

**ANKARA YILDIRIM BEYAZIT UNIVERSITY**  
**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**



**LIE DETECTION FROM VOICE AND HEART RATE**

**M.Sc. Thesis by**  
**Şerife ÖZDAMARLAR**

**Department of Computer Engineering**

**May, 2019**

**ANKARA**

# **LIE DETECTION FROM VOICE AND HEART RATE**

**A Thesis Submitted to**

**The Graduate School of Natural and Applied Sciences of**

**Ankara Yıldırım Beyazıt University**

**In Partial Fulfillment of the Requirements for the Degree of Master of Science  
in Computer Engineering, Department of Computer Engineering**

**by**

**Şerife ÖZDAMARLAR**

**May, 2019**

**ANKARA**

## M.Sc. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**LIE DETECTION FROM VOICE AND HEART RATE**” completed by **ŞERİFE ÖZDAMARLAR** under the supervision of **ASSIST. PROF. DR. ÖZKAN. KILIÇ** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Özkan KILIÇ

Supervisor

Assist. Prof. Dr. Gül TOKDEMİR

Jury Member

Assist. Prof. Dr. Hilal KAYA

Jury Member

Prof. Dr. Ergün ERASLAN

Director

Graduate School of Natural and Applied Sciences

## **ETHICAL DECLARATION**

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information and documents are obtained in the framework of academic and ethical rules,
- All information, documents and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of above statements, I accept to renounce all my legal rights.

**Date: 21.05.2019**

**Signature: .....**

**Name & Surname: Şerife ÖZDAMARLAR**

## **ACKNOWLEDGMENTS**

Firstly, I would like to express my sincere gratitude to my supervisor, Assist. Prof. Dr. Özkan KILIÇ for his tremendous support and motivation during my study. His immense knowledge and precious recommendations constituted the milestones of this study. His guidance assisted me all the time of my research and while writing this thesis.

I would like to express my greatest and deepest gratitude for my dear father Mustafa Özdamarlar, my dear mother Hatice Özdamarlar and my dear sisters Havva and Döne Özdamarlar for their support, motivation and understanding during my study.

Finally, I would like to thank my friends for their support and motivation.

**2019, 21 May**

**Şerife ÖZDAMARLAR**

# **LIE DETECTION FROM VOICE AND HEART RATE**

## **ABSTRACT**

A number of recent studies have shown that much of the information obtained from voice data has a variety of applications in research, including emotion analysis, estimation of gender and age, and lie detection. The aim of this study is to detect a lie by using voice and heart rate data and to compare accuracies of different machine-learning algorithms. Although there are many studies on lie detection in the literature, our study focuses on lie detection from not only voice but also heart rate data. 15 women and 15 men speakers were selected for this study. Each speaker was asked 10 questions and a total of 300 sound files were recorded using an external professional microphone. A smart wristband was used to collect heart rate data from each speaker. Although our main focus was voice and heart rate, a personality test was also given to the participants. Twelve features from the voice data were extracted by using MFCC. Minimum, maximum and average heart rate were calculated for each sound file. Lie detection results calculated from support vector machine, artificial neural network, logistic regression and decision tree methods were compared in the current study. The best accuracy (66,67%) was achieved with Logistic Regression which had 19 features among other used methods. When accuracy rates of all methods were taken into consideration, it was concluded that the better accuracy rate was achieved after the addition of new features which are heart rate, gender and personality type to sound data for lie detection. The main finding of this study was that only the voice data and the pulse data failed to yield any reliable results in detecting the lies. However, it was found that every new added attribute could not improve the results. As a result of this study, we have learned the importance of the selection of attributes and of using the attributes together.

**Keywords:** Artificial Neural Network, Decision Tree, Heart Rate, Lie Detection, Logistic Regression, MFCC, Personality Traits, Support Vector Machine

# SES VE KALP ATIŞ HIZINDAN YALAN TESPİTİ

## ÖZ

Son zamanlarda yapılan duygu analizi, cinsiyet tahmini, yaş tahmini ve yalan tespiti çalışmalarının da içinde olduğu araştırmalar ses verisinden birçok bilgi elde edildiğini ortaya koymuştur. Bu çalışmanın amacı ses ve nabız verilerini kullanarak yalan tespiti yapmak ve farklı makine öğrenmesi algoritmalarında uygulanarak başarı sonuçlarının karşılaştırılmasıdır. Literatürde yalan tespiti üzerine birçok çalışma bulunmasına rağmen bizim çalışmamız yalan tahmininde sadece ses üzerine değil nabız verisine de odaklanmaktadır. Bu çalışma için 15 kadın 15 erkek konuşmacı seçilmiştir. Her bir konuşmacıya 10 soru sorulmuştur ve toplamda 300 ses dosyası harici profesyonel bir mikrofon kullanılarak toplanmıştır. Her konuşmacıdan nabız verilerinin alınması için akıllı bileklik kullanılmıştır. Asıl odak noktamızın ses ve nabız olmasına rağmen katılımcılara kişilik testi de uygulanmıştır. Mel frekans kepsitrum katsayıları (MFCC) kullanılarak ses verisinden 12 tane özellik çıkarılmıştır. Her ses dosyası için en düşük, en yüksek ve ortalama nabız verileri hesaplanmıştır. Bu çalışmada destek vektör makineleri, yapay sinir ağları, lojistik regresyon ve karar ağaçları metotlarının yalan tespiti sonuçları karşılaştırılmıştır. Kullanılan yöntemler arasında lojistik regresyon algoritmasında 19 öznitelik ile %66,67 oranında en yüksek doğruluk oranı elde edilmiştir. Tüm metotların doğruluk oranları göz önünde bulundurulduğunda ses verisine nabız, yaş, cinsiyet ve kişilik tipleri eklendikten sonra yalan tespiti yapmak için daha iyi sonuçlar elde edildiği sonucuna varılmıştır. Bu çalışmanın temel bulgusu yalnızca ses ve nabız verisinin yalan tespitinde etkili sonuçlar vermediğidir. Bununla birlikte, eklenen her yeni öznitelik sonuçları iyileştirmedeği görülmüştür. Bu çalışma sonucunda öznitelik seçiminin önemini ve özniteliklerin bir arada kullanılmasının önemi öğrenilmiştir.

**Anahtar Kelimeler:** Yapay Sinir Ağları, Karar Ağaçları, Kalp Atışı, Yalan Tespiti, Lojistik Regresyon, MFCC, Kişilik Tipleri, Destek Vektör Makineleri

## CONTENTS

|                                               |           |
|-----------------------------------------------|-----------|
| M.Sc. THESIS EXAMINATION RESULT FORM.....     | ii        |
| ETHICAL DECLARATION .....                     | iii       |
| ACKNOWLEDGMENTS .....                         | iv        |
| ABSTRACT.....                                 | v         |
| ÖZ .....                                      | vi        |
| NOMENCLATURE.....                             | ix        |
| LIST OF FIGURES .....                         | x         |
| LIST OF TABLES .....                          | xi        |
| <b>CHAPTER 1 - INTRODUCTION.....</b>          | <b>1</b>  |
| 1.1 Purpose of Thesis .....                   | 1         |
| 1.2 Research Questions and Hypotheses .....   | 1         |
| 1.3 Organization of the Thesis .....          | 2         |
| 1.4 Related Works .....                       | 3         |
| 1.5 Lying .....                               | 12        |
| 1.5.1 Definition of Lie.....                  | 12        |
| 1.5.2 Factors of lying .....                  | 12        |
| 1.6 Heart Rate .....                          | 13        |
| 1.7 Big Five Personality Test .....           | 14        |
| 1.7.1 Big Five Personality Traits .....       | 14        |
| 1.8 Performance Criteria .....                | 16        |
| 1.8.1 Confusion Matrix .....                  | 16        |
| 1.8.2 Accuracy .....                          | 17        |
| 1.8.3 Precision.....                          | 18        |
| 1.8.4 Recall.....                             | 18        |
| 1.8.5 F1 - Score .....                        | 18        |
| <b>CHAPTER 2 - RESEARCH METHODOLOGY .....</b> | <b>19</b> |
| 2.1 Dataset .....                             | 19        |
| 2.1.1 Pre-Experiment .....                    | 19        |
| 2.1.2 Fundamental Dataset.....                | 20        |

|                                                                |                                          |           |
|----------------------------------------------------------------|------------------------------------------|-----------|
| 2.2                                                            | Methods .....                            | 29        |
| 2.2.1                                                          | Feature Extraction (MFCC) .....          | 30        |
| 2.2.2                                                          | Classification Methods .....             | 31        |
| <b>CHAPTER 3 - APPLICATION &amp; EXPERIMENTAL RESULTS.....</b> |                                          | <b>37</b> |
| 3.1                                                            | Application .....                        | 37        |
| 3.1.1                                                          | Development Environment and Methods..... | 37        |
| 3.1.2                                                          | Application of Learning Methods .....    | 38        |
| 3.2                                                            | Experimental Results.....                | 40        |
| 3.2.1                                                          | Support Vector Machine Results .....     | 40        |
| 3.2.2                                                          | Logistic Regression Results .....        | 45        |
| 3.2.3                                                          | Decision Tree Results .....              | 50        |
| 3.2.4                                                          | Artificial Neural Network Results .....  | 54        |
| 3.2.5                                                          | Comparison of All Methods.....           | 59        |
| <b>CHAPTER 4 - CONCLUSION &amp; FUTURE WORK.....</b>           |                                          | <b>64</b> |
| 4.1                                                            | Conclusion.....                          | 64        |
| 4.2                                                            | Future Work .....                        | 65        |
| <b>REFERENCES.....</b>                                         |                                          | <b>67</b> |
| <b>APPENDICES .....</b>                                        |                                          | <b>72</b> |
| APPENDIX A. Written Documents of the Main Experiment.....      |                                          | 72        |
|                                                                | Scenario .....                           | 72        |
|                                                                | Experiment Questions.....                | 73        |
|                                                                | Informed Consent Form.....               | 74        |
|                                                                | Personality Test.....                    | 75        |
| <b>CIRRICULUM VITAE .....</b>                                  |                                          | <b>76</b> |

## **NOMENCLATURE**

### **Acronyms**

ANN Artificial Neural Network  
CSC Columbia-SRI-Colorado  
FNIR Functional Near Infrared Spectroscopy  
GMM Gaussian Mixture Model  
GSR Galvanic Skin Response  
HMM Hidden Markov Modeling  
HR Heart Rate  
LVA Layered Voice Analysis  
MFCC Mel Frequency Cepstrum Coefficients  
SVM Support Vector Machine  
VSA Voice Stress Analyzer

## LIST OF FIGURES

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| <b>Figure 1.1</b> Average heart rates for truthful and lying response.....            | 5  |
| <b>Figure 1.2</b> Result comparison between lie and true average.....                 | 9  |
| <b>Figure 1.3</b> Big Five Factor . ....                                              | 15 |
| <b>Figure 1.4</b> Confusion Matrix.....                                               | 17 |
| <b>Figure 2.1</b> Steps to create the dataset .....                                   | 20 |
| <b>Figure 2.2</b> Example of getting spectrum .....                                   | 21 |
| <b>Figure 2.3</b> Heart Rate Fluctuated Graph Example .....                           | 22 |
| <b>Figure 2.4</b> Heart Rate Fixed Graph Example.....                                 | 22 |
| <b>Figure 2.5</b> Lie and Truth Cards .....                                           | 24 |
| <b>Figure 2.6</b> Steps of creating the dataset file .....                            | 27 |
| <b>Figure 2.7</b> MFCC Block Diagram.....                                             | 30 |
| <b>Figure 2.8</b> A simple artificial neural network . ....                           | 34 |
| <b>Figure 3.1</b> Application Flow Diagram .....                                      | 38 |
| <b>Figure 3.2</b> The Graph of SVM Results .....                                      | 43 |
| <b>Figure 3.3</b> The Graph of LR Results .....                                       | 48 |
| <b>Figure 3.4</b> The Graph of Decision Tree Results.....                             | 52 |
| <b>Figure 3.5</b> The Graph of Artificial Neural Network Results .....                | 57 |
| <b>Figure 3.6</b> Comparison of All Methods Accuracy Results With 20% Test Size ..... | 61 |
| <b>Figure 3.7</b> Comparison of All Methods Accuracy Results With 25% Test Size ..... | 62 |

## LIST OF TABLES

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| <b>Table 1.1</b> Normal resting heart rates .....                     | 13 |
| <b>Table 2.1</b> Heart Rate List. ....                                | 23 |
| <b>Table 2.2</b> Example of Extracted 12 MFCC. ....                   | 27 |
| <b>Table 2.3</b> Last Column Numbers .....                            | 35 |
| <b>Table 3.1</b> Results of Support Vector Machines.....              | 41 |
| <b>Table 3.2</b> - SVM Heart Rate Results .....                       | 43 |
| <b>Table 3.3</b> SVM MFCC Results .....                               | 44 |
| <b>Table 3.4</b> SVM Heart Rate and MFCC Results .....                | 45 |
| <b>Table 3.5</b> Results of Logistic Regression .....                 | 45 |
| <b>Table 3.6</b> Logistic Regression Heart Rate Results .....         | 48 |
| <b>Table 3.7</b> Logistic Regression MFCC Results .....               | 49 |
| <b>Table 3.8</b> Logistic Regression Heart Rate and Mfcc Results..... | 50 |
| <b>Table 3.9</b> Results of Decision Tree .....                       | 50 |
| <b>Table 3.10</b> Decision Tree Heart Rate Results .....              | 53 |
| <b>Table 3.11</b> Decision Tree MFCC Results .....                    | 53 |
| <b>Table 3.12</b> Decision Tree Heart Rate and MFCC Results .....     | 54 |
| <b>Table 3.13</b> Results of Artificial Neural Network .....          | 55 |
| <b>Table 3.14</b> Artificial Neural Network Hr Results .....          | 58 |
| <b>Table 3.15</b> Artificial Neural Network MFCC Results.....         | 58 |
| <b>Table 3.16</b> Artificial Neural Network Hr and Mfcc Results ..... | 59 |
| <b>Table 3.17</b> Results of All Features with 25% Test Size.....     | 60 |
| <b>Table 3.18</b> Results of 19 Feature with 20% Test Size.....       | 60 |
| <b>Table 3.19</b> Hypothesis Results .....                            | 63 |

# CHAPTER 1

## INTRODUCTION

### 1.1 Purpose of Thesis

The aim of this study is to determine whether a person lies or not by using sound, heart rate, age, gender, personality traits and duration of the recorded sound. For this purpose, machine learning methods, namely support vector machines (SVM), artificial neural network (ANN), decision tree and logistic regression, are applied. Thirty people participate voluntarily in a lie detection experiment. Their speech data is categorized into two groups for training and testing purposes. Ultimately, we search for the best method to detect lying with respect to these data.

Although there are various studies in this area, unlike this study, they do not use both voice and heart rate as a complementary modality. Moreover, detection of lying depends on other factors including age, gender and personality types. Personality types are determined thanks to the five-factor personality tests in the scope of work [1].

### 1.2 Research Questions and Hypotheses

Detection of lying may require many different characteristics due to variation in the participants. For this reason, the effect of different features on lies is studied in this thesis. The following research questions and corresponding hypotheses are pursued:

Q1: Is there an effect of voice and heart rate on detection of lying?

H1: The author hypothesizes that voice and heart rate improve lie detection

Q2: Can lies be detected only by voice?

H2: It is hypothesized that voice-only models will have more than 50% accuracy rate.

Q3: Can lies be revealed only by heart rate?

H3: It is hypothesized that the heart rate is not enough to detect lies successfully.

Q4: Does age affect the detection of lying?

H4: The author hypothesizes that age has an effect on lie detection.

Q5: Does gender have an effect on lying?

H5: The author hypothesizes that gender has an effect on lie detection.

Q6: Can personality traits affect the detection of lying?

H6: We hypothesize that personality trait improves the accuracy of lie detection.

In order to test our hypothesis, the author used machine learning techniques as explained in the following chapters.

### **1.3 Organization of the Thesis**

This thesis consists of five chapters. The rest of the thesis is arranged as follows:

Chapter 2 is the section where the generated dataset is explained. The chapter contains pre-experiment and fundamental dataset. This chapter also explains the steps to create the fundamental dataset: collecting voice data, collecting heart rate data, creating dataset and big five personality test and compiling data files.

Chapter 3 describes the methods employed. First, the main step of machine learning application is described in this section. MFCC, the feature extraction method, is explained. Secondly, the classification methods are explained. In this section, machine learning algorithms are mentioned: support vector machine, logistic regression, decision tree and artificial neural network algorithms.

Chapter 4 consists of application steps and the results of the experiments. The first section of this chapter describes the application steps and clarifies development environments. Next section is the discussion of experimental results, including

results of support vector machine, logistic regression, decision tree and artificial neural network algorithms. Finally, this section compares the results of all methods.

Chapter 5, the last chapter, contains the conclusion of this thesis and recommendations for future works.

## 1.4 Related Works

It is often possible to detect when a close friend or acquaintance is lying based on their gestures, facial expressions, and voice tone and speaking shape. The same clues may not be helpful for to judge a stranger's truthfulness. This study investigates whether access to other data, such as the person's age, gender, personality type, and heart rate can indicate whether a person lying. In order for a computer to make such a judgement, a dataset and corpus are required.

There is a plethora of corpora and studies on speech processing. Feature extraction is the first step for machine learning; this subject has been intensely processed in the literature. One study compares different feature extraction methods such as Mel frequency cepstrum coefficients (MFCC), Gaussian Mixture Model (GMM), and Hidden Markov Model (HMM) [2]. MFCC is the one with the best accuracy at 95%. This result encouraged us to use MFCC as the feature extraction method in the current study.

There is an important study which tries to detect a lie using Columbia-SRI-Colorado (CSC) corpus [3]. MFCC was used as the feature extraction method and six different classification methods were investigated. These classification methods were logistic regression, support vector machine, gradient boost classifier, multi-layer perceptron classifier, bagging classifier and recurrent neural network. Support vector classifier produces the best result with the accuracy between 54% and 61%. This study contributes to our study in terms of using MFCC and comparing classification methods.

Another paper in the field of psychology on lying studies the effects of gender and perceived benefits [4]. This paper makes inferences about the situations in which people lie. Taking advantage of this study, we selected the questions for the experiment carried out within the scope of our study. We benefited from this study in the situations in which people lie.

Ayçay has another study in the field of psychology which compares genders on lying which determined that women are more powerful than men and that they can keep the lies more easily in mind [5]. Throughout the study, it was observed that most women had no difficulty lying and did not reduce their normal speaking speed. On the other hand, male participants were found to have more difficulties.

In an important book in the field of psychology on detecting lies and deceit, Vrij defines deception, types of lies, reasons of lying, the frequency of lying, and personal characteristics of the liar [6]. This book also presents information on polygraph tests, the scientific content analysis criteria, and the guilty knowledge test, important for us to obtain information on the personal characteristics of the liars.

Mishra and Vaidya used voice analysis to detect different stress situations, namely, neutral, happy, sad and angry [7]. Their participants were mainly South Asians/Native Hindi speakers speaking English. The authors report two accuracy results. First, the average efficiency of the human stress recognition system is more than 95%. Second, the average efficiency of the automatic stress recognition system is approximately 85%. Since lying creates stress, this paper suggests that heart rates measure would be useful.

