PREDICTION OF PARKINSON DISORDER BASED ON KINECT ARM MOVEMENT BY USING MACHINE LEARNING TECHNIQUES

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering
in
Computer Science & Engineering

Submitted By

Sagarjit Dash

(Roll No. 801532045)

Under the supervision of

Dr. Ashutosh Mishra

Assistant Professor

Thapar University, Patiala

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

July 2017

Certificate

Certificate

I hereby certify that the work which is being presented in the entitled, "Prediction of Parkinson disorder based on Kinect arm movement by using Machine Learning Techniques", in partial fulfillment of the requirements for the award of degree of Master of Engineering in Computer Science & Engineering submitted in Computer Science and Engineering Department at Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of Dr. Ashutosh Mishra and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

S. Dash

(Sagarjit Dash)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. Ashutosh Mishra)
Assistant Professor,
CSED

Abstract

Parkinson Disorder is a gradually increasing neurodegenerative disease. In this, motor and non-motor symptoms can impact on human basic functions to a wide variety of range. This article describes the required features and predicts the Parkinson Disorder in a patient with higher efficacy. In this work, mean vector velocity (MVV) was performed to analyze the essential features and crossed verified it by the Rank of Correlation. Previously many scientists analyzed Parkinson disorder using the fundamental classification models and also the different techniques to identify the postural tremor in the arm. It was found that they concentrated on the slow movement of the patients. But it has been proven that some abnormalities are also present in the medium movement. In this work the main tasks involve are (I) detecting four arm-features (II) calculating the mean instantaneous velocity of that four features and (III) performance evaluation of the four classification models under Artificial Neural Network. Finally we can able to predict the Parkinson disorder for patients with having good accuracy and sensitivity.

Acknowledgement

Acknowledgement

First of all I would like to thank the Almighty, who has always guided me to work on the right path of the life. It is a great privilege to express my gratitude and admiration towards my respected supervisor and PG coordinator **Dr. Ashutosh Mishra**, Assistant Professor, Computer Science & Engineering Department, Thapar University, Patiala. He has been an esteemed guide and great support behind achieving this task. This work would not have been possible without the encouragement and able guidance of him. I have got all required information time to time, I am also thankful to my supervisor for his time, patience, discussions and valuable comments. His enthusiasm and optimism made this experience both rewarding and enjoyable. I am truly grateful to him for extending his total co-operation and understanding whenever I needed help and guidance from him. I am also heartily thankful to **Dr. Maninder Singh**, Associate Professor and Head, Computer Science & Engineering Department, for motivation and providing uncanny guidance and support throughout the preparation of the thesis report.

I will be failing in my duty if I do not express my gratitude to **Dr. S. S. Bhatia**, Senior Professor and Dean of Academic Affairs, for making provisions of infrastructure such as library facilities, computer labs equipped with net facilities, immensely useful for the learners to equip themselves with the latest in the field.

I am also thankful to the entire faculty and staff members of Computer Science and Engineering Department for their direct-indirect help, cooperation, love and affection, which made my stay at Thapar University memorable. Last but not least, I would like to thank my family for their wonderful love and encouragement, without their blessings none of this would have been possible.

S Dash

Sagarjit Dash

801532045

Table of Contents

List of Figures

List of Tables

Abbreviation

PD – Parkinson Disease

ML – Machine Learning

ANN – Artificial Neural Network

DNN – Deep Neural Network

SM – Slow Movement

MM – Medium movement

FM – Fast Movement

Shld – Shoulder

Elb – Elbow

Wrst – Wrist

Hnd – Hand

J9 – Joint9

J10 – Joint10

J11 – Joint11

J12 – Joint12

WTH – Wrist to Hand

ETW – Elbow to Wrist

STH – Shoulder to Hand

Chapter 1
Introduction

In this chapter, we have discussed about the prediction of the Parkinson's disease affected patients using the analysis of different machine learning approaches. This chapter includes the main motivation and objectives of the research work. This study has been carried out by the prediction of 4 essential joints of arm. We have also discussed about the most affected region in arm for this neurodegenerative disease.

