ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

MARKETING CAMPAIGN MANAGEMENT USING MACHINE LEARNING TECHNIQUES: AN UPLIFT MODELING APPROACH

Ph.D. THESIS

Meltem SANİSOĞLU

Department of Management

Management Programme

JUNE 2024

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

MARKETING CAMPAIGN MANAGEMENT USING MACHINE LEARNING TECHNIQUES: AN UPLIFT MODELING APPROACH

Ph.D. THESIS

Meltem SANİSOĞLU
(403172012)

Department of Management

Management Programme

Thesis Advisor: Prof. Dr. Şebnem BURNAZ

JUNE 2024

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PAZARLAMA KAMPANYASI YÖNETİMİ: ARTIMLI MODELLEME YAKLAŞIMI

DOKTORA TEZİ

Meltem SANİSOĞLU
(403172012)

İşletme Anabilim Dalı

İşletme Programı

Tez Danışmanı: Prof. Dr. Şebnem BURNAZ

HAZİRAN 2024

Meltem SANİSOĞLU, a Ph.D. student of İTU Graduate School student ID 403172012, successfully defended the dissertation entitled "MARKETING CAMPAIGN MANAGEMENT USING MACHINE LEARNING TECHNIQUES: AN UPLIFT MODELING APPROACH", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Şebnem BURNAZ** İstanbul Technical University	....
Jury Members :	**Prof. Dr. Tolga KAYA** Istanbul Technical University	....
	Assoc. Prof. Dr. Meltem KIYGI ÇALLI Kadir Has University	....
	Assoc. Prof. Dr. Umut ASAN Istanbul Technical University	....
	Prof. Dr. Oylum KORKUT ALTUNA Istanbul University	....

Date of Submission : 9 May 2024
Date of Defense : 28 June 2024

To my son,

FOREWORD

As I reflect on the journey that has led to the completion of this dissertation, I am filled with a profound sense of gratitude and pride. The process of researching, writing, and refining this work has been both challenging and rewarding, and I am aware that I did not undertake it all by myself.

First and foremost, I am deeply grateful to my supervisor, Prof. Dr. Şebnem BURNAZ, whose guidance, support, and unwavering belief in my abilities have been invaluable throughout this entire process. She has not only served as my academic supervisor, but also as a mentor, offering expertise, encouragement, and constructive feedback that pushed me to strive for excellence and find strength to overcome this process and complete my thesis.

Additionally, I am thankful to Prof. Dr. Tolga KAYA for opening the door to a new world of research for me with insightful advice on the subject and for providing thoughtful suggestions that significantly contributed to the rigor and quality of this work. I would also like to express my sincere appreciation to all the members of my dissertation committee for sharing their insightful comments to support this work.

I must also acknowledge the support of my family, my husband and my son whose love, encouragement, and patience sustained me during moments of doubt and uncertainty. I am profoundly grateful to have them.

Finally, I know that its not common to thank oneself, yet I have made a solemn promise: I extend heartfelt gratitude not only to those who supported me but also to myself for my own perseverance and determination. I pledge to continue along this path knowing that each step forward is a tribute to the journey thus far.

9 May 2024 — Meltem SANİSOĞLU
(Engineer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

B2B	: Business to Business
B2C	: Business to Consumer
CATE	: Conditional Average Treatment Effect
CV	: Cross validation
CVT	: Class Variable Transformation
ED	: Euclidean distance
FN	: False negative
FP	: False positive
KL	: Kullback-Leibler distance
MMS	: Multimedia Messaging Service
ROAS	: Return on Advertising Spend
ROC	: Receiving Operating Characteristic
ROI	: Return on Investment
SMS	: Short Message Service
TN	: True negative
TP	: True positive
R	: R programming language
VIF	: Variance inflation factor

SYMBOLS

$\boldsymbol{B}$	: Bootstrap samples
$\boldsymbol{C}$	: Control
$\boldsymbol{D}$	: Divergence measure
$\boldsymbol{K}$	: Divided parts of data
$\boldsymbol{M^U}$	: Expected uplift
$\boldsymbol{N}$	: Number of observations
$\boldsymbol{P^C}$	: Probabilities of control group
$\boldsymbol{P^T}$	: Probabilities of treatment group
RNR	: Response-Non-Response ratio
RR	: Response rate
$\boldsymbol{T}$	: Treatment
$\boldsymbol{Y}$	: Response variable
$\boldsymbol{X}$	: Input variables
$\boldsymbol{Z}$	: Transformed response variable

LIST OF TABLES

Page

LIST OF FIGURES

Page

MARKETING CAMPAIGN MANAGEMENT USING MACHINE LEARNING TECHNIQUES: AN UPLIFT MODELING APPROACH

SUMMARY

In order to engage customers and increase sales, businesses in today's dynamic business environment devote a large amount of resources to a variety of marketing methods. Businesses utilize a variety of strategies to attract customers' attention and encourage them to make a purchase. However even with a wide range of marketing strategies and channels, a fundamental issue still remains: how can businesses effectively evaluate how their marketing campaigns influence consumer behavior?

By focusing on the significance of uplift modeling in determining the true impact of marketing initiatives, this research directly addresses this issue. By identifying the incremental impact of marketing initiatives, uplift modeling provides a more sophisticated approach than common predictive analytics which only forecasts customer behavior. The purpose of this research is to explore the limitations of conventional predictive analytics in marketing, investigate the application of prescriptive analytics specifically uplift modeling and develop a framework for implementing uplift modeling in business-to-business (B2B) marketing instances by examining real-world data from the Turkish telecom industry.

In this study, three uplift modeling methodologies (two model approach, class variable transformation and modeling uplift directly) performed on a real-world B2B marketing campaign dataset and shown that the marketing campaign can be optimized by predicting the incremental impact more precisely with uplift models than the conventional predictive models. The results revealed how uplift modeling, which enables for the targeting of customers whose behavior is most likely to be positively influenced by marketing efforts, is helpful in improving resource allocation. Out of all the uplift models, the model that used the class variable transformation approach was able to capture 46% of uplift while targeting only the half of the campaign audience. This finding confirms the earlier research in the uplift modeling literature and shows that uplift models are successful in forecasting the truly responsive customers for a direct marketing campaign.

It has been demonstrated that conventional response models are inadequate to distinguish which customers are positively impacted by a marketing treatment whereas uplift models successfully identify which customers will make a purchase due to being influenced by the marketing treatment. It is also shown that in a direct marketing campaign, focusing on a larger target audience may not always yield the greatest or most effective outcomes. Instead, different uplift modeling strategies combined with machine learning algorithms can yield higher uplifts.This study makes significant contribution as being the first to introduce uplift modeling in Turkish literature and being one of the few studies to apply uplift modeling in B2B context in the world-wide academic literature that predominantly focused on researches in B2C. Further, it provides valuable managerial insights for marketers to gain deeper customer insights and foster stronger relationships with customers by leveraging uplift modeling.

MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PAZARLAMA KAMPANYASI YÖNETİMİ: ARTIMLI MODELLEME YAKLAŞIMI

ÖZET

Günümüzün dinamik iş dünyasında şirketler, müşterilerin ilgisini çekmek ve satışları artırmak için çeşitli pazarlama stratejilerine önemli kaynaklar yatırmaktadır. Basılı medya ve satış aramaları gibi geleneksel yöntemlerden çevrimiçi reklamlar ve e-posta pazarlaması gibi dijital yaklaşımlara kadar işletmeler, tüketicilerin dikkatini çekmek ve satın alma davranışını teşvik etmek için çeşitli taktikler kullanmaktadır. Ancak pazarlama kanallarının ve stratejilerinin çoğaldığı bir ortamda pazarlama kampanyalarının gerçek etkisinin tüketici davranışındaki doğal değişikliklerden ayırt edilmesinde büyük bir zorluk ortaya çıkmaktadır. Bu çalışma, pazarlama faaliyetlerinin artan etkilerini ayırt etmede öngörücü analitiğin, özellikle de artımlı modellemenin rolüne odaklanarak bu zorluğu ele almayı amaçlamaktadır. Artımlı modelleme, yalnızca müşteri davranışını tahmin etmekle kalmayıp aynı zamanda pazarlama aksiyonlarının gerçek etkisine de ışık tutarak geleneksel tahmine dayalı analitiğe göre bir avantaj sunmaktadır. Bu yaklaşım, verimli kaynak tahsisi ve yatırım getirisinin artırılması için kilit öneme sahiptir. Tüketici temelli (B2C) bağlamlarda artımlı modelleme üzerine yapılan kapsamlı araştırmalara rağmen, işletmeden işletmeye (B2B) pazarlamadaki uygulaması kapsamlı bir şekilde araştırılmamıştır. Bu çalışma, artımlı modellemeyi B2B pazarlamaya uygulayarak ve gerçek kampanya verilerini kullanarak çeşitli teknikleri analiz ederek bu boşluğu doldurmayı amaçlamaktadır.

Geleneksel tahmine dayalı kestirimsel analitiğin, müşteri davranışlarını tahmin etmede ve pazarlama kampanyalarını buna göre tasarlamak için karar vericilere gelecekteki olaylar hakkında tavsiyelerde bulunmada yararlı olduğu bilinmektedir. Ancak bu analizler, “Ne olacak?” sorusu üzerinden tahmin sağladıkları için pazarlamanın müşterilerin davranışları üzerindeki etkisini ölçmede yetersiz kalmaktadır. Veri bilimi çalışmalarının çoğu iş kararlarını iyileştirmeyi amaçlasa da, iş değerini optimize etmek yerine bir tahmin problemini çözmektedir. Alternatif olarak, öngörücü analitik, hedef kitle üzerindeki etkilerini en üst düzeye çıkarmak ve yatırım getirisini artırmak için pazarlama eylemlerinin etkisini belirlemek için kullanışlıdır. Bir öngörücü analitik yöntemi olan artımlı modelleme, hangi müşterilerin hangi aksiyonu alacağını belirlemek için hem uygulama hem de kontrol gruplarını modelleyerek pazarlama etkisinin müşterilerin gelecekteki davranışları üzerindeki artan etkisini tahmin eden yeni bir tekniktir. Hangi müşterilerin ilgili pazarlama kampanyasına gerçekten yanıt verdiğini belirleyerek, işletmelerin kaynaklarını daha verimli bir şekilde tahsis edebilmelerini ve bu sayede daha iyi yatırım getirisi sağlamalarını amaçlamaktadır. Bu çalışmada, Bu araştırmanın amacı, pazarlama kampanyalarının tüketici davranışları üzerindeki etkisini ölçmek için daha doğru bir model geliştirmek ve pazarlama uygulamaları yoluyla müşteriler üzerinde etkili olan eylemleri ayırt etmeye odaklanmaktır. Bunu başarmak için, bu çalışmada pazarlamadaki geleneksel tahmine dayalı analitiğin sınırlamaları ortaya konmakta, pazarlama faaliyetlerinin artan

etkilerinin belirlenmesinde öngörücü analitiğin, özellikle de artımlı modellemenin uygulanması araştırılmakta, artımlı modellemenin hangi müşterilerin eylemlerinin gerçekten pazarlama kampanyalarının bir sonucu olduğunu tespit etmek için uygulama ve kontrol gruplarının davranışlarını karşılaştırılmakta ve davranışları belirli pazarlama faaliyetlerinden etkilenme olasılığı en yüksek olan müşterileri hedefleyerek kaynak tahsisini optimize etmeyi ve yatırım getirisini artırmayı amaçlayan gerçek pazarlama kampanyası senaryosunda artımlı modellemenin uygulanması için bir çerçeve geliştirilmektedir. Bu yaklaşım ile, işletmelere pazarlama stratejilerinin etkinliğini değerlendirmek ve geliştirmek için daha hassas bir araç sunarak kestirimsel ve öngörücü analitik arasındaki boşluğun doldurulması amaçlanmaktadır.

Pazarlamada tüketiciler ile iletişim, fiyat, yer, ürün ve promosyon gibi temel unsurlardan oluşan pazarlama karmasında hayati bir rol oynar. Bunlar arasında tutundurma, reklam, satış, halkla ilişkiler ve satış promosyonu gibi iletişim unsurlarını yaygın olarak paylaşan pazarlama iletişimi ile güçlü bir şekilde ilişkilidir. Tutundurma karması aracılığıyla, bir ürün tüketici zihninde konumlandırılmadan önce farkındalık yaratılır ve böylece ürün veya hizmet ile satış yapan şirket hakkında olumlu tutumlar geliştirilerek başarılı bir satış gerçekleştirilebilir. Şirketlerin kitleleriyle bağlantı kurmak için kullandıkları yaklaşımlardan biri olarak doğrudan pazarlama, hedeflenen potansiyel müşteriler veya müşterilerle kişiselleştirilmiş, etkileşimli iletişimi kapsamaktadır. Doğrudan pazarlamanın nihai amacı, kampanya alıcılarının davranış değişikliğini olumlu yönde artırarak pazarlama faaliyetlerinin performansını optimize etmek, böylece tüm spesifik pazarlama faaliyetlerinden elde edilen geri dönüşümü ve satışı artırırken, kampanya alıcılarının davranışları üzerindeki olumsuz etkilerini azaltmaktır.

Bir öngörücü analitik yöntemi olarak artımlı modelleme, birbirini dışlayan iki veya daha fazla ayrı sınıfın koşullu ortalama eylemini, ilgili uygulama grubundan elde edilen sonucu kontrol grubundan elde edilen sonuçtan çıkararak modellemeye çalışır. Bu sayede, artımlı modelleme işletmelerin pazarlama kampanyası iletişiminden hangi kitlenin en çok pozitif etkileneceğini ayırt ederek örneklemin daha az bir kısmını hedefleyebilmelerini ve pazarlama giderlerinden tasarruf sağlarken geliri korumalarını sağlayabilmektedir. Artımlı modelleme, müşteri kitlesini dört gruba ayırarak “ikna edilebilirler”, “kayıp vakalar”, “kesin olanlar” ve “rahatsız edilmemesi gerekenler” olarak sınıflandırmakta ve buna göre işletmelerin “ikna edilebilirler” grubunu hedefleyerek pazarlama kampanyalarını optimize etmelerine olanak sağlamayı amaçlamaktadır.

Bu çalışmada üç artımlı modelleme yöntemi (iki model yaklaşımı, sınıf değişkeni dönüşümü ve artımın doğrudan modellenmesi) gerçek bir B2B pazarlama kampanyası verisi üzerinde uygulanmıştır. İki model yaklaşımı, artımlı modellemede en basit yaklaşım olarak uygulama ve control grupları ile eğitilen iki ayrı model oluşturmakta ve uygulama grubu verilerini içeren model tarafından tahmin edilen sınıf olasılıklarından kontrol grubu verilerini içeren model tarafından tahmin edilen sınıf olasılıklarının çıkartılarak artım hesaplanır. İki model yaklaşımı herhangi bir sınıflandırma problemine uygulanabilir ve doğrudan pazarlama kampanyalarının çoğunda yükselişi tahmin etmek için uygulanması kolaydır. Bununla birlikte, literatürde iki model yaklaşımının artımı tahmin etmek için oldukça kusurlu bir yöntem olduğu tartışılmaktadır. Hiçbir model mükemmel tahmin gerçekleştiremediğinden, iki iyi modelden elde edilen sonuçların birbirinden çıkartılmasıyla iyi bir model elde edileceğinin garantisi bulunmamaktadır. İkinci yöntem olarak sınıf değişkeni dönüşümü yöntemi, satın alma davranışı ve uygulama ya da kontrol grubu aidiyetine göre sınıf değişkenini dönüştürerek tek bir model kurmaktadır. İki modelli yaklaşımdan farklı olarak, sınıf değişkeni dönüşümü yalnızca bir tedavi için hedeflenecek en uygun grup olan "ikna edilebilirler" grubuna odaklanmamaktadır. Bunun yerine, "kayıp vakalar" ve "rahatsız edilmemesi gerekenler"i belirleyerek, pazarlama kampanyası ile bu müşteri gruplarına temas riskini ortadan kaldırmaktadır. Üçüncü yöntem olarak artımın doğrudan modellenmesi, verileri alt popülasyonlara ayırmak ve sınıflandırmak için ağaç tabanlı algoritmalar uygular ve bu da onları uygulama ve kontrol grubu arasındaki farklılıkları modellemek için kullanışlı hale getirir. Karar ağaçlarındaki bölme kriterlerinin özü, doğrudan artımın modellenmesi için uygun olmalarını sağlar. Literatürde, artımı en iyi şekilde tahmin etmek için bölme kriterine uyum sağlamak üzere önerilen çok sayıda yöntem vardır. Ancak, hangi yöntemin en etkili yöntem olduğu konusunda bir fikir birliği bulunmamaktadır. Bu çalışmada, doğrudan artım modellenmesi, random forest ve karar ağacı öğrenicileri kullanılarak gerçekleştirilmiş ve Guelman (2014) tarafından geliştirilen özel R paketi (*uplift*) kullanılmıştır. Bölme kriteri için, Kullback-Leibler ve kareli Öklid uzaklığı kullanılmış ve sonuçları karşılaştırılmıştır.

Veri bilimine dayalı modellerin etkinliğini değerlendirmek için literatürde çeşitli yöntemler kullanılmaktadır. Bu alanda özellikle sınıflandırma modellerini değerlendirmek için Roc eğrisi (işlem karakteristik eğrisi) yaygın olarak kullanılan yararlı bir yöntemdir. Ancak artım modellemesinde, bir birey hem uygulama hem de kontrol grubuna ait olamayacağından, roc eğrileri performansı değerlendirmek için yeterli değildir. Bu nedenle, artım modellerindeki bu ikilemi çözmek ve belirli uygulamaların gerçek etkisini ölçmek için Qini eğrileri kullanılmaktadır. Qini eğrisi, grup başına hesaplanan artan satın alma oranını göstermektedir. Segment grubu başına satın alma artışını hesaplamak için, uygulama ve kontrol grubunun kümülatif satın alma sayısı çıkarılır. Qini eğrisinde dikey eksen kümülatif artan satın alma oranını gösterirken, yatay eksen ise uygulama yapılan grup büyüklüğünü göstermektedir. Bu şekilde oluşturulan eğri, pazarlama uygulamasının potansiyel olumsuz etkisinin görselleştirilmesini sağlamaktadır. Nüfusun bir alt grubuna yapılan uygulamanın, nüfusun tamamına yapılmasından daha yüksek bir satın alma oranı gösteren bir qini eğrisi, iyi bir artım modeli göstergesidir. Dolayısıyla, başarılı bir artım modelinde Qini eğrisi tüm nüfus yerine uygulama grubunun daha küçük bir kısmı hedeflendiğinde daha fazla iyileştirme ve artış sağlamalıdır. Bu çalışmada modellerin performansını değerlendirmek için qini katsayısı ve qini eğrileri kullanılmıştır.

Bu çalışmada kullanılan veri seti, bir mobil telekomünikasyon şirketinden anonim olarak toplanan ve Aralık 2020-Ocak 2021 dönemini kapsayan 21.439 B2B müşterinin demografik ve davranışsal özelliklerine dayanmaktadır. Aylık tarife planına ek olarak satın alınabilecek haftalık bir internet paketi ürünü sunulan indirimli bir çapraz satış kampanyasının verilerini içermektedir. Kampanya, müşteriler tarafından okunduğu varsayılan çevrimiçi anlık bildirim ile müşterilere iletilmektedir. Dolayısıyla, uygulama, kampanya anlık bildirimini almak olarak tanımlanmıştır. Bir müşteri anlık bildirim aldığında uygulama grubunda olduğu kabul edilmiş, bildirimin gönderildiği andan itibaren bir haftalık dönem içerisinde satın alma gerçekleştirdiyse bu uygulama sonucu satın alma gerçekleştirdiği varsayılmıştır. Uygulama ve kontrol grubundaki toplam satın alma oranları sırasıyla %6.3 ve %3'tür ve kontrol grubu büyüklüğü, uygulama grubu büyüklüğünden on kat fazla olduğu için artım modellemesi için uygundur (Radcliffe and Surry, 1999). Veri seti gerçek bir pazarlama kampanyasından elde edildiği için modellemeye uygun hale getirmek adına çeşitli veri önişleme adımları gerçekleştirilmiştir. Eksik ve aykırı değerler veri setinin çarpıklığına göre uygun imputasyon yöntemleri ile doldurulmuş veya veri setinden çıkartılmıştır. Veri setinde yer alan satın alan ve almayanların dengesiz dağılımı ise literatürde kullanılan çeşitli örnekleme yöntemleri değerlendirilerek doğruluk ve kesinlik açısından en iyi performansı gösteren "aşırı örnekleme" tekniği ile dengelenmiştir. Aynı zamanda ısı haritası ve varyans artış faktörü yöntemleri ile modele dahil edilmesi gereken değişkenler belirlenmiş ve çoklu bağlantı önlenmiştir. Veri ön işleme adımları sonrası artım modellemesi yöntemleri olan iki model yaklaşımı, sınıf değişkeni dönüşümü ve doğrudan artım modellemesi uygulanmıştır.

İki model yaklaşımında hem uygulama hem de kontrol grubu için oluşturulan iki model tüm gözlemlere uygulanarak iki satın alma tahmini elde edilmiştir. Artımın tahmin edilmesi için kontrol grubu model sonuçları, uygulama grubu model sonuçlarından çıkartılmış ve elde edilen artım modelinin sonuçları qini eğrisi ile değerlendirilmiştir. Sonuç olarak bu yöntem ile hesaplanan qini değeri -0.0065 olarak elde edilmiştir. Sonuçlar, literatürdeki diğer çalışma sonuçları ile uyumlu olarak iki modelin birbirinden çıkarılmasının bir artış bulmakta başarısız olduğunu göstermektedir. Negatif qini değeri, belirli bir uygulama yapmanın hiçbir şey yapmamaktan daha kötü sonuçları olduğunu göstermektedir (Guelman, 2014).Qini eğrisi, hedeflenen nüfus oranı arttıkça kümülatif artan kazanımların da arttığını göstermektedir. Ancak, nüfusun neredeyse tamamı hedeflendiğinde maksimum artışa ulaşılmaktadır.

Sınıf değişkeni dönüşümü (CVT) yönteminde, bağımlı değişken uygulama grubu aidiyetine göre bir z değişkenine dönüştürülmüştür. Eğitilen model tüm test setine uygulanarak, satın alma oranları tahmin edilmiştir. Sonuçlara göre qini değeri 0.2094 olarak hesaplanmakta ve çizilen qini eğrisinde, pazarlama uygulamasının yapıldığı örneklemin büyüklüğü arttıkça satın alma oranı artmakta, daha sonra hafifçe düzleşmekte ve son segmentlerdeki olumsuz etkiyi yakalayarak azalmaktadır. Bu durum, artım modellemesinde tanımlanan "rahatsız edilmemesi gereken" müşterilerin olumsuz etkisinin model ile tespit edildiğini göstermektedir. CVT yönteminde, uygulama yapılması gereken optimum grup büyüklüğü, popülasyonun yarısı olarak ortaya çıkmaktadır. Bu da pazarlama kampanyası ile hedeflenecek kitlenin yarısının hedeflenmesinin, tüm grubun hedeflenmesiyle elde edilebilecek artıştan daha fazla sonuç vereceğini göstermektedir.

Doğrudan artım modellemesinde kullanılan iki bölme kriteri (KL ve ED) için qini sonuçları sırasıyla 0.0913 ve 0.1045 olarak hesaplanmaktadır. Qini değerinin pozitif işaretli olması ve qini eğrisinin hedeflenen nüfus oranı arttıkça kümülatif artan kazançların da arttığını göstermesi dikkat çekicidir. Bununla birlikte, CVT yaklaşımında da ortaya çıktığı gibi, ağaç temelli doğrudan artım modelleri "rahatsız edilmemesi gereken" müşteri grubunu yakalamayı başararak grubun son kesimlerinde artan kazancın azaldığını göstermektedir. Hedef kitlenin %70'ine pazarlama uygulaması gerçekleştirilerek optimum satın alma oranına erişileceği qini eğrisinde gösterilmektedir.

