T.C. DOĞUŞ UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY

COMPUTER ENGINEERING

Understanding Users' Neuro Physical Conditions for Minimizing Human Oriented Risks

Ph.D. Thesis by

Aykut Güven

2014194001

Thesis Supervisor

Prof. Dr. Mitat Uysal

İstanbul, Aug 2019

	DOKTORA TEZ SAVUNMA SINAV TUTANAĞI	Doküman No	FR.1.46
		Yürürlük Tarihi	1.11.2017
		Revizyon Tarihi	1.11.2017
		Revizyon No	1
		Sayfa	1/1

SOSYAL BİLİMLER / FEN BİLİMLERİ ENSTİTÜSÜ

Tarih : 06.12.2019

Anabilim Dalı : Bilgisayar Mühendisliği

Öğrencinin Adı Soyadı : Aykut GÖKEN

Öğrenci No : 2014194001

Tez Danışmanının Adı Soyadı : Mitat UYSAL

İkinci Tez Danışmanının Adı Soyadı : Selim AKYOKUŞ

Tezin Başlığı : Understanding User's Neuro Physical Conditions for Monitoring Human Orientation Risk

Doğuş Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin 44.Maddesi uyarınca yapılan değerlendirmeler sonunda;

☒ **tezin kabul edilmesine** ☐ **tezde düzeltme verilmesine** ☐ **tezin reddedilmesine**

oy birliği / ~~oy çokluğu~~ ile karar verilmiştir. Gereği için arz olunur.

Prof. Dr. Mitat Uysal

Danışman Üye

Üye

İbrahim Sogukpınar

Prof. Dr.

Üye

Dr. Öğr. Üyesi Zeynep Hilal Kilimci

Üye

Prof. Dr. Selim Akyokuş

Üye

Rayim Dapr

Dr. Öğr. Üyesi

Anabilim Dalı Başkanı Onayı:

Dr. Öğr. Üyesi Yeşim Kaygısız

(Form No: FR.0001, Revizyon Tarihi: 30.09.2016, Revizyon No:00)

DECLARATION

I declare that this thesis my own work and has not been submitted in any form for another degree or diploma at any university or institution. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Aykut Güven

Without them nothing would be important.

Dedicated to My Precious Family …

ACKNOWLEDGEMENTS

I would like to thank my advisor professor, Prof. Dr. Mitat Uysal for his continuous support and guidance. His contributions were invaluable during my thesis project.

I would also like to thank Prof. Dr. Selim Akyokuş who was my secondary advisor. He supported many areas of my study.

I must extend my appreciation for Dr. Zeynep Hilal Kilimci who acted as my mentor throughout this project. with her valuable scientific contributions to my study, especially on journal works. I am very happy and proud to have worked with such a valuable colleague.

I have to thank to my colleagues in Idea Technology. They were part of this study during data collection stage. Especially, Mrs. Şeyma Taşçı, her precious editorial support was valuable during this thesis.

Thanks to Joga Bonito team for their support. They were a part of this thesis by supplying data for a long time.

Finally, special thanks goes to my family members, my wife Özlem Güven and my sons Mehmet Efe Güven and Ali Kuzey Güven for their continuous support and encouragement and through this long and challenging journey.

İstanbul, 2019 Aykut Güven

ABSTRACT

Today, smart devices are capable to collect their users' data using a variety of different sensors. Devices are usually online and they share their data using internet. This device to device communication is called as IoT (Internet of Things). IoT backbones store data in Data Centres (DC). Stored data contains valuable information about the users. Analysing the data gives detailed information about the users. In this thesis, a small model of this above IoT structure is simulated. Data is collected and emotion analysis of the users is improved by enriching the physical characteristics of users such as sleep, heartbeat, mobility, etc. with neuro physical parameters such as keystroke patterns, and keystroke error count measurement. The contribution of this work to the literature is the combine usage of physical and neuro physical parameters. Keystroke patterns, and keystroke error count measurement are studied as neuro physical parameters while sleep quality, energy, mobility/movement, and heart pulse are analysed as physical parameters. Another novelty is that the classification task is performed by both conventional(classical) machine learning algorithms and deep learning models to analyse the emotion of the users. For this purpose, feedforward neural network (FFNN), convolutional neural network (CNN), recurrent neural network (RNN), and long short-term neural network are employed as deep learning methodologies while multinomial naive Bayes (MNB), support vector regression (SVR), decision tree (DT), random forest (RF), and decision integration strategy (DIS) are evaluated as conventional machine learning algorithms. To the best of our knowledge, this is the very first attempt to analyse the neuro physical conditions of the users by evaluating deep learning models for sensitivity analysis and enriching physical characteristics with neuro physical parameters. The dataset is collected with the usage of smart devices and sensors from the users during one-year time period. Experimental results demonstrate that the utilization of deep learning methodologies and the combination of both physical and neuro physical parameters enhance the classification success of the system to interpret the sensitivity of the users. A wide range of comparative and extensive experiments show that the proposed model exhibits noteworthy results compared to the state-of-art studies.

Keywords: Biometric, Emotions, Physical, Neuro Physical, ANN, Ensemble Learning, Machine Learning, Decision Integration Strategy, Deep Learning.

ÖZET

Günümüzde akıllı cihazlar barındırdıkları sensorler vasıtasıyla kullanıcıları hakkında birçok veriyi toplamaktadırlar. Akıllı cihazları kullanan yüzmilyonlarca insan olduğu düşünüldüğünde toplanan bu veriler büyük bir IoT ağı ortaya çıkarmaktadır. Bu kadar büyük hacimli veriler, klasik veri saklama yöntemlerinin ötesinde büyük veri mimarilerine uygun bir şekilde veri merkezlerinde toplanmaktadırlar. Toplanan verinin analiz edilerek yorumlanması sonucu ortaya kullanıcılar hakkında kıymetli bilgilerin çıkartılması mümkündür. Günümüzün en aktif araştırma konularından bir tanesi duygu analizidir. Duygu analizi ile kişilerin belli konular hakkında düşünce ve duyguları anlaşılmaya çalışılır. Duygu analizinde genel yöntem metin analizi yöntemi ile kişilerin vermiş oldukları tepkileri anlamaya çalışmaktır. Bu tez kapsamında kullanıcıların duygusal durumlarının analizi farklı bir yöntem izlenerek yapılmaya çalışılmıştır. Kullanıcıların fiziksel özellikleri olan uyku, kalp atışı, hareketlilik vs. ile davranışsal özellikleri olan yazma ritmi algılama, yazı hata oranları gibi parametereleri toplanarak duygu durumlarının anlaşılmaya çalışıldığı bir model oluşturulmuştur. Çalışmanın literature en büyük katkılarından bir tanesi hem fiziksel hemde davranışsal kişi özelliklerini aynı anda analiz etmesidir. Diğer bir yenilik ise klasik makine öğrenmesi yöntemlerinin derin öğrenme algoritmaları ile birlikte kullanılarak zengin içerikli bir makine öğrenme yaklaşımının benimsenmesidir. Klasik makine öğrenme algoritmaları olarak Multinominal Naïve Bayes, Destek Vektör Regresyon, Karar Ağaçları, Rastsal Karar Ağaçları yöntemleri kullanılmıştır. Derin öğrenme algoritmaları olarak da İleri Yönlü Yapay Sinir Ağı (FFNN), Evrişimsel Sinir Ağları (CNN), Tekrarlayan Yapay Sinir Ağları (RNN), Uzun kısa-zaman hafızalı yapay sinir ağları algoritmaları (LSTM) kullanılmıştır. Sonuçlar Kollektif öğrenme (Ensemble Learning) algoritması kullanılarak tekilleştirilmiştir. Çalışma kapsamında 15 kişilik bir denek grubundan akıllı saatler ve özel bir mobil uygulama kullanılarak 1 sene boyunca bilgiler toplanmıştır. Toplanan bu bilgiler hem klasik makine öğrenmesi hemde derin öğrenme yöntemleri kullanılarak analiz edilmiştir. Analiz sonuçları algoritma bazında yorumlanmıştır.

Anahtar Kelimeler: Biometri, Duygu Analizi, Fiziksel Özellikler, Davranışsal Özellikler, Yapay Sinir Ağları, Ensemble Öğrenme, Makine Öğrenmesi, Karar Entegrasyon Stratejisi, Derin Öğrenme.

CONTENTS

TABLE LIST

FIGURE LIST

ABBREVIATIONS

ANN : Artificial Neural Network

CNN : Convolutional Neural Network

DC : Data Centers

DIS : Decision Integration Strategy

DNA : Deoxyribo Nucleic Acid

DT : Decision Trees

EC : Error Count

EDM : Euclidean Distance Measure

EEG : Electro Encephalo Graphy

FBI : Federal Bureau of Investigation

FNNN : FeedForward Neural Network

GNF : Generic Normalization Factor

HRV : Heart rate variability

IoT : Internet of Things

k-NN : k-Nearest Neighbours

KPD : Key Press Delay

KSD : Key Stroke Delay

LSTM : Long Short-Term Memory Neural Network

MIT : Massachusetts Institute of Technology

MNB : Multinomial Naive Bayes

MTL : Multi-Task Learning

NB : Naïve Bayes

NoSQL : Not Only SQL

NPT : Neuro Physical Testing

NSF **:** National Sleep Foundation

RBF **:** Radial Basis Function

RF **:** Random Forest

RNN **:** Recurrent Neural Network

SHI **:** Sleep Health Index

SVM **:** Support Vector Machines

SVR **:** Support Vector Regression

1.INTRODUCTION

1.1 Emotions and Physical / Neuro Physical Biometric Parameters

Intelligent and integrated devices become one of the essential parts of our daily life both social and business. Especially, computers, phones, tablets, sensors, and cloud services have been part of our private and public domains all around the world in the last decades. These devices gather many parameters about their users' including mobility (walking, running, climbing), sleep time, places where the users visited. Moreover, developing technology integrates devices to the biological body of its user and this creates an online human being accessible by the world. It was thought that being online generally increases tension/stress level of the users and lowers his/her immunity against regular life/business problems. Nowadays users are more technology friendly and accepting those devices as a part of their daily life. Changing positive perceptions on technology and increasing device capabilities create many opportunities in order to use these technologies in many areas. There are many applications those use collected parameters exist. These applications usually for health care and fitness purposed applications. Is it possible to estimate emotions of a user if these devices capable to collect many parameters? This thesis focuses on answering the question. Understanding emotional states of users are quite important for people who work at critical positions and many countries have strong regulations. These regulations have strict obligations so that; they regularly need to take Neuro Physical Testing (NPT) to ensure mental health is good to avoid risks depending on stress. The application period of NPT is usually once a year or twice which is not sufficient. Managing such risks should be a part of daily operation which is possible to implement this with intelligent and integrated devices. Otherwise, user-oriented mistakes may cause dangerous situations to be not compensated. According to Quorum Disaster Recovery Report, Q1 2013 disaster created by users is 22% while hardware failures 55%, software failures 18%, natural disasters 5%, merely. Furthermore, understanding users' neuro physical conditions using devices has a long history. Earlier studies have a limited input devices and the popular one is keyboard. This is why many researchers concentrate on the understanding user's keyboard usage. The first study is implemented in 1980 (Gaines et al.1980). The Study demonstrates that the keystroke can be used to supply a security in computer systems. In this thesis, keystroke parameters are also evaluated as a feature of the proposed study.

Sensors technology and autonomous communication capabilities creates a broad opportunity for the future. Objects which can be alive or machines can communicate each other without any limitations. This is called as Internet of Things(IoT). IoT becomes another popular research field in computer sciences. Although custom devices are designed for IoT purposes, Regular Smart Devices are also used for IoT applications. Smart Devices are capable to measure various parameters of their users such as heart data, sleep time, mobility, blood pressure, body temperature, location data and many known and unknown parameters behind by means of sensors. Some of these parameters are also used in our study. These physical parameters can be related to emotional state of a person.

There are many types of supervised classifiers used in different research areas such as text categorization, image classification, pattern recognition. A review and comparison of supervised algorithms is presented in papers (Aggarwal et al.2012; Sebastiani,2002). Some of the commonly used supervised algorithms include naïve Bayes (NB), k-nearest neighbours (k-NN), decision trees (DT), artificial neural networks (ANN), and support vector machines (SVMs).

In this thesis, we propose to analyse neuro physical conditions of the users by combining with physical conditions and evaluating deep learning models. For this purpose, a mobile application is designed in order to gather the parameters from users. Application is collected and recorded those physical and neuro physical parameters from each user and recoded to a central repository for the purpose of analysing. Architecture section (1.2) gives detailed information about how application works. One year of data is collected from fifteen users to demonstrate the efficiency of this work. After that, feedforward neural network (FFNN), convolutional neural network (CNN), recurrent neural network (RNN), long short-term neural network (LSTM) as deep learning methodologies; multinomial naive Bayes (MNB), support vector regression (SVR), decision tree (DT), random forest (RF), and decision integration strategy (DIS) as traditional machine learning algorithms are employed for classifying and understanding neuro physical/emotional condition of each user. To our knowledge, this is the very first attempt to combine the physical characteristics and neuro physical parameters of the users. Another novelty is to utilize the deep learning models by interpreting emotion analysis of the users for the classification task in this domain. Experiment results indicate that the consolidation of both physical and neuro physical parameters and the utilization of deep learning methodologies boost the classification success of the system to understand the emotion of the users. Related Work (1.3) section gives detailed information about studies done related to our study.

The rest of this thesis is organized as follows: Section 2 gives a detailed information about biometrics and history of biometrics. Section 2 also focuses on definition of emotions and history of emotional studies. Section 3 explains Neuro Physical Parameter and Physical Parameters used in thesis. Section 4 explains algorithms used in this study. Section 4 has two main parts first part focuses on classical machine learning algorithms and second part focuses on Deep Learning algorithms. Deep Learning algorithms are known ANN models. Section 5 explains experimental results of the study. Mainly performance of each algorithm are given in detail. Section 6 is conclusion and future study part.

1.2 Architecture & Developed Software

Collecting users' parameters has been difficult part of this thesis. Data collection period was long and needed daily attention for the users. Instead of using third party applications in the smart devices, we decided to create a mobile application for data collection. At the earlier stages of the data collection period, users were trained for the mobile application usage and they were also informed about tasks to do daily basis. Application usage behaviours of the users were observed for the first month to see if they were able to use the mobile application properly. Feedbacks were given to the users if there was an improper usage of the mobile application. Data quality and flow became stabile after the first month and formal data collection process started. Sensors on Vestel Smart Watch and Apple iWatch were used as main physical data suppliers. Users were asked to use given smart watches in order to collect their physical data. Smart Watches collects physical data by default as a part of their functionality. This data is pulled by the mobile application. The mobile application sends the data to the backend repository. A middleware program at the server side manages data manipulation. It makes necessary quality checks on the data and places incoming data to a proper place in the repository.

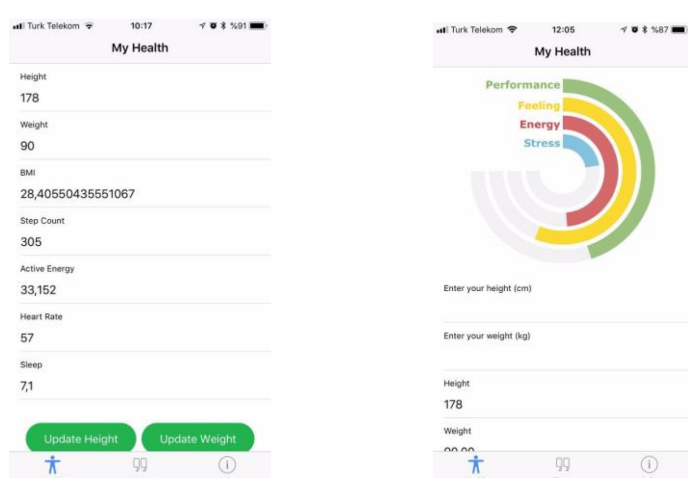


Figure 1.1 Created Mobile Application

Besides the physical parameters, the mobile application has an interface to collect neuro physical parameters using keystrokes. Aim of collecting keystroke patterns is to measure attention. The mobile application gives a text input screen to the users and asks to type a given text regularly to collect keystroke patterns. The given text is generated randomly by the mobile application. Random texts eliminate memorization risk. The users have to make a fresh effort to type text on each time.

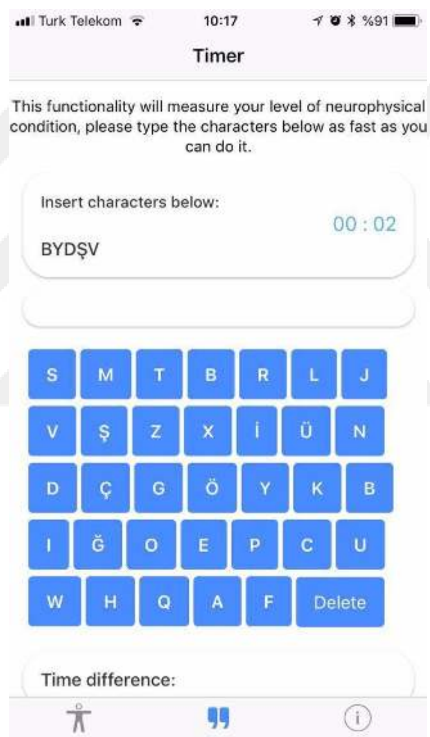


Figure 1.2 Mobile Application Keystroke Patterns Screen

Keyboard layout is also randomized. Randomized keyboard is a special type of keyboard and order of the letters randomly change on each iteration. The mobile application logs keystroke times and typing errors when a user types the given text. A detailed explanation of keystroke patterns is presented in Section 3.1.