1.1 Background

Our human brain is responsible for performing different tasks. In the brain there are different areas like cerebrum, cerebellum, frontal lobe, temporal lobe and these are devoted to control mainly vision, voluntary muscle movements, memory etc. In each area a huge amount of neurons are available which are handling sensational works **[1].** For this reason, the synchronization and proper functioning between each neuron (Axon and Dendron) is very much crucial **[2, 3].** In many cases like malnutrition, inborn problems or for the age, the proper works of neurons are hampered. The processes (thinking, observing, leg movement, arm movement etc) between neurons are not happened perfectly in the motor circuit due to the interruption of the normal signals flow in central nervous system. This indicates an abnormality in the motor system which leads to Neurodegenerative disorder (like Parkinson's disease) **[4, 5].**

Neurodegenerative means the destruction of neuron cell which can't be reproduced. Once the structure of the cell is damaged, it is not possible to recover it back **[6].** Neurodegenerative diseases are not curable. There are different types of neurodegenerative diseases like Parkinson's disease (PD), Alzheimer's disease, Huntington's disease, Lewy Body Dementia (LBD), Motor Neuron Disease (MND) etc.

This work deals with the prediction on Parkinson's disorder which is now a day's a highly alerted insolvable, incurable disease. The symptoms of the Parkinson's disease are Tremors, Bradykinesia, rigidity, postural instability, Akinesia, Digestion problem, depression, low blood pressure, leg discomfort. The main symptoms among these are Bradykinesia, tremors.

1.2 Motivation

These days Parkinson disease is the most popular neurodegenerative disorder in all over the world. There are huge amount of brain diseases which have been discovered and predicted by many scientists using different techniques and have often produced good results. But this time much better approaches which have been introduced in the market, can be used for tighter and better prediction (may be the accuracy will be greater or same as previous prediction). It is caused by the loss of brain cells (neurons) in a part of the brain called the substantia nigra, which produces the chemical messenger dopamine. As the cell dies, less dopamine is produced and transported to the striatum, the area of the brain that co-ordinates movement. When more than 80% of dopamine has been lost then PD symptoms occur. Approximately 4 million people worldwide suffer from Parkinson's disease. Around 120,000 people in the UK have the condition of PD. There is no cure for Parkinson's disease, but treatments can help control the symptoms and maintain quality of life. The risk of developing Parkinson's disease increases with age. Symptoms usually occur after the age of 50. Around 1 in 20 people are diagnosed under the age of 40 years. Almost 30% of the people are facing this common but incurable problem **[7].** So this is a very serious issue in the world to predict the Parkinson in patient and improve the efficacy whereas possible.

So, these diseases can be re-analyzed based upon some newer techniques (modeling approach) which are far better than the older. Machine learning modeling and simulation approaches helps to find correct patients who are suffering from Parkinson abnormalities.

1.3 Objectives

The main purpose of this study is to improve the prediction efficiency and efficacy (the ability to produce desired or intended results) that would be beneficial for the patients who are suffering from Parkinson and the percentage ratio will be reduced. First stage Parkinson can be cured by the proper treatment. So it's important to identify the PD at the early stage for the betterment of the patients. The main objective of this research work is to find the 4 essential joints of the arm which are mainly responsible for the arm movement and machine learning models which are able to identify PD in patients. There are some novel points which have been maintained throughout the analysis for specific purposes and those are given below:-

(i) To predict the four essential attributes which are encountered for arm movement disorder among 20 joints of all body parts using physics concept.

(ii) To analyze the mathematical concept to cross check the correctness of the calculated features.

(iii) To analyze different ML algorithms, uses the attribute's values at its best and thus shows performance better in terms of accuracy.

(iv) To train the different classification algorithms and to identify the actual suitable algorithm for this prediction problem which can be correctly classified the PD patients.

1.4 Symptoms Categories

The symptoms of the Parkinson disease broadly divided into two categories.

