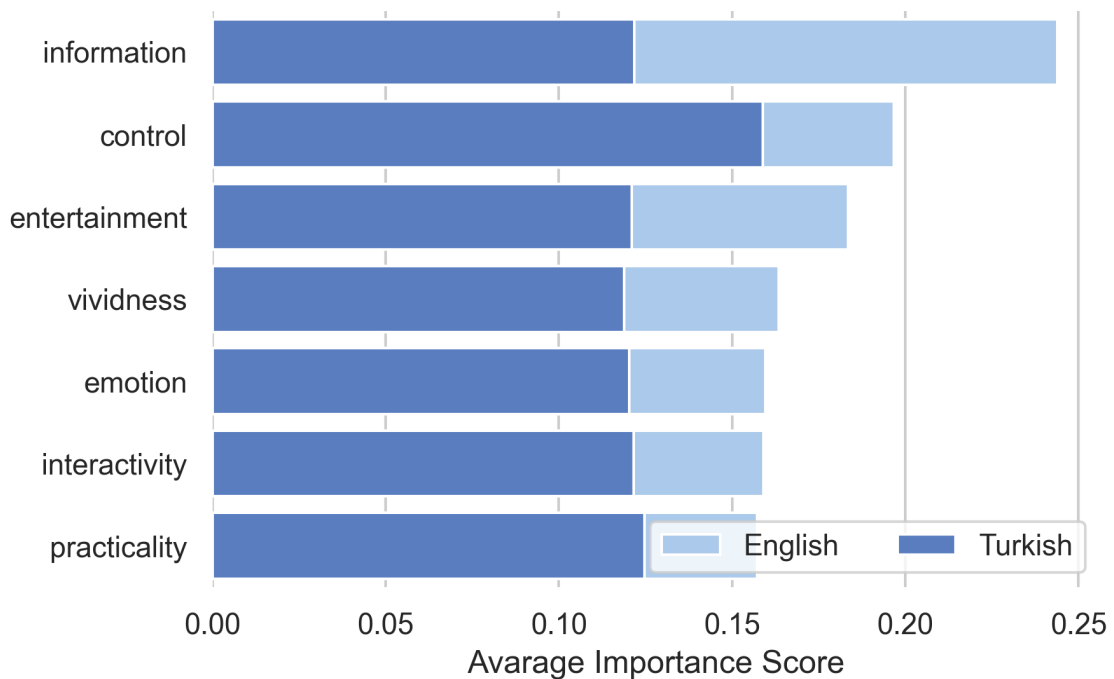


Figure 4.7 Average Importance Scores by Feature/Variable Category
Sorted in decreasing order by English importance scores.

Final estimates and statistics for variables following the bootstrap procedure in details are shown in Figure 4.8. The highest importance scores belongs to the variables in **Control** category, demonstrating the necessity of controlling the industry and the number of followers. Regarding Industry variables, for English, Logistics and Food Service industries have significant effect, while for Turkish all control variables except Financial Services are significant. Naturally, number of followers increase mean post interactions for both languages. Food service has a contrasting effect in the Turkish and English data. In the Turkish data, it shows a negative effect with a mean of -1.211, while in the English data, it demonstrates a positive effect with a mean of 2.035. The p-values indicate that this difference is statistically significant in both datasets.

Regarding *Practicality* of Twitter brand posts, messages posted in day time decrease mean post interactions in Turkish, where the effect is negative, however less significant in English. For both languages messages posted during weekdays does not effect post interactions significantly. Text length demonstrates a different pattern in the Turkish and English data. In the Turkish data, it has a negative effect with a mean of -0.916, while in the English data, it has a positive effect with a mean of 0.392. However, the p-values indicate that there is no statistically significant difference in the effect of text length between the Turkish and English data. The variable Holiday has a positive coefficient (mean = 0.778), indicating that tweets posted during holidays have higher interactions for English. The p-value (0.009) suggests this relationship is statistically significant for English, while not significant for Turkish. Lastly, the variable Position shows a negligible effect in the Turkish data, while it has a positive effect in the English data, meaning messages staying at the top position of messages are more likely to be liked, shared or commented in English. While aligning with the earlier investigations conducted by De Vries et al. and later by Schultz on the effects of position and holiday variables, there is not enough evidence to derive the effect of text length regarding English messages according the the results of this study.

Among the variables in *Interactivity* category, number of links in Turkish, and Questions in English have a negative significant effect on mean post interactions. All other variables demonstrate negligible effects in both the Turkish and English datasets, as indicated by the p-values. Regarding Question variable in English, the results are similar with (De Vries et al., 2012), estimating a negative effect of questions on post interactions. While conflicting results were reported on the effect of interactive content in previous work (De Vries et al., 2012; Pletikosa Cvijikj & Michahelles, 2013; Schultz, 2017), no significant relation is observed regarding the variables of this group except the Question variable in this study.

Vividness features on overall are not significantly effecting mean interactions for Turkish. However, for English, positive coefficients and p-values of images and emojis imply that incorporating vivid or lively elements in the content may enhance engagement and interactions on Twitter, particularly among English-speaking users. While aligning with the earlier investigations conducted by De Vries et al. in regarding vivid posts getting more interactions, this study reveals divergent outcomes with Schultz, introducing complexities that expand the understanding of the vivid features and its impacts on brand post engagements.

Within *Emotion* category, happy and sad messages increase the interactions significantly, while the effect is negative for angry emotional appeal in English messages. In

Turkish data, while similar effect is observed for sad emotional appeal, the opposite effect is observed for messages containing angry tone. Along with their importance scores, emotional appeal variables are shown to be useful on explaining post interactions, also in line with studies of Pinto & Yagnik and (Karpinska-Krakowiak & Modlinski, 2020) highlighting the significance of the emotions in shaping the post engagement of users. Findings of this study extend and strengthen the existing body of knowledge surrounding the effect of emotional appeal, showcasing consistent associations in more detail with discrete emotions.

Messages mentioning *Information* about their products and services effect post interactions positively for both languages. The variable holds relatively more importance in the English context indicating that focusing on informative content may lead to higher engagement and interactions on Twitter, especially when targeting English-speaking audiences. The conclusions drawn in Mazloom et al.'s study are closely aligned with the results presented, however conflicting with earlier work of Antoniadis et al.; De Vries et al. and Karpinska-Krakowiak & Modlinski who found no significant relation between the informative content and popularity of brand posts

Entertainment, on the other hand, exhibits a positive effect in both the Turkish and English data. However, the effect is not significant in Turkish. It suggests that focusing on entertaining content may be influential in capturing attention and generating interactions on Twitter, particularly among English-speaking users. Consistent with the observations made by earlier studies of De Vries et al.; Pletikosa Cvijikj & Michahelles, findings reinforce the notion that entertaining posts draw more attention of users.