The mobile application contains a large emotion embedded questions pool. Three of these questions are randomly chosen. These questions are asked users daily basis in the morning and the evening. Answers are logged to understand the users' initial mood when they start the day(morning). At the end of the day, another randomly chosen questions are asked to users for understanding end of the day(evening) mood. A mood calculation is done depending on the answers. These calculations are used in the machine learning models as input parameters.

All of the parameters explained above are sent to backend repository and stored in a NoSQL structure. The mobile application pulls the data hourly basis. Each hour many parameters and measurements are collected. Some of the parameters may have multiple measurements during the related period. For instance, Heartbeat parameter may have 15 different values measured by the smart device. Number of samples can change depending on device or user activity. This is why a NoSQL repository is quite suitable to store unstructured data. Unstructured data can be thought as raw data. Raw data is processed by a backend process and a summarized clean data is created.

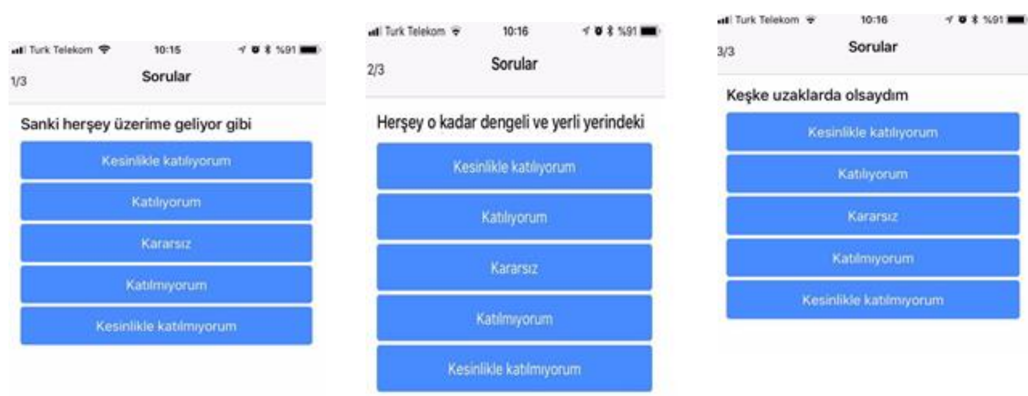


Figure 1.3 Emotion Embedded Questions

There is approximately 1 year of data from 15 users in the repository. Daily basis, System calculates a summarized data for each user. This summarized data is main input for machine learning/deep learning algorithms.

Data Structure of the backend system is given in Figure 1.4. First column represents a data structure of backend repository. Each data is timestamped and may contain multiple measurements as expressed above. Second column shows a data sample stored in data repository. Data is filled by the mobile application via middleware layer. These are raw data and needs pre-processing before machine learning/deep learning stage.


```
{
    "_id" : ObjectId("5b45e50c605ab1086ccae120"),
    "pdId" : "e6d46eb3b54165ef59ad0d1879fd08cb",
    "name" : "D298F0BF-BD86-453E-9483-77AE76B7BDC3",
    "type" : "steps",
    "startDate" : ISODate("2018-07-11T00:00:00.000+0000"),
    "stepData" : [
        {
            "hour" : NumberInt(0),
            "stepCount" : NumberInt(28)
        },
        {
            "hour" : NumberInt(1),
            "stepCount" : NumberInt(0)
        },
        {
            "hour" : NumberInt(2),
            "stepCount" : NumberInt(0)
        },
        {
            "hour" : NumberInt(3),
            "stepCount" : NumberInt(0)
        },
        {
            "hour" : NumberInt(4),
            "stepCount" : NumberInt(0)
        },
        {
            "hour" : NumberInt(5),
            "stepCount" : NumberInt(0)
        },
        {
            "hour" : NumberInt(6),
            "stepCount" : NumberInt(43)
        },
        {
            "hour" : NumberInt(7),
            "stepCount" : NumberInt(17)
        },
        {
            "hour" : NumberInt(8),
            "stepCount" : NumberInt(110)
        },
        {
            "hour" : NumberInt(9),
            "stepCount" : NumberInt(3629)
        },
        {
            "hour" : NumberInt(10),
```

```
{
    "_id" : ObjectId("5b449d2a605ab10ca450c1fc"),
    "hdId" : NumberInt(1531223443),
    "name" : "D298F0BF-BD86-453E-9483-77AE76B7BDC3",
    "timeDiffs" : [
        3.399,
        2.555,
        1.486,
        2.163,
        1.971
    ],
    "totalTime" : 11.574,
    "avgTimeDiffs" : 2.3148,
    "maxTimeDiff" : 3.399,
    "minTimeDiff" : 1.486,
    "errorCount" : NumberInt(7)
}
{
    "_id" : ObjectId("5b44b2b3605ab10cc0164ea0"),
    "hdId" : NumberInt(1531228955),
    "name" : "D298F0BF-BD86-453E-9483-77AE76B7BDC3",
    "timeDiffs" : [
        2.174,
        9.934,
        0.957,
        1.196,
        4.151,
        1.442
    ],
    "totalTime" : 19.854,
    "avgTimeDiffs" : 3.3089999999999997,
    "maxTimeDiff" : 9.934,
    "minTimeDiff" : 0.957,
    "errorCount" : NumberInt(7)
}
{
    "_id" : ObjectId("5b44b345605ab10cc0164ea2"),
    "hdId" : NumberInt(1531229102),
    "name" : "D298F0BF-BD86-453E-9483-77AE76B7BDC3",
    "timeDiffs" : [
        1.418,
        2.987,
        2.283,
        1.517,
        2.96,
        2.133
    ],
    "totalTime" : 1.3297999999999998E1,
    "avgTimeDiffs" : 2.2163333333333333,
```


Figure 1.4 Data Structure of Backend System

A presentation of processed data is given in Table 3.1. Format of Table 3.1 is main input format for machine learning/deep learning layer.

1.3 Related Work

Biometrics are active research area of computer science. In this section related researches those may only contribute to our study has been inspected. Because of the nature of the biometrics most of the studies focuses on security. Earlier studies research either recognition of patterns or identification of the users. Main research areas like Fingerprint, Face Recognition, Keystroke Recognition, Iris Recognition, Handwritten Recognition mainly does the similar things using different approaches. High performance recognition algorithms are main focus. Although pattern recognition is separate and huge area of computer science, this thesis scope is limited to Keystroke Patterns. This is why researches on Keystroke pattern recognition are on focus. Especially studies on keystroke including emotions analysis are detailed inspected because of Keystroke Patterns used as an input parameter in this thesis.

As it is mentioned above, Keystroke Patterns like the other biometric methods started to be investigated for security purposes. Most of the former studies focus on recognition. These former studies mainly focus on a few methods. These methods can be grouped as:

- Minimum distance based methods (Bergadano, Gunetti & Picardi,2002; Bleha,1988; Bleha,1990; Joyce & Gupta,1990)
- Statistical based methods (Song & Perrig,1997; Monrose & Rubin,1997; Monrose & Rubin,2000)
- Data Mining based methods (Gutierrez et al.2002)
- Artificial Neural Networks based methods (Anagun & Cin,1998; Güven, Akyokus & Uysal,2007; Obaidat & Sadoun,1997)

These methods are employed to determine the keystroke dynamics in the state-of-the-art studies. Security first approach is dominant in keystroke pattern recognition. Hence, the usability of keystroke dynamics for hardening password and encryption is represented in (Monrose, Reiter & Wetzel, 2002). As a result of the studies, it is accepted that every user has a unique pattern while using keyboard. This unique pattern is called as Typing DNA. Typing DNA is utilized to differentiate users. The study in (Güven & Soğukpınar,2003), proposes to understand users' behaviours using Key Press Delay (KPD) and Key Stroke Delay (KSD) techniques. The experiment results exhibit that KPD perform well to differentiate users while KSD is competitive. Measuring KPD and KSD patterns gives a glue o the researchers about users' neuro physical conditions. Abnormal situations are observed in KPD as change in pattern as well in KSD. Results obtained in Keystroke pattern recognition show that Keystrokes can be used more than identification. If physical conditions or emotional conditions effect keystrokes than it is also a proof that there is a link between keystroke and emotions. To prove this connection recent studies are conducted. One of the good examples that combines keystroke patterns and emotions researches is the study done by Khanna and Sasikumar (Khanna & Sasikumar,2010). They obtained that Positive/Negative mood of the users' effect keystroke patterns. As a result of the study there is a connection between emotions and keystroke patterns. Emotions are more than Positive and Negative states and can be enhanced as described in Plutchik's wheel of emotions. Another study shows these emotional states detailed(Epp,2011). Epp's study shows confidence, hesitance, nervousness, relaxation, sadness, and tiredness emotional states and they reach 77 to 88% accuracy values as they stated in their study.

Keystrokes and emotions based studies measure keystroke patterns which are basically key delay times while users are typing a text. Anomalies on typing patterns are identified and mapped to emotions. In this thesis, there is another keystroke parameter called Error Count (EC) derived. EC parameter is also employed to enhance the emotion recognition procedure of the users. In addition to keystroke parameters, error counts are also measured to get a better classification. Keystroke pattern parameters are used as behavioural/Neuro Physical part of this study.

At the age of sensors there are many parameters obtained on the regular devices. Cameras or Electroencephalography (EEG) devices are used in many studies to predict users' behaviours. MIT researchers (Vondrick, Pirsiavash & Torralba,2016) developed an algorithm that predicts how two people approaches each other will interact by viewing the images given. The algorithm tries to estimate the next movement of two people by analysing given images. Patterns are recognized and the final movement such as hug, high five, handshake is estimated with the proposed algorithm. This study employs images to forecasts the features are extracted from images and calculated probabilities of the final movement. Electroencephalography (EEG) devices are also utilized in the literature studies. Measuring EEG signals gives an opinion about human emotional conditions. Discrimination between calm, exciting positive and exciting negative emotional states are observed when EEG is used (Giannakaki et al.2017). Nowadays, EEG like parameters can be obtained using Heartbeat data. According to this assumption, heartbeat parameters are measured using ordinary smart watches as a part of this thesis. Heartbeats are also used to evaluate users' emotions. It is known that heartbeat changes depending on conditions. It beats faster in extraordinary conditions and gets slow when in a relax environment. Even listening classical music gives an overall insight about user emotional condition because of affecting heartbeat (Ramirez et al.2018). While first studies use electromechanical film (the EMFi chair) sensors to get parameters (Anttonen & Surakko,2005), recent technology is capable to obtain these parameters using ordinary smart watches. In another study (Harper & Southern,2019), authors focus on the just heartbeat dataset to interpret the emotions of user. For this purpose, Bayesian algorithm is employed as classification algorithm. They conclude that the usage of heartbeat dataset and Bayesian algorithm for the classification purpose perform 90% accuracy value. In (Valderas et al.2019), extensive experiments are conducted on the heartbeat dataset to identify psycho-neural illnesses. Heart rate variability (HRV) is used to differentiate several emotional states which may be an indicator for illness. In another study (Zhu et al.2019), researchers conclude their study in that the heartbeat is a good indicator for identifying disorders and evaluating emotions is a key value to identify emotional disorders.

Sleep quality is one of the parameters of this study. There is a strong relation between sleep quality and emotional condition. The study (Rehman et al.2018) demonstrates that insufficient sleep duration or sleep quality may cause many problems including negative emotions, perceptual anomalies even paranoia. In another study (Garett et al.2016), authors concentrate on to show the relation between social media usage and sleep quality by observing twitter messages of college students for the purpose of analysing emotional states of students. Experiment results demonstrate that the usage of social media is associated with sleep quality among students. Sleep quality has also similar effects even in older ages. In (Becker et al.2018), it is claimed that increasing sleep quality creates mediation effect in older adults.

Motion/Energy and walking posture parameters are also considered to identify the user emotions. Walking posture is easily observable and changeable when people is under stress. In the study (Hackford et al.2019), authors emphasize that walking style also impacted physiological states during stress and changing the way a person walks may improve their responses to stress. Human gait analysis is also one of the popular research area in emotion identification. In (Venture et al.2014), researchers clarify that it is possible to identify several emotion conditions like neutral, joy, anger, sadness, and fear. The experimental results represent that the use of the gait analysis characterize each emotion which is supported by accuracy results.

Sentiment analysis is recently popular research field. There are too many researches to specify and classify users' feelings by using social media platforms. It is usually difficult to analyse texts for interpreting emotions. Textual information can be generally divided into two main types: Facts and Opinions. Facts are objective expressions about items, events and their properties. Opinions are usually subjective entities that describe people’s sentiments, appraisals or feelings toward entities, events and their properties

(Liu,2010). In this study, the main focus is not to implement sentiment analysis. As a part of our thesis, similar to sentiment analysis, predefined emotion attached questions are asked to the users. Answers are collected and logged. Thus, not the opinions but the facts are obtained by interacting with the users. The users' answers are mapped to the emotions and correlated with physical and neuro physical parameters.

Our work in this thesis differs from the above mentioned literature studies. Our main purpose is to approach the problem not only as a computer science problem but also a psychology problem. This makes the thesis an interdisciplinary study. There are many researches exist in literature under psychology. Most of them are well known theories. On the other hand, number of studies on emotional analysis in computer science are also increasing. We believe that combination of disciplines may create huge impact on the field. Section 2 has detailed explanation about Emotions from psychology perspective. Section 2 is also covers a large history of Biometrics.

Other differentiator of our study is that there are many different types of parameters collected in order to increase the classification accuracy of the study. In this thesis, not only physical but also behavioural/neuro physical parameters, and the combination of them are evaluated while literature studies above are usually focuses on one or few parameters. Moreover, inclusion of deep learning models in addition to the traditional machine learning algorithms boosts the classification performance of the proposed system. This is why our study is the very first attempt to analyse neuro physical states of the users.

2. HISTORY OF BIOMETRICS & EMOTIONS THEORY

History and important milestones of biometrics as a part of computer science and emotions as a part of psychology is investigated in this section. Although they are quite different fields it can be seen that they have very old and long history in their life time.

2.1 Definition of the Biometrics

The term "biometrics" is derived from ancient Greek words "bio" (life) and "metrics" (to measure). Modern biometric systems exist for the last a few decades. Developments in computer technology led development of automated biometric systems. Many of these automated methods are based on ideas that were originally conceived hundreds of years ago.

Human face may be the first biometric identifier for thousands of years. Since the beginning of civilization, people have used faces to identify unknown (unfamiliar) and known (familiar) individuals. When population increased, it became more challenging task to understand new individuals. Some populations used special signs or colours to differentiate each other's. Using such signs was an indicator but seen not sufficient by time. Behavioural biometrics is also used in human-to-human recognition. Speaker or gait recognitions are used for this purposes. People use these behavioural characteristics to recognize known people on a day-to-day basis.

By the mid-1800s, there was a need to recognise people faster because of rapid population growth of cities due to industrial progress/revolution and developing techniques in farming. A Search to find a new method started in South America, Asia, and Europe. It was especially a need for police departments. At the end of 1800s a method was developed to use fingerprints in identification. It was similar to Bertillon's method but that was using only fingerprint patterns instead of other body measurements. The first system used by a government to index fingerprints was developed in India by Azizul Haque. He developed a system for Edward Henry who was working as inspector at Police department of Bengal, India. This system, named as Henry's System, and its variations are still used for classifying purposes.

True biometric systems started to spread in second half of the twentieth century, depending on the emergence of computer systems. Biometric applications has made an explosion of activity in the 1990s and began to be popular part of everyday applications after 2000s.

2.2 A Brief History of the Biometrics Technologies

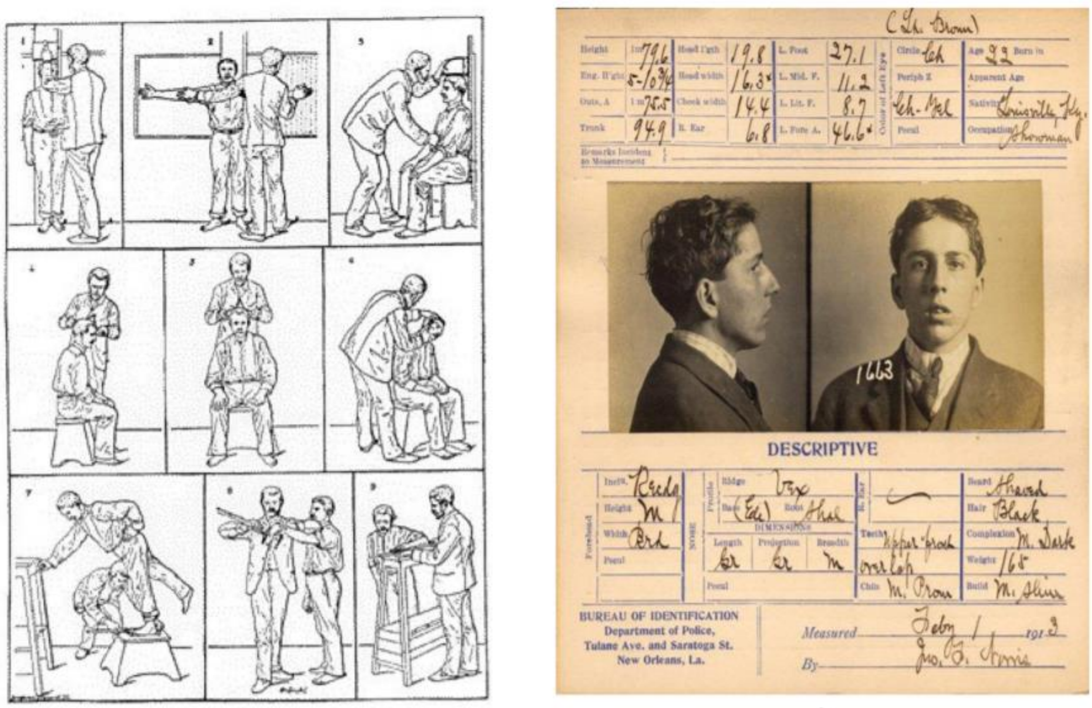
Tom Topol has made a remarkable timeline for history of Biometrics. (Tom Topol,2012). The time line gives a perfect chronological information about each biometrics method. Following time line is based on Tom Topol's study.