Motor symptoms   Non-motor symptoms

1.4.1 **Motor symptoms**- This is a symptoms where any voluntary action involved. It's indicates the movement related disorder like tremor, rigidity, freezing, Bradykinesia or any voluntary muscle movement.

1.4.2 **Non-motor symptoms**- This is a symptoms which are not related to movements such as tiredness, depression, pain etc.

In addition with there are two other categories of PD which are divided by doctors:

Primary symptom   Secondary symptom

1.4.3 **Primary symptom**- It is the most noticeable or important symptom. The three primary symptoms of Parkinson's are all motor symptoms: tremor, rigidity or stiffness and slowness of movement (Bradykinesia). Balance and posture are also affected as Parkinson's progresses, so postural imbalance is sometimes seen as the fourth primary symptom.

1.4.4 **Secondary symptom**- It is less obvious symptom which still has an impact on quality of life. These can be either motor or non-motor.

Parkinson's is a very individual condition, with symptoms varying a great deal from person to person. A very wide range of symptoms is associated with Parkinson's, and not everyone will experience every symptom.

1.5 PD symptoms

Symptoms tend to appear gradually, normally in just one side of the body at first, although both sides will eventually be affected. As Parkinson's progresses, symptoms appear and develop at different rates in different people. In many people symptoms also fluctuate from day to day, with the person experiencing 'good' and 'bad' days. The main symptoms for the Parkinson disorder are tremor, Bradykinesia and rigidity which will be explained in details below.

1.5.1 Bradykinesia

Bradykinesia is one of the main and important symptoms in Parkinson disease. "Bradykinesia means slowness of movement and is one of the cardinal manifestations of Parkinson's disease" **[8].**

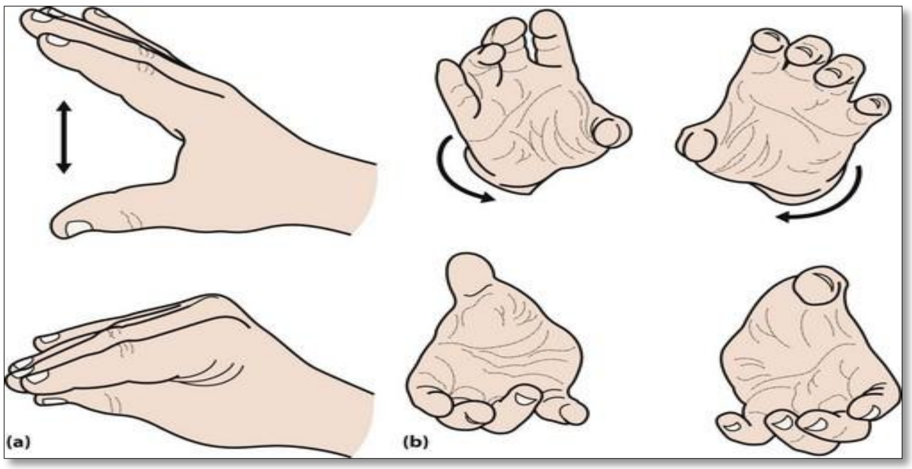

Fig1.1: Tests for Bradykinesia. (a) Ask the patient to tap their fingers together quickly. (b) Ask the patient to twist their hand as if polishing door knobs.

The above figure describes a test to prove that if the patient is suffering from Bradykinesia or not. For that the patient was told to tap their fingers together in speed (fig1: a) and in second figure (fig1: b) ask to twist their hand and result was observed. When a person has this symptom they may have a hard time to in completing the movement once the attempt has been initiated because of the impact on the function of their basal ganglia, which is the part of the brain that is responsible for motor learning and control.