Tüm durumlar için en iyi sonucu veren tek bir artımlı modelleme yöntemi yoktur (Röbler vd., 2021). Bu çalışmada verilere üç artımlı modelleme tekniği uygulanmış ve bir B2B pazarlama kampanyasının, geleneksel tahmin modellerine kıyasla artımlı modeller ile artan etkiyi daha kesin bir şekilde tahmin ederek optimize edilebileceği gösterilmiştir. Bulgular, davranışları pazarlama uygulamalarından olumlu etkilenme olasılığı en yüksek olan müşterilerin hedeflenmesini sağlayarak kaynak tahsisini optimize etmede artımlı modellemenin etkinliğini ortaya koymuştur. Artım modelleri arasında, sınıf değişkeni dönüştürme yöntemine sahip model, kampanya kitlesinin yalnızca yarısını hedeflerken artımın %46'sını yakalamayı başarmıştır. Artımlı modelleme literatüründeki önceki çalışmalarla uyumlu olan bu sonuç, artımlı modellerin bir doğrudan pazarlama kampanyası için gerçekten duyarlı müşterileri tahmin etme konusundaki başarısını göstermektedir. Bu çalışma ile Türkçe literatüre artımlı modelleme ilk kez kazandırılmış ve dünya literatüründe B2B bağlamında artımlı modelleme uygulayan az sayıda çalışmadan biri olarak önemli bir katkı sağlamıştır. Geleneksel tahmin modellerinin hangi müşterilerin bir pazarlama uygulamasından olumlu etkilendiğini ayırt etmekte yetersiz kaldığı, buna karşılık artımlı modellemenin hangi müşterilerin pazarlama uygulamasından etkilenerek satın alma yapacağını başarılı bir şekilde belirlediği gösterilmiştir. Bir doğrudan pazarlama kampanyasında daha fazla müşteriyi hedeflemenin her zaman en iyi veya en verimli sonuçları vermediğini ve makine öğrenimi algoritmalarıyla uygulanan çeşitli artımlı modelleme teknikleriyle daha yüksek artışlar elde edilebileceğini göstermek de önemlidir. Dolayısıyla, bu çalışma, doğrudan pazarlama çabalarında gerçekten duyarlı müşterileri hedeflemek için öngörücü analitikte makine öğrenimi uygulamalarını kullanarak pazarlama kampanyalarını iyileştirerek alandaki ilgili çalışmaları genişletmektedir. Bu araştırmadan elde edilen içgörüler, alandaki pazarlama uygulayıcıları için önemli yönetimsel çıkarımlar sunmaktadır. Mevcut rekabet ortamında pazarlama analitiği, işletmelerin büyük hacimli verileri anlamasını ve pazarlama stratejisi için bunlardan yararlanmasını sağlayarak veri odaklı karar vermenin önemli bir parçası haline gelmiştir. Pazarlama yöneticileri, olumlu yanıt vermesi muhtemel müşteriler ile pazarlama faaliyetinden bağımsız olarak satın alma yapacak müşterileri ayırt etmek için artımlı modellemeyi benimseyerek karar vermeyi kolaylaştırabilir ve kaynakların pazarlama uygulamalarından etkilenme olasılığı en yüksek olan bireylere daha iyi tahsis edilmesini sağlayarak kampanya sonuçlarının optimize edebilir. Aynı zamanda geleneksel hedefleme yöntemlerinin gözden kaçırmış olabileceği müşterilerle etkileşim kurarak daha güçlü müşteri ilişkilerini teşvik edebilir ve memnuniyet ve sadakatin artmasını sağlayabilir. Son olarak, artımlı modelleme işletmelerin pazarlama stratejilerinin etkinliğini değerlendirmek ve geliştirmek için işletmelere daha hassas bir araç sunmaktadır. Bu çalışma ile uygulayıcılara, pazarlama analitiğinde makine öğrenimi uygulamalarının anlaşılmasını artırmak için öngörücü analitik ve artımlı modellemeden yararlanmaya yönelik eğitim ve geliştirme girişimlerine yatırım yapmaları tavsiye edilmektedir. Bu çalışmanın sınırlamaları ve gelecekteki araştırmalar için öneriler de çalışma sonucunda sunulmaktadır.

1. INTRODUCTION

In today's business landscape, companies extensively invest in a variety of marketing strategies, ranging from traditional methods like print ads and telecalls to digital approaches such as online ads and email marketing. These efforts aim to boost customer engagement and sales, a critical aspect of business growth (Röbler et al., 2021; Huang and Tsui, 2016). However, a major challenge emerges in distinguishing the true impact of these marketing campaigns from the natural changes in consumer behavior. This study seeks to address this challenge by focusing on the role of prescriptive analytics, particularly uplift modeling, in discerning the incremental effects of marketing actions. Uplift modeling offers an advantage over traditional predictive analytics by not just forecasting customer behavior but also by shedding light on the actual influence of marketing efforts (Cao et al., 2017; Olaya et al., 2020). This approach is key to efficient resource allocation and enhancing the return on investment (ROI). Despite the extensive research on uplift modeling in consumer-based (B2C) contexts, its application in business-to-business (B2B) marketing has not been as thoroughly explored. This study aims to bridge this gap by applying uplift modeling in B2B marketing and analyzing various techniques using real-world data. The importance of this research lies in its potential to improve B2B marketing campaign performance, providing businesses with strategic advantages in how they allocate resources and optimize campaigns.

This study aims to evaluate the effectiveness of uplift modeling in B2B marketing. It will examine the limits of conventional predictive analytics in marketing and analyze how uplift modeling can be applied effectively in B2B contexts. The study will also develop a practical framework for using uplift modeling in real-world marketing scenarios.

The structure of the thesis is as follows: The introduction sets the stage by outlining the background, significance, and aims of the study. This is followed by a section on the theoretical background and related work, which discusses marketing communications, B2B interactions, and a review of relevant literature. The methodology section describes the uplift modeling techniques, along with classification, prediction methods, and evaluation techniques used in the study. The data section provides details on the dataset from a Turkish telecommunications sector, variables, and preprocessing steps. The findings section presents the results of the uplift modeling methods and a comparative analysis of the uplift models. Finally, the conclusion offers a comprehensive assessment of the findings, discussing both theoretical and practical implications, limitations of the study, and potential directions for future research.

1.1 Background

Companies invest a considerable number of resources in planning and executing marketing campaigns to engage with customers and increase their sales. These marketing campaigns may include traditional actions such as print ads, telecalls or door-to-door marketing and digital marketing actions as online ads, email marketing, push notifications, text messages or offers in e-commerce (Röbler et al., 2021; Huang and Tsui, 2016). This marketing approach that depends on the level of customer response confounds the impact of a marketing action with other natural shifts in consumer behavior. In particular it disregards the possibility that some customers may show positive response even if they were not targeted, whereas some customers may show negative response even though they were targeted with the marketing actions (Hu, 2023). Determining if these marketing efforts are successful in persuading customers to take specific actions is challenging since companies need to ensure that their marketing campaigns target the right customers who are likely to respond to the campaign and take the desired action. Otherwise, the marketing campaign can end up being a waste of resources in terms of time, effort and money. It is crucial for businesses to target relevant customers who are likely to buy due to the campaign and not because they already had the intention to buy (Shimizu et al., 2021; Siegel, 2011).

Conventional predictive analytics has been proven useful in predicting customer behavior and advising decision-makers on future events to design marketing campaigns accordingly. By depending on forecasting methods, predictive analysis aid marketing professionals in making predictions as opposed to managing previous marketing efforts (Paetz et al., 2022). However, these analyses fall short in measuring the impact of marketing influence on customers' behavior, since they provide prediction over “What will happen?” Even though, most data science studies aim to improve business decisions, they solve a prediction problem instead of optimizing business value (Shteingart et al., 2022). Alternatively, prescriptive analytics is useful to determine the impact of marketing actions to maximize their effect on the target audience and improve the return on investment. Among prescriptive analytics, uplift modeling is an emerging technique that predicts the incremental effect of the marketing influence on customers’ future behavior by modeling both the treatment and control groups to determine which customers would have taken the action anyway (Cao et al., 2017; Olaya et al., 2020). By identifying which customers are truly responsive to the related marketing campaign, businesses can allocate their resources more efficiently, resulting in better ROI.

1.2 Problem Definition

The effectiveness of marketing campaigns often gets confounded with natural shifts in consumer behavior. A critical challenge in marketing analytics is distinguishing between customers who respond positively to marketing efforts and those whose actions are independent of such campaigns. This distinction is vital because some customers might respond positively without being targeted, while others might react negatively despite being the focus of marketing efforts. The inability to accurately measure the true impact of marketing actions on customer behavior leads to inefficient resource allocation for companies potentially wasting time, effort, and money. Understanding the real influence of marketing campaigns on consumer behavior is crucial for optimizing return on investment. If businesses fail to identify and target the customers most likely to be influenced by their campaigns, they risk directing resources toward individuals who would have made purchases regardless, or worse, alienating potential customers. This misallocation not only reduces the effectiveness of marketing efforts but also impacts the overall profitability and market positioning of the company.

The aim of this research is to develop a more accurate model for measuring the impact of marketing campaigns on consumer behavior with the help of prescriptive analytics, focusing on distinguishing the influenced actions on customers via marketing treatment. To achieve it, this study will:

1. Explore the limitations of conventional predictive analytics in marketing, which focuses on forecasting customer behavior rather than measuring the influence of marketing efforts.
2. Investigate the application of prescriptive analytics, specifically uplift modeling, in identifying the incremental effects of marketing actions. Uplift modeling contrasts the behavior of treatment and control groups to ascertain which customers' actions are truly a result of marketing campaigns.
3. Develop a framework for implementing uplift modeling in real-world marketing scenario, aiming to optimize resource allocation and enhance ROI by targeting customers whose behavior is most likely to be influenced by specific marketing actions.

This approach seeks to bridge the gap between predictive and prescriptive analytics, offering businesses a more precise tool for assessing and enhancing the effectiveness of their marketing strategies.

1.3 Research Objective

This study mainly aims to evaluate the effectiveness of uplift modeling as a prescriptive analytics method in enhancing the performance of B2B marketing campaigns. Research objectives include an investigation of uplift modeling in B2B contexts, addressing the notable gap in literature predominantly centered around B2C applications. It is planned to compare various uplift modeling techniques using a real-world B2B campaign data, aiming to identify most effective methods specific to B2B marketing scenarios. Additionally, this study intends to demonstrate practical applications of uplift models in optimizing B2B marketing communication, providing actionable insights and managerial recommendations. Finally, this study not only seeks to advance the practical understanding of uplift modeling in B2B campaigns but also aims to make a significant contribution to the academic area in marketing analytics and opening doors for further research in this evolving field.

1.4 Significance

While uplift modeling methods combined with machine learning techniques have gained attention in literature in the last decade, almost all studies are in B2C context for optimizing direct marketing campaigns (Hansotia and Rukstales, 2002; Rzepakowski and Jaroszewicz, 2012; Cao et al. 2017; Diemert et al.,2018; Gubela et al., 2019;Shimizu et al., 2019; Goldenberg et al., 2020; Olaya et al., 2020; Gubela and Lessmann, 2021; Baier and Stöcker, 2022; Nyberg and Klami, 2022; Hu, 2023; Proença and Moraes, 2023) and retention campaigns for churn management (Radcliffe and Simpson, 2007; Röbler et al., 2021; Devriendt et al., 2021). However, uplift modeling with machine learning techniques can also provide significant improvements in B2B marketing campaign performance by resulting insights into which marketing channels are most effective and which customer segments are most responsive to the specific campaigns. Hence, this approach can offer a strategic advantage in refining future campaigns and optimizing resources allocation for maximum impact.

This study significantly aims to contribute to the academic literature which includes previous research on uplift modeling in B2C but overlooked applications in B2B context with the purpose of improving B2B marketing campaign performance by determining the most responsive customer segment to target. By demonstrating three different methodologies to model uplift by using a real-world data provided anonymously from a mobile telecom company to explore the usage and feasibility of uplift modeling compared to traditional direct marketing campaign setting, it is aimed to answer the question if uplift modeling can be used to target better and increase the efficiency of a cross-sell direct marketing campaign. This study offers practical managerial implications and encourages both researchers and practitioners to use machine learning to own advantage in B2B marketing to create more targeted, efficient and effective marketing campaigns.

1.5 Outline

This study follows an outline beginning with an introduction that provides the background information and highlighting the significance and purpose of the study. The theoretical background and related work section explain the marketing communications and the role of direct marketing in marketing communications, followed by the B2B interactions and targeting in a direct marketing setting with the literature review. Then the methods for measuring the performance of marketing communications is demonstrated, followed by the concept of causal inference and uplift modeling. The methodology section outlines the uplift modeling techniques and various classification and prediction methods along with evaluation techniques. The data section provides insight into dataset, variables and preprocessing steps for the application in a Turkish GSM sector. The findings section focuses on the outcomes of the uplift modeling methods and demonstrate a comparative model analysis. Finally, the study is concluded with a comprehensive assessment, offering both theoretical and practical implications while also acknowledging its limitations and suggesting avenues for future research.

2. THEORETICAL BACKGROUND AND RELATED WORK

This section is dedicated to providing a comprehensive understanding of several aspects in the marketing. To begin, the concept of communicating with consumers in marketing is explored by including strategies and methods used by companies to connect with consumers effectively. Next, the evolution of direct marketing is covered by tracing its historical development and adaptation to changing technologies and consumer behaviors in order to provide insight on how direct marketing has changed over time. Then, the marketing communications in B2B is examined to discuss how marketing strategies differ when targeting other businesses, considering the unique dynamics and communication channels involved.

In the direct marketing domain, it is crucial to measure the performance of marketing communication. Hence, how companies measure the effectiveness of marketing communication and the various methods used in literature are demonstrated in order to assess whether these marketing actions are effective and efficient. Causal inference concept is explored to understand the cause-and-effect relationships in marketing, enabling better decision-making. Lastly, uplift modeling, the method that provides valuable insights into how marketing campaigns can be tailored to maximize their effectiveness, is introduced. By studying these topics and discussing related previous literature, it is aimed to provide a solid theoretical foundation on direct marketing and grasp the core ideas and principles on enabling more successful marketing strategies.

2.1 Communicating with Consumer in Marketing

Man has been trying to communicate from the beginning of time. At the most fundamental level, communication is the most significant component of interpersonal social interaction. As the time passed and more significantly developments such as primitive printing techniques, it became possible to broaden communications to reach a larger audience. Early printed materials offer illustrations of the creation of a new medium of communication intended to inform a larger audience about the accesibility of goods and services. It was the start of marketing communications era (Yeshin, 1998).

Marketing communication refers to the engagement with customers and potential customers through multiple communication channels, including direct mail, print media like newspapers and magazines, television, radio, billboards, telemarketing and the Internet. It plays a vital role in the marketing mix, which comprises the essential elements of price, place, product and promotion. Among these, promotion is strongly related to marketing communications that commonly shares the elements of communication as advertising, selling, public relations and sales promotion. Through the promotion mix, awareness is first created before a product is positioned in the consumer minds to develop positive attitudes about the product or service and the company that is selling that may result in a successful sale (Kayode, 2014).

In the basic model of communication process, a message is created by the sender or source to elicit a particular response from the recipient. The message is then interpreted based on the recipient's perception and comprehension, and feedback is given to the source in the form of an action. The fundamental model of communication by Shannon and Weaver (1949) is represented in Figure 2.1.

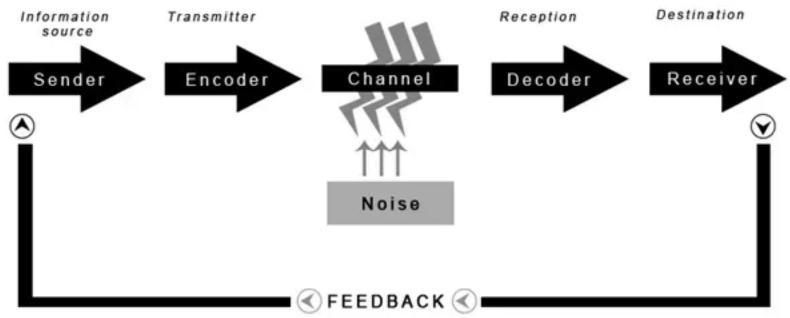

Figure 2.1 : Shannon and Weaver’s model of communication

As it is shown in Fig. 2.1, the process of communication is represented as a linear process and consists of the sender which is the organization or a person that shares information that is encoded into a message form and transmitted by using a channel. The decoding process consists of the interpretation of the information by the receiver which is the individual that the message is intended to reach. Through the transmission, noise factors include barriers which could adversely affect communication between the source and the receiver. At last, as a reaction to the message, feedback to the source could be transmitted (Kayode, 2014). This model which has been influential in the fields of information theory and communication studies, is valuable for understanding the basic elements of communication and the potential breakdowns in the communication process.

There are four fundamental elements of marketing communication mix such as advertising, sales promotion, publicity and personal selling. Advertising constitutes any sort of non-personal, paid presentation and promotion of concepts, products or services by a designated sponsor. To encourage the buying of the products or the services sales promotions are used as incentives. Publicity is the planned commercially important news about the products or the services in a publicly available medium to impact the creation of demand for the product. Lastly, personal selling covers the oral presentation during a discussion with one or more potential customers in an effort to close deals (Kayode, 2014). In summary, these four fundamental elements, advertising, sales promotion, publicity, and personal selling, form the core of the marketing communication mix, each playing a distinct role in promoting products or services and engaging with customers.

Mass marketing and direct marketing are two distinct approaches companies use to connect with their audiences, each with its own set of strategies and objective. Mass marketing is what marketing experts employ to attract the broadest segment of the market by using uniform product offerings, pricing, distribution and promotional methods to target all consumers, irrespective of their demographics. The primary aim in mass marketing strategy is to reach the maximum potential customer base, catering to various consumer needs with products that can be valuable to a diverse audience. This strategy emphasizes selling products at competitive prices to generate a substantial volume of sales and seek to achieve widespread product visibility through a unified offer (Indeed Career Guide, n.d.). However, in direct marketing, companies and organizations aim to create and nurture a direct connection with their customers and target them specifically for certain relevant product and service offers. Companies utilize vast amount of customer and market data to select and target consumers for a specific campaign by contacting directly via mail or other communication channels. Today, an increasing number of companies particularly those in the service industry such as financial services and insurance are adopting this approach as their primary strategy for engaging with their customers (Javaheri, 2007). The differences among mass and direct marketing is summarized in Table 2.1 below:

Table 2.1: Comparison of mass and direct marketing (Roberts and Berger, 1999)

Mass Marketing	Direct Marketing
Using mass media to reach a mass audience	Directly communicating with the prospects or customers
Impersonal communication	Personalized communication
One way communication among advertiser and consumer	Interactive communication
Highly visible promotional programs	Relatively less visible promotional programs
Promotion is determined by the available budget	The size of budget can be determined depending on the success of the promotion
The action that is desired could be postponed or unclear	Specific action such as inquiry or purchase is requested
No sample data or incomplete data for marketing decision making	Comprehensive data for driving marketing actions
Segment level analysis is available	Individual level analysis is available

In summary, mass marketing relies on mass media to reach a broad audience with impersonal, one-way communication, often driven by a predetermined budget. In contrast, direct marketing involves personalized and interactive communication with prospects or customers, allowing for specific actions like inquiries or purchases. It utilizes comprehensive data for marketing decisions, providing individual-level analysis and flexibility in budget allocation based on promotion success.

Direct marketing indicates the marketing activities that are customized and target specific customer groups rather than aiming larger groups with the same activity as it is applied in mass marketing (Bose and Chen, 2009). The ultimate objective of direct marketing is to optimize the marketing activities' performance by boosting the behavioral change of campaign receivers in a positive manner by improving conversion, sales and non-churn from all specific marketing activities while decreasing their negative impact on campaign receivers' behavior such as churn and scattering loss. Direct marketing activities include both traditional marketing actions such as prints, sales calls or door-to-door rendition and digital marketing actions such as online advertisements, e-mail marketing, product offers in e-commerce, cross-sell and up-sell offerings (Röbler et al., 2021; Huang and Tsui, 2016).

Primarily most of the direct marketing campaigns aim to shift the customer behavior to encourage more purchase or the amount of purchase or purchasing products with higher margins (Hansotia and Rukstales, 2002). Companies simply aim for the campaign to produce additional revenue with the best possible return on investment (ROI) that requires calculation of incremental number of purchases or sales through conscientious use of control group (Cao et al., 2017). However, the problem in direct marketing is that it is not known which individual the company should contact for communicating the campaign, in order to achieve a return on direct marketing campaign investment that would exceed the company’s cost for the investment (Radcliffe and Surry, 1999; Hansotia and Rukstales, 2002). Even though there are additional objectives such as brand awareness and generating customer sincerity, most of the direct marketing activities are mainly assessed by the profitability or other kind of return-on-investment calculation. Hence, if the primary focus is to achieve incremental revenue with the direct marketing campaign, the measurement of success is not simple due to the complexity of not knowing that level of sales would have been achieved if the marketing activity in question would not have been initiated. In order to evaluate the incremental impact of a marketing campaign, the most suitable way that is widely recognized is to use a control group and compare the performances of the treated and the untreated control group chosen from the target population (Radcliffe and Surry, 2011). Currently, most of the marketing activities are using traditional response models that attempt to classify customers according to the highest likelihood of purchasing by characterizing who have already bought before, bought in a recent historical period or bought in apparent response to a direct marketing activity such as a targeted email or call from sales person (Radcliffe and Surry, 2011; Cao et al., 2017). In the business-to-consumer centric campaigns, conventional models have been more commonly utilized for direct marketing campaign management. In summary, the primary objective of direct marketing is to stimulate consumer behavior toward increased purchasing, emphasizing higher-margin products. Companies strategically seek these campaigns with the underlying goal of achieving optimal return on investment (ROI), precisely assessing incremental purchases through control group methodologies. However, a critical challenge persists in the identification of target individuals for campaign communication, complicating the realization of a ROI exceeding the campaign's incurred costs.

2.1.1 The evolution of direct marketing

Over the past two decades, there has been a significant transformation in the marketing landscape. A notable shift has occurred in terms of market composition and consumer behavior. This, along with several other factors, has led to a decline in the effectiveness of traditional marketing communications. Consequently, a more favorable environment for direct marketing activities has emerged, viewed positively from both the consumer and company perspectives (Evans et al., 1996). Exploring the changes in direct marketing over time reveals an interesting journey shaped by new strategies, technology, and how people interact with it. Direct marketing, which used to be a simple way to communicate with the customers, has changed a lot to fit into today's complex market. With the development of the technology, traditional marketing practices had also experienced an immense transformation. In order to adapt to the fastly changing environment, companies begun to change from traditional marketing strategies to mutual digital programs (Barnes and Scornavacca, 2004). Instead of using fragmented traditional efforts, most of the existing businesses modified and restructured their marketing tactics by using novel and meaningful digital capabilities. It's not just about reaching people one-on-one anymore; it now involves personal communication, smart analysis, and new technologies. Knowing how direct marketing has evolved is essential for businesses trying to make sense of today's marketing challenges.

Digital marketing, the specific term for describing the marketing of products and services using digital channels, evolved into an umbrella term to define the processes that utilize digital technologies for customer acquisition, constructing customer preferences, brand promotion, retention and growth (Kannan and Li, 2017). Further it is defined as “the activities, institutions and processes facilitated by digital technologies for creating, communicating and delivering value for customers and other stakeholders” by the American Marketing Association. Digital marketing incorporates both direct marketing by acknowledging customers not just by their individual traits but also by their behavior, and interactive marketing by speaking to an individual, collecting and recalling the respond of the individual (Deighton and Kornfeld, 2009). E-mail, web, databases, mobile and digital TV are the application areas for digital marketing that promotes interactive and non-interactive marketing initiatives with the goal of profitably acquiring and retaining customers (Chaffey, 2010).

Digital marketing varies in many aspects compared to the traditional marketing approach. A few significant elements comparison has been shown in Table 2.2 below:

Table 2.2 : Comparison of traditional and digital marketing

Traditional Marketing	Digital Marketing
Involves print, broadcast, postal mail & tele-marketing	Involves e-mail marketing, social media, text messages & online advertisements
No interaction with the audience	Interaction with the audience
Campaigns planned over a long period of time	Campaigns planned over a short period of time
Time-consuming and costly process	Less costly and faster process
Success is determined according to reaching large audiences	Success is determined according to reaching specific target audiences
One campaign stands in place for a long time	Campaigns can be changed and replaced efficiently
Responses can take place during work hours	Responses can take place anytime

As the most recognizable form of marketing, traditional marketing includes print, broadcast, direct mail and telephone where no interaction with the audience occurs. Most of the campaigns are costly in terms of money and time due to long planning duration. Measuring the success of traditional marketing actions is mainly decided on the basis of reaching a large local audience which is also limited due to limited number of customer technology. On the other hand, direct marketing strategies involve email marketing, social media, text messages and other online advertising mediums which enables interaction with the audience by enabling response or feedback to occur anytime. Digital marketing campaigns are fairly less costly and easy to plan and implement over the short period of time. The success of a digital marketing action is measured by the profit and reaching specific number of audience that is targeted (Yasmin et al., 2015).

