5 Latest Research Papers On ML You Must Read in 2020

These papers provide a breadth of information about **Machine Learning** that is generally useful and interesting from a data science perspective.

Contents

- Smart Health Monitoring and Management Using Internet of Things, Artificial Intelligence with Cloud Based Processing
- Internet of Things with BIG DATA Analytics − A Survey

Cricket Analytics and Predictor

Mr. Suyash Mahajan
Student, Department of Information Technology,
Walchand Institute of Technology, Solapur, India

Ms. Gunjan Kandhari
Student, Department of Information Technology,
Walchand Institute of Technology, Solapur, India

Ms. Salma Shaikh
Student, Department of Information Technology,
Walchand Institute of Technology, Solapur, India

Ms. Rutuja Pawar
Student, Department of Information Technology,
Walchand Institute of Technology, Solapur, India

Mr. Jash Vora
Student, Department of Information Technology, Walchand Institute of Technology, Solapur, India

Ms. A. R. Deshpande
Assistant Professor, Department of Information Technology.
Walchand Institute of Technology, Solapur, India. anjali.deshpande10@gmail.com

Abstract—Cricket is one the most watched sport now-a-days. Winning in cricket depends on various factors like home ground advantage, performances in the past matches, experience of the players, performance at the specific venue, performance against the specific team and the current form of the team and the player. In the recent past, a lot of research has been done which measures the player's performance and predicts the winning percentage. This article briefs about the factors that cricket game depends on and discusses various researches which predicted the winning of a team with an advent of statistical modeling in sports. Cricket is one of the most popular team games in the world. With this article, we embark on predicting the outcome of Indian Premier League (IPL) cricket match using a supervised learning approach from a team composition perspective. Our work suggests that the relative team strength between the competing teams forms a distinctive feature for predicting the winner. Modeling the team strength boils down to modeling individual player's batting and bowling performances, forming the basis of our approach. We use statistics and recent performance of a player to model him. Player independent factors have also been considered in order to predict the outcome of a match. Machine learning is used in predicting the outcome of a cricket match before and during a match.

INTRODUCTION

Sports analytics play a major role in various problems associated with sport. Some of these problems are the ranking of individual players and their specialized skills, the composition of teams with an optimal balance of specialized skills, the ranking of teams, the negotiation of contracts, their potential revenue streams, the planning of both physical and mental training, the development of strategies for winning games and tournaments, assessing the effectiveness of coaches and referees, the medical aspects of sports injuries (health and insurance), the analysis and improvisation of rules, the quality of equipment and technology, the determination of awards, historical records and the generation of odds for gambling activities.

Related to above information, the coherent statistical presentation of both raw data and its inference to the decision makers is to facilitate successful planning and implementation. Furthermore, the media and the public have a great appetite for well visualized statistics. New opportunities for sports analytics have arisen due to the advent and availability of detailed and high quality data. For example, in Major League Baseball (MLB), the systems have provided comprehensive data on pitching and fielding. These systems record every play while also tracking the exact movements of all players on the field. Using these data sources, we can make very useful prediction, and various Statistics for improvement purposes.

Today's level of sports analytics has evolved where both the technology which provides data, and the statistical methodologies which provide the tools for analyzing data, improved very rapidly. Though sports analytics has been rapidly developing, it has not been the case with cricket. Due to historical reasons where cricket was perceived as a leisurely gentleman's game played without remuneration to players (until recently), cricket was not subject to large financial transactions.

This has changed in the last few years with the introduction of shorter formats of the game. The shortest and newest format, known as T20, generates intense interest and vast

sums of money, especially in the Indian subcontinent. The demand for cricket analytics has increased accordingly and the main website for cricket information and data is cricinfo.com. Cricket is a sport that originated in England in the 16th century and later spread to her colonies. The first international game however did not feature England but was played between Canada and the United States in 1844 at the grounds of the St George's Cricket Club in New York. In time, in both of these countries, cricket took a back seat to other, faster sports like ice-hockey, basketball, and baseball. International cricket is played today by a number of British Commonwealth countries; the main ones being Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri Lanka, West Indies and Zimbabwe. These teams are members of the International Cricket Council (ICC). A second rung of international teams called Associates includes numerous countries including Canada. Statistical modeling has been used in sports since decades and has contributed significantly to the success on field. Various natural factors affecting the game, enormous media coverage, and a huge betting market have given strong incentives to model the game from various perspectives. However, the complex rules governing the game, the ability of players and their performance on a given day, and various other natural parameters play an integral role in affecting the final outcome of a cricket match. This presents significant challenges in predicting the accurate results of a game. The game of cricket is played in three formats - Test Matches, ODIs and T20s. We focus our research on IPL. To predict the outcome of IPL cricket matches, we propose an approach where we first estimate the batting and bowling potentials of the 22 players using their career statistics and active participation in recent games.

I. RELATEDWORK

[1]In this paper, a methodology for identifying promising batting orders in one-day cricket was presented. In particular, they suggested some batting orders that have never been tried by the Indian team and contradict prevailing wisdom. As a byproduct of investigation, a simulation procedure was developed for generating first innings runs against an average opponent. The simulation procedure was based on estimates from a Bayesian log-linear model. Finally, methods were developed with the intention of finding optimal or nearly optimal batting orders at the start of a team's innings.

[2]In this paper, two methodologies have been used. MySQL database is used for storing data whereas Java for the GUI. The algorithm used is Clustering Algorithm for prediction. The steps followed are as-

1. Begin with a decision on the value of k being the number of clusters.
2. Put any initial partition that classifies the data into k clusters.
3. Take every sample in the sequence; compute its distance from centroid of each of the clusters. If sample is not in the cluster with the closest centroid currently, switch this sample to that cluster and update the centroid of the cluster accepting the new sample and the cluster losing the sample.
4. Repeat above step until convergence is achieved, that is until a pass through the training sample causes no new assignments.

[3]In this paper, they have featured the various types of possibilities for big data analysis in various fields (Programming Language, Statistical Solutions, and Visualization Tools), also, endeavor to recognize which one of them is more prominent to use than others, and they discovered that R is a normal programming language to use for data scientists. SPSS is great as statistical apparatus for non-analysts clients, and Tableau Public is best suitable visualization instrument to introduce data and break down it in graphical path, yet for web visualization reason D3 will be the best decision.

[4]This paper presents the Usage of the Duckworth-Lewis technique to decide assets staying, toward the end of each finished over; the anticipated run aggregate of the batting group could be refreshed to give a more precise expectation of the match result. Finally, it was discovered that the triumphant probabilities were allocated to the contending groups in ODI matches. With the utilization of D-L approach, this procedure can be promptly adjusted to deliver 'in the run' forecasts.

[5]This paper introduces a model that has three segments which focuses on diverse contemplations developing out of a more profound examination of T20 cricket. The models are made utilizing Data Analytics strategies from machine learning area. In this work 5 highlights of IPL vocation and 5 highlights of International T20 Career have been thought about for both batsmen and bowlers yet in future work more highlights can be made and considered.

[6]In this paper they have clarified the instruments of different techniques utilized for resetting target scores that are interfered with one-day cricket matches. Each of these strategies yields a reasonable focus in a few circumstances. None has demonstrated palatable in inferring a reasonable focus under all conditions. We have introduced a strategy which gives a reasonable reconsidered target score under all conditions.

A two-factor relationship has been determined which gives the normal numbers of runs which might be scored from any mix of these two assets and henceforth have inferred a table of extents of an innings for any such blend. This empowers the extent of the assets of the innings of which the batting groups are denied when overs are lost because of the stoppage in the play to be computed essentially and subsequently a reasonable revision to the objective score to be made. The parameters of relationship may change, for example, change in principles or conceivably changes in group choice and playing procedure.

[7]In this paper, they have discussed about the Duckworth-Lewis technique for target forecast in the session of cricket and clarified traps in the strategy. During analytics they utilized Correlation based subset assessment technique. As opposed to the conviction of the Duckworth/Lewis strategy, the setting of the diversion and strategic overs influences the expectation. In spite of the inadequate idea of the dataset,

relapse calculations and closest neighbor calculation gave the approximately correct results. Thus a half and half approach of utilizing quadratic relapse display with KNN as a smoothening capacity was utilized as an indicator and in addition the Duckworth/Lewis technique of having 1/1000 of the data was considered. At last the idea of expectation with energy of the amusement as a component was presented.

[8]This paper focuses on the execution of players as what number of runs will every batsman score and what number of wicket will every bowler take for both the groups. Both the issues are focused as grouping issues where number of runs and number of wickets are ordered in various reaches. The utilization of Naïve Bayes, random forecast, multiclass SVM and choice tree classifiers produce the expectation models for both the issues were made. Random Forest classifier was observed to be the most precise for both the issues. Four multiclass grouping calculations were utilized and thought about. Random Forest ended up being the most precise classifier for both the datasets with an exactness of 90.74% for foreseeing runs scored by a batsman and 92.25% for anticipating wickets taken by a bowler. Consequences of SVM were amazing as it accomplished an exactness of only 51.45% for foreseeing runs and 68.78% for anticipating wickets.

[9]The paper displays a data visualization and prediction device in which an open source, circulated, and non-social database, H-Base is used to keep the data identified with IPL (Indian Premier League) cricket matches and players. This data is then utilized for picturing the past execution of players' execution. Moreover, the data is utilized to anticipate the result of a match through different machine learning approaches. The proposed instrument can demonstrate the group administrations in the player barters for choosing the correct group. Finally, it was concluded that the novelty of the proposed approach lies in addressing the problem as a dynamic one and using a suitable non-relational database, H-Base for scalability of application. Out of all the machine learning algorithms used, KNN has been observed to be the most accurate.

[10]This Paper specifies the various factors that affect the game, winning in Cricket relies upon different variables like home group advantage, exhibitions before, involvement in the match, execution at the particular setting, execution against the particular group and the present type of the group and the player. Amid the previous couple of years part of work and research papers have been distributed which measure the player execution and their triumphant predictions. This article briefs about the variables that cricket diversion relies upon and focuses on couple of other research papers that anticipated the cricket wining.

[11]With statistical displaying in sports, foreseeing the result of an amusement has been built up as a central issue. Cricket is a standout amongst the most prevalent group recreations on the plane. It is observed that the relative group quality between the contending groups, frames an unmistakable component for foreseeing the victor. The utilization of profession insights and also the ongoing exhibitions of a player are shown. Player autonomous variables have additionally been considered with a specific end goal to anticipate the result of a match. It was demonstrated that the K Nearest Neighbor (KNN) calculation yields better outcomes when contrasted with different classifier. That is, the paper tends to the issue of anticipating the result of an ODI cricket match utilizing the measurements of 366 matches. The oddity of this approach lies in tending to the issue as a dynamic one and utilizing in taking part of players as key component in anticipating the prediction of the match.

[12]In this paper, they focus on anticipating the best appropriate Team to be lined for a specific match. We propose statistical displaying way to deal with the ideal players for the match to be played. This work recommends that the relative group quality between the contending groups frames an unmistakable component for foreseeing the victory. Demonstrating the group quality comes down to displaying singular player batting and rocking the bowling alley exhibitions, framing the premise of approach utilized. Vocation insights and also the ongoing exhibitions of a player have been utilized to demonstrate. Player free factors have additionally been considered keeping in mind the end goal to foresee the result of a match. Exploratory investigation was performed utilizing Hadoop and Hive for Indian players. Results show up to 91% exactness when contrasted with the genuine outcomes accessible over web. Finally, Making strategies of order of the batting innings or the bowling order can be sorted with these scores.

[13]In this Research paper it is intended to distinguish the variables which assume a key part in anticipating the result of an ODI cricket match and furthermore decide the exactness of the prediction made utilizing the method of data mining. In this examination, statistical hugeness for different factors which could clarify the result of an ODI cricket match is investigated. Home field advantage, winning the hurl, approach (batting first or handling first), match write (day or day and night), contending group, setting commonality and season in which the match is played will be key highlights considered for the examination. For motivations behind model building, three calculations are focused: Logistic Regression, Support Vector Machine and Naïve Bayes. Logistic regression is connected to data as of now acquired from beforehand played matches to distinguish which includes independently or in a mix with different highlights assumed to be a part in the prediction. SVM and Naïve Bayes Classifier are utilized for display preparing and prescient examination. Graphical portrayal and Perplexity frameworks are used to examine the models. An offering situation is likewise considered to clarify the choices that can be taken after the model has been constructed. Impact of this choice on the cost and result of the model is additionally examined.

[14]The paper tends to the issue of foreseeing the after effect of an ODI cricket match using the bits of knowledge of 5000 matches. The interest of this approach lies in tending to the issue as a dynamic one, and using the consequences of the past matches as the key component in foreseeing the prediction of the match. It was observed that basic features can yield especially reassuring outcomes. Predicting the winner of the matches utilizing distinctive

administered calculations has been accomplished and now we can anticipate the upcoming matches. There may be some more calculations coming in future which give better outcomes at that point utilized as a part of this paper.

II. METHODOLOGY

The work of our project focuses on two models.
The two models are:

1. Descriptive model
2. Predictive model

DESCRIPTIVE MODEL:

The descriptive model focuses mainly on two aspects:
It describes the data and statistics of the previous information i.e. batting, balling or all-rounder.
It gives the past information of the matches played by the IPL teams.

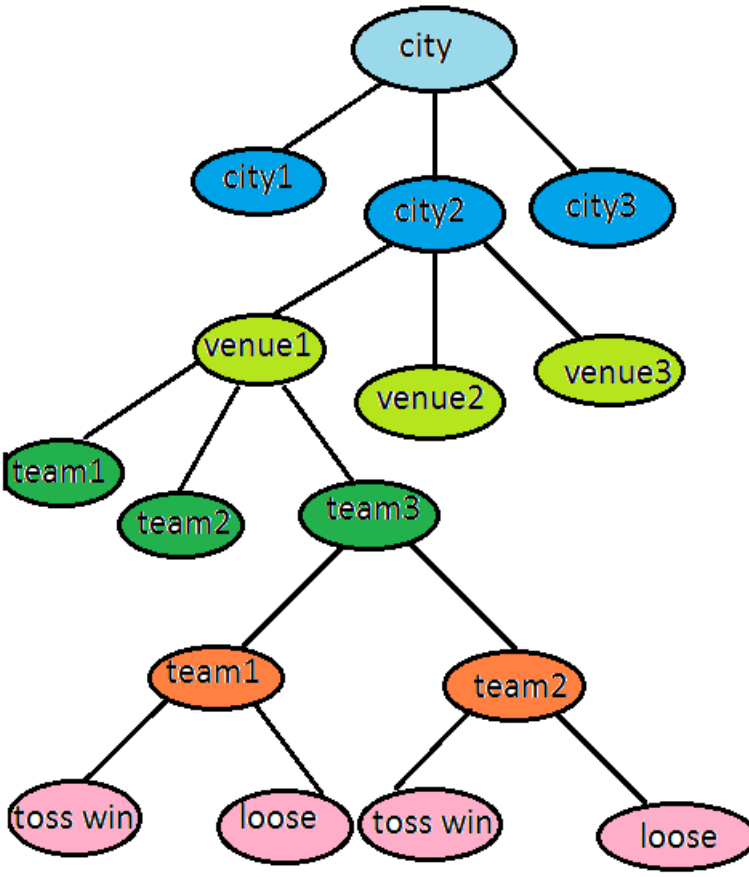
PREDICTIVE MODEL:

The predictive model focuses on predicting the winning percentage of the team. The ranking of the players is displayed as well.
The user has the liability to choose the two teams playing against each other. The selection of the teams works on the criteria as:

1. If the players are batsmen then, sorting is done according to the strike rate of the batsmen.
2. If the players are bowlers then, sorting is done according to the average rate of the bowler.
3. If the players are all-rounder then, sorting is done considering both strike rate as well as average rate.

The algorithm used for this model is Decision Tree Classifier. A decision tree is built using top-down approach. In this algorithm the root node i.e. the prior factor considered is the 'city' where the match is being played. The tree is built according to the prominent factors (city, venue, teams, toss decision) considered in the match.

Decision Tree Classifier Diagram:



Algorithm:

1. Start
2. Select the root node as 'city'.
3. Choose one of the cities from 'city1','city2','city3', etc.
4. Select the venues among one of the venue ('venue1','venue2','venue3') present in the city.
5. A team is selected and compared against the other teams.
6. Toss decision is made and the result is predicated upon the win/loose criteria.
7. End

III. RESULT

The user has the option of sign up/login. After the successful sign up/ login procedure, the user has liability to access two models that is Descriptive model that shows the statistics of the player and the Predictive model that predicts the winning percentage of the team that the user has selected.
The user can also read the latest tweets and news on the website.

IV. CONCLUSION

The website developed is an authorized website. This website is beneficial for the coach as he can rank the

players on their priority from the previous data, it is beneficial to the owner to get the details of the IPL match played and the users who predict the winning percentage of the team and get the statistics of the player.

V. REFERENCES

[1] T. B. Swartz, P. S. Gill, D. Beaudoin, and B. M. Desilva, "Optimal batting orders in one-day cricket," Computers & Operations Research, vol.33, no. 7, pp. 1939–1950, 2006.

[2] Preeti Satao, "Cricket Score Prediction System (CSPS) Using Clustering Algorithm." Technical Research Organization India, Vol. 3, no. Issue 4, 2016, pp. 2394–0697., troindia.in/journal/ijcesr/vol3iss4/43-46.pdf.

[3] Tamanna Siddiqui, Mohammad Alkadri, Najeeb Ahmad Khan, "Review of Programming Languages and Tools for Big Data Analytics", International Journal of Advanced Research in Computer Science, vol.8,no.5,May-June 2017.

[4] "Predicting The Match Outcome in One Day - jssm.org." [Online]. Available: http://www.jssm.org/volume05/iss4/cap/jssm-05-480.pdf&p=DevEx.LB.1,5063.1. [Accessed: 11-Aug-2018].

[5] C. Deep, C. Patvardhan, and C. Vasantha, "Data Analytics based Deep Mayo Predictor for IPL-9," International Journal of Computer Applications, vol. 152, no. 6, pp. 6–11, 2016.

[6] F. C. Duckworth and A. J. Lewis, "A Fair Method for Resetting the Target in Interrupted One-Day Cricket Matches," Operational Research Applied to Sports, pp. 128–143, 2015

[7] Vijay Ramakrishnan, Sethuraman K, and Parameswaran R, "Target Score Prediction in the game of Cricket." [Online]. Available: https://people.ucsc.edu/~praman1/static/pub/ML_Project_CS7641_report.pdf. [Accessed: 11-Aug-2018]

[8] K. Passi and N. Pandey, "Increased Prediction Accuracy in the Game of Cricket Using Machine Learning," International Journal of Data Mining & Knowledge Management Process, vol. 8, no. 2, pp. 19–36, 2018.

[9] S. Singh and P. Kaur, "IPL Visualization and Prediction Using H-Base," Procedia Computer Science, vol. 122, pp. 910–915, 2017.

[10] "Analysis on Attributes Deciding Cricket Winning," Scribd. [Online]. Available: https://www.scribd.com/document/357690109/Analysis-on-Attributes-Deciding-CricketWinning. [Accessed: 11-Aug-2018].

[11] Madan Gopal Jhawar and Vikram Pudi, "European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases."

[12]S. Agarwal, L. Yadav, and S. Mehta, "Cricket Team Prediction with Hadoop: Statistical Modeling Approach," Procedia Computer Science, vol. 122, pp. 525–532, 2017.

[13] Mehvish Khan and Riddhi Shah, "Role of External Factors on Outcome of a One Day International Cricket (ODI) Match and Predictive Analysis" International Journal of Advanced Research in Computer and Communication Engineering, vol. Vol. 4, no. Issue 6, pp. 192–197, Jun. 2015.

[14]Geddam Jaishankar Harshit and Rajkumar S, "A Review Paper on Cricket Predictions Using Various Machine Learning Algorithms and Comparisons Among Them," International Journal for Research in Applied Science & Engineering Technology (IJRASET) , vol. Vol 45, no. 98, pp. 27–32.

REAL TIME SLEEP/DROWSINESS DETECTION

Submitted in partial fulfillment of the requirements

of the degree of

Bachelor of Engineering

By

Roshan Shantaram Tavhare

(65)

Guide:

Dr. Varsha Shah

Computer Engineering Department

Rizvi College of Engineering

University of Mumbai

2018-2019

CERTIFICATE

This is to certify that the project entitled **"Real Time Sleep/Drowsiness Detection"** is a bonafide work of **"Roshan Shantaram Tavhare (65)"** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **"Bachelor of Engineering"** in **"Computer Engineering"**.

Dr. Varsha Shah
Guide

Prof. Shiburaj Pappu
Head of Department

Dr. Varsha Shah
Principal

Project Report Approval for B.E.

This Project report entitled **Real Time Sleep/Drowsiness Detection** by **Roshan Shantaram Tavhare (65)** is approved for the degree of Bachelor of Computer Engineering.

Examiners

1.----------

2.----------

Guide

1.----------

2.----------

Date: 22/04/2019

Place: Mumbai

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

Roshan Shantaram Tavhare

Date: 22/04/2019

ABSTRACT

The main idea behind this project is to develop a nonintrusive system which can detect fatigue of any human and can issue a timely warning. Drivers who do not take regular breaks when driving long distances run a high risk of becoming drowsy a state which they often fail to recognize early enough. According to the expert's studies show that around one quarter of all serious motorway accidents are attributable to sleepy drivers in need of a rest, meaning that drowsiness causes more road accidents than drink-driving. This system will monitor the driver eyes using a camera and by developing an algorithm we can detect symptoms of driver fatigue early enough to avoid the person from sleeping. So, this project will be helpful in detecting driver fatigue in advance and will give warning output in form of alarm and pop-ups.

Moreover, the warning will be deactivated manually rather than automatically. For this purpose, a de- activation dialog will be generated which will contain some simple mathematical operation which when answered correctly will dismiss the warning. Moreover, if driver feels drowsy there is possibility of incorrect response to the dialog. We can judge this by plotting a graph in time domain. If all the three input variables show a possibility of fatigue at one moment, then a Warning signal is given in form of text and sound. This will directly give an indication of drowsiness/fatigue which can be further used as record of driver performance.

Keywords- Drowsiness, Supervised Learning, Unsupervised Learning, Machine Learning.

Index

List of Figures

Chapter 1
Introduction

Real Time Drowsiness behaviors which are related to fatigue are in the form of eye closing, head nodding or the brain activity. Hence, we can either measure change in physiological signals, such as brain waves, heart rate and eye blinking to monitor drowsiness or consider physical changes such as sagging posture, leaning of driver's head and open/closed state of eyes.

The former technique, while more accurate, is not realistic since highly sensitive electrodes would have to be attached directly on the driver' body and hence which can be annoying and distracting to the driver. In addition, long time working would result in perspiration on the sensors, diminishing their ability to monitor accurately. The second technique is to measure physical changes (i.e. open/closed eyes to detect fatigue) is well suited for real world conditions since it is non-intrusive by using a video camera to detect changes. In addition, micro sleeps that are short period of sleeps lasting 2 to 3 minutes are good indicators of fatigue. Thus, by continuously monitoring the eyes of the driver one can detect the sleepy state of driver and a timely warning is issued

Chapter 2
Review of Literature

In this section, we have discussed various methodologies that have been proposed by researchers for drowsiness detection and blink detection during the recent years.

Manu B.N in 2016, has proposed a method that detect the face using Haar feature-based cascade classifiers. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier that will detect the object. So along with the Haar feature-based classifiers, cascaded Adaboost classifier is exploited to recognize the face region then the compensated image is segmented into numbers of rectangle areas, at any position and scale within the original image. Due to the difference of facial feature, Haar-like feature is efficient for real-time face detection. These can be calculated according to the difference of sum of pixel values within rectangle area and during the process the Adaboost algorithm will allow all the face samples and it will discard the non-face samples of images.

Amna Rahman in 2015, has proposed a method to detect the drowsiness by using Eye state detection with Eye blinking strategy. In this method first, the image is converted to gray scale and the corners are detected using Harris corner detection algorithm which will detect the corner at both side and at down curve of eye lid. After tracing the points then it will make a straight line between the upper two points and locates the mid-point by calculation of the line, and it connects the mid-point with the lower point. Now for each image it will perform the same procedure and it calculates the distance ‘d’ from the mid-point to the lower point to determine the eye state. Finally, the decision for the eye state is made based on distance ’d’ calculated. If the distance is zero or is close to zero, the eye state is classified as “closed” otherwise the eye state is identified as “open”. They have also invoked intervals or time to know that the person is feeling drowsy or not. This is done by the average blink duration of a person is 100-400 milliseconds (i.e. 0.1-0.4 of a second).