Another study analyzes fluctuations in heart rate, respiration, and perspiration as physiological indicators of deception [8]. The study questions whether or not there are physiological diversities while lying. The paper reports different results for heart rate. However, heart rate increases when a person lies. The ultimate result of this study is that the heart rate has no important difference according to basic evaluation as shown in Figure 1.1.

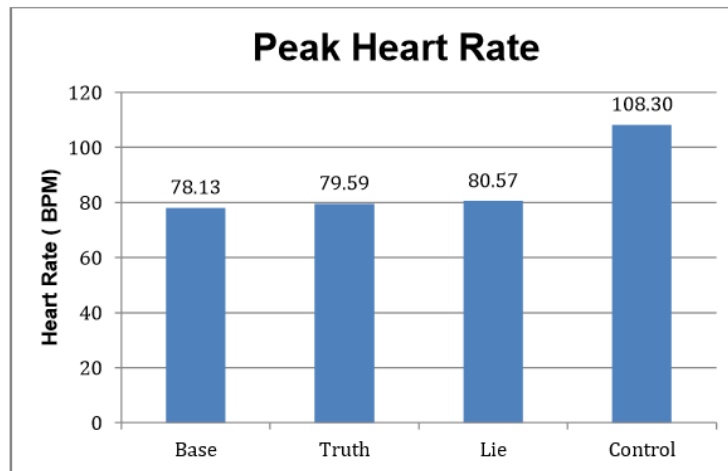


**Figure 1.1** Average heart rates for truthful and lying response [8]

A study proposed voice activity detection which used MFCC features and Support Vector Machine (SVM) [9]. Their study used three datasets: NIST2005, Bus Stop, and Lab datasets. The dataset was separated as 30% training and 70% testing. According to this study, compared to traditional methods, SVM is simpler to re-adjust the new datasets than traditional methods so long as there is a short training voice sample.

Another study showed that the combination of MFCCs and SVMs is a powerful tool in recognizing the speaker gender [10]. They used 124 female voices and 122 male voices for gender detection. There is a comparison between different MFCC numbers. These are 12, 19, 26, 34 and 40 MFCCs. They also studied RBF and polynomial kernel functions, revealing that RBF kernel and 26 MFCCs has the best prediction.

Another study on kernel function selection for the solution of classification problems via SVMs considered whether the differences between the performance of the kernel functions are determined independently of the influences of the data sets [11]. The results indicated that the kernel function with the most successful performance is radial basis function, and that there is no difference between the polynomial kernel

function and performances. However, it was seen that the polynomial kernel function is costly in the implementation of the algorithm.

Thanks to these two studies, radial based kernel function is used in our study.

Gilke et al. proposed a paper which is MFCC-based vocal emotion recognition using ANN [12]. In this paper, MFCC was used as a feature extraction method and ANN was used as a classification method. This study carried out for three different languages: English, Hindi, and Marathi. Five different emotional classes were also studied: neutrality, happiness, sadness, anger, and surprise. The best accuracy -- 85% -- was achieved in the in anger emotional class in the Hindi language. In the Marathi language, anger had maximum accuracy with 83,5%. Finally, the maximum accuracy of the English language was 78,6% in surprise emotional class. This study provides a useful perspective to apply MFCC and ANN algorithms.

Gershenson's paper offers an introduction to and general description of ANNs and involves a basic artificial neural network model [13]. Furthermore, it mentions weights and a backpropagation algorithm. This paper helped us understand how ANN works.

Gupta et al. analyze various decision tree algorithms for classification in data mining and describe various algorithms of the decision tree, their features, advantages, and disadvantages [14]. It also comprises applications of decision trees in e-commerce, industry, intelligent vehicles, medicine, and business. Although this article is in the area of data analysis, it is also useful for our study.

Another work on decision tree methods (applications for classification and prediction) explains the basic concepts of decision tree algorithms. [15]. These concepts are nodes, branches, splitting, stopping, and pruning. The authors give an example of the decision tree algorithm and describe its advantages. This study enabled us to understand the basics of the decision tree algorithm.

DePaulo et al. proposed an experimental study on lying in everyday life [16]. An experiment was conducted among 77 college students and 70 community members and compared males to females. This study also calculated a rate of lying for both college students and community members. There was no important influence of the gender of participants on lying. In addition, this study presented a comparison of social interactions in lying, kinds of lies and characteristics of the lies. This study is important for us in explaining sex differences in lying.

Given that heart rate is one of the most important issues for this thesis, MacGill's paper helpfully explained the definition of heart rate and normal resting heart rate [17]. It included a table of normal resting heart rates at different ages, according to the National Institutes of Health. Using the data from the paper, we judged whether or not the observed pulse changes are normal.

A study that developed mobile application for lie detection is more successful for truth detection with 96% precision and 95% recall than lie detection with 65% precision and 59% recall [18]. This study was executed in three stages. The first stage was simple lies, carried out with 100 participants. The second stage was the ultimatum game, executed with 41 participants. The last stage was the Yatzy game, which involved 51 participants. One of the results of this study was the duration of speech while telling the truth is shorter than that of telling a lie. This study showed that dishonesty influences user interaction with mobile devices. The study further proposed an exploration of lie detection among mobile devices, irrespective of the input content.

Vicianova proposed historical techniques of lie detection [19]. Early methods of lie detection were defined in 1000 BC in China, where a test was executed with rice, suggesting that people lie when they dry their mouths. Greek physicist and physician proposed lie detection by calculating the heart rate. Vicianova mentioned the trial by ordeal method, also known as the Judgement of God. It was believed that various experiments using divine forces could ascertain whether or not a person was lying. These methods were used up to the 16th century. Modern methods of lie detection

emerged in the 1870s, with a great number of lie detecting machines appearing after 1881. The most important of these was the polygraph. After the 1980s, brain-based lie detection methods were improved. These methods form the basis of lie detecting techniques.

Another study aimed at emotion recognition analyzed speech under stress using linear techniques and non-linear techniques [20]. The linear technique involved pitch, vocal tract spectrum, formant frequencies, duration and MFCC. Because emotion recognition is used in lie detection, this is an important issue for us. This study, which used a German database, was executed for neutrality, happiness, disgust, sadness, boredom and anger. The mean values of MFCC neutral emotion are lower than those of MFCC achieved in boredom emotion.

Sondhi et al. proposed a voice-analysis study to detect deception [21]. This study tried to detect voice parameters for a reliable mark of stress owing to deception. Sondhi et al. used PRAAT software for analysis and studied the fundamental and formant frequencies. They reported that fundamental frequency and the first two formant frequencies are related to stress. As a result of this study, the fundamental frequency could be used as a stress indicator for lie detection because stress could increase formant frequencies. However, this increase was not a sufficient indicator. At the beginning, we thought that formant frequencies would be an effective indicator for detecting lies. However, we decided not to use formant frequencies thanks to studies of Sondhi et al. and our experimental results.

Another voice stress detection study presented stress or speech style as an identification mechanism [22]. The aim of this study was to study the influence of physical stress on timewise patterns of speech based on yes/no questions. The participants experienced no stress in this study because there was no real pressure on participants. The authors compared voice stress analysis technology and polygraph systems, concluding that the identification of stress can be more reliable in voice stress analysis technology than polygraph systems.

Pan et al. proposed a study that applies fractional Mel cepstral coefficient in deceptive speech detection [23]. The study reported that the accuracy of MFCC in detecting deception through linear discriminant analysis or hidden Markov model is lower than the fractional Mel cepstral coefficient. Using hidden Markov model, the accuracy was 71% for males and 70,2% for females.

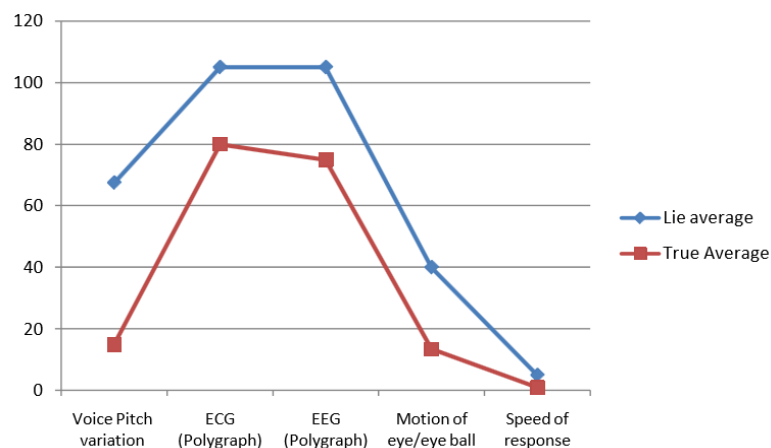


**Figure 1.2** Result comparison between lie and true average

Suresh and Shaheen put forward a study that analyzed speech stress based on lie detector for loyalty test [24]. They used artificial neural network method for lie detection. Figure 1.2 shows the difference between lie and true average. The study concluded that lying increases the parameters. Accuracy was around 92% in high-stress situations. Compared to other studies, this accuracy rate is quite high.

Another study on deception detection in the field of psychology presented lie detection approaches and methods and examined recent studies and proposed their results [25]. One of these accuracy rates reported from a recent study was 77% in real time measurement. Furthermore, the author describes a study which reported accuracy in lie detection and truth detection. The author also compared laboratory results with and field study results. As a result, this study focused on the development of lie detection in applied fields.

Bustami et al. proposed lie detection system from voice using bi-directional associative memory algorithm [26]. This lie detection system processed sound and used some words like 'go', 'roads' and 'move'. They used artificial neural network algorithm for this system. The lie detection system had maximum accuracy rate 90% for the verb 'go' and a 70% minimum accuracy rate for the verb 'move'. This study presented the use of bidirectional associative memory for lie detection from voice.

Street presented ALIED (Adaptive Lie Detector theory), a theory of lie detection [27]. This study researched the ALIED theory, which argued that people make smart and conscious judgements. If judgments regarding whether someone is lying are reliable, then they decide accordingly. In fact, this is like educated learning. This study tries to understand how people lie or make real judgments. In this study, people were given 10 dollars to encourage them to lie. Researchers observed how the participants identified the lies. Perez-Rosas et al. developed verbal and nonverbal clues for real-life deception detection [28]. They developed a dataset of 118 videos from real life scenarios (television street interviews and trials). The study used both facial-gesture expression and speech transcription. They executed a study for the human capability for lie detection using both verbal and nonverbal features, as well as SVMs, DT and random forest classification algorithms. This study had an accuracy rate between 77% and 82%. The authors argue that their system is better than the human ability to detect lying.

Vrij et al. proposed a cognitive approach for both verbal and non-verbal lie detection and provide a meta-analysis [29]. This approach involved three techniques: embarking cognitive load, fostering interviewees to tell more and asking surprise questions. This approach had 67% accuracy for lie detection. Vrij et al. reported 47% accuracy for a traditional standard approach. This study mentioned that telling the truth is mentally easier than lying according to all available theories. In addition, this approach helped to determine the experiment questions in our study.

Another study presents a comparison between total detail and proportion of complications within-subjects verbal lie detection measures [30]. 27 truth-tellers and

26 liars participated in this study. The truth tellers describe a trip they have taken, while the liars concocted one. This study reported that there is no relation between the amount of information for truth tellers or liars. Vrij et al. also report that liars added fewer complications to their stories and that their stories were more scripted than the truth tellers'.

A study in the field of science and pseudoscience on voice stress analyses mentioned scientific forensic methodology, voice stress analyzer (VSA) and layered voice analysis (LVA)[31]. The study reported that exploring speech samples is difficult which is like low quality and adverse acoustic environment. We, too, experienced this difficulty; we were able to record correctly only two times. Lacerda reported that VSA and LVA are not good options for criminal investigators.

A study on lies, liars and lie detection presents the relation between lying and emotion, as well as complexity of lying.[32]. The author describes signs of deception and explains lie detection. This study mentions situations that increase and decrease the lying. According to this study, halted speech and delay between the question and answer signify a lie. This study also argues that an objective observer is able to perceive a lie. This study reported that police officers and college students are more skilled at detecting deception.

One study discusses performance criteria for lie detection through a videotaped interview [33]. A two-stage experiment was conducted in this study. Its first perspective was law enforcement and the second was non-professional interviewers. This study reported that judges' ability to detect lies is greater than their ability to detect truths. This study argued that it may be good to use scientific studies rather than the observations of the judges to decide whether the witness is telling the truth.

## 1.5 Lying

### 1.5.1 Definition of Lie

There are lots of definition of lying. According to Turkish Language Society, a lie is a failure to tell the truth and stands for, unrealistic promise and, crisp [34]. In addition to Turkish Language Society's definition, according to the Stanford Encyclopedia of Philosophy, lying may be separated into two types: The first type defines lying as a form of cheating. The second describes lying and cheating as not morally good. Thanks to these definitions, we can summary the lie as a false statement or deception. This deception stems from different intentions which can be damaging, punishment, rejection and loss fear. Thus, lie types can be categorized as false intention, big lie, economical lie, white lie, bluff, and such. Actually, lie can be seen as a kind of defense mechanism.

### 1.5.2 Factors of lying

Many factors cause lying. Features of lying vary according to gender or personality types.

Women tend to tell more lies in order not to disappoint people they lie to their girlfriends because girls want to hear good things. Women also tell lies on behalf of others. On the other hand, men tell lies to protect their pecuniary interests. Women's tendency to lie is socially motivated while men's is for personal benefit [16].

Lying is generally associated with personality characteristics. Those who lie can be understood from their personality traits. Although lying is for defending oneself under difficult conditions, personal characteristics can also lead to lying. The properties of the good liar are [6]:

- 1) Being a natural actor
- 2) Being well-prepared
- 3) Having a creative and natural mind.

- 4) Thinking quickly
- 5) Not being mentally forced
- 6) Having a good memory to recall the lies.
- 7) Not feeling guilty while lying
- 8) Being like an actor

Thus, it is quite difficult to classify the lies of people with the above-mentioned characteristics.

## 1.6 Heart Rate

The heart rate is defined as measurements of the number of times the heart beats per minute [17]. Heart rate is affected by physical activity, threats to safety, and emotional situations. The following table shows the distribution of age and normal heart rates.

**Table 1.1** Normal resting heart rates

| Age                 | Normal heart rate (bpm) |
|---------------------|-------------------------|
| Up to 1 month       | 70 to 190               |
| From 1 to 11 months | 80 to 160               |
| From 1 to 2 years   | 80 to 130               |
| From 3 to 4 years   | 80 to 120               |
| From 5 to 6 years   | 75 to 115               |
| From 7 to 9 years   | 70 to 110               |
| Over 10 years       | 60 to 100               |

According to Table 1.1, heart rates of healthy people who are over 10 years are between 60 and 100. In this thesis, participants are between 25 and 62 years. It is expected that participants' heart rate should be between 60 and 100. Since heart rate cannot be controlled by a person's own desire, this is an important indicator that can be used for lie detection.

## **1.7 Big Five Personality Test**

Five important personality traits or “big five factors” -- openness, conscientiousness, extraversion, agreeableness and neuroticism -- will be explained in this part.

### **1.7.1 Big Five Personality Traits**

Big five-factor personality test is an explanatory model that defines people’s personal features [35]. Under the same circumstances, different people might behave differently, depending on their personality traits. In the big five personality test, participants use five-point Likert scale to indicate to what extent they agree with 120 statements.

The big five-factor test is generally used as a hiring tool.

The big five factors are openness, conscientiousness, extraversion, agreeableness and neuroticism. OCEAN is the abbreviation of the big five-factor model. The following figure briefly shows features of this model.

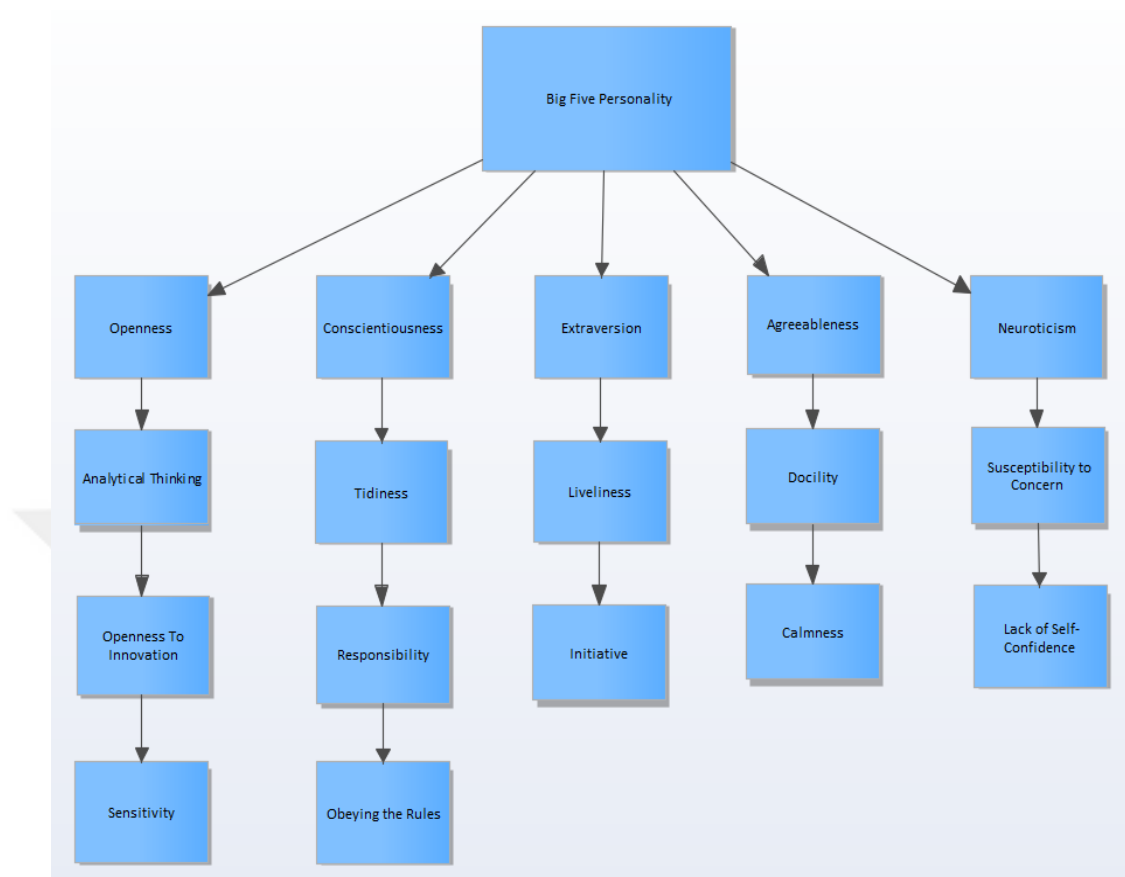


**Figure 1.3** Big Five Factor [36]

#### 1.7.1.1 Openness

This type can be defined as openness to experience, intelligence, or culture. A person that has high point in the openness factor has attributes like analytic, complex, curious, creative, untraditional, original, imaginative, and artistic.

#### 1.7.1.2 Conscientiousness

This factor is related to self-control, responsibility and messiness. Personas who score high on this personality type have some drawbacks like perfectionism and workaholism. Responsible persons are self-organized, planned and diligent. They are more likely to obey rules.