In today's marketing world, digital platforms such as mobile, online channels, and social media is used by businesses to communicate with their customers. As one of the digital marketing strategies, mobile marketing utilizes mobile devices to reach targeted customers for introducing, promoting and marketing specific products and services. The interaction between businesses and customers has increased along with the use of smartphones and other mobile devices. The main components of today's mobile marketing channels are SMS, MMS, creating a native mobile application, company business websites and emails (Kumar and Mittal, 2020). Companies now access their target customers using mobile marketing channels since consumers of all ages, students and workers, members of civil society and business people, as well as employees of the private sector and government agencies are mobile. Because of this flexible technology, businesses can contact their clients independent of location or time. Smartphones have had the greatest impact on customers' lives by altering the communication and the behavior while serving as their best and irreplacable companion. It is remarkable because smartphones combine all these innovations and expand them by the aspect of independence across space and time or, what is known as ubiquity (Shankar and Balasubramanian, 2009).

As one of the mobile marketing mediums, push notification has been substantially involved in the consumers' lives with the introduction of the smartphones. Push notifications are the messages that are sent to users to their mobile phone home screen, whether or not they are engaged in the application. These notifications are used by companies to drive users back into their mobile app, create engagement and send promotional content. In the perspective of companies, push notifications are the primary function of mobile apps which can deliver brief messages to encourage consumers to make a purchase or informed about the new material without even opening the relevant application (Wohllebe et al., 2021).

A basic illustration of a push notification on a mobile phone home screen is shown in Figure 2.1 below. The left screen in the figure displays a general layout of a push notification on a mobile phone screen; whereas the right screen in the figure illustrates how it seems with fabricated content. Mobile users who receive push notifications, are quickly drawn to them and may make impulsive purchases as a result. These are not only encouraging users to make purchases but also encouraging them to return to using the application and receiving the messages that the marketers aim to deliver (Taylor et al., 2011; Liu et al., 2012). Hence, including push notifications into mobile marketing strategy is greatly beneficial for businesses.

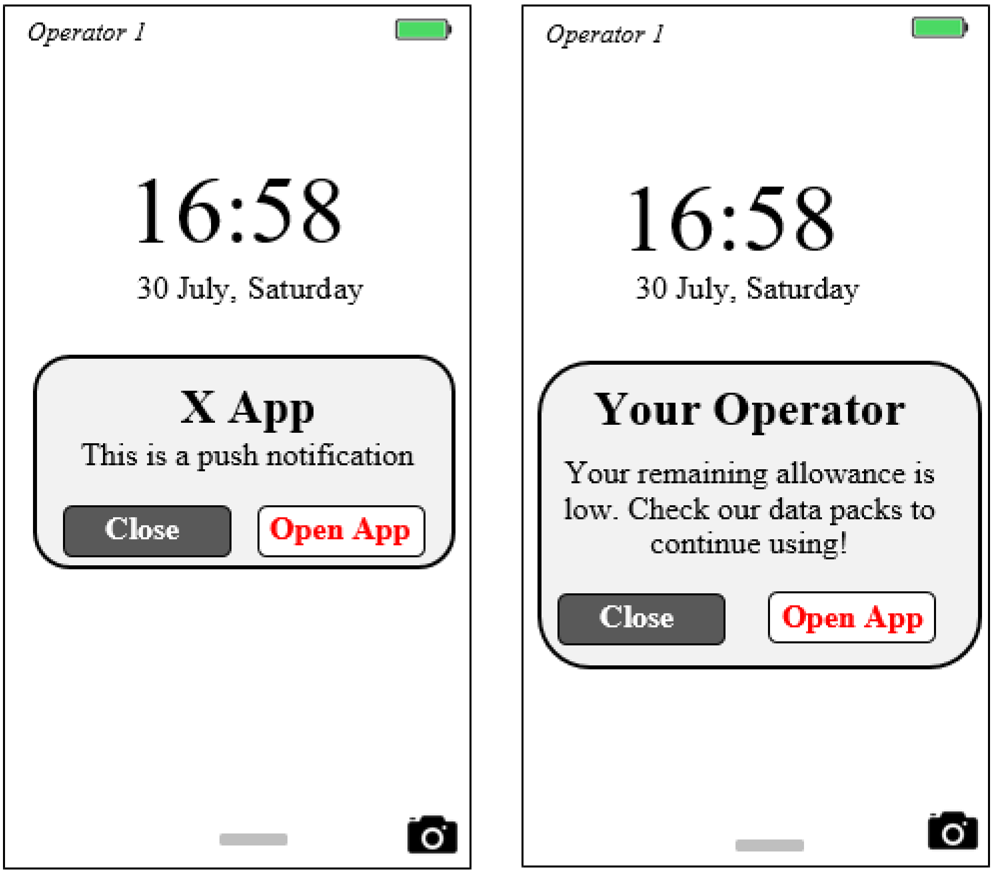

Figure 2.2 : Mobile push notification examples (own illustration)

The perception of mobile push notifications as a direct marketing medium is compared to other direct marketing tools by Lachner et al. (2020). Multiple dimensions from the literature were considered while evaluating mobile push notifications. The most notable elements of mobile push notifications, in the opinion of users, are their capacity to incorporate location-based features and customers' consent to receive them. Further, the widespread availability and the ability to quickly react to time-sensitive content reveal interesting and promising aspects of mobile push notifications. (Lachner et al., 2020). Among the direct marketing strategies, utilizing push notifications to target customers stand out as advantageous with many superior aspects and ease on practice for marketing professionals.

As another medium, social media is an effective strategy for digital marketing when integrated into a company's marketing communication strategy involving multiple platforms such as Facebook, Twitter, Instagram, Youtube and etc. (Medina and Correia, 2012). Social media refers to internet-based technologies that enable online discussions, spanning various platforms like social networking sites, blogs, companysponsored forums, chat rooms, consumer-to-consumer emails, product/service rating sites, internet discussion boards, and multimedia-rich sites. Recently, official company and brand websites have seen a decline in audience, attributed to the growing prevalence of brands engaging in social media marketing (Hollensen and Raman, 2014). This shift is indicative of the evolving landscape, where brands find value in connecting with audiences through diverse social media channels, prompting a reevaluation of traditional company websites. Hence, marketing professionals are exploring avenues to integrate social media into their interactive marketing strategies. The conventional communication approach, dependent on the traditional promotional mix for shaping interactive marketing strategies, needs to evolve. A new paradigm must emerge, encompassing all types of social media as viable tools in the creation and execution of interactive marketing strategies (Hollensen and Raman, 2014). This change emphasizes the need for a wider view, including various social media platforms in interactive marketing strategies, challenging the limits of traditional promotion methods.

Digital marketing is a crucial part of today's direct marketing plans, transforming how businesses engage with their target audiences. It involves using online channels like social media, websites, and mobile devices to connect with consumers. Unlike traditional methods, digital marketing enables precise targeting, personalized messages, and real-time interaction. It includes various strategies such as content marketing, emails, and social media ads. The importance of digital marketing lies in its adaptability to changing consumer behaviors, offering businesses data-driven insights to improve their approaches. It not only boosts brand visibility but also fosters meaningful customer relationships, playing a vital role in overall marketing strategies for businesses in the digital era. For a digital marketing action to be successful, sales improvement which would lead to higher profits and return on investments needed to be achieved. Furthermore, it is strongly suggested that all digital marketing strategies should be built and managed in accordance with the company's culture, structure, mission and code of conduct in order to succeed in realizing the company's vision (Barutcu, 2008). In order to convince targeted customers' to purchase a certain product or service with the direct marketing action, it is essential to contact them with the appropriate message which would increase market value for both the consumer and the company. Modern form of direct marketing has evolved into digital marketing which uses online channels to target, send personalized messages, track responses and engage directly with the consumers. This shift represents a dynamic change in how businesses connect with customers, making digital marketing an essential part of today’s direct marketing strategies.

2.1.2 Direct marketing in B2B

As direct marketing is widely used by businesses to establish direct connection with consumers and drive sales; companies have also initiated direct marketing campaigns to foster business-to-business (B2B) relationships. Adapting to changing dynamics of commerce and communication, B2B direct marketing has undergone a historical evolution. For nearly three decades, companies utilized various direct marketing methods aiming to generate leads and develop lasting relationships with businesses seeking to cultivate loyalty. In the early stages, B2B direct marketing primarily involved direct mail campaigns where companies sent catalogs and promotional materials directly to other businesses, mainly aiming to promote products or services tailored to specific industries. With the developing telecommunications, telemarketing became a prominent B2B direct marketing strategy. Companies used phone calls to reach out to other businesses, offering products, services and solutions directly. However, the efficiency of B2B direct marketing has shown a decline, as direct mail and telemarketing faced challenges such as decreasing response rates and increasing reliance on voicemail. In e-mail programs, it has become a fortunate standard for companies to observe 1% to 3% response rates in the last few years due to the spam e-mail tagging and receiving excessive amount of emails per day (Coe, 2004; Zimmermann and Blythe, 2013). Also in telemarketing, there is a continuing decline in response rates caused by more and more people not answering calls and relying on voicemails for time management. Although the volume of marketing messages has increased greatly, businesspeople are busier today to respond or even hear most marketing messages (Coe, 2004). Hence, the direct marketing efforts in B2B to ensure consistent communication with businesses has become more costly since the overall profitability of B2B direct marketing strategies have been suffered.

The digital revolution brought a significant transformation in B2B direct marketing. Marketing communications in B2B has evolved with the advent of digital marketing tools and social media platforms which facilitate faster and more personalized interactions between customers and suppliers, thereby deepening relationships (Huotari et al., 2015). Many goods and services that were previously difficult for B2B customers to obtain are now widely accessible due to the growth of digital media that makes it easier for B2B buyers to make educated selections. In order to manage customer relationships across both online and offline channels, B2B businesses benefit from establishing a trustworthy online presence through websites, blogs, online business communities or social media platforms (Pandey et al., 2020). The growth of online platforms, business communities and social media further expanded the avenues for B2B direct marketing, enabling businesses to establish and nurture relationships with customers.

In the contemporary landscape, B2B direct marketing is characterized by a multi-channel approach. Businesses leverage a combination of email marketing, content marketing, social media marketing and targeted advertising to connect with other businesses. Recently, digital marketing stands out as offering accurate targeting of potential customers (Pandey et al. 2020). Although digital marketing is included in B2B businesses commercial strategies, this field of study is still developing as a research area. In fact, until recently, the majority of the businesses considered that digital marketing was only effective in B2C, where its advantages, such as its capacity to boost sales, save expenses, generate consumer insight, foster customer relationships, offer value, and improve the brand, have been extensively studied (Kim and Moon, 2021; Setkute and Dibb, 2022). The view is progressively changing, though, as B2B companies such as Cisco and IBM share their digital marketing success stories, and businesses start to understand how digital marketing in B2B improves client trust and information flow (Venkatesh et al., 2019; Pandey et al., 2020). Hence, the historical journey of B2B direct marketing, underscores the continuous adaptation to technological advancements and the importance of a holistic and personalized approach to building and maintaining business relationships.

Although direct marketing approaches are evolving in B2B, the existing body of literature on this subject is still notably limited, indicating a scarcity of comprehensive research in this area. Pandey et al. (2020) studied this emerging issue by concentrating on how much research has been done so far on the use of digital marketing in B2B and what would be the prospective study scope for the future. They reviewed various research articles on digital marketing usage in B2B and identified emerging themes which were confirmed with semi-structured interviews with industry experts. The results revealed that the majority of the research were based in developed countries and the organizations aim to achieve higher sales productivity via digital marketing usage by improving customer relationship, branding and higher return on investment. The study suggested valuable academic and managerial implications and future study directions in the use of digital marketing in B2B organizations. Machine learning technologies to redefine digital marketing investment decisions; developing a roadmap to adopt and leverage social media marketing in a seamless and methodological approach; benefiting open-innovation to manage customer relationship; conducting a study to understand which social networking platform has the greatest influence on B2B sales; exploring analytics tools to effectively collect and analyze the data and more region-specific research to examine customers needs and preferences across different regions are suggested as future study directions (Pandey et al., 2020).

Further, Karjaluoto et al. (2015) examined the role of digital channels in industrial marketing communication to determine the actions that industrial firms have taken to increase the use of digital elements in their marketing communication and the challenges that the digital marketing communication faces. The extensive analysis showed that the B2B marketers aim to get strategic level benefits such as cost-effectiveness and competitive advantage with digital marketing communications. Although a successful digital marketing strategy requires innovation, many organizations utilize essential forms of digital marketing communication as website, e-mail and digital sales materials due to preferring personal communication and organizational complexity of business relationships. The findings support that customers are increasingly participating actively in the communication process despite there is not any significant change in B2B communication. Digital marketing campaigns for general brand awareness may not provide the expected results because the message is not reaching its target audience or because the target audience is receiving the message but does not find the broad content to be relevant (Karjaluoto et al., 2015). Hence, effectively reaching the appropriate audience with tailored content and prioritizing cost-effectiveness are important elements in creating successful digital marketing strategies within a B2B setting.

Many B2B organizations increasingly employ direct marketing activities to enhance customer acquisition and retention. However, a considerable amount of businesses face challenges due to a lack of profound understanding of leading B2B digital marketing strategies, impacting their return on investment (Wertime and Fenwick, 2011). The optimization of potential profits from digital marketing actions poses a common challenge for organizations, therefore making outstanding marketing communication has become a strategic management goal for successful businesses (Cornelissen, 2004). Businesses can leverage machine learning techniques to enhance their marketing communication effectiveness by targeting potential customers to achieve higher conversion rates (Zulaikha et al., 2021). By analyzing vast amount of data, machine learning algorithms can identify patterns and trends, helping businesses understand customer behavior and preferences. This insight enables personalized content recommendations and targeted messaging, improving the relevance of communication.

The academic literature in B2B digital marketing is still developing, with recent studies indicating its early stages. The most promising areas for further research in B2B digital marketing are probably those involving analytics, machine learning and performance assessment. Given the substantial investment in direct marketing campaigns by companies aiming to boost sales and enhance customer relationships, a critical aspect is analyzing the success and effectiveness of these campaigns. Although there are many research that put traditional direct marketing and response modeling techniques into practice and show successful marketing campaigns with higher response rates, the fundamental issue that concerns the organizations is to decide who to target as an audience with the marketing campaign so that the targeted customer will most likely respond due to the campaign. A potential solution to this challenge is found in prescriptive analytics which provide businesses the potential to optimize decision-making and enhance overall marketing performance. However, there are challenges associated with understanding and implementing its methods. The ability to take intelligent actions takes more than only predicting, it also requires a precise understanding of what actions to take and when to take them. (Lepenioti et al., 2020). Despite types of prescriptive analytics such as incremental response models being extensively discussed in the literature on B2C direct marketing, including digital marketing strategies, their application in B2B setting remains notably absent. There is only one identified study proposing a model for retention management in B2B (Caigny et al., 2021) and the literature review reveals a lack of exploration on the prescipritve analytics approach in B2B direct marketing campaign management. More explanation on prescriptive analytics and methods to enhance direct marketing campaign performance is provided in Sections 2.3 and 2.4. These sections elaborate on the principles and strategies associated with models in prescriptive analytics, offering insights into its application for informed decision-making and improved overall campaign effectiveness.

2.2 Measuring Marketing Communications Performance

Evaluating the effectiveness of communication techiques in marketing is crucial for businesses looking to maximize their outreach efforts in the constantly evolving field of marketing. Understanding the effectiveness of marketing communications becomes a complex task that demands a comprehensive exploration of diverse metrics and analytical approaches. This sections explores the complicated field of assessing the effectiveness of marketing communications, providing insight into the various approaches used in the literature to measure the marketing communications performance.

A common way to assess marketing campaign performance is through prediction models which estimate the expected performance with a candidate prediction model applied on unseen data. The standard model evaluation methodology follows:

i. Evaluating the model's performance on an independent test set which is a labeled data but has been separated in the beginning and not used to train the model.

ii. Adjusting the models' scores to align with the overall population distribution, especially if it is anticipated to differ from the sample distribution used in model training and testing.

When strategizing a campaign, the main goal is to target individuals who most likely respond to it. However, there is a limit to the number of individuals that could be targeted by the campaign due to the budget restrictions. Hence, choosing a suitable model for the target segment is crucial and how well it performs with the remaining population has little to no impact. Typically, the success of this model is measured with the amount of responders that are captured within the targeted indiviudals (p et al., 2001).

In literature three fundamental performance measures are utilized for evaluating marketing communication performance. Table 2.3 below represents the list of these measures along with their respective calculation methods, accompanied by the following explanations. Let,

- *A, B* — the total number of responders and non-responders, respectively.
- A_x, B_x — the total number of responders and non-responders in the *x*-th top quartile, respectively.
- *x* (A+B)* or ($A_x + B_x$) — all cases in the top *x*-th quartile.
- *A/(A+B)* — overall response rate

Table 2.3 : List of three performance measures

	Response Rate	Lift	Response Non-Response Ratio
Definition	Percentage of customers reached in a campaign among those approached	The ratio of the response rate to the overall response rate	The ratio of the percentage of all responders to the percentage of all non-responders in the top *x*-th quartile
Calculation	RRj = Aj/(Aj+Bj)	Lift (j) = RRj / (A/(A+B))	*RNRj = (Aj/A) / (Bj/B)*

The response rate (RR), which is also commonly known as conversion rate in literature, represents the percentage of responders the campaign captures among all approached customers. This metric addresses how frequently we expect to encounter a responder during the marketing campaign execution. For instance, in a telemarketing campaign, we might be interested in understanding the likelihood of encountering a responder in a given number of calls. Although this measure is valuable for estimating campaign profitability, it is highly dependent on the overall response rate. It decreases linearly with a decline in the overall response rate in the population. Hence, proper normalization is required to compare models developed on populations with different response rates. Moreover, if the response rate in the future population for a campaign is unknown, obtaining a reliable estimate of the response rate becomes challenging (Rosset et al., 2001). This metric measurement may vary depending on the nature of marketing communication. For instance, measuring click-through rate involves calculating the percentage of people who clicked on a link in the marketing message;

engagement rate reflects the level of interaction and engagement with the marketing campaign encompassing actions as likes, comments and shares on social media or the net promoter score serves as an indicator of the likelihood that customers will recommend the brand to others.

Secondly, the lift is calculated as the ratio between the response rate (RR) and the overall response rate. It provides how much more effective we are choosing the target audience with our model than if we were to choose it at random. For instance, by contacting the top model scored 2% of the population, we could reach 16% of the respondents as opposed to just 2% of the responses when engaging randomly 2% of the population. In this instance, the random model is being improved 8 times, or the lift is at 2% is 8. It directly indicates the improvement a campaign based on a model would offer compared to a randomly selected campaign. As an alternative perspective to the response rate, the lift serves as a valuable measure for campaign evaluation. While mildly influenced by the overall response rate, it exhibits a less pronounced increase as the overall response decreases, making it a common and intuitive criterion for assessment.

Additionally, the response non-response ratio (RNR) is determined by the ratio between the percentage of all responders and the percentage of all non-responders in the top *x*-th quartile. Due to being independent of the overall response rate, RNR is easy to compare for evaluating model performance. However, its lack of intuitive application for campaign planning and difficulty in illustrating campaign effectiveness pose challenges. When comparing two models through their ROC curves, the comparison essentially involves assessing their RNR at all possible cutoff points simultaneously (Rosset et al., 2001). Further explanation on ROC curves is provided in Section 3.3.1

In addition to these metrics, another avenue for evaluating the effectiveness of a marketing campaign is to scrutinize its cost-effectiveness. Return on marketing investment (ROI) and return on advertising spend (ROAS) are widely used two key indicators for evaluating the efficiency of the direct marketing treatments. ROI is defined as the increase in sales generated by marketing as a percentage of the company's marketing spend and ROAS represents the additional revenue generated by the direct marketing media divided by its cost (Reutterer et al., 2006). Focusing on such metrics related to marketing investments can have significant benefits including decrease in inefficient spending, reallocation of funds to more effective treatments, accelerated marketing process times, growth in the top line and higher profit for companies. All investments made to generate additional revenue and profit are included in marketing cost. Hence, by developing a market model with reliable, impartial marketing statistics and metrics, marketing effectiveness aims to maximize marketing expenditure both in the short and long term, supporting and harmonizing with the brand strategy (Powell, 2008).

Marketing professionals may assess the success of direct marketing campaigns and make data-driven decisions to improve future initiatives. Setting clear objectives, choosing relevant performance measures, gathering precise data, evaluating the outcomes, and making any adjustments are all necessary for measuring marketing campaigns effectively. One of the key differences between B2C and B2B marketing campaigns is the way that effectiveness is measured. B2C campaigns often focus on metrics like conversion rate, click-through rate, and engagement rate, which are indicators of the consumer's interest and willingness to take action. In contrast, B2B campaigns often focus on metrics like lead generation, cost per lead, and revenue generated, which reflect the business's bottom line and the impact on sales. To measure the effectiveness of a marketing campaign, marketers should start by defining clear goals and objectives that align with the overall business strategy. These goals could include increasing brand awareness, generating leads, driving sales, or improving customer loyalty. Once the goals have been established, appropriate performance metrics should be selected to track progress towards achieving those goals. The choice of metrics will depend on the specific goals of the campaign and the target audience. Next, accurate data should be collected to track the chosen metrics. This may involve using analytics tools to track website traffic, email opens and clicks, social media engagement, or other key performance indicators. It's important to ensure that the data collected is accurate and reliable, as inaccurate data can lead to incorrect conclusions and ineffective decision-making. Once the data has been collected, it's time to analyze the results to determine the effectiveness of the marketing campaign. This involves comparing the actual performance metrics against the goals that were set at the beginning of the campaign. Based on the analysis, adjustments may need to be made to the campaign to improve its effectiveness.

Predictive analytics has traditionally been used in marketing to forecast future trends and identify patterns in customer behavior. However, in recent years, there has been a shift towards prescriptive analytics as a more effective way to assess marketing campaign performance. Prescriptive analytics goes beyond predicting what is likely to happen and instead focuses on identifying the best course of action to achieve a desired outcome (Lepenioti et al., 2020). It can be particularly useful in marketing because it allows marketers to optimize campaigns in real-time based on the insights generated from the data by improving campaign performance, reduce costs and increase ROI. The next section delves deeper into prescriptive analytics with introducing the causal inference concept.

2.3 Causal Inference

In direct marketing applications it is vital to understand the causal effect of a marketing treatment to run a successful campaign. The impact of an intervention such as mobile advertisement needs to be predicted on several individuals using observational data. Which customers, for instance, can be persuaded by a product promotion? Which clients should be treated with a certain type of treatment? These are the type of causal queries that concern a treatment's effect or change in outcomes. Traditional classification techniques suffer to address these queries since they only utilize relationships to forecast and predict outcomes (Li et al., 2021).

There are three levels of data science tasks as descriptive, predictive and prescriptive that are shown in Figure 2.3. The descriptive approach states what happened by providing actual results which is often a value, chart or dashboard, such as the average sales over the last 6 months. The predictive sciences investigate further to predict what will happen by providing a model, such as giving the probability of a user's purchase. Different from these two approaches, however, the prescriptive science aims to optimize the outcome by prescribing what to do such as which action to take to persuade user to purchase.