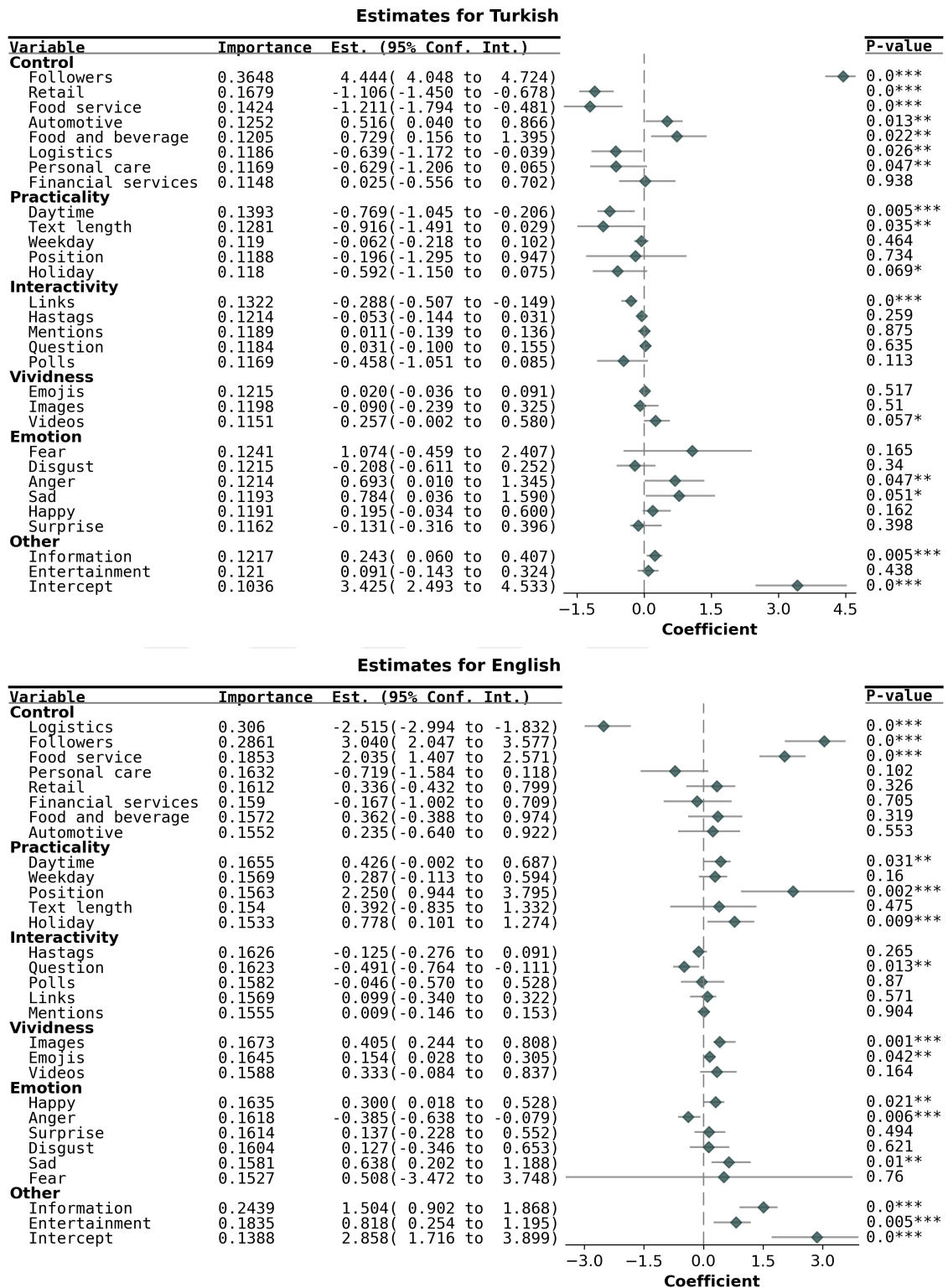


Figure 4.8 Random Subspace Bootstrap Estimates

Each category is sorted according to importance score in decreasing order.

These findings provide insights not only into the relative importance of the variables in driving Twitter post interactions considering different language contexts, but also into the differences and similarities of the variables' effects on the outcomes in the

Turkish and English data. The variations in importance scores, directions of effect and significance values highlight the potential differences in user preferences and engagement factors across languages.

4.4 Conclusion

This study aimed to present the factors that contribute to higher user engagement regarding multinational brands' Twitter posts. The focus was the structural and textual factors/ features of Twitter posts categorized as vividness, practicality, interactivity, information, entertainment, as well as the emotional appeal classified as happiness, anger, sadness, surprise, fear and disgust. Brand post popularity was conceptualized by a modified version of the conceptual model proposed by previous studies, and operationalized as sum of likes, shares and comments. Different than previous studies, multinational brands' Twitter posts in English and Turkish languages were compared for a cross cultural perspective on the analysis of features. For scoring and measuring their significance, a bootstrap resampling procedure using random subspace of features was proposed.

Consequently, aimed at answering the research question *RQ: What are the most influential factors, including emotions, explaining brand post popularity across cultures?*, the key findings can be summarized as:

- Informative and entertaining posts about products and services are more influential in driving user engagement in English-speaking context than Turkish.
- As of their rank, while emotions are less important factors than information, entertainment and vividness regarding English speaking users, their importance is similar along all feature categories regarding Turkish speakers deriving brand post interactions. Regarding both languages, brand posts with a sad emotional appeal are more likely to be interacted. Happy messages are favored by English speakers. While messages having an angry are favoured by Turkish users, the effect is opposite regarding English speaking users.
- More vivid posts with emojis and images play a significant role in capturing attention or generating engagement in English-speaking setting, while they are less effective in Turkish-speaking setting.
- Messages posted during day time are less likely to be interacted than messages

posted during night time for Turkish speaking users. Oppositely, day time messages are more favored by English speaking users.

- On overall, interactive features of posts like hashtags, links, mentions questions or polls are less significant factors effecting popularity of brand posts regarding both languages. Questions for English speaking users, and Links for Turkish speaking users are less favored than messages not containing them.

Along with the aim and key findings, this work contributes to previous research to resolve the three issues not adequately addressed. First, the random subspace regression using bootstrap is shown to be useful not only to estimate coefficients, confidence intervals and p-values of features, but also to evaluate their importance. The model found in 1000 iterations has higher explained variance and lower error rate relative to the proposed models in previous studies, validating the method proposed in this study. The approach may be followed in analyzing data where assumptions of the underlying model are not met by data.

Second, this is the first work comparing the estimates attained from random subspace regression models on English and Turkish data, exploring the effect of culture on the antecedents of brand post popularity. As shown by the findings, an obvious cultural variation on the effects of vividness, practicality, interactivity, information and entertainment features by language exists, supporting the fact that these features in different cultural settings may not be perceived similarly by consumers.

Third, as well as evaluating the cross cultural differences on vividness, practicality, interactivity, information and entertainment features of brand Twitter posts, the emotional tone of a brand message is also analyzed across cultures. As shown by the findings, a cultural variation on the effects of happiness, sadness, anger, surprise, fear and disgust by language exists, supporting the fact that emotions in different cultural settings may also not be similarly perceived by consumers.

From the managerial point of view, by the proposed approach of this study brands can optimize their social media strategies to enhance brand post popularity. By leveraging vividness, practicality, interactivity, entertainment, and information, managers can create more engaging and impactful brand posts that resonate with the target audience. Understanding the dynamics of these factors and their combined influence will enable managers to drive higher engagement levels, increase brand visibility, and ultimately achieve their marketing objectives. Moreover, harnessing emotional appeal to establish a deeper connection, Brand posts that evoke emotions are more likely to resonate with users and gain popularity. Managers should aim to create content that elicits emotional responses, such as happiness or

sadness. By tapping into users' emotions, brand managers can establish a deeper connection with the audience, leading to increased engagement and brand affinity.

A key limitation of this study is the lack of cultural generalizability due to the comparison of only two languages. As the analysis focuses on English and Turkish data, the findings may not be applicable to generalization for patterns in eastern or western cultures. Future research should aim to include a more diverse range of languages to capture cross-cultural variations in brand post popularity and interactions.

Another limitation is the time scale of the data used in the analysis. The dataset may have a specific time range, which might not reflect the most recent trends and dynamics in brand post popularity. Future research may validate the approach by updating the data with a more current time scale and provide a more accurate understanding of how brand post interactions evolve over time. Moreover, the current study does not explicitly examine the effect of time on brand post popularity. Considering the temporal dynamics and trends in social media engagement would be valuable for future research. Investigating how the popularity of brand posts changes over time, taking into account factors such as seasonality, trending topics, or specific events, would provide deeper insights into the dynamics of online brand interactions.

This study only analyzes a subset of features and emotions associated with brand posts. To obtain a more comprehensive understanding of the factors influencing post popularity, future research should consider incorporating additional features such as sentiment analysis of accompanying photos or videos, as well as exploring the values communicated through brand posts. By expanding the range of analyzed features, a more nuanced picture of the determinants of brand post popularity can be obtained by future studies.

In conclusion, while this study provides valuable insights into the cross-cultural analysis of brand post popularity, several limitations should be acknowledged. Future research should aim to address these limitations by including a more diverse set of languages, utilizing more recent data, expanding the range of analyzed features and emotions, and considering the effect of time on brand post interactions. By addressing these limitations, researchers can further enhance our understanding of the complex dynamics surrounding brand engagement in cross-cultural contexts.

5. CONCLUSION

The adoption of an interdisciplinary approach in this dissertation has been instrumental in investigating cross-cultural differences in the impact of emotional appeals within brand post engagements. By combining methodologies from diverse research areas, this study has provided a comprehensive analysis of the complex relationship between emotions and branding communication in social media. Cross-cultural differences concerning the impact of emotional appeals within brand post engagements is investigated through adoption an approaches from branding communication, machine learning and information theory to address several of these aspects. Thus, along with the key findings their implications must further be discussed with their contribution to theory, methodology and management realms.