Chapter 3

Proposed Methodology

3.1) The different types of methodologies have been developed to find out drowsiness.

3.1.1) Physiological level approach: This technique is an intrusive method wherein electrodes are used to obtain pulse rate, heart rate and brain activity information. ECG is used to calculate the variations in heart rate and detect different conditions for drowsiness. The correlation between different signals such as ecg (electrocardiogram), EEG (electroencephalogram), and EMG (electromyogram) are made and then the output is generated whether the person is drowsy or not.

3.1.2) Behavioral based approach: In this technique eye blinking frequency, head pose, etc. of a person is monitored through a camera and the person is alerted if any of these drowsiness symptoms are detected.

3.2) The various technology that can be used are discussed as:

3.2.1) TensorFlow: IT is an open-source software library for dataflow programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production.

TensorFlow computations are expressed as stateful dataflow graphs. The name TensorFlow derives from the operations that such neural networks perform on multidimensional data arrays. These arrays are referred to as "tensors".

3.2.2) Machine learning: Machine learning is the kind of programming which gives computers the capability to automatically learn from data without being explicitly programmed. This means in other words that these programs change their behavior by learning from data. Python is clearly one of the best languages for machine learning. Python does contain special libraries for machine learning namely scipy, pandas and numpy which great for linear algebra and getting to know kernel methods of machine learning. The language is great to use when working with machine learning algorithms and has easy syntax relatively.

3.2.3) OpenCV: OpenCV stands for Open Source Computer Vision. It's an Open Source BSD licensed library that includes hundreds of advanced Computer Vision algorithms that are optimized to use hardware acceleration. OpenCV is commonly used for machine learning,

image processing, image manipulation, and much more. OpenCV has a modular structure. There are shared and static libraries and a CV Namespace.

In short, OpenCV is used in our application to easily load bitmap files that contain landscaping pictures and perform a blend operation between two pictures so that one picture can be seen in the background of another picture. This image manipulation is easily performed in a few lines of code using OpenCV versus other methods. OpenCV.org is a must if you want to explore and dive deeper into image processing and machine learning in general.

3.2.4) Kivy: Kivy is an open source Python library for developing mobile apps and other multitouch application software with a natural user interface (NUI). It can run on Android, iOS, Linux, OS X, and Windows. Distributed under the terms of the MIT license, Kivy is free and open source software. Kivy is the main framework developed by the Kivy organization, alongside Python for Android, Kivy iOS, and several other libraries meant to be used on all platforms.

3.3) System Description:

3.3.1) Login/ Sign up: The first step in the system is the simple and flexible credential interface where the user/driver must enter the login credentials, or he/she can Sign up in order to be a part of the system, the same is shown in the fig .1 as follows

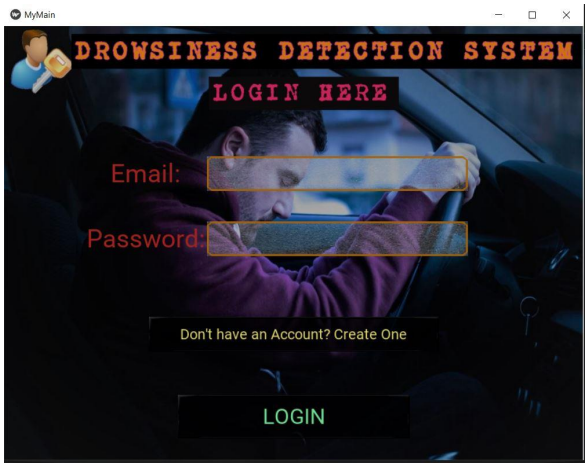


Fig.3.1

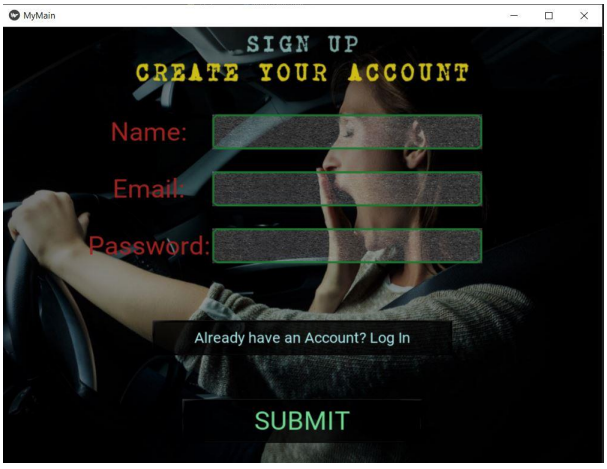


Fig.3.2

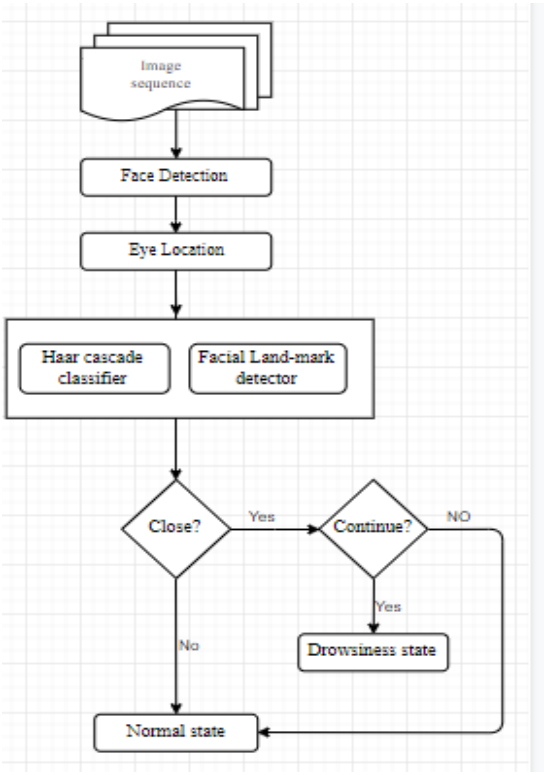


Fig.3.3**:** Flowchart of system

II. FLOWCHART AND ALGORITHM:

The various detection stages are discussed as:

3.3.2) Face Detection: For the face Detection it uses Haar feature-based cascade classifiers is an effective object detection method proposed by Paul Viola and Michael Jones in their paper, "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images.

Here we will work with face detection. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier. Then we need to extract features from it. For this, Haar features shown in the below image are used. They are just like our convolutional kernel. Each feature is a single value obtained by subtracting sum of pixels under the white rectangle from sum of pixels under the black rectangle. Fig. 3.4 represents five haar like features & example is shown in Fig.3.5

Fig. 3.4

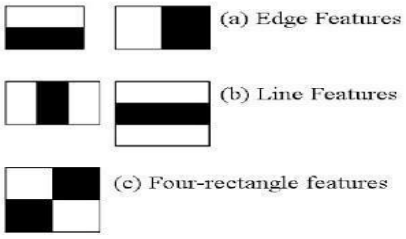


Fig. 3.5

A cascaded Adaboost classifier with the Haar-like features is exploited to find out the face region. First, the compensated image is segmented into numbers of rectangle areas, at any position and scale within the original image. Due to the difference of facial feature, Haar-like feature is efficient for real-time face detection. These can be calculated according to the difference of sum of pixel values within rectangle areas. The features can be represented by the different composition of the black region and white region. A cascaded Adaboost classifier is a strong classifier which is a combination of several weak classifiers. Each weak classifier is trained by Adaboost algorithm. If a candidate sample passes through the cascaded Adaboost classifier, the face region can be found. Almost all of face samples can pass through and nonface samples can be rejected

3.3.3) Eye detection: In the system we have used facial landmark prediction for eye detection Facial landmarks are used to localize and represent salient regions of the face, such as:

- Eyes
- Eyebrows
- Nose
- Mouth
- Jawline

Facial landmarks have been successfully applied to face alignment, head pose estimation, face swapping, blink detection and much more. In the context of facial landmarks, our goal is detecting important facial structures on the face using shape prediction methods. Detecting facial landmarks is therefore a twostep process:

- Localize the face in the image.
- Detect the key facial structures on the face ROI.

Localize the face in the image: The face image is localized by Haar feature-based cascade classifiers which was discussed in the first step of our algorithm i.e. face detection.
Detect the key facial structures on the face ROI: There are a variety of facial landmark detectors, but all methods essentially try to localize and label the following facial regions:

- Mouth
- Right eyebrow
- Left eyebrow
- Right eye
- Left eye
- Nose

The facial landmark detector included in the dlib library is an implementation of the One Millisecond Face Alignment with an Ensemble of Regression Trees paper by Kazemi and Sullivan (2014).

This method starts by using:

1. A training set of labeled facial landmarks on an image. These images are manually labeled, specifying specific (x, y)-coordinates of regions surrounding each facial structure.
2. Priors, of more specifically, the probability on distance between pairs of input pixels.

The pre-trained facial landmark detector inside the dlib library is used to estimate the location of 68 (x, y)-coordinates that map to facial structures on the face.

The indexes of the 68 coordinates can be visualized on the image below:

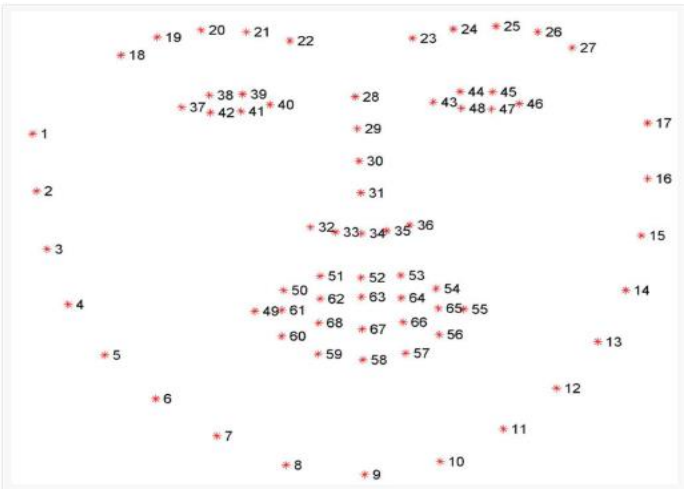


Fig.3.6: Visualizing the 68 facial landmark coordinates

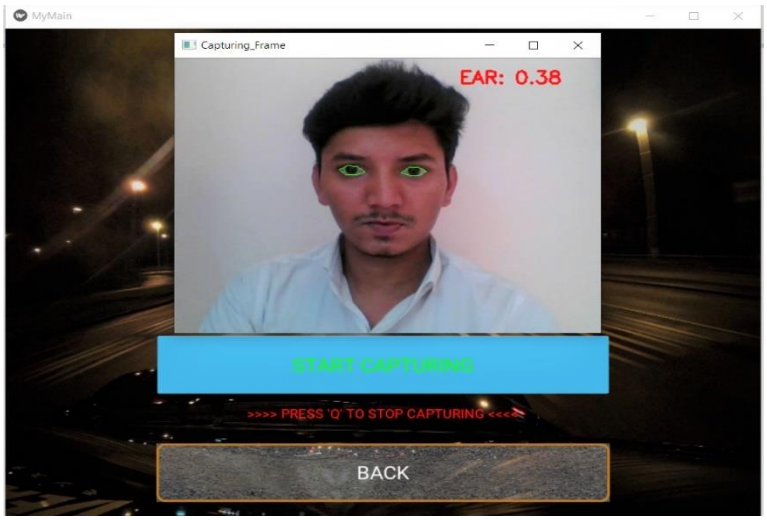


Fig.3.7: Detection of both the eyes

We can detect and access both the eye region by the following facial landmark index show below

- The right eye using [36, 42].
- The left eye with [42, 48].

These annotations are part of the 68 point iBUG 300-W dataset which the dlib facial landmark predictor was trained on. It's important to note that other flavors of facial landmark detectors exist, including the 194 point model that can be trained on the HELEN dataset.

Regardless of which dataset is used, the same dlib framework can be leveraged to train a shape predictor on the input training data.

3.3.4) Recognition of Eye's State:

The eye area can be estimated from optical flow, by sparse tracking or by frame-to-frame intensity differencing and adaptive thresholding. And Finally, a decision is made whether the eyes are or are not covered by eyelids. A different approach is to infer the state of the eye opening from a single image, as e.g. by correlation matching with open and closed eye templates, a heuristic horizontal or vertical image intensity projection over the eye region, a parametric model fitting to find the eyelids, or active shape models. A major drawback of the previous approaches is that they usually implicitly impose too strong requirements on the setup, in the sense of a relative face-camera pose (head orientation), image resolution, illumination, motion dynamics, etc. Especially the heuristic methods that use raw image intensity are likely to be very sensitive despite their real-time performance.

Therefore, we propose a simple but efficient algorithm to detect eye blinks by using a recent facial landmark detector. A single scalar quantity that reflects a level of the eye opening is derived from the landmarks. Finally, having a per-frame sequence of the eye-opening estimates, the eye blinks are found by an SVM classifier that is trained on examples of blinking and non-blinking patterns.

Eye Aspected Ratio Calculation:

For every video frame, the eye landmarks are detected. The eye aspect ratio (EAR) between height and width of the eye is computed.

$$EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2||p1 - p4||} \quad (1)$$

where p1, . . ., p6 are the 2D landmark locations, depicted in Fig. 1. The EAR is mostly constant when an eye is open and is getting close to zero while closing an eye. It is partially person and head pose insensitive. Aspect ratio of the open eye has a small variance among individuals, and it is fully invariant to a uniform scaling of the image and in-plane rotation of the face. Since eye blinking is performed by both eyes synchronously, the EAR of both eyes is averaged.

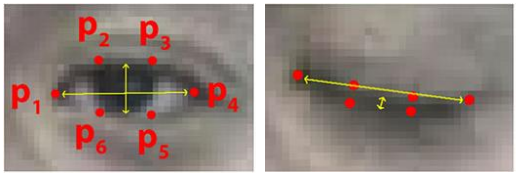


Fig.3.8

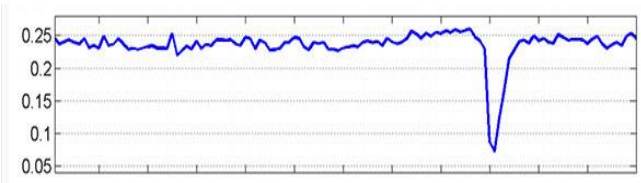


Fig.3.9 EAR for single blink

Fig 3.8: Open and closed eyes with landmarks p(i) automatically detected. The eye aspect ratio EAR in Eq. (1) plotted for several frames of a video sequence.

User flexibility Set up Features: This interface helps the user to adjust the EAR accordingly ranging from low to high percent. It also allows the user to set the timer sensitivity shown in the fig 3.10.

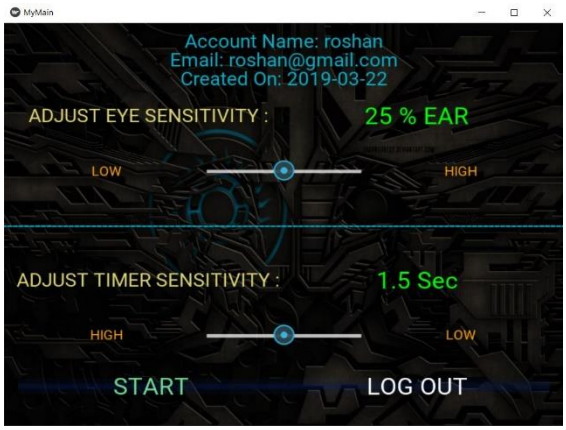


Fig. 3.10

3.3.5) Eye State Determination:

Finally, the decision for the eye state is made based on EAR calculated in the previous step. If the distance is zero or is close to zero, the eye state is classified as “closed” otherwise the eye state is identified as “open”.

3.3.6) Drowsiness Detection:

The last step of the algorithm is to determine the person’s condition based on a pre-set condition for drowsiness. The average blink duration of a person is 100-400 milliseconds (i.e. 0.1-0.4 of a second). Hence if a person is drowsy his eye closure must be beyond this interval. We set a time frame of 5 seconds. If the eyes remain closed for five or more seconds, drowsiness is detected and alert pop regarding this is triggered.

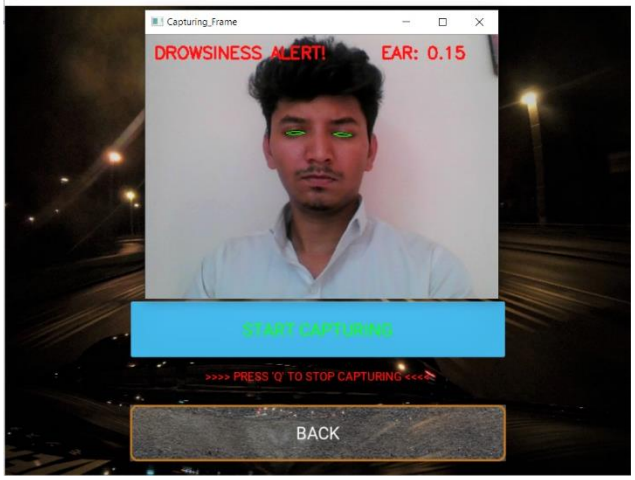
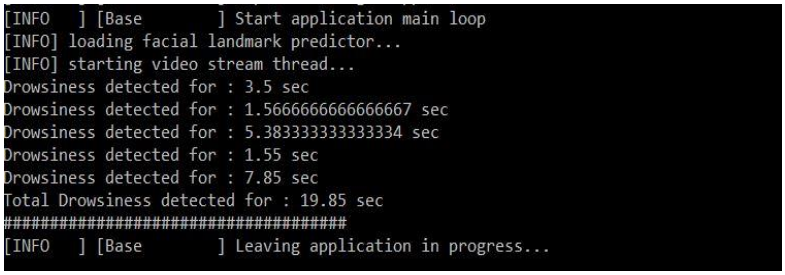


Fig.3.11 Drowsiness state

Fig .3.12 Console information



Chapter 4

Modelling diagram (project Work Flow)