1.5.2 Tremor

Tremor is rhythmic and voluntary muscle movement to-and-fro in one or more body parts **[9]** .It may occur in any age but it is most common in middle-aged and older persons. Tremor affects men and women equally. There are different types of tremors based on (I) environments in which it is shown: at resting condition, while managing certain positions; (II) the different body parts like arms, hand, leg and other organs; (III) the frequency at which the tremor manifests: lower(less than 4Hz), average(4 to 7 Hz) or higher(greater than 7Hz) bandwidths. If we broadly classify the tremors then we will get the following types.

(a) Resting Tremor (b) Postural Tremor

a) **Resting Tremor**: Resting tremor(RT), the important characteristic of Parkinson disorder that appears at 4Hz to 6Hz frequency range**[10]** and disappears when voluntary actions are performed like arm movement, leg movement etc. Basically resting tremor can be visualized in resting situation like sitting or sleeping. In the early stages of the disease, about 70 percent of people experience a slight tremor in the hand or foot on one side of the body, or less commonly in the jaw or face. A typical onset is tremor in one finger. The tremor consists of a shaking or oscillating movement, and usually appears when a person's muscles are relaxed, or at rest, hence the term "resting tremor."

b) **Postural Tremor:** Postural tremor (PT), an action that occurs during any kind of movement of an affected body parts. It is generally appears in more than 6Hz frequency range and there is no limitation to the upper value **[11].** PT can be visualize in dynamic situation like walking, running, or throwing something or anything which is related to motion. Our prediction and analysis has been carried out with the postural tremor where arm movement is the main concern and it will determine if any patient is suffering from Parkinson’s disease or not and in which portion of the arm is mostly affected.

Fig 1.2: (a) patient are experiencing resting tremor (b) second patient are experiencing postural tremor by upward the hands against gravity.

1.5.3 Rigidity

Rigidity causes stiffness and inflexibility of the limbs, neck and trunk. Muscles normally stretch when they move, and then relax when they are at rest. In Parkinson's rigidity, the muscle tone of an affected limb is always stiff and does not relax, sometimes contributing to a decreased range of motion. People with PD most commonly experience tightness of the neck, shoulder and leg. A person with rigidity and Bradykinesia tends to not swing his or her arms when walking. Rigidity can be uncomfortable or even painful.

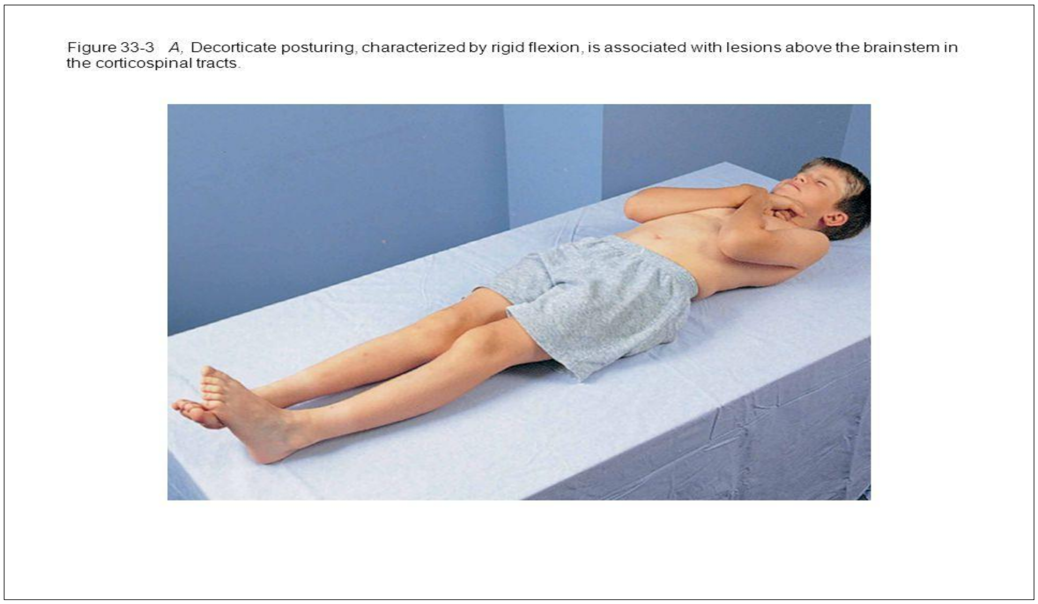

Fig 1.3: Example of muscle stiffness (rigidity)