5.1 Discussion on Key Findings

The findings discussed in 2 underscore the importance of emojis as a powerful tool for brand communication and customer engagement. Emojis offer a unique way to express emotions, which can enhance positive word-of-mouth and foster a deeper connection with consumers. For managers and marketers, the implications are clear: integrating emojis into marketing campaigns and social media strategies can lead to improved brand perception, increased customer loyalty, and higher willingness to purchase. However, it is essential to consider cultural nuances and language preferences when using emojis, as the study reveals differences in emoji usage between Turkish and English-speaking communities. To ensure effective communication, brands should tailor their emoji selection to resonate with the specific audience they are targeting. Moreover, understanding the social reasons behind these differences can aid brands in crafting more contextually appropriate messages and avoiding potential misinterpretations.

The convergence of emoji preferences to a smaller subset across both language communities signifies a growing universal understanding of certain emojis. For managers, this implies the value of incorporating commonly recognized emojis in marketing efforts to ensure messages transcend linguistic and cultural barriers. Leveraging information-theory methodologies for emoji analysis can provide deeper insights into consumer emotions and behaviors, enabling brands to refine their communication strategies and deliver more impactful messages. Moreover, the finding that negative emojis are used similarly across cultures emphasizes the need for caution when employing such emojis. Managers should consider the potential implications of using negative emojis, as they can evoke strong emotions and influence brand perception. By being sensitive to cultural differences and context, brands can harness the full potential of emojis as a valuable instrument for fostering customer engagement, driving brand loyalty, and ultimately boosting business success.

As discussed in 3, the extraction of emotions from various forms of data, such as text, images, and videos, holds significant implications in social media branding communication. Emotion extraction can offer valuable insights into the attitudes and opinions of the target audience towards a social policy or a particular product. By understanding the emotions expressed by the crowd, brands and policymakers can gauge the level of acceptance or resistance towards their offerings. This knowledge can enable them to make data-driven decisions and tailor their strategies accordingly. For instance, if analysis reveals negative emotions associated with a product, a brand can identify the pain points and work on product improvements or communication strategies to prevent negative word-of-mouth and boost customer satisfaction. On the other hand, positive emotions can be capitalized on to reinforce and amplify the message's impact.

A benchmark data source annotated by 6 basic emotions of happiness, surprise, sadness, anger, fear, disgust, and neutrality is an outcome of the dissertation serving this purpose for social media campaign managers to be able to measure emotions in a context, language and time independent way. Moreover, the finding that deep learning classifiers outperform conventional classifiers in emotion extraction for both Turkish and English has important implications for developing effective emotion detection tools. Investing in the development and implementation of advanced deep learning models can lead to more accurate emotion recognition systems, allowing companies to obtain real-time and precise emotional feedback from their customers and followers. This knowledge can be harnessed to customize the emotional content of social media messages to elicit stronger positive reactions, ultimately enhancing customer engagement and loyalty. Moreover, understanding that the performance of neural networks is positively correlated with their complexity, managers can make

informed decisions regarding the trade-off between model complexity and training time. Striking the right balance ensures efficient emotion extraction processes while maintaining acceptable computational costs.

In addition to evaluation of statistical and deep learning methods for classifying emotions in English and Turkish, the end-to-end framework proposed in 3 serves the purpose of understanding customer sentiment, refining marketing strategies, and enhancing brand-customer interactions. The ability to identify and analyze emotions expressed in textual data and other media enables businesses and policy-makers to make informed decisions, improve their offerings, and engage with their target audience more effectively. By leveraging advanced deep learning classifiers and recognizing the importance of model complexity, organizations can stay ahead in the competitive landscape and establish stronger emotional connections with their stakeholders.

The findings of this study highlight the significance of considering cultural differences and emotional appeal in social media brand messages to enhance user engagement as discussed in 4. The observed preference for informative and entertaining posts about products and services by English-speaking users suggests that companies should focus on crafting engaging and informative content tailored to the preferences of this audience. Additionally, the emphasis on emotional appeal for Turkish users indicates that brands should invest in understanding and tapping into the emotions that resonate with this audience to create more effective campaigns, as emotions are similarly effective as the other post features. This could involve using language and imagery that evoke specific emotions to elicit a stronger response, especially anger and sad messages for Turkish and happy messages for English.

The differences in emotional preferences between English and Turkish users, such as English speakers favoring happy messages while Turkish users engage more with posts having an angry emotional appeal, underline the importance of cultural sensitivity in branding communication. Companies must be mindful of the emotional nuances of different cultures and tailor their messages accordingly. This might involve conducting market research to understand the cultural norms and values of the target audience and aligning brand messages with these cultural preferences.

Interactive features having less impact on the popularity of brand posts in both languages suggests that businesses should focus on crafting content that is visually appealing and culturally relevant rather than solely relying on interactive elements. Companies can benefit from tracking and analyzing user emotions to create culture-oriented social media campaigns that strike a chord with their target audience. By understanding the emotional triggers of users from different cultures, brands can

refine their social media strategies to create content that resonates more deeply and drives higher user engagement, measured through likes, shares, and comments.

This study provides valuable insights for social media marketers and brand managers, emphasizing the importance of emotional appeal and cultural sensitivity in crafting effective social media campaigns. By considering the differences in emotional preferences and cultural norms between different language-speaking audiences, companies can create more impactful content that drives higher user engagement. This approach can ultimately lead to stronger brand loyalty, increased brand reach, and enhanced brand-consumer relationships, contributing to the overall success of the company's social media marketing efforts.

5.2 Theoretical and Methodological Implications

The variance in emoji selection and usage across different cultures, specifically among English-speaking and Turkish-speaking Twitter users within the context of brand social media communication is assessed. Detailed in Chapter 2, in order to address three research questions regarding emoji preferences and popularity across these two languages, alternative methodologies rooted in information theory and distributional measures are employed. In addition to building upon prior research on the cultural utilization of emojis, this study sampled from a more specific Twitter message context spanning a broader time period. The findings presented indicate that there are both similarities and notable differences in the aggregate preferences of these two communities. However, when analyzing the preferences over time, it becomes evident that the usage patterns are becoming more similar in terms of divergence and popularity. As it is shown, combination of interdisciplinary approaches to such big data advance the theory on consumer behavior, observing the actual behavior of users instead of limiting the scope to a number of participants and languages. It is also shown that emojis represent a combination of emotions that varies according to culture and language.

A comprehensive framework capable of annotating textual data and extracting emotions across any language, context, or time is developed. The study described in Chapter 3 focused specifically on the Turkish and English languages, utilizing a curated data set of social media messages in the context of branding communication. While annotated data sets for emotions are readily available in English, they are

limited when it comes to more specific languages such as Turkish. Furthermore, even the available data sets in English or Turkish lack contextual specificity. In addition to addressing these gaps, the study evaluates state-of-the-art classifiers not only based on their performance but also their training time, in order to enhance the overall structure of the proposed framework. The utilization of information from other domains, such as lexicons, to create annotated data is demonstrated.

Factors that contribute to higher user engagement with multinational brands' Twitter posts are identified. Analyzing both the structural and textual aspects of Twitter posts, which were categorized into vividness, practicality, interactivity, information, entertainment, and emotional appeal (happiness, anger, sadness, surprise, fear, and disgust), Brand post popularity is measured using a modified conceptual model and operationalized as the sum of likes, shares, and comments. Unlike previous studies, this research compares multinational brands' Twitter posts in English and Turkish languages, providing a cross-cultural perspective on the analysis of these features. To assess the significance of the factors, a bootstrap resampling procedure using a random subspace of features is proposed. The key findings of this study described in Chapter 4 indicate that there are significant differences in the most prominent determinants of brand post popularity across English and Turkish speaking markets, and moreover, some of the estimates are in contrast with previous work. This shows the necessity for marketing researchers to be cautious on the results of previous studies, and re-align their models according to specific cultural contexts.

The proposed approach in Chapter 4 yields higher explained variance and lower error rates compared to previous models, validating its effectiveness. This method can be applied to non-normally distributed data where the assumptions of the OLS model are not met. Examining the cultural impact on the factors influencing brand post popularity, the findings reveal the variations in the effects of vividness, practicality, interactivity, information, and entertainment features across languages, suggesting that these features may be perceived differently by consumers in different cultural settings. Finally, evaluating cross-cultural differences in the features of brand Twitter posts, the study also analyzes the emotional tone of brand messages across cultures. The findings demonstrate cultural variations in the effects of happiness, sadness, anger, surprise, fear, and disgust based on language, indicating that emotions may be perceived differently by consumers in diverse cultural settings.