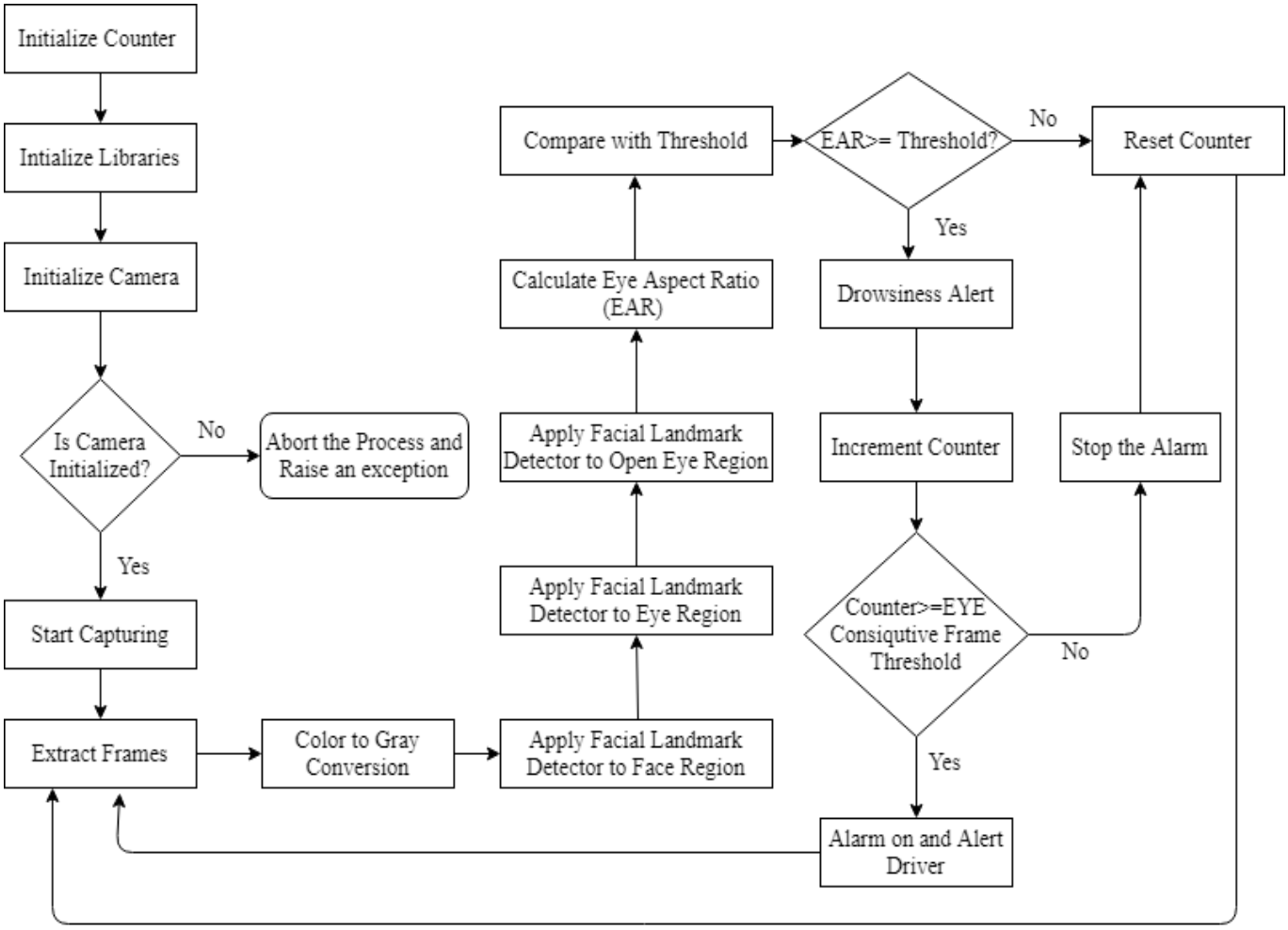


Fig. 4.1 Flow chart of System

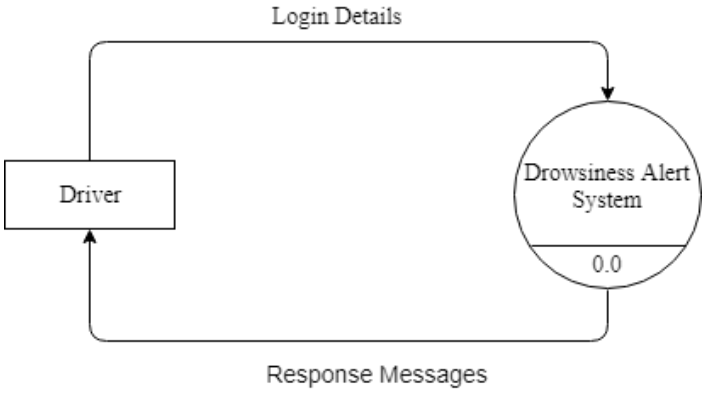


Fig.4.2 DFD level 0

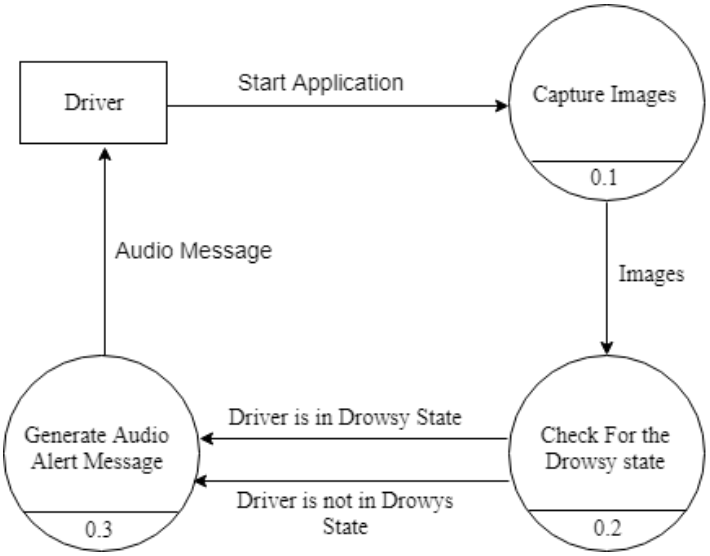


Fig 4.3 DFD level 1

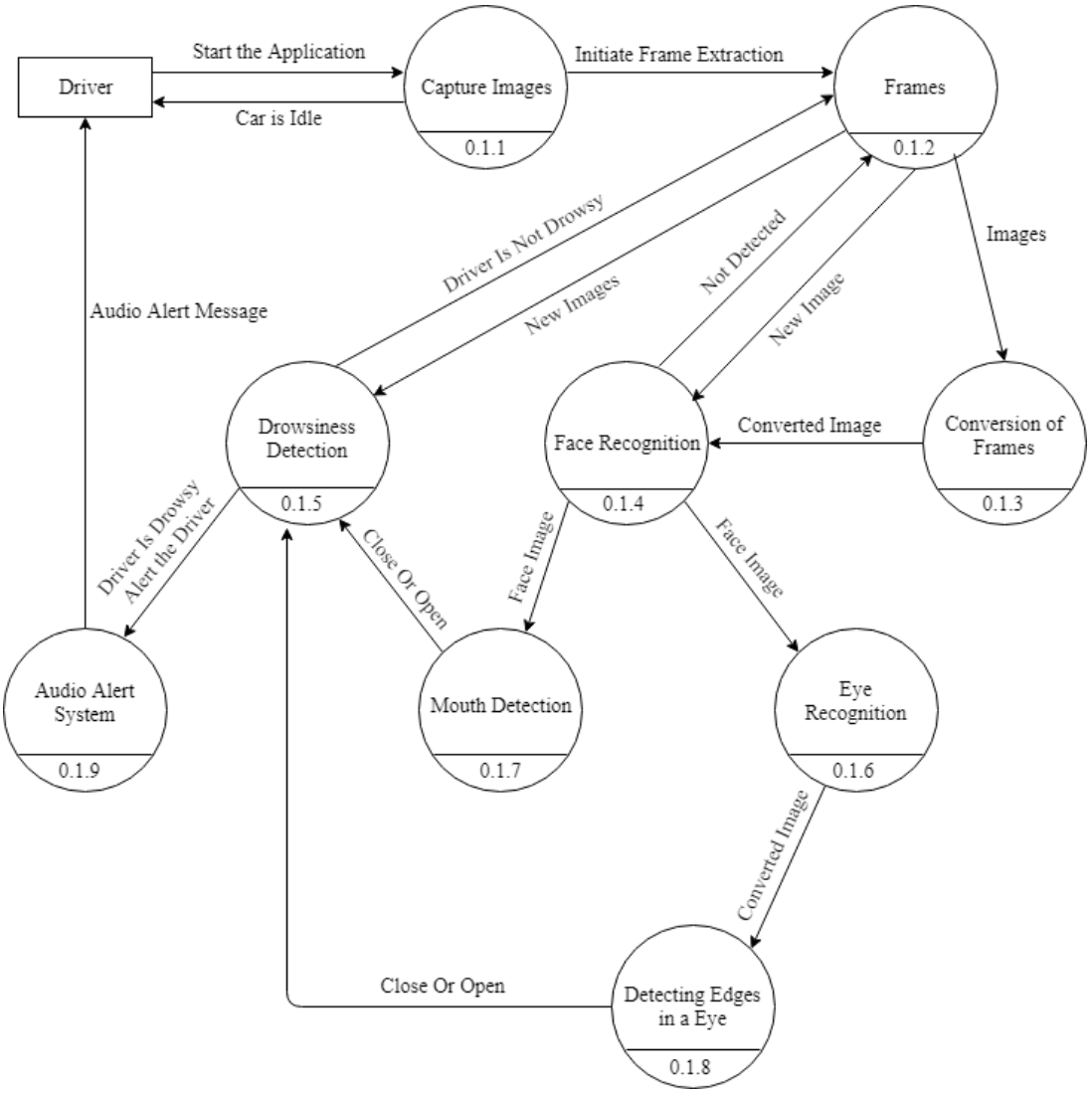


Fig 4.4 DFD level 2

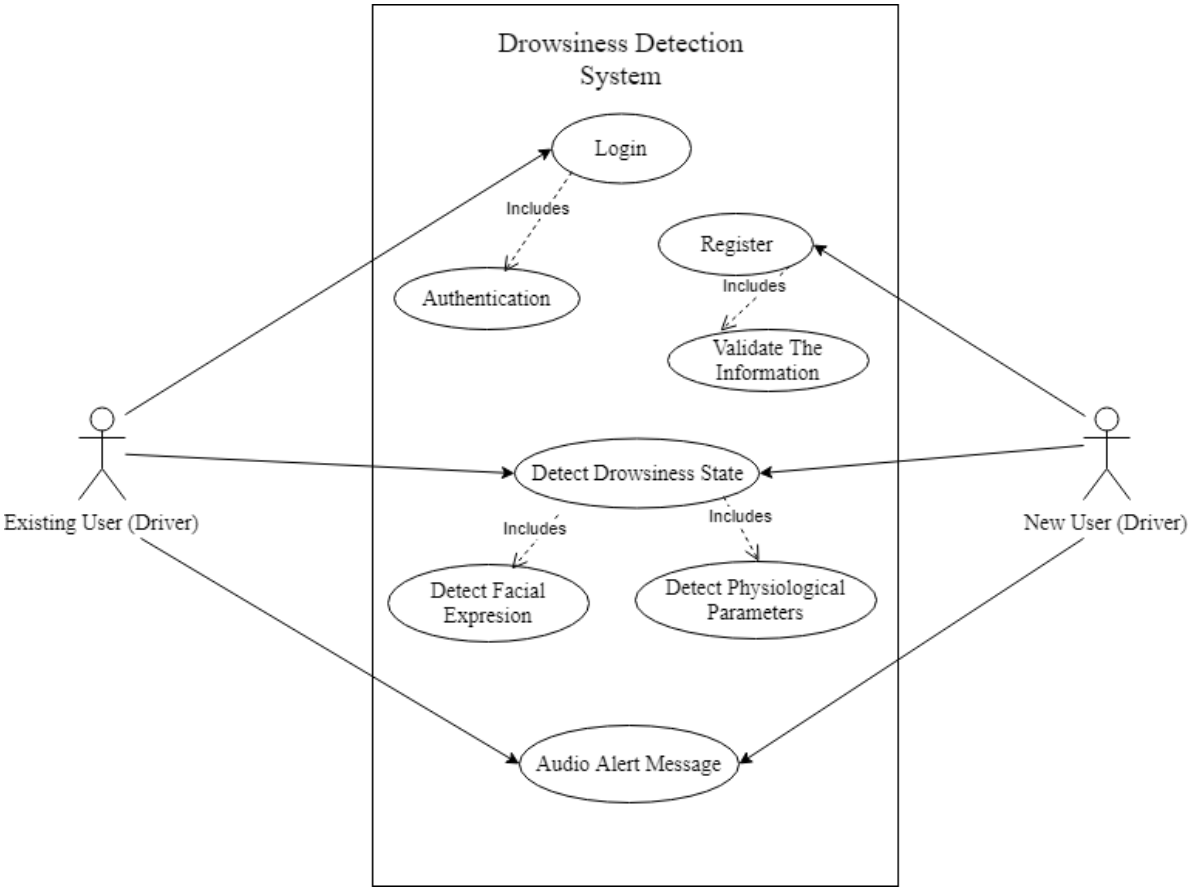


Fig 4.5 use-case Diagram

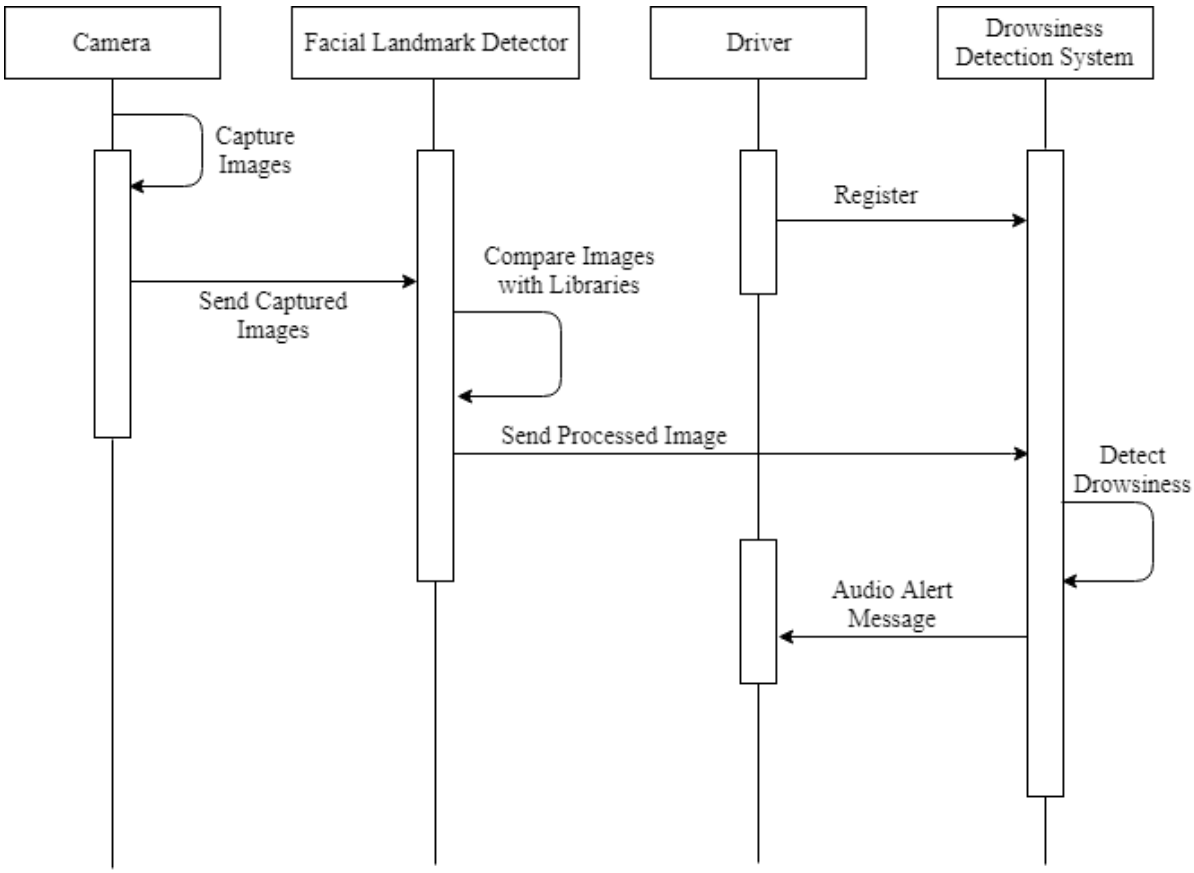


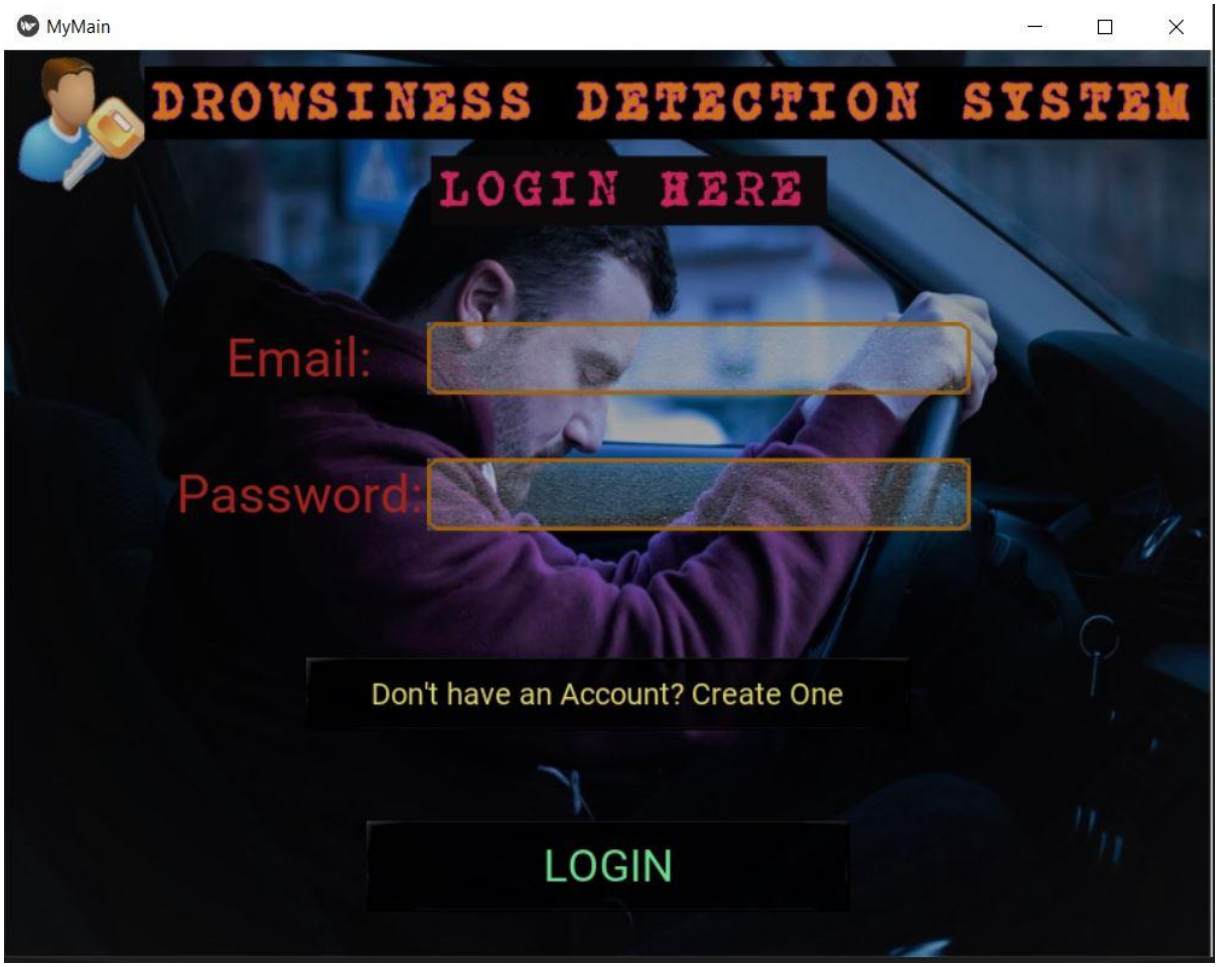
Fig 4.6 Sequence Diagram

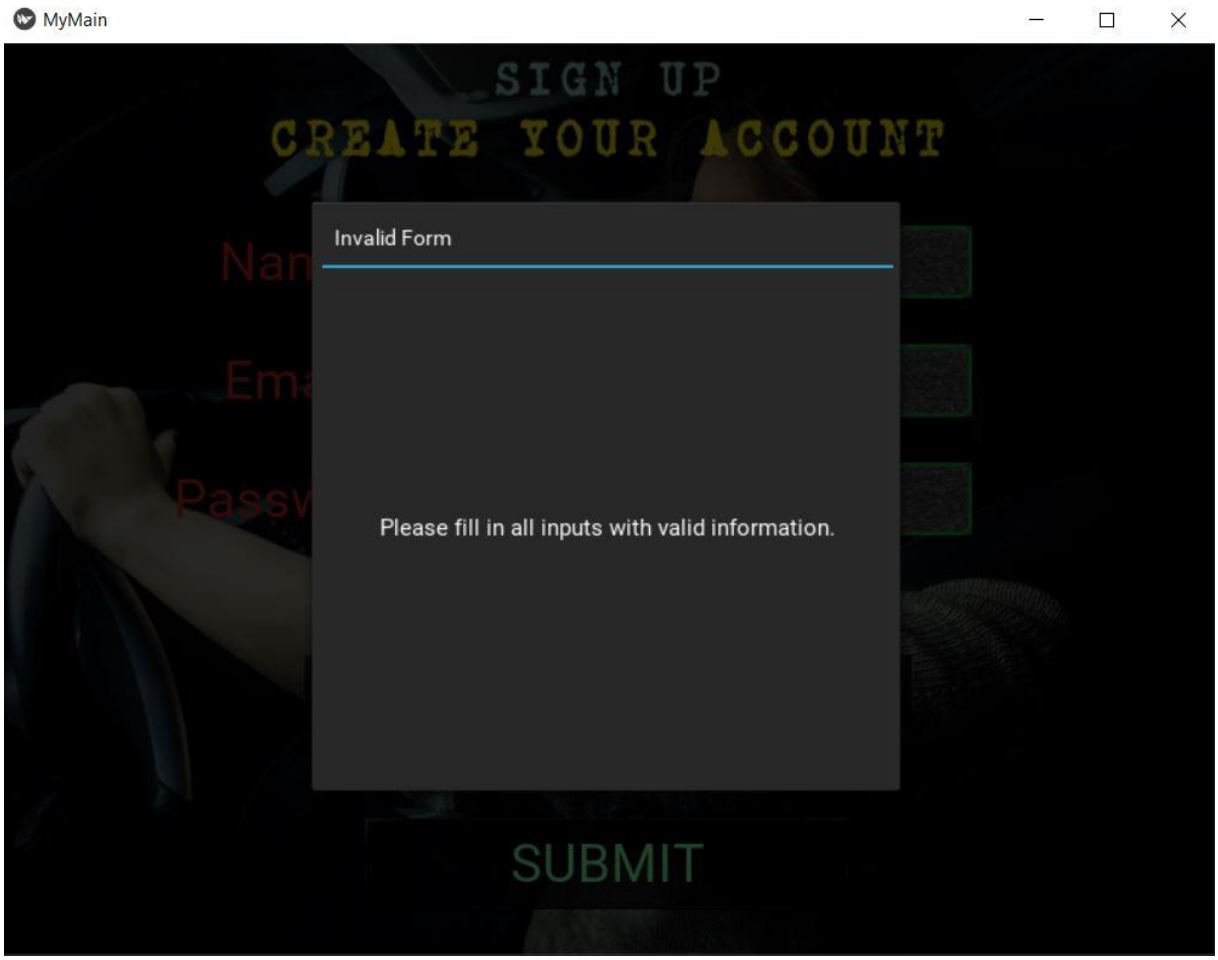
Chapter 5

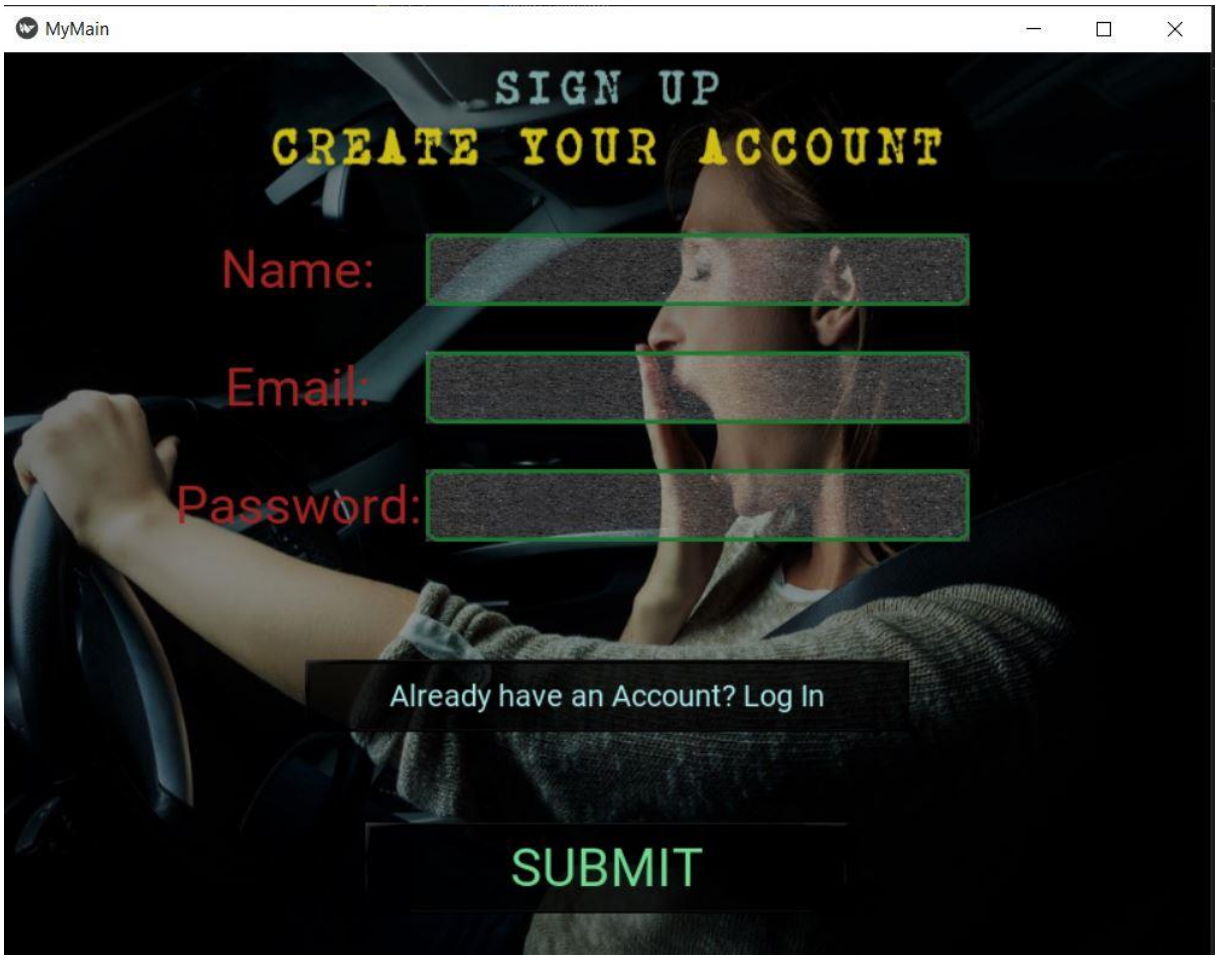
Results and Discussions

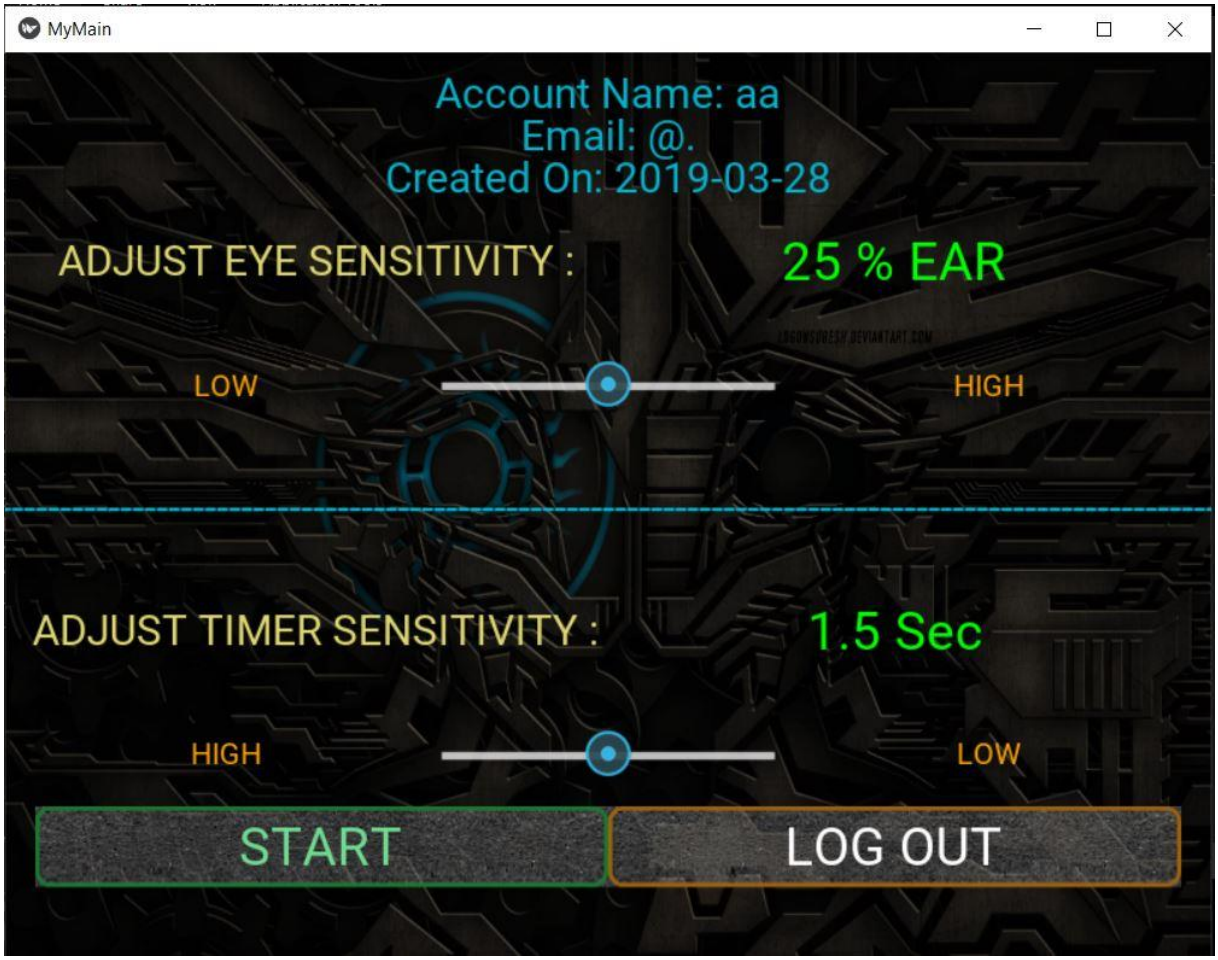
Implementation of drowsiness detection with Python and OpenCV was done which includes the following steps: Successful runtime capturing of video with camera. Captured video was divided into frames and each frame were analyzed. Successful detection of face followed by detection of eye. If closure of eye for successive frames were detected, then it is classified as drowsy condition else it is regarded as normal blink and the loop of capturing image and analyzing the state of driver is carried out again and again. In this implementation during the drowsy state the eye is not surrounded by circle or it is not detected, and corresponding message is shown.