1858 – Hand Images Recorded to Identify People

Sir William Herschel who is working for Civil Service of India used to track workers' activity using their handprint and finger print records. This was first employee tracking system which is using biometric identification. Workers payments were done prior to records in this identification system at the end of the month.

1870 – Bertillon used body measurements to identify individuals

Alphonse Bertillon used individuals' body measurements to identification. This is called as "Bertillon age" or anthropometries. Bertillon made body measurements of individuals when they were arrested. Bertillon assumed that although they could change their identities, they could not change their body shapes. Police authorities throughout the world utilized in the system. The system lost its popularity when it was realized that some people could have the same metrics.



a b

Figure 2.1 Biometric measurements at earlier times

a Shows how measurements were taken. b shows identification card based on measurements.

1892 – Galton develops the first classification system using fingerprints

First fingerprint classification system developed by Sir Francis Galton. He presented a new classification method using prints of all ten fingers. These characteristics are also called minutiae. These characteristics are often referred to as Galton's details. Galton system is still used today.

1894 – Fingerprint concept used in a novel of Twain

Increasing popularity of fingerprints affected literature. First time, Mark Twain mentioned the use of fingerprints for identification. In his novel called Tragedy of Pudd'nhead Wilson, a man proved his innocence using fingerprint comparison. Fingerprints have been a part of many stories and movies by the time.

1896 – Henry develops a fingerprint classification system for Government

Bertillon system was difficult to implement and needed too many body part measurements. Sir Edward Henry who was Inspector General of the Bengal Police, was in search of a method of identification to replace Bertillon system. They decide to replace Bertillon system after Galton's study on fingerprints. One of Henry's workers, Azizul Haque, developed a method of classifying and storing the information so that searching could be performed easily and efficiently. Developed system has also been implemented in London as the first British fingerprint system. The Henry Classification System became a base for many systems used by the governmental institutions like the Federal Bureau of Investigation (FBI). FBI used a similar system for many years to perform fingerprint searches.

1903 – NY State Prisons use fingerprints for identification

Commission of The New York Civil Service started to use fingerprint identification system to prevent examination frauds. They stopped examination fraud using fingerprint identification. This practice was also used by the New York state prison system. Criminals' fingerprints were taken and these records were used for the identification of criminals in 1903. After the success of the practice, system started to be used many parts of the United States. Many penitentiaries like Leavenworth, Kansas, St. Louis, Missouri used the same system. Growing need and popularity of using fingerprint systems increased use of fingerprint systems in many other police departments and investigation bureaus. As a result of this trend, the US Congress established the Identification Division of the FBI.

1903 – Bertillon System fails

Bertillon system is based on body measurements. Two men who were the exactly same measurements, were sentenced to the US Penitentiary at Leavenworth, Kansas. This was a failure for Bertillon system and created too many questions for reliability of the system.

1936 – First conceptual design for using Iris identification

An Ophthalmologist named Frank Burch claimed that iris pattern is unique for all individuals. He proposed a concept study to use iris patterns as a new recognition system.

1960s – First computer aided Face recognition system

The first semi-automatic face recognition system was designed by Woodrow W. Bledsoe. Bledsoe used pictures as main input. System was using face parts like ears, nose, eyes etc. as features. These face parts and their calculated ratios or distances among them were used as feature points.

1960 – First speech identification system

First acoustic speech identification study was published by Guntar Fant. Guntar tried to understand physiological components of speech. He used x-rays to understand these components. He recorded images of individuals while they were making phonic sounds. These findings were crucial to speaker recognition and these findings used to understand dynamics of speech recognition.

1963 – Fingerprint automation proposed by Hughes

In 1963, Hughes proposed used of fingerprints for automation purposes. An automatic system which recognises fingerprints using an internal database was proposed.

1965 –First computer-aided signature recognition system

The first signature recognition system is established in 1965 by North American Aviation. This was one of the first publicly available system.

1969 – FBI automates fingerprint process recognition

In 1969, the Federal Bureau of Investigation (FBI) decided to shortened process of fingerprint identification system. It was becoming destructive and needed many man-hours. National Institute of Standards and Technology (NIST) is assigned to solve this problem by FBI. NIST identified two key problems:

1- Scanning fingerprint pictures and extracting fingerprint patterns
2- Compare and matching of fingerprint patterns

1970s – First face recognition system

The first automatically face recognition system is proposed by Goldstein, Harmon, and Lesk. Earlier solutions were using manual computations. They used 21 different markers to identify face. Hair colour, lip thickness, nose length etc. were used as one of these 21 markers.

1970 – First behavioural parts of speech recognition modelled

First acoustic speech recognition was developed in 1960. It was using mouth pattern while speech. This study is extended by Dr. Joseph Perkell. He used x-rays to model speech by adding jaw and tongue movements. The model was used to understand complex behavioural and biological components of speech.

1974- First hand geometry system for commercial

After development of fingerprint recognition systems, the first commercial hand geometry systems became available at the market in 1970s. Hand geometry systems were used three areas:

1- Access control for physical areas
2- Time management and attendance control
3- Identification of People

1975 – FBI funds development of sensors technology

The biggest problem of biometric systems was to store extracting minutiae. Devices those store minutiae were capacitive and had limitations to store all of the features. Especially image storing and compression was problematic area. To solve this problem FBI worked with NIST. NIST developed the first algorithm called M40. M40 algorithm was the first algorithm used by FBI to search fingerprint data. M40 was storing images in a smaller size thus search and classification time was shortening.

1976 – First speaker recognition system is established

Texas Instruments designed and developed a prototype speaker recognition system. Developed system's tests were done by the US Air Force and the Mitre Corporation

1977 – Dynamic signature information is patented

An apparatus named "Personal identification apparatus" has been designed by Veripen Inc. The company awarded a patent for their device. System was collecting dynamic pressure information. Device was used to capture of dynamic characteristics of an individual's signature patterns. This is a pioneer for handwriting recognition systems. This system is developed for the Electronic Systems Division of the United States Air Force.

1980s – NIST Speech Group is formed

The National Institute of Standards and Technology (NIST) Speech Group is established to improve speech processing techniques. NIST Speech Group is hosted and event called "NIST Speaker Recognition Evaluation Workshop" yearly to review evaluations since 1996.

1985 – Concept of irides are unique

Ophthalmologists, Drs. Aran Safir and Leonard Flom are proposed a concept. According to their study irides are unique and no two irides are alike.

1985 – First patent for hand geometry recognition

Although first commercial applications of hand geometry are based on the early 1970s by University of Georgia in 1974, The US Army began testing this technology to use in banking operations. Studies showed that hand geometry recognition has a potential to be used in banking industry. Hand identification was patented by David Sidlauskas.

1986 – Exchange of fingerprint data standard is released

Usage of fingerprint data all around the world created a need to interchange fingerprint data among organisations. In order to standardize data exchange NIST created a standard called ANSI/ NBS-I CST 1-1986. This was the first version of current fingerprint interchange standard. Enhanced version of this standard is still used today.

1986 – Patent for iris identification

Drs. Aran Safir and Leonard Flom were awarded a patent for iris recognition concept. It was known that iris was a unique property for each individual like fingerprints. Dr. Flom approached Dr. John Daugman to develop an algorithm for iris recognition.

1988 – First computer-aided facial recognition system is established

The Lakewood Division of the Los Angeles County Sheriff's Department started to use pictures and video images to search suspects from a database. Suspects' pictures were digitalised and checked with references in the database.

1988 – Face recognition uses Eigenface technique

Kirby and Sirovich proposed a new approach to face recognition problem. They used linear algebra technique called Eigenface to recognise a face. This was a milestone because It shortened face recognition time. Using less than 100 measurement points was making possible face recognition.

1991 –Real-time face recognition

Concept of residual error could be used to detect faces in images. Turk and Pentland discovered residual error technique while they were working on the eigenfaces techniques. As a result of this approach, real time automated face recognition was possible.

1992 – Biometric Consortium is established within US Government

Biometric Consortium is established by The National Security Agency in 1995. Biometric Consortium was limited to government agencies. Industry and academia were not a part of the consortium. They were attending as observer. With the explosion of biometric activities in the early 2000s, consortium expanded and accepted those non-governmental organisations as members. The Consortium itself remains active as a key liaison and discussion forum between government, industry, and academic communities.

1993 – Face Recognition Technology (FERET) program started

The Face Recognition Technology (FERET) program was started by the Defence Advanced Research Products Agency (DARPA) and the DoD Counterdrug Technology Development Program Office. Main purpose of the program was to support face recognition technologies from beginning to commercial stage.

1994- First iris recognition algorithm is patented

First iris recognition algorithm is patented by Dr. John Daugman. This patent is used by many iris based commercial products.

1994 – Competition for Integrated Automated Fingerprint Identification System (IAFIS)

A competition was held to establish the Integrated Automated Fingerprint Identification System (IAFIS). There were three major challenges;

1- Acquisition of Digital fingerprints
2- Extraction of Local ridge characteristics
3- Matching of Ridge characteristics

As winner of the competition, Lockheed Martin was winner of the competition and selected as contractor to build the FBI's IAFIS.

1994 – Benchmarking of Palm System

According to a common assumption, the first palm print based system was designed by a Hungarian company named Recoware. Their system was called RECOderm™. In order to get a fast technology transfer, Lockheed Martin Information Systems bought Recoware in 1997.

1994 – INSPASS is implemented

Biometric systems are used by The Immigration and Naturalization Service Passenger Accelerated Service System (INSPASS). Travellers who have biometric cards were no need to wait for long inspector control while they were entering to the US. When their cards stored their biometric information and it was easy to check identification when they have cards with them. This service ended in late 2004.

1995 – Iris prototype becomes commercial

A joint project between Defence Nuclear Agency(DNA) and Iriscan company created the first commercial iris recognition product.

1996 – Hand geometry is implemented at the Olympic Games

1996 was one of the important year for biometric history. One of the biggest public usage of hand geometry occurred at the Atlanta Olympic Games. System were used by 65.000 people and more than 1.000.000 transactions occurred in a period of 28 days.

1996 – NIST supports speaker recognition

the National Institute of Standards and Technology (NIST) established a Speech Group. They started to host yearly evaluations in 1996 for speech recognition. Aim of the group is to track improvements on the field and enlarging speaker recognition community.

1997 – Generic biometric interoperability standard is published for commercial applications

In order to synchronize several vendors' efforts in the biometric field a generic biometric interoperability standard released. This standard was sponsored by NSA and named as the Human Authentication API (HA-API). It was a major regulation for biometric technologies.

1998- COOlS (DNA forensic database) is created by FBI

The FBI created Combined DNA Index System (CODIS) to digitally store DNA information. Using this system DNA information can be queried for forensic purposes.

1999 – Machine readable travel documents embeds biometrics

The International Civil Aviation Organization's (ICAO) Technical Advisory Group on Machine Readable Travel Documents (TAG/MRTD) started a study to implement biometric parameters into passports and other valuable documents. Aim of this study was create papers those store its user's biometric properties inside.

1999 – FBI's IAFIS major components are live

The FBI's ten-fingerprint system (IAFIS) designed by Lockheed Martin is operational in 1998. System is capable to store large amount of fingerprints and search capabilities.

2000 – Vendor Test (FRVT 2000) for First Face Recognition

Multiple US Government organisations sponsored the Face Recognition Vendor Test (FRVT) in 2000. Purpose of FRVTs was to evaluate commercially exists technologies on the field. FRVTs held many times after 2000. FRVTs in 2003 evaluated fingerprints and iris recognition in 2006.

2000 – First study published on vascular patterns for recognition

This paper proposed usage of blood vessel pattern for identification purposes. According to the study, blood vessels are unique for each individual.

2000 – West Virginia University opens biometrics degree program

West Virginia University (WVU) opened the first bachelor's degree program in Biometric Systems in 2000. Biometric-related courses exist many years but it was the first program opened independently. WVU encourages program participants to obtain a dual-degree in Computer Engineering and Biometric Systems as the biometric systems degree is not accredited.

2001 – Face recognition is used at the Super Bowl

A face recognition system was installed at the Super Bowl in January 2001 in Tampa, Florida. System was mainly looking for individuals who are not allowed to enter the stadium. System made some errors and this created privacy concerns into the consciousness of the general public.

2002 – Biometrics in ISO/IEC standards committee

The International Organization for Standardization (ISO) established the ISO/IEC JTC1 Subcommittee 37 (JTC1 /SC37). Main purpose of subcommittee is to create standards for generic biometric technologies. The Subcommittee also develops standards for interoperability. Interoperability is a key concept for data interchange between applications and systems.

2002 – M1 Technical Committee on Biometrics is established

The M1 Technical Committee on Biometrics is formed. This technical committee reports to the Inter National Committee on Information Technology Standards (INCITS), an accredited organization of the American National Standards Institute (ANSI).

2002 – Palm Print Staff Paper is submitted

A staff paper about palm print technology and Integrated Automated Fingerprint Identification System (IAFIS) was submitted to the governmental organizations including Identification Services (IS) Subcommittee, Criminal Justice Information Services Division (CJIS) Advisory Policy Board (APB). Adding strength of palm print capabilities to IAFIS system has been approved and the FBI started a new initiative called Next Generation IAFIS (NGI). Main part of NGI was to develop requirements for an integrated National Palm Print Service.

2003 – US Government coordinates biometric activities

The National Science & Technology Council (NST), a US Government cabinet-level council, formed a subcommittee on biometrics to coordinate biometrics Research and Development, policy, outreach, and international collaboration.

2003 – ICAO accepts to integrate biometrics into MRTDs

The International Civil Aviation Organization (ICAO) started to a study to integrate biometric identification information to passports on 2003. Other Machine Readable Travel Documents (MRTDs) were also kept in scope. Face recognition was selected as the globally interoperable biometric property.

2003 – European Biometrics Forum is formed

The European Biometrics Forum is an independent organisation of European Commission. Main purpose of the forum is to make EU a leader in Biometric technologies. The forum also acts as the driving force for coordination, support and strengthening of the national bodies.

2004 – US-VISIT program becomes operational

The United States Visitor and Immigrant Status Indication Technology (US-VISIT) program is the cornerstone of the DHS visa issuance and entry exit strategy. Biometric parameters are used in identification as a part of this system. US-VISIT creates a control mechanism on travellers to understand if the traveller had any security issue. Former records are stored with biometric information to track pervious border transactions of the user.

2004 – DOD implements ABIS

The Automated Biometric Identification System (ABIS) is a Department of Defence (DoD) system. System has ability to track and identify national security threats. System collects many biometric parameters from enemy combatants.

2004 – Government-wide personal identification card for all federal employees

President Bush issued Homeland Security Presidential Directive 12 (HSPD-12) in 2004. This directive gives an obligation to all federal employees and contractors to have government-wide personal identification card. This cards holds biometric information of holders.

2004 – First state-wide automated palm print databases in the US

California, Connecticut and Rhode Island established state-wide palm print databases in 2004. These databases help governmental agencies in several states to search and share palm prints among them. Database holds all information about usual suspects and their corresponding palm print patterns.

2004 – Face Recognition Grand Challenge begins

US Government sponsored a face recognition contest called "The Face Recognition Grand Challenge (FRGC)". Participating researchers analyse the given data, try to solve the problems, and then reconvene to discuss various approaches and their results.

2005 – US patent for iris recognition concept expires

The US patent which is protecting the iris recognition concept expired in 2005. This brought new opportunities for the companies those are working on iris recognition area. However, implementation patent of iris identification would expire on 2011. Implementation patent is owned by lrisCodes® company.

2005 – Capture Iris data while walking

SRI International company introduce a new system in 2005. System enables the collection of iris images from users walking through a portal. This is a novelty because the system has capability of collecting data from moving individuals.

2008 – U.S. Government coordinates biometric database

Fingerprint and Face recognition measurement algorithms and related technological developments were finalized. An iris identification algorithm was also developed. The FBI and Department of Defence also started a next generation database design to combine all of the biometric information.

2010 – Use of biometrics for terrorist tracking

A fingerprint evidence collected from 9/11 event area was matched with a GITMO detainee. Same technique was also used to identify other 9/11 suspects. Integrated Automated Fingerprint Identification System (IAFIS) system shortened to evaluation process.

2011 – Biometric identification used to identify body of Osama bin Laden

Osama bin Laden identified with 95 percent certainty using facial recognition and DNA recognition technologies after his death.

2013 – Apple includes fingerprint scanners into regular smartphones

Apple Inc. has placed many biometric sensors into its products. Fingerprint recognition feature allows users to unlock their devices event make purchases in the various Apple digital media stores (iTunes Store, the App Store, iBookstore).

Today, there are many sensors are a part of our daily life. Even cell phones are capable to recognize not only our fingerprints but also face or even our voice. Speech-to-text capabilities of the devices became an ordinary feature. Most of us are not using old fashioned password but instead we are using out biometric properties which are much more reliable for identification.

2.3 Understanding Emotions

Emotions are basically focus of psychology. There have been many studies conducted last a few hundred years. Detection of emotions is quite old area. One of the earliest studies is done by Charles Darwin in 1872, just thirteen years after his famous book "The Origin of Species". Darwin has contribution to many fields of science, from evolution to geology to botany, are well-known. Darwin used photography to classify emotions. This approach is called Darwin's Camera. Results of the study was published by Darwin as "The Expression of the Emotions in Man and Animals" (Darwin,1872). It was one of the first scientific texts to use photographic illustrations. In his study, Darwin investigated emotions and effect of emotions to human face. There are many photos of different emotional states taken and classified. Some of the photos are given below.