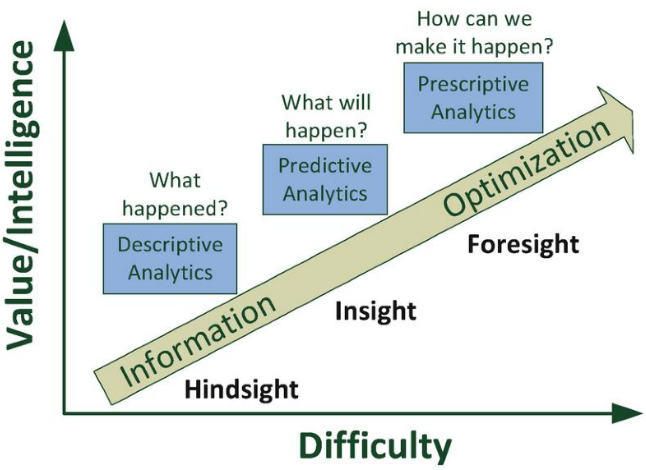

Figure 2.3 : Three levels of data science tasks

Although the majority of data science initiatives attempt to optimize business decisions, they frequently address forecasting or prediction issues. The main challenge of presciptive modeling is that it cannot be known what would have occured if a different course of actions had taken or what the counterfactual result would have been. (Shteingart et al., 2022). The causal relationships should be understood in order to model the counterfactual.

Causal inference is the group of techniques that are known to allow one to determine the precise relationship between an action and its result (Shteingart et al., 2022). The "basic identity of the causal inference" demonstrates that comparing the actual outcome (what occurs to the treated) with the counterfactual (what would have occured if they had not been treated) is the key conceptualization for understanding causality (Rubin, 1974; Varian, 2016). As Rubin introduced in his study, the counterfactual had to be estimated since one can not actually observe what would have happened to the treated group if the individuals had not been treated (Rubin, 1974). Hence, estimating the counterfactual prediction of what would have occured in the absence of the treatment is the crucial step in any causal analysis.

The goal of causal classification models is to predict whether a treatment would alter a person's outcome. When a product promotion advertising is the treatment in marketing application, causal classification is used to identify customers who are more likely to buy the product as a result of having seen the advertisement (Li et al., 2021). In order to understand the distinction between causal classification and normal classification, the difference between observed and potential outcomes needed to be understood. In Rubin's (1974) potential outcome framework, for a treatment T, each individual has two potential outcomes, denoted as follows:

Y^1 , outcome of the individual that is treated (T = 1)

Y^0 , outcome of the individual that is controlled (T = 0)

where Y denotes the observed outcome. Only one potential outcome can be observed for an individual at a certain time point. If an individual is observed purchasing the product after viewing an advertisement, $Y = 1 \mid T = 1$, the potential outcome when $T = 1$ is the same as the observed outcome ($Y^1 = 1$). However, the counterfactual outcome (Y^0), the purchase occurence when the advertisement is not viewed (T = 0) is not observed (Li et al., 2021).

Traditional classification predicts whether an individual will experience the intended outcome ($Y^1 = 1 \mid T = 1$) which is observed after the treatment, unconcerned about this outcome is due to treatment or not. On the other hand, causal classification aims to predict the difference in potential outcomes, $Y^0 = 0$ and $Y^1 = 1$, focusing on the change in the outcomes due to treatment (Li et al., 2021). The causal effect of the treatment that is the difference in the outcome is denoted as:

$$T_i \;\; = Y_i^1 - \; Y_i^0 \qquad (2.1)$$

where Y_i^1 shows the reaction of individual i, that has received the treatment and Y_i^0 shows the reaction of individual i, that has not been exposed to treatment. The individuals in the both groups, treatment and control, are compared to measure the treatment impact. It is important that these two groups are as similar as they can be for the comparison to be reliable (Karlsson, 2019). Conditional Average Treatment Effect (CATE) is used to estimate the expected causal effect of treatment, which is defined as follows in equation 2.2 where X_i denotes the vector of features (Athey and Imbens, 2016).

$$CATE := T\;(X_i) = \mathbb{E}\;[Y_i(1) - \; Y_i(0) |\; X_i] = \; \mathbb{E}[Y_i(1) |\; X_i] - \; \mathbb{E}[Y_i(0) |\; X_i] \qquad (2.2)$$

An effective approach to causal inference is uplift modeling which is the set of methods that a business can use to model CATE (Gutierrez and Gerardy, 2016; Shteingart, 2022). Uplift, also known as the predicted difference in result between the treated and control groups is the source of uplift modeling which is a common assessment measure in marketing (Shteingart, 2022). For instance, a business person in a telecommunications company is curious to know the impact of a promotional email that is sent to various customer profiles on their likelihood to renew their phone plans in the upcoming months. The manager would be in a better position to effectively target customers with such uplift information at hand.

Predicting customer uplift is both a causal inference and machine learning issue. It is a causal inference issue, due to the need of predicting the difference between two results that are peculiar to the individuals as a customer can only receive a promotional email or not. Uplift modeling is significantly depending on randomized experiments in order to solve this counterfactual issue by randomly assigning customer to either treatment or control groups. Hence, uplift modeling is also a machine learning problem due to the need of training different models and selecting the model that results the most reliable uplift prediction according to the performance metrics that demand reasonable cross-validation strategies along with potential feature engineering (Gutierrez and Gerardy, 2016).

As being in the center of the 2021 Economics Nobel prize, causal inference was recently publicly acknowledged and considered to be one of the challenges of data science (Hermann et al., 2022; Wing, 2020). Leading companies such as Linkedin, Uber, Netflix and Amazon use the causal framework extensively since it shows that concentrating on causal effect yields a superior return on investment (Shteingart et al., 2022). Predicting causal impact using observational data has recently gained popularity as a promising study area, opposed to randomized controlled trials due to large amount of accessible data and less funding requirements. Numerous causal impact estimating approaches for observational data have emerged, embracing the quickly growing machine learning field (Yao et al., 2022). As an approach to causal inference, next sections are covering uplift modeling and the literature review in this subject in more detail.

2.4 Uplift Modeling

Uplift modeling has been first appeared in the literature with the millenium shift and had been entitled with many names such as incremental value modeling, differential response analysis and true lift models. It indicates to the problem where it is needed to recognize the applicant that leads to the most desirable outcome from a set of treatments (Fang, 2017). Uplift models attempt to model the conditional average treatment of two or more separate mutually exclusive classes by subtracting the outcome from the respective treatment group by the outcome from the control group. Rubin's framework (1974) of modeling causality has been regulary employed in the literature for modeling causality with the propensity scores and diverse algorithms has been used further to model uplift (Karlsson, 2019).

Currently, most of the marketing activities are using traditional response models that attempt to classify customers according to the highest likelihood of purchasing. However, propensity models on their own are not certainly able to provide which customers are most likely to contribute to incrementality. It is necessary to use different statistical models that focus on customers whose propensities of sales or response are significantly influenced by incentives (Cao et al., 2017). Hence, uplift modeling has been proposed to predict the difference in response scores of the customer if treated and if not treated.

Given definite resources and a variety of alternatives for direct marketing communications as a treatment, uplift modeling can help businesses decide which course of action will result in the largest incremental response and return on marketing efforts. Since the aim is to target only the customers who are most likely to respond to the direct marketing campaign, uplift modeling is apt to optimize treatment action of direct communications and the related time and expenses invested (Munting, 2020). Hence, the uplift modeling could enable businesses to be able to target a lesser portion of the sample by determining which customers are more likely to buy before treatment is administered, which would save marketing expenses while retaining or even increasing revenue (Siegel, 2011).

Uplift modeling solves a fundamental issue that a customer may either receive or not receive a treatment, hence they cannot be both in treatment and control groups. Therefore, the uplift of an individual customer is measured by estimating the similar customer groups' conversion rates and comparing them with the conversion rate in control groups' (Pok, 2020). By defining population into four groups, uplift modeling predicts the buying behavior of customers treated and non-treated as shown in Figure 2.4. From top to bottom, the first dimension categorizes which customers will respond when contacted, and the second dimension further categorizes which customers will make a purchase even if not contacted (Kane, Lo and Zheng, 2014).

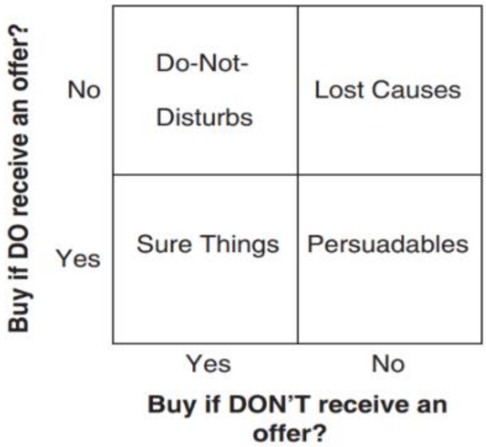

Figure 2.4 : Four types of marketing target groups (Kane, Lo and Zheng, 2014)

The lower-right part of the Figure 2.4 shows the customers that are worthy of expending the cost of contact and labelled as "Persuadables" who will purchase if contacted but won't purchase if not contacted (Siegel, 2011). These are the individuals that the uplift modeling aims to determine for direct marketing treatment to provide an incremental benefit to the company by turning the response from no purchase to purchase with the treatment. The lower-left quadrant shows the "Sure Things" which consists of the customers that would purchase the product or the service in subject with or without the treatment. "Lost causes" which are positioned in the upper-right part of the Figure 2.4 includes the customers who would not make any purchase with or without the treatment (Karlsson, 2019). Hence, the direct marketing effort should not be expended to the "Sure things" and "Lost causes" because treatment would be ineffective in their purchase decision. Lastly, the "Do-not-disturbs" which is also named as "Sleeping dogs" in literature (Siegel, 2011; Rzepakowski and Jaroszewicz, 2012) represents the customers who will be disturbed by the treatment and change their purchase decision to no-purchase so that the treatment will cause an opposite effect than the intended. In order to avoid customer churn, companies should not target their marketing treatments for the customers in "Do-not-disturb" and let them being "Sleeping Dogs".

It is important to note that a customer can not be in both treatment and control groups, hence it is not possible to know in which of the four region the customer belongs, only the row or the column is known. Uplift modeling can provide a solution to this problem. The original idea behind uplift modeling is to use two training data sets, one set including the control group and the other containing the treatment group. The treament group contains the individuals who are communicated of the marketing campaign and the control group contains the individuals who were not targeted and received any marketing treatment. Instead of modeling a single class probability, the uplift modeling seeks to model the difference among the conditional class probabilities in the control and treatment groups (Börthas and Sjölander, 2020). As a result, the real profit from targeting a specific customer group can be modeled. Although uplift modeling promises a clear opportunity to improve marketing campaign efficiency, it is challenging to build both from practical and technical standpoint. The effectiveness of uplift modeling in sales is quite difficult and requires a well-balanced experimental design with appropriate test and control groups, constant marketing incentive and a notice on the seasonality as a confounding factor (Cao et al., 2017).

In the last several years, large number of uplift modeling approaches and algorithms are applied and proposed to literature. Table 2.4 gives an overview of previous studies with uplift modeling approach, application area and business objective with the contribution. It can be seen that many of the studies aim at predicting uplifts for binary outcomes such as conversion, purchase, visit or churn. More recently, revenue uplift modeling approaches have also attracted attention in order to meet the financial goals as well (Gubela and Lessmann, 2021; Albert and Goldenberg, 2021).

Table 2.4 : Summary of previous studies on uplift modeling

Authors	Modeling Approach	Application	Contribution
Hansotia and Rukstales (2002)	Two model approach, Modeling uplift directly	Real-life marketing campaign	By using uplift modeling, generating an incremental visit from a customer who would not be as likely to visit in the absence of a promotion.
Lo (2002)	Modeling uplift directly	Synthetic data-set	Proposing a methodology to identify the customers whose decisions will be positively influence
Radcliffe and Simpson (2007)	Two model approach, Modeling uplift directly	Real-life retention campaign	Showing that both retention campaigns can be improved using uplift modeling to model savability and positive impact on the profitability of the campaign with the overall achieved retention level.
Jaskowski and Jaroszewicz (2012)	Two model approach, Class variable transformation	Clinical trial data	Demonstrating uplift modeling may still be capable of selecting a subgroup of patients for which the treatment is successful even if the treatment overall is not beneficial.
Guelman *et al*. (2012)	Modeling uplift directly	Real-life retention campaign	Proposing a new procedure that can be used to identify the target customers who are likely to respond positively to a retention activity.
Rzepakowski and Jaroszewicz (2012)	Modeling uplift directly	Publicly available datasets: clinical trial	Presenting tree-based classifiers designed for uplift modeling in both single and multiple treatment cases and showing significant improvement over previous uplift techniques.
Su *et al*. (2012)	Modeling uplift directly	Publicly available datasets: training program	Introducing facilitating score to account for both the confounding and interacting impacts of covariates on the treatment effect.
Zaniewicz and Jaroszewicz (2013)	Modeling uplift directly	Publicly available datasets: marketing campaign, clinical trial	Presenting the adaptation of support vector machine methodology to uplift modeling problem
Athey and Imbens (2015)	Two model approach, Class variable transformation	Synthetic data-set	Introducing methods for constructing trees for causal effects

Table 2.4 (continued) : Summary of previous studies on uplift modeling

Soltys et al. (2015)	Two model approach, Modeling uplift directly	Publicly available datasets: marketing campaign, clinical trial	Providing a thorough analysis of ensemble methods in uplift modeling and showing excellent performance.
Gutierrez and Gerardy (2016)	Two model approach, Class variable transformation, Modeling uplift directly	Synthetic data-set	Providing a comparison of the three approaches and contributing to the literature by providing a unified review of uplift literature
Kondareddy *et al*. (2016)	Modeling uplift directly	Real-life marketing campaign	Proposing a two-step approach for incremental response modeling
Cao *et al*. (2017)	Two model approach	Real-life marketing campaign	Proving selection of customers with the largest response rate does not guarantee maximum incrementality.
Diemert *et al.* (2018)	Two model approach, Class variable transformation	Publicly available datasets: marketing campaigns	Highlighting the need for large scale benchmarks for uplift models and not one uplift modeling method performs best in all datasets.
Rudas and Jaroszewicz (2018)	Two model approach, Class variable transformation	Publicly available dataset: children physical activity	Analyzing two uplift modeling approaches, identifying the situations when each model performs best and proposing third model for uplift linear regression.
Gubela *et al.* (2019)	Two model approach, Class variable transformation	Real-life marketing campaign	Consolidating prior work in uplift modeling, empirically showing the performance of uplift modeling vs. conversion modeling and benchmarking strategies.
Zhao and Harinen (2019)	Two model approach	Synthetic and real-world datasets	Evaluating uplift modeling approaches for two common practical challenges: (1) when there are multiple treatment groups present and (2) when there are different costs associated with the treatments.
Shimizu *et al.* (2019)	Class variable transformation	Real-life marketing campaign	Achieved a cost reduction with uplift modeling with only a negligible reduction in the number of acquired customers who make non-coupon purchases.

Table 2.4 (continued) : Summary of previous studies on uplift modeling

Goldenberg *et al*. (2020)	Two model approach, Class variable rtransformation	Real-life marketing campaign	Introducing an optimization framework to solve the personalized promotion assignment and showing that a personalized treatment assignment can turn an under-performing promotions campaign into a campaign with significant uplift in ROI.
Olaya *et al*. (2020)	Two model approach, Class variable transformation	Real-life student retention campaign	Demonstrating uplift modeling framework maximizes the effectiveness of retention efforts in higher education institutions
De-Caigny *et al*. (2021)	Modeling uplift directly	Real-life retention campaign	Contributing to B2B retention management literature by demonstrating with the uplift algorithm the return on marketing investment is maximized by targeting only risky customers for whom the retention campaign is effective by proposing a segmented uplift approach
Röbler *et al*. (2021)	Two model, Class variable transformation, Modeling uplift directly	Real-life marketing and retention campaigns	Suggesting a weighted ensemble approach to reduce the volatility in uplift modeling
Gubela and Lessmann (2021)	Two model, Class variable transformation, Modeling uplift directly	Real-life marketing campaign	Introducing new evaluation metrics to achieve higher profit across different uplift models
Devriendt *et al*. (2021)	Two model approach, Modeling uplift directly	Real-life retention campaign	Introducing a new evaluation metric for achieving maximum potential profit with an uplift model and showing that uplift models outperform predictive models and lead to improved profitability of retention campaigns.
Albert and Goldenberg (2021)	Two model approach, Modeling uplift directly	Real-life marketing campaign	Providing an adaptive and real-time solution to optimal promotion selection by ensuring budget constraints compliance
Baier and Stöcker (2022)	Two model approach, Modeling uplift directly	Real-life marketing campaign	Providing a new approach to model profit uplift at the individual level without needing an unrelated second step to transform modeled outcomes

Table 2.4 (continued) : Summary of previous studies on uplift modeling

Nyberg and Klami (2022)	Two model approach, Class variable transformation, Modeling uplift directly	Publicly available dataset: marketing campaign	Proposing new undersampling methods for handling class imbalances for uplift modeling.
Sanisoglu *et al.* (2022)	Two model approach	Real-life marketing campaign	Demonstrating that uplift modeling can be used to determine the most favorable customer segment in terms of targeting for a cross-sell marketing campaign in telecommunications sector.
Hu (2023)	Two model approach, Class variable transformation, Modeling uplift directly	Real-life marketing campaign	Proposing two step feature selection approaches for removing irrelevant and redundant features before building uplift models.
Proença and Moraes (2023)	Class variable transformation, Modeling uplift directly	Synthetic dataset	Introducing incremental profit per conversion as a novel uplift measure of promotional campaign efficiency.

3. METHODOLOGY

There are many approaches to uplift modeling. In this section, the formulation for uplift modeling problem and three common approaches to the uplift problem is introduced. Further, classification methods and evaluation techniques are explained in detail.

3.1 Uplift Modeling

As described in Section 2.4, in uplift modeling there are two groups as Treatment and Control groups consisted of the observations that received treatment or not. Treatment and control groups are identified with the superscript ^{T} that states treatment group calculations and superscript ^{C} that states control group calculations. The probabilities in the treatment and control groups are shown as P^T and P^C respectively. The input variables are same for all models that include treatment and control group data. The response variable which shows the result of the purchase activity takes on values as $Y \in \{0,1\}$ where 1 accounts for the positive response as buying the product and 0 accounts for the negative response as not buying the product. The input variables are denoted as $X \in \{X_1,...,X_p\}$ where p corresponds to the number of the variables in the data whereas x represents a single observation. M^U is the notation for the expected uplift which is defined as the difference among conditional probabilities of the treatment and control groups:

$$M^U = P^T(Y = 1|X = \boldsymbol{x}) - P^C(Y = 1|X = \boldsymbol{x}) \tag{3.1}$$

If the result is negative in equation (3.1), then it states that the probability of a control group customer making a purchase is higher than the customer who is treated (Börthas and Sjölander, 2020). Hence, it shows that the campaign had negative effect on customers.

Uplift modeling can be considered as a regression or a classification assignment according to the nature of the objective. It can be applied as a regression task if the conditional net gain is taken as a numerical quantity to be measured, or as a classification task to anticipate the class of a particular customer who will react to a treatment positively or not (Börthas and Sjölander, 2020). In this thesis, uplift modeling is evaluated as a classification task to predict the customers who will purchase when treatment is applied. Uplift can be modeled by three different methods as two model approach, class variable transformation method and modeling uplift directly with tree-based methods which are explained in detail in the following subsections.

3.1.1 Two model approach

The simplest approach to uplift modeling is the Two Models Approach. This method builds two models that are trained separately by the treatment and the control group. The first one is trained on observations that received treatment and the second is trained on observations that did not receive treatment (in control group). The uplift is calculated by subtracting the class probabilities predicted by the model that includes the control group data from the class probabilities predicted by the model that includes treatement group data as shown in Eq. (3.1) (Radcliffe and Simpson, 2007). Figure 3.1 demonstrates the two models approach in uplift modeling that took place in three steps. First step, is to build a response model for the control group. Second step is to build a similar model for the treatment group. Finally, the last step is calculating the incremental response by subtracting one model score from the other.

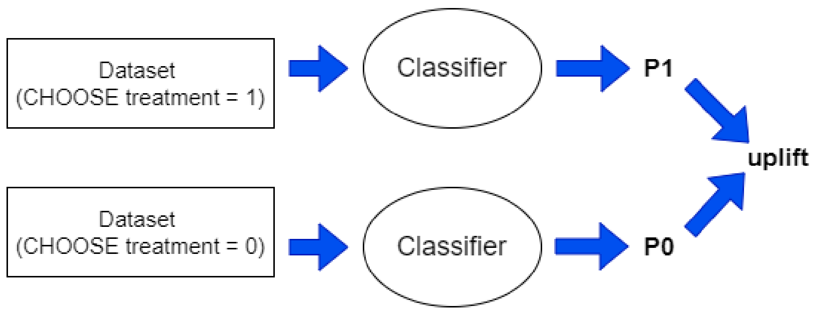

Figure 3.1 : Two Models Approach where P1=P(purchase|T=1) and P0=P(purchase|T=0) (Ludziejewski et al., 2020)

The two model approach can be defined mathematically as shown in equation 3.1, demonstrating that for each classified observation, the class probabilities predicted by the model involving the control group data is subtracted from the class probabilities predicted by the model involving the treatment group data. The input variables which are denoted as $X \in \{X_1,...,X_p\}$ in equation 3.1 is same for both models but derive from different datasets indicating that the model parameters in P^T is different from the model parameters in P^C.

Literature has posit several studies that applied two model approach as listed in Table 2.2. Hansotia and Rukstales (2002), applied two model in a real-life promotion campaign and aimed to optimize the campaign design by improving return on direct marketing investment. They demonstrated that with the presence of a promotion, an incremental visit from a customer who would not be likely to visit is generated (2002). Similarly for a real-life campaign setting, Cao et al. (2017) used two model approach to identify customers that will create incremental transaction and purchase and showed that only selecting the individuals with largest response rate does not ensure maximum gain. Further more recent studies by Goldenberg et al. (2020); Albert and Goldenberg (2021) and Baier and Stöcker (2022) utilized two model approach to compare the model results with other uplift modeling methods such as class variable transformation and tree-based uplift modeling. It has been shown that it is possible to optimize the real-life promotion campaigns in budget limitations by selecting each customer which promotion to present in order to maximize the purchase (Albert and Goldenberg, 2021). For retention prevention campaigns also, two model approach is utilized as one of othe uplift modeling methods by Radcliffe and Simpson (2007) and Devriendt et al. (2021) showing that retention campaigns can be improved by using uplift modeling in terms of predicting customer churn with maximum profit.

Two model approach can be applied to any classification problem and easy to implement for most of the direct marketing campaigns to estimate the uplift. However, it is argued in the literature that two models approach is rather a flawed method for predicting uplift. Since no model performs perfect prediction, there is no guarantee that the substraction of the result from two good models would become a good model of uplift (Radcliffe, 2007; Rzepakowski and Jaroszewicz, 2012). In this study, for the purpose of having a simple approach to compare with more advanced approaches, two models approach is applied with logistic regression. Oversampled training data is used to train the model and tested on the validation set. Two model approach results are then evaluated and compared with other uplift modeling methods.

3.1.2 Class variable transformation

Class variable transformation is first proposed by Jaskowski and Jaroszewicz (2012). This method constructs a single model that directly models the uplift rather than separately modeling treatment and control groups probabilities as two model approach does. The group belonging (treatment and control group) and purchase is transformed into a single feature defined as $Z \in \{0, 1\}$ such that

$$Z = \begin{cases} 1 \;\; if\; Y = 1\; and\; T, \\ 1 \;\; if\; Y = 0\; and\; C, \\ 0 \qquad otherwise. \end{cases} \tag{3.2}$$

where treatment group is denoted as *T* and control group is denoted as *C*. Equation (3.2) transforms the probabilistic classification model into an uplift prediction model. If the individual is in the treatment group and made a purchase (*Y = 1* and *T),* then this individual is either in the *Sure Things or Persuadables* groups (see Figure 2.3) and *Z* is set to 1. In other respects, if the individual is in the control group and did not make a purchase (*Y = 0* and *C)*, then this individual is either in the *Lost causes* or *Persuadables groups* and *Z* is also set to 1. For all other cases the *Z* is set to 0, which makes all individuals in these cases belong to *Do-not-disturbs* and eliminating the risk of contacting them with the treatment (Jaskowski and Jaroszewicz, 2012).