1.6 Thesis Organization

This research work proposes an effective methodology for identify the Parkinson's disease based on some mathematical and physical concept. Different results of the classification algorithms of machine learning are analyzed and final outcome is made based on the accuracy of the model. The rest of the thesis is as organized below:

The first chapter provides a basic introduction of Parkinson disorder and different types of symptoms on it. The **second chapter** gives an account of the literature survey and the various approaches of machine learning for the analysis of PD. The **third chapter** deals with the problem statement. This chapter deals with the main aim of carrying this research work and the objectives of the thesis. The **fourth chapter** identifies the tools and techniques employed to do the proposed work. This chapter acquaints with the RapidMiner tool and its features, the machine learning tasks. The predictive machine learning model like classification, descriptive machine learning model like clustering are also explained. In **fifth chapter**, we have presented the methodology on the proposed approach. This chapter includes the description of workflow, the three algorithms used: Deep Neural Network, Neural Network, AutoMLP and Perceptron. In the **sixth chapter**, we have shown the implementation and results to predict the disease in patients. It includes the procedure applied, experimental setup, results and observations. The **seventh chapter** concludes the present work done under different constraints. The **eighth chapter** is the indication about the limitation of the work and provides possible improvements in future.

Chapter 2
Literature Review

In this chapter, we discussed about various approaches for Parkinson disorder prediction and measurements using various techniques of machine learning, proposed by various researchers. From last four decades, Parkinson disorder has become a serious issue in all over the world. At the very beginning there are several techniques which were used to detect and for treatment. There are several models and measurements which have been followed those are described below:

From several years, various machine learning models have been constructed and executed to analyze and measure the Parkinson disease. In 2012, G.Yadav et al. they performed a comparative analysis of different classifiers like tree, statistical methods and Support Vector Machine. This paper was analyzed on speech articulation difficulty symptoms of PD affected people and they got high sensitivity (.62, .97, .91 respectively) **[12].**

There are many other symptoms like leg abnormality symptoms, hand abnormality symptoms, gait disorder. But this paper only analyzed on the voice abnormality. One of the main symptoms is the arm movement which should be analyzed on the basis of machine learning techniques. Most of the people which are middle aged or late middle aged, are suffering the most from arm and leg tremor as early detection of PD for them are difficult. Therefore it's necessary to predict the arm movement disorder for the PD patients.

In 2011, Juan Sanchez-Ramos et al. worked on postural tremor and essential tremor (ET). They performed quantitative analysis of tremor in Welders. Here they have tried to analysis to distinguish the tremor severity among the workers. There manganese was chronically exposed in Welding fumes to get the effect on workers. They established that mean centre frequencies (CF) of Welders and patients with ET were significantly higher than the mean CF of PD subjects **[13].**

In 2014, Mohamed Elgendi et al. did a preliminary study on healthy patients and this evaluation was performed on arm movement and data were collected from Kinect sensor. This was done in order to check the efficacy. They were able to classify three different (slow, medium and fast) arm movements with an overall deviation 5.43% for interclass classification and .49% for intraclass classification **[14].**

In 2011, Mohamed Elgendi et al. proposed an algorithm which was able to classify the different velocity of arm. They established abnormality is associated with slowness while the fatness indicated the healthy movement. They were able to classify the three speeds by a developed algorithm with 4.94% error rate **[15].**

In 2001, H.Ben-Pazi et al. proposed the prediction the resting tremor of Parkinson disorder by the evidence of oscillators. This limb based prediction was carried out through accelerometer sensor data, where Peak frequencies have been taken for arm, leg and chin those are similar. Records were digitized and spectral analysis was performed. Coherence between different axes of individual accelerometers and between different segments of the same limb was high. However, coherence between tremors in different limbs was low. There was no consistent pattern across patients of coherence. These data suggest that tremor in PD is generated by multiple oscillatory circuits, which operate on similar frequencies **[16].**