5.3 Managerial Implications

The study illuminates the intricate nature of emojis as representations that encapsulate a combination of emotions, highlighting their propensity to vary across diverse cultures and languages. From a managerial standpoint, emojis can serve as valuable indicators of sentiment or emotion within consumer messages, providing insights into consumers' emotional states and attitudes. However, it is crucial for managers to recognize the dynamic nature of emojis and continuously adapt their computational methods to account for potential shifts in their interpretation and cultural meaning.

Furthermore, the study makes a significant contribution by introducing a comprehensive framework that automates the extraction of emotions. This framework goes beyond sentiment analysis and can be applied across various managerial contexts, including customer relationship management, brand management, and public opinion mining. By leveraging this framework, managers can streamline and enhance their understanding of the emotional content embedded within consumer communications. This, in turn, enables them to make informed decisions, develop effective strategies, and tailor their marketing efforts to better resonate with their target audience.

The findings of the study underscore the existence of cultural variations in the effects of specific emotions such as happiness, sadness, anger, surprise, fear, and disgust. It becomes evident that emotions are perceived and experienced differently by consumers across diverse cultural settings. These cultural nuances necessitate a deep understanding of the cultural context when crafting marketing strategies and designing brand communications. By considering cultural differences, managers can adapt their approaches, ensuring that their messages align with the emotional sensibilities and preferences of their target audience, leading to increased resonance and engagement.

Chapter 4 of the study offers invaluable insights from a managerial perspective. It proposes an approach that empowers brand managers to optimize their social media strategies and enhance brand post popularity. By incorporating factors such as vividness, practicality, interactivity, entertainment, and information, brand managers can craft posts that are more engaging and impactful, effectively capturing the attention and interest of their target audience. Understanding the dynamics and combined influence of these factors provides managers with the tools to drive higher levels of engagement, increase brand visibility, and achieve their marketing objectives. Moreover, the integration of emotional appeal in brand posts establishes a deeper connection with users and contributes to increased popularity. By creating content that elicits emotional responses, such as happiness or sadness, brands can

forge a stronger emotional bond with their audience, fostering brand loyalty, positive associations, and a sense of shared values.

In conclusion, the study highlights the cultural and linguistic nuances surrounding the interpretation and impact of emotions in social media branding communication, while also providing managerial insights for leveraging their potential in understanding consumer sentiments and optimizing brand strategies. By recognizing and adapting to cultural variations, and by harnessing emotional appeal in social media brand communications, managers can create meaningful and resonant experiences for their audiences, ultimately driving greater engagement, loyalty, and success in the marketplace.

While this research makes notable contributions to various aspects of the field, it is essential to acknowledge its limitations and identify potential avenues for future research.

5.4 Limitations and Future Research Directions

While strong implications are expected as a result of the contributions and findings of this dissertation, certain limitations should still be acknowledged. First, it is important to note that this study focuses exclusively on the communication platform of Twitter. It is possible that brands and consumers may exhibit different behaviors and interactions on other social media platforms such as Facebook or Instagram. However, this narrow focus on Twitter can also be seen as advantageous. Due to the platform's character limit of 280 (previously 140 until 2017), users may rely more on non-verbal cues, such as emojis, within the constrained communication environment. This heightened usage of emojis allows researchers to observe and explore a diverse range of emotional patterns.

Second, the choice of the Turkish language in this study represents a cultural context that lies between English-speaking countries and Asian-speaking countries in terms of individualism and collectivism. Considering that previous models assessing the individualism-collectivism dimensions of countries have relied on samples that may not be fully representative, often consisting of well-educated white-collar participants, it is important to note that Turkey has been undergoing a shift from institutional collectivism to in-group collectivism since 1995 (Pagda et al., 2021). This implies that Turkish individuals may exhibit less cohesion with society as a whole

but demonstrate stronger loyalty to sub-societal groups such as family, resulting in a unique blend of individualistic and collectivist tendencies. Just as preferences may differ across different social media platforms, they may also vary among countries like Turkey, which occupy a middle ground in terms of collectivism. Consequently, the first and second limitations discussed above make it challenging to generalize the findings of this study. To address this, future research can aggregate data from alternative platforms and languages to explore if similar patterns exist across various cultures through the lens of social media. The methodology employed in this study can be readily applied to other platforms and languages, requiring no additional efforts apart from expanding the data set as long as the data is publicly accessible.

Third limitation is the limited cultural generalizability resulting from the comparison of only two languages. Since the analysis primarily focuses on English and Turkish data, the findings may not be readily applicable for generalizing patterns in other Eastern or Western cultures. To address this limitation, future research should strive to incorporate a more diverse range of languages to capture cross-cultural variations in brand post popularity and interactions.

Fourth limitation pertains to the specific time scale of the data used in the analysis, which might not fully reflect the most recent trends and dynamics in brand post popularity. Future research can enhance the validity of the approach by updating the data with a more current time scale, enabling a more accurate understanding of how brand post interactions evolve over time. Additionally, the current study does not explicitly examine the impact of time on brand post popularity. Considering the temporal dynamics and trends in social media engagement would be valuable for future investigations. Exploring how the popularity of brand posts changes over time, while considering factors such as seasonality, trending topics, or specific events, would yield deeper insights into the dynamics of online brand interactions.

In this study, observational Twitter data is utilized to examine the dynamics of English and Turkish speaking markets. While the findings shed light on some intriguing similarities and differences in their behaviors and preferences, future research should consider incorporating complementary methodologies such as interviews and focus groups. These additional methods can provide a more comprehensive understanding of the underlying factors influencing consumer behavior, brand perceptions, and emotional responses in both markets. By delving deeper into the perspectives and experiences of individuals through qualitative approaches, researchers can better validate and elucidate the intricate dynamics that shape the distinct characteristics of each market. Moreover, such an integrated approach would not only enhance the credibility of the findings but also pave the way for more tailored and effective

marketing strategies in both English and Turkish speaking regions.

Finally, this study only analyzes a subset of features and emotions associated with brand posts. To obtain a more comprehensive understanding of the factors influencing post popularity, future research should consider incorporating additional features, such as sentiment analysis of accompanying photos or videos, and explore the values conveyed through brand posts.

In conclusion, this dissertation has successfully illuminated various aspects of brand post popularity, shedding light on the intricate relationship between emotions and consumer responses. The insights derived from this research offer valuable contributions to the field, enhancing our understanding of the factors that influence brand post engagement. However, to further enrich our knowledge in this domain, it is imperative for future studies to broaden their scope by incorporating a more diverse range of languages. By encompassing a wider linguistic landscape, researchers can capture the nuances of cross-cultural variations in brand post popularity, exploring how cultural and linguistic factors interplay to shape consumer perceptions and responses. By including languages from different regions and cultures, future studies can provide a more comprehensive and nuanced understanding of the intricate dynamics at play in brand post interactions. This broader language representation would not only deepen our knowledge of cultural influences on consumer behavior but also enable marketers to develop more effective and culturally sensitive social media marketing strategies.

In today's dynamic and interconnected digital landscape, analyzing the factors that influence consumer engagement with social media brand posts is of paramount importance for marketers. However, the complexity deepens when considering cross-cultural variations in consumer behavior and preferences. Factors such as vividness, interactivity, practicality, information, entertainment, and emotional appeal play vital roles in determining the success of brand posts across different cultures. By research conducting cross-cultural analyses, marketers can gain valuable insights into how these factors impact consumer engagement in diverse cultural contexts. This knowledge will empower brands to create social media content that resonates with consumers from various cultural backgrounds, leading to increased brand visibility, stronger customer engagement, and ultimately, a competitive advantage in the global marketplace.

5.5 Concluding Remarks

The dissertation aims to provide valuable perspectives on comprehending the substantial differences in consumer preferences between a more individualistic (English speaking) and more collectivist (Turkish speaking) markets concerning emotions and branding communication on social media platforms. Examining of cross-cultural differences concerning the impact of emotions in social media branding communication, an interdisciplinary set of approaches from Branding Communication, Machine Learning and Information Theory are employed to answer the overarching research question of:

RQ: What is the impact of emotions on brand communication in social media across cultures/languages?

While answering the main RQ, the empirical gap in cultural perspective in earlier studies in the context of branding communication is also addressed analyzing a large sample of Twitter messages of multinational brand accounts from various industries in English and Turkish languages from different aspects.