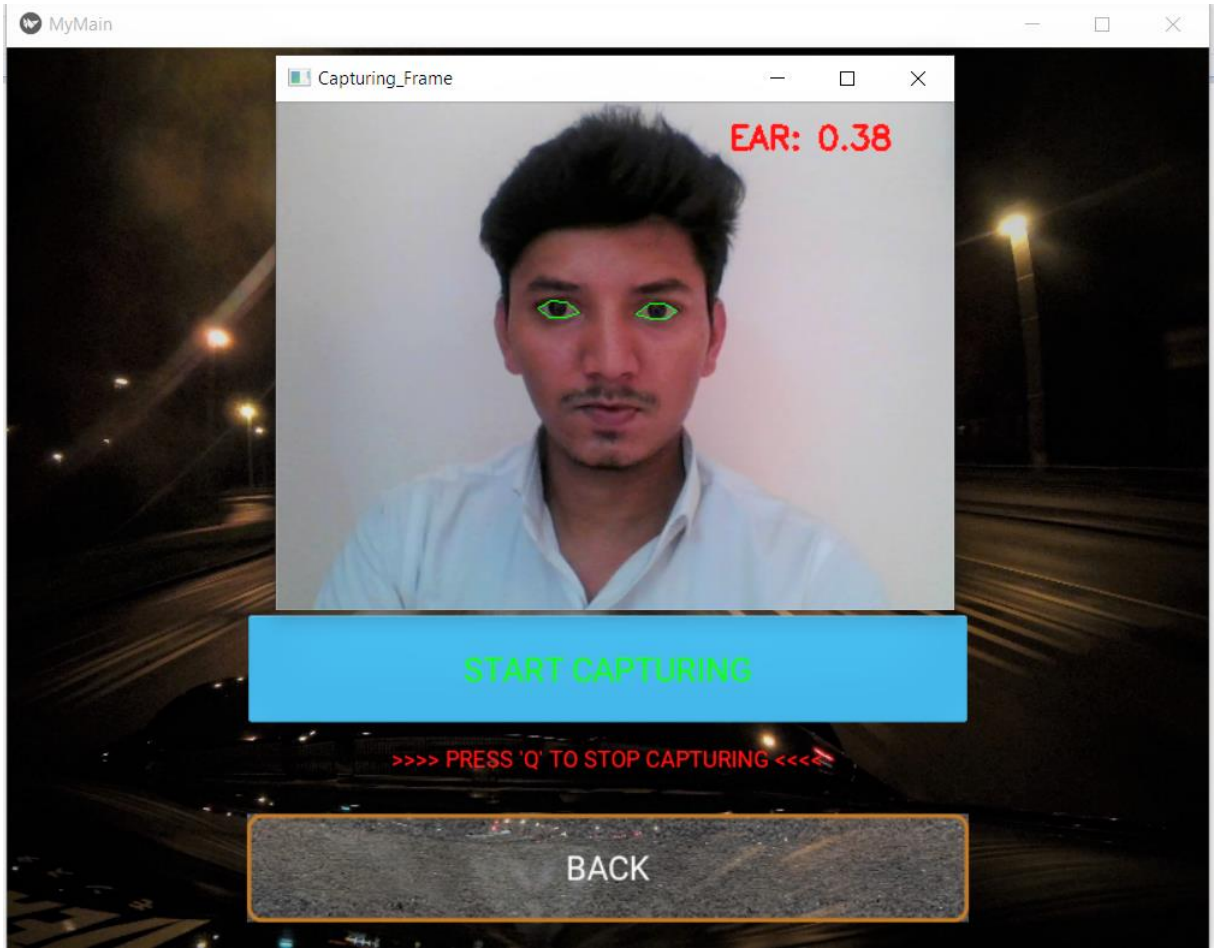
Screenshots of the system shown below:

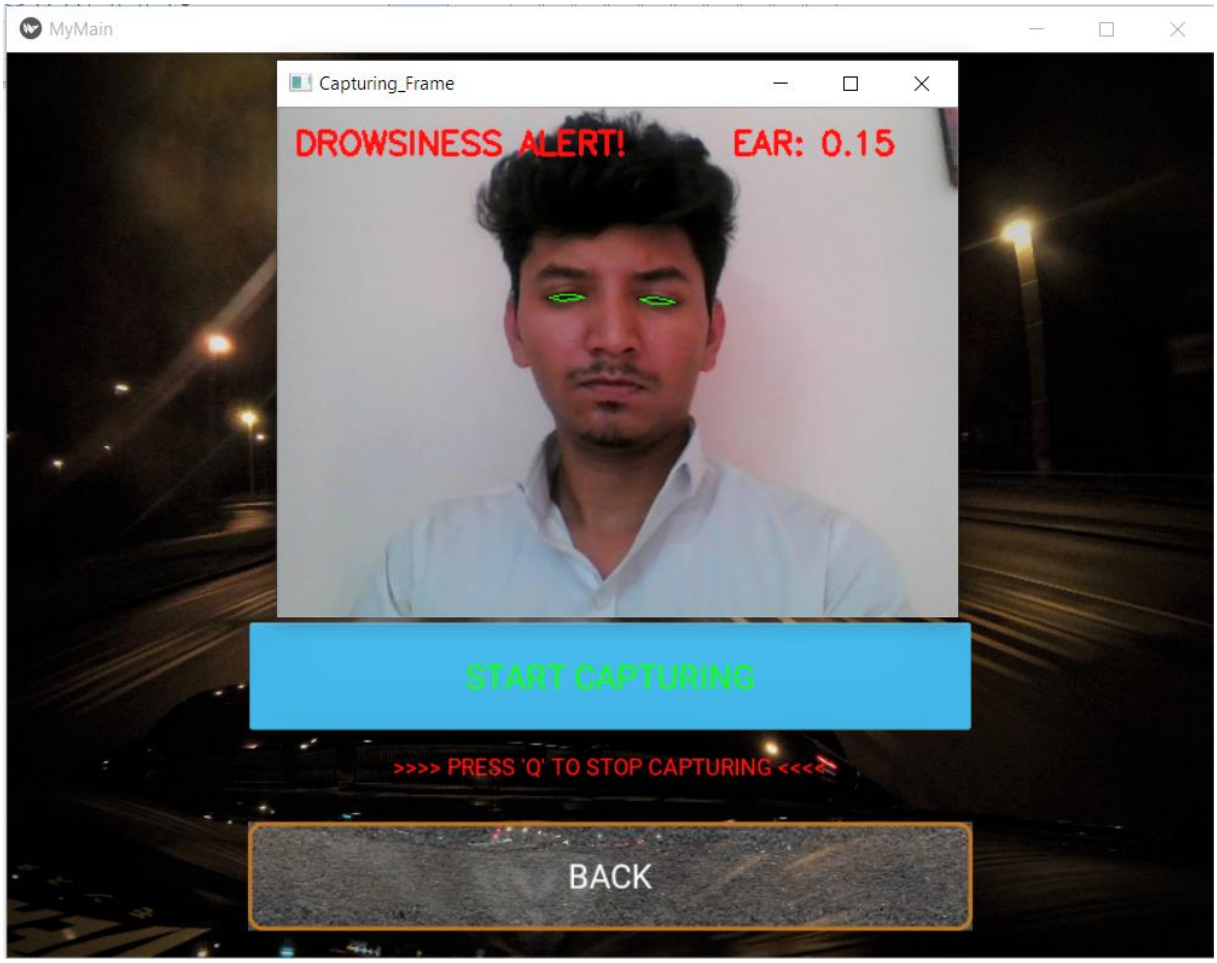



MyMain
SIGN UP
CREATE YOUR ACCOUNT
Invalid Form
Please fill in all inputs with valid information.
SUBMIT


MyMain
SIGN UP
CREATE YOUR ACCOUNT
Name:
Email:
Password:
Already have an Account? Log In
SUBMIT


MyMain
Account Name: aa
Email: @.
Created On: 2019-03-28
ADJUST EYE SENSITIVITY :
25 % EAR
LOW
HIGH
ADJUST TIMER SENSITIVITY :
1.5 Sec
HIGH
LOW
START
LOG OUT


MyMain
Capturing_Frame
EAR: 0.38
START CAPTURING
>>>> PRESS 'Q' TO STOP CAPTURING <<<<
BACK




```
[INFO   ] [Base        ] Start application main loop
[INFO] loading facial landmark predictor...
[INFO] starting video stream thread...
Drowsiness detected for : 3.5 sec
Drowsiness detected for : 1.5666666666666667 sec
Drowsiness detected for : 5.383333333333334 sec
Drowsiness detected for : 1.55 sec
Drowsiness detected for : 7.85 sec
Total Drowsiness detected for : 19.85 sec
####################################
[INFO   ] [Base        ] Leaving application in progress...
```

Future Work:

Our model is designed for detection of drowsy state of eye and give and alert signal or warning in the form of audio alarm. But the response of driver after being warned may not be enough to stop causing the accident meaning that if the driver is slow in responding towards the warning signal then accident may occur. Hence to avoid this we can design and fit a motor driven system and synchronize it with the warning signal so that the vehicle will slow down after getting the warning signal automatically.

We can also provide the user with an Android application which will provide with the information of his/her drowsiness level during any journey. The user will know Normal state, Drowsy State, the number of times blinked the eyes according to the number of frames captures. Which can be shown in fig 6.1

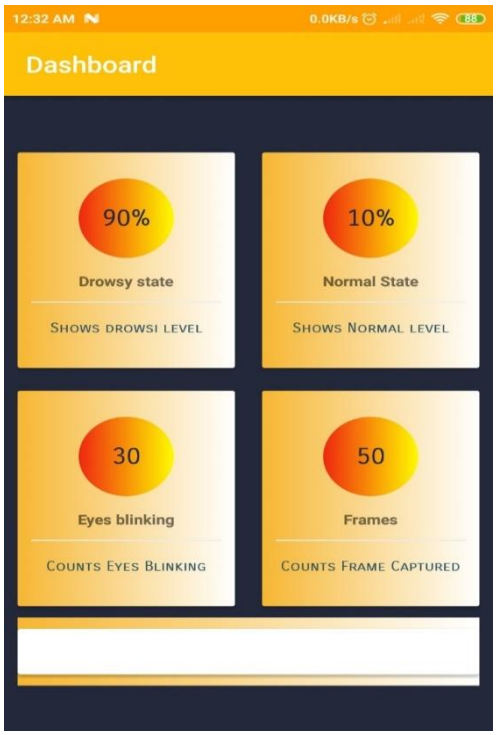


Fig 5.1 Android Application for Future scope

Chapter 6

Conclusions

A real-time eye blink detection algorithm was presented. We quantitatively demonstrated that Haar feature-based cascade classifiers and regression-based facial landmark detectors are precise enough to reliably estimate the positive images of face and a level of eye openness. While they are robust to low image quality (low image resolution in a large extent) and in-the-wild.

Limitations:

Use of spectacles: In case the user uses spectacle then it is difficult to detect the state of the eye. As it hugely depends on light hence reflection of spectacles may give the output for a closed eye as opened eye. Hence for this purpose the closeness of eye to the camera is required to avoid light.

Multiple face problem: If multiple face arises in the window then the camera may detect more number of faces undesired output may appear. Because of different condition of different faces. So, we need to make sure that only the driver face come within the range of the camera. Also, the speed of detection reduces because of operation on multiple faces.

Chapter 7

References

IEEE standard

Journal Paper,

[1] Facial Features Monitoring for Real Time Drowsiness Detection by Manu B.N, 2016 12th International Conference on Innovations in Information Technology (IIT) [Pg. 78-81]
https://ieeexplore.ieee.org/document/7880030

[2] Real Time Drowsiness Detection using Eye Blink Monitoring by Amna Rahman Department of Software Engineering Fatima Jinnah Women University 2015 National Software Engineering Conference (NSEC 2015)
https://ieeexplore.ieee.org/document/7396336

[3] Implementation of the Driver Drowsiness Detection System by K. Srijayathi International Journal of Science, Engineering and Technology Research (IJSETR) Volume 2, Issue 9, September 2013

Names of Websites referred

https://www.codeproject.com/Articles/26897/TrackEye-Real-Time-Tracking-Of-Human-Eyes- Using-a

https://realpython.com/face-recognition-with-python/

https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/

https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/

https://www.pyimagesearch.com/2017/04/10/detect-eyes-nose-lips-jaw-dlib-opencv-python/

https://www.codeproject.com/Articles/26897/TrackEye-Real-Time-Tracking-Of-HumanEyesUsing-a

https://docs.opencv.org/3.4/d7/d8b/tutorial_py_face_detection.html

https://www.learnopencv.com/training-better-haar-lbp-cascade-eye-detector-opencv/

Appendix

1.Coding

Importing our required Python packages.

detect_drowsiness.py

```
# import the necessary packages
        from scipy.spatial import distance as dist
        from imutils.video import VideoStream
        from imutils import face_utils
        from threading import Thread
        import numpy as np
        import playsound
        import argparse
        import imutils
        import time
        import dlib
        import cv2
```

Sound Alarm

```
        def sound_alarm(path):
# play an alarm sound
        playsound.playsound(path)
```

eye_aspect_ratio function

```
        def eye_aspect_ratio(eye):
```

```
# compute the euclidean distances between the two sets of
# vertical eye landmarks (x, y)-coordinates
        A = dist.euclidean(eye[1], eye[5])
        B = dist.euclidean(eye[2], eye[4])

# compute the euclidean distance between the horizontal
# eye landmark (x, y)-coordinates
        C = dist.euclidean(eye[0], eye[3])

# compute the eye aspect ratio
        ear = (A + B) / (2.0 * C)

# return the eye aspect ratio
        return ear
```

Parsing command Line Argument

```
        # construct the argument parse and parse the arguments
        ap = argparse.ArgumentParser()
        ap.add_argument("-p", "--shape-predictor", required=True,
        help="path to facial landmark predictor")
        ap.add_argument("-a", "--alarm", type=str, default="",
        help="path alarm .WAV file")
        ap.add_argument("-w", "--webcam", type=int, default=0,
        help="index of webcam on system")
        args = vars(ap.parse_args())
```

Defining EYE_AR_THRESH

```
        EYE_AR_THRESH = 0.3
        EYE_AR_CONSEC_FRAMES = 48
        COUNTER = 0
        ALARM_ON = False
```

Facial landmark predictor

```
# initialize dlib's face detector (HOG-based) and then create
# the facial landmark predictor
        print("[INFO] loading facial landmark predictor...")
        detector = dlib.get_frontal_face_detector()
        predictor = dlib.shape_predictor(args["shape_predictor"])
```

Extracting the eye regions

```
# grab the indexes of the facial landmarks for the left and
# right eye, respectively
        (lStart, lEnd) = face_utils.FACIAL_LANDMARKS_IDXS["left_eye"]
        (rStart, rEnd) = face_utils.FACIAL_LANDMARKS_IDXS["right_eye"]
```

Instantiate VideoStrem

```
# start the video stream thread
        print("[INFO] starting video stream thread...")
        vs = VideoStream(src=args["webcam"]).start()
        time.sleep(1.0)
 # loop over frames from the video stream
    while True:
        frame = vs.read()
        frame = imutils.resize(frame, width=450)
        gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
 # detect faces in the grayscale frame
       rects = detector(gray, 0)
```

Facial landmark detection to localize each of the important regions of the face:

```
# loop over the face detections
    for rect in rects:
# determine the facial landmarks for the face region, then
# convert the facial landmark (x, y)-coordinates to a NumPy
# array
    shape = predictor(gray, rect)
    shape = face_utils.shape_to_np(shape)


# extract the left and right eye coordinates, then use the
# coordinates to compute the eye aspect ratio for both
eyes
    leftEye = shape[lStart:lEnd]
    rightEye = shape[rStart:rEnd]
    leftEAR = eye_aspect_ratio(leftEye)
    rightEAR = eye_aspect_ratio(rightEye)
# average the eye aspect ratio together for both eyes
     ear = (leftEAR + rightEAR) / 2.0
```

Visualize each of the eye regions

```
# compute the convex hull for the left and right eye, then
   leftEyeHull = cv2.convexHull(leftEye)
   rightEyeHull = cv2.convexHull(rightEye)
   cv2.drawContours(frame, [leftEyeHull], -1, (0, 255, 0),1)
   cv2.drawContours(frame, [rightEyeHull], -1, (0, 255, 0),1)
```

Check to see if the person in our video stream is starting to show symptoms of drowsiness:

```
# check to see if the eye aspect ratio is below the blink
# threshold, and if so, increment the blink frame counter
        if ear < EYE_AR_THRESH:
        COUNTER += 1

# if the eyes were closed for a sufficient number of
# then sound the alarm
        if COUNTER >= EYE_AR_CONSEC_FRAMES:
# if the alarm is not on, turn it on
        if not ALARM_ON:
        ALARM_ON = True

# check to see if an alarm file was supplied,
# and if so, start a thread to have the alarm
# sound played in the background
        if args["alarm"] != "":
        t = Thread(target=sound_alarm,
        args=(args["alarm"],))
        t.deamon = True
        t.start()

# draw an alarm on the frame
        cv2.putText(frame, "DROWSINESS ALERT!", (10, 30),
        cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 0, 255), 2)

# otherwise, the eye aspect ratio is not below the blink
# threshold, so reset the counter and alarm
        else:
        COUNTER = 0
        ALARM_ON = False
```

Displaying the output frame:

```
# draw the computed eye aspect ratio on the frame to help
# with debugging and setting the correct eye aspect ratio
# thresholds and frame counters
        cv2.putText(frame, "EAR: {:.2f}".format(ear), (300, 30),
        cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 0, 255), 2)
 # show the frame
        cv2.imshow("Frame", frame)
        key = cv2.waitKey(1) & 0xFF
 # if the `q` key was pressed, break from the loop
        if key == ord("q"):
        break
 # do a bit of cleanup
        cv2.destroyAllWindows()
        vs.stop()
```

Acknowledgements

I am profoundly grateful to **Dr. Varsha Shah** for his expert guidance and continuous encouragement throughout to see that this project rights its target.

I would like to express deepest appreciation towards Dr. Varsha Shah, Principal RCOE, Mumbai and Prof.Shiburaj Pappu HOD Computer Department whose invaluable guidance supported me in this project.

At last I must express my sincere heartfelt gratitude to all the staff members of Computer Engineering Department who helped us directly or indirectly during this course of work.

Roshan Tavhare

In Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM, 360-367, 2019, Prague, Czech Republic.
Link to paper

A Study of Various Text Augmentation Techniques for Relation Classification in Free Text

Praveen Kumar Badimala Giridhara[1,2,*], Chinmaya Mishra[1,2,*], Reddy Kumar Modam Venkataramana[1,*], Syed Saqib Bukhari[2] and Andreas Dengel[1,2]
[1]*Deprtment of Computer Science, TU Kaiserslautern, Gottlieb-Daimler-Straße 47, Kaiserslautern, Germany*
[2] *German Research Center for Artificial Intelligence (DFKI), Kaiserslautern, Germany*
{*praveen_kumar.badimala_giridhara, saqib.bukhari, Andreas.Dengel*}*@dfki.de,* {*mishra, modam*}*@rhrk.uni-kl.de*

Keywords: Relation Classification, Text Data Augmentation, Natural Language Processing, Investigative Study

Abstract: Data augmentation techniques have been widely used in visual recognition tasks as it is easy to generate new data by simple and straight forward image transformations. However, when it comes to text data augmentations, it is difficult to find appropriate transformation techniques which also preserve the contextual and grammatical structure of language texts. In this paper, we explore various text data augmentation techniques in text space and word embedding space. We study the effect of various augmented datasets on the efficiency of different deep learning models for relation classification in text.

1 INTRODUCTION

Relation classification is an important task in processing free text using Natural Language Processing (NLP). The basic idea behind relation classification is to classify a sentence into a predefined relation class, given the two entities in the sentence. For instance, in the sentence; $< e1 >$ *Bill Gates* $< /e1 >$ *is the founder of* $< e2 >$ *Microsoft* $< /e2 >$,"Bill Gates" and "Microsoft" are the two entities denoted by "$< e1 >< /e1 >$" and "$< e2 >< /e2 >$" respectively. The relation classification process should be able to classify this sentence as belonging to the relation class *"founderOf"* (which is a predefined class). One of the major limitations in the field of NLP is the unavailability of labelled data. It takes a great deal of time and effort to manually annotate relevant datasets. Hence, it has become increasingly necessary to come up with automated annotation methods. This has led to the development of many semi-supervised annotation methods for generating annotated datasets. But, they fall short when compared to the quality and efficiency of a manually annotated dataset.

One workaround would be to perform data augmentations on manually annotated datasets. Data augmentation techniques are very popular in the field of image processing because of the ease in generating augmentations using simple image transformations. However, it is very challenging to find an appropriate method for text data augmentation as it is difficult to preserve grammar and semantics.

As per our knowledge, there have been no work that studies different text data augmentation techniques on relation classification for free text over different Deep Learning Models as yet. In this paper, we investigate various text data augmentation techniques while retaining the grammatical and contextual structure of the sentences when applying them. We also observe the behavior of using augmented datasets with the help of two different deep learning models namely, CNN (Zeng et al., 2014) and Attention based BLSTM (Zhou et al., 2016).

2 RELATED WORKS

Data Augmentation in the field of Image Processing and Computer Vision is a well-known methodology to increase the dataset by introducing varied distributions and increase the performance of the model for a number of tasks. In general, it is believed that the more the data a neural network gets trained on, the more effective it becomes. Augmentations are performed by using simple image transformations such as rotation, cropping, flipping, translation and addition of Gaussian noise to the image. Krizhevsky et

*contributed equally

al., (Krizhevsky et al., 2012) used data augmentation methods to increase the training data size for training a deep neural network on ImageNet dataset (Deng et al., 2009). The increase in training data samples showed reduced overfitting of the model (Krizhevsky et al., 2012) and increased the model performance. These techniques enable the model to learn additional patterns in the image and identify new positional aspects of objects in it.

On similar lines, data augmentation methods are explored in the field of text processing for improving the efficacy of models. Mueller and Thyagarajan (Mueller and Thyagarajan, 2016) replaced random words in a sentence with their respective synonyms, to generate augmented data and train a siamese recurrent network for sentence similarity task. Wang and Yang (Wang and Yang, 2015) used word embeddings of sentences to generate augmented data for the purpose of increasing data size and trained a multi-class classifier on tweet data. They found the nearest neighbour of a word vector by using cosine similarity and used that as a replacement for the original word. The word selection was done stochastically.

For information extraction, Papadaki (Papadaki, 2017) applied data augmentation techniques on legal dataset (Chalkidis et al., 2017). A class specific probability classifier was trained to identify a particular contract element type for each token in a sentence. They classified a token in a sentence based on the window of words/tokens surrounding them. They used word embeddings obtained by pre-training a word2vec model (Mikolov et al., 2013) on unlabeled contract data. Their work examined three data augmentation methods namely; interpolation, extrapolation and random noise. The augmentation methods manipulated the word embeddings to obtain a new set of sentence representations. The work by Papadaki (Papadaki, 2017) also highlighted that interpolation method performed comparatively better than the other methods like extrapolation and random noise. The work by Zhang and Yang (Zhang and Yang, 2018) explored various perturbation methods where they introduced random perturbations like Gaussian noise or Bernouli noise into the word embeddings in text related classification tasks such as sentence classification, sentiment classification and relation classification.

One of the recent works by Kobayashi (Kobayashi, 2018) trained a bi-directional language model conditioned on class labels where it predicted the probability of a word based on the surrounding context of the words. The words with best probability values were taken into consideration while generating the augmented sentences wherein the words in the sentences were replaced in a paradigmatic way.

However, to the best of our knowledge, there have been no work that specifically focuses on studying the different text data augmentation techniques on Relation Classification task in free text so far.