Different from the Two-model approach, the Class variable transformation does not solely focus on *Persuadables* which would be the optimal group to target for a treatment. Due to the fundamental problem of causal inference, it is not possible to observe the outcome of a single individual in both the treatment and control group (Börthas and Sjölander, 2020). Hence, the class variable transformation can not be used to particularly target *Persuadables.*

The transformation proposed by Jaskowski and Jaroszewicz demonstrated that modeling uplift is available by the transformed variable *Z* and equation 3.2 if there is a binary outcome and the size of treatment and control group is balanced. This method is functionally advantageous because that the single linear models are much simpler to apply and the direction and strength of the impact of each variable on the uplift can be efficiently determined with a single set of coefficients (Jaskowski and Jaroszewicz, 2012). Another advantage of class variable transformation approach is that the single target feature which is the transformed class can be modeled by any statistical machine learning methods to predict conditional class probabilities.

Jaskowski and Jaroszewicz (2012), proposed and applied CVT to clinical trial data aiming to discover groups of patients for which the treatment would be most beneficial. It has been shown that uplift modeling with CVT could select a subgroup of patient for which the treatment is successful even if the treatment overall is not beneficial. Further, it has been demonstrated by Shimizu et al. (2019) that in a direct marketing campaign, coupons could be distributed to the most promising customers and the campaign effectiveness could be improved by achieving a cost reduction with uplift modeling. Other recent studies by Goldenberg et al.(2020); Röbler et al. (2021) and Gubela and Lessman (2021) utilized CVT for real-life promotion campaigns for optimizing the incremental outcome subject to the required ROI constraints.

Class variable transformation is applied in this study with logistic regression and compared with other proposed uplift modeling methods in order to determine the best suited approach for the problem that is subject to this thesis.

3.1.3 Modeling uplift directly

Modeling the uplift directly with tree-based methods by adapting existing machine learning algorithms is another approach that has been widely used in uplift modeling. This approach applies tree-based algorithms to separate and classify the data into subpopulations which make them useful for modeling discrepancies among the treatment and control group (Radcliffe and Surry, 2011). The essence of the splitting criteria in the decision trees allow them to be convenient for modeling uplift directly.

In order to ensure that the uplift is captured, the splitting criterion is modified in the tree-based algorithm. There are numerous methods proposed in the literature to accommodate the splitting criterion to best estimate the uplift. However, there isn't any agreement on which method is the most effective one. Hansotia and Rukstales (2002) proposed a splitting criterion which attempts to maximize the variance between the differences in treatment and control groups' probabilities for the left and right subnodes. Further, Rzepakowski and Jaroszewicz (2012) suggested that the splitting criteria should be based on conditional distribution divergence as a way to quantify how two probability distributions differ from one another. With this method, the test is chosen at each level of the tree so that, after a split has been made the divergence between the class distributions in the treatment and control groups is maximized. In this study, as the splitting criterion two distribution divergence measures, the Kullback-Leibler divergence and the squared Euclidean distance, from Rzepakowski and Jaroszewicz (2012) is used. Given the probabilities $P = \{p_1, p_2\}$ and $Q = \{q_1, q_2\}$, the divergences are defined respectively as:

$$KL(P:Q) = \sum_i p_i \log\frac{p_i}{q_i} \tag{3.3}$$

$$E(P:Q) = \sum_i (p_i - q_i)^2 \tag{3.4}$$

i is equal to 1 and 2 for binary classification as it is in the problem covered in this thesis where the outcome has two classes $Y \in \{0, 1\}$. p_1 and p_2 are the treatment probabilities $P^T(Y = 0)$ and $P^T(Y = 1)$; q_1 and q_2 are the control probabilities $P^C(Y = 0)$ and $P^C(Y = 1)$.

Given a divergence measure D, the splitting criterion is defined in equation (3.5) decides the split of that node with the largest value of D_{gain}.

$$D_{gain} = D_{aftersplit}\left(P^T(Y), P^C(Y)\right) - D_{beforesplit}\left(P^T(Y), P^C(Y)\right) \tag{3.5}$$

P^T and P^C show the class probabilities in the treatment and control groups before and after the split is performed. The resulting divergence measure after a split has been made is demonstrated in equation (3.6).

$$D_{aftersplit}\left(P^T(Y), P^C(Y)\right) = \sum_{a=a_1}^{a_2} \frac{N_a}{N} D\left((P^T(Y|a), (P^C(Y|a)\right) \tag{3.6}$$

N is the number of observations before the split has been made, $a \in \{a_1, a_2\}$ denotes the left and right lead of that split and N_a is the number of observations in each leaf after the split (Rzepakowski and Jaroszewicz, 2012). If the split is created from a binary variable, $A \in \{0, 1\}$, the left leaf a_1 represents $A = 0$ and the right leaf represents $A = 1$ in equation (3.6).

Tree-based algorithms for modeling uplift directly is prevalent in the uplift modeling literature and used by Hansotia and Rukstales (2002); Radcliffe and Surry (2011); Rzepakowski and Jaroszewicz (2012); Guelman et al. (2015); Röbler et al. (2021); Devriendt et al. (2021) and Albert and Goldenberg (2021) to mention a few. One advantage of this method is that modeling uplift with this method is available for both binary and continous target features (Radcliffe, 2007). In this study, modeling uplift directly is implemented using decision tree learners with random forests and the specific R-package (uplift) by Guelman (2014) is used. For the splitting criterion, both of the divergence measures, Kullback-Leibler and the squared Euclidean distance, are utilized and compared.

3.2 Classification and Prediction

Classification and prediction are integral elements of data science, used to interpret and forecast data behavior. Classification involves categorizing data into predefined groups based on shared characteristics and prediction forecasts future trends based on existing data using methods such as regression analysis. This section will focus on how logistic regression and random forests, two powerful techniques, are used in classification and prediction, providing insights into their methodologies and applications.

3.2.1 Logistic regression

Logistic regression is the simplest method for classification purposes since it is a generalized linear model and easy to implement. In this study, logistic regression is used for a binary response, aiming to estimate the conditional probability of an individual purchasing the product.

The linear logistic model models the log-likelihood ratio as:

$$log\frac{P(Y=1|X=x)}{P(Y=0|X=x)} = \beta_0 + \beta^T \boldsymbol{x} \tag{3.7}$$

where β_0 is the intercept term, $\beta \in R^p$ denote the vector of regression coefficients and x denotes one of the observations (Hastie et al., 2015). For conditional probability estimation the equation (3.7) can be transformed into equation (3.8) as follows:

$$P(Y=1|X=\boldsymbol{x}) = \frac{e^{\beta^0+\beta^T x}}{1+e^{\beta^0+\beta^T x}} \tag{3.8}$$

Among uplift modeling methods; two model approach and class variable transformation are applied with logistic regression and the models are fitted with a K-fold cross validation procedure to maximize the usage of training data. Cross validation (CV) is a prevalent resampling method which regularly takes samples from a training set and refit a model on each sample to learn more about the model. In this study, K-fold cross validation is applied.

In K-fold cross validation the data is split into K equal parts where one part is set aside and used as the test set and the other parts ($K - 1$) are used for training. The resulting model is used to obtain a performance estimation on the first part. Setting aside a different part in each time, this process is repeated K times and the performance estimation over the K parts are then averaged (Hastie et al., 2008).

For choosing K value, the bias-variance trade off should be considered. If K gets larger, the size difference between the training set and the entire data set decreases causing cross-validation estimators' variance increasing. When compared to the parts that are not as highly correlated, the variance of the mean of highly correlated similarly assigned parts is higher. Hence, $K = 5$ is chosen as a good compromise.

3.2.2 Random forests

Random Forests is a statistical machine learning method that builds many decision trees and then average them for reaching a final prediction. Random forests work by selecting B bootstrap samples from the training data, then using this samples to train B different decision trees. The name "random forests" refer to the fact that each tree is constructed using random input variables before each split. This averaging technique helps to avoid overfitting and improve stability. The algorithm for random forests is explained further in Algorithm 1 (Hastie et al., 2008):

Table 3.1 : Algorithm 1

Algorithm 1 Random Forests

1. **For** $b = 1$ *to* B **do**
   (a) Draw a bootstrap sample **Z*** of size N from the training data with replacement.
   (b) Grow a random-forest tree T_b to the bootstrapped data **Z***, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached.
      i. Select m variables at random from the p variables.
      ii. Pick the best variable/split-point among the m.
      iii. Split the node ino two daughter nodes.
   end
2. Output the ensemble of trees $\{T_{b=1}\}^B$

3. To make a prediction to a new point **x:**

   Regression:

$$f_{rf}^{B}(\boldsymbol{x}) = \frac{1}{B}\sum_{b=1}^{B} T_b(\boldsymbol{x})$$

   Classification: Let $C_b(\boldsymbol{x})$ be the class prediction of the bth random forest tree

$$C_{rf}^{B}(\boldsymbol{x}) = majority\ vote\ \{C_b(\boldsymbol{x})\}_1^B$$

A *random* sample of m random variables are selected as split candidates from the p predictors when a split is made. In classification problems, m is set to be approximately $\sqrt{p}$ due to the less value of m reducing the variance when there are many correlated variables (Hastie et al., 2008).

For uplift modeling, an algorithm of random forests is proposed by Guelman et al. (2012) in which the training data set both contains data from treatment and control groups. The uplift prediction of the individual trees are averaged and an uplift is achieved. The number of trees in the forest and the number of predictions in the random subset of each node is used as two tuning parameters. Based on Algorithm 1, the proposed algorithm for uplift is shown in Algorithm 2 (Guelman et al., 2012).

In order to grow a tree, a criterion is needed for making recursive binary splits. Among the several splitting criteria available for this purpose, the most used two splitting criteria that are based on conditional distribution divergences (Euclidean-distance and Kullback-Leibler) from Rzepakowski and Jaroszewicz (2012) is used in this study.

Table 3.2 : Algorithm 2

Algorithm 2 Random Forests for *Uplift modeling*
1. **For** $b = 1$ *to* B **do** (a) Draw a bootstrap sample **Z*** of size N from the training data with replacement. (b) Grow an uplift decision tree UT_b to the bootstrapped data **Z***, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached. i. Select m variables at random from the p variables. ii. Pick the best variable/split-point among the m.The split criterion should be based on a conditional divergence measure. iii. Split the node ino two daughter nodes. **end**
2. Output the ensemble of trees $\{UT_{b=1}\}^B$
3. The predicted uplift for a new data point $\boldsymbol{x}$ is achieved by averaging the uplift predictions of the individual trees in the ensemble: $$f_{uplift}^{B}(\boldsymbol{x}) = \frac{1}{B}\sum_{b=1}^{B} UT_b(\boldsymbol{x})$$

3.3 Evaluation Methods

Evaluation methods are critical for assessing the effectiveness of predictive models in data science. The Receiver Operating Characteristic Curve (ROC) is a key tool in this area especially useful for evaluating classification models. Similarly, the Qini curve is vital in uplift modeling for measuring the additional impact of certain treatments or actions. These curves both provide a visual and a quantitative way to understand and improve model performance. This section will focus on explaining how ROC and Qini curves are essential in enhancing the effectiveness of models.

3.3.1 Roc curve

Receiver Operating Characteristic Curve (ROC) is one of the most popular visual chart that is used to evaluate classification models. It simply evaluates the classification model's performance by calculating the categorization of positive cases as positives and negative cases as negatives. In classification models, the predicted and the actual response for each individual may typically be compared according to four different outcomes shown in a confusion matrix. An example of a confusion matrix is demonstrated in the Figure 3.2 below.

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Figure 3.2 : A confusion matrix (Draelos, 2019)

The columns labeled as “Actually Positive” and “Actually Negative” refer to the true classes in the test set; where as the rows labeled as “Predicted” shows the predicted class for the observations in the test set by the model. True positives (TP) is when the model correctly classifies the positive case as positive and True negatives (TN) is when the model correctly classifies the negative case as negative. On the other hand, False positives(FP) occurs when the model classifies the case as positive but the actual result is negative and False negatives (FN) occurs when the model classifes the case as negative but the actual result is positive. Hence, in FP and FN the model classifies the observation to the wrong class.

In order to plot a ROC curve, the metrics from the confusion matrix is used. On the y-axis of the ROC curve, the true positive rate (TP rate) that is also named as sensitivity or recall shown in Equation (3.9) is computed:

$$TP\ rate = \frac{TP}{P} = \frac{number\ of\ positives\ correctly\ classified}{total\ number\ of\ positives} \qquad (3.9)$$

On the x-axis of the ROC curve, the false positive rate (FP rate) shown in Equation (3.10) is computed:

$$FP\ rate = \frac{FP}{N} = \frac{number\ of\ negatives\ incorrectly\ classified}{total\ number\ of\ negatives} \qquad (3.10)$$

The other important performance metrics for assessing classifiers is the accuracy and precision values that can be calculated as shown in Equation (3.11) and (3.12) respectively.

$$accuracy = \frac{TP + TN}{P + N} \qquad (3.11)$$

$$precision = \frac{TP}{TP + FP} \qquad (3.12)$$

For a better performance than a random classifier, the classification model should give a result in the upper triangle above the diagonal line in ROC curve plot. The area under the curve (AUC) is also used to compare different classifier's performance which should be greater than 0.5.

In uplift models, since the true answer of the uplift for each individual can never belong to both treatment and control group simultaneously, ROC curves is not adequate to evaluate the performance. Hence, to solve this dilemma for uplift models, the commonly used method is Qini curve which is explained further in the next section.

3.3.2 Qini curve

Qini coefficient and Qini curve are used as a generalization of the uplift curve that is developed by Radcliffe (2007). The Qini curve plots the incremental purchase computed per segment and group which is computed according to Equation (3.13):

$$u_{kp} = r_k^t - \frac{r_k^c n_k^t}{n_k^c} \qquad (3.13)$$

where u_{kp} denotes the predicted uplift per segment; r_k^t and r_k^c represent the number of purhases per segment in treatment and control groups; n_k^t and n_k^c is the group sizes per segment for treatment and control groups respectively. An example of computation of incremental purchase is shown in Table 3.3. For each individual the model returns a score and the data is sorted in decreasing order. Then the data is splitted into k segments and the number of individuals and the number of purchases in each segment is computed for treatment and control groups (Karlsson, 2019). In the first row of Table 3.3, segment 1 has 20 purchases in the treatment group and 5 purchases in the control group. Hence, the cumulative incremental purchases for segment 1 is 20 – 5.(100/100) = 15.

Table 3.3: An example of incremental purchase computation

	Targeted		Control		Treated - Control
Segment (k)	Cumulative purchases ($r^t{}_k$)	Cumulative targeted ($n^t{}_k$)	Cumulative purchases ($n^c{}_k$)	Cumulative targeted ($r^c{}_k$)	Cumulative extra purchases (u_{ka})
1 (10%)	20	100	5	100	15
2 (20%)	30	200	10	200	20
3 (30%)	44	300	22	300	22
…	…	…	…	…	…

In order to compute the uplift per segment, the cumulative number of purchases for the treatment and the control group is subtracted. The y-axis of the Qini curve is the uplift that is the cumulative number of incremental purchases and the x-axis shows the number of segments that is treated. This incremental purchase method makes the Qini curve to show the potential negative impact of the treatment. An example of Qini curve is illustrated in Figure 3.3.

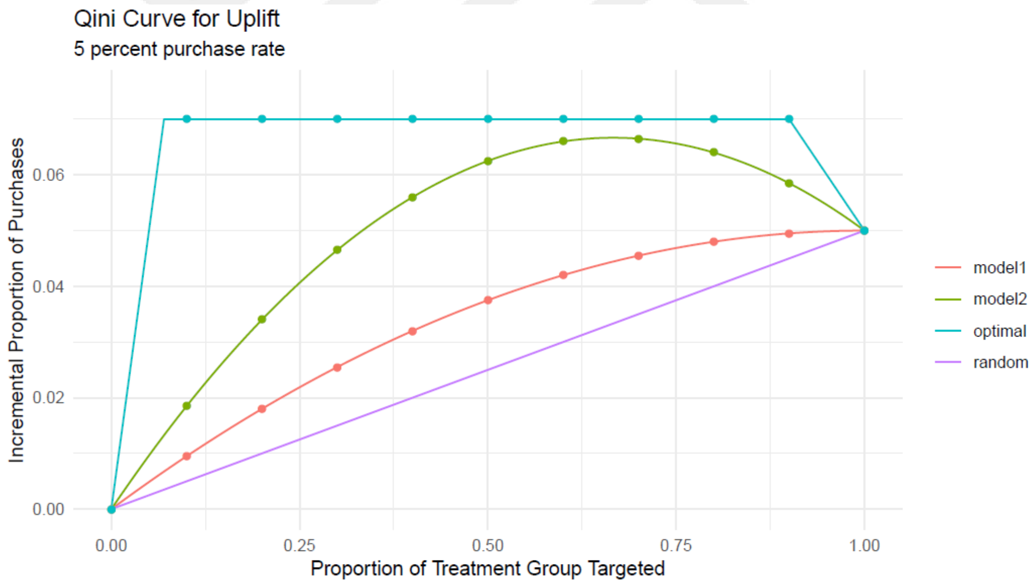

Figure 3.3 : Qini curve demonstrating optimum curve vs. random (Karlsson, 2019)

The blue line shows the *optimal* curve which theoretically represents the best possible outcome where each individual that can be persuaded by treatment is identified and given treatment first, assuming all the "Sure things" are "Persuadables". The purple diagonal line is the *random* curve representing the outcome if all individuals were given treatment in random order. The red and green curves are representing two uplift models, among which model 2 showed with green curve outperforms the other model which is represented with red curve. It can be seen from the example as well that if the entire sample is targeted for treatment, the uplift model and random allocation gives the same result due to giving treatment to each individual. In an ideal uplift model, “Persuadables” should receive the highest score and the “Do-not-disturbs” should get the lowest score.

A Qini curve with an uplift that is higher when treating a subgroup of the population than when treating the full population is indicative of a good model performance (Radcliffe, 2007). Hence, in a successful uplift model a Qini curve should yield greater uplift when a smaller part of the treatment group is targeted rather than the entire population. In the example Qini curve in Figure 3.3, the total purchase rate for the population is 5%, meaning that only 5% of the customers in the treatment group purchases without any negative impact of treatment is present or it may be possible that 8% of the customers in the treatment group make a purchase and 3% of the customers that are treated determine not to purchase due to the treatment. Hence, it is possible with uplift model to yield a higher purchase rate by decreasing the treatment group size by excluding the “Do-not-disturbs” from the treatment group. In the example of Figure 3.3, the highest possible uplift could be captured by model 2, suggesting that approximately 65% of the treatment group should be given treatment to gain approximately 6.5% uplift.

The Qini value is defined as the area between the Qini curve and the random model curve. A positive Qini coefficient is desired since a negative value suggests that doing a certain action will have worse consequences than doing nothing (Guelman, 2014).

4. DATA

In this study real-world data is used to perform uplift models. This section will provide details of the dataset with variables, the ethics principles of using real-world data and the data preprocessing actions that are applied step-by-step.

4.1 Dataset

The data set used in this study was based on demographic and behavioral characteristics of 21.439 customer comprising December 2020- January 2021 period, collected anonymously from a mobile telecommunications company. It contains the data of a discounted cross-sell campaign which offers a weekly data add-on product that could be purchased additionally on top of the monthly tariff plan. The offered product and the price is same for all customers, thus there is not any customization in the customer level. The campaign is communicated to the customers with the online push notification that is assumed to be read by them. Hence, the treatment is defined as receiving a campaign push notification. When a customer received push notification, disregarding if the customer clicked or not, they are considered to be in the treatment group. If the customer purchases the discounted product in the following week the notification is sent, it is considered to be a purchase as a result of the treatment. Table 4.1 below shows the total purchase rate in the treatment and control group which are 6.3% and 3.0% respectively. The purchase rate is calculated by n/N where n is the number of purchasers and N is the total sample size. Hence, the realized treatment effect is approximately 3.3% (6.3% - 3.0%).

Table 4.1 : Data description

Groups	No Purchase	Purchase	Total # of customers	Purchase Rate
Treatment	1,773	120	1,893	6.3%
Control	18,956	590	19,546	3.0%
Total # of customers	20,729	710	21,439	3.3%

The selection of the treatment group is not completely random but applied according to the information of previous campaign results and frequent users. The service provider defined frequent users as the customers who had purchased additional data packs before without specifying any other rule set. The reason why all of the customers are not included in the treatment is that there are many diverse campaign notifications that are sent to customers each day, and the contact policy regulation restricts providers to send more than one campaign notification to the same customer in the two consecutive days. Thus, it is aimed to identify the most “persuadable” customers to target specifically. For this campaign the control group size is sufficient for an uplift modeling analysis since it is ten times larger than the treatment group (Radcliffe and Surry, 1999).

4.1.1 Ethics

Permission marketing is defined as “the privilege (not the right) of delivering anticipated, personal and relevant messages to people who actually want to get them” by Seth Godin (2008).

By giving permission to marketing communication, the customers have accepted that the data is used within the provider to improve their services to their customers. All the information used in this study is anonymized before it has been given to the author and it is ensured that there is no risk of identifying a single individual since the features do not contain any individual information that would make anyone distinguishable.

4.2 Variables

The total data contains 26 attributes. All features are summarized in the Table 4.2 below with variable names in the model, description, class and information on value that the variables could take. All categorical variables are encoded and numeric values are scaled with R programming software as part of data preprocessing before dividing data into training and validation sets and building uplift models.

Table 4.2 : Variable description

Variable	Description	Class	Values
Purchase	Indicator if a customer purchased the addon with credit card within the campaign window (1 week) from treatment. This is the feature that is modeled.	Binary	1 if customer purchased the addon, 0 if not purchased
Treatment	Indicator if customer is in treatment or control group.	Binary	1 if customer belongs to the treatment group, 0 if customer belongs to the control group
Segment	Business segment of the customer	Categorical	Corp, Sme or Soho
Tariffclass	Main category of the customer's monthly tariff plan	Categorical	Employee, Holding, High-class or Soho
Dataoveruser	Indicator if customer has used over their tariff plan in the 2 month period that includes campaign duration	Binary	1 if customer overused, 0 if not
Tenuremonths	Duration of the customer lifetime in *months*	Numeric	Numeric
Dataallowance	Data amount of customer's monthly tariff plan in *GB*	Numeric	Numeric
Pricepergb	Monthly average fee according to monthly data allowance (*TL*)	Numeric	Numeric
Avgdatausage	Average monthly data usage in the last 6 month period that includes campaign duration (GB)	Numeric	Numeric
Capped	Indicator if the customer has a capped tariff plan which stops data usage if the customer is out of tariff plan allowance	Binary	1 if the customer has a capped tariff plan, 0 if not
Pbtwelve	Indicator if the customer purchased addon in the last 12 months	Binary	1 if customer purchased at least 1 addon, 0 if not
Pbnine	Indicator if the customer purchased addon in the last 9 months	Binary	1 if customer purchased at least 1 addon, 0 if not
Pbsix	Indicator if the customer purchased addon in the last 6 months	Binary	1 if customer purchased at least 1 addon, 0 if not
Pbthree	Indicator if the customer purchased addon in the last 3 months	Binary	1 if customer purchased at least 1 addon, 0 if not

Table 4.2 (continued) : Variable description

Frtwelve	The number of addons purchased in the last 12 months	Numeric	Numeric
Frnine	The number of addons purchased in the last 9 months	Numeric	Numeric
Frsix	The number of addons purchased in the last 6 months	Numeric	Numeric
Frthree	The number of addons purchased in the last 3 months	Numeric	Numeric
Abatwelve	Average allowance of purchased addons in the last 12 months	Numeric	Numeric
Abanine	Average allowance of purchased addons in the last 9 months	Numeric	Numeric
Abasix	Average allowance of purchased addons in the last 6 months	Numeric	Numeric
Abathree	Average allowance of purchased addons in the last 3 months	Numeric	Numeric
Aptwelve	Average paid price for bought addons in the last 12 months	Numeric	Numeric
Apnine	Average paid price for bought addons in the last 9 months	Numeric	Numeric
Apsix	Average paid price for bought addons in the last 6 months	Numeric	Numeric
Apthree	Average paid price for bought addons in the last 3 months	Numeric	Numeric

4.3 Data Pre-processing

The real-world data is usually not perfect and includes a lot of errors such as missing values and outliers. In data preprocessing, it is prevalent to apply methods to finalize dataset by removing noisy data or imputing the missing data (Garcia et. al, 2016). The following sections will describe how missing values are treated, class imbalance in the dataset is handled and relevant variables are selected.