First, emojis, as visual representations of emotions, encompass a combination of feelings that can vary significantly depending on cultural and linguistic factors. The cultural and linguistic diversity across different regions and languages implies that the interpretation and significance of emojis may differ across cultures. For instance, a specific emoji might evoke a particular emotion in one culture but be associated with a different emotion in another culture. This highlights the importance of considering cultural context when analyzing the emotional impact of emojis in cross-cultural studies.

Additionally, the use of emojis provides a convenient and language-independent approach for annotating textual data. Unlike traditional text-based annotations, emojis offer a visual representation that transcends language barriers, making them accessible and easily understood by individuals regardless of their language proficiency. This allows for automated annotation and analysis of textual data in a standardized manner across various languages, times, and contexts, enhancing the efficiency and comparability of research findings.

Furthermore, empirical findings indicate significant cultural variations in the effects of specific emotions such as happiness, sadness, anger, surprise, fear, and disgust. These findings suggest that the perception and interpretation of emotions differ among consumers in diverse cultural settings. For instance, a brand's use of an emoji conveying happiness might elicit different emotional responses and consumer attitudes in one culture compared to another. These cultural variations in emotional responses emphasize the importance of considering cultural nuances and contextual

factors when evaluating the emotional effects of various stimuli, including emojis. By acknowledging and accounting for these cultural differences, marketers and researchers can better tailor their strategies and communications to effectively engage and resonate with diverse audiences across different cultural backgrounds.

In conclusion, this dissertation has provided valuable insights into the realm of emotions and their impact on social media branding communication across English and Turkish speaking markets. The social media data utilized in this research has unveiled intriguing similarities and differences in emotional responses among consumers in both cultural settings. Nevertheless, it is evident that these emotions are subject to significant cultural variations, which underscore the importance of considering cultural nuances and contextual factors in future research and marketing strategies. Investigating the role of cultural values and societal norms on emotional responses, and exploring how cultural differences influence the effectiveness of various emotional appeals in marketing campaigns requires further attention by future research to understand consumer behaviors and emotional responses in diverse cultural settings. By acknowledging and embracing these cultural differences, marketers can better connect with their target audiences, fostering more authentic and relatable brand experiences. Implementing culturally sensitive communication approaches will undoubtedly yield more successful and impactful social media branding campaigns in the diverse and ever-evolving global marketplace.

BIBLIOGRAPHY

- Affi, A., Kotlerman, J. B., Ettner, S. L., & Cowan, M. (2007). Methods for improving regression analysis for skewed continuous or counted responses. *Annu. Rev. Public Health*, 28, 95–111.
- Al-Omari, H., Abdullah, M. A., & Shaikh, S. (2020). Emotet2: Emotion detection in english textual dialogue using bert and bilstm models. In *2020 11th International Conference on Information and Communication Systems (ICICS)*, (pp. 226–232). IEEE.
- Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, (pp. 579–586).
- Altmay, G. & Turhan, Ç. (2022). Emotion analysis on turkish texts: A systematic mapping study. In *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, (pp. 523–527). IEEE.
- Aman, S. & Szpakowicz, S. (2007). Identifying expressions of emotion in text. In *Text, Speech and Dialogue: 10th International Conference, TSD 2007, Pilsen, Czech Republic, September 3-7, 2007. Proceedings 10*, (pp. 196–205). Springer.
- Antoniadis, I., Paltsoglou, S., & Patoulidis, V. (2019). Post popularity and reactions in retail brand pages on facebook. *International Journal of Retail & Distribution Management*.
- Anusha, V. & Sandhya, B. (2015). A learning based emotion classifier with semantic text processing. In *Advances in intelligent informatics*, (pp. 371–382). Springer.
- Asghar, M. Z., Khan, A., Bibi, A., Kundi, F. M., & Ahmad, H. (2017). Sentence-level emotion detection framework using rule-based classification. *Cognitive Computation*, 9(6), 868–894.
- Baayen, R. H. (2001). *Word frequency distributions*, volume 18. Springer Science Business Media.
- Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A systematic review of emoji: Current research and future perspectives. *Frontiers in Psychology*, 10, 2221.
- Balabantaray, R. C., Mohammad, M., & Sharma, N. (2012). Multi-class twitter emotion classification: A new approach. *International Journal of Applied Information Systems*, 4(1), 48–53.
- Banerjee, S. & Chua, A. Y. (2019). Identifying the antecedents of posts’ popularity on facebook fan pages. *Journal of Brand Management*, 26, 621–633.
- Bao, Y., Quan, C., Wang, L., & Ren, F. (2014). The role of pre-processing in twitter sentiment analysis. In *Intelligent Computing Methodologies: 10th International Conference, ICIC 2014, Taiyuan, China, August 3-6, 2014. Proceedings 10*, (pp. 615–624). Springer.
- Baziotis, C., Athanasiou, N., Chronopoulou, A., Kolovou, A., Paraskevopoulos, G., Ellinas, N., Narayanan, S., & Potamianos, A. (2018). Ntua-slp at semeval-2018 task 1: Predicting affective content in tweets with deep attentive rnns and transfer learning. *arXiv preprint arXiv:1804.06658*.
- Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-lstm: Convolutional

- lstm model for sentiment analysis in social big data. *Information Processing & Management*, 58(1), 102435.
- Ben-Ze'ev, A. (2001). *The subtlety of emotions*. MIT press.
- Bengio, Y. et al. (2009). Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, 2(1), 1–127.
- Berry, W. D. (1993). *Understanding regression assumptions*, volume 92. Sage.
- Bishara, A. J. & Hittner, J. B. (2017). Confidence intervals for correlations when data are not normal. *Behavior research methods*, 49, 294–309.
- Boot, T. & Nibbering, D. (2019). Forecasting using random subspace methods. *Journal of Econometrics*, 209(2), 391–406.
- Boynukalin, Z. & Karagoz, P. (2013). *Emotion analysis on Turkish texts*, (pp. 159–168). Springer, Charm.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24, 123–140.
- Calvo, R. A. & Mac Kim, S. (2013). Emotions in text: dimensional and categorical models. *Computational Intelligence*, 29(3), 527–543.
- Can, F., Kocberber, S., Balcik, E., Kaynak, C., Ocalan, H. C., & Vursavas, O. M. (2008). Information retrieval on turkish texts. *Journal of the American Society for Information Science and Technology*, 59(3), 407–421.
- Chaffar, S. & Inkpen, D. (2011). Using a heterogeneous dataset for emotion analysis in text. In *Advances in Artificial Intelligence: 24th Canadian Conference on Artificial Intelligence, Canadian AI 2011, St. John's, Canada, May 25-27, 2011. Proceedings 24*, (pp. 62–67). Springer.
- Chang, C. (2016). Efl reviewers' emoticon use in asynchronous computer- mediated peer response. *Computers and Composition*, 40, 1–18.
- Chatterjee, A., Narahari, K. N., Joshi, M., & Agrawal, P. (2019). Semeval-2019 task 3: Emocontext contextual emotion detection in text. In *Proceedings of the 13th international workshop on semantic evaluation*, (pp. 39–48).
- Chu, S.-C. & Choi, S. M. (2011). Electronic word-of-mouth in social networking sites: A cross-cultural study of the united states and china. *Journal of Global Marketing*, 24(3), 263–281.
- Chua, A. Y. & Banerjee, S. (2015). How businesses draw attention on facebook through incentives, vividness and interactivity. *IAENG International Journal of Computer Science*, 42(3), 275–281.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4), 661–703.
- Colombetti, G. (2009). From affect programs to dynamical discrete emotions. *Philosophical Psychology*, 22(4), 407–425.
- Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston, MA.
- Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273–297.
- Coyle, M. A. & Carmichael, C. L. (2019). Perceived responsiveness in text messaging: The role of emoji use. *Computers in Human Behavior*, 99, 181–189.
- Crosby, L. & Johnson, S. (2007). Experience required. *Marketing Management*,