3 AUGMENTATION METHODS

In this Section, we describe various types of augmentation techniques that have been used in our experiments. As discussed briefly in Section 1, it is very challenging to create manually annotated datasets due to which we only have a few publicly available datasets with acceptable number of training and test data. Due to the constraints of cost and effort, most of the manually annotated datasets have less number of samples and it becomes difficult to optimally train deep learning models with the limited amount of data.

The use of distant supervision methods Mintz et al., (Mintz et al., 2009) to annotate data is able to compensate for the lack of quantity but is often susceptible to inclusion of noise when generating the dataset. This in turn constrains the performance of training models while performing relation classification. We consider applying augmentations on a manually augmented dataset as one of the ways to workaround the problem.

In our experiments, the training data was augmented at two levels; the text level and the word embedding level. Word Similarity and Synonym methods were used to generate new texts whereas interpolation and extrapolation methods (Papadaki, 2017) were used to generate embedding level augmentations. In order to apply the augmentation techniques, we tagged the sentences using NLTK (Bird et al., 2009) POS tagger. We restricted the augmentations only to nouns, adjectives and adverbs in each sentence. It was observed that by applying the restriction, we were in a better position to preserve the grammatical and semantic structures of the original sentences as compared to randomly replacing the words. GloVe word vectors (Pennington et al., 2014), which are a collection of pre-trained word vectors were used as word embeddings for words in our experiments.

We describe each of the augmentation methods in the following subsections.

3.1 Similar Words Method

This method by Wang and Yang (Wang and Yang, 2015) of text data augmentation works by exploiting the availability of similar words in the word embed-

ding space. We replaced words with their respective top scored similar words to generate new sentences. An example input sentence and the resulting augmented sentence are given below:

- *Input Sentence* : The winner was $< e1 >$ Marilyn Churley $< /e1 >$ of the $< e2 >$ Ontario $< /e2 >$ New Democratic Party .
- *Augmented Sentence* : The winners was $< e1 >$ Monroe Churley $< /e1 >$ of the $< e2 >$ Manitoba $< /e2 >$ York Democrat parties

Here "Ontario" is replaced with "Manitoba", "New Democratic Party" with "York Democrat parties" among others.

3.2 Synonym Method

We followed the work by Mueller and Thyagarajan (Mueller and Thyagarajan, 2016) where they randomly replaced words with their synonyms obtained from WordNet Synonym Dictionary (Miller, 1995). We restricted the words to nouns, adjectives and adverbs and replaced them with their respective top scored synonym words to generate new sentences. A sample input sentence and the resulting augmented sentence are shown below:

- *Input Sentence* : Dominguez was not hotheaded said $< e1 >$ Woods $< /e1 >$ a former $< e2 >$ Arizona $< /e2 >$ attorney general.
- *Augmented Sentence* : Dominguez was not hotheaded said $< e1 >$ Wood $< /e1 >$ a former $< /e2 >$ Arizona $< /e2 >$ lawyer general.

Here "Woods" is replaced with "Wood", "attorney" with "lawyer" among others.

3.3 Interpolation Method

As described in work by Papadaki (Papadaki, 2017), we first obtained the top three nearest neighbours of a word in the embedding space. Then, we found the centroid from the embedding vectors of the nearest neighbours.

We calculated the new word embedding vector from the centroid and the original word embedding using the formula below:

$$w_j^{'} = (w_k - w_j)\lambda + w_j \quad (1)$$

In equation 1, $w_j^{'}$ denotes the new word embedding, w_k denotes the centroid, w_j denotes the original word embedding and λ is a parameter within the range of [0-1]. It controls the degree of interpolation. The new word embedding vector tends to move away from the original word embedding in the direction towards the centroid based on the λ value. The λ value that was used in our experiments is 0.25. Figure 1 depicts a graphical representation of the procedure.

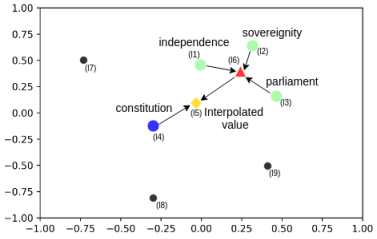


Figure 1: Interpolation between the centroid and word embedding to create a new embedding vector. The green points (l1,l2 and l3) represent the selected nearest neighbors, red indicator mark (l6) is the centroid obtained from the nearest neighbors, blue point (l4) depicts the original embedding vector and the yellow indicator mark (l5) depicts the resulting new word embedding vector. The black points represent other word embeddings present in the embedding space.

After obtaining the new word embedding vectors for the words in a sentence, we replaced them in place of their original embeddings and generated a new word embedding list for the sentence.

3.4 Extrapolation Method

Similar to the interpolation method (subsection 3.3), we calculated the new word embedding vector from the centroid and the original word embedding for a word using the formula given below (Papadaki, 2017):

$$w_j^{'} = (w_j - w_k)\lambda + w_j \quad (2)$$

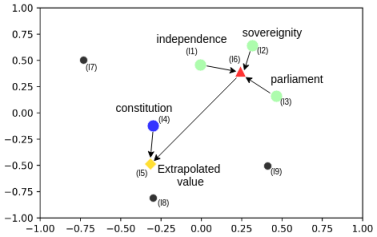


Figure 2: Extrapolation between the centroid and word embedding to create a new embedding vector. The labels are the same as figure 1.

In equation 2, $w_j^{'}$ denotes the new word embedding, w_k denotes the centroid, w_j denotes the old word embedding and λ is a parameter with value in the range [0-∞] that controls the degree of extrapolation. The new word embedding vector tends to move away from the original word embedding in the direction opposite to the centroid constrained by the λ value. The

λ value that was used in our experiments is 0.25 for SemEval2010 and 0.5 for KBP37. Figure 2 depicts a graphical representation of the procedure.

3.5 Random Noise Method

As described in the work by Papadaki (Papadaki, 2017), in this method, instead of generating new augmented sentences we inserted perturbations to the word embeddings in the embedding layer of our model. For each word embeddings we added a Gaussian noise value to them as per the equation below:

$$w_j^{'} = w_j + x_j \quad (3)$$

where, w_j is the original word embedding, $w_j^{'}$ is the resulting word embedding and x_j is the random noise value. The values for each vector element of noise was sampled from the truncated normal distribution with $\mu = 0$, $\sigma = 1$ and with range between 0 and 0.3. Later we randomly selected the vector elements with a probability of 0.3 and rest of the vector elements were set to zero. The resulting vector is considered as noise vector. The values were kept in a small range so as to keep the resulting word embedding from moving too far away from the contextual word embedding space. One exception in this method to that of the other augmentation methods is that, we inserted perturbations in all the words regardless of its POS tag.

We have tried to explore the strengths and weaknesses of these methods in the experiments.

4 DEEP LEARNING METHODS

We have used two deep learning models namely; CNN and Attention-Based Bidirectional LSTM, for our experiments. Both the models have been briefly introduced in the subsections below.

4.1 CNN

Convolutional Neural Networks (CNN) were first implemented by (LeCun et al., 1998) on MNIST dataset. The basic principle behind CNN is to apply convolutions using multi-dimensional filters over an image as a sliding window to extract feature maps. CNNs have proved to be very efficient in the field of image processing as they are able to exploit the spatial structure of pixel intensities in an image. Recently, it has been observed that CNNs have also been able to extract lexical and sentence level features. We decided to use the text CNN model proposed by Zeng et al., (Zeng et al., 2014) which is widely used for relation classification tasks.

4.2 Attention-Based Bidirectional LSTM

Recurrent Neural Networks (RNNs) are frequently used in the NLP domain. RNNs capture the sequential word patterns since they are inherently able to remember the previous states. RNNs are good at modelling sequential data (Boden, 2002). RNNs suffer from vanishing gradient problem. Long Short Term Memory (LSTM), which is a subset of RNN, is used to overcome the vanishing gradient problem by making use of a gated approach. One of the recent variants of LSTM for language specific tasks is Attention based LSTMs. They are able to capture the key part of the sentence by assigning higher weights to the relevant parts of an input sentence. Attention mechanism in neural networks have shown success in solving different tasks, like question answering (Hermann et al., 2015) and machine translation (Bahdanau et al., 2014). The attention mechanism in a model allows us to capture the most important semantic information in a sentence. In our experiments we are using the Attention-Based Bidirectional LSTM model proposed by Peng Zhou and Xu (Peng Zhou and Xu, 2016).

5 DATASETS

We used two datasets to run the augmentation experiments. The first dataset is **SemEval2010** (Hendrickx et al., 2009) which has a total of 10 unique relation classes (ignoring the direction aspect). It is a manually annotated public dataset which is generally accepted as a benchmark for relation classification tasks. The second dataset is the **KBP37** dataset which was created by Dongxu and Dong (Dongxu and Dong, 2016) and has a total of 19 unique relation types (ignoring direction aspect). We decided to use this (KPB37) dataset because it has more relation types and is more realistic than the SemEval2010 dataset (Dongxu and Dong, 2016). We wanted to experiment how different augmentation techniques worked on these datasets.

Each relation type (except *'other'* and *'no_relation'*) is further split into two relation classes for both the datasets, which considers the two directions w.r.t the positioning of the entities in a sentence. As an example, the relation *'org:subsidiaries'* is further split into *'org:subsidiaries(e1,e2)'* and *'org:subsidiaries(e2,e1)'* based on the location of the two entities e1 and e2 in a sentence. A few statistical details about both the datasets are provided in table 1.

Table 1: Details about SemEval2010 and KBP37 datasets.

Dataset: SemEval2010	No. of sentences in train: 8000 No. of sentences in test: 2717 No. of relation types: 19 No. of relation classes: 10	
	Classes: Cause-Effect	Instrument-Agency
	Product-Producer	Content-Container
	Entity-Origin	Entity-Destination
	Component-Whole	Member-Collection
	Communication-Topic	Other
Dataset: KBP37	No. of sentences in train: 15917 No. of sentences in test: 3405 No. of relation types: 37 No. of relation classes: 19	
	Classes: per:alternate_names	org:alternate_names
	per:origin	org:subsidiaries
	per:spouse	org:top_members/employees
	per:title	org:founded
	per:employee_of	org:founded_by
	per:countries_of_residence	org:country_of_headquarters
	per:stateorprovinces_of_residence	org:stateorprovince_of_headquarters
	per:cities_of_residence	org:city_of_headquarters
	per:country_of_birth	org:members
	no_relation	

6 EXPERIMENTS

As mentioned in Section 4, we implemented the text CNN model proposed by (Zeng et al., 2014) and attention-based BLSTM model proposed by Zeng et al., (Zhou et al., 2016). These two models were considered as the benchmark models for our experiments. We used PF (Position Features) method as introduced by (Zeng et al., 2014), where we specify the distance of each word in a sentence to the two entities (nominals) and feed them to the CNN and the attention-based BLSTM models along with the input sentences. For instance, in the sentence *"The winner was* $<e1>$ *Marilyn Churley* $</e1>$ *of the* $<e2>$ *Ontario* $</e2>$ *New Democratic Party."*, words between $<e1></e1>, <e2></e2>$ denote the two entities w_1, w_2 respectively. We find the distance of the current word to w_1 and w_2 and use the resulting distance vectors d_1 and d_2 in the models.

For each dataset, we randomly extracted 75% samples from the train data and used them as a reduced dataset (which we refer to as 75% dataset in the rest of the paper) for the experiments. While re-sizing the datasets, we took care to maintain the "data to class ratio" of the original datasets by selecting a proportional number of samples from each of the classes. Augmentation methods were applied on each of the datasets, namely the 100% and the 75% datasets. While generating augmented data, the same PF and class labels were used as that of the source sentences. The resulting augmented data was appended to the original data in order to generate new train data. For example, on the 75% KBP37 train data, after running the synonym method we obtain a new list of sentences. This list was then appended to the original 75% train data and the resulting list was used as a training file for our models.

As mentioned in Section 3, we only considered nouns, adjectives and adverbs for data augmentations. It came to our notice that in many cases, replacing verbs/prepositions with similar words/synonyms resulted in making the sentences ungrammatical and also in some cases changed the contextual meaning of the sentence. The resulting sentences were no longer a logical fit for their respective relation classes. However, replacing nouns, adjectives and adverbs retained the meaning even though at times it was not 100% grammatical. We were able to retain the context and generate augmented data by restricting the replacements to the above mentioned POS tags for each sentence in the datasets.

When training the models, we first found out the optimal hyper-parameters for each of the datasets (original, and 75%) for both KBP37 and SemEval2010 and used them as baseline parameters. Since not every hyper-parameter detail and implementation was available for the considered models (section 4), we implemented and optimized the mod-

Table 2: F1 scores obtained for KBP37 dataset.

Deep Learning Methods	Experiments	F1 scores and Train Data Size			
		100% Training Samples		75% Training Samples	
		F1 score	No. of Training Samples	F1 score	No. of Training Samples
CNN	Baseline	50.74	15917(-)	49.22	11922(-)
	Similar Words	50.86 ↑	31834(+)	49.45 ↑	23844(+)
	Synonym	51.68 ↑	31834(+)	49.28 ↑	23844(+)
	Interpolation	50.89 ↑	31834(+)	50.25 ↑	23844(+)
	Extrapolation	50.36 ↓	31834(+)	48.75 ↓	23844(+)
	Random Noise	50.45 ↓	31834(+)	49.90 ↑	23844(+)
att-BLSTM	Baseline	51.84	15917(-)	49.05	11922(-)
	Similar Words	51.84 =	31834(+)	50.40 ↑	23844(+)
	Synonym	51.60 ↓	31834(+)	49.84 ↑	23844(+)
	Interpolation	50.01 ↓	31834(+)	49.49 ↑	23844(+)
	Extrapolation	51.51 ↓	31834(+)	49.87 ↑	23844(+)
	Random Noise	51.19 ↓	31834(+)	49.63 ↑	23844(+)

(-) denotes the original number of training samples. (+) denotes that augmented data has been added to the original training samples. (=) denotes same F1 score with respect to the corresponding baseline score in the same column. (↑) denotes an increase in F1 score with respect to the corresponding baseline score in the same column. (↓) denotes a decrease in F1 score with respect to the corresponding baseline score in the same column.

els. Once we obtained an accuracy closer to the benchmark after several runs, we froze the hyper-parameters. Then, for each dataset we applied the augmentation methods and generated the corresponding new augmented data. In order to verify the effect of augmentations on the results, we used the same hyper-parameters and ran the models with the new augmented train data. We did not do any cross validations during training. We used the pre-trained, 300-dimensional GloVe word vectors for all the experiments. The embedding layer was kept as *non-trainable* in all the models while training. Test data was kept separate and no augmentations or data resizing was done on it. For all the trained models we used the same test data, i.e, for all experiments run on KBP37 dataset, we used the test set provided for KBP37 and the same was done for SemEval2010.

7 EXPERIMENTAL RESULTS

Table 2 shows the F1 scores for the experiments conducted on KBP37 dataset. The results are split for 100% training samples and 75% training samples. For testing, we used the same test set for all the trained models which had 3405 sample sentences as provided in the KBP37 dataset (1).

Table 3 shows the F1 scores for the experiments conducted on SemEval2010 dataset. As described above for KBP37 dataset, the results are split for 100% training samples and 75% training samples and we also mention the number of training samples used in each of the experiments. For testing, we used the same test set for all the trained models which had 2717 sample sentences as provided in SemEval2010 dataset (1).

For KBP37 dataset, we take the results obtained by Dongxu and Dong (Dongxu and Dong, 2016) for PF(Position Feature) experiments as the state of the art, where they obtained an F1-score of 51.3% for CNN model and 54.3% for RNN model. Similarly, for SemEval2010 dataset we consider F1 score of 78.9% obtained by Zeng et al., (Zeng et al., 2014) for CNN and F1 score of 84.0% obtained by Zhou et al., (Zhou et al., 2016) for att-BLSTM as state of the art results. The best results obtained after multiple runs have been mentioned under baseline results. Since the embedding layers were kept as non-trainable in all the models, a slight dip is observed in the F1 scores with respect to the state of the art results considered.

A decrease of up to 3% is observed between the baseline F1 scores for 75% data samples and the F1 scores for 100% data samples. The F1 scores for baseline and the F1 scores obtained from models trained on augmented data vary by approximately $\pm 2\%$. We observe an increase in the F1 scores for all deep learning models across both the datasets, when training on augmented data generated with synonym method over 75% training samples.

One key observation is that for augmentation experiments performed over 75% dataset, there is at least one augmentation method where we are able to achieve an F1 score similar to that of F1 score for baseline (100% data). This holds true for both the

Table 3: F1 scores obtained for SemEval2010 dataset.

Deep Learning Methods	Experiments	F1 scores and Train Data Size			
		100% Training Samples		75% Training Samples	
		F1 score	No. of Training Samples	F1 score	No. of Training Samples
CNN	Baseline	79.51	8000(-)	77.31	5994(-)
	Similar Words	78.35 ↓	16000(+)	77.72 ↑	11988(+)
	Synonym	77.97 ↓	16000(+)	77.64 ↑	11988(+)
	Interpolation	79.84 ↑	16000(+)	76.70 ↓	11988(+)
	Extrapolation	79.52 ↑	16000(+)	78.08 ↑	11988(+)
	Random Noise	76.54 ↓	16000(+)	76.56 ↓	11988(+)
att-BLSTM	Baseline	79.89	8000(-)	78.00	5994(-)
	Similar Words	79.41 ↓	16000(+)	78.03 ↑	11988(+)
	Synonym	79.04 ↓	16000(+)	78.42 ↑	11988(+)
	Interpolation	78.40 ↓	16000(+)	75.78 ↓	11988(+)
	Extrapolation	78.08 ↓	16000(+)	76.59 ↓	11988(+)
	Random Noise	78.79 ↓	16000(+)	78.69 ↑	11988(+)

(-) denotes the original number of training samples. (+) denotes that augmented data has been added to the original training samples. (↑) denotes an increase in F1 score with respect to the corresponding baseline score in the same column. (↓) denotes a decrease in F1 score with respect to the corresponding baseline score in the same column.

SemEval2010 and KBP37 datasets. For instance, we can consider the F1 score for baseline (100% data) for KBP37 dataset. For CNN model we get an F1 score of 50.74%. For interpolation experiment performed on 75% dataset, we observe an F1 score of 50.25% which is very close to the 100% baseline result of 50.74%. The same can also be observed for att-BLSTM model on KPB37 dataset, where we get a 100% baseline F1 score of 51.84% and 50.40% for Similar Words method performed on 75% train data. All such results have been highlighted in blue in both the results tables (3, 2).

With these observations, we extrapolate that adding meaningful augmentations which preserve the grammatical and the semantic structures of the original sentences enable us to obtain results similar to that of the original dataset even with less data. However, one limitation in the augmentation techniques used is the lack of ability to introduce new variations with respect to the positions of entities in the original sentences. For example, for an input Sentence: *"Dominguez was not hotheaded said* $< e1 >$ *Woods* $< /e1 >$ *a former* $< e2 >$ *Arizona* $< /e2 >$ *attorney general."*, on applying synonym method we obtain, *"Dominguez was not hotheaded said* $< e1 >$ *Wood* $< /e1 >$ *a former* $< /e2 >$ *Arizona* $< /e2 >$ *lawyer general."*. It can be clearly seen that even though we are able to generate a new sentence which preserves the grammatical and semantic structure of the original sentence, we do not change the positions of the entities denoted by $< e1 >, < /e1 >$ and $< e2 >< /e2 >$ respectively. The position features contribute significantly to the learning while training the models. A lack of new entity ordering results in an almost saturated F1 score for augmentation experiments run on 100% data. Additionally, the interpolation and extrapolation methods may perform better if we used word vectors that are more relevant to these datasets thereby capturing their data distributions, instead of using pre trained GloVe word vectors. We have not explored all the variations of hyper-parameter values for each augmentation method. Further tuning of these hyper-parameters may yield better results.

8 CONCLUSION

In this paper, we investigated the available techniques for text data augmentations with respect to Relation Classification in free text. As per our knowledge, there have been no work that studies different text data augmentation techniques on relation classification for free text over different Deep Learning Models as yet. We implemented five text data augmentation techniques and explored the ways in which we could preserve the grammatical and the contextual structures of the sentences while generating new sentences automatically using data augmentation techniques. We discussed the importance of manually annotated data and why it is crucial to build new methods to augment data on smaller manually annotated datasets. The experiments are carried on one subsampled dataset (75%) and it will be interesting to observe the effects of augmentations in even more detail on further subsampled datasets.

From our experiments, we observed that we are

able to mimic the performance of an original dataset by adding augmented data to a small dataset. We also highlighted the inability of the existing text data augmentation techniques to introduce new features for learning. This is a limitation which bars automatic data generation from being as diverse and contextual as a manually annotated dataset with real texts. Hence, it is crucial to develop new augmentation methods which can introduce diversity and at the same time retain the grammar and context in a sentence. As a future work, one can experiment different combinations of these methods as an hybrid approach and see the possible improvement of accuracy for different NLP related tasks and also generate synthetically similar sentences using Generative Adversarial Networks (GANs) which are widely used in image domain to obtain synthetic data.

REFERENCES

Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Bird, S., Klein, E., and Loper, E. (2009). *Natural Language Processing with Python*. OReilly Media Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 1st edition.

Boden, M. (2002). A guide to recurrent neural networks and backpropagation. In *The Dallas Project, SICS Technical Report T2002:03*.

Chalkidis, I., Androutsopoulos, I., and Michos, A. (2017). Extracting contract elements. In *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*, pages 19–28. ACM.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE.

Dongxu, Z. and Dong, W. (2016). Relation classification: Cnn or rnn? In *Natural Language Understanding and Intelligent Applications*, pages 665–675, Cham. Springer International Publishing.

Hendrickx, I., Kim, S. N., Kozareva, Z., Nakov, P., Ó Séaghdha, D., Padó, S., Pennacchiotti, M., Romano, L., and Szpakowicz, S. (2009). Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions*, pages 94–99. Association for Computational Linguistics.

Hermann, K. M., Kocisky, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., and Blunsom, P. (2015). Teaching machines to read and comprehend. In *Advances in Neural Information Processing Systems*, pages 1684–1692.

Kobayashi, S. (2018). Contextual augmentation: Data augmentation by words with paradigmatic relations. In *NAACL-HLT*.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. In *Gradient-based learning applied to document recognition.Proceedings of the IEEE, 86(11)*, pages 2278–2324. IEEE.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, pages 3111–3119. Curran Associates Inc.

Miller, G. A. (1995). Wordnet: A lexical database for english. In *Communications of the ACM Vol. 38. No. 11*, pages 39–41. ACM.

Mintz, M., Bills, S., Snow, R., and Jurafsky, D. (2009). Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 1003–1011. Association for Computational Linguistics.

Mueller, J. and Thyagarajan, A. (2016). Siamese recurrent architectures for learning sentence similarity. In *AAAI*, pages 2786–2792.

Papadaki, M. (2017). *Data Augmentation Techniques for Legal Text Analytics*. Department of Computer Science, Athens University of Economics and Business, Athens.

Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Wang, W. Y. and Yang, D. (2015). That's so annoying!!!: A lexical and frame-semantic embedding based data augmentation approach to automatic categorization of annoying behaviors using# petpeeve tweets. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2557–2563.

Zeng, D., Liu, K., Lai, S., Zhou, G., and Zhao, J. (2014). Relation classification via convolutional deep neural network. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics*, pages 2335–2344.

Zhang, D. and Yang, Z. (2018). Word embedding perturbation for sentence classification. *arXiv preprint arXiv:1804.08166*.

Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., and Xu, B. (2016). Attention-based bidirectional long short-term memory networks for relation classification. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 207–212.

Smart Health Monitoring and Management Using Internet of Things, Artificial Intelligence with Cloud Based Processing

Prateek Kaushik, Student

Dept. of Computer Science & Engineering

G D GOENKA UNIVERSITY, Gurugram, India

(pratikkaushik14@gmail.com/prateek.kaushik@gdgu.org)

Abstract: This paper discuss the latest technology which is transforming Health industry in an exceptional way. Internet of things and Artificial Intelligence has transformed Health Industry in an exponential manner. Machine Learning techniques such as supervised machine learning techniques and unsupervised machine learning techniques, Deep Learning such as Auto-Encoder, Restricted Boltzmann Machine and Convolutional Neural Network, Electro-Cardiogram Sensors and Actuators. Such skyrocketing technology applications help monitoring personal health using wearable sensors which offers smart and reliable solutions. With exponential increase of using such technologies increases chances of getting diagnoses early and in real-time.

Keywords: Internet of things, Cloud Platforms, Artificial Intelligence, Supervised machine learning techniques, Unsupervised machine learning techniques, Deep learning, Convolutional Neural Networks, Sensors and Actuators, Microcontrollers and Electro-Cardiogram.