4.3.1 Missing values

Few of the variables in the dataset include missing data. Before applying any imputation method, the variables that include missing data is checked for skewness. The number of missing values with features and imputation methods are listed below in Table 4.3.

Table 4.3 : Missing values and imputation method

Variable	Number of missing values	Proportion of dataset	Imputation method
Avgdatausage	509	2,37%	Imputed with the median
Dataallowance	1	0,00%	Removed from the data
Pricepergb	176	0,82%	Imputed with the median

Both of the variables that kept in the data set (average data usage and price per gb) are right-skewed, so the missing values are replaced with the median value of the customers average data usage and price per gb amount. When the data is skewed, it is better to consider using the median value for imputing the missing values because it mitigates the effect of the outliers. The single observation that has missing data allowance is removed from the dataset due to being negligible compared to the data size.

4.3.2 Handling imbalanced data

Imbalance of classes are common in real-world data that is used mostly to identify a rare but significant case. Many of the real-world cases consist of highly imbalanced data such as identifying fraudulent calls, detecting unreliable customers, text classification and so on (Kotsiantis et al., 2006). It has gained more attention in the last several years and many solutions are suggested for class-imbalance problem. In this section, the data level solutions are explained and the methodology for handling imbalance in the data for this study is demonstrated.

There are various ways of resampling to handle class imbalances such as undersampling, oversampling and combination of these techniques. Undersampling, is a method to balance class distribution by randomly eliminating majority class observations. It is a reasonable method to balance the dataset in order to overcome the idiosyncrasies of machine learning algorithm but also it can remove conceivably valuable information that can be significant for class assignment (Kotsiantis et al., 2006). Oversampling, is another method to handle class imbalances by randomly duplicating minority class observations. Although it preserves the original data, it is argued that it may lead to overfitting due to exact replication of minority class observations. Combination of over and under-sampling methods are also used as a hybrid method to handle class imbalances. There isn't any clear evidence in literature that determines which method is the best. However, undersampling and oversampling approaches continue to be popular due to their practical implementation (Yap et al. 2013).

The dataset used in this study is imbalanced due to the distribution of purchasers and non-purchasers. Figure 4.1 demonstrates that the class distribution is clearly skewed into one particular class, non-purchasers who hold 97% of the data as a majority class whereas purchasers is the minority class holding only 3% of the data.

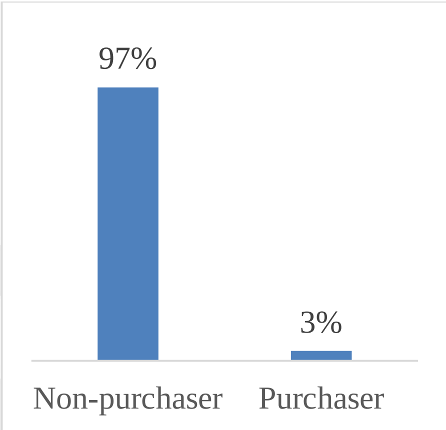

Figure 4.1 : Class distribution of the dataset

Since it is aimed to identify purchasers, with these class distributions in the data the predicted models' accuracy of predicting non-purchasers will be higher compared to the purchasers. Therefore, resampling techniques are evaluated to be used before building machine learning models. Oversampling, undersampling and combination of these methods are applied to a training set which holds 80% of the original dataset and the random forest performances are compared. The confusion matrix result of each method is presented in Appendix A. Oversampling method performed better in terms of accuracy and precision scores which represents more effective result of predicting purchasers. Hence, the data that is divided for training machine learning models is balanced by random over-sampling minority class examples with ROSE package of the R programming software (Rbloggers, n.d). Figure 4.2 shows the distribution of the classes after oversampling. The non-purchasers and purchasers are balanced with holding 50.2% and 49.8% of the training data respectively.

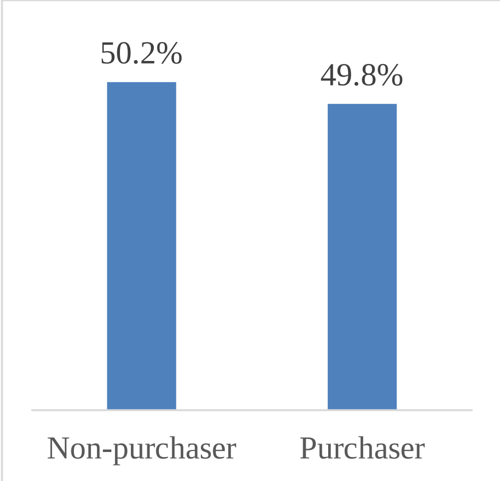

Figure 4.2 : Class distribution of the over-sampled training dataset

4.3.3 Variable selection

Variable selection is crucial to solving real-world challenges since we do not completely understand the underlying data model. While weak, redundant or irrelevant predictors may lead to overfitting and limit the model's interpretability, additional predictors may add fresh data to the prediction (Fang, 2018). Hence, selecting variables is a critical step in data pre-processing.

In statistical machine learning models, the relationship between the response variable and the predictors are illustrated with the function below where ε represents the error:

$$Y = f(X) + \varepsilon \qquad (4.1)$$

The statistical model struggles to find the underlying function and becomes overfitted if there are too many predictors in relation to the size of the training data due to the *curse of dimensionality* (Bellman, 1061). The accuracy of the model can be increased by including factors that are actually related to the response. However, in practice, this is typically not the case since most of the predictors are noisy and partially providing explanation for the response (Börthas and Sjölander, 2020). Hence, including more noisy variables in the model cause poor performance on the unobserved test data. Although some statistical learning models such as decision-tree based models do not need variable selection since it is already a part of the modeling process; the methods used in uplift modeling need variable selection due to relatively small uplift amount compared to outcomes risking overfitting to the data (Radcliffe and Surry, 2011).

In this project for the variable selection and dimension reduction several techniques offered in literature is applied and explained in the following subsections.

4.3.3.1 Multicollinearity

Multicollinearity is a common issue in statistical modeling and machine learning where two or more independent variables in a dataset are highly correlated with each other. This problem could cause to biased and unstable estimates of the model coefficients, hence can lead to inaccuracy of the model predictions. Multicollinearity can be detected and handled using various methods that are commonly used in literature such as correlation matrix, heatmap and variance inflation factor scores.

Heatmap

Heatmap is a powerful visual tool used in variable selection to identify and avoid multicollinearity. By using a heatmap, the correlation between variables can be easily identified and any strong relationships that could lead to multicollinearity issues can be determined. This visual helps to identify which variables are highly correlated with each other and allows practitioners to make informed decisions on which variables to include or exclude in the analysis. Each cell in the matrix represents the strength of the correlation between two variables. The color gradient indicates the level of correlation, with darker shades representing stronger correlations.

Similar to the heatmap, a correlogram is a useful tool for visualizing the correlation structure of a dataset. In a correlogram, variables are represented as rows and columns in a matrix, and the cells of the matrix are color-coded to represent the strength and direction of the correlations between variables.

In order to check the collinearity between predictors the correlation matrix is calculated and correlation is plotted using R programming software. Both correlogram and heatmap is visualized with R and shown in Figure 4.3 and Figure 4.4 below respectively.

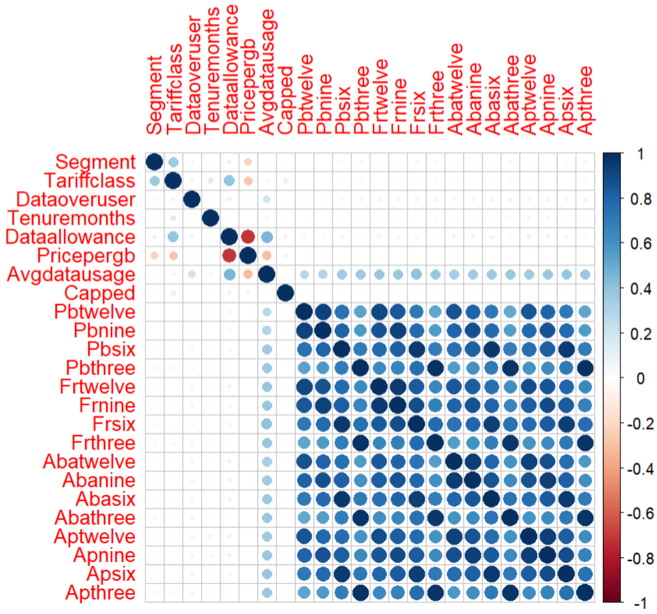

Figure 4.3 : The correlogram of variables

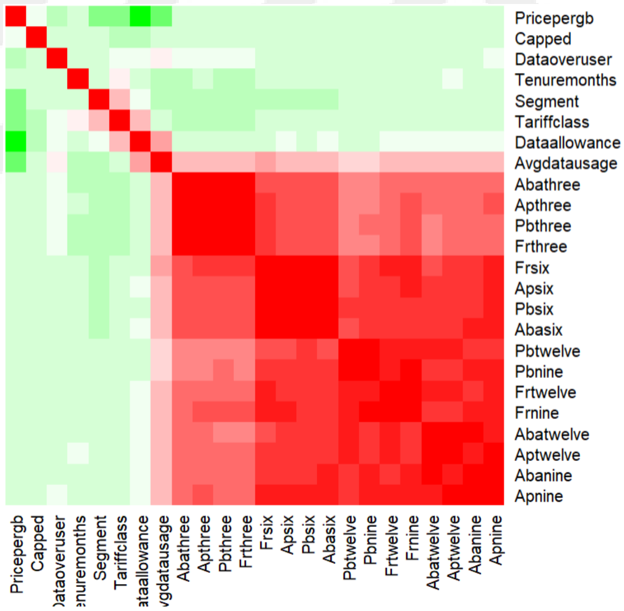

Figure 4.4 : The heatmap of the variables

The collinearity among variables can be easily identified from both of the figures. In Figure 4.3 the positive correlations are displayed in blue scale whereas the negative correlations are shown in red scale. Further, it can be seen from the Figure 4.4 right-bottom variables show high collinearity as displayed in red scale. As expected, the variables in the right-bottom that represent similar information show greater collinearity, hence need to be removed from the dataset before training models. To further decide which variables to remove from the dataset, variance inflation factor scores and net information value scores are also calculated in the following sections.

Variance inflation factor

As another multicollinearity measure, variance inflation factor (VIF) quantifies the severity of multicollinearity in a set of variables. It assesses how much the variance of an estimated regression coefficient increases if the predictors are correlated. VIF is calculated for each predictor in a regression model, and a high VIF value (generally above 10) indicates that the associated predictor may be problematic due to multicollinearity. To calculate the VIF, equation 4.2 is used.

$$VIF_1 = \frac{1}{1-R^2} \qquad (4.2)$$

R^2 represents the coefficient of determination from the linear regression model that has X_1 as dependent variable and X_2 and X_3 as independent variables in the linear regression model shown in equation 4.3.

$$X_1 = \beta_0 + \beta_1 x\, X_2 + \beta_2 x X_3 + \epsilon \qquad (4.3)$$

When R^2 is close to *1* (i.e. X_1 can be highly predicted by X_2 and X_3) the VIF will be very large. Whereas when R^2 is close to *0* (i.e. X_2 and X_3 are not related to X_1): the VIF will be close to *1*. Hence, the range of VIF is between 1 and infinity.

Generally, a VIF score above 10 is interpreted as indicating a level of multicollinearity among model predictors that is not tolerable, suggesting that the predictors are highly correlated with each other. However, according to James et al. (2013), a VIF score above 5 can already be indicative of high correlation among predictors. This suggests a more conservative threshold for identifying problematic levels of multicollinearity (James et al., 2013).

The VIF scores are calculated for variables with R programming language package "*vif*" (Fox and Nilsson, n.d). Scores are listed in Table 4.4 below.

Table 4.4 : VIF scores of variables

Variable	VIF Score
Apnine	85,7206
Abanine	71,8412
Apsix	50,4136
Aptwelve	48,6489
Abatwelve	43,3945
Abasix	38,8277
Frnine	32,1943
Apthree	31,7596
Pbnine	28,5679
Pbtwelve	22,6911
Abathree	20,7242
Frtwelve	17,0410
Frsix	16,2874
Pbsix	12,1292
Pbthree	7,1904
Frthree	4,9184
Avgdatausage	1,6340
Dataallowance	1,5806
Tariffclass	1,3416
Pricepergb	1,2328
Segment	1,1793

Table 4.4 (continued) : VIF scores of variables

Tenuremonths	1,0797
Dataoveruser	1,0455
Capped	1,0239

According to the heatmap and the variance inflation factor scores, the highly correlated variables with VIF scores above 5 are removed from the dataset. The resulting variable list is shown below in Table 4.5.

Table 4.5 : VIF scores of remaining variables

Variable	VIF Score
Frthree	1,4735
Avgdatausage	1,7089
Dataallowance	1,5308
Tariffclass	1,3926
Pricepergb	1,3289
Segment	1,1704
Tenuremonths	1,0237
Dataoveruser	1,0828
Capped	1,0157

After variable selection the correlogram and heatmap of remaining variables are shown in Figure 4.5 and 4.6., respectively

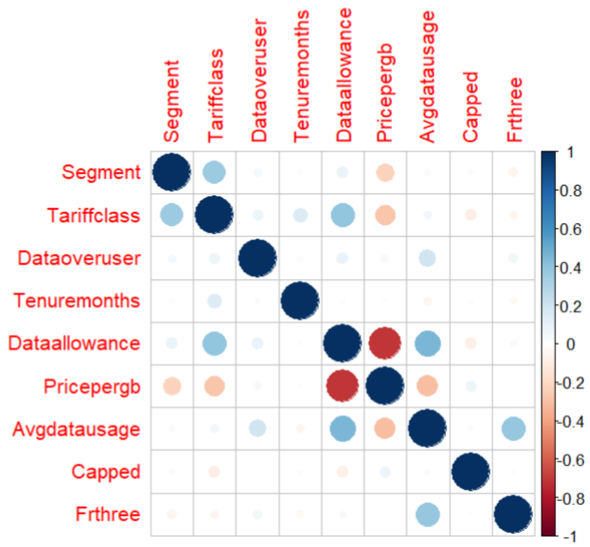

Figure 4.5 : The correlogram of final variables

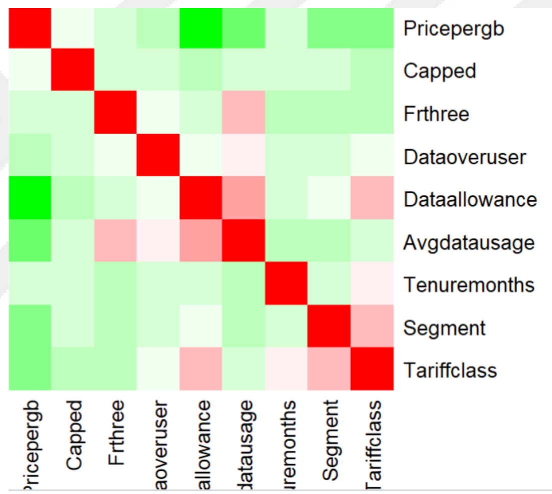

Figure 4.6 : The heatmap of final variables

5. FINDINGS

After data is preprocessed as described in the section above, each model is trained on training datasets and validated on the test sets. Next, the prediction performance of each model is assessed by ROC and Qini curves. The performance of the three alternative uplift modeling methods are visualized and a summary of all methods are demonstrated.

As it is shown in previous studies in literature that no method is consistently outperforming the other, so the aim is to try and test various algorithms to determine which is best for the particular problem in this thesis (Guelman et al, 2012; Jaskowski and Jaroszewicz, 2012). In example, modeling uplift directly with random forests can outperform class variable transformation in one particular data set, but the opposite can occur on another dataset. Hence, various uplift methodologies are applied in this study as explained in section 3 and evaluated in order to determine which performs best for the relevant problem represented with the dataset.

5.1 Two Model Approach

Two model approach in uplift modeling is based on the subtraction of two models as mentioned in detail in section 3.1.1. After preprocessing steps, each training set in treatment and control groups are fitted logistic regression. The summary of two independent models (treatment model and control model) are demonstrated in Appendix B. The confusion matrix results of each model with performance metrics are shown and compared in Table 5.1 below:

Table 5.1 : Confusion matrix results of treatment and control models

	Treatment model	Control model
Accuracy	0,6912	0,7082
Sensitivity	0,6976	0,7010
Specificity	0,6145	0,7142
Precision	0,9562	0,6741

The results show that the treatment model performs well on measuring the probability of purchasers to be actually the customers who make purchase than the control group when the precision values are compared. The ROC curve of both models are shown together in Figure 5.1. The AUC is 0,550 for treatment model and 0,707 for control model.

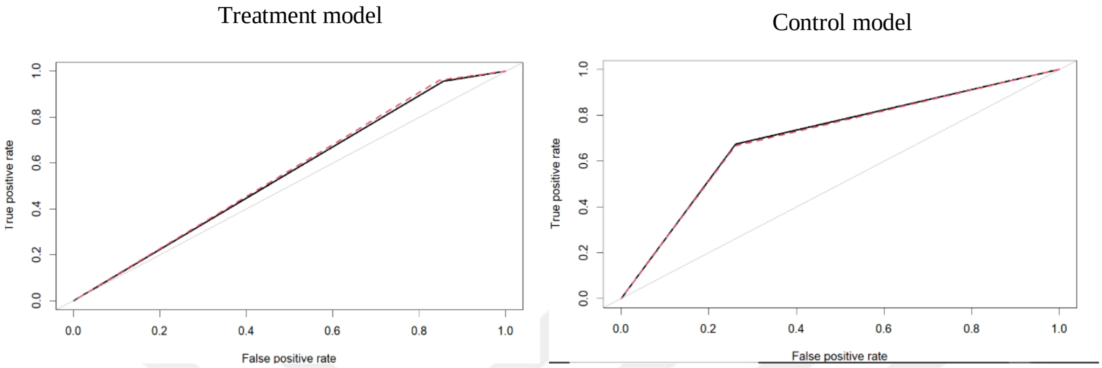

Figure 5.1 : ROC curve for treatment model(left) and control model(right)

After each model is applied to all of the observations and two estimations are performed; one with the treatment model and the other with the control model. In order to estimate the uplift, the control model score is subtracted from the treatment model score as formulated in equation (3.1). The resulting model's confusion matrix results is shown in Table 5.2 and the ROC curve is visualized in Figure 5.2.

Table 5.2 : Confusion matrix results of the subtracted model

	Subtraction of two models
Accuracy	0,4853
Sensitivity	-
Specificity	0,4918
Precision	-

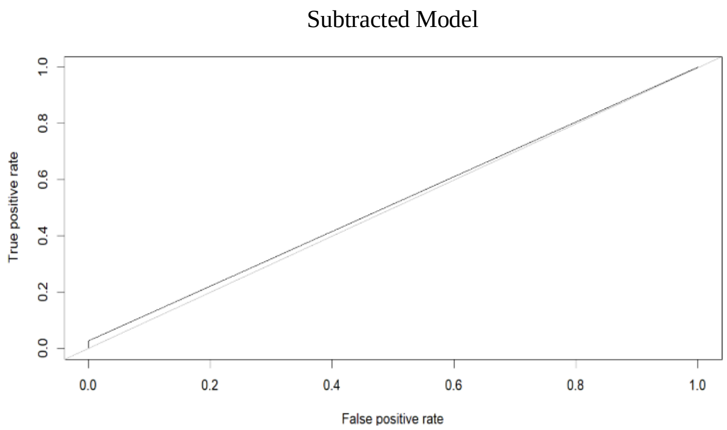

Figure 5.2 : ROC curve for the subtracted model

The resulting qini curve which plots the incremental purchase according to groups is computed with equation 3.13 and shown in Figure 5.3. The red solid line in the Figure 5.3 represents a random classifier, the blue dotted line is the subtracted model's qini curve and the black point demonstrates the greatest value for the model's curve. The qini value which is computed with the area between the model's qini curve and the random model curve is calculated as -0,0065.

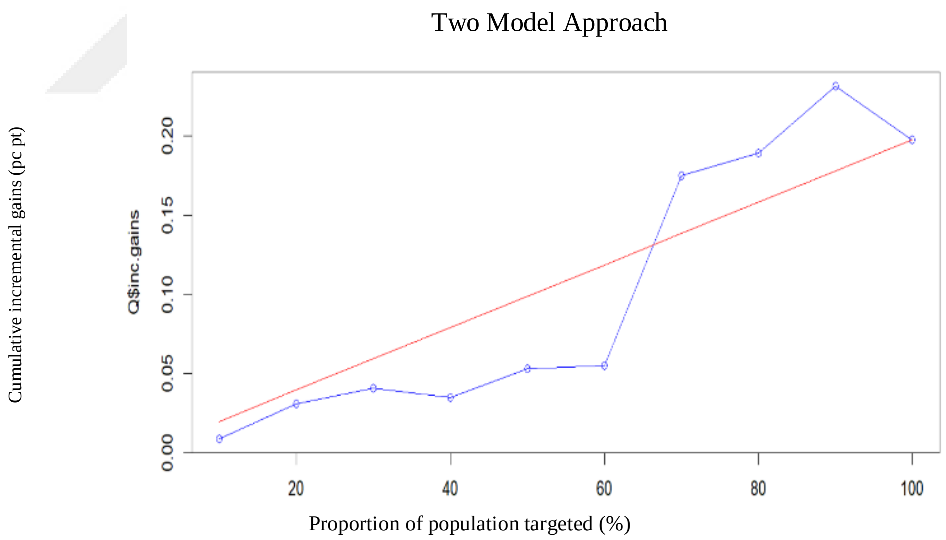

Figure 5.3 : Two model approach qini curve

The results demonstrate that the subtraction of the two models fails to find an uplift. Negative qini value suggests that doing a certain action has worse consequences than doing nothing (Guelman, 2014). The qini curve shows that cumulative incremental gains increases while the targeted population proportion increases. However, the maximum uplift is reached when almost entire population is targeted.

This result is consistent with the literature as it is highlighted in Radcliffe and Simpson (2007) and Radcliffe and Surry (2011) that two model approach tends to fail specifically when the uplift is not certainly large. Even if the two models were excellent in terms of their prediction performance, it is not guaranteed that subtraction of these two good models will result a good uplift model. According to Radcliffe and Surry (2011) there are four reasons that lead two model approach to fail in real-world cases. Firstly, since this technique builds two separate models that are built on the same premise and use the same model form with candidate predictors; nothing in the fitting process is particularly pursuing to fit the difference in behavior between the two groups. Second, for uplift modeling application, the size of the uplift is commonly small compared with the size of main impact which leads the modelling process to focus on the main effect. Further, building a model commonly starts with variable selection, therefore, theoretically it is possible that the most predictive variables for modelling the non-uplift outcome could be disparate than the uplift model's predictive variables. Hence, this could cause the two model approach to fail capturing important predictors for uplift. Lastly, fitting two models separately could cause fitting errors in the two models to be blended unfavorably and decrease the quality of fit which may be avoided by applying a direct uplift modeling approach (Radcliffe and Surry, 2011).

Although two model approach is worth trying due to its simplicity, it is usually a not accurate solution to capture the uplift in real-world cases as the one presented in this study.

5.2 Class Variable Transformation

In Class Variable Transformation, the outcome feature is transformed according to Equation (3.2) and the training data set is fit with logistic regression using the new transformed feature z as response variable. The result of the trained model is applied to test set and the scores are predicted. The resulting roc curve is presented in the Figure 5.4 as the "black" curve along with the direct uplift modeling resulting roc curves.

The area under the curve (AUC) is 0.64, demonstrating that the model performs better than a random classifier. However, the uplift performance could not be assessed solely according to roc curve as described in Section 3.3. The qini curve of the model is plotted according to incremental gains per segment in Figure 5.5. The qini value is computed as 0.2094.