- 16(4), 20–28.
- Darwin, C. & P., P. (1998). *The expression of the emotions in man and animals*. USA: Oxford University Press.
- Das, G., Wiener, H. J., & Kareklas, I. (2019). To emoji or not to emoji? examining the influence of emoji on consumer reactions to advertising. *Journal of Business Research*, 96, 147–156.
- DasGupta, A. (2008). *Asymptotic theory of statistics and probability*, volume 180. Springer.
- Davis, M. & Holbrook, N. (2021). Unicode emoji. unicode technical reports. Report.
- De Vries, L., Gensler, S., & Leeftang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83–91.
- Delia, E. B. & Armstrong, C. G. (2015). sponsoring the frenchopen: An examination of social media buzz and sentiment. *Journal of Sport Management*, 26, 184–199.
- Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018a). BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018b). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Diaz-Diaz, N., Aguilar-Ruiz, J. S., Nepomuceno, J. A., & Garcia, J. (2005). Feature selection based on bootstrapping. In *2005 ICSC Congress on Computational Intelligence Methods and Applications*, (pp. 6–pp). IEEE.
- Dresner, E. & Herring, S. C. (2010). Functions of the nonverbal in cmc: Emoticons and illocutionary force. *Communication theory*, 20(3), 249–268.
- Edelman, D. (2010). Branding in the digital age. *Harvard Business Review*, 88(12), 62–69.
- Efron, B. & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Ekinici, Y., Yoon, T. H., & Oppewal, H. (2005). An examination of the brand relationship quality scale in the evaluation of restaurant brands. in j. chen (ed.). *Advances in hospitality and leisure*, 1, 189–197.
- Ekman, P. (2005). *Basic Emotions*, (pp. 45–60).
- Endres, D. M. & Schindelin, J. E. (2003). A new metric for probability distributions. *IEEE Transactions on Information theory*, 49(7), 1858–1860.
- Evans, V. (2017). *The Emoji Code: how smiley faces, love hearts and thumbs up are changing the way we communicate*. Michael O’Mara Books.
- Gawne, L. & G., M. (2019). Emoji as digital gestures. *Language@ Internet*, 17(2).
- Geurin, A. N. & Burch, L. M. (2017). User-generated branding via social media: An examination of six running brands. *Sport Management Review*, 20(3), 273–284.
- Gill, A. J., French, R. M., Gergle, D., & Oberlander, J. (2008). Identifying emotional characteristics from short blog texts. In *30th Annual conference of the cognitive science society*, (pp. 2237–2242). Citeseer.
- Glikson, E., Cheshin, A., & Kleef, G. A. V. (2018). The dark side of a smiley: Effects of smiling emoticons on virtual first impressions. *Social Psychological and Personality Science*, 9(5), 614–625.
- Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for large-scale

- sentiment classification: A deep learning approach. In *28th International Conference on Machine Learning*.
- Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N project report, Stanford*, 1(12), 2009.
- GobalStats (2023). Social media stats may 2022-2023.
- Goodrich, K. & De Mooij, M. (2014). How ‘social’ are social media? a cross-cultural comparison of online and offline purchase decision influences. *Journal of marketing communications*, 20(1-2), 103–116.
- Green, E. G., Deschamps, J. C., & Paez, D. (2005). Variation of individualism and collectivism within and between 20 countries: A typological analysis. *Journal of cross-cultural psychology*, 36(3), 321–339.
- Gudykunst, W. B. (1997). Cultural variability in communication: An introduction. *Communication research*, 24(4), 327–348.
- Gudykunst, W. B. (2005). *Theorizing about intercultural communication*. Sage.
- Gudykunst, W. B., Matsumoto, Y., Ting-Toomey, S., Nishida, T., Kim, K., & S., H. (1996). The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures. *Human communication research*, 22(4), 510–543.
- Guntuku, S. C., Li, M., Tay, L., & Ungar, L. H. (2019). Studying cultural differences in emoji usage across the east and the west. In *International AAAI Conference on Web and Social Media*.
- Hajar, M. (2016). Using youtube comments for text-based emotion recognition. *Procedia Computer Science*, 83, 292–299.
- Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, (pp. 929–932).
- Harris, Z. S. & Zellig, S. (1981). *Co-occurrence and transformation in linguistic structure.*, (pp. 143–210). Dordrecht: Springer.
- Hasan, M., Rundensteiner, E., & Agu, E. (2014). Emotex: Detecting emotions in twitter messages.
- Herring, S. C. & Dainas, A. R. (2018). Receiver interpretations of emoji functions: A gender perspective. In *1st International Workshop on Emoji Understanding and Applications in Social Media (Emoji2018)*.
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
- Hinde, J. & Demétrio, C. G. (1998). Overdispersion: models and estimation. *Computational statistics data analysis*, 27(2), 151–170.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8), 832–844.
- Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Hockenbury, D. & Hockenbury, S. (2007). *Discovering Psychology*. New York: Worth Publishers.
- Hofstede, G. (2011). Dimensionalizing cultures: The hofstede model in context. *Online readings in psychology and culture*, 2(1).
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind* ((Rev. 3rd ed.) ed.). New York: McGraw-Hill.
- Horowitz, J. L. (2019). Bootstrap methods in econometrics. *Annual Review of*

- Economics*, 11, 193–224.
- Hsieh, S. H. & Tseng, T. H. (2017). Playfulness in mobile instant messaging: Examining the influence of emoticons and text messaging on social interaction. *Computers in Human Behavior*, 69, 405–414.
- Huang, A. H., Yen, D. C., & Zhang, X. (2008). Exploring the potential effects of emoticons. *Information Management*, 45(7), 466–473.
- Huang, C., Trabelsi, A., & Zaïane, O. R. (2019). Ana at semeval-2019 task 3: Contextual emotion detection in conversations through hierarchical lstms and bert. *arXiv preprint arXiv:1904.00132*.
- Huang, Y.-H., Lee, S.-R., Ma, M.-Y., Chen, Y.-H., Yu, Y.-W., & Chen, Y.-S. (2019). Emotionx-idea: Emotion bert—an affectional model for conversation. *arXiv preprint arXiv:1908.06264*.
- Hudson, S., Roth, M. S., Madden, T. J., & Hudson, R. (2015). The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism management*, 47, 68–76.
- Hwangbo, H., Sharma, V., Arndt, C., & TerMaath, S. (2023). A randomized subspace-based approach for dimensionality reduction and important variable selection. *Journal of Machine Learning Research*, 24, 1–30.
- Höpfner, R. (2014). *Asymptotic Statistics*. Berlin, Boston: De Gruyter.
- Jaeger, S. R., Roigard, C. M., Jin, D., Vidal, L., & Ares, G. (2019). Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food Research International*, 119, 895–907.
- Jaeger, S. R., Xia, Y., Lee, P. Y., Hunter, D. C., Beresford, M. K., & Ares, G. (2018). Emoji questionnaires can be used with a range of population segments: Findings relating to age, gender and frequency of emoji/emoticon use. *Food Quality and Preference*, 68, 397–410.
- Jain, V. K., Kumar, S., & Fernandes, S. L. (2017). Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. *Journal of computational science*, 21, 316–326.
- Jin, X., Li, Y., Mah, T., & Tong, J. (2007). Sensitive webpage classification for content advertising. In *Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising*, (pp. 28–33).
- Johnson, W. (1944). I. a program of research. *Psychological Monographs*, 56(2), 1.
- Karpinska-Krakiwicz, M. & Modlinski, A. (2020). Popularity of branded content in social media. *Journal of Computer Information Systems*, 60(4), 309–315.
- Kavanagh, B. (2010). A cross-cultural analysis of japanese and english non-verbal online communication: The use of emoticons in weblogs. *Intercultural Communication Studies*, 19(3), 65–80.
- Kejriwal, M., Wang, Q., Li, H., & Wang, L. (2021). An empirical study of emoji usage on twitter in linguistic and national contexts. *Online Social Networks and Media*, 24.
- Kim, J. & Gupta, P. (2012). Emotional expressions in online user reviews: How they influence consumers’ product evaluations. *Journal of Business Research*, 65(7), 985–992.
- Kingma, D. P. & Ba, J. (2017). Adam: A method for stochastic optimization.
- Koplenig, A. (2019). A non-parametric significance test to compare corpora. *Plos one*, 14(9: e0222703).
- Krohn, F. B. (2004). A generational approach to using emoticons as nonverbal