In 2015, Ondřej Ťupa et al. worked on gait features estimation to indentify Parkinson Disorder using Kinect. This paper enables the detection of selected gait attributes by Microsoft (MS) Kinect image and depth sensors to track movements in three-dimensional space **[17]**. Digital signal-processing methods for rejecting gross data-acquisition errors, segmenting video frames, and extracting gait features. The proposed algorithm describes methods for estimating the leg length, normalized average stride length (SL), and gait velocity (GV) of the individuals in the given sets using MS Kinect data. Here the achieved accuracy was 97.2%.

In 2016, Reza Fazel-Rezai1 et al. performed tremor quantification of PD that was a pilot study where the tremor severity was predicted from the wearable inertial sensors. Classification model was built by the machine learning approach where SVM was taken into the consideration. They found that index finger more accurately predict tremor severity compared to signals from a sensor placed on the wrist. The final outcome of this analysis was 88.6% to 88.9% for resting tremor and 78.8% to 81.8% for postural tremor **[18].**

In 2000, MegMorris et al. produced a work on postural instability of Parkinson Disorder- a comparison based study with concurrent task. The purpose of this investigation was to determine the effects of dual task performance on postural instability in subjects with idiopathic Parkinson's disease (PD) compared with healthy elderly people **[19].**

Chapter 3
Problem Statement

3.1 Research Gap Analysis

These time neurodegenerative disorder is creating a serious issue all over the world and PD is one of the main disease under this category. Every year, this disease is encountered mostly in middle aged people. Early age detection of Parkinson is quite tough as there is no such direct methodology invented as per the knowledge. So it's very much important to predict the PD with higher efficacy in the middle aged people. If this is disease correctly countered then the PD affected patient might be cured otherwise wrong detection may lead the patients to permanently incurable state. Many researchers have focused mainly on gait abnormality, voice recognition and proposed some algorithms by using data mining, machine learning models or different estimation techniques on gait to predict Parkinson disorder. But very few of them concentrated on leg and arm movement which show major abnormalities in PD patients. Those who worked on arm movement they used different classification techniques to identify the disease and they calculated accuracy of the trained models. There are some better approaches which can be followed to get the better efficacy though there may a chance to get lower accuracy.

3.2 Problem formulation and approach

- Prediction of the Parkinson disease is calculated so that the patients can understand that if they are facing this incurable disease or not.
- It's needed to take the arm movements according to time of each patient separately to analyze the abnormality with equal importance.
- This work proposes an effective methodology (MVV) to get the essential features of the whole arm which will indicate the abnormality. Besides this, different types of classification models of machine learning are analyzed and trained and final outcome is made based on the accuracy of the model.
- This work also signifies that Deep Neural Network is the best approach among four techniques to predict the PD with its efficiency.

3.3 Objectives

- To study for better understanding of the existing techniques and tools applied on Parkinson disorder in the machine learning concepts.
- To study and implement the various algorithms which one uses the attribute's values at its best and thus performs better in terms of accuracy.
- To estimate the four essential feature of the arm from the dataset for each individual subjects.
- To perform and detect the PD based on two steps processfor each subject separately.

3.4 Scope of the research

This Parkinson disorder analysis by using machine learning approach will provide an easierway to diagnose the disease. In this work we have analyzed only the arm movement to detect PD. But same concept can be applied on the leg movement and better result can be achieved. The various other machine learning techniques and algorithms can be taken into the consideration so that the better outcome will produce.