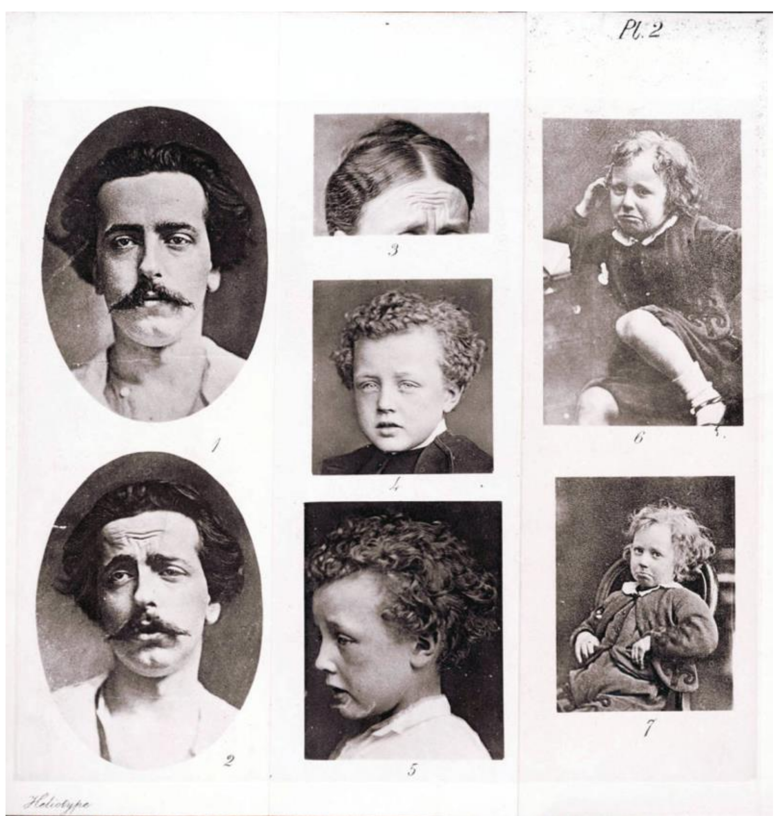


Figure 2.2 People who are tired and natural

Figure 2.2 shows how people look like when they are natural and under tired. Their faces change when they are not in natural state. Darwin used different categories of human beings in his study. Figure 2.3 shows infants in several emotional states.

Figure 2.3 Infants: Suffering and Weeping

Elder people emotions are also used in Darwin’s study. Aging may effect emotional behaviours. While infants may cry while they suffer older people may show different reactions.

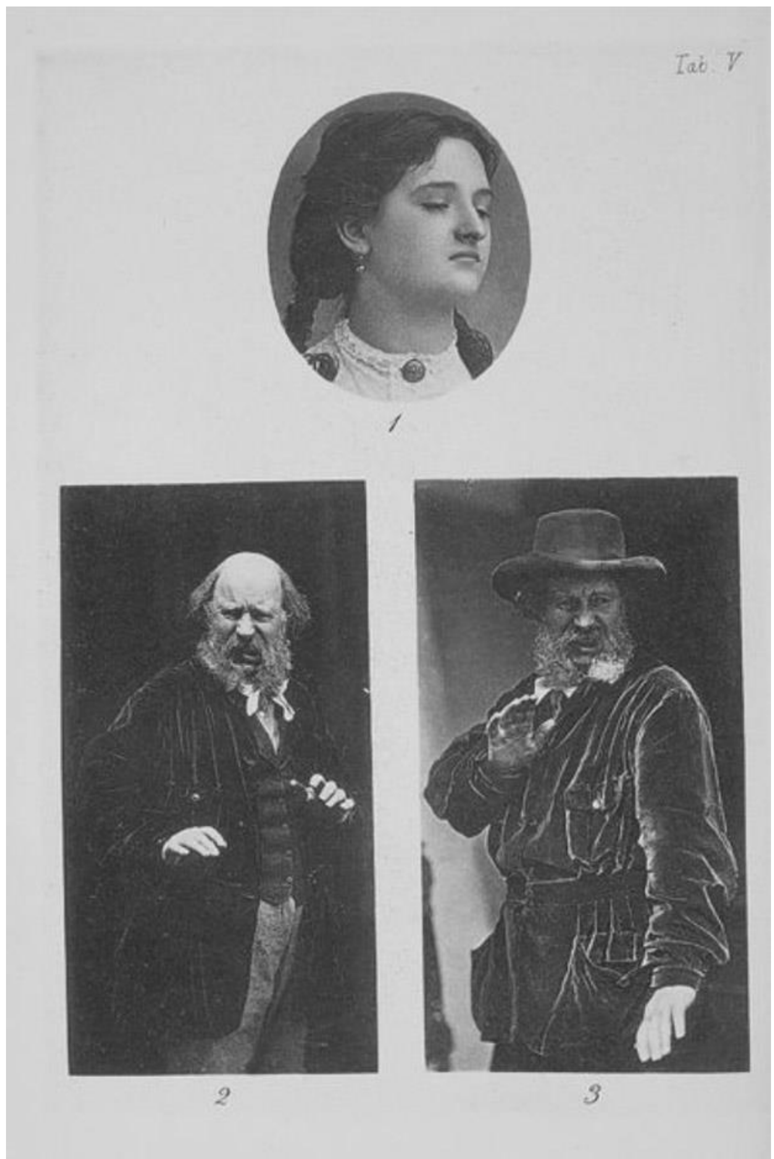


Figure 2.4 Disdain, Contempt, and Disgust

Darwin showed in 1872 that emotions are predictable when analyse people or animals' physical reactions. Darwin tried to show that some circumstances may create similar reactions in human beings and animals. His study was generally in physical domain. On the other hand emotion is a mental state associated with the nervous system brought on by chemical changes variously associated with thoughts, feelings, behavioural responses, and a degree of pleasure or displeasure. Not only physical condition but biometric parameters may effect emotional conditions. To relate biometric parameters to emotions is difficult because of the nature of the emotions. It is also difficult to model or measure emotions. It is needed to have a scale to understand or measure emotions. This scale is developed by Robert Plutchik. He was a psychologist and developed a psychoevolutionary theory of emotion.

Plutchik's theory offers a method to understand primary emotions, and dyads. His theory of basic emotions has 10 postulates:

1. Emotions are applicable to all animals and humans.
2. Emotions have many forms in species and historically evolves.
3. Emotions help organisms to keep them alive.
4. Emotions can be expressed differently depending on species but still they might have certain common elements.
5. There is a small number of basic, primary, or prototype emotions.
6. Emotions can be combination of other emotions/feelings. Combinations may create new emotions.
7. Primary emotions are assumptions or idealized statuses of some characteristics can only be inferred from various kinds of evidence.
8. Primary emotions can be understood in terms of pairs of polar opposites.
9. All emotions vary in their degree of similarity to one another.
10. Each emotion may have many intensity or levels.

Emotions can be thought as states of feelings. Combination of feelings creates emotions. For example, happiness can be described as a combination of joy and pleasure. According to Psychologist Robert Plutchik, number of basic emotions are eight. They are: trust, joy, surprise, fear, sadness, anticipation, disgust and anger. Plutchik invented the wheel of emotions, which illustrates the various relationships among the emotions.

Although Plutchik defines only eight primary emotions, as it can be seen in the wheel there are several different degrees corresponding to each emotion. Besides, combination of emotions can create new emotions. This makes emotions complex than most people realize.

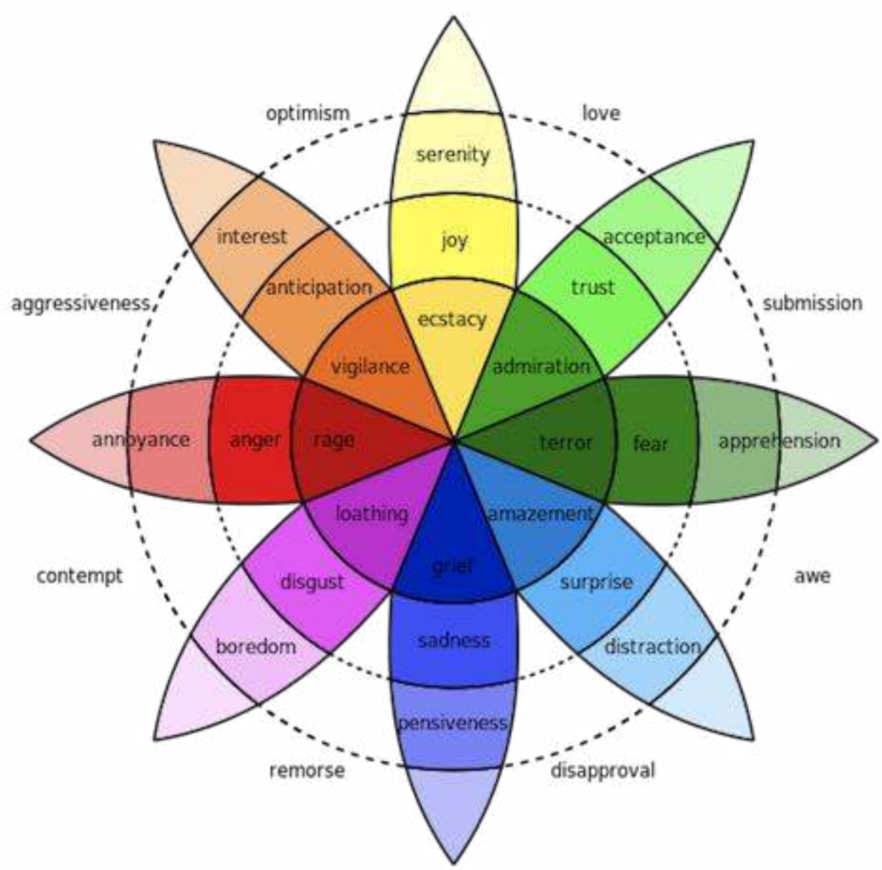


Figure 2.5 Plutchik's Wheel of Emotions

As we stated earlier, Plutchik's eight basic emotions are fear, joy, trust, anticipation surprise, sadness, disgust and anger. Each basic emotion also has a corresponding(polar) opposite, they are:

- Sadness is the opposite of Joy.
- Anger is the opposite of Fear.
- Surprise is the opposite of Anticipation.
- Trust is the opposite of Disgust.

Definition of polar emotions and inter emotions between basic emotions gives a perfect scale to work on emotions. When one wing of the wheel is checked we can access different level of the same emotion; from strong towards weak. Centre of the wing indicates a stronger feeling while edges indicate weaker emotions.

Emotions also creates new emotions when they have an interaction(dyads). For instance, if joy + trust occurs at the same time this may create “love”. A Full presentation of dyads is given in Figure 2.6.

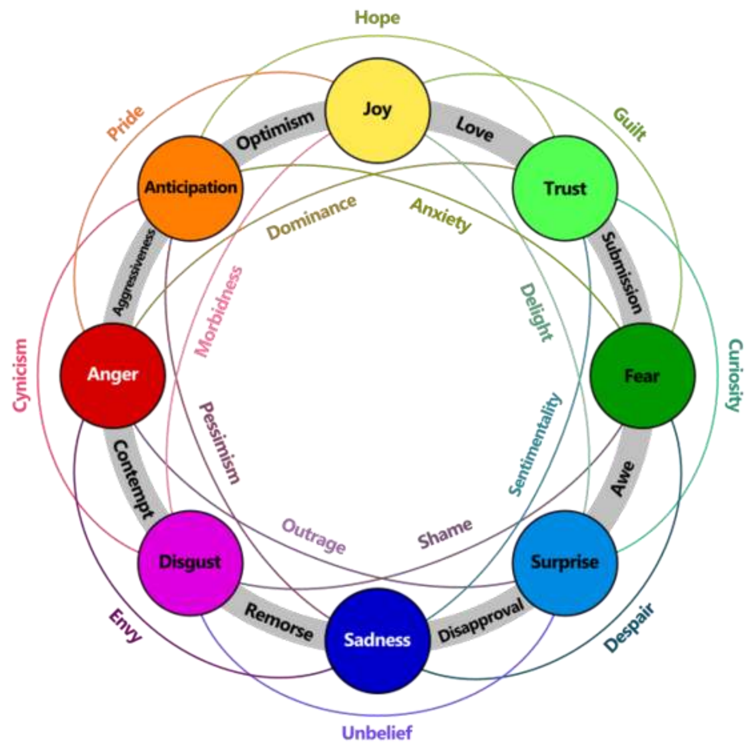


Figure 2.6 Dyads/Combination of emotions.

8 main feelings are like alphabet letters. When they come together they create new emotions as a result of feeling combinations. In Figure 2.6 main feelings Disgust and Fear creates Shame emotion. Positive and negative feelings create stronger emotions than their underlying feelings. Opposite feelings may be combined. When a positive + negative feeling is combined new emotions are created. For instance, when Anger and Joy is combined Pride emotion occurs or another example as an edge emotion Disgust and Joy combines to create morbidness emotion.

A detailed examples of dyads with several combinations of feelings are given in Table 2.1.

Table 2.1 Dyads Combination of Feelings/Emotions

Human Feelings	Emotions	Polar Feelings	Emotions
Optimisim	Anticipation & Joy	Disapproval	Surprise & Sadness
Hope	Anticipation & Trust	Unbelief	Surprise & Disgust
Anxiety	Anticipation & Fear	Outrage	Surprise & Anger
Love	Joy & Trust	Remorse	Sadness & Disgust
Guilt	Joy & Fear	Envy	Sadness & Anger
Delight	Joy & Surprise	Pessimism	Sadness & Anticipation
Submission	Trust & Fear	Contempt	Disgust & Anger
Curiosity	Trust & Surprise	Cynicism	Disgust & Anticipation
Sentimentality	Trust + Sadness	Morbidness	Disgust & Joy
Awe	Fear & Surprise	Aggresiveness	Anger & Anticipation
Despair	Fear & Sadness	Pride	Anger & Joy
Shame	Fear & Disgust	Dominance	Anger & Trust

Table 2.1 also represents feelings and polar feelings. Feelings are combinations of emotions. Dyad and polar dyads shows that there is a transition among emotions and feelings. This structural approach makes Plutchik's theorem valuable to understand emotions and their relationships.

3. PARAMETERS OF THE STUDY

Proposed model for the study is given in Figure 3.2. At first, data collection is performed by collecting different type of information, namely physical and neuro physical from the users. The dataset is composed of data gathered from fifteen users in 25-35 age range for nearly 365 days. A smart watch which is connected to intelligent phone is used in physical data collection from sensors. The mobile application is used to measure emotional states of the users. Emotion embedded questions are asked to the users periodically. Each question has a hidden emotional value. It is evaluated as 1 if the user has positive answer to the question, and negative answer is evaluated as -1. Emotional value is between -4 and 4. It is calculated as sum of these answers as seen in Figure 3.1. Emotional value and corresponding feeling is given Figure 3.1.

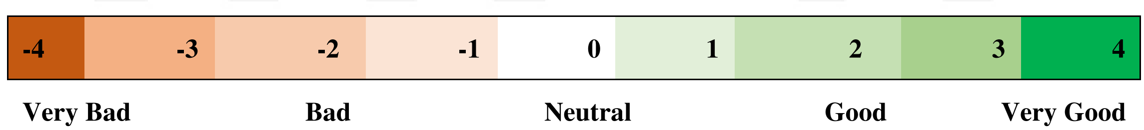


Figure 3.1 The range of feeling values.

The mobile application sends all the collected data to central repository. Central repository keeps data with a timestamp by adding an anonymous user id. The raw data is collected hourly basis. Before machine learning stage, summarization task runs and data is summarized to daily basis. Summarization task creates one row of data for each user for each day. Approximately 365 x 15 matrix of data is created in a year for all user. This matrix has all physical and behavioral data for all 15 users as it is shown in Table 3.1.

Table 3.1 Summary of samples on collected data in the repository.

Total Time(sec)	EC	Avg(sec)	Energy(kcal)	Sleep (hour)	Step	Heart Beat	Feeling
15	3	1.875	550	9	1550	64	-2
12	2	1.714	330	6	1660	63	2
13	3	2.166	302.5	6	1605	67	-1
9	2	1.5	825	3	1825	74	2
16	3	2.285	357.5	4	1715	63	-1
12	2	1.5	302.5	8	1605	58	-3
7	0	1.000	256.6	7	1770	82	2
7	0	1.166	550	6	1550	77	2
11	2	1.375	330	3	1660	76	-1
16	3	2.285	550	8	1550	60	-1
15	2	2.112	425	7	1500	65	1
18	1	2.312	575	8	1600	63	3
21	3	2.335	482	8	1490	67	2
12	1	1.812	678	7	1790	69	1

Data samples are scalar format. They are collected and processed as seen in the flow at Figure 3.2. Data Flow goes as it is shown on the figure. Raw Data coming from sensors and users becomes information by processing on each stage.