### **1.7.1.3 Extraversion**

Extraversion includes features like excitability, cheerfulness, talkativeness, sociability and liveliness. Actually, it is related to be social. Sociability and liking being with people, entertainment, leadership, power, willingness and friendly behavior also represent this type of five personality traits.

### **1.7.1.4 Agreeableness**

Agreeableness involves features like gentle, respectful, reliable, flexible and merciful. There is a tendency to think of others, to be calm in interpersonal relations, and to be sympathetic and friendly. A person who has a high score for this type is loved by others. On the other hand, a low score for this type indicates a jealous, quarrelsome and anomalous person in personal relationships.

### **1.7.1.5 Neuroticism**

Neuroticism can be said as emotional instability. This type is defined with features which are anxious, distrustful, nervous and worried. People who take a high score for this type create high-level goals and underestimate their own performances. Conversely, a low score suggests a person who is relaxed, unflappable and calm.

## **1.8 Performance Criteria**

There are multiple metrics for machine learning applications: accuracy, precision, sensitivity, F1 score and the confusion matrix.

### **1.8.1 Confusion Matrix**

A confusion matrix compares the classifications made by an algorithm to the existing values. Existing conditions are determined in training dataset. There is a sample 2x2 confusion matrix shown in Figure 1.4.

|                  |          | Existing Values |          |
|------------------|----------|-----------------|----------|
|                  |          | Positive        | Negative |
| Predicted Values | Positive | TP              | FP       |
|                  | Negative | FN              | TN       |

**Figure 1.4** Confusion Matrix

This figure shows the number of classified values as positive or negative. The existing value shows the number of labels that are true or false. When an algorithm classifies a sample that is true in real life as positive, this is calculated in the TP cell. If a negative sample from real life is inaccurately classified as positive, it is added up in the FP area. If a positive sample from real life is classified as negative, this is calculated in the FN cell. Finally, if negative data is correctly classified as negative, it is estimated in the TN region. Accuracy, precision, recall and F1 score are estimated using TP, TN, FP and FN values according to the confusion matrix.

### 1.8.2 Accuracy

Accuracy is the most important metric in machine learning applications. High accuracy means a successful algorithm. It is the ratio of accurate estimates given by the algorithm to all estimates. An accuracy rate should be greater than 50% in a lie detection system to be better than chance. The following formula shows the estimation of accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (1.1)$$

### 1.8.3 Precision

Precision is the success of a condition which is classified as positive. The following formula shows the calculation of precision.

$$Precision = \frac{TP}{TP+FP} \quad (1.2)$$

### 1.8.4 Recall

Recall indicates how successful the positive states are classified. The following equation shows for the estimation of recall.

$$Recall = \frac{TP}{TP+FN} \quad (1.3)$$

### 1.8.5 F1 - Score

F1-Score is the harmonic average of the recall and precision. The next statement is to measure it.

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (1.4)$$

# CHAPTER 2

## RESEARCH METHODOLOGY

### 2.1 Dataset

The dataset was created in two stages. First, 4 women and 5 men uttered 5 true and 5 false statements. To create the fundamental dataset, 15 women and 15 men speakers took part in the study. These two datasets will be explained in the following sections.

#### 2.1.1 Pre-Experiment

In this section, 4 women and 5 men took part in the study. This dataset was created using only the speakers' voices. The speakers were given the set of questions before their voices were recorded. For the first recording, the participants responded truthfully to five questions. Afterward, the participants told lies for the same five questions. These questions are:

- 1) What are your name and surname?
- 2) What are your age and gender?
- 3) What is your job?
- 4) Do you like your job?
- 5) Are you satisfied with your salary?

After collecting the data, we extracted five formant frequencies of each voice record. However, there was no considerable difference between the results for lies and truth.

In this situation, the participants knew before the recordings started whether they would be expected to lie, which gave them time to prepare. Consequently, no stress or excitement was observed when the speakers' voices were examined. Thus, it was decided to collect data in a different way to introduce unexpectedness and unpreparedness.

### 2.1.2 Fundamental Dataset

If a person is not a professional liar or an actor, a change occurs in his/her voice and heart rate. Accordingly, it was decided to collect heart rate data while recording the participants' voices. To create the second dataset, 15 women and 15 men speakers whose ages varied 25 and 62 years participated in the second session. Furthermore, both their heart rate and sounds were collected.

There is a sequence diagram for creating the dataset.

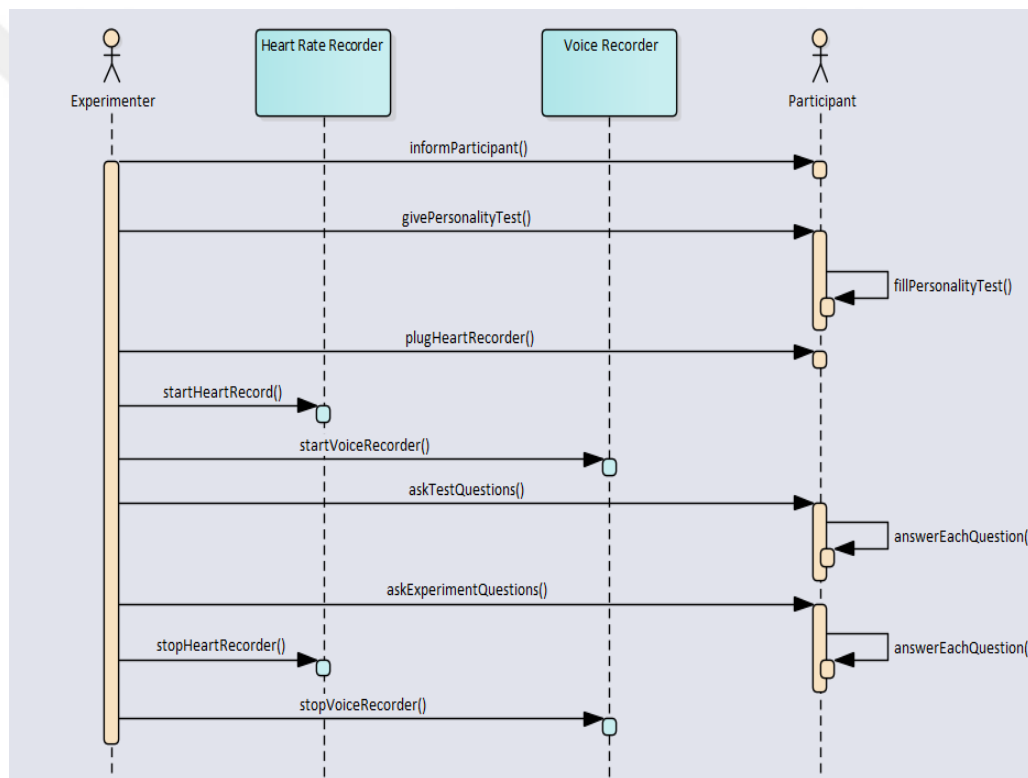


**Figure 2.1** Steps to create the dataset

Figure 2.1 shows steps to create fundamental dataset. There are two devices in this experiment: a heart rate recorder a voice recorder. First, the participant is informed about the experiment. Then a personality test is administered (See Appendix A). Then, the participant wears a wristband to record their heart rate. Next, the experimenter starts heart rate recorder and voice recorder. The experimenter asks test

questions and participant answer each question. Then, the experimenter asks the main experiment questions and participant answer each question. The experimenter stops the heart rate recorder and voice recorder. This experiment takes between 5 and 10 minutes.

#### 2.1.2.1 Collecting Voice Data

While collecting the voice recordings, an external microphone (Blue Microphones Snowball Ice) was used. This microphone provides a voice recording with minimal noise. For better sound quality, the speakers were set to a distance of up to approximately 20 cm. An example spectrum of the records can be seen in Figure 2.2. This spectrum shows a single record.

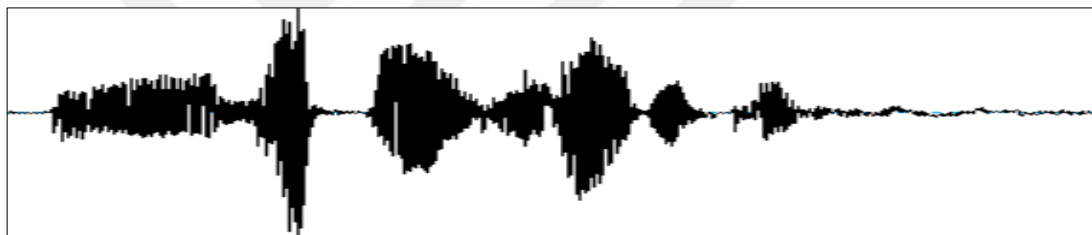


Figure 2.2 Example of getting spectrum

Each speaker had 10 voice recordings, each of which consisted of one question's answer. The dataset has a total of 300 voice recordings. Duration of these voice recordings varies between 2 to 35 seconds.

#### 2.1.2.2 Collecting Heart Rate Data

Speakers' heart rates were also recorded during data collection. In other words, the voice recordings and heart rate data were collected simultaneously. Heart rate can be affected by any stressful stimuli. This effect generally decreases while telling a lie in the literature [37]. To capture this phenomenon, heart rate was saved during the second experiment via a smart wristband (Veryfit Id107 Hr Plus). This device provided instant recording of heart rate. This wristband can be controlled from a mobile phone thanks to its mobile application. The participants put the wristband on

before they started to the experiment. Owing to remote control, they did not need to do anything by themselves to start the recording.

We have two type of heart rate graphs: fluctuated and fixed. Figure 2.3 shows the fluctuated type which is an example of the heart rate record graph taken from one of the participants.

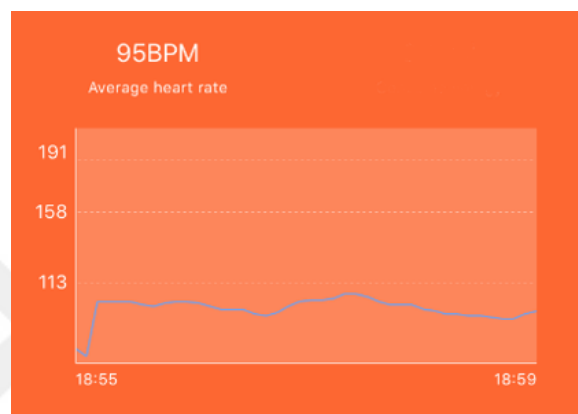


**Figure 2.3** Heart Rate Fluctuated Graph Example

As shown in the figure above, an average heart rate can be deduced directly from the graphic. It shows fluctuations at any time. While the speaker tells a lie, there is a decline in the heart rate. Each speaker has a graph similar to the one in the figure.

The following figure shows the fixed type.

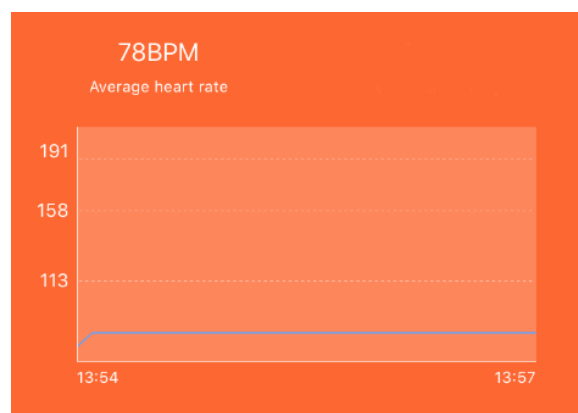


**Figure 2.4** Heart Rate Fixed Graph Example

Figure 2.4 represents a graph little change in the pulse graph for some participants, as in the graph. Three of all participants have a stable graph type. This type of graph is seen in people whose lie is difficult to classify. These can be considered skilled/effective liars (see 1.5.2).

Minimum, maximum and average heart rates were extracted by observing the graphs of each speaker for each question. As a result, one speaker has three heart rate values for each of the 10 questions as shown in Table 2.1.

**Table 2.1** Heart Rate List

| Question# | Length | HR                          | Min | Max | Avg |
|-----------|--------|-----------------------------|-----|-----|-----|
| 0         | 3      | 99                          | 99  | 99  | 99  |
| 1         | 3      | 100                         | 100 | 100 | 100 |
| 2         | 11     | 97-98-99-100                | 97  | 100 | 99  |
| 3         | 4      | 95-93-92                    | 92  | 95  | 94  |
| 4         | 24     | 91-94-96-98-100-101-102-103 | 91  | 103 | 97  |
| 5         | 3      | 105                         | 105 | 105 | 105 |
| 6         | 5      | 103-102-101-100             | 100 | 103 | 102 |
| 7         | 10     | 97-95-93-91                 | 91  | 97  | 94  |
| 8         | 5      | 90-89                       | 89  | 90  | 90  |

Table 2.1 shows question numbers, duration in seconds, heart rate values, and minimum, maximum and average heart rates in beat per minute. Average heart-rate values were rounded up because fractions have been neglected. Eventually, heart-rate values are prepared for the dataset.

### 2.1.2.3 Stages Creating Dataset

The participants answered five warm-up questions to become accustomed to the study environment before they answered the actual ten experiment questions. While answering all questions, a microphone and a smart wristband recorded their voice and heart rate. Unlike in the first experiment, the speakers did not know whether they would lie or tell the truth; thus, they had no time to prepare to lie. Two cards were created indicating truth or lie as shown in Figure 2.5.



**Figure 2.5** Lie and Truth Cards

When the participant was asked a question, they were also shown one of these cards. This was an attempt to create an instant lying situation because we did not want participants to have time to prepare. Five warm-up questions were posed before the experiment started. These test questions allowed participants to become familiar with the cards. These five questions are:

- 1) What is your name?
- 2) How old are you?
- 3) What is your gender?

- 4) What is your job?
- 5) Where is the capital of Turkey?

After this test, ten questions were asked to participants. These ten questions are:

- 1) What do you do when your phone rings while driving?
- 2) What do you do if you find money on the road?
- 3) What do you do when you go to an all-inclusive hotel?
- 4) What are your hobbies?
- 5) Tell me how you robbed the bank?
- 6) What do you use as a crime tool if you want to make a robbery?
- 7) We learned that you lie regularly. When was the last time you lied?
- 8) What is your last lie?
- 9) Your partner/friend received something new, but you do not like it. What is your comment?
- 10) If you win the lottery, what is the first thing you will do to change your life?

The questions were not related to the personal benefits of the participants. They were asked to respond instantaneously to these questions. Some questions implied a false scenario. The reactions in those cases were measured. When the participants heard the questions, they were surprised. An attempt was made to direct people for lying with these scenarios. Moreover, we selected these questions in such a way that people would not refrain from answering.

#### **2.1.2.4 Big Five Personality Test**

We believe that a participant's personality affects his physiological reactions while telling a lie. In order to test this, we checked each participant's personality with the Big Five Personality Test. The Big Five Personality Test was applied to all participants (See Appendix A). Although this test consists of 120 statements, in this study ten were selected to calculate the personality factor. These were chosen so in

order to save time. Each personality factor has two expressions for this study. These statements are:

- 1) I see myself as an introvert
- 2) I often see myself as a reliable person
- 3) I see myself as someone who tends to be lazy
- 4) I see myself as comfortable coping with stress
- 5) I see myself as less interested in art
- 6) I see myself as an outgoing, social person
- 7) I see myself as a person prone to finding others' mistakes
- 8) I see myself as a man who does his work conscientiously
- 9) I see myself as angry
- 10) I see myself as someone with an effective imagination

Possible options for each of the statements were strongly-disagree, disagree, neutral, agree and strongly-agree. These are enumerated from 1 to 5. These statements are classified as openness, conscientiousness, extraversion, agreeableness, neuroticism. Statements 5 and 10 are in openness. In addition, conscientiousness involves statements 8 and 3. Further, extraversion contains statements 6 and 1. Moreover, statements 2 and 7 pertain to agreeableness. Finally, neuroticism comprises statements 9 and 4. The score is calculated with using factors' statements values. The score is calculated with the following equations:

$$\textit{Openness} = \textit{statement5} - \textit{statement10}$$

$$\textit{Conscientiousness} = \textit{statement8} - \textit{statement3}$$

$$\textit{Extraversion} = \textit{statement6} - \textit{statement1}$$

$$\textit{Agreeableness} = \textit{statement2} - \textit{statement7}$$

$$\textit{Neuroticism} = \textit{statement9} - \textit{statement4}$$

According to the results of these equations, personality factor is determined as open, conscientious, extraverted, agreeable and neurotic. Their first letters are added to the file as personality result.

### 2.1.2.5 Compiling Data Files of the Experiment

A Comma Separated Values (CSV) file was created for each speaker and a table which is collected data was obtained for each participant. Each CSV file contains speakers' features.

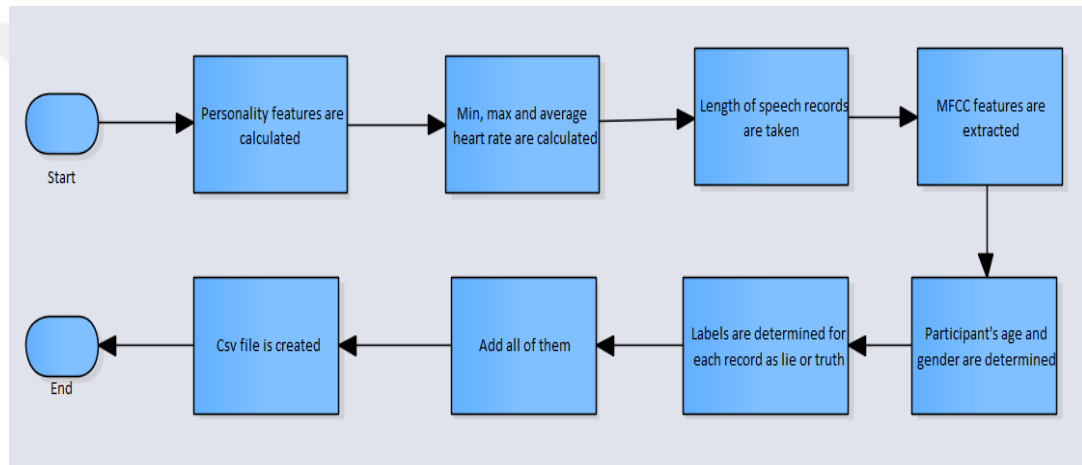


**Figure 2.6** Steps of creating the dataset file

First, each participant was enumerated from 1 to 30. Second, questions were numbered from 0 to 9. Third, the speaker's gender and personality types were added to the file. Next, as demonstrated in Table 2.2, extracted 12 MFCC values were added to the CSV file. MFCC values are rounded up to demonstrate in this table.