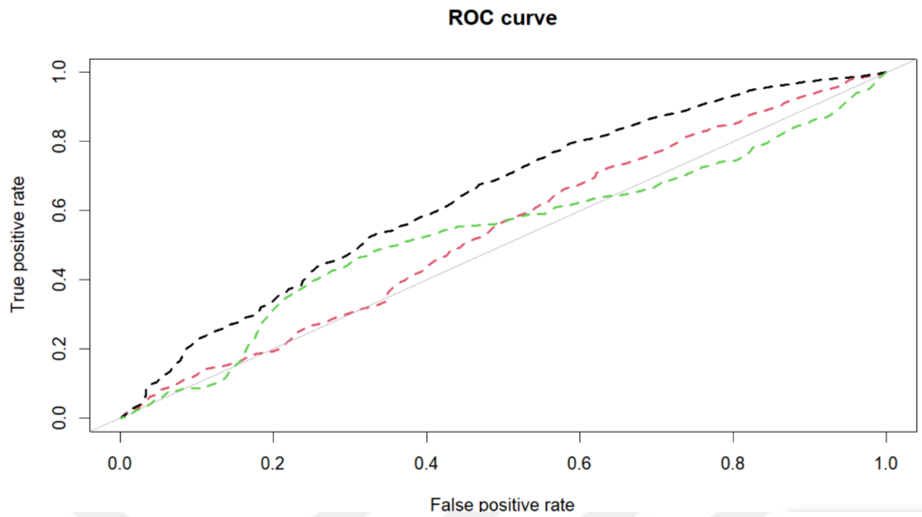

Figure 5.4 : Roc curves for CVT (black) and Direct uplift models with ED (red) & KL (green)

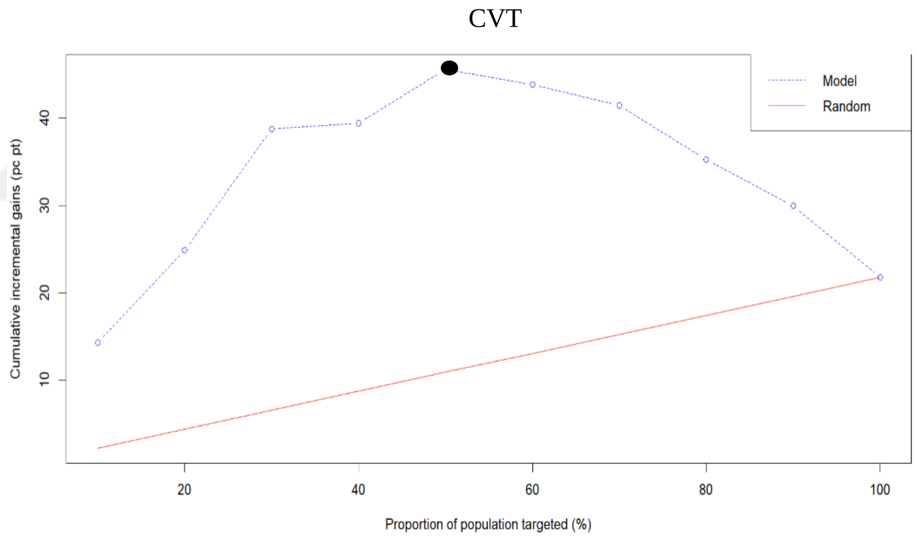

Figure 5.5 : Class variable transformation method qini curve

Clearly it can be seen from the plot that, qini curve is increasing as the proportion of the sample that is given treatment increase, then flattens out slightly and decreases by capturing the negative effect in the last segments. This is a definite sign that the uplift model captures the treatment's negative impact on individuals which are defined in the section 2.3 as "Do-not-disturbs". CVT method used uplift model is successful to identify more purchasers earlier in the sample, that the black dot demonstrates the highest point in the uplift curve that gives the optimum proportion of the treatment group that should be given treatment. It is seen that the optimum proportion is at the half of the treatment population, indicating targeting half of the treatment group would result greater than the uplift that could be achieved by targeting the entire group.

In previous studies in literature, there have been cases shown that the CVT outperforms two model approach as being used for uplift modeling purposes without being generalizable for all data sets. Kane et al. (2014) compared two model approach with two class variable transformation approach on three cases from financial services, online merchandise and retail business and showed that CVT outperformed the other. On the contrary, Devriendt et al. (2018) applied all three uplift modeling methods on four data sets and CVT showed poor performance on three of them but gave a good result in only one of the case. Comparison of uplift modeling methods are discussed further in section 5.4.

5.3 Modeling Uplift Directly

In modeling uplift directly with random forest according to Algorithm 2 explained in Table 3.2 in section 3.1.3, two distributional divergence based on Kullback-Leibler and Euclidean distance is applied with equation 3.3 and equation 3.4 respectively, while fitting the model to the training data set. The number of trees in each forest has been selected 1000 and nodes with less than 50 observations are not been subject for splitting. The R package called “uplift” that is provided by Guelman (2014) is used for building models and computing performance metrics.

The resulting models are applied to test data set and the ROC curves are illustrated above in Figure 5.4 with “red” curve for ED and “green” curve for KL. The area under the curve is 0.539 for ED and 0.534 for KL. The qini curves of the two models are plotted according to incremental gains per segment in Figure 5.6.

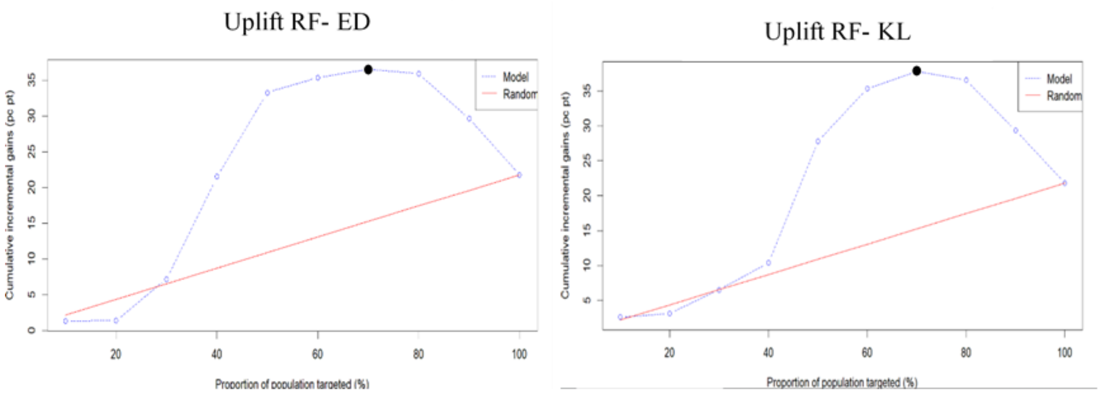

Figure 5.6 : Direct uplift modeling with RF qini curves with ED (left) and KL (right)

The qini score for the model with ED is 0,1045 and for the model with KL is 0,0913. The tree-based direct uplift modeling methods result similar qini scores and curves for both distributional divergence methods. It is noteworthy that the qini value is positive signed and the qini curve shows that cumulative incremental gains increases while the targeted population proportion increases. However, as it is also resulted in CVT approach, the tree-based direct uplift models incremental gain declines in the last segments of the group by managing to capture the "Do-not-disturbs". The black dots in Figure 5.6, indicates that the maximum uplift could be achieved by targeting the 70% of the treatment group.

The results show that direct uplift modeling manages to capture uplift in the data. This is agreeing with the literature which posit several cases that applied tree-based direct uplift modeling methods. Guelman et al. (2015), applied uplift random forests to an insurance campaign for customer retention and demonstrated that uplift model could be used to target the customers who are most inclined to respond favorably to an activity for retention. Rzepakowski and Jaroszewicz (2012), compared direct uplift modeling with traditional response models and other uplift modeling techniques on a real direct marketing campaign data. The results demonstrate that tree-based method outperforms other uplift modeling methods. Olaya et al. (2020), applied direct uplift modeling framework for student retention problem and showed that identifying students who are more likely to be retained is possible. Further, Devriendt et al. (2021) found similar results on some data sets that direct uplift approaches perform better but on some data sets other approaches provided improved results. Gubela and Lessman (2021) also evaluated different uplift modeling approaches on several direct marketing campaigns including direct uplift modeling methods and showed that no single uplift modeling method consistently outperformed the others. A study from Röbler et al. (2021) combined all three uplift modeling methods as two model approach, class variable transformation and direct uplift approaches, into a weighted ensemble and applied on nine real-world data sets. Their study showed that this ensemble approach could be used to overcome volatility issues in uplift prediction. Moreover, a recent study of Röbler et al. (2022) compared several tree-based uplift modeling methods on three real-world data sets based on performance metrics such as qini curve and qini value and proposed a new tree-based uplift method.

Direct uplift modeling approaches are increasingly gaining attention in literature being applied with other uplift modeling methods as it is applied in this study, or investigated with various classification and prediction methods to achieve an uplift model with excellent performance. For the data set studied in this thesis, the outcome is similar for both of the direct uplift modeling approaches with two different splitting criterion. However, the overall performance of the tree-based uplift models is not greater than CVT, which is compared further in detail in the next section.

5.4 Comparison of models

There is not a single uplift modeling method that works best for all cases (Röbler et al., 2021). The three uplift modeling techniques is applied to the same data set and compared according to the performance metrics of uplift models such as qini curve and qini values. In Table 5.3 the trained models and the qini score for each model is listed.

Table 5.3 : Models' Qini scores

Method	Model	Qini Score
Two Model Approach	Logistic Regression	-0,0065
Class Variable Transformation	Logistic Regression	0,2094
Modeling Uplift Directly	Random Forest Euclidean Distance	0,1045
Modeling Uplift Directly	Random Forest Kullback-Leibler Distance	0,0913

According to the results, all of the uplift models except two model approach has positive qini value, showing that the CVT and direct uplift models predict a positive gain of the treatment. It can be seen from the value of the qini scores that the CVT method outperforms the other uplift modeling methods with the highest qini score. Modeling uplift directly with Euclidean distance is the second best model according to the qini scores. Two model approach is the worst performing model with a negative qini value, indicating that giving a treatment will have worse results than doing nothing for most of the population.

For interpretation purposes, the qini curves plotted separately in previous sections are combined in Figure 5.7. As the qini curves compared, it is clearly seen that all models except two model approach return a qini curve above the random model and increase as the proportion of the sample that is given treatment increase, then flatten out slightly and decrease by capturing the negative effect in the last segments. This shows that the models capture the treatment's negative effect on individuals which are defined in the first part of this thesis as "Do-not-disturbs". It can be clearly seen that the Class Variable Transformation method used uplift model outperforms the others as it manages to identify more purchasers earlier in the sample. Also, it is noteworthy that all uplift models manage to identify "Do-not-disturbs", the observations that their possibility to purchase gets lower when treated.

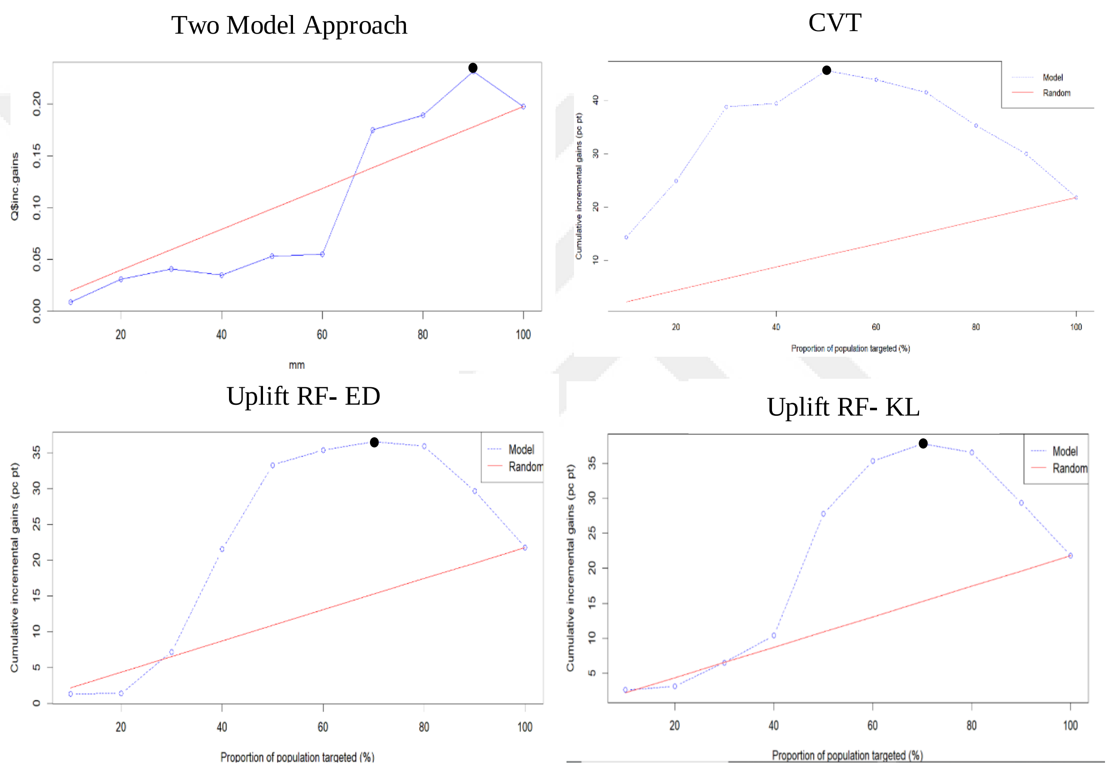

Figure 5.7 : Qini curves of all uplift modeling methods

The total amount of purchasers can be substantially increased by the uplift modeling techniques without giving treatment to the the entire treatment groups. The highest point in the uplift curve results the optimal uplift and gives the optimum proportion of the treatment group that should be given treatment. The Table 5.4 below shows for each model the portion of group that should be treated to reach the maximum incremental gain.

Table 5.4 : Treatment group amount for maximum incremental gain

Method	Model	Percentage of Treated	Max Incremental Gains
Two Model Approach	Logistic Regression	90%	0,2318
Class Variable Transformation	Logistic Regression	50%	0,4559
Modeling Uplif Directly	Random Forest Euclidean Distance	70%	0,3651
Modeling Uplif Directly	Random Forest Kullback-Leibler Distance	70%	0,3779

For the best performing method CVT, it is seen that the optimum proportion is at 50% of the treatment group that would result approximately 46% uplift which is dramatically greater than the 18,7% uplift that would be achieved if the entire treatment group was given treatment. Modeling uplift directly with random forests suggest to give treatment to 70% of the group by yielding approximately 37% uplift which also give a better result than treating the entire group. Lastly, even in two model approach, a 23% uplift which is slightly more than giving treatment to the entire group, could be achieved by only targeting 90% of the treatment group.

In this study the CVT method gave the best results compared to the other uplift modeling techniques. However, it is shown in previous studies that no method significantly outperforms the others on all datasets. For example, Guelman et al. (2014) applied four different uplift modeling techniques on a Canadian insurance company retention campaign and compared the results according to the incremental gains curve. The results showed that direct uplift modeling with random forests performed best on that case especially for low-target volumes by identifying 30% of the customers who the retention campaign would be most effective. Rzepakowski and Jaroszewicz (2012) compared tree-based approach for a single treatment and multiple treatments for several clinical trials data sets. The findings revealed that it is better to apply two treatments in different order as suggested by the model than using a single averaged treatment. Devriendt et al. (2018) studied ten distinct techniques on four real-world cases and put forward that no approach consistently outperformed the others highlighting the instability of uplift models and lack of an application-oriented approach to evaluate uplift models. Karlsson (2019) applied class variable transformation and direct uplift modeling approaches in a study for an insurance marketing campaign. According to qini values, it is demonstrated that random uplift forest with Euclidean distance splitting criterion is the winning model with highest qini value. Further, Röbler et al. (2021) compared two model approach, CVT and direct modeling approaches to show that none of the models outperformed the others on all data sets. Hence, a combination of three methods is proposed to achieve best results.

In this study, the three uplift modeling approaches applied to model the uplift in the real-world direct marketing campaign data which resulted as CVT outperforming the other uplift modeling methods.

6. CONCLUSION

Businesses today spend an extensive number of resources on a variety of marketing efforts, from digital approaches like email marketing and web advertisements to more traditional approaches like print advertisements. It is challenging to assess the real impact of these efforts in the context of inevitable changes in consumer behavior. While traditional predictive analytics has shown to be effective at predicting consumer behavior and notifying decision-makers about upcoming events, it is not as effective at determining the actual impact of marketing efforts on consumer behavior. The main focus of this study is the prescriptive analytics, emphasizing on uplift modeling which identifies the incremental benefits of marketing initiatives and results in more effective resource allocation and enhanced return on investment. In order to develop a more accurate model for measuring the impact of marketing campaigns on consumer behavior, the limitations of conventional predictive analytics in marketing are explored; prescriptive analytics, specifically uplift modeling, in identifying the incremental effects of marketing actions is investigated and by using a real-world B2B Turkish GSM sector data, a framework for implementing uplift modeling in a direct marketing campaign is developed. Three uplift modeling methodologies are applied to the data and shown that a B2B marketing campaign can be optimized by predicting the incremental impact more precisely with uplift models than the conventional prediction models. The findings revealed the effectiveness of uplift modeling in optimizing resource allocation by enabling to target customers whose behavior is most likely to be positively influenced by marketing initiatives. Among the uplift models, the model with the class variable transformation method managed to capture 46% of uplift while targeting only the half of the campaign audience. As agreeing with the previous studies in the uplift modeling literature, this result demonstrates the success of uplift models' prescriptive performance of predicting the truly responsive customers for a direct marketing campaign.

This study suggests various implications, both theoretical and practical. It contributes to filling the notable gap in uplift modeling literature, offering actionable insights and managerial recommendations through a real-world example. Subsequent sections will provide a detailed explanation of these contributions.

6.1 Theoretical Implications

This study has several significant research implications. Firstly, it contributes to the growing body of literature in marketing analytics by strengthening the theoretical foundation with the integration of the various aspects of marketing theory and related work. By incorporating insights from the exploration of communicating with consumers, the evolution of direct marketing and the dynamics of marketing communications in B2B setting, this study offers a comprehensive understanding of complexities involved in marketing analytics and underscores the need for an approach to marketing strategy development. The emerging topics of marketing analytics that stated recently in literature include optimization of marketing budgets, effective resource allocation, use of machine learning to improve effectiveness of marketing campaigns and improve customer experiences (Petrescu and Krishen, 2023). With the integration of predictive and prescriptive analytics in marketing strategies, this study highlights the importance of distinguishing between forecasting customer behavior and measuring the impact of marketing efforts, deepening our understanding of marketing analytics principles.

Although there is an extensive academic literature that have touched prescriptive analytics in marketing, this study addresses a notable gap in the existing body of knowledge on uplift modeling in marketing. While previous research predominantly focuses on uplift modeling's application in business-to-consumer (B2C) contexts as shown in Table 2.4, this study extends the understanding of uplift modeling to the domain of business-to-business (B2B) marketing. By exploring the effectiveness of uplift modeling in B2B contexts, the study advances theoretical insights into the application of prescriptive analytics in direct marketing setting. Further, it contributes to the research field of uplift modeling by performing a thorough benchmark of uplift modeling techniques on a real-world B2B direct marketing campaign, demonstrating how prescriptive analytics enable targeting the most responsive customer segment, hence optimizing the marketing campaign management. Among the commonly studied uplift modeling techniques as two model approach, class variable transformation and direct uplift modeling, the results are compared according to their performance of capturing the uplift in the marketing campaign data set. From the comparison of uplift modeling methods, the best performing model which utilized class variable transformation technique provides empirical evidence that this simple approach can outperform random targeting in the analyzed B2B marketing campaign. This result is agreeing with few of the prior studies and contributing to the practical knowledge on uplift modeling even though in various studies it has shown that no uplift modeling method outperforms others on all examples (Devriendt et al., 2018; Olaya et al. 2020). Hence, this study manages to develop a framework for implementing uplift modeling in real-world marketing scenario emphasizing on optimizing resource allocation by targeting customers whose behavior is most likely to be influenced by specific marketing actions.

In summary, this thesis has made significant contribution as being the first to introduce uplift modeling in Turkish literature and being one of the few studies to apply uplift modeling in B2B context in the world-wide academic literature. It has been demonstrated that conventional response models are inadequate to distinguish which customers are positively impacted by a marketing treatment whereas uplift models successfully identify which customers will make a purchase due to being influenced by the marketing treatment. It is also important to show that targeting more customers in a direct marketing campaign does not always give best or efficient results, and higher uplifts can be achieved with various uplift modeling techniques applied with machine learning algorithms. Hence, this study extends related studies in the domain by improving marketing campaigns with utilizing machine learning applications in prescriptive analytics to target truly responsive customers in direct marketing efforts.

6.2 Managerial Implications

The insights derived from this research offer key managerial implications for marketing practitioners in the field. In the current competitive landscape, marketing analytics has become an essential part of data-driven decision making, enabling businesses to understand large volumes of data and leverage them for marketing strategy. It is becoming more and more expected of marketing managers to rely on marketing data while making decisions and to employ sophisticated data analysis techniques to extract meaningful information from data for use in marketing models (Petrescu and Krishen, 2023). In particular, successful marketing activites are predicated on the utilization of appropriate models for data estimation and the accurate predictions derived from the estimated model outcomes (Paetz et al., 2022). The primary objective of marketing practitioners is to maximize financial benefit of the campaign by achieving high prediction rates for the positive responders to the campaign. Additionally, they strive to eliminate negative responders from the target list in order to save corporate resources by achieving acceptable prediction rates for these negative responder groups (Marinakos and Daskalaki, 2017). Marketing managers can facilitate decision making by adopting uplift modeling to distinguish customers who likely to respond positively and those who would make a purchase regardless of the marketing action. Uplift models may enable marketing practitioners to tailor their approaches efficiently and as a result better allocation of resources towards individuals most likely to be influenced by marketing treatments, thereby optimizing campaign outcomes. Rather than relying on conventional response models, which may inadvertently waste company resources on customers who would not make a purchase or worse that could be influenced negatively by the marketing effort and give up buying, uplift models can precisely identify customers who are truly responsive to the campaign. Hence, this targeted approach would ensure maximum impact with minimal expenditure, improving the overall efficiency of marketing operations.

Furthermore, uplift modeling enhances customer relationship management strategies. By focusing on customers who likely to benefit from the marketing treatments, businesses can provide targeted offers and personalized experiences. This engagement with customers who might have been overlooked by traditional targeting methods, may foster stronger customer relationships, leading to enhanced satisfaction and loyalty. This engagement could be easily reflected in long-term profitability and growth in the business side.

In addition, the uplift model reveals the relative importance of several variables in predicting customer behavior during the marketing campaign period. As shown in Table 6.1, the most significant predictor Tenuremonths, indicates that customers with longer tenure are more likely to respond positively to the campaign. This suggests that loyalty or familiarity with the service plays a crucial role in purchase decisions.

Avgdatausage is also highly influential, implying that customers who use more data are more receptive to the campaign. Pricepergb and Dataallowance further highlight the importance of cost and data plan features in shaping customer decisions. Tariffclass and Dataoveruser suggest that both the type of tariff plan and recent overusage behavior are relevant, although to a lesser extent. The segment variable indicates some variation in responses across different business segments, while Frthree has minimal impact, suggesting recent addon purchase history is less critical. Overall, these findings suggest that targeting long-term, high-data-usage customers with attractive pricing and data allowances will likely yield the best results for marketing campaigns. Marketing professionals should consider these key variables to tailor their strategies and maximize the campaign's effectiveness.

Table 6.1 : Relative importance of variables in uplift model

Variable	Relative importance (CVT uplift model)
Tenuremonths	42,08
Avgdatausage	21,13
Pricepergb	10,00
Dataallowance	9,47
Tariffclass	6,05
Dataoveruser	4,75
Segment	4,62
Frthree	1,90

This study study suggests that the evaluation of uplift models offers businesses a more precise tool for assessing and enhancing the effectiveness of their marketing strategies. Even with the simplest approach of uplift modeling, the incremental impact of marketing treatments can be measured and enable marketing managers to measure the effectiveness of their campaigns more accurately and refine their strategies based on empirical evidence. With this study, it is highly suggested for practitioners to invest in training and development initiatives for leveraging prescriptive analytics and uplift modeling to enhance understanding of machine learning applications in marketing analytics. It is important to employ advanced data analysis techniques, compare various methods and evaluate campaign outcomes for continuous improvement and better results over time for the company.

6.3 Limitations and Future Research

While this study provides valuable insights into the effectiveness of uplift modeling in direct marketing campaigns, it is essential to acknowledge certain limitations. Firstly, the marketing treatment in this study is defined as receiving a push notification from the company regarding the cross-sell marketing campaign. This definition of treatment presents a potential limitation, because the individuals who received the notification are presumed to be engaged with it and classified as the treatment group. However, in this treatment group there might be individuals who did not read and engage with the marketing notification, or simply were not aware of such marketing treatment. This raises the question whether the uplift model would yield better results if those individuals who did not engage with the treatment could be identified and excluded from the analysis. By restricting the treatment group to individuals only who actively engaged with the marketing notification, the models may more accurately identify responsive customers, thus simplifying the task of distinguishing between “persuadables” and “sleeping dogs”.