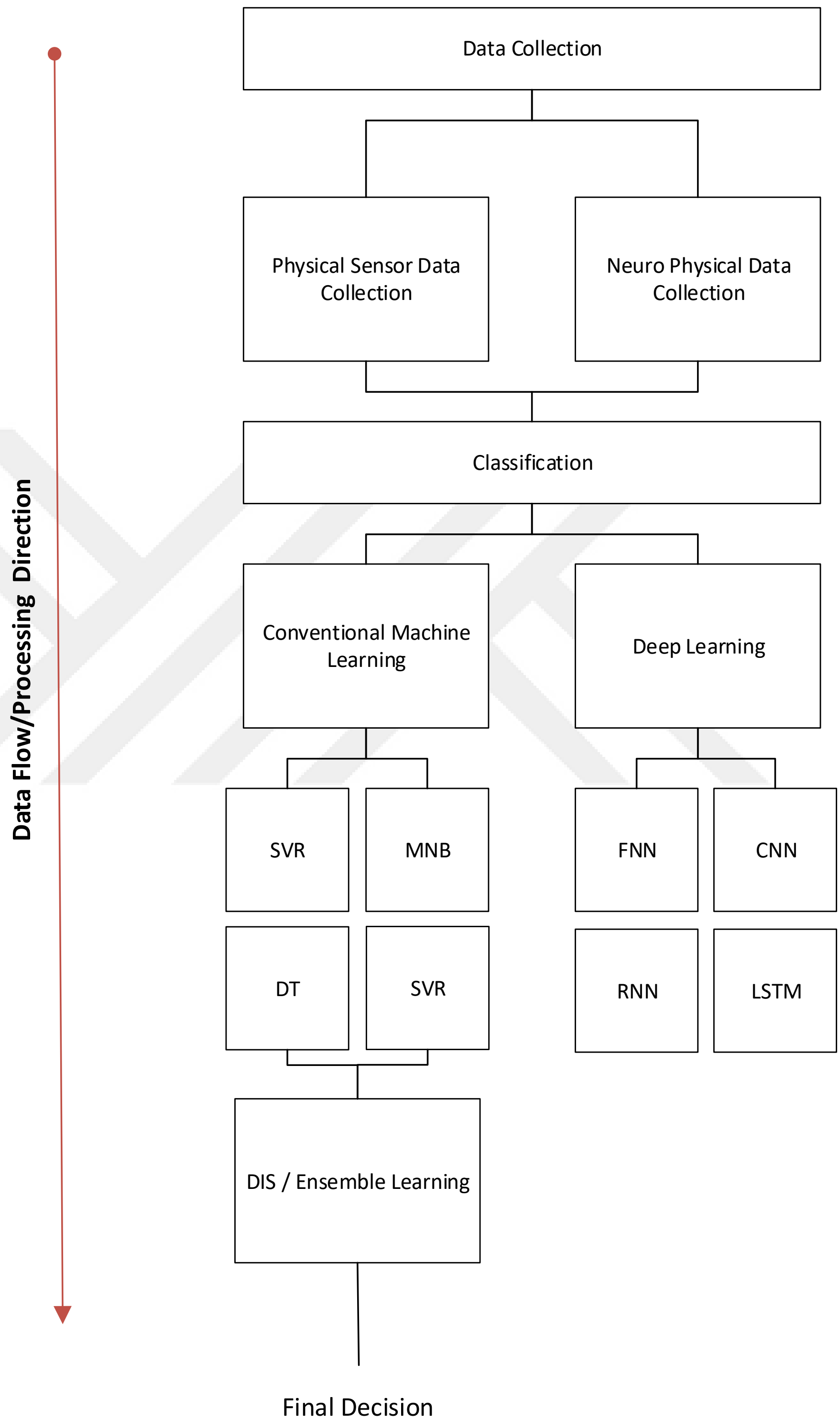


Figure 3.2 Flowchart of the Proposed system

After gathering collection of the data as physical and neuro physical, second step is to process employing machine learning libraries included in Python. Before processing data with machine learning algorithms, dataset is divided into two halves, randomly. 33% of the dataset is employed for the training process and the remaining part is used as the test data. After the data separation step, multinomial naive Bayes, support vector regression, decision tree, random forest, feedforward neural network algorithms, and decision integration strategy algorithms are performed on the data. Our main purpose is that we intend to measure classification performance of artificial neural network against to all machine learning algorithms and decision integration strategy to analyze neuro physical conditions of the users.

3.1. Neuro Physical Parameters

In this thesis, we concentrate on the keystroke dynamics and error count measurements as neuro physical parameters.

3.1.1 Keystroke Dynamics Measurement: Keystroke Dynamics is about to measurement of keystroke timings. Identification is one of the main study field. There are three approaches for keystroke identification, namely Euclidean distance measure (EDM), non-weighted probability, weighted probability. These are the most common methods used in the literature. Distance between pattern vectors is measured with EDM approach. Assume that there are two pattern vectors where *X* stands for real time vector and *M* signifies the reference vector of user. Mathematical equation of two *N* dimensional vectors written as follows:

$$X = [x_1,x_2,x_3,x_4,x_5,\ldots,x_n] \tag{3.1}$$

$$M = [m_1,m_2,m_3,m_4,m_5,\ldots,m_n] \tag{3.2}$$

$$D(X,M) = \sqrt{\sum(x_i - m_i)^2} \tag{3.3}$$

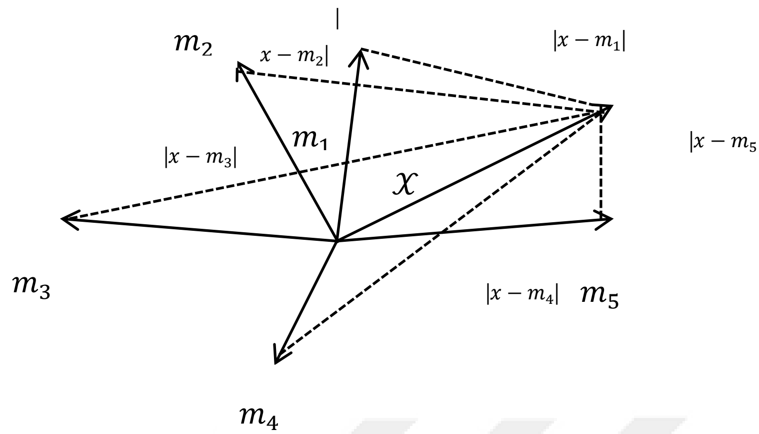


Figure 3.3 Sum of error calculation between vectors in EDM.

In this study we used Keystroke dynamics to calculate TotalTime parameter. Sum of measured KPD and KSD parameters are stored as TotalTime parameter.

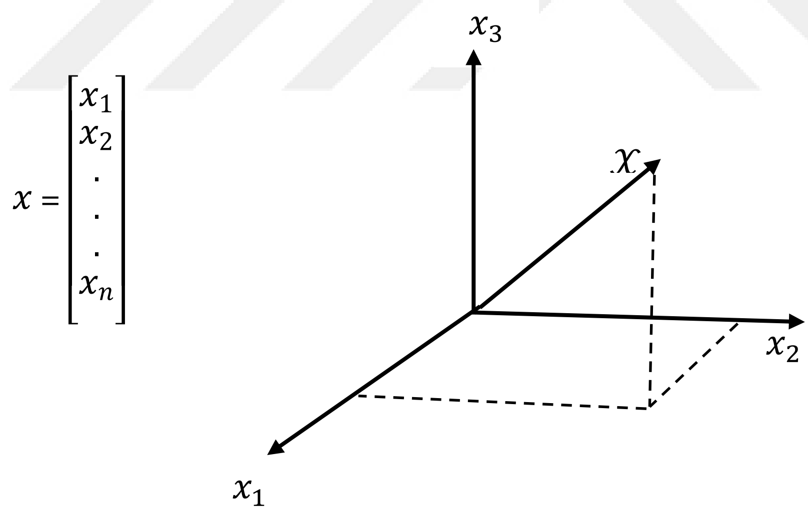


Figure 3.4 A vector created using biometric values

In Figure 3.4, $<x_1, x_2, x_3, \dots, x_d>$ are keystroke patterns. KSD and KPD are in this data format. Basically, they are array of measured keystroke times.

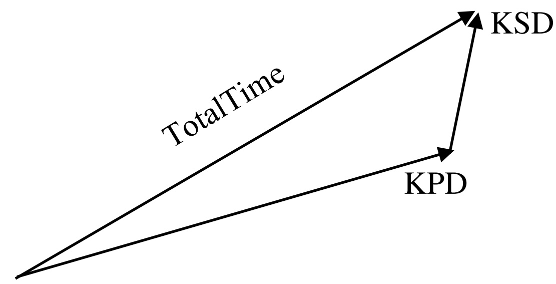


Figure 3.5 TotalTime calculation

Assume that, KPD and KSD are N dimensional vectors which are seen in Figure 3.5 and expressed as follows:

$$KPD = \sum u_i I_i \tag{3.4}$$

$$KSD = \sum x_i I_i \tag{3.5}$$

$$\text{TotalTime} = \sqrt{(KPD + KSD)^2} \tag{3.6}$$

TotalTime parameter is obtained using the mobile application which is created for this study. Application asks users to type some random texts and measures text typing time. TotalTime parameter is calculated as shown in formula 3.6

3.1.2 Error Count Measurement: Users under abnormal conditions are open to make mistakes. For this reason, measurement of errors as a parameter can be a key indicator to understand emotional condition of a user. Error count is gathered during Keystroke pattern analysis. Users are periodically asked to type a text by the application. Any mistakes in typing is count as 1 error. Classification process in the algorithms uses error count (EC) as an input parameter. Ψec indicates the error count. ω is length of the text to be written. in this study, EC is the number of errors in a window period. Error count (Ψec) indicator is defined as below in order to employ it as a probabilistic function. Error count can be scalar or normalized as in (Equation 3.7). The value of Ψec ranges from 0 to 1. When there is not any error, 1 is produced. The exact opposite 0 is generated if all samples are in error. Scalar value is used during the calculations in experimental study.

$$\Psi ec = \frac{(\omega - Ec)}{\omega} \tag{3.7}$$

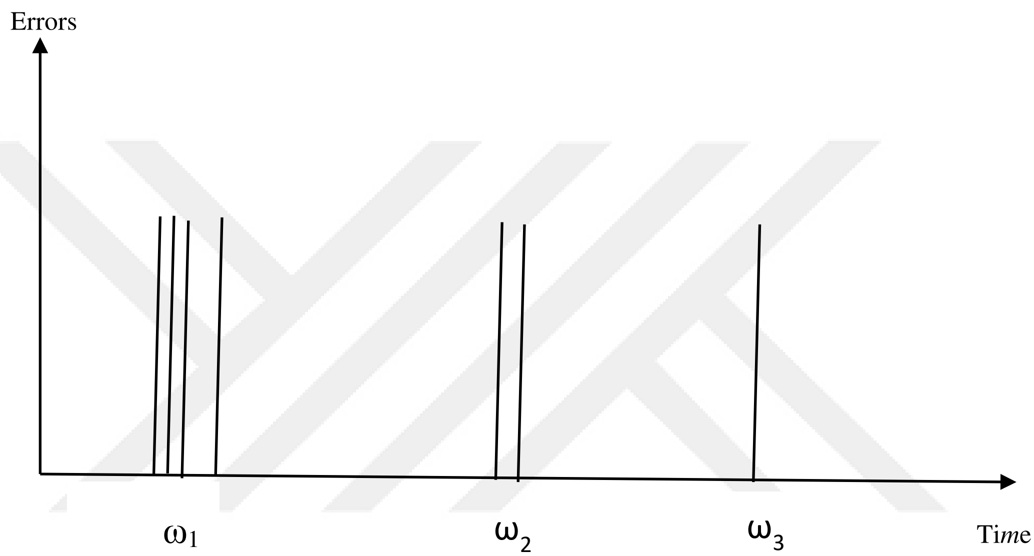


Figure 3.6 Error windows in time

In Table 3.2, corresponding Ψ_{ec} values are presented for each ω values for a distribution like in Figure 3.6.

Table 3.2 Corresponding Ψec calculation for three states.

ω order	**ω size**	**E_c**	**Ψ_{ec}**
1	10	4	0.6
2	10	2	0.8
3	10	1	0.9

Error Count parameter is obtained using the custom application which is created for this study. Application asks users to type some random texts and collects Errors if exists.

3.2. Physical Parameters

There are many physical parameters which can be obtained using standard intelligent devices. Sleep quality, energy, mobility/movement, and heart pulse rates are accepted as physical parameters in this work. Parameters are normalized using Ψ formulation. Ψ is named Generic Normalization Factor (GNF) in this thesis. Ψ generates a normalized value between 0-1 for all type of parameters.

3.2.1 Sleep Quality Measurement: Sleep Quality is a significant component of physical and mental health, as well as overall well-being. There are different indicators to measure the quality of sleep such as time in sleep. Although time is not the only parameter, it gives a general perspective about the quality of sleep. In Table 3.2, recommended sleep durations by age are given. Table 3.3 shows type of people by ages and their corresponding sleep durations recommended by National Sleep Foundation (NSF) in 2005 which is called Sleep Health Index (SHI). SHI is the first index developed by a professional organization for age-based expected sleep durations based on a rigorous, systematic review of the world scientific literature relating sleep duration to health, performance, and safety (Knutson et al.2017).

Table 3.3 Recommended sleep durations by age.

Type	**Age**	**Sleep(hour)**	**Average(hour)**
Newborns	0 to 3 months	14 to 17	15.5
Infants	4 to 11 months	12 to 15	13.5
Toddlers	1 to 2 years	11 to 14	12.5
Preschoolers	3 to 5 years	10 to 13	11.5
School age children	6 to 13 years	9 to 11	10
Teenagers	14 to 17 years	8 to 10	9
Younger adults	18 to 25 years	7 to 9	8
Adults	26 to 64 years	7 to 9	8
Older adults	>65 years	7 to 8	7.5

Today, smart devices have capability of measuring the sleep time. Sleep analysis is commonly used by people to understand and interpret their sleeping habits like "Time to go bed", "In bed duration", "Sleep duration", etc. In this study, sleep quality is calculated with Equation (Equation 3.8). Ψ_{sq} identifies GNF for sleep quality parameter which produces a value between 0 and 1. The lower values represent a bad indicator of sleep quality; higher values indicate a better sleep quality. Sc is an average expected sleep hour depending on SHI, and St is an actual sleep time.

$$\Psi_{sq} = 1 - (| S_c - S_t |/ S_c) \tag{3.8}$$

Sleep Quality parameter is obtained using smart watch. Sleep Quality parameter can be scalar or normalized value as in (Equation 3.8). Scalar value is used in experimental study. Smart watch measures and sends sleep time daily basis. Sleep time is an input parameter for machine learning/deep learning models.

3.2.2 Energy Measurement: Energy is evaluated in two main parts such as active energy and resting/passive energy. Passive energy is the minimum level of energy to keep alive a human being and used as base level of energy usually used by body itself. On the other hand, an active energy is exposed by burning when a human being exhibits an extra effort. In this study, total energy which is the sum of both is employed. Ψ_{eq} parameter is GNF parameter for the energy quality whose value between 0 and 1. Se is utilized as an average energy, and Ste is an actual energy.

$$\Psi_{eq} = 1 - (| S_e - S_{te} |/ S_e) \tag{3.9}$$

Energy parameter is obtained using smart watch. Energy parameter can be scalar or normalized value as in (Equation 3.9). Scalar value is used in experimental study. Smart watch measures and sends energy daily basis. Energy is an input parameter for machine learning/deep learning models.

3.2.3 Mobility/Movement Measurement: Mobility/Movement is a measure of steps taken for each day. Depending on life style of each individual, it exhibits an average value. Depending on illness or a rush day, this value can vary and a good indicator to distinct an ordinary and unordinary day. Ψ_{mq} is GNF parameter for mobility measurement whose value varies between 0 and 1. S_m is an average steps count, and S_{me} is an actual step count.

$$\Psi_{mq} = 1 - (|\, S_m - S_{me} \,| / S_m) \tag{3.10}$$

Mobility/Movement parameter is obtained using smart watch. Mobility/Movement parameter can be scalar or normalized value as in (Equation 3.10). Scalar value is used during the calculations in experimental study. Smart watch measures and sends step counts daily basis Mobility/Movement is an input parameter for machine learning/deep learning models.

3.2.4 Heart Pulse Measurement: Heart pulse is a measure for heart beats for an individual and has an average value for each user. Heart pulse is also differentiator to understand if someone is having an ordinary day or not. Ψ_{hq} is GNF parameter for heart pulse identifier, which relies between 0 and 1. S_h is an average heart beat value, and S_{he} is an actual heart beat value.

$$\Psi_{hq} = 1 - (|\, S_h - S_{he} \,| / S_h) \tag{3.11}$$

Heart Pulse parameter is obtained using smart watch. Heart Pulse parameter can be scalar or normalized value as in (Equation 3.11). Scalar value is used in experimental study. Smart watch measures and sends heart pulse counts daily basis. Heart pulse is an input parameter for machine learning/deep learning models.

Each of the parameters will be used in understanding of emotions of the users. Classification algorithms will use these parameters in train and test stages. Table 2.1 shows some of these parameters.

4. ALGORITHMS OF THE STUDY

In this thesis, we concentrate on the supervised learning and unsupervised/deep learning algorithms. Nature of problem is suitable for algorithms in both group. Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF) algorithms are inspected under Supervised Learning Algorithms. Artificial Neural Network (ANN) methods are inspected under Unsupervised Learning / Deep Learning section. Following Neural Network models are used;

- Feedforward Neural Network (FFNN)
- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- Long Short-Term Memory Network (LSTM)

The Study aimed to measure performance of algorithms under each group. Section 4.1 presents conventional algorithms, Section 4.2 presents deep learning algorithms. These algorithms are implemented in Experimental Results section (Section 5).

4.1 SUPERVISED LEARNING ALGORITHMS

Supervised learning is a machine learning method. A learning function that maps an input to an output depending on sample input-output pairs. It generates a function using labeled training data. In supervised learning, each sample is a pair. Pair contains an input object and a desired output object. A supervised learning algorithm analyzes the training data and produces a kernel function. This function can be used for generating outputs for new inputs those are unknown.

Figure 4.1 is a presentation to show how Supervised learning works. When a new data comes, classification algorithm checks and makes a decision made to classify new data. Data is placed to the correct group(cluster) using the logic used in Supervised learning algorithm.

Figure 4.1 shows very ideal situation in classification. Because data points easily separated by a linear line. Usually real life samples are not so ideal and needs more complicated approaches to separate data into groups.