**Table 2.2** Example of Extracted 12 MFCC

| Mfcc1    | Mfcc2   | Mfcc3  | Mfcc4  | Mfcc5  | Mfcc6  | Mfcc7  | Mfcc8  | Mfcc9  | Mfcc10 | Mfcc11 | Mfcc12 |
|----------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -495.670 | 133.498 | 39.723 | 35.635 | 39.072 | 45.962 | 45.667 | 39.065 | 24.401 | 29.058 | 28.675 | 34.914 |
| -467.798 | 139.168 | 23.525 | 34.109 | 35.061 | 47.953 | 46.754 | 39.414 | 20.936 | 24.656 | 26.240 | 29.982 |

**Table 2.2 (continued):** Example of Extracted 12 MFCC

|          |         |        |        |        |        |        |        |        |        |        |        |
|----------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -533.677 | 125.692 | 49.089 | 45.348 | 42.539 | 53.907 | 49.180 | 43.459 | 29.358 | 30.805 | 27.748 | 36.552 |
| -498.856 | 127.090 | 39.464 | 42.969 | 40.870 | 46.914 | 47.536 | 38.200 | 26.150 | 27.665 | 27.378 | 34.619 |
| -427.122 | 110.221 | 39.127 | 40.878 | 35.932 | 38.697 | 39.841 | 33.139 | 20.780 | 18.771 | 23.266 | 26.469 |
| -571.998 | 133.563 | 59.759 | 45.458 | 42.911 | 56.782 | 49.329 | 44.824 | 30.845 | 35.779 | 32.529 | 40.218 |
| -462.872 | 130.467 | 39.137 | 40.067 | 37.666 | 49.775 | 38.968 | 33.180 | 22.662 | 20.643 | 19.819 | 30.488 |
| -497.795 | 141.503 | 49.980 | 41.481 | 36.113 | 48.457 | 48.003 | 36.908 | 20.961 | 21.308 | 23.927 | 29.154 |
| -502.060 | 118.624 | 39.519 | 55.017 | 34.143 | 52.230 | 47.974 | 41.321 | 29.883 | 28.630 | 22.121 | 32.598 |
| -504.513 | 110.827 | 39.234 | 55.956 | 44.680 | 55.558 | 45.315 | 36.684 | 28.899 | 23.910 | 25.909 | 37.542 |

Duration of speech was also added to the file. Then, minimum, maximum and average heart rate were appended to the file for each voice record. The final data field is for each voice record labeled as 0 or 1. Zero represents the lie and 1 represents the truth. As a result, a CSV file consists of speaker number, question number, gender, personality types, 12 MFCC values, duration of speech, minimum HR, maximum HR, average HR values and labels.

Finally, 19 CSV files are created as the following list:

- All features are gender, personality type, MFCC values, duration of speech, minimum heart rate, maximum heart rate, age.
- Heart rate values
- Heart rate values and age
- Heart rate values, age and gender
- Heart rate values, age, gender and personality types

- Heart rate and gender
- Heart rate and personality
- Heart rate and MFCC values
- Heart rate, MFCC and age values
- Heart rate, MFCC, age and gender
- Heart rate, MFCC, age, gender and personality
- Heart rate, MFCC and gender
- Heart rate, MFCC and personality
- MFCC values
- MFCC and age values
- MFCC, age and gender
- MFCC, age, gender and personality
- MFCC and gender
- MFCC and personality

These files are used to compare the results of each method. For each file, test and train data are selected with ten different random states and this is applied for all methods. Different random states provide a selection of different part of data from the dataset for testing and training. Each random state's accuracy results are found and the best result is selected.

## 2.2 Methods

Feature extraction is the most important step in developing machine learning applications. The existing data are divided into training and testing data, which are used to develop a predictive model. Machine learning steps are as follows:

- 1) Feature selection,
- 2) Feature extraction,
- 3) Classification

### 2.2.1 Feature Extraction (MFCC)

MFCC is a cepstral method which converts speech into parameters according to Mel-scale [38]. In addition to the sound information, the audio signal also includes information that can be used to classify age, gender and emotional state. The main aim of feature extraction is to facilitate detection by summarizing sound data. The most common feature extraction method is MFCC for sound analysis processes. The sound range of human hearing varies between 20 and 20,000 Hz. MFCC is based on the human voice detection mechanism. MFCC is used for both feature extraction and speaker recognition. The formula is shown in the following equation:

$$Mel(f) = 2595 * \log \left( 1 + \frac{f}{700} \right) \quad (2.1)$$

Specific points of each voice are found by using Mel frequency and these points give cepstrum coefficients. Extraction of Mel frequency cepstrum coefficients steps is shown in the following figure.

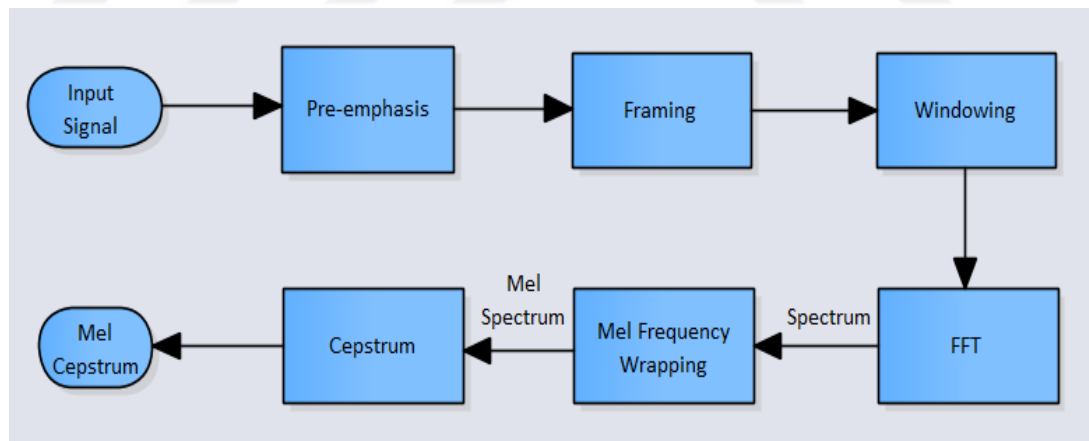


**Figure 2.7** MFCC Block Diagram [38]

As shown in Figure 2.7, the signal is received as input. Then, MFCC performs six steps: pre-emphasis, framing, windowing, fast fourier transform (FFT), Mel frequency wrapping and cepstrum. As a result, Mel spectrum is achieved as output.

We used Librosa, an audio and music processing library in Python, to extract attributes. Librosa's load method requires the audio file as a parameter. Thanks to this library, waveform and sampling rate were achieved. In our study, we used Librosa's feature method which was used for feature extraction and manipulation for MFCCs extraction.

```
1. librosa.feature.mfcc(y=y, sr=sr, n_mfcc=12)
```

As shown in the above code block, *y* is the raw signal [39]; *sr* is sample rate. In addition, *n\_mfcc* represents the number of MFCC and the number of coefficients of MFCC selected as 12. The output of this method is the matrix MFCC.

## 2.2.2 Classification Methods

A variety of classification methods are used in machine learning applications. Some of the most commonly used are support vector machines, logistic regression, decision trees and artificial neural networks.

### 2.2.2.1 Support Vector Machine

Support vector machines are a supervised learning method that tries to find the best line to separate two different classes. SVMs are composed of positive and negative vectors.

The method, which is used to classify both linear and non-linear data, is very easy to implement, and requires no parameters. This algorithm is often preferred because of its high accuracy. Besides, there is no overfitting problem in the algorithm. Classifying non-linear data in support vector machines is more difficult than classifying linear data [9].

Kernel functions are used to classify nonlinear data. One of the common kernel functions is the radial basis function kernel (RBF kernel), also called Gaussian

kernel. The following formula represents RBF kernel. Parameters of the kernel function are  $\sigma$  and C [10].

$$K(x,y) = \exp \left( - \frac{\|x - z\|^2}{2\sigma^2} \right) \quad (2.2)$$

In our study, the positive vectors are labeled as 0 for the lie and the negative vectors are labeled as 1 for truth. We consider radial basis function as the kernel function. Gamma value is not specified, and we used the default values of the function. Gamma default is auto which is used as follows:

$$\frac{1}{n_{features}} \quad (2.3)$$

C is a regularization parameter which controls changing between the complexity of the decision function and the experimental risk. It is also called as penalty parameter of the error term and is selected as 1.0 in our study. In addition, we used the sklearn SVM library for application of the SVM method.

#### 2.2.2.2 Logistic Regression

Logistic regression is one of the supervised learning methods. This algorithm classifies the probability of a result with two categories. The method aims to make parameter classifications according to the logistic models created. Logistic regression produces classified logistic values between 0 and 1. Although the logistic method is quite easy to implement, there should not be lost data as in decision trees.

In our study, there is no specified parameter for this method. We used Python's sklearn linear\_model library to apply the logistic regression method.

#### 2.2.2.3 Decision Tree

Decision trees are a supervised learning method [14]. Decision trees, a commonly used algorithm, can be used for every data type. They are used to divide data sets containing large numbers of records into smaller pieces. This method is easy to understand and implement and can work with small amounts of data. Moreover, this

method can be used to process both numeric and categorical data. The important thing is that there be no data loss since decision trees do not support data loss. As well as the mentioned advantages to this method, there is a disadvantage: overfitting. Techniques such as pruning can also be used to solve the disadvantage. There are different algorithms in decision trees. Entropy, the most commonly preferred, is used for categorical variables. Entropy, a measure of uncertainty about data, is calculated with the following equations:

$$H(T) = - \sum p(x) \log p(x) \quad (2.4)$$

In this equation, indicated  $p(x)$  represents the percentage of the group belonging to a class. While creating a decision tree, the gain is measured for each attribute using the entropy formula. A decision tree is obtained by finding the category with the lowest gain.

In our study, we used entropy as the type of decision tree algorithm because most features had better accuracy in entropy algorithm in our experiments. Also, we used Python's sklearn tree library and DecisionTreeClassifier method to execute the algorithm.

#### **2.2.2.4 Artificial Neural Network**

Artificial Neural Network (ANN) is based on the structure and operation of nerve cells in the human brain. The artificial neural network consists of neurons and their connections to each other. ANN includes inputs, activation functions and outputs as shown in Figure 2.8. Inputs are first multiplied by weights. Then, the activation function, which is a mathematical function, determines the activation of the neuron [13].

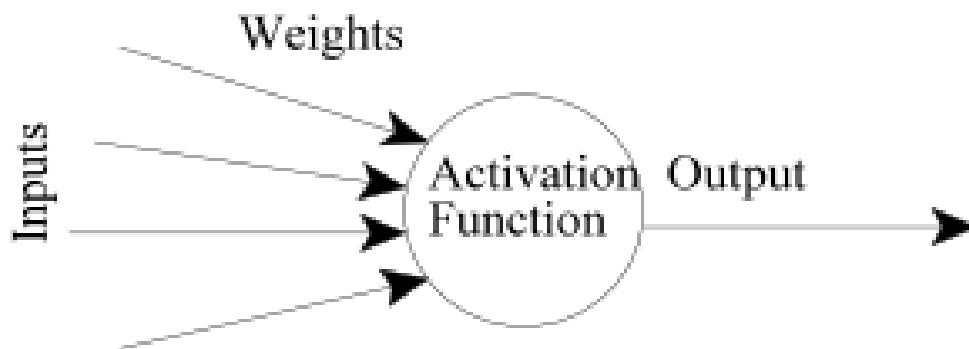


**Figure 2.8** A simple artificial neural network [13]

In this study, one input layer, two hidden layers and one output layer in the artificial neural network are used. There are 19 files to apply the algorithm, all of whose parameters are different.

Because of this reason, we determined a common structure for the application of ANN. Parameters are shown as follows:

- 1) Input Layer: last\_column\_number-1
- 2) Hidden Layer 1
  - a. Number of neurons: last\_column\_number-1
  - b. Activation function: Rectified Linear Unit Function (Relu)
- 3) Hidden Layer 2
  - a. Number of neurons: (last\_column\_number-1) / 2
  - b. Activation function: Relu
- 4) Output Layer
  - a. Number of nodes: 1
  - b. Activation function: Sigmoid

Features represented with the following descriptions:

- Hr: Minimum, maximum and average heart rate values
- HrAge: Minimum, maximum, average heart rate and age values

- HrGender: Minimum, maximum, average heart rate and gender values
- HrPersonality: Minimum, maximum, average heart rate and personality type values
- HrAgeGender: Minimum, maximum, average heart rate and gender values
- HrAgeGenderPers: Minimum, maximum, average heart rate, age, gender and personality values
- Mfcc: 12 Mel Frequency Cepstrum Coefficients
- MfccAge: Mfcc and age values
- MfccGender: Mfcc and gender values
- MfccPersonality: Mfcc and personality type values
- MfccAgeGender: Mfcc, age and gender values
- MfccAgeGenderPers: Mfcc, age, gender, personality type values
- HrMfcc: Minimum, maximum, average heart rate and Mfcc values
- HrMfccAge: Minimum, maximum, average heart rate, Mfcc and age values
- HrMfccGender: Minimum, maximum, average heart rate, Mfcc and gender values
- HrMfccPersonality: Minimum, maximum, average heart rate, Mfcc and personality type values
- HrMfccAgeGender: Minimum, maximum, average heart rate, Mfcc, age and gender values
- HrMfccAgeGenderPers: Minimum, maximum, average heart rate, Mfcc, age, gender and personality values
- All: Gender, personality type, MFCC, duration of speech, minimum, maximum, average heart rates, and age values

Table 2.3 shows that mentioned last\_column\_number values for each file.

**Table 2.3** Last Column Numbers

| Features | Features Size |
|----------|---------------|
| Hr       | 3             |
| HrAge    | 4             |

**Table 2.3 (continued):** Last Column Numbers

|                     |    |
|---------------------|----|
| HrGender            | 4  |
| HrPersonality       | 4  |
| HrAgeGender         | 5  |
| HrAgeGenderPers     | 6  |
| Mfcc                | 12 |
| MfccAge             | 13 |
| MfccGender          | 13 |
| MfccPersonality     | 13 |
| MfccAgeGender       | 14 |
| MfccAgeGenderPers   | 15 |
| HrMfcc              | 15 |
| HrMfccAge           | 16 |
| HrMfccGender        | 16 |
| HrMfccPersonality   | 16 |
| HrMfccAgeGender     | 17 |
| HrMfccAgeGenderPers | 18 |
| All                 | 19 |

We used the Adam algorithm as an optimization function for application of all dataset to ANN. In addition, we used binary cross-entropy as a loss function. Accuracy was selected as the evaluation metric. We selected 100 as the number of epoch and 20 as batch size. Finally, 0 represents a lie and 1 represents a truth as output.

# CHAPTER 3

## APPLICATION & EXPERIMENTAL RESULTS

### 3.1 Application

In this section, the development environment, development language, and its methods will be discussed. In addition, steps of application of mentioned four machine learning algorithms will be explained.

#### 3.1.1 Development Environment and Methods

PyCharm and Spyder editors are used as development environment [40]. Firstly, PyCharm is used to create the dataset. Spyder, a scientific python development environment, was used because it provided a graphical interface for outputs for execution of methods.

Python language is used as the programming language. Python 3.5 is selected as the python version. Python has many libraries for machine learning applications and can be applied quickly and easily. Pandas, Librosa, Keras, and Sklearn libraries are used in this study. Panda is used for reading and writing CSV files. Another library, Librosa extracts MFCC features from a sound file according to the number of coefficient parameters. The next library is Keras, which has packages for classification of an artificial neural network. Sklearn, the last and the most important library, supplies the most significant packages for this study. These packages are preprocessing, model\_selection, metrics, SVM, tree and linear\_model methods. Thanks to these libraries and their packages, feature extraction, selection, model fitting and evaluation can be done easily.

### 3.1.2 Application of Learning Methods

In this study, lies are detected using sound and heart rate for four different algorithms, the results of which are compared. This study consists of six steps for each algorithm as shown in Figure 3.1.

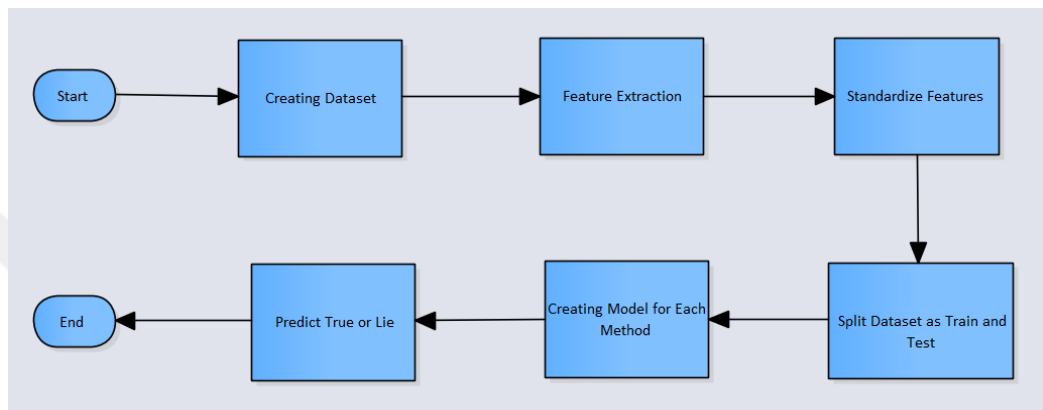


**Figure 3.1** Application Flow Diagram

The first step is creating the dataset as in Figure 3.1, which was mentioned in the previous chapter. The next step is feature extraction. Twelve features are extracted using MFCC. In addition, personality test results are calculated to determine the personality type of each participant as openness, conscientiousness, extraversion, agreeableness and neuroticism. The pulse data collected with each voice recording is taken as minimum, maximum and average. Gender, age and duration of speech are likewise added to the dataset. The next step is standardization of the features. That is, personality type and gender values are not a numeric value, but other features are numeric. In this situation, features need to be standardized to compare all features. The dataset is then separated as training and testing. In this study, this separation is made in two different shapes and their results are compared. First, the training set is selected as 75% of all dataset. Second, the training set is chosen as 80% of all dataset. The following step is the creation of a model for each method with these train and test data. Then, lie or truth is classified according to the model. Finally, a confusion matrix is created for each model and accuracy, precision, recall and F1-score are estimated.

The pseudo code of learning application is as follows:

```
1.
2. BEGIN
3.
4. READ dataset
5. SET max accuracy to 0
6. SET test size to 0.20 or 0.25
7.
8. specify input and output of dataset
9. Encode categorical data
10.
11. FOR iteration = 0 to 10
12.     split the dataset into the training set and test set
13.
14.     Scale feature
15.
16.     Case based on Model
17.     Case "SVM"
18.         Fitting SVM with parameters to the Training Set
19.
20.     Case "DT"
21.         Fitting DT with parameters to the Training Set
22.
23.     Case "LR"
24.         Fitting LR with parameters to the Training Set
25.
26.     Case "ANN"
27.         Initialize the ANN
28.         ADD input layer and first hidden layer
29.         ADD the second hidden layer
30.         ADD the output layer
31.         Compile the ANN
32.         Fit the ANN to the training set
33.     END Case
34.
35.     Make classifications
36.     Evaluate the model
37.
38.     IF accuracy > max accuracy
39.         SET max accuracy to accuracy
40.     END IF
41.
42. END FOR
43.
44. Print results to the console
45. Print results to the file.
46.
47. END
48.
```

According to this pseudo code, all steps are executed in all methods. This pseudo code is also applied to all datasets.