- communication. *Journal of technical writing and communication*, 34(4), 321–328.
- Kuha, J. (2004). Aic and bic: Comparisons of assumptions and performance. *Sociological methods research*, 33(2), 188–229.
- Kujur, F. & Singh, S. (2016). Social networking sites as a multimedia tool for brand popularity—an exploratory study. *Kujur, F., & Singh, S.(2016). Social Networking Sites as a Multimedia Tool for Brand Popularity—An Exploratory Study. Indian Journal of Science and Technology*, 9, 45.
- Laverie, D. A., Kleine, R. E., & Kleine, S. S. (1993). Linking emotions and values in consumption experiences: An exploratory study. *Advances in Consumer Research, Association for Consumer Research*, 20, 70–75.
- Lazemi, S. & Ebrahimpour-Komleh, H. (2018). Multi-emotion extraction from text using deep learning. *International Journal of Web Research*, 1(1), 62–67.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436–444.
- Lee, C. & Hsieh, H. (2019). Engaging consumers in mobile instant messaging: the role of cute branded emoticons. *Journal of Product Brand Management*.
- Lee, C. & Kahle, L. (2016). The linguistics of social media: Communication of emotions and values in sport. *Sport Marketing Quarterly*, 25(4).
- Li, F., Pan, S. J., Jin, O., Yang, Q., & Zhu, X. (2012). Cross-domain co-extraction of sentiment and topic lexicons. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (pp. 410–419).
- Li, H., Lee, Y.-C., Zhou, Y.-C., & Sun, J. (2011). The random subspace binary logit (rsbl) model for bankruptcy prediction. *Knowledge-Based Systems*, 24(8), 1380–1388.
- Li, M., Chng, E., Chong, A. Y. L., & See, S. (2018). Multi-class twitter sentiment classification with emojis. *Industrial Management Data Systems.*, 118(9).
- Li, M., Chng, E., Chong, A. Y. L., & See, S. (2019). An empirical analysis of emoji usage on twitter. *Industrial Management Data Systems.*, 119(8), 1748–1763.
- Li, Y. M., Lin, L., & Chiu, S. W. (2014). Enhancing targeted advertising with social context endorsement. *International Journal of Electronic Commerce*, 19(1), 99–128.
- Lin, H.-C., Swarna, H., & Bruning, P. F. (2017). Taking a global view on brand post popularity: Six social media brand post practices for global markets. *Business Horizons*, 60(5), 621–633.
- Lindner, T., Puck, J., & Verbeke, A. (2020). Misconceptions about multicollinearity in international business research: Identification, consequences, and remedies.
- Lita, L. V., Schlaikjer, A. H., Hong, W., & Nyberg, E. (2005). Qualitative dimensions in question answering: Extending the definitional qa task. In *Proceedings of the national conference on artificial intelligence*, volume 20, (pp. 1616). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.
- Liu, K. L., Li, W. J., & Guo, M. (2012). Emoticon smoothed language models for twitter sentiment analysis. In *AAAI Conference on Artificial Intelligence*, volume 26, (pp. 1678–1684).
- Ljubešić, N. & Fišer, D. (2016). A global analysis of emoji usage. In *10th web as corpus workshop*, (pp. 82–89).
- Louppe, G. & Geurts, P. (2012). Ensembles on random patches. In *Machine Learning*

- and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24-28, 2012. *Proceedings, Part I* 23, (pp. 346–361). Springer.
- MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.
- Manganari, E. E. (2021). Emoji use in computer-mediated communication. *The International Technology Management Review*, 10(1), 1–11.
- Mazloom, M., Rietveld, R., Rudinac, S., Worring, M., & Van Dolen, W. (2016). Multimodal popularity prediction of brand-related social media posts. In *Proceedings of the 24th ACM international conference on Multimedia*, (pp. 197–201).
- McCracken, M. W. & Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574–589.
- McElduff, F., Cortina-Borja, M., Chan, S. K., & Wade, A. (2010). When t-tests or wilcoxon-mann-whitney tests won’t do. *Advances in physiology education*, 34(3), 128–133.
- Miller, D. A., Cronin, T., Garcia, A. L., & Branscombe, N. R. (2009). The relative impact of anger and efficacy on collective action is affected by feelings of fear. *Group Processes Intergroup Relations*, 12(4), 445–462.
- Mohammad, S., Bravo-Marquez, F., Salameh, M., & Kiritchenko, S. (2018). Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, (pp. 1–17).
- Mohammad, S. & Turney, P. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, (pp. 26–34).
- Mohammad, S. M. & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence*, 29(3), 436–465.
- Newman, M. E. (2005). Power laws, pareto distributions and zipf’s law. *Contemporary physics*, 46(5), 323–351.
- Ng, S. (2017). Opportunities and challenges: Lessons from analyzing terabytes of scanner data.
- Nisbett, R. E. & Masuda, T. (2003). Culture and point of view. In *National Academy of Sciences*, volume 100, (pp. 1163–11170).
- Novak, K., Smailović, J., Sluban, B., & Mozetič, I. (2015a). Emoji sentiment ranking. http://kt.ijs.si/data/Emoji_sentiment_ranking/. Available at:.
- Novak, K., Smailović, J., Sluban, B., & Mozetič, I. (2015b). Sentiment of emojis. *PloS one*, 10(12:e0144296).
- Oflazer, K. (1994). Two-level description of turkish morphology. *Literary and linguistic computing*, 9(2), 137–148.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological bulletin*, 128(1), 3.
- Pagda, Z., Bayraktar, S., & Jimenez, A. (2021). Exploring culture and leadership after 23 years: A replication of globe project in turkey. *Journal of International Management*, 27(1).
- Pak, A., Paroubek, P., et al. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, (pp. 1320–1326).

- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.
- Park, J., Baek, Y. M., & Cha, M. (2014). Cross-cultural comparison of nonverbal cues in emoticons on twitter: Evidence from big data analysis. *Journal of Communication*, 64(2), 333–354.
- Park, J., Barash, V., Fink, C., & Cha, M. (2014). Emoticon style: Interpreting differences in emoticons across cultures. In *International AAAI Conference on Web and Social Media*, volume 7, (pp. 466–475).
- Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, (pp. 1532–1543).
- Perikos, I. & Hatzilygeroudis, I. (2013). Recognizing emotion presence in natural language sentences. In *Engineering Applications of Neural Networks: 14th International Conference, EANN 2013, Halkidiki, Greece, September 13-16, 2013 Proceedings, Part II 14*, (pp. 30–39). Springer.
- Pinker, S. (1999). How the mind works. *Great Issues for Medicine in the Twenty-First Century*, 882, 119–127.
- Pinto, M. B. & Yagnik, A. (2017). Fit for life: A content analysis of fitness tracker brands use of facebook in social media marketing. *Journal of Brand Management*, 24, 49–67.
- Pletikosa Cvijikj, I. & Michahelles, F. (2013). Online engagement factors on facebook brand pages. *Social network analysis and mining*, 3, 843–861.
- Plutchik, R. (1984). In search of the basic emotions. *Contemporary Psychology: A Journal of Reviews*, 29(6), 511–513.
- Prada, M., Rodrigues, D. L., Garrido, M. V., Lopes, D., Cavalheiro, B., & Gaspar, R. (2018). Motives, frequency and attitudes toward emoji and emoticon use. *Telematics and Informatics*, 35(7), 1925–1934.
- Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J., & Chen, C. (2010). Dasa: Dissatisfaction-oriented advertising based on sentiment analysis. *Expert Systems with Applications*, 37(9), 6182–6191.
- Rabbi, F., Khalil, A., Khan, I., Almuqrin, M. A., Khalil, U., & Andualem, M. (2022). Robust model selection using the out-of-bag bootstrap in linear regression. *Scientific Reports*, 12(1), 10992.
- Rahman, Z., Suberamanian, K., Zauddin, H., Moghavvemi, S., & Nasir, M. (2017). Fanpage viral metrics analysis: Study on frequently posted contents. *Journal of Engineering and Applied Sciences*, 12(16), 4039–4046.
- Rao, Y., Pang, J., Xie, H., Liu, A., Wong, T.-L., Li, Q., & Wang, F. L. (2017). Supervised intensive topic models for emotion detection over short text. In *Database Systems for Advanced Applications: 22nd International Conference, DASFAA 2017, Suzhou, China, March 27-30, 2017, Proceedings, Part I 22*, (pp. 408–422). Springer.
- Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL student research workshop*, (pp. 43–48).
- Redmond, M., Salesi, S., & Cosma, G.
- Refaee, E. & Rieser, V. (2014). Can we read emotions from a smiley face? emoticon-based distant supervision for subjectivity and sentiment analysis of arabic twitter feeds. In *5th International Workshop on Emotion, Social Signals, Sen-*