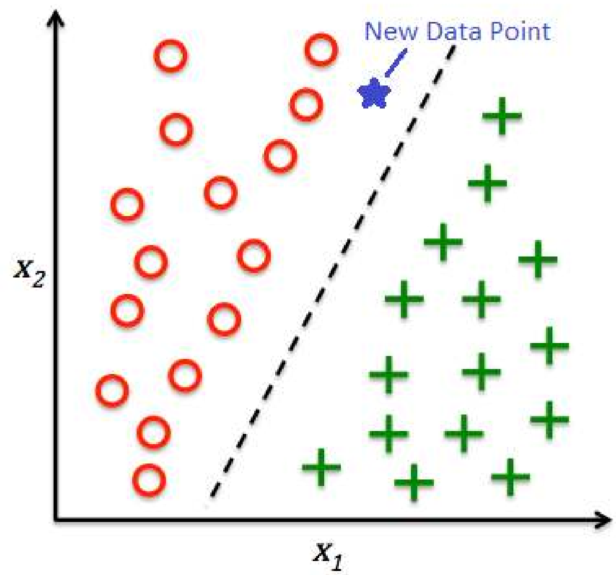


Figure 4.1 Supervised Learning

Following supervised learning algorithms are used in the study:

- Naive Bayes Algorithm (NB)
- Support Vector Regression Algorithm (SVR)
- Decision Tree Algorithm (DT)
- Random Forest Algorithm (RF)
- Ensemble Learning/Decision Integration Strategy (DIS)

Each of supervised machine learning algorithms are explained in detail as a part of this section.

4.1.1 Naive Bayes Algorithm (NB): History of Bayes algorithm is based on 1960s and was introduced into the text retrieval community. It has been a popular method for text classification and categorization. It had a good performance to classify many types of documents such as spam or legitimate, sports or politics, etc. Using appropriate pre-processing, it is competitive with more advanced methods like support vector machines. Naive Bayes algorithm also finds a large application in automatic medical diagnosis.

In computer science and statistics literature, there are many Bayes models exist. These models are named as Simple Bayes or Independence Bayes. All these names reference the same Bayes theorem with minor changes. Naive Bayes is also variant of Bayes theorem. This method assumes each feature's value as independent and will not consider any correlation or relationship between the features.

A naive Bayes classifier is a probabilistic machine learning model which is used for classification tasks. The classifier is based on the Bayes theorem as demonstrated in Equation (4.1).

$$P(A\backslash B) = \frac{P(B\backslash A)P(A)}{P(B)} \tag{4.1}$$

Using Bayes theorem, it is found the probability of *A* happening, given that *B* has occurred. Here, *B* is the evidence, and *A* is the hypothesis. The assumption made here is that the predictors/features are independent. That is presence of one particular feature does not affect the other. Hence, it is called naive. Bernoulli naive Bayes, Gaussian naive Bayes, multinomial naive Bayes, complement naive Bayes are the different types of naive Bayes approach. Bayes approach gives better results when parameters are independent or have weak dependency. In this work, multinomial naive Bayes algorithm is used.

4.1.2 Support Vector Regression Algorithm (SVR): Support Vector Regression (SVR) is a regression algorithm that has derived from Support Vector Machine. Theory of this approach is developed and presented by Vladimir Vapnik (Vapnik,1998). The main point of SVR is very similar to linear regression but SVR adds the margin concept. With this concept, any point in space must be reachable with the mainline and its sum with support vector. This sum operation is the margin concept. The way of finding support vector depends on the dimension of used space. In linear spaces used formula is:

$$y = \sum_{i=1}^{N} (\alpha_{\mathrm{i}} - a_i^*) . < x_i, x > +b \tag{4.2}$$

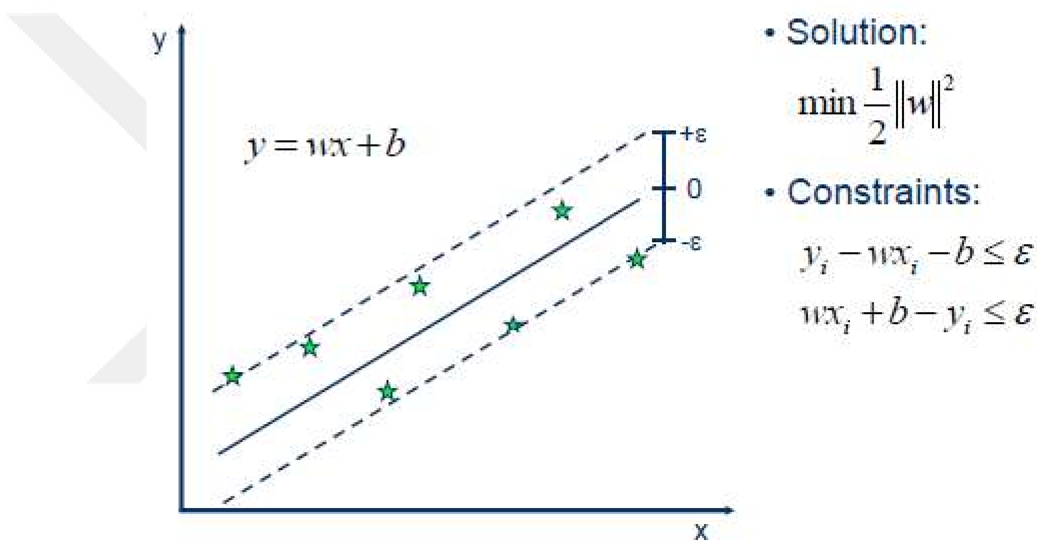


Figure 4.2 Linear Kernel Function for SVR

In our case, space is in 8th dimensions. Because there are 8 parameters exists. Equation (4.2) defines formula of Support Vector Machine. To increase performance of SVR approach many kernel functions can be used. These kernel functions may be non-linear functions and may perform better depending on problem.

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \cdot \langle \varphi(x_i), \varphi(x) \rangle + b$$

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b$$

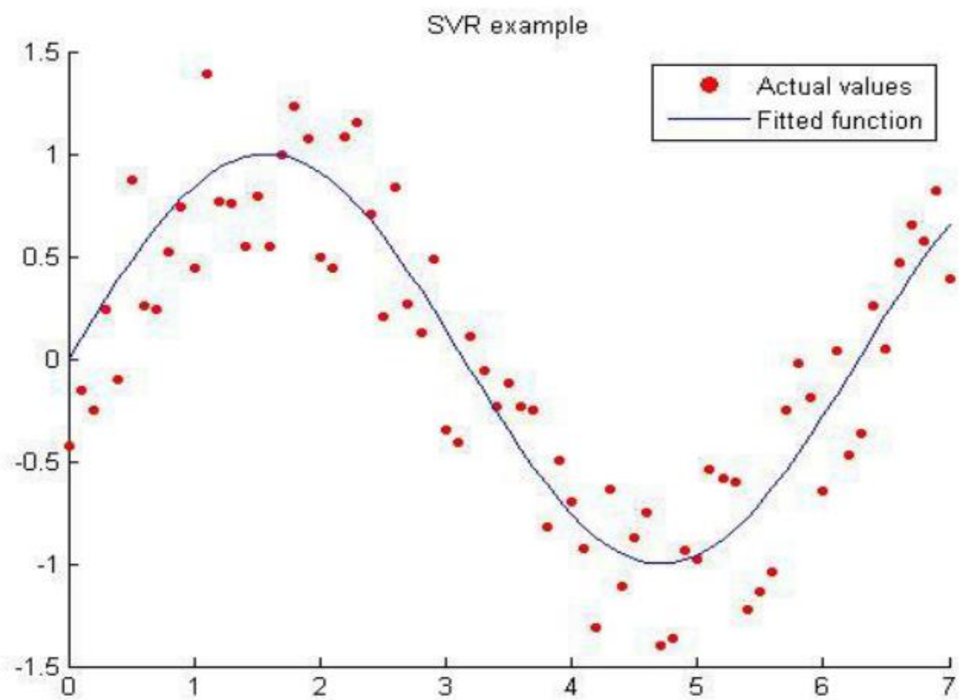


Figure 4.3 Non Linear Kernel Function for SVR

Choosing a kernel function is strongly dependent type of the problem and distribution of the data. As a result; the kernel functions are used to map the original dataset (linear/non-linear) into a higher dimensional space with view to making it linear dataset.

In this study, a non-linear kernel is used. Experimental studies show that using radial basis function (RBF) kernel gives the best results in SVR classification.

4.1.3 Decision Tree Algorithm (DT): Decision tree is one of the popular supervised learning algorithm for classification and prediction. Tree based learning algorithms are considered to be one of the best and mostly used supervised learning methods. A Decision tree can be thought like a flowchart in tree structure. Each internal node represents a conditional state on an attribute, each branch shows an outcome of the condition, and each leaf node holds a class label or a value. It is proposed by J. Ross Quinlan (Quinlan,1975).

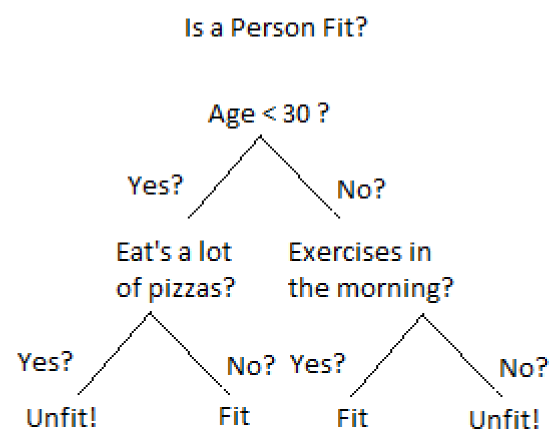


Figure 4.4 Decision Tree

Decision tree works similar to human brain. As a result; Decision tree algorithm breaks down the data into smaller subset until the last subset has only similar objects. During this process, separated subsets get attached to questions which have certain responses. These questions make it possible to find the right subset which the new data belongs to.

4.1.4 Random Forest Algorithm (RF): Random forests or random decision forests algorithm is an enhanced version of decision tree algorithm. It is proposed by Breiman (Breiman,2001). Having only one tree may not give the best result in some problems. Random forests algorithm creates smaller individual decision trees and each decision tree makes its own computation. These computed values are combined and the best answer is selected by a predefined rule. Random forests are also a good example of ensemble learning algorithms. Another main advantage of random forests is to correct decision trees' habit of over fitting to their training set (Korting,2013).

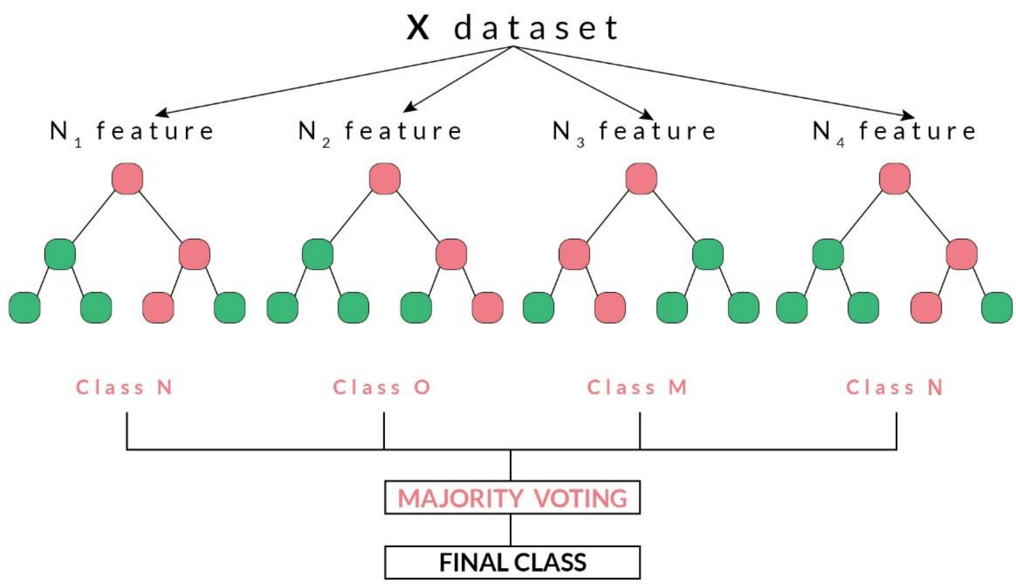


Figure 4.5 Random Forest Algorithm

Splitting a decision tree to a random forest is a random process. Each forest has own data when forest construction is over. Each forest is independent creates its own voting depending on data it holds. In this work, random forest algorithm is employed with ten estimators.

4.1.5 Ensemble Learning/Decision Integration Strategy (DIS): Decision integration is also machine learning method where same or different machine learning classifiers are used for decision making process. It is also named ensemble learning. Instead of using only one algorithm to make a decision, DIS uses multiple classifiers. Each classifier learns from training data and makes its own decisions. Majority voting is a known approach to combine decisions made by classifiers. Using Majority voting, algorithm creates only one final decision.

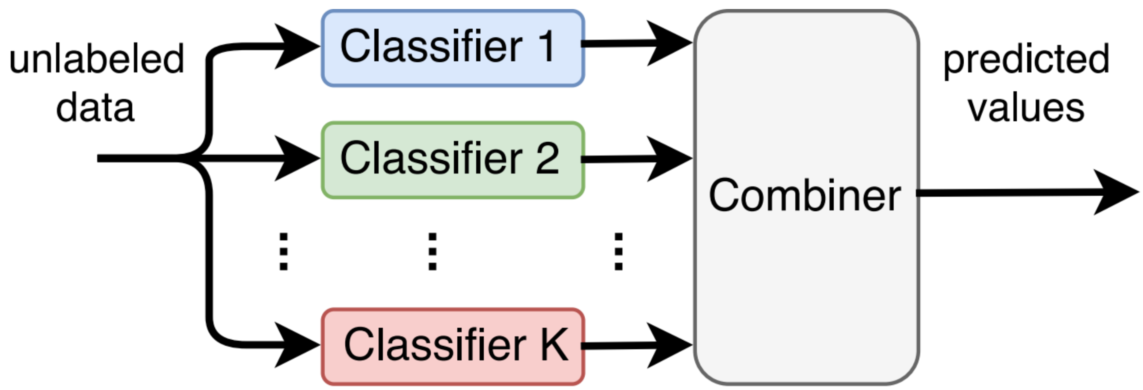


Figure 4.6 Structure of DIS/Ensemble Learning to make decision.

Ensemble Learning for conventional algorithms is used as DIS strategy as a part of this thesis. There are three main ensemble learning techniques.

- Bagging
- Boosting
- Stacking

Bagging use similar learners. Mean value of the results are taken and used as final decision. Boosting use iteration in decision making. It evaluates last decisions done and if a decision done by a classifier is correct it increases weight value for that classifier at the next iteration and vice versa. Stacking use different learners and combines their decision depending on rules defined.

Stacking technique is preferred in this study because studies show that combination majority voting is robust and efficient for classification (Liu & Wang,2010).

Machine learning algorithms those are explained above used as classifiers in Ensemble algorithm. A Voting Classifier is created and machine learning classifiers are defined within Voting Classifier. This method helped to create only one hypothesis by combining many results come from several machine algorithms.

4.2 UNSUPERVISED LEARNING / DEEP LEARNING ALGORITHMS

ANN models are used in Deep learning algorithms. ANN consists of cells which are similar to neurons in human brain. A neuron is a part of human nervous system. Brain is decision maker part of human body. Neurons are basic structure of brain. Most of activities those makes us human are executed by neurons. Neurons carries information throughout the brain. Neurons are also named nerve cells. %10 of human brain consists of neurons. The rest consists of glial cells and astrocytes that support and nourish neurons.

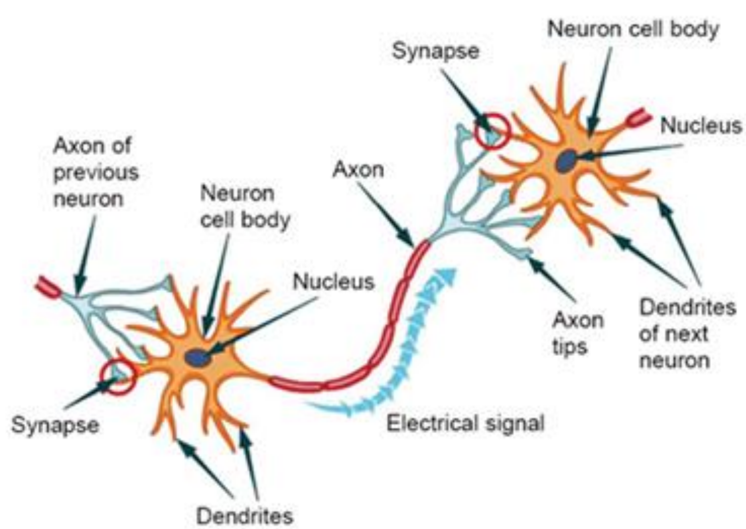


Figure 4.7 Structure of a Neuron

Neurons are connected each others via axons. Each neuron has mainly inputs, cell and output. Millions of them works together to make human body functional. ANN simulates this biological phonemea. Figure 4.8 shows structure of an artifial neuron. It has inputs, weights, cell and activation function. Weights are multiplied by input values and calculation is done in the cell. Depending on threshold value of the activation function calculated value is transfered to the next neuron.