## 3.2 Experimental Results

Four different algorithms were applied in the scope of this study. There are result tables for each algorithm. Each table includes 19 feature sets of results. As shown in the following four tables, there are nine columns. The first column contains taken features names. The next column shows the features' size, followed by the method name and the number of iterations. For each feature set, ten different random states are executed. The number of iterations that give the best results from the outputs for every 10 iterations is written in this column. Test data set are selected according to two values: 20% and 25% of the whole dataset. That is, the training data sets are 80% and 75% respectively for these test sizes. These values are shown in the next column which is named test size. The same features yield a different result with different test sizes. The next columns involve accuracy, precision, recall and F1-score results.

### 3.2.1 Support Vector Machine Results

There is a result table for support vector machine algorithms as well as for heart rate feature sets, MFCC feature sets and combining of heart rates and MFCC feature sets.

Figure 3.1 indicates the application of support vector machine algorithm results for each feature set. As shown in Figure 3.1, the best accuracy result is 61.67% with 12 MFCC values and 20% test size. The second-best result is 60% for four feature sets. The first set contains 12 MFCC and personality types. The second set involves 12 MFCC and age values. The other consists of 12 MFCC, minimum heart rate, maximum heart rate, average heart rate and age values. The last but most important set is all features which have 12 MFCC, minimum, maximum, and average heart rates, gender, age, personality type and duration of speech values.

For this algorithm personality type, age and heart rate values do not contribute to improved accuracy. In fact, these features reduce the accuracy rate by 1.67%. The lowest accuracy rate is 51.67% by using 16 features that consist of 12 MFCC, minimum, maximum and average heart rate and age. Another feature set which includes 12 MFCC and gender also reflects the lowest accuracy.

**Table 3.1** Results of Support Vector Machines

| Features                   | Features Size | Method | #Iteration | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|------------|-----------|----------|-----------|--------|----------|
| All                        | 19            | SVM    | 1          | 0.20      | 56.67    | 55.26     | 56.67  | 54.86    |
| All                        | 19            | SVM    | 1          | 0.25      | 60.00    | 59.42     | 59.42  | 59.42    |
| Hr                         | 3             | SVM    | 2          | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| Hr                         | 3             | SVM    | 2          | 0.25      | 57.33    | 28.67     | 50.00  | 36.44    |
| HrAgeGenderPersonality     | 6             | SVM    | 2          | 0.20      | 58.33    | 54.46     | 51.14  | 43.16    |
| HrAgeGenderPersonality     | 6             | SVM    | 2          | 0.25      | 57.33    | 53.87     | 50.80  | 41.52    |
| HrAgeGender                | 5             | SVM    | 2          | 0.20      | 58.33    | 54.46     | 51.14  | 43.16    |
| HrAgeGender                | 5             | SVM    | 1          | 0.25      | 53.33    | 50.29     | 50.22  | 48.01    |
| HrAge                      | 4             | SVM    | 2          | 0.20      | 56.67    | 28.81     | 48.57  | 36.17    |
| HrAge                      | 4             | SVM    | 1          | 0.25      | 57.33    | 78.38     | 51.52  | 39.15    |
| HrGender                   | 4             | SVM    | 2          | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrGender                   | 4             | SVM    | 2          | 0.25      | 57.33    | 28.67     | 50.00  | 36.44    |
| HrMfccAgeGenderPersonality | 18            | SVM    | 6          | 0.20      | 53.33    | 51.44     | 51.18  | 49.76    |
| HrMfccAgeGenderPersonality | 18            | SVM    | 3          | 0.25      | 54.67    | 54.05     | 53.75  | 53.26    |
| HrMfccAgeGender            | 17            | SVM    | 3          | 0.20      | 53.33    | 52.29     | 52.19  | 52.00    |
| HrMfccAgeGender            | 17            | SVM    | 3          | 0.25      | 56.00    | 55.49     | 55.00  | 54.41    |
| HrMfccAge                  | 16            | SVM    | 3          | 0.20      | 51.67    | 50.36     | 50.34  | 49.99    |

**Table 3.1 (continued):** Results of Support Vector Machines

|                          |    |     |   |      |       |       |       |       |
|--------------------------|----|-----|---|------|-------|-------|-------|-------|
| HrMfccAge                | 16 | SVM | 3 | 0.25 | 60.00 | 59.75 | 59.29 | 59.12 |
| HrMfccGender             | 16 | SVM | 1 | 0.20 | 55.00 | 54.00 | 53.96 | 53.96 |
| HrMfccGender             | 16 | SVM | 3 | 0.25 | 53.33 | 52.45 | 52.14 | 51.11 |
| HrMfccPersonality        | 16 | SVM | 3 | 0.20 | 55.00 | 54.69 | 54.71 | 54.69 |
| HrMfccPersonality        | 16 | SVM | 3 | 0.25 | 57.33 | 56.93 | 56.61 | 56.40 |
| HrMfcc                   | 15 | SVM | 3 | 0.20 | 55.00 | 54.41 | 54.38 | 54.38 |
| HrMfcc                   | 15 | SVM | 3 | 0.25 | 53.33 | 52.55 | 52.32 | 51.65 |
| HrPersonality            | 4  | SVM | 2 | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| HrPersonality            | 4  | SVM | 2 | 0.25 | 57.33 | 28.67 | 50.00 | 36.44 |
| MfccAgeGenderPersonality | 15 | SVM | 3 | 0.20 | 56.67 | 55.56 | 54.21 | 52.38 |
| MfccAgeGenderPersonality | 15 | SVM | 3 | 0.25 | 56.00 | 56.09 | 54.11 | 50.98 |
| MfccAgeGender            | 14 | SVM | 3 | 0.20 | 58.33 | 57.92 | 55.72 | 53.69 |
| MfccAgeGender            | 14 | SVM | 3 | 0.25 | 57.33 | 57.86 | 55.54 | 52.90 |
| MfccAge                  | 13 | SVM | 6 | 0.20 | 60.00 | 60.00 | 57.58 | 56.04 |
| MfccAge                  | 13 | SVM | 6 | 0.25 | 60.00 | 59.13 | 57.47 | 56.60 |
| MfccGender               | 13 | SVM | 3 | 0.20 | 51.67 | 50.69 | 50.67 | 50.55 |
| MfccGender               | 13 | SVM | 3 | 0.25 | 54.67 | 54.34 | 52.50 | 47.92 |
| MfccPersonality          | 13 | SVM | 6 | 0.20 | 56.67 | 55.54 | 54.55 | 53.35 |
| MfccPersonality          | 13 | SVM | 6 | 0.25 | 60.00 | 59.00 | 58.12 | 57.83 |
| Mfcc                     | 12 | SVM | 6 | 0.20 | 61.67 | 61.51 | 59.76 | 59.11 |
| Mfcc                     | 12 | SVM | 6 | 0.25 | 57.33 | 56.49 | 56.39 | 56.40 |

The graph of support vector machines results can be seen in Figure 3.2. The figure shows features size and their accuracy results with 20% test size. This graph shows any incremental change. These results are not consistent.

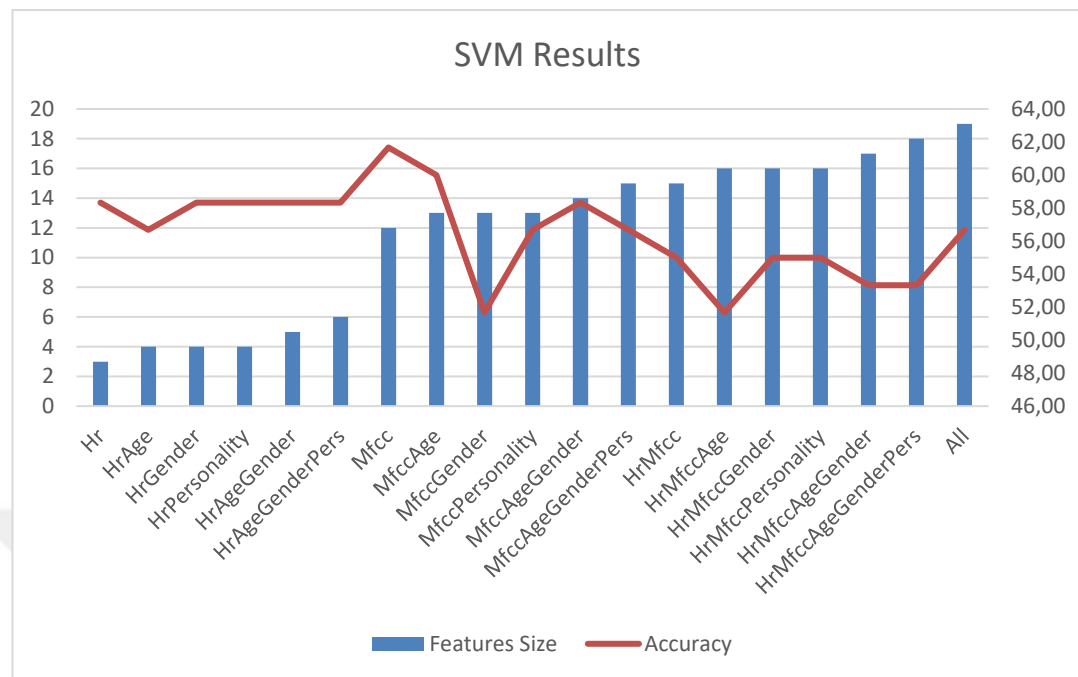


**Figure 3.2** The Graph of SVM Results

There are six different results in Table 3.2. Heart rate features consist of minimum, maximum and average heart rates. Their accuracy result is 58,33% and thus, H3 which is mentioned in 1.2 is proved according to SVM algorithm because heart rate values are insufficient to detect lies successfully. Similar to heart rate accuracy result, heart rate and gender features have 58.33% accuracy result. Hence, H5 is not proved. In addition, heart rate and personality features have the same results, which shows that H6 is not proved. As shown in Table 3.2, accuracy results are only affected by the age factor since there are a lot of different age values. Thus, H4 is proved, but it has a negative effect.

**Table 3.2** SVM Heart Rate Results

| Features | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------|---------------|--------|-----------|----------|-----------|--------|----------|
| Hr       | 3             | SVM    | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrAge    | 4             | SVM    | 0.20      | 56.67    | 28.81     | 48.57  | 36.17    |

**Table 3.2 (continued): SVM Heart Rate Results**

|                        |   |     |      |       |       |       |       |
|------------------------|---|-----|------|-------|-------|-------|-------|
| HrGender               | 4 | SVM | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| HrPersonality          | 4 | SVM | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| HrAgeGender            | 5 | SVM | 0.20 | 58.33 | 54.46 | 51.14 | 43.16 |
| HrAgeGenderPersonality | 6 | SVM | 0.20 | 58.33 | 54.46 | 51.14 | 43.16 |

MFCC results can be seen in Table 3.3 for the support vector machine algorithm. According to the table, H2 is proved because of only MFCC feature set has 61.67% accuracy. The table shows that age, gender and personality influence accuracy, precision, recall and F1-score outcomes. Thus, H4, H5 and H6 are not proved because age, gender and personality have a negative impact. Since MFCC contain more attributes than the heart rate values, it has been found to be effective in changing the results as the attributes increase.

**Table 3.3 SVM MFCC Results**

| Features                 | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|--------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Mfcc                     | 12            | SVM    | 0.20      | 61.67    | 61.51     | 59.76  | 59.11    |
| MfccAge                  | 13            | SVM    | 0.20      | 60.00    | 60.00     | 57.58  | 56.04    |
| MfccGender               | 13            | SVM    | 0.20      | 51.67    | 50.69     | 50.67  | 50.55    |
| MfccPersonality          | 13            | SVM    | 0.20      | 56.67    | 55.54     | 54.55  | 53.35    |
| MfccAgeGender            | 14            | SVM    | 0.20      | 58.33    | 57.92     | 55.72  | 53.69    |
| MfccAgeGenderPersonality | 15            | SVM    | 0.20      | 56.67    | 55.56     | 54.21  | 52.38    |

Table 3.4 shows heart rate values and MFCC results with test size 20%. As mentioned in Table 3.2, gender and personality type have no impact on results in Table 3.4. Therefore, H5 and H6 are not proved. In addition, age affects the results, but it reduces accuracy, precision, recall and F1-scores; thus, H4 is not proved. According to this table, when the accuracy rate is compared with Table 3.2 and Table

3.3, mentioned H1 is not proved because the accuracy rate decreases by 3.88% and 6.67% according to the previous two tables.

**Table 3.4** SVM Heart Rate and MFCC Results

| Features                   | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| HrMfcc                     | 15            | SVM    | 0.20      | 55.00    | 54.41     | 54.38  | 54.38    |
| HrMfccAge                  | 16            | SVM    | 0.20      | 51.67    | 50.36     | 50.34  | 49.99    |
| HrMfccGender               | 16            | SVM    | 0.20      | 55.00    | 54.00     | 53.96  | 53.96    |
| HrMfccPersonality          | 16            | SVM    | 0.20      | 55.00    | 54.69     | 54.71  | 54.69    |
| HrMfccAgeGender            | 17            | SVM    | 0.20      | 53.33    | 52.29     | 52.19  | 52.00    |
| HrMfccAgeGenderPersonality | 18            | SVM    | 0.20      | 53.33    | 51.44     | 51.18  | 49.76    |

### 3.2.2 Logistic Regression Results

There is another results table for the logistic regression algorithm. This section gives results for all feature sets and accuracy result graphs of the algorithm.

The best accuracy is 66.67%, with 19 features and test size of 25% of the whole dataset in Table 3.5. The next best accuracy result is 63.33%, with the same features and test size of 20%. Another feature set which consists of MFCC, minimum, maximum and average heart rates, gender and personality type has the lowest accuracy with 53.33%.

**Table 3.5** Results of Logistic Regression

| Features | Features Size | Method | #Iteration | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------|---------------|--------|------------|-----------|----------|-----------|--------|----------|
| All      | 19            | LR     | 7          | 0.20      | 63.33    | 64.51     | 63.74  | 62.96    |
| All      | 19            | LR     | 4          | 0.25      | 66.67    | 64.00     | 63.30  | 63.53    |

**Table 3.5 (continued):** Results of Logistic Regression

|                            |    |    |   |      |       |       |       |       |
|----------------------------|----|----|---|------|-------|-------|-------|-------|
| Hr                         | 3  | LR | 2 | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| Hr                         | 3  | LR | 2 | 0.25 | 57.33 | 28.67 | 50.00 | 36.44 |
| HrAgeGenderPersonality     | 6  | LR | 2 | 0.20 | 55.00 | 45.37 | 48.29 | 41.28 |
| HrAgeGenderPersonality     | 6  | LR | 2 | 0.25 | 57.33 | 56.18 | 53.14 | 48.54 |
| HrAgeGender                | 5  | LR | 2 | 0.20 | 58.33 | 54.46 | 51.14 | 43.16 |
| HrAgeGender                | 5  | LR | 2 | 0.25 | 57.33 | 54.23 | 52.00 | 47.09 |
| HrAge                      | 4  | LR | 2 | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| HrAge                      | 4  | LR | 2 | 0.25 | 57.33 | 53.77 | 50.40 | 39.15 |
| HrGender                   | 4  | LR | 2 | 0.20 | 60.00 | 63.16 | 52.57 | 44.10 |
| HrGender                   | 4  | LR | 2 | 0.25 | 61.33 | 65.86 | 55.49 | 49.76 |
| HrMfccAgeGenderPersonality | 18 | LR | 2 | 0.20 | 60.00 | 58.24 | 57.71 | 57.70 |
| HrMfccAgeGenderPersonality | 18 | LR | 2 | 0.25 | 53.33 | 51.00 | 50.91 | 50.48 |
| HrMfccAgeGender            | 17 | LR | 2 | 0.20 | 60.00 | 58.24 | 57.71 | 57.70 |
| HrMfccAgeGender            | 17 | LR | 9 | 0.25 | 54.67 | 54.07 | 54.02 | 54.00 |
| HrMfccAge                  | 16 | LR | 6 | 0.20 | 61.67 | 61.25 | 60.10 | 59.78 |
| HrMfccAge                  | 16 | LR | 6 | 0.25 | 61.33 | 60.48 | 59.63 | 59.49 |
| HrMfccGender               | 16 | LR | 2 | 0.20 | 58.33 | 56.58 | 56.29 | 56.28 |
| HrMfccGender               | 16 | LR | 9 | 0.25 | 57.33 | 56.96 | 56.96 | 56.96 |
| HrMfccPersonality          | 16 | LR | 6 | 0.20 | 60.00 | 59.34 | 58.59 | 58.33 |
| HrMfccPersonality          | 16 | LR | 2 | 0.25 | 61.33 | 65.86 | 55.49 | 49.76 |
| HrMfcc                     | 15 | LR | 6 | 0.20 | 58.33 | 57.54 | 57.07 | 56.88 |
| HrMfcc                     | 15 | LR | 6 | 0.25 | 60.00 | 59.11 | 58.77 | 58.76 |
| HrPersonality              | 4  | LR | 1 | 0.20 | 56.67 | 28.33 | 50.00 | 36.17 |

**Table 3.5 (continued):** Results of Logistic Regression

|                          |    |    |   |      |       |       |       |       |
|--------------------------|----|----|---|------|-------|-------|-------|-------|
| HrPersonality            | 4  | LR | 1 | 0.25 | 56.00 | 28.00 | 50.00 | 35.90 |
| MfccAgeGenderPersonality | 15 | LR | 2 | 0.20 | 56.67 | 54.58 | 54.29 | 54.17 |
| MfccAgeGenderPersonality | 15 | LR | 7 | 0.25 | 57.33 | 58.00 | 57.11 | 56.01 |
| MfccAgeGender            | 14 | LR | 2 | 0.20 | 58.33 | 56.58 | 56.29 | 56.28 |
| MfccAgeGender            | 14 | LR | 9 | 0.25 | 58.67 | 58.02 | 57.68 | 57.58 |
| MfccAge                  | 13 | LR | 6 | 0.20 | 60.00 | 59.34 | 58.59 | 58.33 |
| MfccAge                  | 13 | LR | 6 | 0.25 | 58.67 | 57.54 | 56.93 | 56.70 |
| MfccGender               | 13 | LR | 2 | 0.20 | 56.67 | 54.99 | 54.86 | 54.86 |
| MfccGender               | 13 | LR | 7 | 0.25 | 58.67 | 59.68 | 58.43 | 57.17 |
| MfccPersonality          | 13 | LR | 6 | 0.20 | 60.00 | 59.34 | 58.59 | 58.33 |
| MfccPersonality          | 13 | LR | 6 | 0.25 | 60.00 | 59.00 | 58.12 | 57.83 |
| Mfcc                     | 12 | LR | 6 | 0.20 | 61.67 | 61.51 | 59.76 | 59.11 |
| Mfcc                     | 12 | LR | 6 | 0.25 | 61.33 | 60.54 | 59.31 | 58.97 |

As seen in Figure 3.3, there is not a regular increase or decrease in the accuracy results, but there is an overall increase. The graph shows that the maximum accuracy is 63.33% because the test size is 20%.