Secondly, the utilization of data from a real-world marketing campaign introduces limitations related to group sizes. Due to the distribution of purchasers as the minority class, group sizes are naturally imbalanced. In order to perform a successful uplift model and produce findings that can be interpretable, the sizes of treatment and control groups are significant. In this study, the control group size was sufficient to perform uplift modeling, since control group size was ten times larger than the treatment group size (Radcliffe and Surry, 2011). However, using a bigger control group may result in stronger uplift models.

Lastly, the generalizability of the study findings is constrained by the fact that the data evaluated is from a single country and sector. Variations in customer behavior across different markets and sectors targeted for direct marketing actions may lead to differing outcomes. Therefore, the results of this study may not be fully applicable to B2B direct marketing contexts in other regions or industries.

This study has demonstrated that the number of purchasers targeted with a direct marketing campaign could be significantly increased while the adverse impact on non-responsive individuals could be minimized. However, future research can further enhance the utilization of uplift models by exploring diverse objectives and scenarios.

Firstly, uplift models could be applied to simulate various objectives beyond purchase or retention scenarios. For example, future studies could investigate the use of uplift models to increase customer spending, boost the number of store or website visits, predict electoral voting behavior or improve retention efforts. By expanding the uplift modeling objectives, researchers could uncover new insights into consumer behavior and decision-making process.

Secondly, future research could explore the ways to increase company revenue by optimizing the cost-effectiveness of marketing treatments. This could involve identifying customers who would spend more as a result of marketing treatment, thereby offsetting the cost of treatment and maintaining or increasing overall earnings. In this study the target feature is binary and defined as “purchase”. In future studies, “spending” could be defined as a continuous target feature and provide more extended insights. Furthermore, an interesting extension to this work would be to investigate multi-treatment marketing campaigns. By offering varying discounted products to separate segments of customers, researchers can identify the most effective offer strategy for various target groups. This approach could help companies tailor their marketing efforts more precisely to different customer segments, ultimately improving campaign effectiveness and ROI.

Lastly, future studies could address feature selection in uplift modeling methodologies. By employing advanced classification methodologies and comparing them with existing approaches, researches can identify optimal feature sets for predicting uplift. This could lead to the development of more accurate and robust uplift models, enhancing their practical utility in real-world marketing scenarios. In conclusion, future research directions in uplift modeling should focus on expanding the scope of objectives, optimizing revenue generation strategies, exploring multi-treatment campaigns and improving methodological approaches. By addressing these areas, the understanding of uplift modeling and its application in marketing strategies could be advanced.

REFERENCES

Albert, J., Goldenberg, D. (2021). E-commerce promotions personalization via online multiple- choice knapsack with uplift modeling. *Association for Computing Machinery.* Retrieved June 3, 2022 from https://arxiv.org/pdf/2108.13298.pdf

American Marketing Association (n.d). *Definitions of Marketing.* Retrieved Jul 22, 2022 from https://www.ama.org/the-definition-of-marketing-what-is-marketing/.

Athey, S., Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. In: *Proceedings of the National Academy of Sciences, 113.27, 7353 – 7360.*

Barutcu, S. (2008). Consumers' attitudes towards mobile marketing and mobile commerce in consumer markets. *Ege Academic Review,* 8(1), 15 – 32.

Bellmann, R. E. (1961). *Adaptive Control Processes: A Guided Tour.* Princeton University Press, Princeton, NJ.

Bose, I., Chen, X. (2009). Quantitative models for direct marketing: A review from systems perspective. *European Journal of Operational Research, 195**,** 1-16.*

Börthas, L., Sjölander, J. K. (2020). Machine learning based prediction and classification for uplift modeling (Degree project). *KTH Royal Insitute of Technology School of Engineering Sciences.*

Cao, Y., Xu, C. and Hairong, G. (2017). Untangle customers' incrementality using uplift modeling with a case study on direct marketing. *MWSUG, Paper BF03.*

Chaffey, D. (2010). Applying organizational capability models to assess the maturity of digital-marketing governance. *Journal of Marketing Management, **26:3-4,** 187 – 196.*

Coe, J. M. (2004). The integration of direct marketing and field sales to form a new B2B sales coverage model. *Journal of Interactive Marketing, 18**,** 62 – 74.*

Cornelissen, J. (2004). *Corporate Communication: Theory and Practice,* Sage Publications, London.

De-Caigny, A., Coussement, K., Verbeke, W., Idbenjra, K. and Phan M., (2021). Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach. *Industrial Marketing Management, 99**,** 28-39.*

Deighton, J. and Kornfeld, L. (2009). Interactivity's unanticipated consequences for marketers and marketing, *Journal of Interactive Marketing, 23***,** *4 – 10.*

Devriendt, F., Moldovan, D. and Verbeke, W. (2018). A literature survey and experimental evaluation of the statof-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics, *Big Data, 6(1), 13 – 41.*

Devriendt, F., Berrevoets, J. and Verbeke, W. (2021). Why you should stop predicting customer churn and start using uplift models. *Information Sciences, 548***,** *497 – 515.*

Draleos, R. (2019). Measuring performance: the confusion matrix. Retrieved Dec 20, 2022 from https://glassboxmedicine.com/2019/02/17/measuring-performance-the-confusion-matrix/

Fang, X. (2018). *Uplift modeling for randomized experiments and observational studies*. (PhD Thesis). Massachusetts Institute of Technology, U.S.A.

Fox, J. and Nilsson, H. (nd). "vif{car}". R package version 3.1.1 Retrieved Dec 13, 2022 from https://cran.r-project.org/package=car

Garcia, S., Ramirez-Gallego, S., Luengo, J., Benitez, J. M. and Herrera F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1**,** 9.

Godin, S. (2008). Permission Marketing. *Seth's Blog*. Retrieved May 21, 2022, from https://seths.blog/2008/01/permission-mark/

Goldenberg, D., Albert, J., Bernardi, L. And Estevez, J. (2020). Free lunch! Retrospective uplift modeling for dynamix promotions recommendation within ROI contraints. In: *Fourteenth ACM Conference on Recommender Systems (RecSys'20), September 22-26, 2020, Virtual Event, Brazil.* ACM, New York, NY, USA.

Gubela, R. M. and Lessmann, S. (2021). Uplift modeling with value-driven evaluation metrics. *Decision Support System, 150, 113648.*

Guelman, L. (2014). "uplift: Uplift Modeling". R package version 0.3.5 Retrieved Nov 6, 2022 from https://cran.r-project.org/package=uplift

Guelman, L., Guillen, M. and Perez-Marin, A.M. (2012). Random forests for uplift modeling: An insurance customer retention case. In: *International Conference on Modeling and Simulation in Engineering, Economics and Management, 123 – 133.*

Guelman, L., Guillen, M. and Perez-Marin, A.M. (2015). Uplift random forests. *Cybernetics and Systems, 46:3-4,* 230 – 248.

Gutierrez, P. and Gerardy, J. (2016). Causal inference and uplift modeling a review of the literature. In: *JMLR: Workshop and Conference Proceedings 67:1-13.*

Hair Jr., J. F. and Sarstedt, M. (2021). Data, measurement, and causal inferences in machine learning: opportunities and challenges for marketing. *Journal of Marketing Theory and Practice, 29-1*, 65 - 77.

Hansotia, B. and Rukstales, B. (2002). Incremental value modeling. *Journal of Interactive Marketing, 16***,** 3.

Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations.* First Edition. CRC Press.

Hastie, T., Tibshirani, R. and Friedman, J. (2008). *The Elements of Statistical Learning Data Mining, Inference and Prediction.* Second Edition. Springer.

Hermann, Z., Horvath, H. and Lindner, A. (2022). Answering causal questions using observational data-achievements of the 2021 Nobel Laureates in economics. *Financial and Economic Review 21, 1, 141 – 163.*

Hollensen, S., **Raman, A.** (2014). Social media marketing: from 'bowling' to 'pinball'. *Journal of the Canadian Institute of Marketing, 10, 8-12.*

Huang, E.Y., Tsui, C. (2016). Assessing customer retention in B2C electronic commerce: An empirical study. *Journal of Marketing Analytics, 4, 172-185.*

Huotari, L., Ulkuniemi, P., Saraniemi, S. and Malaska, M. (2015). Analysis of content creation in social media by B2B companies. *Journal of Business and Industrial Marketing, 30(6),* 761-770.

Indeed Career Guide (n.d.). *A Definitive Guide to Mass Marketing (With Pros and Cons).* [online] Available at: https://www.indeed.com/career-advice/career-development/mass-marketing.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R.* Springer Science & Business Media.

Jaskowski, M. and Jaroszewicz, S. (2012). Uplift modeling for clinical trial data. *ICML 2012 Workshop on Clinical Data Analysis.*

Kane, K., Lo, V. S. and Zheng, J. (2014). Mining for the truly responsive customers and prospects using true-lift modeling: Comparison of new and existing methods, *Journal of Marketing Analytics, 82(4),* 218 – 238.

Kannan, P. K. and Li, H. A. (2017). Digital marketing: a framework, review and research agenda, *International Journal of Research in Marketing, 34, 22 – 45.*

Karjaluoto, H., Mustonen, N. and Ulkuniemi, P. (2015). The role of digital channels in industrial marketing communications. *Journal of Business and Industrial Marketing, 30(6), 703 – 710.*

Karlsson, H. (2019). *Uplift Modeling: Identifying optimal treatment group allocation and whom to contact to maximize return on investment.* (Master's Thesis). Linköping University, SWEDEN.

Kotsiantis, S., Kanellopoulos, D. and Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering, 30.*

Kim, K. H. and Moon, H. (2021). Innovative digital marketing management in B2B markets. *Industrial Marketing Management, 95, 1 – 4.*

Kumar, V. and Mittal, S. (2020). Mobile marketing campaigns: practices, challenges and opportunities. *International Journal of Business Innovation and Research, 21, 523–539.*

Lachner, M., Arnold, A.R. and Wangenheim, F. (2020). Mobile push notifications: a conceptualization from a direct marketing perspective. In: *Four Essays on Mobile Push Notifications*, Retrieved Aug 2, 2022 from https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/419316/1/DISS_final_Michael_Lachner_woCV.pdf

Larsen, K. (2009). Net lift models. In: *M2009- 12th Annual SAS Data Mining Conference, October 26-27, Las Vegas, NV.*

Lepenioti, K., Bousdekis, A., Apostolou, D. and Mentzas, G. (2020). Prescriptive analytics: literature review and research challenges. *International Journal of Information Management, 50, 57–70.*

Li, J., Zhang, W., Liu, L., Yu, K., Le, T.D and Liu, J. (2021). A general framework for causal classification. *International Journal of Data Science and Analytics, 11, 127 – 139.*

Liu, H., Gan, Y., Yang, J., Sighom, S., Wang, Y., Chen, Y. and Ye, F. (2012). Push the limit of WiFi based localization for smartphones. In: *18th Annual International Conference on Mobile Computing and Networking - Mobicom '12.*

Lo, V. S. (2002). The true-lift model: A novel data mining approach to response modeling in database marketing. *ACM SIGKDD Explorations Newsletter*, 4(2), 78 – 86.

Ludziejewski, J., Tomaszewska, P., Zalewska A. (2020). XAI Stories Case studies for explainable artificial intelligence. *Warsaw University and Warsaw University of Technology*. Retrieved May 28, 2022, from https://pbiecek.github.io/xai_stories/index.html

Marinakos, G. and Daskalaki, S. (2017). Imbalanced customer satisfaction for bank direct marketing. *Journal of Marketing Analytics, **5**, 14 – 30.*

Medina, I., and Correia, P. (2012). The Importance of Social Media for Commerce. A Case Study in Madeira (Portugal). *International Journal of Interactive Mobile Technologies*, 6, 37-42.

Munting, M. (2020). *An explorative study towards the feasibility of uplift modeling within a direct marketing setting and a web-based setting.* (Master's Thesis). University of Twente, NETHERLANDS.

Olaya, D., Vasquez, J., Maldonado, S., Miranda J. and Verbeke, W. (2020). Uplift modeling for preventing student dropout in higher education. *Decision Support Systems, 134, 113320.*

Paetz, F., Steiner, W.J. and Hruscha, H. (2022). Advanced data analysis techniques with marketing applications. *Journal of Business economics, 92, 557 – 561.*

Pandey, N., Nayal, P., Rathore, A.S. (2020). Digital marketing for B2B organizations: structured literature review and future research directions. *Journal of Business & Industrial Marketing, 35/7, 1191 – 1204.*

Petrescu, M. and Krishen, A.S. (2023). A decade of marketing analytics and more to come: JMA insights. *Journal of Marketing Analytics, 11, 117 – 129.*

Pok, W. (2021). How uplift modeling works, Retrieved Sep 26th from https://ambiata.com/blog/2020-07-07-uplift-modeling

Powell, G. (2008). *Marketing Calculator: Measuring and Managing Return on Marketing Investment.* Wiley & Sons (Asia).

Proença, H.M. and Moraes, F. (2023). Incremtnal profit per conversion: a response transformation for uplift modeling in e-commerce promotions. In *KDD'23, August 6-10.* Long Beach: KDD.

Radcliffe, N. J. (2007). Using control groups to target on predicted lift: building and assessing uplift models. *Direct Marketing Journal, 1, 14 – 21.*

Radcliffe, N. J. and Simpson, R. (2007). Identifying who can be saved and who will be driven away by retention activity. *Journal of Telecommunications Management, 1.2, 168 – 176.*

Radcliffe, N. J. and Surry, P. D. (1999). Differential response analysis: Modeling true response by isolating the effect of a single action. Credit Scoring and Credit Control VI.

Radcliffe, N. J. and Surry P. D. (2011). Real-world uplift modelling with significance-based uplift trees. Portrait Technical Report TR-2011-1, Stochastic Solutions.

Rbloggers (n.d). *Class imbalance-handling imbalance data in R.* Retrieved May 28, 2022 from https://www.r-bloggers.com/2021/05/class-imbalance-handling-imbalanced-data-in-r/

Reutterer, T., Mild, A., Natter, M. and Taudes, A. (2006). A dynamic segmentation approach for targeting and customizing direct marketing campaigns. *Journal of Interactive Marketing, 20, 43 – 57.*

Roberts, M. L. and Berger, P. D. (1999). *Direct Marketing Management.* Upper Saddle River, N.J., Prentice Hall.

Rosset, S., Neumann, E., Eick, U., Vatnik, N. and Idan, I. (2001). Evaluation of prediction models for marketing campaigns. In: *KDD'01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining.*

Röbler, J., Tilly, R. and Schoder, D. (2021). To treat, or not to treat: Reducing volatility in uplift modeling through weighted ensembles. In: *Proceedings of the 54th Hawaii International Conference on System Sciences.*

Rubin, D. (1974). Estimating causal effects of treatment in randomized and non-randomized studies. *Journal of Educational Psychology, 66(5), 688 – 701.*

Rudas, K. and Jaroszewicz, S. (2018). Linear regression for uplift modeling. *Data Mining and Knowledge Discovery, 32(5), 1275 – 1305.*

Rzepakowski, P. and Jaroszewicz, S. (2012). Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems, 32, 303 – 327.*

Setkute, J. and Dibb, S. (2022). "Old boys' club": Barriers to digital marketing in small B2B firms. *Industrial Marketing Management, 102, 266 – 279.*

Shankar, V. and Balasubramanian, S. (2009). Mobile marketing: a synthesis and prognosis. *Journal of Interactive Marketing, 23***,** *118 – 129.*

Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication*, Urbana: University of Illinois Press.

Shimizu, A., Togashi, R., Lam, A. and Huynh, N. V. (2019). Uplift modeling for cost effective coupon marketing in c-to-c e-commerce. In: *2019 IEEE 31st International Conference on Tools with Artifical Intelligence (ICTAI).*

Shteingart, H., Oostra, G., Levinkron, O., Parush, N., Shabat, G. and Aronovich, D. (2022). ML prescriptive canvas for optimizing business outcomes. In: *Proceedings of 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD2022).*

Soltys, M., Jaroszewicz, S. and Rzepakowski, P. (2015). Ensembe methods for uplift modeling. *Data Mining and Knowledge Discovery, 29***,** *1531 – 1559.*

Su, X., Kang, J., Fan, J., Levine, R.A. and Yan, X. (2012). Facilitating score and causal inference trees for large observational studies. *Journal of Machine Learning Research,* ***13****, 2955 – 2994.*

Taylor, D. G., Voelker, T. A. and Pentina, I. (2011). Mobile application adoption by young adults: a social network perspective. *WCOB Faculty Publications.*Paper 1. Retrieved Jul 30, 2022 from http://digitalcommons.sacredheart.edu/wcob_fac/1?utm_source=digitalcommons.sacredheart.edu%2Fwcob_fac%2F1&utm_medium=PDF&utm_campaign=PDFCoverPages

Thomas, A. R. (2007). The end of mass marketing: or, why all successful marketing is direct marketing. *Direct Marketing: An International Journal*, vol.1, no.1, 6-16.

Venkatesh, R., Mathew, L. and Singhal, T.K. (2019). Imperatives of business models and digital transformation for digital service providers. *International Journal of Business Data Communications and Networking, 15(1), 105 – 124.*

Wertime, K. and Fenwick, I. (2011). *DigiMarketing: the essential guide to new media and digital marketing.* John Wiley&Sons, Hoboken, NJ.

Wing, J. M. (2020). Ten research challenge areas in data science. *Harvard Data Science Review 2, 3.* Retrieved Sep 4, 2022 from https://hdsr.mitpress.mit.edu/pub/d9j96ne4/release/3

Wohllebe, A., Hübner, D., Radtke, U. and Podruzsik, S. (2021). Mobile apps in retail: effect of push notification frequency on app user behavior. *Innovative Marketing, 17(2), 102 – 111.*

Yap, B.W., Rani, K.A., Rahman, H.A.A., Fong, S., Khairudin, Z., Abdullah, N.N. (2014). An Application of Oversampling, Undersampling, Bagging and Boosting in Handling Imbalanced Datasets. In: *Herawan, T., Deris, M., Abawajy, J. (eds) Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013).* Lecture Notes in Electrical Engineering, vol 285. Springer, Singapore.

Yasmin, A., Tasneem, S. and Fatema, K. (2015). Effectiveness of digital marketing in the challenging age: an empirical study. *International Journal of Management Science and Business Administration, 1***,** *69 – 80.*

Yeomans, K.A. and Golder, P.A. (1982). The Guttmann-Kaiser criterion as a predictor of the number of common factors. *The Statistician, vol. 31, No.3.*

Zaniewicz, L. and Jaroszewicz, S. (2013). Support vector machines for uplift modeling. In *2013 IEEE13th international conference on data mining workshops*, 131-138. New York: IEEE.

Zhao, Y. and Harinen, T. (2019). Uplift modeling for multiple treatments with cost optimization. In *IEEE international conference on data science and advanced analytics*, 1 – 10. New York: IEEE.

Zimmermann, A., Blythe J. (2013). *Business to Business Marketing Management A Global Perspective.* Routledge Taylor & Francis Group London and New York.

Zulaikha, S., **Mohamed, H., Kurniawati, M., Rusgianto, S. and Rusmita, S.A.** (2021). Customer predictive analytics using artificial intelligence. *The Singapore Economic Review.* World Scientific Publishing Company.

APPENDICES

APPENDIX A: The Confusion Matrix results of oversampling, undersampling and combination models

APPENDIX B: The Summary of Two Model Approach models

APPENDIX A: The Confusion Matrix results of oversampling, undersampling and combination models

Table A.1 : Confusion matrix results of sampling methods

	Accuracy	**Precision**	**Sensitivity**	**Specificity**
Oversampling	0.861	0.1126	0.4362	0.8763
Undersampling	0.637	0.0759	0.8454	0.6293
Both	0.8353	0.1061	0.5034	0.8472

APPENDIX B: The Summary of Two Model Approach models

Summary of Treatment Model

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.6912	-1.2845	0.7254	0.8644	1.8404

Coefficients:

	Estimate	Std.Error	z value	Pr(>|z|)
(Intercept)	0.20748	0.10099	2.054	0.039930 *
Segment2	0.04609	0.09764	0.472	0.636911
Segment3	-0.72547	0.18014	-4.027	5.64e-05 ***
Tariffclass2	0.19709	0.16104	1.224	0.221014
Tariffclass3	-14.84599	248.89178	-0.060	0.952436
Tariffclass4	-15.39579	624.19342	-0.025	0.980322
Dataoveruser1	0.37432	0.19070	1.963	0.049657 *
Pbtwelve1	-0.42739	0.28744	-1.487	0.137038
Pbnine1	-1.27877	0.38138	-3.353	0.000800 ***
Pbsix1	2.10665	0.28668	7.348	2.01e-13 ***
Pbthree1	0.39811	0.11724	3.396	0.000684 ***
PC1	-0.04082	0.02245	-1.818	0.069031 .
PC2	0.17680	0.02102	8.412	< 2e-16 ***
PC3	-0.14918	0.03964	-3.763	0.000168 ***
PC4	-0.14445	0.04811	-3.003	0.002677 **

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 5489.7 on 4344 degrees of freedom
Residual deviance: 5112.4 on 4330 degrees of freedom
AIC: 5142.4

Number of Fisher Scoring iterations: 13

Summary of Control Model

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.8995	-0.8380	-0.5876	0.9459	2.1289

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.169405	0.050285	-23.255	< 2e-16 ***
Segment2	0.110056	0.033265	3.308	0.000938 ***
Segment3	0.744754	0.047285	15.750	< 2e-16 ***
Tariffclass2	0.070554	0.044096	1.600	0.109598
Tariffclass3	-0.091485	0.089842	-1.018	0.308540
Tariffclass4	0.335333	0.360216	0.931	0.351894
Dataoveruser1	1.201640	0.039456	30.456	< 2e-16 ***
Pbtwelve1	0.126025	0.079899	1.577	0.114725
Pbnine1	-0.205560	0.090508	-2.271	0.023136 *
Pbsix1	0.771414	0.060240	12.806	< 2e-16 ***
Pbthree1	0.852787	0.043176	19.751	< 2e-16 ***
PC1	0.077996	0.008452	9.228	< 2e-16 ***
PC2	0.203996	0.011702	17.432	< 2e-16 ***
PC3	0.136502	0.012258	11.135	< 2e-16 ***
PC4	-0.239872	0.013814	-17.364	< 2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 40036 on 28928 degrees of freedom
Residual deviance: 33349 on 28914 degrees of freedom
AIC: 33379

Number of Fisher Scoring iterations: 4

CURRICULUM VITAE

Name Surname : Meltem SANİSOĞLU

EDUCATION :

- **B.Sc.** : 2012, Istanbul Technical University, Faculty of Chemical-Metallurgical Eng., Chemical Engineering
- **M.Sc.** : 2017, Istanbul Technical University, Faculty of Management Engineering, Management Eng.

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2023 - cont. Senior Product Manager at Trendyol Group
- 2014 - 2023 Product Manager at Vodafone Turkey

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Sanisoğlu, M., Kaya, T., Burnaz, Ş.** (2022). Marketing Campaign Management Using Machine Learning Techniques: An Uplift Modeling Approach. In: Intelligent and Fuzzy Systems. INFUS 2022. Lecture Notes in Networks and Systems, vol 505. Springer, Cham. https://doi.org/10.1007/978-3-031-09176-6_17
- **Sanisoglu, M., Burnaz, S. & Kaya, T.** (2023). A gateway toward truly responsive customers: using the uplift modeling to increase the performance of a B2B marketing campaign. *Journal of Marketing Analytics* (2023). https://doi.org/10.1057/s41270-023-00254-2

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- **Ulubaşoğlu G., Şenel M., and Burnaz S.** (2017). To Switch or Not? Analyzing the Question for Consumers in Turkish Mobile Telecommunications. Marketing at the Confluence between Entertainment and Analytics. Developments in Marketing Science: Proceedings of the Academy of Marketing Science. Springer, Cham. https://doi.org/10.1007/978-3-319-47331-4_193.