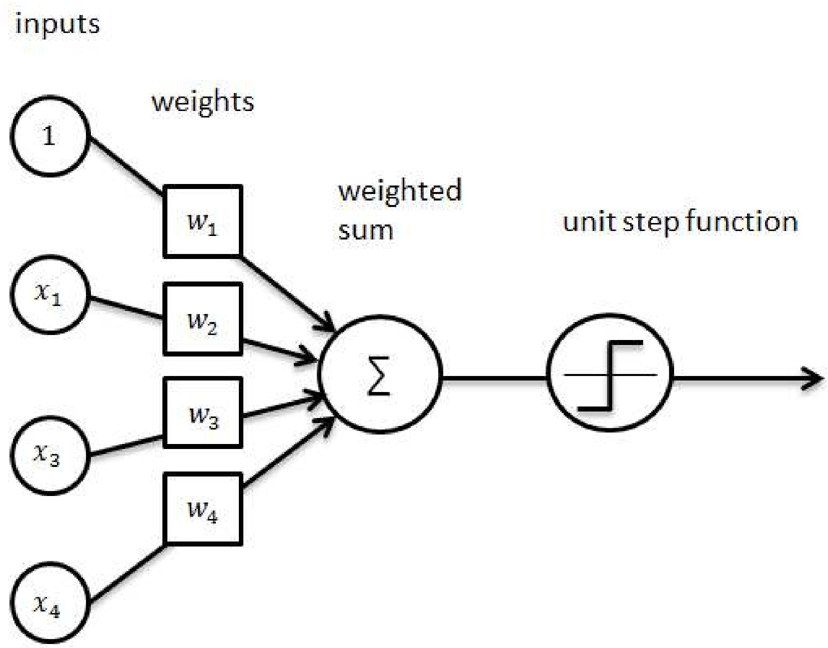


Figure 4.8 Structure of an Artificial neuron

Artificial Neural Networks can be trained. It is also called as learning process. When a train data exists ANN can arrange network parameters adaptive and change **w** values, cell function or parameters of activation function. Training is an iterative process needs some time to complete training process. Loss function makes a backpropagation and starts reinitialization of ANN parameters. This is an iterative process and proceeds until to obtain best hyper parameters those minimize the error. Figure 4.9 represents the process.

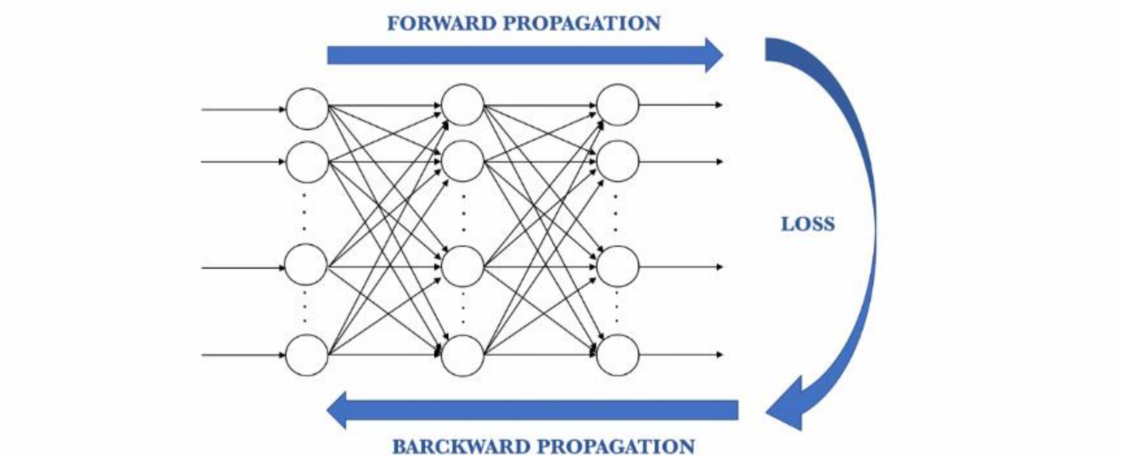


Figure 4.9 Learning Process via backward propagation

Base of Deep learning is ANN. Rina Dechter proposed the term deep learning in 1986 to machine learning community (Dechter,1986). Igor Aizenberg and his colleagues extend the concept and proposed a new term called “Artificial Neural Networks” in 2000 (Aizenberg,2000). Deep learning and Artificial Neural Networks offers much more than conventional machine learning algorithms. These techniques make machine learning closer to concept of Artificial Intelligence. Deep learning techniques like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Feed Forward Neural Networks (FFNN) etc. have been implemented many fields in computer science including natural language processing, audio/speech recognition or social network filtering. Today, Artificial Neural Network (ANN) techniques are commonly used to solve complicated problems. ANN algorithms offer an iterative process so that parameters of a model can be changed depending on results observed. In deep learning, each layer extracts its own data and pass this information to the next layer to be processed. Finally, a data flow from raw data to abstract information can be achieved. For example, a picture consists of pixels and first layer of an ANN model extracts pixels from a given picture as matrixes. Second layer handles corners using pixel data came from the first layer. Third layer handles objects using corners data came from second layer. Increasing number of layers improves power of the ANN model.

Unlabelled data is more than labelled data. Deep learning algorithms have capability of feature extraction from unlabelled data. This is a major benefit of deep learning algorithms over conventional machine learning algorithms. Deep learning algorithms are pretty suitable for unsupervised learning tasks.

Following neural network model are used as a part of the study;

- Feed Forward Neural Network (FFNN)
- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- Long Short-Term Memory Network (LSTM)

4.2.1 Feedforward Neural Network (FFNN): In (McCulloch & Pitts,1943), researchers proposed a study which offered a computational model for neural networks with a threshold logic. It was the first network model that can be named Feed Forward Neural Network. Feed Forward Neural Networks are also known as multi-layered network of neurons (MLN).

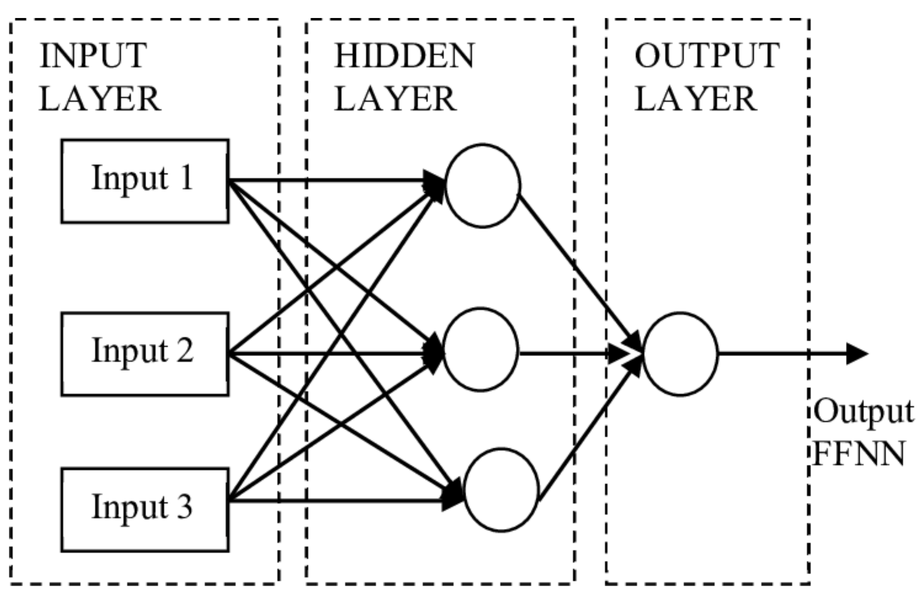


Figure 4.10 FFNN Network Model

As shown in Figure 4.10; FNN network model are named Feedforward because it transmits the information forward direction only. It may have several hidden layers. Each layer gets the information from previous layer, process it and pass it to the next layer. The hidden layers are capable to handle the complex non-linearly separable relations between inputs and the outputs. There is no feedback mechanism exists. It is the simplest model of ANNs.

4.2.2 Convolutional Neural Network (CNN): CNN is one of the most popular deep learning network algorithms. CNN is also a type of Feed Forward Neural Network type. It is commonly used in image processing area. Hidden layers consists of convolutional layers and pooling layers.

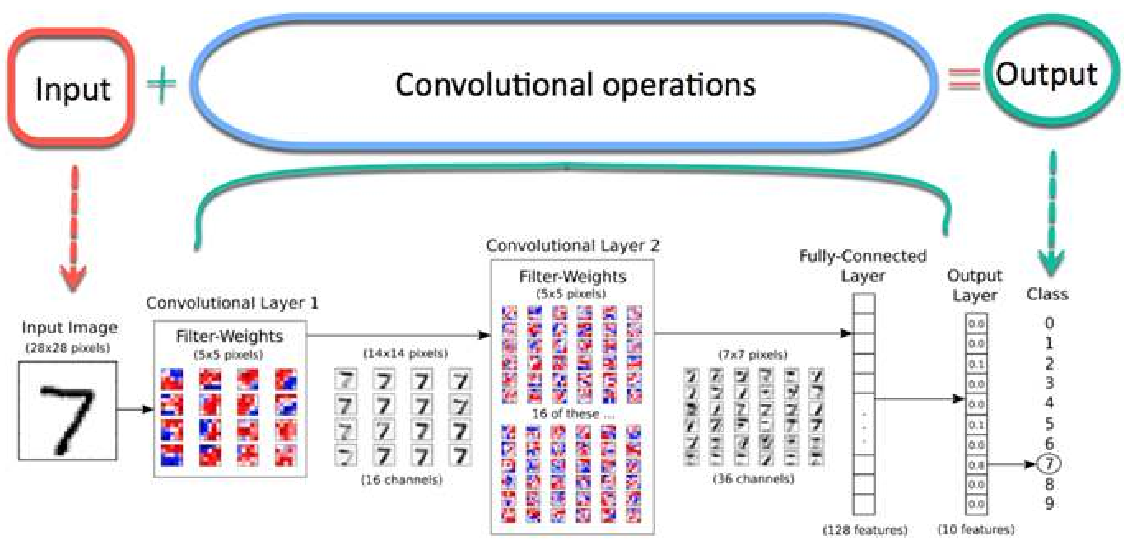


Figure 4.11 A Presentation of CNN model

Main part of a CNN is convolutional layer. Convolutional layer contains a convolution filter. Filter matrix traverse all the input values in input matrix. At the end, this operation generates a new generated matrix with the size of filter matrix. Filter matrix can be thought as an object definition. These object definitions can be thought as reference objects like car, dog or cat. If we would like to find a reference object on a pixel representation of picture, we apply corresponding filter. If searching object exists in the source filter will find that object in the picture. Same logic can be used in many areas not only in image processing but text searching, speech recognition etc. Figure 4.11 gives a high level presentation of CNN algorithm.

4.2.3 Recurrent Neural Network (RNN): Recurrent Neural Networks (RNN) is powerful algorithm to model sequential data. FFNN neural networks does not have a feedback mechanism but RNN have. An RNN hidden state is a function of all previous hidden states of recurrent nets (Elman,1990; Lipton et al.2015). RNN output does not only depend on current value but also past information. This property differs RNN from FFNN and even CNN type of artificial neural networks.

When iterations are independent and each iteration are calculated separately FFNN works well. But when iterations are related each other especially value of previous iteration is needed for calculation FFNN is not a good ANN model. RNN is capable to hold previous value. RNN is an enhanced model of FFNN. Figure 4.12 explains how RNN works.

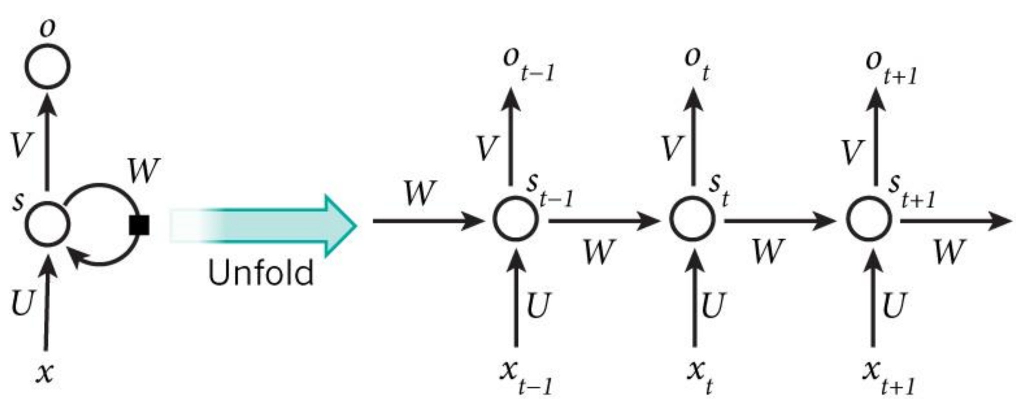


Figure 4.12 RNN Model

At the left hand side folded presentation of RNN is seen. Unfolded RNN model can be seen at the right hand side. RNN may have gradient exploding problems. If data has long term dependencies, RNN may have difficulties to learn past data. Values can vanish at the following iterations if values at previous layers gets smaller. This is called vanishing gradient problem. On the other hand, values can be larger at the following iterations if values at previous layers gets bigger. This is called exploding gradient problem. To deal with exploding problems, several methods like gradient clipping or appropriate activation functions can be used.

4.2.4 Long Short-Term Memory Network (LSTM): Recurrent Neural Networks (RNN) are limited to store only one iteration value. It might be necessary to store more values in artificial neural network model. Long Short-Term Memory Network (LSTM) solves this problem and capable to store large amount of data and use them as inputs in calculations. LSTM is a variant of RNN. LSTM is also more successful to solve vanishing/exploding problems of RNNs (Greff et al.2017; Kent & Salem,2019). RNN was developed for sequential data processing problems. LSTM holds long-distance dependencies than RNN. Each LSTM unit includes input, forget and output gates to control which parts of information to remember, forget and pass to the next step. Figure 4.13 shows a high level presentation of LSTM cell.

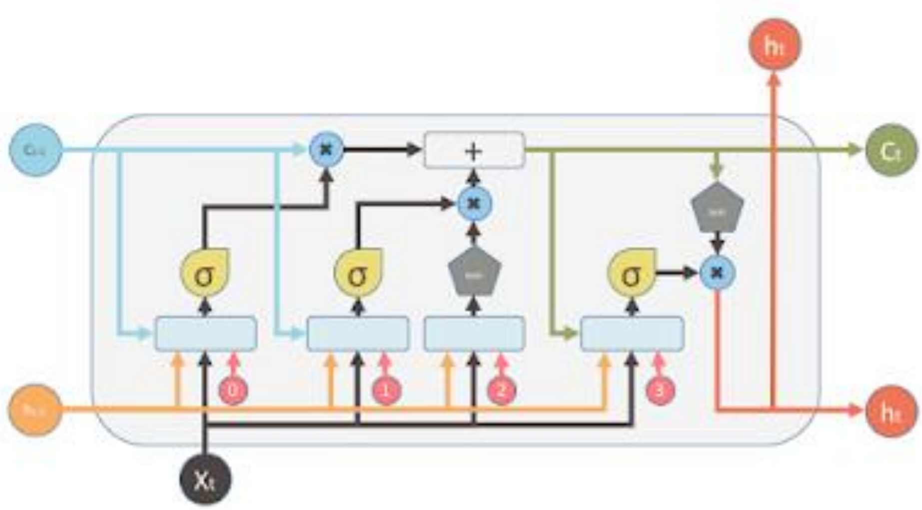


Figure 4.13 LSTM Cell Structure

5. EXPERIMENTAL RESULTS

In this thesis, the comprehensive experiments are carried out to analyze the neuro physical conditions of users using conventional machine learning algorithms and deep learning models. In order to demonstrate the contribution of our work, accuracy is employed as an evaluation metric in the experiments. We employ 67% of the dataset for the training process and the remaining part is the test data. The widely used 10-fold cross validation with 5x2 approach is applied on the dataset. Abbreviations mentioned in the tables and explanations are employed for the traditional machine learning algorithms and deep learning models as follows: MNB: Multinomial Naive Bayes, RF: Random Forest, SVR: Support vector Regression, DT: Decision Tree, DIS: Decision Integration Strategy, FFNN: Feed Forward Neural Network, CNN: Convolutional Neural Network, RNN: Recurrent Neural Network, LSTM: Long Short-Term Neural Network. The best accuracy results are acquired for each user is indicated with boldface.

As a first step, the classification performances of both conventional machine learning algorithms and deep learning methods are analyzed in terms of each user as given in Table 5.1 and 5.2, respectively. Experiments are carried out on two different approaches.

In the first approach, each user data is isolated and evaluated in the algorithms. Algorithms are trained and tested with each users' data. There are 365 rows of data and a row of data for each user is created for each day. Physical and neuro physical parameters are allocated in each row for the user. As a summary, the training and test set for each user is obtained with its own data and processed from each other independently in the first approach.

Table 5.1 Accuracy percentages of ML algorithms and DIS for each user.

User	MNB	RF	SVR	DT	DIS
1	32.23	60.43	56.65	**78.23**	81.44
2	26.45	60.56	36.34	**76.21**	67.21
3	27.78	62.66	29.56	**77.45**	70.36
4	31.67	65.78	43.64	**79.21**	76.60
5	26.95	56.87	37.68	**75.34**	70.84
6	20.42	57.44	43.32	**81.32**	71.56
7	32.56	58.63	66.67	**78.39**	75.44
8	29.78	58.56	34.78	**78.61**	71.83
9	24.75	60.76	40.98	**81.21**	66.05
10	25.56	60.66	39.54	**81.01**	75.13
11	27.87	59.78	38.56	**73.11**	73.09
12	23.36	56.51	43.43	73.13	**73.67**
13	27.67	57.32	43.65	**81.17**	72.24
14	36.43	60.69	39.27	**78.23**	70.88
15	30.23	63.71	31.61	**75.44**	66.42
Average	28.25	60.02	41.71	**77.87**	72.18

In Table 5.1, the accuracy percentages of conventional machine learning algorithms and decision integration strategy in terms of each user is presented. It is clearly observed that the mean classification accuracy of DT is the best performing machine learning algorithm with 77.87% accuracy result among others. Moreover, DT exhibits superior classification performance compared to the DIS which is competitive with the 72.18% accuracy success, by boosting the success of the proposed system. It is followed by RF with 60.02%, SVR with 41.71%, and MNB with 28.25%. MNB puts forth the poor classification performance among others with 28.25% of accuracy value. Fundamentally, naive Bayes algorithm makes an assumption that the features are conditionally independent. When parameters are somehow dependent each other Bayes algorithm may give poor results such as this classification performance of MNB. In this work, parameters collected, especially physical ones, are dependent each other. We consider that the cause behind of poor performance of MNB is dependency of features. As a result of Table 5.1, the success order of the classifiers is summarized as follows:

DT > DIS > RF > SVR > MNB

Another novelty of our thesis is to use deep learning algorithms to make a comparison with conventional algorithms. Same dataset is used to obtain performance of best known deep learning algorithms. As a known fact deep learning algorithms have layers and much more complicated algorithms prior to conventional machine learning algorithms. In deep learning part, Adam optimization with a learning rate of 0.0001 is used. Hyperbolic tangent as the nonlinear activation function is employed in the experiments. To alleviate the problem of over-fitting, dropout is applied and set to 0.5. Number of layers is adjusted to 7, number of epochs is set to 100. For deep learning algorithms, CNN, RNN, and LSTM models have 7 layers. In CNN, the first one is an input layer. The second layer is convolution layer with 64 filters whose kernel size is 3. The third layer is max-pooling layer whose max-pool size 2. The fourth and fifth layers are the same as second convolution layer, and third layer, respectively with the same characteristics. The sixth layer is full-connected layer with 16 neurons. The last layer is output softmax layer with 5 neurons. In RNN model, the first layer is an input layer. It is followed by the implementation of RNN with 2 Simple RNN layers each with 32 RNN cells followed by 2 time distribute dense layers for 5 class classification. Dropout rate on the fully connected layers is set to 0.5. Tangent activation function is used at the last layer. In LSTM model, the first layer is an input layer. It is followed by the implementation of LSTM with 2 layers of 32 LSTM cells followed by 2 time distribute dense layers for 5 class classification. Dropout rate on the fully connected layers is set to 0.5. Tangent activation function is used at the last layer.