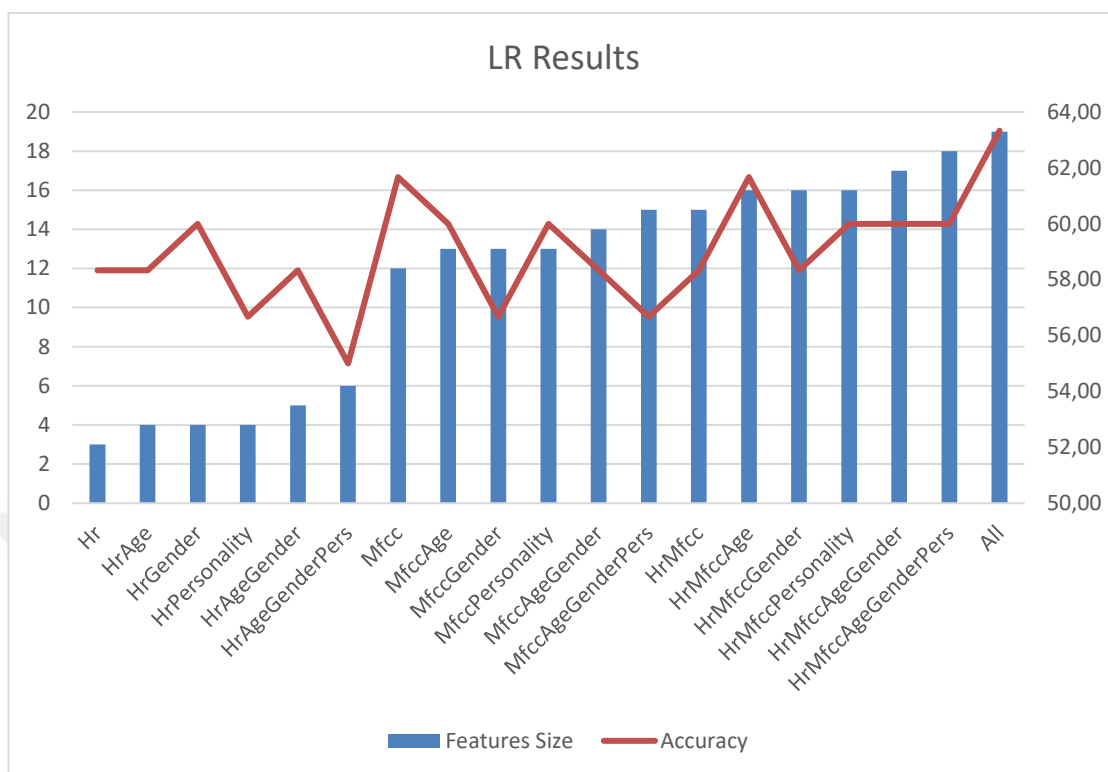


**Figure 3.3** The Graph of LR Results

As shown in Table 3.6, heart rate has 58.33% accuracy, but it is insufficient and H3 is proved. When age is added to heart rate values, there is no rise in the accuracy rate in this algorithm. Consequently, H4 is refuted. However, the addition of gender increases by 1.67% and H5 is verified. Also, personality type decreases accuracy rate and H6 is denied.

**Table 3.6** Logistic Regression Heart Rate Results

| Features | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------|---------------|--------|-----------|----------|-----------|--------|----------|
| Hr       | 3             | LR     | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrAge    | 4             | LR     | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrGender | 4             | LR     | 0.20      | 60.00    | 63.16     | 52.57  | 44.10    |

**Table 3.6 (continued):** Logistic Regression Heart Rate Results

|                        |   |    |      |       |       |       |       |
|------------------------|---|----|------|-------|-------|-------|-------|
| HrPersonality          | 4 | LR | 0.20 | 56.67 | 28.33 | 50.00 | 36.17 |
| HrAgeGender            | 5 | LR | 0.20 | 58.33 | 54.46 | 51.14 | 43.16 |
| HrAgeGenderPersonality | 6 | LR | 0.20 | 55.00 | 45.37 | 48.29 | 41.28 |

MFCC has the maximum accuracy with 61.67% according to Table 3.7. Therefore, H2 is proved. In this algorithm, MFCC results are affected by the addition of age, gender and personality. These decrease accuracy, precision, recall and F1-score results as shown in the table; H4, H5 and H6 are refuted.

**Table 3.7** Logistic Regression MFCC Results

| Features                 | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|--------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Mfcc                     | 12            | LR     | 0.20      | 61.67    | 61.51     | 59.76  | 59.11    |
| MfccAge                  | 13            | LR     | 0.20      | 60.00    | 59.34     | 58.59  | 58.33    |
| MfccGender               | 13            | LR     | 0.20      | 56.67    | 54.99     | 54.86  | 54.86    |
| MfccPersonality          | 13            | LR     | 0.20      | 60.00    | 59.34     | 58.59  | 58.33    |
| MfccAgeGender            | 14            | LR     | 0.20      | 58.33    | 56.58     | 56.29  | 56.28    |
| MfccAgeGenderPersonality | 15            | LR     | 0.20      | 56.67    | 54.58     | 54.29  | 54.17    |

As seen in Table 3.8, the best accuracy is 61.67% for heart rate, MFCC and age feature set, so H1 and H4 are justified. There are three identical results for the logistic regression algorithm with 60.00% accuracy in the table. These reflect the addition of personality, age and gender to heart rate and MFCC values. Personality type and gender features diminish the results of feature set which consists of heart rate and MFCC values. Hence, H5 and H6 are declined.

**Table 3.8** Logistic Regression Heart Rate and Mfcc Results

| Features                   | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| HrMfcc                     | 15            | LR     | 0.20      | 58.33    | 57.54     | 57.07  | 56.88    |
| HrMfccAge                  | 16            | LR     | 0.20      | 61.67    | 61.25     | 60.10  | 59.78    |
| HrMfccGender               | 16            | LR     | 0.20      | 58.33    | 56.58     | 56.29  | 56.28    |
| HrMfccPersonality          | 16            | LR     | 0.20      | 60.00    | 59.34     | 58.59  | 58.33    |
| HrMfccAgeGender            | 17            | LR     | 0.20      | 60.00    | 58.24     | 57.71  | 57.70    |
| HrMfccAgeGenderPersonality | 18            | LR     | 0.20      | 60.00    | 58.24     | 57.71  | 57.70    |

### 3.2.3 Decision Tree Results

A decision tree algorithm is applied to all feature sets and result tables obtained. Results will be examined for the following tables.

The best accuracy is 65% with minimum, maximum and average heart rate, age and gender in decision tree algorithm and test size is 20% (Table 3.9). The next best accuracy rate is 61.67% in two feature sets. The first feature set has minimum, maximum and average heart rate and personality type. Another feature set has 12 MFCC and gender. The lowest accuracy of decision tree algorithm is 49.33% with heart rates, age and gender and test size of 25%.

**Table 3.9** Results of Decision Tree

| Features | Features Size | Method | #Iteration | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------|---------------|--------|------------|-----------|----------|-----------|--------|----------|
| All      | 19            | DT     | 6          | 0.20      | 60.00    | 59.43     | 59.26  | 59.28    |
| All      | 19            | DT     | 1          | 0.25      | 57.33    | 56.49     | 56.39  | 56.40    |
| Hr       | 3             | DT     | 1          | 0.20      | 51.67    | 46.89     | 47.85  | 44.92    |
| Hr       | 3             | DT     | 1          | 0.25      | 57.33    | 55.82     | 54.76  | 53.70    |

**Table 3.9 (continued): Results of Decision Tree**

|                            |    |    |   |      |       |       |       |       |
|----------------------------|----|----|---|------|-------|-------|-------|-------|
| HrAgeGenderPersonality     | 6  | DT | 4 | 0.20 | 58.33 | 58.33 | 59.16 | 57.37 |
| HrAgeGenderPersonality     | 6  | DT | 8 | 0.25 | 57.33 | 57.85 | 56.52 | 55.02 |
| HrAgeGender                | 5  | DT | 4 | 0.20 | 65.00 | 63.24 | 64.29 | 63.28 |
| HrAgeGender                | 5  | DT | 8 | 0.25 | 49.33 | 48.20 | 48.50 | 46.59 |
| HrAge                      | 4  | DT | 4 | 0.20 | 55.00 | 54.02 | 54.40 | 53.43 |
| HrAge                      | 4  | DT | 8 | 0.25 | 50.67 | 49.87 | 49.89 | 48.31 |
| HrGender                   | 4  | DT | 4 | 0.20 | 56.67 | 56.17 | 56.78 | 55.43 |
| HrGender                   | 4  | DT | 1 | 0.25 | 53.33 | 51.65 | 51.52 | 51.11 |
| HrMfccAgeGenderPersonality | 18 | DT | 8 | 0.20 | 56.67 | 57.18 | 56.67 | 55.88 |
| HrMfccAgeGenderPersonality | 18 | DT | 0 | 0.25 | 53.33 | 53.49 | 53.67 | 52.80 |
| HrMfccAgeGender            | 17 | DT | 6 | 0.20 | 53.33 | 53.50 | 53.54 | 53.28 |
| HrMfccAgeGender            | 17 | DT | 6 | 0.25 | 60.00 | 59.11 | 58.77 | 58.76 |
| HrMfccAge                  | 16 | DT | 5 | 0.20 | 55.00 | 55.05 | 55.00 | 54.89 |
| HrMfccAge                  | 16 | DT | 7 | 0.25 | 54.67 | 54.86 | 54.48 | 53.67 |
| HrMfccGender               | 16 | DT | 2 | 0.20 | 55.00 | 55.58 | 55.71 | 54.89 |
| HrMfccGender               | 16 | DT | 3 | 0.25 | 56.00 | 55.63 | 55.54 | 55.49 |
| HrMfccPersonality          | 16 | DT | 2 | 0.20 | 53.33 | 54.21 | 54.29 | 53.28 |
| HrMfccPersonality          | 16 | DT | 3 | 0.25 | 57.33 | 56.96 | 56.79 | 56.71 |
| HrMfcc                     | 15 | DT | 8 | 0.20 | 53.33 | 53.97 | 53.33 | 51.39 |
| HrMfcc                     | 15 | DT | 7 | 0.25 | 56.00 | 55.99 | 55.94 | 55.87 |
| HrPersonality              | 4  | DT | 4 | 0.20 | 61.67 | 59.84 | 60.62 | 59.78 |
| HrPersonality              | 4  | DT | 0 | 0.25 | 58.67 | 55.73 | 55.47 | 55.50 |
| MfccAgeGenderPersonality   | 15 | DT | 6 | 0.20 | 58.33 | 57.81 | 57.74 | 57.76 |

**Table 3.9 (continued): Results of Decision Tree**

|                          |    |    |   |      |       |       |       |       |
|--------------------------|----|----|---|------|-------|-------|-------|-------|
| MfccAgeGenderPersonality | 15 | DT | 4 | 0.25 | 56.00 | 54.55 | 54.79 | 54.41 |
| MfccAgeGender            | 14 | DT | 2 | 0.20 | 53.33 | 53.62 | 53.71 | 53.12 |
| MfccAgeGender            | 14 | DT | 4 | 0.25 | 56.00 | 53.92 | 54.07 | 53.90 |
| MfccAge                  | 13 | DT | 8 | 0.20 | 53.33 | 53.75 | 53.33 | 52.00 |
| MfccAge                  | 13 | DT | 4 | 0.25 | 56.00 | 53.92 | 54.07 | 53.90 |
| MfccGender               | 13 | DT | 2 | 0.20 | 61.67 | 62.28 | 62.57 | 61.57 |
| MfccGender               | 13 | DT | 7 | 0.25 | 57.33 | 57.44 | 57.22 | 56.96 |
| MfccPersonality          | 13 | DT | 6 | 0.20 | 56.67 | 56.00 | 55.89 | 55.88 |
| MfccPersonality          | 13 | DT | 7 | 0.25 | 56.00 | 56.11 | 55.87 | 55.49 |
| Mfcc                     | 12 | DT | 2 | 0.20 | 55.00 | 55.58 | 55.71 | 54.89 |
| Mfcc                     | 12 | DT | 7 | 0.25 | 54.67 | 54.69 | 54.55 | 54.27 |

Figure 3.4 shows the distribution of features sizes and their accuracy rates. There is an incremental change for heart rate feature sets and heart rate-MFCC feature sets.

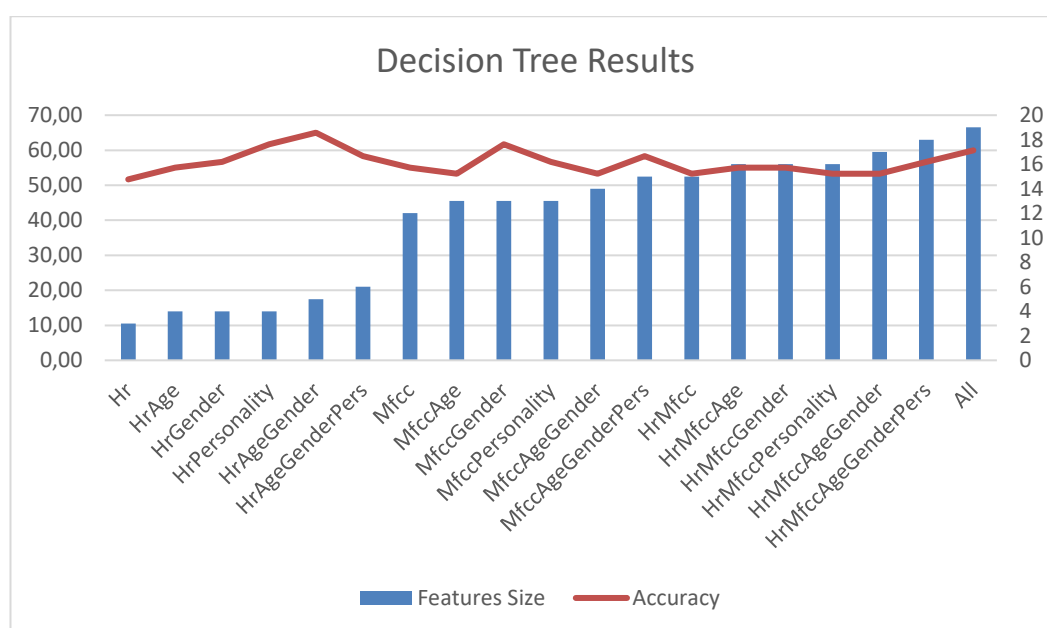
**Figure 3.4** The Graph of Decision Tree Results

Table 3.10 shows that heart rate accuracy is a low 51.67%, proving H3. In the decision tree algorithm, adding new features to heart rate feature set results in outcomes as shown in the table. The addition of personality type, age and gender features to the heart rates increase outcomes. Consequently, H4, H5 and H6 are justified.

**Table 3.10** Decision Tree Heart Rate Results

| Features               | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Hr                     | 3             | DT     | 0.20      | 51.67    | 46.89     | 47.85  | 44.92    |
| HrAge                  | 4             | DT     | 0.20      | 55.00    | 54.02     | 54.40  | 53.43    |
| HrGender               | 4             | DT     | 0.20      | 56.67    | 56.17     | 56.78  | 55.43    |
| HrPersonality          | 4             | DT     | 0.20      | 61.67    | 59.84     | 60.62  | 59.78    |
| HrAgeGender            | 5             | DT     | 0.20      | 65.00    | 63.24     | 64.29  | 63.28    |
| HrAgeGenderPersonality | 6             | DT     | 0.20      | 58.33    | 58.33     | 59.16  | 57.37    |

Table 3.11 represents MFCC results in a decision tree algorithm with test size of 20%. MFCC feature set yields 55% accuracy, but it is a very low rate, proving H2. The maximum accuracy is 61.67% in MFCC and gender features. According to this table, personality type and gender raise outcomes, but age weakens these results. Consequently, H5 and H6 are justified, but H4 is not.

**Table 3.11** Decision Tree MFCC Results

| Features   | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Mfcc       | 12            | DT     | 0.20      | 55.00    | 55.58     | 55.71  | 54.89    |
| MfccAge    | 13            | DT     | 0.20      | 53.33    | 53.75     | 53.33  | 52.00    |
| MfccGender | 13            | DT     | 0.20      | 61.67    | 62.28     | 62.57  | 61.57    |

**Table 3.11 (continued):** Decision Tree MFCC Results

|                          |    |    |      |       |       |       |       |
|--------------------------|----|----|------|-------|-------|-------|-------|
| MfccPersonality          | 13 | DT | 0.20 | 56.67 | 56.00 | 55.89 | 55.88 |
| MfccAgeGender            | 14 | DT | 0.20 | 53.33 | 53.62 | 53.71 | 53.12 |
| MfccAgeGenderPersonality | 15 | DT | 0.20 | 58.33 | 57.81 | 57.74 | 57.76 |

Table 3.12 demonstrates heart rates and MFCC features findings in the decision tree algorithm. Heart rates and MFCC features have minimum accuracy, so H1 is not proved. Still, age, gender and personality raise accuracy rates. However, individually these features are not sufficiently affected. In addition, these are effective when they are used together. According to the table, H4 and H5 are verified, while H6 is rejected.

**Table 3.12** Decision Tree Heart Rate and MFCC Results

| Features                   | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| HrMfcc                     | 15            | DT     | 0.20      | 53.33    | 53.97     | 53.33  | 51.39    |
| HrMfccAge                  | 16            | DT     | 0.20      | 55.00    | 55.05     | 55.00  | 54.89    |
| HrMfccGender               | 16            | DT     | 0.20      | 55.00    | 55.58     | 55.71  | 54.89    |
| HrMfccPersonality          | 16            | DT     | 0.20      | 53.33    | 54.21     | 54.29  | 53.28    |
| HrMfccAgeGender            | 17            | DT     | 0.20      | 53.33    | 53.50     | 53.54  | 53.28    |
| HrMfccAgeGenderPersonality | 18            | DT     | 0.20      | 56.67    | 57.18     | 56.67  | 55.88    |

### 3.2.4 Artificial Neural Network Results

The final section describes the application of algorithm results. This section defines results for the artificial neural network. In addition, there are partial results tables for each feature set which includes heart rates, MFCC and the combination of heart rate and MFCC values.

Table 3.13 shows that the best accuracy is 65% in artificial neural network algorithm. In this method, four different feature sets have the best accuracy. The first is minimum, maximum and average heart rate. The second is heart rate, age and gender values. The next is heart rate and age values. The final feature set is heart rate and gender values. Moreover, the test size of all feature sets is 20%. Four features which consist of heart rates and personality type yield another high accuracy rate of 62.67%. Another feature set provides 49.33% accuracy as the minimum value. This set is composed of 12 MFCC, three heart rates, age and gender.