- timent and Linked Open Data, LREC.*
- Riloff, E. (1996). Automatically generating extraction patterns from untagged text. In *Proceedings of the national conference on artificial intelligence*, (pp. 1044–1049).
- Riloff, E., Wiebe, J., & Phillips, W. (2005). Exploiting subjectivity classification to improve information extraction. In *Proceedings of the 20th national conference on Artificial intelligence*, volume 20, (pp. 1106–1111).
- Rosenthal, S., Farra, N., & Nakov, P. (2019). Semeval-2017 task 4: Sentiment analysis in twitter. *arXiv preprint arXiv:1912.00741*.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161.
- Sabate, F., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R. (2014). Factors influencing popularity of branded content in facebook fan pages. *European management journal*, 32(6), 1001–1011.
- Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv, abs/1910.01108*.
- Schacter, D. L., Gilbert, D. T., Wegner, D. M., & Hood, B. M. (2011). *Psychology (European ed)*. Basingstoke: Palgrave Macmillan.
- Scherer, K. R. & Wallbott, H. G. (1994). Evidence for universality and cultural variation of differential emotion response patterning. *Journal of personality and social psychology*, 66(2), 310.
- Schultz, C. D. (2016). Driving likes, comments, and shares on social networking sites: how post characteristics affect brand interactions in apparel retailing. In *Proceedings of the 18th Annual International Conference on Electronic Commerce: e-Commerce in Smart connected World*, (pp. 1–9).
- Schultz, C. D. (2017). Proposing to your fans: Which brand post characteristics drive consumer engagement activities on social media brand pages? *Electronic Commerce Research and Applications*, 26, 23–34.
- Schweter, S. (2020). Berturk-bert models for turkish. *Zenodo, 2020*, 3770924.
- Seki, Y., Eguchi, K., Kando, N., & Aono, M. (2005). Multi-document summarization with subjectivity analysis at duc 2005. In *Proceedings of the Document Understanding Conference (DUC)*.
- Seyeditabari, A., Tabari, N., Gholizade, S., & Zadrozny, W. (2019). Emotional embeddings: refining word embeddings to capture emotional content of words. *arXiv preprint arXiv:1906.00112*.
- Shaheen, S., El-Hajj, W., Hajj, H., & Elbassuoni, S. (2014). Emotion recognition from text based on automatically generated rules. In *2014 IEEE International Conference on Data Mining Workshop*, (pp. 383–392). IEEE.
- Shehu, H. A., Sharif, M. H., Sharif, M. H. U., Datta, R., Tokat, S., Uyaver, S., Kusetogullari, H., & Ramadan, R. A. (2021). Deep sentiment analysis: a case study on stemmed turkish twitter data. *IEEE Access*, 9, 56836–56854.
- Somasundaran, S., Wilson, T., Wiebe, J., & Stoyanov, V. (2007). Qa with attitude: Exploiting opinion type analysis for improving question answering in on-line discussions and the news. In *ICWSM*.
- Song, K., Feng, S., Gao, W., Wang, D., Chen, L., & Zhang, C. (2015). Build emotion lexicon from microblogs by combining effects of seed words and emoticons in a heterogeneous graph. In *Proceedings of the 26th ACM conference on hypertext social media*, (pp. 283–292).

- Speer, R., Chin, J., & Havasi, C. (2017). Conceptnet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31.
- Stoica, P. & Selen, Y. (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4), 36–47.
- Strapparava, C., . M. R. Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, (pp. 70–74).
- Strapparava, C., Valitutti, A., et al. (2004). Wordnet affect: an affective extension of wordnet. In *Lrec*, volume 4, (pp.40). Lisbon, Portugal.
- Swani, K., Milne, G., & P. Brown, B. (2013). Spreading the word through likes on facebook: Evaluating the message strategy effectiveness of fortune 500 companies. *Journal of Research in Interactive Marketing*, 7(4), 269–294.
- Swani, K. & Milne, G. R. (2017). Evaluating facebook brand content popularity for service versus goods offerings. *Journal of Business Research*, 79, 123–133.
- Takahashi, K., Oishi, T., & Masaki, S. (2017). Is smiling? cross-cultural study on recognition of emoticon’s emotion. *Journal of cross-cultural psychology*, 48(10), 1578–1586.
- Tang, D., Qin, B., Liu, T., & Li, Z. (2013). Learning sentence representation for emotion classification on microblogs. In *CCF International Conference on Natural Language Processing and Chinese Computing*, (pp. 212–223). Springer.
- Terveen, L., Hill, W., Amento, B., McDonald, D., & Creter, J. (1997). Phoaks: A system for sharing recommendations. *Communications of the ACM*, 40(3), 59–62.
- Thongmak, M. (2015). Engaging facebook users in brand pages: Different posts of marketing-mix information. In *Business Information Systems: 18th International Conference, BIS 2015, Poznań, Poland, June 24-26, 2015, Proceedings 18*, (pp. 299–308). Springer.
- Tocoglu, M. A. & Alpkocak, A. (2014). Emotion extraction from turkish text. In *2014 European Network Intelligence Conference*, (pp. 130–133). IEEE.
- Tocoglu, M. A. & Alpkocak, A. (2018). Tremo: A dataset for emotion analysis in turkish. *Journal of Information Science*, 44(6), 848–860.
- Tocoglu, M. A. & Alpkocak, A. (2019a). Lexicon-based emotion analysis in turkish. *Turkish Journal Of Electrical Engineering Computer Sciences*, 27(2), 1213–1227.
- Tocoglu, M. A. & Alpkocak, A. (2019b). Lexicon-based emotion analysis in turkish. *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(2), 1213–1227.
- Tocoglu, M. A., Ozturkmenoglu, O., & Alpkocak, A. (2019). Emotion analysis from turkish tweets using deep neural networks. *IEEE Access*, 7, 183061–183069.
- Tomihira, T., Otsuka, A., Yamashita, A., & Satoh, T. (2020). Multilingual emoji prediction using bert for sentiment analysis. *International Journal of Web Information Systems*, 16(3), 265–280.
- Tossell, C. C., Kortum, P., Shepard, C., Barg-Walkow, L. H., Rahmati, A., & Zhong, L. (2012). A longitudinal study of emoticon use in text messaging from smartphones. *Computers in Human Behavior*, 28(2), 659–663.
- Tweedie, F. J. & Baayen, R. H. (1998). How variable may a constant be? measures of lexical richness in perspective. *Computers and the Humanities*, 32(5), 323–

- Ucan, A., Dörterler, M., & Akçapınar Sezer, E. (2022). A study of turkish emotion classification with pretrained language models. *Journal of Information Science*, 48(6), 857–865.
- Umashanthi, P. & Jacob, E. (2015). Emoticons vs. emojis on twitter: A causal inference approach. *arXiv preprint, arXiv:1510.08480*.
- Uymaz, H. A. & Metin, S. K. (2023). Emotion-enriched word embeddings for turkish. *Expert Systems with Applications*, 120011.
- Vandergriff, I. (2013). Emotive communication online: A contextual analysis of computer-mediated communication (cmc) cues. *Journal of Pragmatics*, 51(1), 1–12.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Vidal, L., Ares, G., & Jaeger, S. R. (2016). Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference*, 49, 119–128.
- Vosoughi, S., Zhou, H., & Roy, D. (2016). Enhanced twitter sentiment classification using contextual information. *arXiv preprint arXiv:1605.05195*.
- Wang, G. & Ma, J. (2011). Study of corporate credit risk prediction based on integrating boosting and random subspace. *Expert Systems with Applications*, 38(11), 13871–13878.
- Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2012). Harnessing twitter" big data" for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, (pp. 587–592). IEEE.
- Wichitaksorn, N., Kang, Y., & Zhang, F. (2023). Random feature selection using random subspace logistic regression. *Expert Systems with Applications*, 119535.
- Xiang, W., Chen, S., Sun, L., Cheng, S., & Bove, J. V. M. (2016). Odor emoticon: an olfactory application that conveys emotions. *International Journal of Human-Computer Studies*, 91, 52–61.
- Ye, R., Du, W., & Lu, Y. (2023). Bootstrap inference for skew-normal unbalanced heteroscedastic one-way classification random effects model. *Journal of Statistical Computation and Simulation*, 1–31.
- Yu, C.-E. & Sun, R. (2019). The role of instagram in the unesco's creative city of gastronomy: A case study of macau. *Tourism Management*, 75, 257–268.
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the east and west? cultural differences in using the eyes and mouth as cues to recognize emotions in japan and the united states. *Journal of Experimental Social Psychology*, 43(2), 303–311.
- Zadeh, A. H. & Sharda, R. (2014). Modeling brand post popularity dynamics in online social networks. *Decision Support Systems*, 65, 59–68.
- Zahiri, S. M. & Choi, J. D. (2017). Emotion detection on tv show transcripts with sequence-based convolutional neural networks. *arXiv preprint arXiv:1708.04299*.
- Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., & Cannataro, M. (2020). Sentiment analysis for mining texts and social networks data: Methods and tools. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(1), e1333.