1. INTRODUCTION

According to World Health Organization, it is estimated that average life expectancy has increased by 5% during 2011 till 2015. Although, more than 16,000 children under age 5 died every day because of not having proper access to medication and diagnosis. As per WHO report, more than 65% people across globe are overweight and suffer from obesity. More than 200 million women in the world do not have access to proper Health care services.

Primary reason for using such technology is to reduce traditional way of diagnosing patients which take several number of long hours and Healthcare professional has to be present with the patient to monitor the activities. To get rid of this, Patient-Oriented Approach is adopted. According to United Nation Report, it is predicted that older people population will increase by 2 billion in 2050 and the report also states that 89% of the aged people will live independently and out of this 89%, more than 65% people suffer from at least one chronic disease making it difficult for them to properly take care of themselves.

To trigger down Health problems is to remotely access the Health because it will give real-time information about Heart-Rate, Pulse Rate, ECG, High Blood Pressure, Low Blood Pressure Statistics in real-time because of wearable technology. Internet of Technology is now the world's most powerful communication paradigm because each object in our daily life has become part of the internet and Internet of Technology offers exceptional communication and computing capabilities.

2. INTERNET OF THINGS

Internet of Things based monitoring devices enables continuous and spontaneous monitoring of people such as obesity, Pulse Rate, Heart-Rate, Blood Sugar Level, Blood Pressure level and hypertension. In 1999, Kevin Ashton, co-founder of the Auto-ID Centre at the Science Institute of Technology, declared RFID chips that could easily enable 'objects' to communicate over the wireless network.

IoT consist physical devices which can communicate over a network based on wireless technology for example sensors and actuators. A sensor is a physical device which can receive and respond to a signal transmitted in the

form of motion, heat, light or chemical reaction. Once it detect the signal, it is able to convert it into analog signal or digital signal. Similarly, a transducer converts one form of energy to another form of energy but transducer do not provide the quantifiable data of conversion whereas sensors converts one form of energy into another form and also provide quantifiable data to measure. Some common examples of sensors are Temperature Sensor, Pulse Rate Sensor, Heart-Rate Sensor, Proximity Sensor and Humidity Sensor. Whereas Actuators are capable of converting one form of energy into mechanical energy or in basic terms it is regarded as 'mover'. It requires electric current, hydraulic fluid pressure or pneumatic pressure which are used to convert any form of signal into mechanical energy.

The most important parameters to assess Health of quality of any patient relies on Blood Pressure Level, Pulse Rate, Heart Beat Rate, Sugar Level and ECG Rate. To monitor all the parameters remotely, doctor will likely to take help of wireless technology such as Wireless Sensor Networks (WSN). It is a group of dedicated sensors for collecting the data and organizing the data and storing them at central location. ZigBee is a type of WSN Technology which is commonly used for Real-time Health Monitoring System.

2.1 PHYSICAL SENSORS

Physical sensors such as Pulse Oximeter Sensor which is used to measure the amount of oxygen dissolved in patient's blood on the basis of Haemoglobin and Deoxyhaemoglobin which are useful in the situation when the patient's oxygen level is unstable. Each Pulse Oximeter Sensor consist two light emitting diode in which one emits red light and other emits infrared light. It also consist of photo-detector which measures the intensity of transmitted light at each wavelength and by reading the differences calculated of the blood oxygen, the probe is then placed on a suitable part of the body to work properly for example: fingertip or ear lobe.[1]

2.1.1 METHOD FOR MONITORING OXYGEN SATURATION IN BLOOD USING PULSE OXIMETER SENSOR

- **TRANSMISSION METHOD:** The light emitting diode, transmitter, receiver and the photo-detector are placed on opposite side of the finger. During this method, most suitable part of the body, usually a fingertip is placed between the LED's and the Photo-Detector. When the fingertip is placed in between, it absorbs light and some part will be absorbed by the photo detector. Now with each heart-beat, the volume of blood flow will increase and in result more light will be absorbed by the fingertip and less light will reach to photo detector.

  Hence, if the graph is plotted of the received light signal, it will consist of peak in between heart beats and trough at each heartbeat. The difference between peak value and trough (bottom) value depicts the blood flow at heart beat.

- **REFLECTANCE METHOD:** The light emitting diode and photo detector are placed on the same side, next to each-other. In this method, fixed light reflection back to the sensor due to fingertip placed on it. More reflection of light will occur if the volume of blood flow increases gradually and reflection will go back to the sensor.
  Hence, if the graph is plotted of the received light signal, it will consist of peaks at each heartbeat and constant reflection will be present in the form of fixed low volume. Difference of the two will result in reflection value due to blood flow at heart beat.

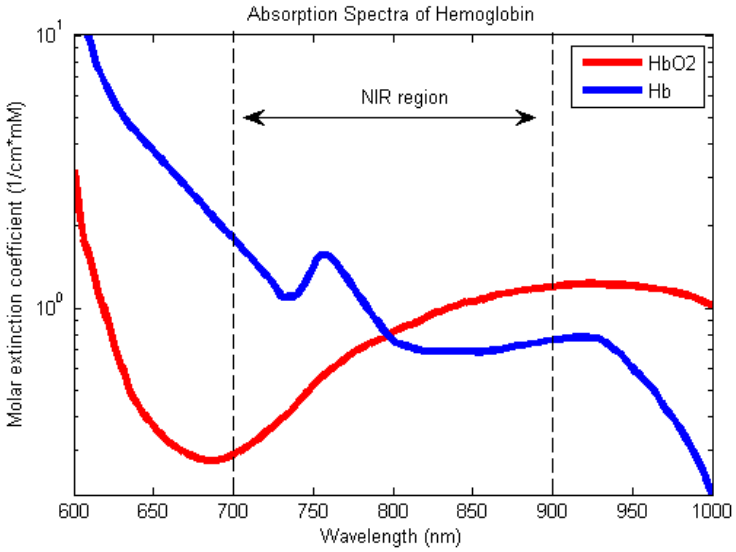


Fig.1 Pulse Oximeter Sensor Graph

The oxygen content can then be calculated easily by comparing the how much red light is absorbed as compared to infra-red light.

2.1.2 PIR SENSOR: Another type of sensor used for measuring the infrared light is called as PIR Sensor. PIR Sensor or Pyroelectric Sensor works in the view in which every object that has a temperature above absolute zero emit heat energy. That heat energy is emitted in the form of radiation radiating at infrared wavelengths which is invisible to the human eye but can be detected using PIR Motion Detectors. [2]

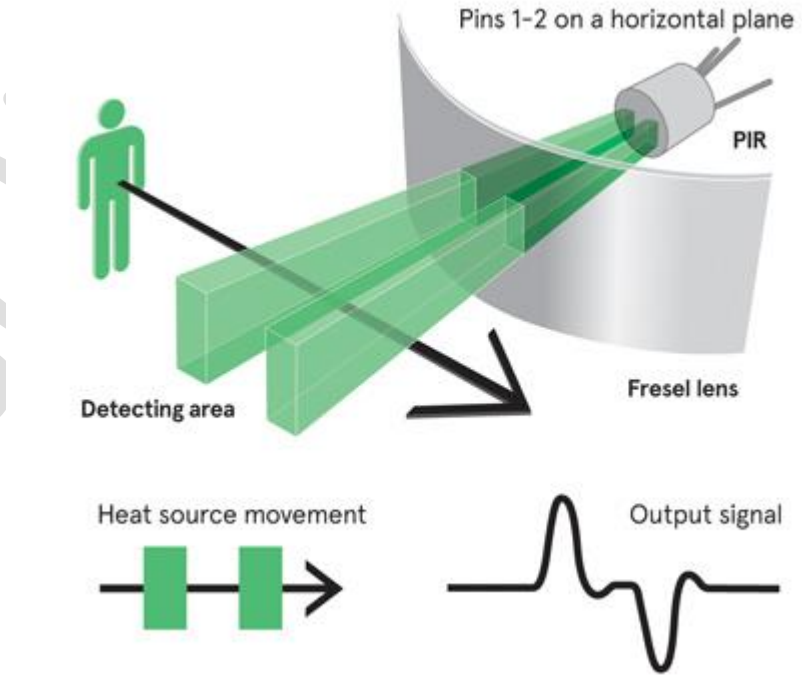


Fig.2 Adapting PIR Sensor Technology

2.1.3 ULTRASONIC SENSOR: This type of sensor is similar to sonar and radar in which radio emits echoes or sound waves emitted by radio or by electronic device to evaluate the attributes of a target by generating high frequency sound waves.

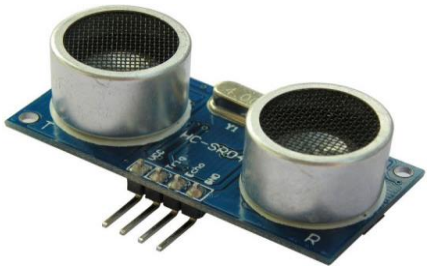


Fig.3 Ultrasonic Sensor Module HC-SR04

2.1.4 TEMPERATURE SENSOR: It collects the data about temperature from a particular source and converts the data into understandable form for a device. There are following types of temperature sensors are available:

2.1.4.1 THERMOCOUPLES: It is low-cost, self-powered and can be used for long distance. It detects the temperature by measuring change in voltage. It consist of two metals, one is open and second is closed. Metals work on Thermo-Electric Effect Principle, when two unequal metals produce a voltage, then a thermal difference between the two metals is calculated. Thus, the temperature goes up and the output voltage of thermocouple increases significantly.

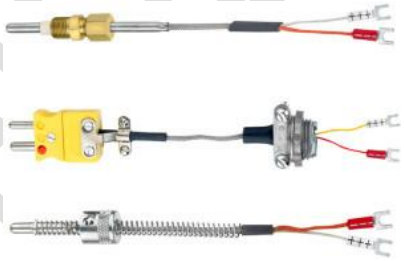


Fig.4 Thermocouple T-50230 -001

2.1.4.2 RESISTOR TEMPERATURE DETECTOR: Also known as RTD sensor which is one of the most accurate sensor in which a resistor temperature detector, the resistance is proportional to the temperature. RTD sensor based on different method such as two-wired, three-wired and four-wired method.

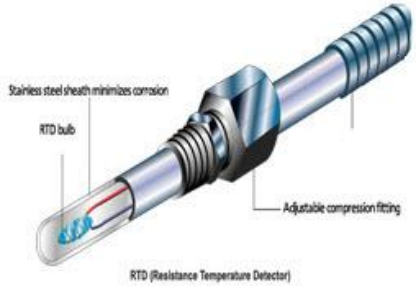


Fig.5 Resistor Temperature Detector

2.1.4.3 **THERMISTORS:** It is low-cost, adaptive to environment and is easy to use. It is capable of changing the resistance when the temperature changes. It consist of negative temperature coefficient which means the resistance get decreased if temperature increases.

Fig.5 Thermistors

2.1.4.4 **IR SENSOR:** It sense temperature by emitting Infrared radiation. It is a type of non-contacting sensors for example by holding an IR Sensor in front of the desk, it detects the temperature of the desk by radiation. This type of sensors are classified as of two types one is thermal infrared sensors and second is infrared sensors.

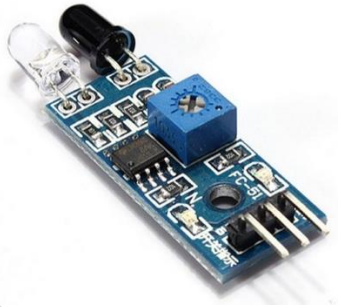


Fig.6 Infrared Radiation Sensor

2.4 COMMUNICATION

Internet of Things consist of smart devices which are able to communicate with each other. It enables them to exchange data over the wireless network on the basis of communication protocols such as BLE stands for Bluetooth Low Energy, IPv6 Internet Protocol Version 6, Low Power Wireless Personal Area Networks 6LoWPAN, ZigBee, Z-Wave, Near-field Communication and RFID Radio Frequency Identification.

2.4.1 **BLE:** It is known as Bluetooth Low Energy device which allows devices to communicate without cables while providing high level of security. It is of low-cost and consumes less power for operating. It follows a process known as pairing to exchange the information. During the process, the Bluetooth enabled device can connect with other devices located in close proximity. It uses the Adaptive Frequency Hopping Technology which enables BLE to achieve robust transmission in the noisy environment. BLE technology has successfully reduced the number of channels to 40 2-Mhz wide channels instead of 79 1-Mhz wide channels.

2.4.2 **IPv6:** Internet Protocol Version 6 offers highly scalable address scheme of 128 bits whereas IPv4 offers only 32 bits address scheme. IPv6 is more secure because it can run end-to-end encryption. The integrity-

checking and 128bit encryption used for standard component in virtual private network, available for all connections and supported by compatible devices and systems.

2.4.3 **6LoWPAN:** It is the name of internet area IETF. The concept of it is to process the capabilities of low power devices to be able to participate in Internet of Things. It follows encapsulation method for transmitting data over the Personal Area Network IEEE 802.15.4 based networks.

2.4.4 **ZigBee:** ZigBee is interoperable and standardize network and application layers in which the devices belonging to different owners can connect. The ideal operation of ZigBee carried out in Home Automation and Smart Energy because of different ZigBee devices can be connected. As the number of ZigBee devices increases, communication paths between devices multiply and eliminate the risk of Single-Point Signal Failure.

2.4.5 **Z-Wave:** Z-Wave is a wireless communication protocol used for Home Automation. It consist of two-way communication using mesh topology, which helps in low-cost wireless connectivity to Home Automation. It consist of Smart Home Hub which is a controller connected to the internet. Whenever a Z-Wave smart controller receives command from Smart Home Application on a digital device, it provides the routing of data across the network of connected devices.

2.4.6 **NFC:** Near Field Communication is a short-range wireless connectivity standard which uses magnetic field lines to transmit the data or to make communication between two electronic devices for exchanging the information. It is based on Peer-to-Peer Approach in which the devices have to be in contact or near about 4cms to make the connection and to transfer the data. It consist of three models such as NFC Reader/Writer, NFC Peer-to-Peer and NFC Card Emulation.

2.4.7 **RFID:** Radio Frequency Identification, an automatic technology to identify objects without the line of sight and able to record data. RFID consists of a reading device called a reader which is a powerful device with ample memory and computational resources. It consist two types of tag in which first tag is Passive tag in which they have limited computational capacity, are able to detect collisions and are able to make communication between two devices feasible. Second tag is Active tag which is able to sense the channel and is also able to detect collisions.

2.5 WIRELESS TRANSCIEVERS

A wireless transceiver works as transmitter and receiver in which an electronic switch allows the transmitter and receiver to be connected to the same network to prevent the transmitter damaging the receiver. Wireless Transceivers are Bluetooth Low Energy Device, Wi-Fi and iBeacons.

2.6 MICROCONTROLLER

It is a printed circuit board (PCB) with circuitry and hardware board features. It includes Bus type, Memory, Port type, Port Number, Processor Type and Operating System.

2.6.1 **Arduino UNO:** It is the most popular open source electronic prototyping platform for creating electronic applications. It consist of ATmega328P microcontroller, 32KB of Flash memory, 5V Operating voltage, Input voltage, Output voltage, Digital I/O Pins, Analog Input Pins, DC Current per I/O and DC Current for 3.3V Pin.

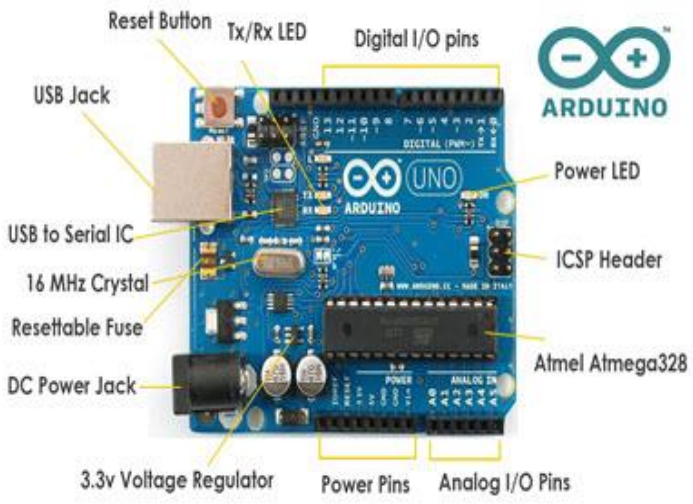


Fig.7 Arduino UNO

2.6.2 **Raspberry Pi Development Board:** It is a low-cost development board and it is of credit-card sized computer that can be plugged into TV or monitor and uses a standard mouse and keyboard. It can be used in digital projects maker and at weather stations. It consist of 1.2GHz, 64bit quad-core Processor, 802.11n Wireless LAN, Bluetooth 4.1, Bluetooth Low Energy, 1GB Ram, 4USN Ports, 40GPIO Pins, Full HDMI Port, Combined 3.5mm audio jack and composite video, Camera interface, Display interface, Micro SD Card slot and VideoCore IV 3D graphics core.

Fig.8 Raspberry Pi Development Board

2.6.3 **BeagleBoon Black Development Board:** It is one of the most popular open source computer and it comes with built-in wireless networking capability. It is easy to use, low-cost and community supported development platform for embedded application developers. It takes upto 5 minutes after booting in Linux to start the development using a single USB cable. It consist of AM335x 1GHz ARM Cortex-A8 Processor, 512 MB DDR3 RAM, 2GB 8bit eMMC on board flash storage, NEON floating-point accelerator, 2x PRU 32-bit microcontroller and 3D Graphics accelerator.

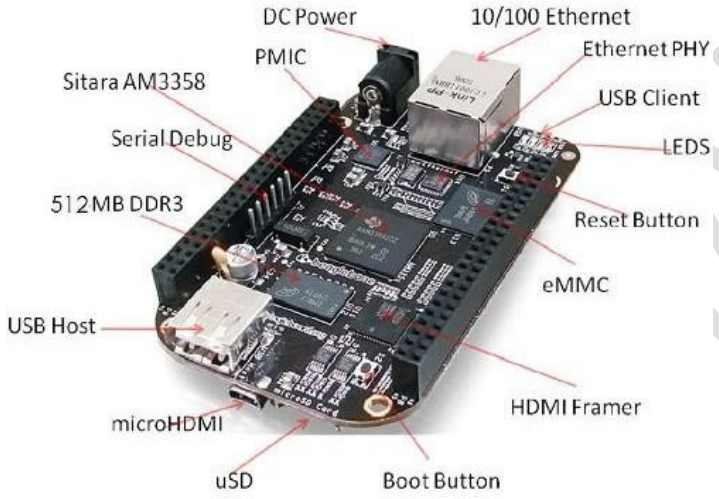


Fig.9 Beagle bone Black Revision C

2.7 Cloud Platforms

Internet of Things consist of various number of cloud platforms for storing data, pre-processing data and analyse the data. It takes the massive amount of data generated through sensors, actuators, applications and store on cloud for initializing the real time access. [4]

2.7.1 **Thingworx:** It is one of the leading Internet of things platform used for Healthcare Industry for storing large amount of data captured through various number of sensors and is stored on cloud. It offers easy connectivity with electronic devices such as RFID and sensors, pre-built gadgets for the dashboard, removes complexity and integrated machine learning.

2.7.2 Microsoft Azure Internet of things Suite: It provides multiple services to Internet of things solutions. It expands profitability and productivity with pre-built connected solutions. It analyses large amounts of processed data. It consist of easy device registration, integrated software like SAP, Oracle and etc., dashboards, visualization tools and real time streaming.

2.7.3 Google Cloud Internet of things Platform: It is among the best platform built for Internet of things to easily connect, store and manage large amount of data captured through sensors. It provides large amount of storage, cuts cost for server maintenance, business through fully protected, intelligent and responsive Internet of things data, capable of analysing large amount of data, efficient and scalable.

2.7.4 IBM Watson Internet of things Platform: It is a powerful platform powered by IBM Bluemix and hybrid cloud PaaS known as Platform as a Service development platform. It provides real-time data exchange, secure communication, cognitive systems and also provides weather data services.

2.7.5 AWM Internet of things Platform: Amazon also provided cloud services for Internet of things applications providing various features like device management, secure gateway for devices, authentication and encryption and device shadow.

2.7.6 Cisco Internet of things Cloud Connect: It helped digital transformation and provides mobility, cloud based suits. It provides easy deployment options for electronic devices, data and voice connectivity, device and IP session report, Billing is customizable and flexible deployment options.

2.7.7 Oracle Internet of things Platform: It provides real-time Internet of Things data analysis which is captured using various sensors. Most important feature is, it provide high-speed messaging in real-time. It is PaaS (Platform as a Service), cloud based platform. It is secure, scalable, provides real-time insights, integrated and provides faster market solutions.

3. ARTIFICIAL INTELLIGENCE

AI roots are found in 1300 CE when Roman Llull's theory of reasoning machine proposed but it became public interest in late 1950s when China invested billions of dollars for the development of AI. Since then other nations also started investing in AI. Now AI is helping Health Industry in every aspect such as real-time monitoring of Pulse rate, Heart beat rate, ECG rate and Blood pressure level etc. AI can be defined as "*the simulation of human intelligence by learning, reasoning and self-correction by machines such as computer systems."*

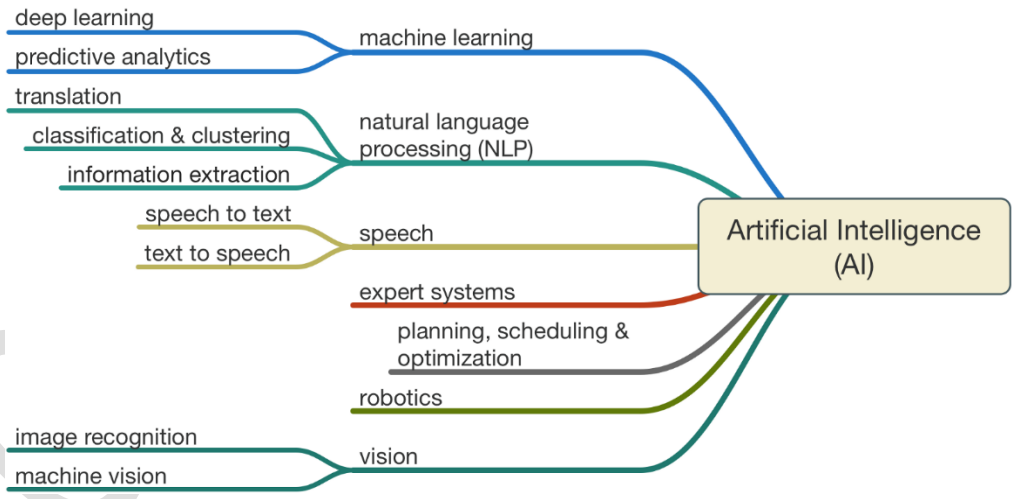


Fig.10 Artificial Intelligence(AI)

AI main aim is to replicate human intelligence which is accomplished by learning and reasoning which is termed as machine learning. Machine learning is the modelling of different situations and improve through experience. Unequalled amount of data is generated through various number of sensors around the world which has made machine learning most essential technology to be used in modern world today. Machine learning techniques are able to analyse hidden insights in the data. Working principle of machine learning is to supply more amount of data and more accurate hidden insights can be analysed. Machine learning is composed of two learning algorithms Supervised Machine Learning Algorithm and Unsupervised Machine Learning Algorithm.

3.1 SUPERVISED MACHINE LEARNING ALGORITHM

The supervised machine learning algorithms are able to classify objects on the basis of their label.