**Table 3.13** Results of Artificial Neural Network

| Features                   | Features Size | Method | #Iteration | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|------------|-----------|----------|-----------|--------|----------|
| All                        | 19            | ANN    | 3          | 0.20      | 61.67    | 61.11     | 60.77  | 60.78    |
| All                        | 19            | ANN    | 7          | 0.25      | 56.00    | 56.03     | 55.90  | 55.72    |
| Hr                         | 3             | ANN    | 4          | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| Hr                         | 3             | ANN    | 0          | 0.25      | 61.33    | 30.67     | 50.00  | 38.02    |
| HrAgeGenderPersonality     | 6             | ANN    | 2          | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrAgeGenderPersonality     | 6             | ANN    | 2          | 0.25      | 57.33    | 28.67     | 50.00  | 36.44    |
| HrAgeGender                | 5             | ANN    | 4          | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrAgeGender                | 5             | ANN    | 2          | 0.25      | 57.33    | 28.67     | 50.00  | 36.44    |
| HrAge                      | 4             | ANN    | 4          | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrAge                      | 4             | ANN    | 0          | 0.25      | 61.33    | 30.67     | 50.00  | 38.02    |
| HrGender                   | 4             | ANN    | 4          | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrGender                   | 4             | ANN    | 2          | 0.25      | 57.33    | 28.67     | 50.00  | 36.44    |
| HrMfccAgeGenderPersonality | 18            | ANN    | 8          | 0.20      | 55.00    | 55.01     | 55.00  | 54.99    |
| HrMfccAgeGenderPersonality | 18            | ANN    | 8          | 0.25      | 50.67    | 50.64     | 50.64  | 50.63    |

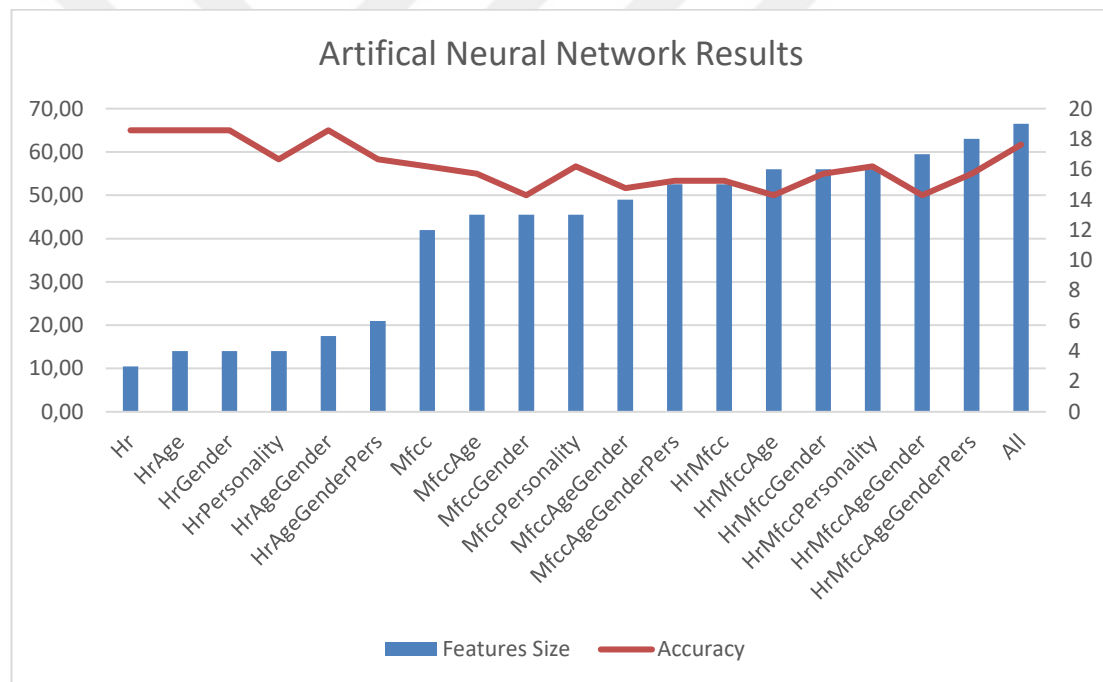
**Table 3.13 (continued):** Results of Artificial Neural Network

|                          |    |     |   |      |       |       |       |       |
|--------------------------|----|-----|---|------|-------|-------|-------|-------|
| HrMfccAgeGender          | 17 | ANN | 1 | 0.20 | 50.00 | 50.00 | 50.00 | 47.12 |
| HrMfccAgeGender          | 17 | ANN | 8 | 0.25 | 49.33 | 49.02 | 49.04 | 48.89 |
| HrMfccAge                | 16 | ANN | 3 | 0.20 | 50.00 | 49.49 | 49.49 | 49.49 |
| HrMfccAge                | 16 | ANN | 8 | 0.25 | 54.67 | 54.59 | 54.59 | 54.59 |
| HrMfccGender             | 16 | ANN | 1 | 0.20 | 55.00 | 55.72 | 55.77 | 54.99 |
| HrMfccGender             | 16 | ANN | 3 | 0.25 | 52.00 | 51.79 | 51.79 | 51.79 |
| HrMfccPersonality        | 16 | ANN | 3 | 0.20 | 56.67 | 56.00 | 55.89 | 55.88 |
| HrMfccPersonality        | 16 | ANN | 1 | 0.25 | 54.67 | 54.91 | 54.98 | 54.59 |
| HrMfcc                   | 15 | ANN | 6 | 0.20 | 53.33 | 52.57 | 52.53 | 52.49 |
| HrMfcc                   | 15 | ANN | 7 | 0.25 | 54.67 | 54.65 | 54.59 | 54.46 |
| HrPersonality            | 4  | ANN | 0 | 0.20 | 58.33 | 29.17 | 50.00 | 36.84 |
| HrPersonality            | 4  | ANN | 4 | 0.25 | 62.67 | 31.33 | 50.00 | 38.52 |
| MfccAgeGenderPersonality | 15 | ANN | 8 | 0.20 | 53.33 | 53.35 | 53.33 | 53.28 |
| MfccAgeGenderPersonality | 15 | ANN | 6 | 0.25 | 56.00 | 55.77 | 55.84 | 55.72 |
| MfccAgeGender            | 14 | ANN | 8 | 0.20 | 51.67 | 51.67 | 51.67 | 51.65 |
| MfccAgeGender            | 14 | ANN | 8 | 0.25 | 56.00 | 55.89 | 55.88 | 55.87 |
| MfccAge                  | 13 | ANN | 6 | 0.20 | 55.00 | 53.75 | 53.37 | 52.78 |
| MfccAge                  | 13 | ANN | 6 | 0.25 | 60.00 | 59.13 | 57.47 | 56.60 |
| MfccGender               | 13 | ANN | 6 | 0.20 | 50.00 | 49.83 | 49.83 | 49.78 |
| MfccGender               | 13 | ANN | 6 | 0.25 | 53.33 | 52.22 | 52.16 | 52.11 |
| MfccPersonality          | 13 | ANN | 6 | 0.20 | 56.67 | 56.00 | 55.89 | 55.88 |
| MfccPersonality          | 13 | ANN | 6 | 0.25 | 52.00 | 51.61 | 51.62 | 51.58 |

**Table 3.13 (continued):** Results of Artificial Neural Network

|      |    |     |   |      |       |       |       |       |
|------|----|-----|---|------|-------|-------|-------|-------|
| Mfcc | 12 | ANN | 3 | 0.20 | 56.67 | 56.84 | 56.90 | 56.62 |
| Mfcc | 12 | ANN | 8 | 0.25 | 54.67 | 55.17 | 55.02 | 54.46 |

Figure 3.5 displays graphically the results of the artificial neural network accuracy with test size of 20%. Increases can be clearly observed with the addition of new features. However, it is not possible to say that the observed results are true for every new feature. While gender can be considered as effective as the heart rate features, it also causes a decline in MFCC features. On the other hand, when both heart rates and MFCC are used together, the gender feature increases accuracy.

**Figure 3.5** The Graph of Artificial Neural Network Results

According to Table 3.14, heart rate yields a 65% accuracy rate. This value is a good result for lie detection, so H3 is not verified for the ANN algorithm. Personality type feature diminishes accuracy, precision and F1-score values. However, age and gender have no impact on heart rate results in the artificial neural network algorithm. Thus, H4, H5, and H6 are not proved.

**Table 3.14** Artificial Neural Network Hr Results

| Features               | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Hr                     | 3             | ANN    | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrAge                  | 4             | ANN    | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrGender               | 4             | ANN    | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrPersonality          | 4             | ANN    | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |
| HrAgeGender            | 5             | ANN    | 0.20      | 65.00    | 32.50     | 50.00  | 39.39    |
| HrAgeGenderPersonality | 6             | ANN    | 0.20      | 58.33    | 29.17     | 50.00  | 36.84    |

According to Table 3.15, MFCC has maximum accuracy with 56.67%, and H2 is proved. In addition, personality type has any effect for accuracy results, but it decreases precision, recall and F1-score values. Gender and age features decrease all results. In this table, the MFCC feature set has maximum value for all results with test size of 20%. Addition of other features diminishes MFCC feature set results for the algorithm. Consequently, H4, H5, and H6 are rejected.

**Table 3.15** Artificial Neural Network MFCC Results

| Features                 | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|--------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| Mfcc                     | 12            | ANN    | 0.20      | 56.67    | 56.84     | 56.90  | 56.62    |
| MfccAge                  | 13            | ANN    | 0.20      | 55.00    | 53.75     | 53.37  | 52.78    |
| MfccGender               | 13            | ANN    | 0.20      | 50.00    | 49.83     | 49.83  | 49.78    |
| MfccPersonality          | 13            | ANN    | 0.20      | 56.67    | 56.00     | 55.89  | 55.88    |
| MfccAgeGender            | 14            | ANN    | 0.20      | 51.67    | 51.67     | 51.67  | 51.65    |
| MfccAgeGenderPersonality | 15            | ANN    | 0.20      | 53.33    | 53.35     | 53.33  | 53.28    |

Table 3.16 shows that heart rate and MFCC has 53.33% accuracy rate. This rate is lower than other results, and thus H1 is not proved. The table proves that heart rates, MFCC and personality types have maximum results. Whereas age feature reduces results, gender and personality type features raise the results. Accordingly, H4 is refuted and H5 and H6 are proved. When gender and age features are added to the heart rates and MFCC, they reduce all results according to heart rates and MFCC feature set. On the other hand, when gender, age and personality types are added to the heart rates and MFCC, they increase the results.

**Table 3.16** Artificial Neural Network Hr and Mfcc Results

| Features                   | Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|----------------------------|---------------|--------|-----------|----------|-----------|--------|----------|
| HrMfcc                     | 15            | ANN    | 0.20      | 53.33    | 52.57     | 52.53  | 52.49    |
| HrMfccAge                  | 16            | ANN    | 0.20      | 50.00    | 49.49     | 49.49  | 49.49    |
| HrMfccGender               | 16            | ANN    | 0.20      | 55.00    | 55.72     | 55.77  | 54.99    |
| HrMfccPersonality          | 16            | ANN    | 0.20      | 56.67    | 56.00     | 55.89  | 55.88    |
| HrMfccAgeGender            | 17            | ANN    | 0.20      | 50.00    | 50.00     | 50.00  | 47.12    |
| HrMfccAgeGenderPersonality | 18            | ANN    | 0.20      | 55.00    | 55.01     | 55.00  | 54.99    |

### 3.2.5 Comparison of All Methods

Each method has a total of 38 feature sets results, of which one half is in 25% test size and the other is in 20% test size. This section focuses only on one feature set result which contains all features in the application of all methods. Furthermore, there are two comparison tables which consists of 25% test size and 20% test size.

As indicated in Table 3.17, logistic regression gives the best result with 66.67% accuracy and 25% test size. Precision, recall and F1-score have a maximum result like accuracy rate. Minimum accuracy is 56% in artificial neural network and its test size is the same with logistic regression.

**Table 3.17** Results of All Features with 25% Test Size

| Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|---------------|--------|-----------|----------|-----------|--------|----------|
| 19            | ANN    | 0.25      | 56.00    | 56.03     | 55.90  | 55.72    |
| 19            | SVM    | 0.25      | 60.00    | 59.42     | 59.42  | 59.42    |
| 19            | DT     | 0.25      | 57.33    | 56.49     | 56.39  | 56.40    |
| 19            | LR     | 0.25      | 66.67    | 64.00     | 63.30  | 63.53    |

Table 3.18 shows that maximum accuracy is 63.33% in logistic regression with 20% test size and this algorithm has the maximum precision, recall and F1-score values. Furthermore, accuracy results from maximum to minimum are logistic regression, artificial neural network, decision tree and support vector machines respectively.

**Table 3.18** Results of 19 Feature with 20% Test Size

| Features Size | Method | Test Size | Accuracy | Precision | Recall | F1-Score |
|---------------|--------|-----------|----------|-----------|--------|----------|
| 19            | ANN    | 0.20      | 61.67    | 61.11     | 60.77  | 60.78    |
| 19            | SVM    | 0.20      | 56.67    | 55.26     | 56.67  | 54.86    |
| 19            | DT     | 0.20      | 60.00    | 59.43     | 59.26  | 59.28    |
| 19            | LR     | 0.20      | 63.33    | 64.51     | 63.74  | 62.96    |

When all results are compared, the best accuracies are achieved with 25% test size in support vector machine and logistic regression algorithms. On the other hand, the decision tree and artificial neural network algorithms have the best results with 20% test size.

It is possible to compare accuracy result for each feature set from the following figures. As shown in the figures, no algorithm yields a maximum accuracy rate for

all feature sets. The line in the figures represent features size, and bars represent accuracy results for SVM, LR, DT and ANN algorithms respectively.

Figure 3.6 shows a comparison of all algorithms' accuracy results with 20% test size. The graph gives an obvious result, that is, the maximum accuracy rate is 65% in the decision tree and ANN algorithms for heart rate, age and gender feature set.

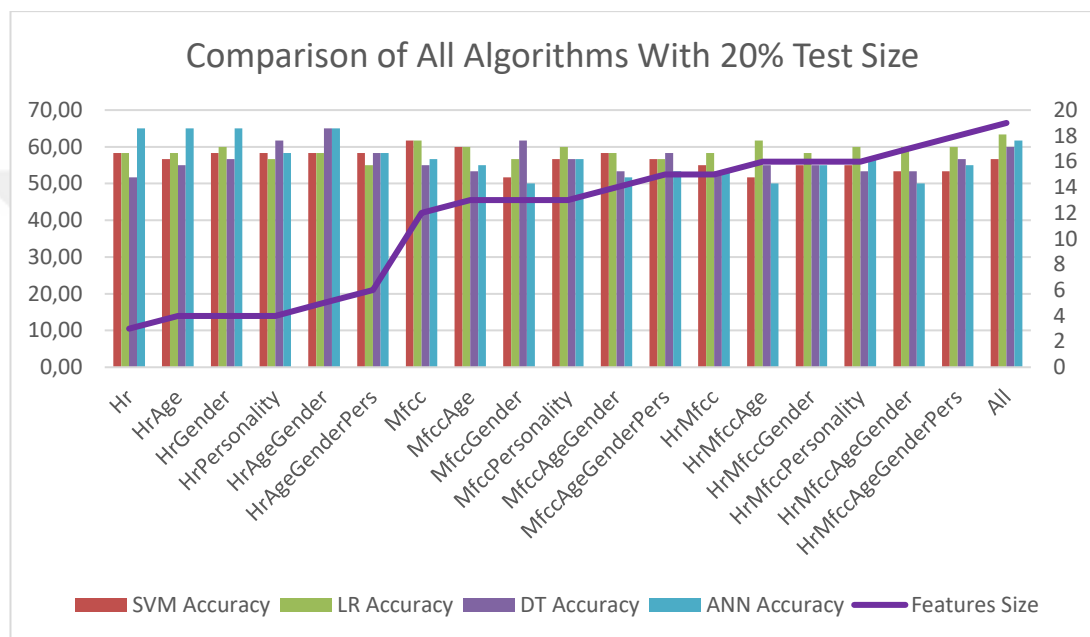


**Figure 3.6** Comparison of All Methods Accuracy Results With 20% Test Size

Figure 3.7 compares all algorithms' accuracy results with 25% test size. Logistic regression method has by far the best accuracy with 66.67% for feature set having all features. Also, the algorithm has maximum accuracy for most of the feature sets.

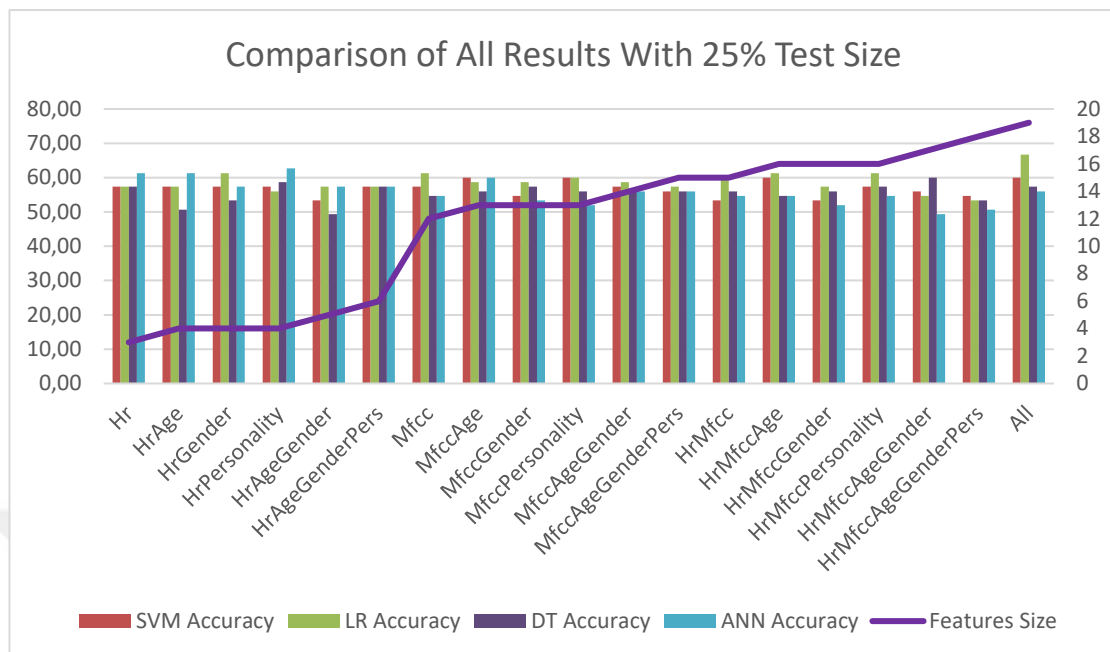


**Figure 3.7** Comparison of All Methods Accuracy Results With 25% Test Size

As a consequence, when all results are compared, the best accuracy is 66.67% in logistic regression algorithm within all algorithms and test size of 25% among all datasets.

The following research hypotheses are evaluated:

H1: The author hypothesizes that voice and heart rate improve lie detection

H2: It is hypothesized that voice-only models will have more than 50% accuracy rate.

H3: It is hypothesized that the heart rate is not sufficient to detect lies successfully.

H4: The author hypothesizes that age has an effect on lie detection.

H5: The author hypothesizes that gender has an effect on lie detection.

H6: We hypothesize that inclusion of personality traits improves the accuracy of lie detection.

As seen in Table 3.19, overall results of hypothesis which are mentioned in 1.2 are shown for all methods.

**Table 3.19** Hypothesis Results

| Hypothesis | SVM | LR | DT | ANN |
|------------|-----|----|----|-----|
| H1         | X   | X  | X  | X   |
| H2         | OK  | OK | OK | OK  |
| H3         | OK  | OK | OK | X   |
| H4         | X   | OK | OK | X   |
| H5         | X   | X  | OK | OK  |
| H6         | X   | X  | OK | X   |

If all results are examined, H1 is refuted but H2 is proved in overall. H3 is verified for most methods. According to the results of H4, H5 and H6, age, gender, and personality – considered individually -- are not sufficient to affect the accuracy, but they can be effective when combined.