Table 5.2 Accuracy percentages of DL algorithms for each user.

User	**FFNN**	**CNN**	**RNN**	**LSTM**
1	72.34	**83.56**	78.49	82.35
2	65.42	74.45	67.30	**75.58**
3	66.90	74.52	68.24	**76.21**
4	69.38	**81.24**	77.05	77.34
5	65.17	**79.90**	75.42	74.80
6	65.88	**78.17**	76.10	75.39
7	73.81	**76.35**	73.86	73.96
8	70.56	**75.28**	74.55	72.47
9	66.34	**76.34**	75.29	73.56
10	71.13	**82.50**	76.92	75.30
11	72.38	**83.09**	76.66	77.12
12	70.15	**81.24**	77.81	79.36
13	69.30	**80.40**	75.20	77.08
14	69.44	**80.51**	74.70	75.22
15	67.27	**78.36**	75.30	76.13
Average	69.03	**79.06**	74.86	76.12

In Table 5.2, the accuracy percentages of deep learning models in terms of each user is given above. We intend to demonstrate the classification capability of neural networks and compare with the traditional machine learning algorithms by implementing FFNN, CNN, RNN, and LSTM. Among the deep learning methods, CNN remarkably exhibits classification success with 79.06% of mean accuracy. Moreover, CNN outperforms other deep learning models while FFNN exhibits the poorest classification performance with 69.03% among the deep learning methods. Furthermore, LSTM maintains approximately 7% improvement considering the success of FFNN while RNN provides nearly 5% enhancement compared to the classification performance of FFNN. As a consequence of Table 5.2, the performance order of deep learning models is abstracted as follows: CNN > LSTM > RNN > FFNN. From a wider perspective, the classification performances of both deep learning models and conventional machine learning techniques are as follows:

CNN > DT > LSTM > RNN > DIS > FFNN > RF > SVR > MNB

Similarly, the best and worst classification performances are carried through CNN with 79.06% of accuracy and MNB with 28.25% of accuracy, respectively. Surprisingly, DT is rather competitive with 77.87% of mean accuracy when the classification successes of deep learning models are considered. It is followed by LSTM with 76.12%, RNN with 74.86%, DIS with 72.18%, FFNN with 69.03%, RF with 60.02%, SVR with 41.71%, and MNB with 28.25% of mean accuracy value.

In the second approach, it is assumed that the users may have similar physical/neuro physical conditions during there are in similar emotional conditions. All users' data are combined to test this assumption by constructing a large single dataset. After that, dataset is divided into two part as: 33% of the dataset is training data and the remaining part is the test. In Table 5.3, the accuracy percentages of all user data based both traditional machine learning algorithms and deep learning models are offered. With this approach, the classification success of both deep learning models and conventional machine learning techniques when considered all user data instead of user based data are the same as:

CNN > DT > LSTM > RNN > DIS > FFNN > RF > SVR > MNB

CNN outperforms others with 84.31% accuracy when the algorithms are evaluated among themselves. In addition to CNN; DT and RNN exhibit approximately 5% improvement compared to the first approach. There is a maximum roundly 16% rise with SVR algorithm in terms of classification success while minimum enhancement to the classification performance is observed as almost 2% with RF in the second approach. Even MNB whose classification success is the worst exhibits nearly 3% contribution. It is followed by FFNN and LSTM with nearly 4% enhancement; and by DIS with roundly 6% improvement in proportion to the first approach. As an another outcome of Table 5.3, it is clearly observed that DT is quite competitive with 82.03% accuracy when considered the classification success of deep learning algorithms. It is also concluded that the best three classification success is performed by deep learning models except DT and DIS. Surprisingly, DIS and RF fail to exhibit the expected classification performance and are underperformed compared to the deep learning algorithms. On the other hand, it is analyzed that the usage of conventional machine learning algorithms as MNB and SVR has a disadvantage because of the lowest classification accuracies. MNB is significantly worst performer algorithm. MNB works better when all the features are completely independent each other's. Some of the features in this study are mainly dependent each other's and this may cause MNB has low performance than others.

Table 5.3 Accuracy percentages of all algorithms.

Algorithm	Accuracy
MNB	31.23
RF	62.56
SVR	57.45
DT	82.03
DIS	79.27
FNN	73.42
CNN	**84.31**
RNN	79.55
LSTM	80.17

It is hard to compare the performance of our results with the state-of-the-art studies because of the lack of works with similar datasets, parameters, and deep learning approaches. Although a comparison of traditional machine learning algorithms and deep learning models is given in this work, we also report the classification accuracies of a number of research works here.

In Wiens' study (Wiens et al.2000), heartbeat parameter is measured as physical parameter to evaluate emotional conditions of users. They evaluate a simple test to evaluate users' heartbeat and emotional conditions. Approximately 50% accuracy result is obtained for the purpose of heartbeat detection and experience of emotions.

In Jaques' study (Jaques et al.2016), multi-task learning for predicting health, stress, and happiness is proposed. Students self-reported their health, stress, and happiness in the morning and evening every day on a scale from 0-100. Multi-task learning (MTL) is performed to the problem of predicting next-day health, stress, and happiness using data from wearable sensors and smartphone logs. Authors report that the proposed model classifies the next-day health, stress, and happiness situations with nearly %62 accuracy result. Of course It has to be considered that corresponding studies have very limited number of parameters. It is difficult to make a well working model when measured parameters are not enough. We can say that accuracy will increase when number of measured parameters increase.

6. CONCLUSION AND FUTURE STUDIES

In this thesis, we intend to enrich physical conditions with neuro physical conditions and evaluate them for the purpose of analyzing emotional conditions/sensitivity of the users. For this purpose, we concentrate on the conventional machine learning algorithms and deep learning methodologies in order to classify sensitivity of the users analyzing both physical and neuro physical parameters. Keystroke patterns, and error count measurements are studied as neuro physical parameters while sleep quality, energy, mobility/movement, and heart pulse are analyzed as physical parameters.

To the best of our knowledge, this is the very first attempt to analyze the neuro physical conditions of the users by evaluating sensitivity analysis and enriching them with the physical characteristics. To analyze the emotional conditions of the users, feedforward neural network (FFNN), convolutional neural network (CNN), recurrent neural network (RNN), and long short-term neural network are employed as deep learning methodologies while multinomial naive Bayes (MNB), support vector regression (SVR), decision tree (DT), random forest (RF), and decision integration strategy (DIS) are evaluated as conventional machine learning algorithms. In order to demonstrate the contribution of this study, the dataset is gathered with the usage of smart devices and sensors from the users. Experiment results demonstrate that the usage of deep learning methodologies and the combination of both physical and neuro physical parameters enhance the classification success of the system to interpret the sensitivity of the users. Especially, CNN exhibits outstanding classification success among others. We conclude the study that the physical and neuro physical parameters are strongly linked to users' emotional conditions. When emotions change users' body gives a reaction to the new situation and this change can be observed by measuring physical/neuro physical parameters. Experimental results also confirm this connection between physical and neuro physical parameters.

As a future work, emotions defined in Plutchik wheel of emotions will be mapped to measured parameters. When parameters are mapped several emotions successfully we will get more details about users' sensitivity. Plutchik's wheel of emotions is an accepted classifier in psychology to identify emotions. There will be new parameters needed to be measure in order to cover map of Plutchik's wheel of emotions. As a main parameter we are planning to enrich this study by employing sentiment analysis of the users which will be obtained from texts in social media platforms in addition to sensitivity analysis. Besides, we also intend to perform decision integration strategy for deep learning methodologies in order to compare decision integration strategy of conventional machine learning algorithms.

Computer Science and Psychology are different disciplines. Recent developments in technology makes closer both area. Many researches as academics and commercial have been started. Prediction of human beings may create a safer world than we have. Especially subject is evaluation of human beings we may need other disciplines rather than computer science. This study makes an attempt to create a link between computer science and psychology to analyze emotions of the users.

REFERENCES

Aggarwal, C.C., Zhai C.X. (2012). *Mining Text Data*, USA: Springer.

Aizenberg, I., (2000). Multi-valued and universal binary neurons. Massachusetts: *Kluwer Academic*.

Anagun, A.S., Cin, I. (1998). A neural network based computer access security system for multiple users. *Computers & Industrial Engineering*, 35, 1-2, 351-354.

Anttonen, J., Surakko, V., (2005). Emotions and heart rate while sitting on a chair, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 491-499.

Becker, N. B., Jesus S. N., Viseu J. N., Stobäus C. D., Guerreiro M. and Domingues R. B., (2018). Depression and quality of life in older adults: Mediation effect of sleep quality, *International Journal of Clinical and Health Psychology*. 18, 1.

Bergadano, F., Gunetti D. and Picardi C., (2002). User authentication through keystroke dynamics, *ACM Transactions on Information and System Security (TISSEC)*, 5, 367-397.

Bleha, S., (1988). *Recognition systems based on keystroke dynamics*, PhD, Dissertation,Univ. Missouri:Colombia.

Bleha, S., Slivinsky, C. and Hussien, B. (1990). Computer-access security systems using keystroke dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 12.

Breiman, L., (2001). Random forests. *Machine Learning*. 45(1), 5–32.

Darwin, C. (1872). *Expression of The Emotions in Man and Animals*, London:John Murray.

Dechter, R., (1986, August). Learning While Searching in Constraint-Satisfaction-Problems, *AAAI-86 Proceedings*, Philadelphia:Pennsylvania.

Elman, J.L., (1990). Finding structure in time, *Cognitive Science*, 14, 179-211.

Epp, C., Lippold, M. and Mandryk R.L. (2011, May). Identifying emotional atates using keystroke dynamics, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Vancouver:Canada.

Gaines, R., Lisowski, W., Press, S. and Shapiro, N., (1980). Authentication by keystroke timing: some preliminary results, Rand Rep. R-2560-*NSF*, Rand Corporation.

Garett, R., Liu, S., Young, S.D., (2016). The relationship between social media use and sleep quality among undergraduate students, *Information, Communication & Society.*

Giannakaki, K., Giannakakis, G., Farmaki, C. and Sakkalis, V., (2017, June). Emotional state recognition using advanced machine learning techniques on EEG data.*30th International Symposium on Computer-Based Medical Systems* (CBMS), IEEE, Thessaloniki.

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., and J. Schmidhuber, (2017). LSTM: *A Search Space Odyssey, Transactions On Neural Networks and Learning Systems,* IEEE, 28 , 10, 2222 – 2232.

Gutierrez, F.J., Ra, Lerma, Salgado Garza L. R., and Cantu, F.J. (2001). Biometrics and data mining: comparison of data mining-based keystroke dynamics methods for identity verification, *Lecture Notes in Computer Science*, 2313, 460-469.

Güven, A., Soğukpınar, I. (2003). Understanding user's keystroke patterns for computer access security, *Elsevier Computers & Security*, 22, 8, 695-706.

Guven, O., Akyokus S., Uysal M., and Guven A., (2007). Enhanced Password Authentication through Keystroke Typing Characteristics, *Artificial Intelligence and Applications*, 549.

Hackford, J., Mackey, A., Broadbent, (2019). The effects of walking posture on affective and physiological states during stress, *Journal of Behavior Therapy and Experimental Psychiatry*, 62,80-87.

Harper, R., Southern J. (2019). A Bayesian Deep Learning Framework for End-To-End Prediction of Emotion from Heartbeat, arXiv.org,arXiv:1902.03043.

Jaques, N., Taylor, S., Nosakhare, E., Sano, A., Picard R. (2016). Multi-task Learning for Predicting Health, Stress and Happiness, *NIPS Workshop*. Barcelona.

Joyce, R., Gupta G. Identity (1990). Authorization Based on Keystroke Latencies. *Communications of the ACM,* 168-176.

Kent, D., Salem, F. M. (2019, August). Performance of three slim variants of the long short-term memory {(LSTM)} layer, *62nd International Midwest Symposium on Circuits and Systems (MWSCAS),* IEEE, Dallas.

Khanna, P., Sasikumar M., (2010). Recognizing emotions from keyboard stroke pattern, *International Journal of Computer Applications*, 11,9.

Knutson, K. L., Phelan, J., Paskow, M. J. (2017). The national sleep foundation's sleep health index, *Journal of the National Sleep Foundation*. 3, 234–240.

Korting, T. S., C4.5 (2013). Algorithm and Multivariate Decision Tree, https://www.academia.edu/1983952/C4._5_algorithm_and_Multivariate_Decision_Trees.

Lipton, Z. C., Berkowitz J., and Elkan C. (2015). A critical review of recurrent neural networks for sequence learning.,1,18, https://arxiv.org/abs/1506.00019

Liu, B. (2010). *Handbook of Natural Language Processing*, London:Chapman&Hall/CRC.

Liu, N., Wang, H. (2010). Ensemble based extreme learning machine, *IEEE Signal Processing Letters*, 17, 8.

McCulloch, W.S, Pitts, Walter (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. 5, 115–133.

Monrose, F. & Rubin, A. (1997, April). Authentication via keystroke dynamics. *Computer and Communications Security*, ACM, Zurich.

Monrose, F., Reiter, M. K. and Wetzel S., (2002). Password Hardening Based on Keystroke Dynamics, *International Journal of Information Security*, 1,69-83.

Monrose, F., Rubin, A. (2000). Keystroke Dynamics.as a biometric for authentication. *Future Generation Computer Systems*, 16, 351-359.

Obaidat, M.S., Sadoun, B., (1997). A simulation evaluation study of neural network techniques to computer user identification, *Information Sciences*, 102, 239-258.

Quinlan, J.R., (1975). Induction of Decision Trees, *Machine Learning*, 1, 81-106.

Ramirez, N., Padilla L., Contreras D. G., Montelongo R. (2018, September). Physiological responses in heart rate with classical music, *15th International Conference on Electrical Engineering*, *Computing Science and Automatic Control* (CCE), Mexico City.

Rehman, A., Gumley, A. and Bielloa, S. (2018). Sleep quality and paranoia: The role of alexithymia, negative emotions and perceptual anomalies, *Elsevier Psychiatry Research*, 259, 216-222.

Sebastiani F., (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34, 1-47.

Song D., Venable P., Perrig A. (1997, April). User Recognition by *Keystroken Latency Pattern Analysis*,
https://users.ece.cmu.edu/~adrian/projects/keystroke/mid.pdf

Topol, Tom, (2012). *A Comprehensive Timeline of 150 Years History in Biometrics*, https://www.passport-collector.com/a-comprehensive-timeline-of-150-years-history-in-biometrics/

Vapnik,V. (1998). *Statistical Learning Theory*. New York, USA: Wiley.

Venture, G., Kadone, H., Zhang, T., Grèzes, J., Berthoz, A. and Hicheur, H., (2014). Recognizing emotions conveyed by human gait, *International Journal of Social Robotics*, 6, 4.

Vondrick, C., Pirsiavash, H. and Torralba, A. (2016, June). Anticipating visual representations from unlabeled video. *Conference on Computer Vision and Pattern Recognition (CVPR),* IEEE, Las Vegas.

Wiens, S., Mezzacappa, E. S. and Katkin, E.S. (2000). Heartbeat detection and exprience of emotions, *Cognition and Emotion,* 14,417-427.

Yamuza, M. T. V., Bolea, J., Orini, M., Laguna, P. Orrite C. (2019). Human emotion characterization by heart rate variability analysis guided by respiration, *IEEE Journal of Biomedical and Health Informatics*, 23,2446-2454.

Zhu, J., Ji, L. and Liu, C. (2019). Heart rate variability monitoring for emotion and disorders of emotion, *Physiological Measurement*, 40, 229-239.

CURRICULUM VITAE

NAME OF AUTHOR: Aykut Güven
PLACE OF BIRTH: Istanbul/ Turkey
DATE OF BIRTH: 07 December, 1974

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED
MSC – Computer Engineering, Gebze Institute of Technology (GYTE)
B. Science – Physics Engineering, Istanbul Technical University (ITU)

PROFESSIONAL EXPERIENCE
Founder Partner – Idea Technology 2004-
Technical Consultant – SPD Consulting 2002-2003
Specialist at Computer Eng. – Doğuş University, 2000-2002
Software Engineer – Alfa Yazılım 1998-2000
Software Engineer – Brain Computer1997-1998