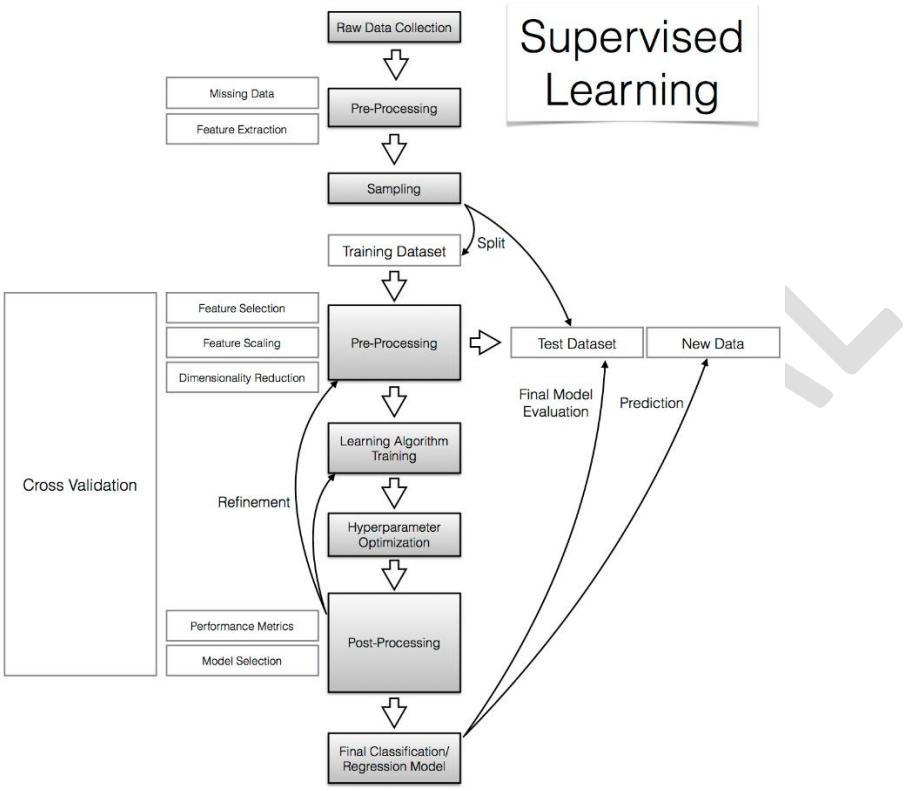


Fig.11 Supervised Machine Learning Algorithm

It composed of several classification algorithm and these are as follows

3.1.1 **LINEAR CLASSIFIER:** The goal of supervised machine learning linear classifier is to group similar objects which share similar features values. It is done by combining the features of the objects linearly to find similar characteristics and hence it is the fastest classification classifier. It can work with large number of dimensions commonly known as columns in real world datasets.

3.1.2 **LOGISTIC REGRESSION:** This type of classifier is used to predict to outcomes either true or false. It uses single estimator to process the multinomial logistic regression model. It is able to make outcomes as more detailed and precise. Logistic regression is more suited to define the relationship between categorical variable rather than numerical variable. The major assumptions carried out in logistic regression are is that a variable should be dichotomous in nature, there should be no outliers in the data and there should not be very high multicollinearity between the dependent and independent variables.

3.1.3 **NAIVE BAYESIAN NETWORK:** Naïve Bayes classifier consist of two nodes. Observed node and unobserved node in which observed node act as children and unobserved node act as parent. It is mostly carried out for large-scale computation, decision-tree induction, and standard benchmark datasets. The main disadvantage of using Naïve Bayesian algorithm is the independency problem which is solved using Averaged one-dependency estimator.

3.1.4 **MULTILAYER PERCEPTRON:** It is a type of neural network in which at least three nodes are present such as input layer, hidden layer and output layer. It composed of quadratic equation for solving the convex and non-convex, unconstrained minimization neural network problems.

Perceptron classifier is used for learning from a batch of dataset composed of training set and test set after train test split suing scikit library.

3.1.5 **SUPPORT VECTOR MACHINE:** Support vector machines or SVM is closely similar to multilayer perceptron neural network in which the data classes is split into halves or equal ratios by creating the largest possible distance between the hyperplane and the instances.

3.1.6 **DECISION TREES:** Decision trees is a type of supervised machine learning algorithm which classify objects on the basis of features values and sorts them. It consist several nodes in which one node represents feature or an instance and each branch represents a value that instance is carrying or that can be assumed. It is generally used in data mining techniques for predictive modelling in which mapping of objects variable or instances is done to predict the target value. It can also be called as regression trees because they are used to predict continuous value or target variables

3.1.7 **RANDOM FOREST ALGORITHM:** It merges several decision trees and the combination of decision trees yields more accurate answer. The most important advantage of using random forest algorithm is that it can be applied to classification problems and regression problems simultaneously. Unlike decision tree model which look for the most important feature among all the features by splitting the data, random forest algorithm finds the best feature among all random features.

3.2 UNSUPERVISED MACHINE LEARNING ALGORITHM

Unlike supervised machine learning algorithm in which labels are present to predict the outcome but in unsupervised machine learning algorithm there is no label present for example finding the topics in a document which can appear in the topics discussed in each document, so the task is to extract topics. Unsupervised machine learning algorithm consist of several algorithms discussed below. [6]

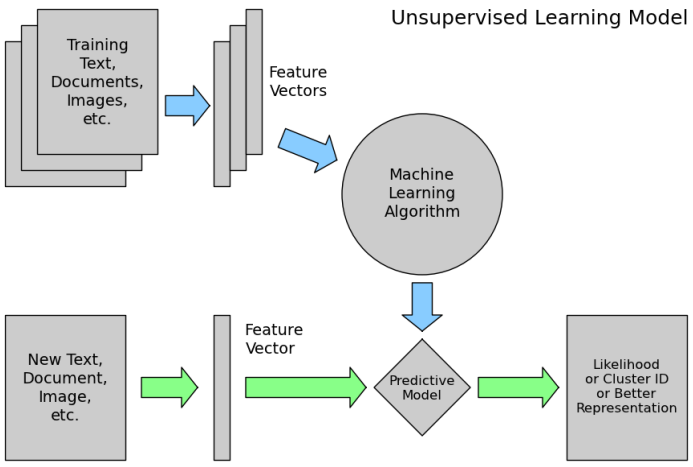


Fig.12 Unsupervised Machine Learning Algorithm

3.2.1 **PRE-PROCESSING AND SCALING:** Pre-processing means adjusting the features in a way so that the data visualization is more suitable for algorithms. It can be done by importing "mglearn" library for pre-processing the data. Types of pre-processing discussed below.

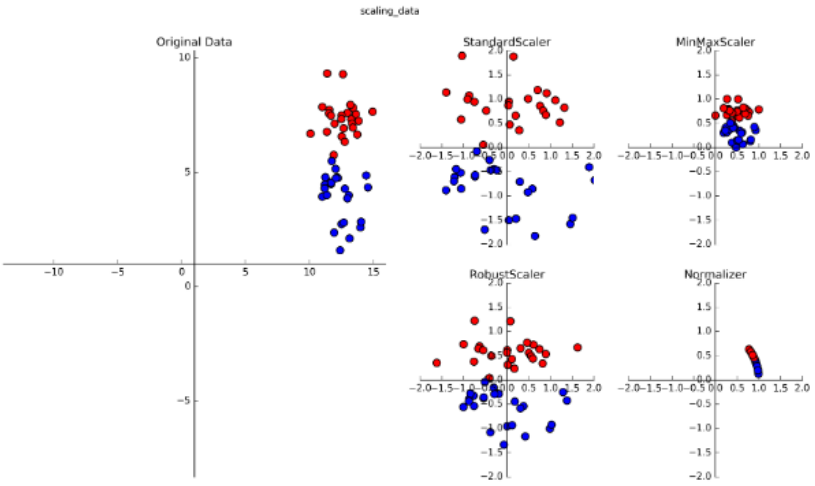


Fig.13 Pre-Processing and Scaling

3.2.1.1 **STANDARD SCALER:** It ensures that the mean of each feature is zero and variance is one for bringing all the features at same magnitude.

3.2.1.2 **ROBUST SCALER:** It also ensures that the features are at the same magnitude. Additionally, it uses quartiles and median values instead of using mean and variance which often ignore noisy values regarded as outliers which disrupt the accuracy of the algorithm.

3.2.1.3 **MIN-MAX SCALER:** It shifts the all the features between binary values 0 and 1. Therefore, plotting a 2Dimensional graph indicating d ata contained in a rectangle like shape between 0 and 1.

3.2.1.4 **NORMALIZER:** It is a type of scaler which is completely different from other scaling techniques because it uses Euclidean distance in which each feature should have Euclidean length of one. This type of normalization is used only the direction of data variables is significant rather than the length of feature vector.

3.2.2 **DIMENSIONALITY REDUCTION:** In real world, datasets contain huge number of dimensions which are difficult and costlier to process for analyses. To reduce the unnecessary dimensions, dimensionality reduction term is coined in which following methods are used.

3.2.2.1 **PRINCIPLE COMPONENT ANALYSIS:** In this method is used, dataset is rotated in such a way so that features are statistically uncorrelated. The rotation is done on the basis of selecting subset of features which are able to define the characteristics of the dataset. By importing "mglearn" library, Principle component analyses can be perform using python command.

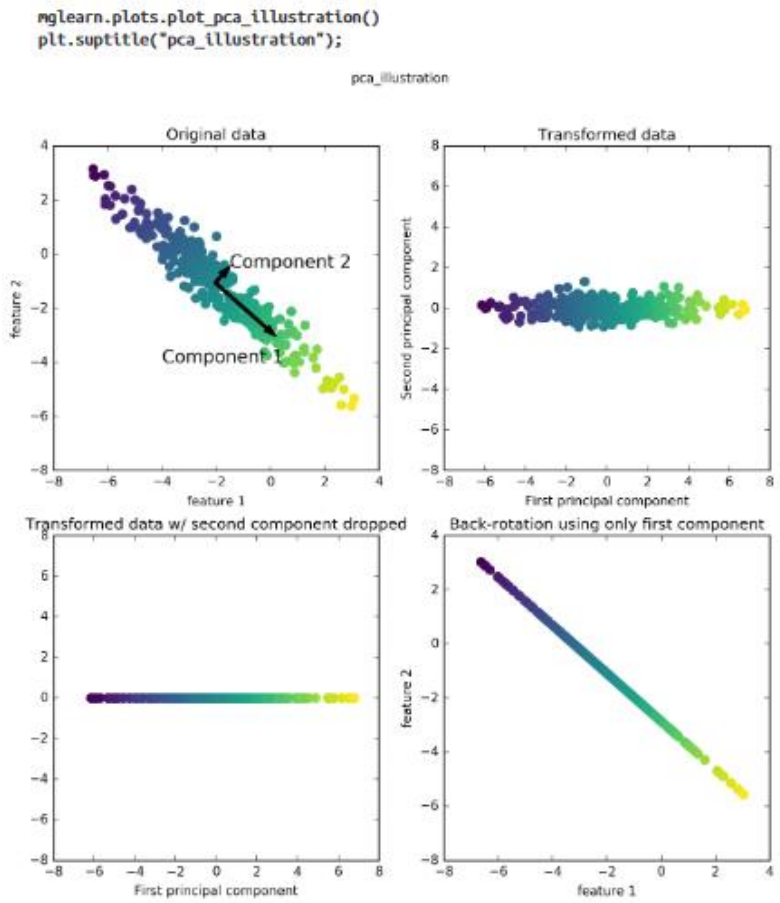


Fig.14 Principle Component Analysis

3.2.2.2 **NON-NEGATIVE MATRIX FACTORIZATION(NMF):** It aims to extract useful features using similar techniques like PCA, unlike PCA in which each data point is a weighted sum of orthogonal components and shows variance of each orthogonal component, in NMF the coefficients and components should be equal to zero or greater than zero that means they should not be negative at all. Due to which, NMF is able to compute sum of non-negative weighted components to identify the original components which makes up the whole dataset. Hence it provides for interpretable results than PCA.

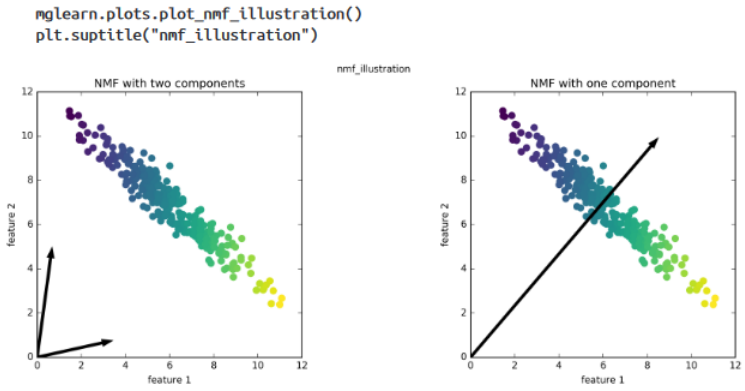


Fig.15 Non-Negative Matrix Factorization (NMF)

3.2.2.3 **MANIFOLD LEARNING WITH t-SNE:** It is a complex method used for providing better scatter plot visualization. Manifold learning t-SNE algorithm used to compute a new representation of the training data. The only exception is it cannot replicate the test set for visualization because it is able to replicate or transform the training set. It starts with a random 2Dimensional representation for each data point and then align them at best possible distance such that closer points should be close enough to closer points in original space and farther points should be farther from farther points in original space.

3.2.3 **CLUSTERING**

Clustering is a type of unsupervised machine learning technique in which partition of dataset is done in clusters which share similar characteristics. There are following types of clustering techniques available

3.2.3.1 **K-MEANS CLUSTERING:** It is one of the simplest algorithm used for clustering data. The working of this algorithm is to assign each data point to the closest cluster and then setting each cluster as them mean of the data points that are assigned to it.

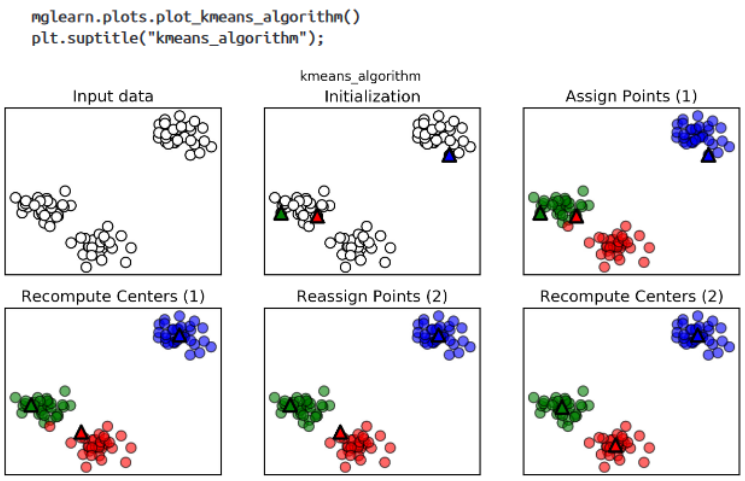


Fig.16 K-Means Clustering

3.2.3.2 **AGGLOMERATIVE CLUSTERING:** Its working rule is similar to KMeans clustering but the only difference is that it merges two similar kind of clusters until the stopping criteria is met which is defined by scikit-learn package.

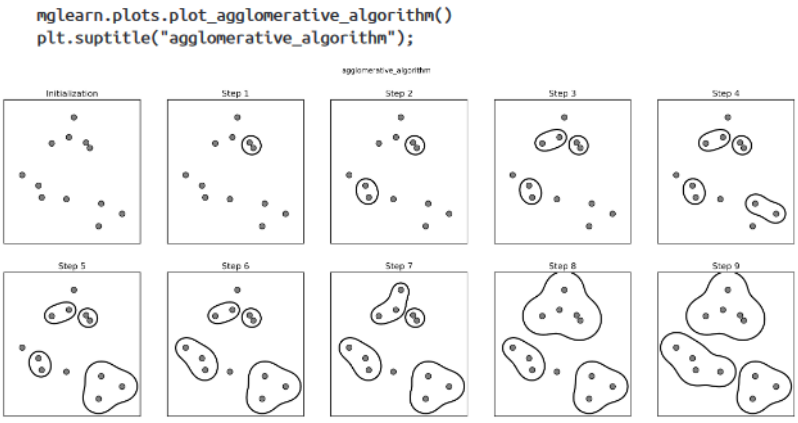


Fig.17 Agglomerative Clustering

3.2.3.3 **DBSCAN:** It is known as "Density based spatial clustering of applications with noise". In all the previous algorithms, it is require to set the parameter number of clusters but in DBSCAN it is not required because it is able to automatically detect the complex clusters shape. It works best in crowded regions of the feature space in which all data points are close to each other.

```
dbscan = DBSCAN()
clusters = dbscan.fit_predict(X_scaled)
# plot the cluster assignments
plt.scatter(X_scaled[:, 0], X_scaled[:, 1], c=clusters, cmap=mglearn.cm2, s=60)
```

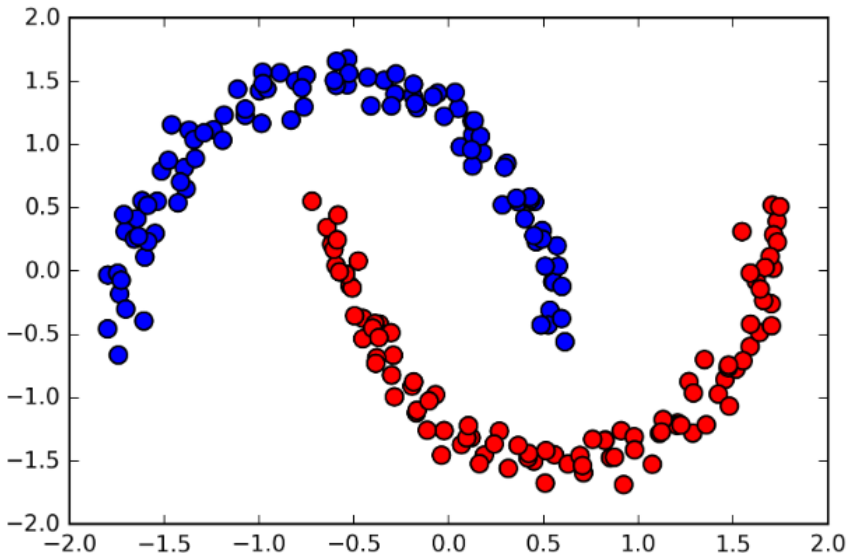


Fig.18 DBSCAN

3.3 **DEEP LEARNING FOR MEDICAL SENSORY DATA:** Deep learning is the fastest growing technology of Artificial Intelligence because of its ease of object recognition and speech recognition. Deep learning is gaining popularity because nowadays deployment of low-cost sensors and their connectivity through LAN, WAN and MAN networks. Deep learning models are of following types discussed below.

3.3.1 **AUTO-ENCODER:** It consist of two phases, encoder and decoder which are defined to learn a new representation of the data by trying to reconstruct the input data. Encoder consist of input **x** and hidden transformation **h** using non-linear mapping.

$$\mathbf{h} = \varphi(\mathbf{W}\mathbf{x} + \mathbf{b})$$

Then decoder uses the same formula for decoding the representation to its original state.

$$\mathbf{z} = \varphi(\mathbf{W}'\mathbf{h} + \mathbf{b}')$$

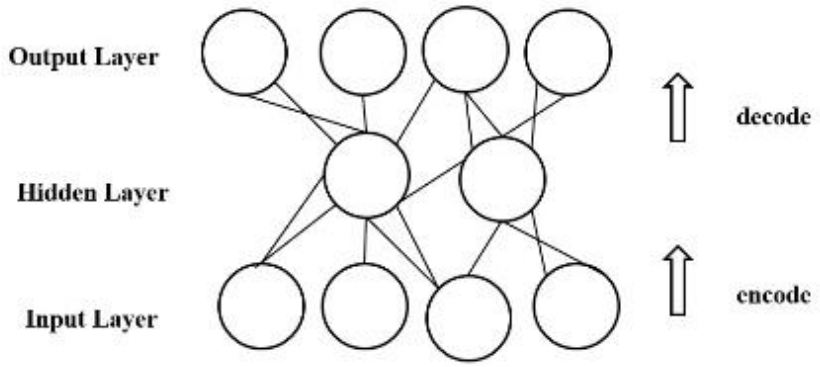


Fig.19 Auto-Encoder

3.3.2 RESTRICTED BOLTZMANN MACHINE: RBM is a two-layer neural network which forms a bipartite graph having symmetric connection in which it is composed of two working units such as visible unit **v** and hidden units **h.**

$$E(\mathbf{v}, \mathbf{h}; \theta) = -\sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij} v_i h_j - \sum_{i=1}^{I} b_i v_i - \sum_{j=1}^{J} a_j h_j$$

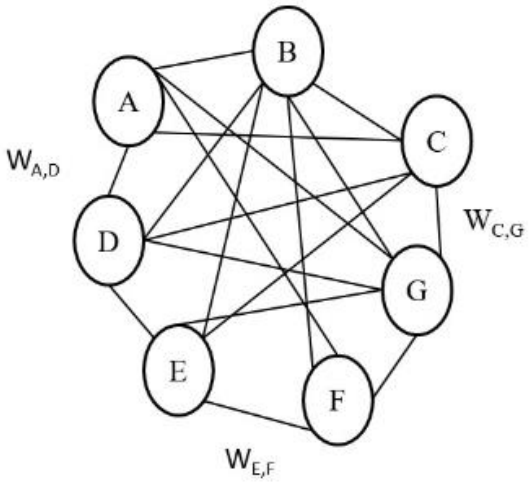


Fig.20 Restricted Boltzmann Machine

3.3.3 CONVOLUTIONAL NEURAL NETWORK: It was proposed by LeCun for image processing having two main properties which are spatially shared weights and spatial pooling. CNN algorithms accepts 2Dimensional input data for computer vision applications.

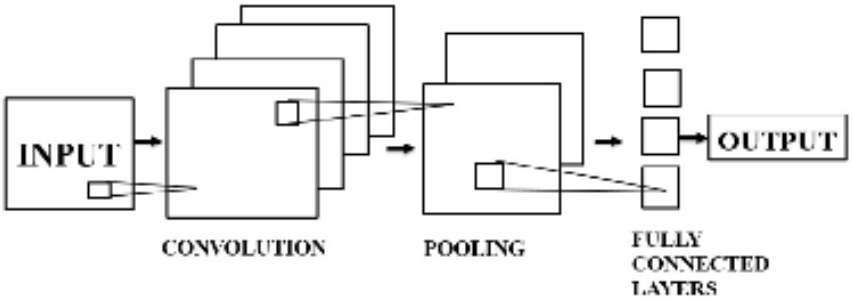


Fig.21 Convolutional Neural Network

3.3.4 MAJOR DEEP LEARNING METHODS NOTATIONS:

Notations	Definition
x	Samples
y	Outputs
v	Visible vector
h	Hidden vector
q	State vector
W	Matrix of weight vectors
M	Total number of units for the hidden layer
w_{ij}	Weights vector between hidden unit h_j and visible unit v_i
S_i	Binary state of a vector
s_i^q	Binary state assigned to unit i by state vector q
Z	Partition factor
d_j	Biased weights for the j-th hidden units
c_i	Biased weights for the i-th visible units
z_i	Total i-th inputs
v_i	Visible unit i
w_{kj}^2	Weight vector from the k-th unit in the hidden Layer 2 to the j-th output unit
w_{ji}^1	Weight vector from the j-th unit in the hidden Layer 1 to the i-th output unit
W_{ji}^1	Matrix of weights from the j-th unit in the hidden Layer 1 to the i-th output unit
$E(q)$	Energy of a state vector q
σ	activation function
$P_r(q)$	Probability of a state vector q
$E(v,h)$	Energy function with respect to visible and hidden units
$pdf(v,h)$	Probability distribution with respect to visible and hidden units

Table 1. Major Deep Learning Notations

4. **ECG**

ECG or EKG both stands for electrocardiogram discovered by Willen Einthoven when he was working in Leiden, used a string galvanometer which was invented in 1901. He found that it is more sensitive than capillary electrometer invented in 1897 by French engineer Clement Ader. The heart is composed of four chambers commonly known as two upper chambers and two lower chambers. The upper chambers contains left atria and right atria and the lower chambers contains left ventricles and right ventricles. The purpose of atria is to receive blood from the body and the purpose of right atrium is to receive oxygen-devoid blood from the body and left atrium receives oxygen-rich blood from the lungs.

The right pump receives the blood from the body and supply it to the lungs. Thus heart is controlled via electrical system in which electrical impulse creates a region in which diffusion of calcium ions, sodium ions and potassium ions across the membrane of cells occurs. The impulse is so strong that it transfers ions to atria by which it get squeezed and pump blood from the heart. Hence, the body system regulates this motion so that the entire cardiovascular system can work properly.

Then comes ECG device which detects the small electrical ions on the skin and amplifier amplifies them when the heart depolarizes at each heartbeat which causes each heart muscle to have a negative charge known as membrane potential.