# CHAPTER 4

## CONCLUSION & FUTURE WORK

### 4.1 Conclusion

The aim of this study is to detect lies by means of voice and heart rate data. This study is based on the application of some machine learning algorithms and a comparison of their results. For this purpose, an experiment was executed with thirty people and a dataset was created. This dataset consists of 19 different CSV files which contain data describing MFCCs, heart rates, gender, age, personality types, and duration of speech.

Four different Python applications were developed, and the results of machine learning that are support vector machine, logistic regression, decision tree and artificial neural network algorithms were compared. These implementations read 19 different CSV files and classify lie for these files. As a result of this classification, evaluation metrics were calculated, and the results compared. In addition, consequences were likewise compared for different test sizes and for different feature sets. According to the results of these comparisons, the most successful method is logistic regression, with 66,67% accuracy, and its test size is 25%.

There are four findings for this study. The first is that adding a new feature can increase the accuracy rate. However, this is not true for every feature. While gender has no effect on heart rate accuracy, it decreases the MFCC accuracy rate. On the other hand, gender raises accuracy when combined with heart rate and MFCC data. Thus, results may change depending on added features. This highlights the importance of feature selection. The second finding is that accuracy rates for voice-only models are greater than 50%, but this rate is not sufficient for lie detection. Nor is heart rate data sufficient to successfully detect lies. Finally, the addition of heart

rates and personality traits, as well as MFCCs obtained from sound data, significantly affected the lie detection results.

Better results could be achieved if this study were conducted with more participants. Even if we had created a spontaneous deceptive situation, there did not exist an environment of real risk. For this reason, the desired success results could not be achieved. Furthermore, participants could give better reactions when questions relate to participants' own benefits. However, experiment questions are not related to their own benefits.

The maximum accuracy rate of this study can be increased by using more participant approximately 1000 people and this rate can approach 90 percent. Also, using of ensemble learning techniques can give the better result than each algorithm.

The contribution of this study is that it shows that lie detection is only possible by means of voice and heart rate data. Still, these results need to be improved. This study can be expanded, and our study can help to understand whether or not criminals are lying in forensic cases. In addition, this study can be used to prevent fraud in phone calls with voice data. Another contribution of this study is that mentioned dataset can be used for any speech detection study.

This study differs from other studies in its focus on both sound and heart rate data. The examination of personality traits constitutes another difference of this study.

## **4.2 Future Work**

If the number of participants is raised, better results might be achieved. Furthermore, if more risk is built into the experiment environment, better responses can be collected. The risk environment changes MFCC and heart rate results, and accuracy rate can be increased. In addition, this study can enhance using the GSR sensor which is Galvanic Skin Sensor. The GSR sensor could be used to achieve better results by measuring the subjects' perspiration. This sensor can provide more features, and results can be improved. Moreover, we can advance this study with

using some devices which use brain activities. These devices are Functional Near Infrared Spectroscopy (FNIRS), EEG (Electroencephalography). FNIRS enables the blood oxygen level and blood volume to be monitored with brain activity and we can enhance results with FNIRS. Furthermore, using electrocardiogram (ECG) can improve this study. ECG ensures to take better heart rate data. It can give better results for heart rate data than wristband.

We could add respiratory rate to the features, thereby enhancing the accuracy rate. If the accuracy ratio is increased, an application can be developed.



## REFERENCES

- [1] S. V. Paunonen and M. C. Ashton, "Big Five factors and facets and the prediction of behavior.," *J. Pers. Soc. Psychol.*, vol. 81, no. 3, pp. 524–539, 2001.
- [2] K. Meena, K. Veena, and A. Professor, "Lie Detection System Using Input Voice Signal," vol. 117, no. 8, pp. 121–125, 2017.
- [3] F. Enos, "Detecting deception in speech," *Diss. Theses*, vol. 70, no. 2, p. 1119, 2009.
- [4] S.Üretmen, "Cinsiyet, Yalan Söyleme ve Çıkar Elde Etmenin Yalana İlişkin Yüklemlere Etkisi," 2011.
- [5] E. Akçay, "LIE ABILITY WITH RESPECT OF MEMORY POWER: COMPARISON BETWEEN GENDERS," *Selçuk Univ. Fac. Commun. Acad. J.*, vol. 3, no. 7, pp. 234–243, 2012.
- [6] A. Vrij, *Detecting lies and deceit: pitfalls and opportunities*. Wiley Series in the Psychology of Crime, Policing and Law, Second Edi. John Willey & Sons, Ltd, 2008.
- [7] A. . Mishra and O. . Vaidya, "Attributes Detection for Allocation of Resources with Stress Detection from Speech," vol. 3, no. 4, pp. 61–66, 2014.
- [8] D. Hipsky, E. Parimanath, J. Mueller, M. McGinnis, and S. Wright, "Analyzing Fluctuations in Heart Rate , Respiration , and Perspiration as Physiological Indicators of Deception," 2014.
- [9] T. Kinnunen and E. Chernenko, "Voice activity detection using MFCC features and support vector machine," *Int. Conf. Speech*, vol. 2, no. 2, pp. 555–561, 2007.

- [10] E. Fokoue, Z. Ma, and E. Fokoué, "Speaker Gender Recognition via MFCCs and SVMs," 2013.
- [11] S. Ayhan and Ş. Erdoğan, "Kernel Function Selection for the Solution of Classification Problems via Support Vector Machines," *Eskişehir Osmangazi Univ. J. Econ. Adm. Sci.*, vol. 9, no. 1, pp. 175–198, 2014.
- [12] M. Gilke, P. Kachare, R. Kothalikar, and V. P. Rodrigues, "MFCC-based Vocal Emotion Recognition Using ANN," *2012 Int. Conf. Electron. Eng. Informatics (ICEEI 2012)*, vol. 49, no. Iceei, pp. 150–154, 2012.
- [13] C. Gershenson, "Artificial Neural Networks for Beginners," vol. 66, no. 6, 2003.
- [14] B. Gupta, A. Rawat, A. Jain, A. Arora, and N. Dhama, "Analysis of Various Decision Tree Algorithms for Classification in Data Mining," *Int. J. Comput. Appl.*, vol. 163, no. 8, pp. 15–19, 2017.
- [15] S. O. N. G. Yan-yan and L. U. Ying, "Decision tree methods: applications for classification and prediction," *Shanghai Arch. Psychiatry*, vol. 27, no. 2, pp. 130–135, 2015.
- [16] B. M. DePaulo, S. E. Kirkendol, D. A. Kashy, M. M. Wyer, and J. A. Epstein, "Lying in Everyday Life," *J. Pers. Soc. Psychol.*, vol. 70, no. 5, pp. 979–995, 1996.
- [17] M. MacGill, "What should my heart rate be?," *Medical News Today*, 2017. [Online]. Available: <https://www.medicalnewstoday.com/articles/235710.php>. [Accessed: 06-Apr-2019].
- [18] A. Mottelson, J. Knibbe, and K. Hornbæk, "Veritaps : Truth Estimation from Mobile Interaction," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI'18*, pp. 1–12, 2018.
- [19] M. Vicianova, "Literature Reviews Historical Techniques of Lie Detection,"

*Eur. J. Psychol.*, vol. 3, no. 11, pp. 522–534, 2015.

- [20] A. A. Khulage and B. . Pathak, “Analysis of Speech Under Stress Using Linear Techniques and Non-Linear Techniques for Emotion Recognition System,” *CoRR*, pp. 285–294, 2012.
- [21] S. Sondhi, R. Vijay, M. Khan, and A. K. Salhan, “Voice analysis for detection of deception,” in *11th 2016 International Conference on Knowledge, Information and Creativity Support Systems, KICSS 2016*, 2017, pp. 1–6.
- [22] V. P. Patil, K. K. Nayak, and M. Saxena, “Voice Stress Detection,” *Int. J. Electr. Electron. Comput. Eng.*, vol. 2, no. 2, pp. 148–154, 2013.
- [23] X. Pan, H. Zhao, and Y. Zhou, “The application of fractional Mel cepstral coefficient in deceptive speech detection,” *PeerJ*, vol. 3, p. e1194, 2015.
- [24] A. Suresh and H. Shaheen, “Speech Stress Analysis based on Lie Detector for Loyalty Test,” *Int. J. Printing, Packag. Allied Sci.*, vol. 4, no. 1, pp. 631–638, 2016.
- [25] J. Masip, “Deception detection: State of the art and future prospects,” *Psicothema*, vol. 29, no. 2, pp. 149–159, 2017.
- [26] Bustami, Fadlisyah, and N. Delemunte, “Lie Detection System with Voice Using Bidirectional Associative Memory Algorithm,” *IJCAT - Int. J. Comput. Technol.*, vol. 2, no. 8, pp. 301–306, 2015.
- [27] C. N. H. Street, “ALIED: A Theory of Lie Detection,” *Cienc. Cogn.*, vol. 10, no. 2, pp. 49–52, 2016.
- [28] V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, Y. Xiao, C. Linton, and M. Burzo, “Verbal and Nonverbal Clues for Real-life Deception Detection,” no. September, pp. 2336–2346, 2015.
- [29] A. Vrij, R. P. Fisher, and H. Blank, “A cognitive approach to lie detection: A

- meta-analysis,” *Leg. Criminol. Psychol.*, vol. 22, no. 1, pp. 1–21, 2017.
- [30] A. Vrij, S. Leal, L. Jupe, and A. Harvey, “Within-subjects verbal lie detection measures: A comparison between total detail and proportion of complications,” *Leg. Criminol. Psychol.*, vol. 23, no. 2, pp. 265–279, 2018.
- [31] F. Lacerda, “Voice stress analyses: Science and pseudoscience,” *J. Acoust. Soc. Am.*, vol. 133, no. 5, pp. 3244–3244, 2013.
- [32] R. Gray, “Lies , Liars , and Lie Detection,” *Fed. Probat.*, no. December, pp. 31–37, 2011.
- [33] A. Curci, T. Lanciano, F. Battista, S. Guaragno, and R. M. Ribatti, “Accuracy , Confidence , and Experiential Criteria for Lie Detection Through a Videotaped Interview,” vol. 9, no. January, pp. 1–14, 2019.
- [34] Turkish Language Society, “No Title,” 2006. [Online]. Available: [http://www.tdk.gov.tr/index.php?option=com\\_gts&kelime=YALAN](http://www.tdk.gov.tr/index.php?option=com_gts&kelime=YALAN). [Accessed: 05-Apr-2019].
- [35] P. T. Costa and R. R. McCrae, “The Five-Factor Model, Five-Factor Theory, and Interpersonal Psychology,” *Handb. Interpers. Psychol. Theory, Res. Assessment, Ther. Interv.*, no. March 2012, pp. 91–104, 2012.
- [36] O. Somer, M. Korkmaz, and A. Tatar, “Beş Faktör Kişilik Envanteri’ nin geliştirilmesi-I: Ölçek ve alt ölçeklerin oluşturulması,” *Turkish Psychol. Assoc.*, vol. 17, no. 49, pp. 21–33, 2002.
- [37] M. Gamer, “Detecting concealed information using autonomic measures,” *Mem. Detect. Theory Appl. Concealed Inf. Test*, pp. 27–45, 2011.
- [38] E. Yucesoy and V. V. Nabiyevev, “Gender identification of a speaker using MFCC and GMM,” in *2013 8th International Conference on Electrical and Electronics Engineering*, 2013, pp. 626–629.

- [39] Librosa Development Team, “Librosa Tutorial,” 2019. [Online]. Available: <https://librosa.github.io/librosa/tutorial.html>. [Accessed: 05-Apr-2019].
- [40] Spyder Project Contributors, “The Scientific Python Development Environment,” 2019. [Online]. Available: <https://pypi.org/project/spyder/>.



# APPENDICES

## APPENDIX A. Written Documents of the Main Experiment

### Scenario

#### Temel Talimatlar:

Çalışmamıza Hoş geldiniz.

Bu çalışmada, başlangıçta 5 tane alıştırma sorusu sorulacaktır. Daha sonra asıl test için gerekli olan 10 soru sorulacaktır ve sizden bu sorulara doğru/yalan cevaplar vermeniz beklenmektedir.

Çalışma başlamadan önce size vereceğimiz akıllı bilekliği kolunuza çok sıkı ve çok gevşek olmayacak şekilde takmanız istenecektir. Bu bileklik çalışma sonlanana kadar bileğinizde takılı olmalıdır.

Çalışma esnasında size sorulan sorulara cevap vermek için mikrofonun yakınında size gösterilen yerde bulunmalısınız.

Bu çalışma başladığında size sorular sorulup doğru/yalan yazılı kartlar gösterilecektir. “Doğru” kartı gösterildiğinde sorulan soruya anlık olarak doğru cevap, “Yalan” kartı gösterildiğinde sorulan soruya o an yalan uydurmanız gerekecektir.

Sorular size sorulup kartı gördükten sonra cevap vermeniz gerekmektedir.

Çalışmamız ortalama 15 dakika sürmektedir.

Hazır olduğunuzda çalışmaya başlayabiliriz.

## Experiment Questions

### Test Soruları

- 1)Adınız?
- 2)Yaşınız?
- 3)Cinsiyetiniz?
- 4)Mesleğiniz?
- 5)Türkiye'nin başkenti neresidir?

### Deney Soruları

- 1)Araba kullanırken telefonunuz çaldığında ne yaparsınız?
- 2)Yolda para bulsanız ne yaparsınız?
- 3)Her şey dahil bir otele gittiğinizde otelden paranızı çıkarmak için neler yaparsınız?
- 4)Hobileriniz nelerdir?
- 5)Bankayı nasıl soyduğunu anlat?
- 6)Soygun yapmak istersen suç aleti olarak ne kullanırsın?
- 7)Düzenli olarak yalan söylediğini öğrendik, en son ne zaman yalan söyledin?
- 8)En son söylediğin yalan nedir?
- 9)Eşiniz/arkadaşınız yeni bir şey almış, siz bunu beğenmediniz ve sizden yorum istediğinde ne dersiniz?
- 10)Büyük ikramiye size çıkarsa hayatınızda değiştireceğiniz ilk şey ne olur?

## Informed Consent Form

### Bilgilendirilmiş Onay Katılım Formu (Informed Consent Form)

Bu çalışma, Ankara Yıldırım Beyazıt Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği programı bünyesinde yürütülmekte olan tez çalışması kapsamında düzenlenmektedir. Çalışma yöneticisi Dr. Özkan Kılıç, yüksek lisans öğrencisi Şerife Özdamarlar' dır.

Çalışma, ses ve kalp ritmi verilerini kullanarak kişilerin yalan söyleyip söylemediğini anlamak için veri toplamayı amaçlamaktadır. Çalışma boyunca kullanılacak materyal genel olarak kişisel rahatsızlık verecek içeriğe sahip değildir. Sizden beklenen sorulan sorulara, size gösterilen kartlara göre doğru/yalan cevaplar vermenizdir. Çalışma esnasında sorulacak sorular sizin kişiliğinizden bağımsız olup sadece çalışmanın daha iyi sonuç vermesi için seçilmiştir. Ancak, katılım sırasında herhangi bir nedenden ötürü kendinizi rahatsız hissederseniz çalışmayı yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda çalışmayı yürüten kişiye, çalışmayı tamamlamadığınızı söylemek yeterli olacaktır. Çalışmanın herhangi riski yoktur.

Çalışmaya katılım bilgilendirilmiş onay (informed consent) esasına dayanmaktadır. Çalışma boyunca, sizden istenecek kimlik bilgileri verilerle eşleştirilmemektedir. Cevaplarınız tamamıyla gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir; elde edilecek bilgiler bilimsel yayınlarda kullanılacaktır. Çalışma sonunda, varsa çalışmayla ilgili sorularınız cevaplanacaktır.

Katılımınız için şimdiden teşekkür ederiz.

(Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

***Bu çalışmaya bilgilendirilmiş olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayınlarda kullanılmasını kabul ediyorum.***

İsim Soyad

Tarih

İmza

---/---/---

## Personality Test

Yönerge: Aşağıdaki ifadelerin sizin için ne kadar doğru olduğunu, ifadenin yanındaki seçeneklerden size uygun olanını işaretleyerek belirtiniz.

|                                                                        | Kesinlikle<br>Katılmıyorum | Kısmen<br>Katılmıyorum | Kararsızım | Kısmen<br>Katılıyorum | Kesinlikle<br>Katılıyorum |
|------------------------------------------------------------------------|----------------------------|------------------------|------------|-----------------------|---------------------------|
| Kendimi içe dönük biri olarak görüyorum                                |                            |                        |            |                       |                           |
| Kendimi genelde güvenilir biri olarak görüyorum                        |                            |                        |            |                       |                           |
| Kendimi tembel olma eğiliminde biri olarak                             |                            |                        |            |                       |                           |
| Kendimi stresle baş edebilen, rahat biri olarak                        |                            |                        |            |                       |                           |
| Kendimi sanatsal ilgileri olan biri olarak görüyorum                   |                            |                        |            |                       |                           |
| Kendimi dışa dönük, sosyal biri olarak                                 |                            |                        |            |                       |                           |
| Kendimi başkalarının hatalarını bulmaya eğilimli biri olarak görüyorum |                            |                        |            |                       |                           |
| Kendimi işlerini özenle yapan biri olarak                              |                            |                        |            |                       |                           |
| Kendimi çabuk sinirlenen biri olarak görüyorum                         |                            |                        |            |                       |                           |
| Kendimi etkin hayal gücüne sahip biri olarak                           |                            |                        |            |                       |                           |

# CIRRICULUM VITAE

## PERSONAL INFORMATION

**Name Surname** : Şerife ÖZDAMARLAR  
**Date of Birth** : 30.07.1991  
**Phone** : 0537 847 0869  
**E-mail** : serifeozdamarlar@gmail.com



## EDUCATION

**High School** : Kecioren Incirli High School  
**Bachelor** : Çankaya University,  
 Mathematics and Computer Science (2009-2014)  
 Çankaya University,  
 Computer Engineering (Double Major) (2012-2014)  
**Master's degree** : Ankara Yildirim Beyazit University,  
 Computer Engineering (2017-2019)

## WORK EXPERIENCE

**Scientific Programs Assistant Specialist** : June 2016 –  
 TÜBİTAK

**Android Software Developer** : July 2014 – June 2016  
 Kardiosis Mechanical and Electronic Products Inc., Ankara

**Software Developer** : January 2014 – May 2014  
 Varatah Technology Solutions, Ankara

**Intern** : August 2013 – September 2013  
 MilSOFT Software Technologies Inc., Ankara

**Intern** : August 2012 – September 2012

Turkey Petroleum Corporation, Ankara

### **TECHNICAL SKILLS**

Programming Languages: Java, C, C++, C#, .Net

UI Frameworks: Pebble, Primefaces

Database Programs: Oracle, Postgres

### **LANGUAGE SKILLS**

English: Intermediate oral and writing, reading skills

### **HOBBIES**

Travelling, swimming

### **CONFERENCE PRESENTATIONS**

Ozdamarlar S., Kılıç Ö. (2019) Lie Detection from Voice and Heart Rate, 6. *Ejons International Journal on Mathematics, Engineering & Natural Sciences 2019, Adana.*