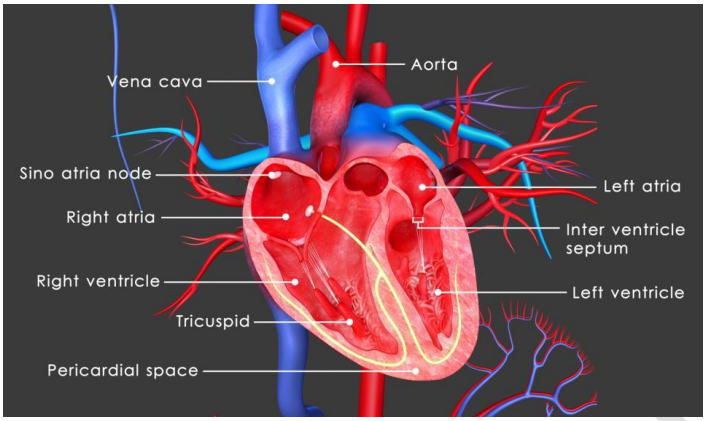


Fig.22 Example Structure of the Heart from Franklin Institute

4.1 **ECG WORKING PRINCIPLE:** When the depolarization happens at each heartbeat in which the negatively charged ions are depolarized by influxing positive charged ions in which each heart muscle gets depolarized. [6]

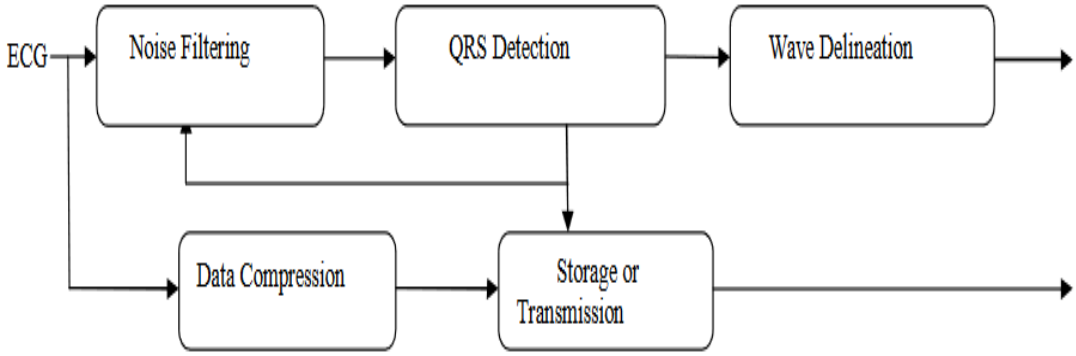


Fig.23 ECG Working Principle

The whole process is carried out when each small ion rises and fall in the polarization and depolarization process in which it strikes with the voltage between two electrodes placed at either side of the heart which shows wave on a screen.

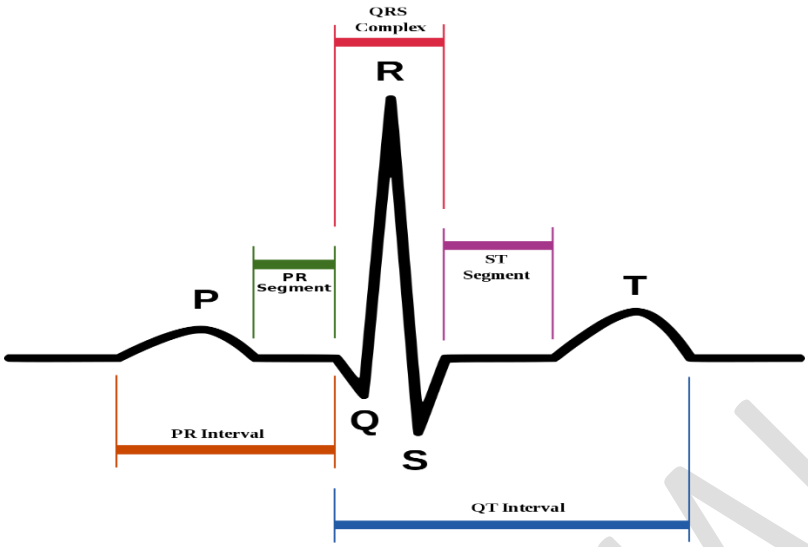


Fig.24 Example of ECG Device Important Parameters

4.2 ECG MEDICAL DEVICE IMPORTANT PARAMETERS: It consist of RR interval, P-wave, PR interval, PR segment, QRS complex, ST segment, ST interval, QT interval explained below.

4.2.1 **RR INTERVAL:** When the heart rate is between 60bpm and 100bpm, the interval is known as current R wave and the next R wave.

4.2.2 **P-WAVE:** When the depolarization is carried out at each heartbeat, an electrical vector is directed from SA node to the AV node and goes from right atrium to the left atrium which known as P-wave and its time is 80 milliseconds.

4.2.3 **PR-SEGMENT:** The PR segment connects P wave and QRS complex when the electrical activity goes down towards the ventricles which shows flat wave on ECG screen. Its duration is 50 milliseconds to 120 milliseconds.

4.2.4 **PR-INTERVAL:** It is the time taken between the P-wave and and QRS complex in which the time taken by electrical impulse to travel from sinus node through AV node to ventricles. Its duration is 120 milliseconds to 200 millisecods.

4.2.5 **QRS-COMPLEX:** It is the time taken of depolarization at left and right ventricles. It has higher amplitude wave than P-wave. Its duration is 80 milliseconds to 120 milliseconds.

4.2.6 **ST-INTERVAL:** It is the time taken from the J point to the end of T-wave. Its duration is 320 milliseconds.

4.2.7 **ST-SEGMENT:** It connects with QRS complex and T-wave where ST-Segment represents the depolarization at ventricles. Its duration is 80 milliseconds to 120 milliseconds.

4.2.8 **QT-INTERVAL:** The total time taken the simulation of QRS complex to T-wave. If the QT interval is prolonged then sudden death take place. Its duration is 420 milliseconds.

5. CONCLUSION

This research paper described a personalized smart health monitoring device using wireless sensors and latest technology. We are able to detect anomaly in real-time if the patient is in danger. Using ECG and other sensor real-time health monitoring can be done even if the distance between the doctor and patient is thousands of miles away using cloud server as discussed above which increases the chances of getting diagnosis at early stage. Due to the evergrowing population, life expectancy has decreased by a great factor because of lack in resources and emergency aid that should be provided when the patient is in danger.

Smart health monitor system can be used for indoor and outdoor pupose and by using latest algorithms discussed above will greatly improve the efficiency to the existing health monitor systems. By this, the accurate and precise measurement of patient's health parameter can be done inn real-time. In this paper, machine learning and deep learning techniques are discussed which works as a catalyst to improve the perfomance of any health montior system such supervised machine learning algorithms, unsupervised machine learning algorithms, auto-encoder, convolutional neural network and restricted boltzmann machine. Hence, AI and IoT have succesfully solved the problems related to patient's health with outcomes comparable to that human clinicians.

REFERENCES

[1] LISHA YU1, WAI MAN CHAN2, YANG ZHAO3, AND KWOK-LEUNG TSU, *"Personalized Health Monitoring System of Elderly Wellness at the Community Level in Hong Kong"*, Department of Systems Engineering and Engineering Management, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong3Center for Systems Informatics Engineering, City University of Hong Kong, Hong Kong

[2] Valerie GAY , Peter LEIJDEKKERS, *"Health Monitoring System Using Smart Phones and Wearable Sensors"*, Faculty of IT, University of Technology Sydney, PO Box 123, Broadway 2007 NSW Australia.

[3] Dr. Deepak Choudhary, Prof. Rakesh Kumar Director, Ms. Neeru Gupta Asst*., "Real-Time Health Monitoring System on WirelessSensor Network".* Associate Professor Galgoatia's, University, India, K.P.Jain engg. College, India, Prof, Mannav Bharti University, India,

[4] Moeen Hassanalieragh, Alex Page, Tolga Soyata, Gaurav Sharma, Mehmet Aktas, Gonzalo Mateos∗Burak Kantarci, Silvana Andreescu, *"Health Monitoring and Management Using Internet-of-Things (IoT) Sensing with Cloud-basedProcessing: Opportunities and Challenges"*, Dept. of ECE,†URMC Dept. of ECE, Chem & BioM Sci University of Rochester Clarkson University Rochester, NY 14627, Dept. of ECE, Chem & BioM Sci University of Rochester Clarkson Postdam, NY 13699.

[5] Nagdev Amruthnath, Tarun Gupta, *"A Research Study on Unsupervised Machine Learning Algorithms for Early Fault Detection in Predictive Maintenance",* Department of IEE and EDMM Western Michigan University Kalamazoo, Michigan, USA

[6] Anand Kumar Joshi, Arun Tomar, Mangesh Tomar, "*A Review Paper on Analysis of Electrocardiograph (ECG) Signal for the Detection of Arrhythmia Abnormalities*", PG Student[ICE], Dept. of Electrical Engineering, Assistant Professor, Dept. of Electrical Engineering. Maharana Pratap College of Technology Gwalior, Madhya Pradesh, India

ISSN 2319 – 1953
International Journal of Scientific Research in Computer Science Applications and Management Studies

Internet of Things with BIG DATA Analytics – A Survey

[1]A.Pavithra, [2]C.Anandhakumar, [3]V.Nithin Meenashisundharam

[1]Assistant Professor, [2]Student, [3]Student,
Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi.
[1]pavithra@stc.ac.in, [2]kumaranand16498@gmail.com, [3]nithinmsv@gmail.com

Abstract- **The INTERNET OF THINGS (IOT) is an internet among things through advanced communication protocols without human's operation. The main idea of IOT is to reach and improve the common goal by their intelligent networking. The IOT is an integrated technology of several sub technologies, such as wireless sensor or semantic and IOT is the way of connecting devices using sensors and monitored by Internet. But the data produced by the IOT is growing rapidly because of the large scale development of various applications. As the data is turned and crossed over terabytes and leading to petabytes, there should be a solution to manage the overwhelming increase in data. Big data is data that are too big (volume), too fast (velocity) and too diverse (variety). Big data is a collection of large data sets that include different types such as structured, unstructured and semi-structured data. Big data is the solution for the data problem and it is considered as the future's data dream. The upcoming IoT will be greatly presented by the enormous quantity of heterogeneous networked embedded devices that generate intensively "Big data". In this article we discuss about Big data on IoT and how it is interrelated to each other along with necessity of implementing Big data with IoT and its benefits, job market.**

Keywords: **Internet of Things, Big Data, Streaming, Spatial, Time Series, Prescriptive Analysis**

I. INTRODUCTION

1.1 Internet of Things

The Internet Of Things (IoT) is a computing concept that describes the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices. The term is closely identified with RFID as the method of communication, although it also may include other sensor technologies, wireless technologies or QR codes.

The IoT is significant because an object that can represent itself digitally becomes something greater than the object by itself. No longer does the object relate just to its user, but it is now connected to surrounding objects and database data. When many objects act in unison, they are known as having "ambient intelligence."

The internet of things is a difficult concept to define precisely. In fact, there are many different groups that have defined the term, although its initial use has been attributed to Kevin Ashton, an expert on digital innovation. Each definition shares the idea that the first version of the internet was about data created by people, while the next version is about data created by things. In 1999, Ashton said it best in this quote from an article in the RFID Journal:

"If we had computers that knew everything there was to know about things – using data they gathered without any help from us – we would be able to track and count everything, and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling, and whether they were fresh or past their best."

Most people think about being connected in terms of computers, tablets and smartphones. IoT describes a world where just about anything can be connected and communicate in an intelligent fashion. In other words, with the internet of things, the physical world is becoming one big information system.

1.2 Big Data

Big data refers to a process that is used when traditional *data* mining and handling techniques cannot uncover the insights and *meaning* of the underlying *data*. *Data* that is unstructured or time sensitive or simply very *large* cannot be processed by relational database engines.

Quite simply, big data reflects the changing world we live in. The more things change, the more the changes are captured and recorded as data. Take weather as an example. For a weather forecaster, the amount of data collected around the world about local conditions is substantial. Logically, it would make sense that local environments dictate regional effects and regional effects dictate global effects, but it could well be the other way around. One way or another, this weather data reflects the attributes of big data, where real-time processing is needed for a massive amount of data, and where the large number of inputs can be machine generated, personal observations or outside forces like sun spots.

Processing information like this illustrates why big data has become so important:

- Most data collected now is unstructured and requires different storage and processing than that found in traditional relational databases.
- Available computational power is sky-rocketing, meaning there are more opportunities to process big data.
- The Internet has democratized data, steadily increasing the data available while also producing more and more raw data.

Data in its raw form has no value. Data needs to be processed in order to be of valuable. However, herein lies

the inherent problem of big data. Is processing data from native object format to a usable insight worth the massive capital cost of doing so? Or is there just too much data with unknown values to justify the gamble of processing it with big data tools? Most of us would agree that being able to predict the weather would have value, the question is whether that value could outweigh the costs of crunching all the real-time data into a weather report that could be counted on.

The Volume: The Volume challenge refers to storing, processing and quickly accessing large amounts of data. While it is hard to quantify the boundary for a volume challenge, common data sets in the order of hundreds of Terabyte or more are considered to be big. In contrast to traditional storage technologies such as relational data base management systems (RDBMS), new Big Data technologies such as Hadoop are designed to easily scale with the amount of data to be stored and processed. In its most basic form, the Hadoop system uses its Hadoop Distributed File System (HDFS), to store raw data. Parallel processing is facilitated by means of its Map Reduce framework that is highly suitable for solving any embarrassingly parallel processing problems. With Hadoop it is possible to scale by simply adding more processing nodes to the Hadoop cluster without the need to do any reprogramming as the framework takes care of using additional resources as they become available. Summarizing the trends in the volume challenge one can observe a paradigm shift with respect to the way the data is handled. In traditional database management systems the database design is optimized for the specific usage requirements, i.e., data is pre-processed and only the information that is considered relevant is kept. In contrast, in a truly data-driven enterprise that builds on Big Data technologies, there is awareness that data may contain value beyond the current use. Thus a master data set of the raw data is kept that allows data scientists to discover further relationships in the data, relationships that may reside beyond the requirements of today. As a side effect it also reduce the costs of human error such as erroneous data extraction or transformation.

The Velocity: Velocity refers to the fact that data is streaming into the data infrastructure of the enterprise at a high rate and must be processed with minimal latency. To this end, different technologies are applicable, depending on the amount of state and complexity of analysis. In cases where only little state is required (e.g., maintaining a time window of incoming values), but complex calculations need to be performed over a temporally scoped subset of the data, Complex Event Processing (CEP) engines offer efficient solutions for processing incoming data in a stream manner. In contrast, when each new incoming data set needs to be related to a large number of previous records, but only simple aggregations and value comparisons are required, noSQL databases offer the necessary write performance. The required processing performance can then be achieved by using streaming infrastructures such as Storm.

The Variety: In a data-driven economy the objective is to maximize the business value by considering all available data. From a technical perspective one could formulate that problem by evaluating a function over all accessible data sets. In practice, however, this approach must confront the challenge of heterogeneous data sources ranging from unstructured textual sources (e.g., social media data) to the wide disparity in the formats of sensor data. Traditionally this challenge is addressed by various forms of data integration. In the context of Big Data there is however a new dimensions to the integration challenge which is the amount of different data sources that need to be integrated. Social media, open (governmental) data sources, and data platforms and markets, result in a data ecosystem of significant variation. As the integration of new data sources requires manual work to understand the source schema, to define the proper transformations and to develop data adapters, existing approaches do not scale effectively.

The Veracity: Apart from the original 3V's described above, an almost inexhaustible list of Big Data V's are discussed. For instance veracity relates to the trust and truthfulness of the data. Data may not fully be trusted, because of the way it has been acquired, for instance by unreliable sensors or imperfect natural language extraction

The Value: When we talk about value, we're referring to the worth of the data being extracted. Having endless amounts of data is one thing, but unless it can be turned into value it is useless. While there is a clear link between data and insights, this does not always mean there is value in Big Data. The most important part of embarking on a big data initiative is to understand the costs and benefits of collecting and analyzing the data to ensure that ultimately the data that is reaped can be monetized.

II. Why Big data with IoT

The first thing that comes to mind when talking about Big Data and IoT is the increase in the volume of data that will hit the data storage framework of companies. Data centers will have to be set up to handle this entire additional data load.

Taking into consideration the enormous impact IoT will on data storage infrastructure, organizations have begun to move towards the Platform-as-a-Service model, a cloud-based solution, as opposed to maintaining their own storage infrastructure. Unlike, in-house data systems that need to be constantly updated as the data load increases, PaaS provides flexibility, scalability, compliance, and a sophisticated architecture to store all valuable IoT data. Cloud storage options include public, private, as well as hybrid models. If a company has sensitive data that is subject to any regulatory compliance requirements that require heightened security, using a private cloud would be the best course of action. For other companies, a public or hybrid cloud can be used for the storage of IoT data.

III. Necessity of Implementing Big Data with IoT

IoT and data remain intrinsically linked together. Data consumed and produced keeps growing at an ever expanding rate. This influx of data is fueling widespread IoT adoption as there will be nearly 30.73 billion IoT connected devices by 2020. The Internet of Things (IoT) is an interconnection of several devices, networks, technologies, and human resources to achieve a common goal. There are a variety of IoT-based applications being used in different sectors and have succeeded in providing huge benefits to the users.

The data generated from IoT devices turns out to be of value only if it gets subjected to analysis, which brings data

analytics into the picture. Data Analytics (DA) is defined as a process, which is used to examine big and small data sets with varying data properties to extract meaningful conclusions and actionable insights. These conclusions are usually in the form of trends, patterns, and statistics that aid business organizations in proactively engaging with data to implement effective decision-making processes.

IV. Merging Data Analytics and IoT will Positively Impact Businesses

Data Analytics has a significant role to play in the growth and success of IoT applications and investments. Analytics tools will allow the business units to make effective use of their datasets as explained in the points listed below.

- **Volume:** There are huge clusters of data sets that IoT applications make use of. The business organizations need to manage these large volumes of data and need to analyze the same for extracting relevant patterns. These datasets along with real-time data can be analyzed easily and efficiently with data analytics software.
- **Structure:** IoT applications involve data sets that may have a varied structure as unstructured, semi-structured and structured data sets. There may also be a significant difference in the data formats and types. Data analytics will allow the business executive to analyze all of these varying sets of data using automated tools and software.
- **Driving Revenue:** The use of data analytics in IoT investments will allow the business units to gain an insight into customer preferences and choices. This would lead to the development of services and offers as per the customer demands and expectations. This, in turn, will improve the revenues and profits earned by the organizations.

V. Types of Analytics can be performed

- **Streaming Analytics:** This form of data analytics is also referred as event stream processing and it analyzes huge in-motion data sets. Real-time data streams are analyzed in this process to detect urgent situations and immediate actions. IoT applications based on financial transactions, air fleet tracking, traffic analysis etc. can benefit from this method.
- **Spatial Analytics:** This is the data analytics method that is used to analyze geographic patterns to determine the spatial relationship between the physical objects. Location-based IoT applications, such as smart parking applications can benefit from this form of data analytics.
- **Time Series Analytics:** As the name suggests, this form of data analytics is based upon the time-based data which is analyzed to reveal associated trends and patterns. IoT applications, such as weather forecasting applications and health monitoring systems can benefit from this form of data analytics method.
- **Prescriptive Analysis:** This form of data analytics is the combination of descriptive and predictive analysis. It is applied to understand the best steps of action that can be taken in a particular situation. Commercial IoT applications can make use of this form of data analytics to gain better conclusions.

VI. The IoT and Big Data job market

Big Data and the Internet of Things are the two most-talked-about technology topics of the last few years. This is one of the chief reasons why they occupy prominent places on analyst firm Gartner's most recent Hype Cycle for Emerging Technologies. These two technologies are set to transform all areas in business as well as everyday life. In the 2015 Internet of Things predictions, IDC notes that over 50% of IoT activity is centered in manufacturing, transportation, smart city, and consumer applications, but that within five years every industry will have rolled out IoT initiatives.

Data Science Central conducted a survey that showed how widespread IoT jobs are, today. This is a list of the top companies that are hiring for IoT related jobs:

1. PTC – The Product Development Company
2. Amazon
3. Continental
4. Savi Group
5. Intel
6. Ayla Networks
7. HP
8. LogMeIn.Inc
9. Red Hat. Inc
10. Honeywell
11. IBM
12. Renesas
13. Cisco Systems. Inc
14. Dell
15. InterDigital

The IoT and Data related positions that companies are hoping to fill with qualified people are:

1. Big Data Lead (IoT)
2. Data Scientist - IoT
3. Data Engineer - Sensors and IoT
4. Data Engineer Sensors and IoT Applications

Given these developments, the opportunities available to certified Big Data professionals in the rapidly growing 'Internet of Things' domain are endless. Simplilearn's Big Data Hadoop Architect Masters Program was designed with the IoT-driven world of the future in mind.
With over 200 hours of high-quality e-learning content, access to CloudLab – a cloud based Hadoop environment lab–on-demand support by Hadoop experts, simulation exams, and a certification to validate your skills, the Big Data Hadoop Architect Masters program will see you ready to take on the challenges and opportunities of a world where the Internet of Things is commonplace.

VII. Benefits of IoT based on Big Data

In literature, various structures for big data analysis and IoT proposed, which can manage the challenges of storage and analysis of high volume data from intelligent buildings. The first presented structure consists of three main components which are big data management, IoT sensor,

and data analysis. These analyzes use are in the realtime management of oxygen level, dangerous gases/soot and the amount of ambient light in smart buildings. In addition to smart building management, IoT devices and sensors for receiving traffic information can be used in real time traffic management with low cost and examine the strengths and weaknesses of existing traffic systems. In smart city management, the big data used in the analysis of data which obtained from different sensors such as water sensors, transportation network sensors, monitoring devices, smart home sensors and smart car park sensors. These data are generated and processed in a multi-stage model and ultimately reached a decision-making stage. These steps are data production, data collection, data integration, data categorization, data processing and decision making. Sometimes it is essential to pay attention to the concepts of web technology in particular proposed framework to investigate the analytical results obtained from the big data in the Internet of Things. In the literature, this topic has devised, and a conceptual framework has been proposed consisting of 5 layers: • Data Collection layer - collected data from various sources, the input layer is the proposed framework. • Extract-transform-load (ETL) layer - provides the ability to change the format of information received from different types of sensors into a defined format. • The semantic reasoning rules layer - an inference engine that acts on the information received from the ETL layer • Learning layer - From the data tailored to the existing extraction data, extract the various specifications and attributes, and finally, Machine learning-based models provided. • Action layer - executes a set of predefined actions by the outputs of the learning layer. Other applications of IoT help with geographic information analysis, cloud computing flow processing, big data analysis, and storage, cloud computing security, clustering mechanisms, health, privacy security, performance evaluation of monitoring algorithms, manufacturing systems, and energy development [17]

VIII. Conclusion

The development of IoT devices, smartphones, and social media provides decision makers with opportunities to extract valuable data about users, anticipate future trends and fraud detection. With the creation of transparent and usable data, big data can create the organizations' values, make the changes clear and expand their performance. The use of data generated from the IoT and the analytical tools creates many opportunities for organizations. These tools use predictive modeling technologies, clustering, classification to provide data mining solutions. IoT improves the decision-making habits of decision-makers. The emergence of IoT and related technologies, such as cloud computing, provides the ability to remove data sources in different domains. Typically, any data is considered useful in the domain itself, and data on shared domains can be used to provide different strategies. Machine learning, deep learning, and artificial intelligence are key technologies that are used to provide value-added applications along with IoT and big data in addition to being used in a stand-alone mode. Before the advent of IoT and cloud computing, the use of these technologies was not possible due to the high amount of data and required computational power. Different data analysis platforms, Business intelligence platforms and analytical applications are emerging platforms that have been introduced to help industries and organizations in transforming processes, improving productivity, and the ability to detect and increase agility. It is anticipated that the speed of technological progress in the next ten years, will be equal to the past thirty years. Therefore, we have to use all our efforts to update our lives to the Internet of Things technology regarding hardware and software.

References

[1] Networking and Information Technology Research and Development (NITRD) Program. (2015). Smart and Connected Communities Framework. [Online]. Available: https://www.nitrd.gov/sccc/materials/ scccframework.pdf.

[2] Y. Sun, Y. Xia, H. Song, and R. Bie, ''Internet of things services for small towns,'' in Proc. Int. Conf. Identificat., Inf. Knowl. Internet Things (IIKI), Oct. 2014, pp. 92–95.

[3] A. J. Jara, Y. Sun, H. Song, R. Bie, D. Genooud, and Y. Bocchi, ''Internet of Things for cultural heritage of smart cities and smart regions,'' in Proc. IEEE 29th Int. Conf. Adv. Inf. Netw. Appl. Workshops (WAINA), Mar. 2015, pp. 668–675.

[4] H. Song, R. Srinivasan, T. Sookoor, and S. Jeschke, Smart Cities: Foundations and Principles. Hoboken, NJ, USA: Wiley, 2016.

[5] U.S. Department of Transportation. (2013). Livability 101. [Online]. Available: http://www.dot.gov/livability/101.

[6] "A survey on Chromecasat digital device" Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162) Published in Volume 5 Issue 10 , October- 2018.

[7] "Multimedia And Its Applications" in International Journal for Research and Development in Technology (ISSN: 2349-3585) Published in Volume 10 Issue 5 on November 2018.

[8] "TensorFlow in Deep Learning" in International Journal of Innovative Research in Technology" (ISSN: 2349-6002) Published in Volume 5 Issue 9 on February 2019.