Chapter 4
Tools and Techniques Used

4.1 Tool used

For the purpose of this research work, RapidMiner Studio tool is employed which provides an integrated environment for data preparation, machine learning, deep learning, text mining, and predictive analytics. It is also used for business and commercial applications as well as research, education, training, rapid prototyping, application development and in many more areas **[20].** Rapid Miner uses a client/server model with the server offered as Software as a Service or on cloud infrastructures. It's analysis as well as user interactive interface are pretty good and produce satisfactory results compare than other data mining software's like Weka, Clementine which are very common. Another advantage of this software is higher leveled complex calculation features are available in the free trail. This software can handle wide variety range of types of the data like nominal values, numerical values, integers, real numbers, text, 2-value nominal, multi-value nominal etc. Prediction of Parkinson's disorder has been performed by RapidMiner Studio7.4.000.

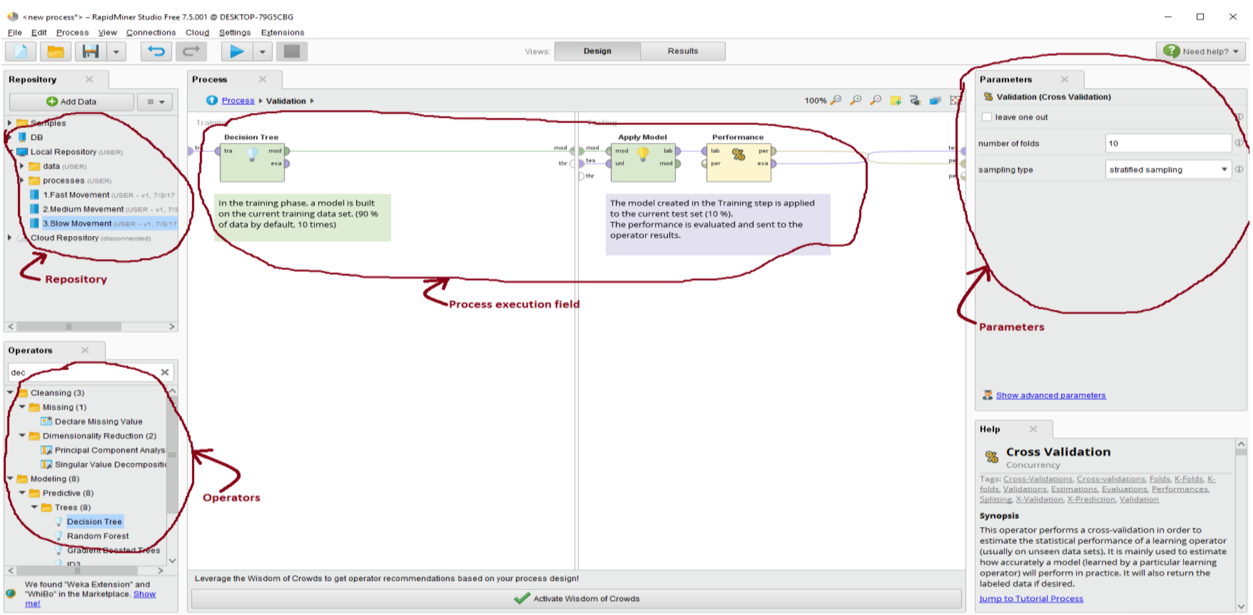

Fig: 4.1: functional window of the Rapid software

This software is mainly divided in 4 blocks:

(a) Repository
(b) Operators
(c) Process
(d) Parameters

Repository area is used to select the existing data or user can add new data file by making new repository under local repository. Operator's area is very much useful where user can select any models, besides models there are several another options like data access, blending, scoring, validation, utility etc. The main area in RapidMiner is the process area where the process structure is defined and linked them with nodes according to the analysis. Another important area which is available at right side of the screen is the setup of parameters. According to the requirements user can tune the parameters for getting different results. Besides this, to the left bottom corner there is one web link from where new packages can be installed in the RapidMiner directly **[21].** So this time RapidMiner is very much advantageous and open source software in the market.

4.2 RapidMiner tool features

- Rapid miner is an environment for machine learning and data mining processes.
- It represents a new approach to design even very complicated problems by using a modular operator concept which allows design of complex nested operator chains for huge number of learning problems.
- Rapid miner uses XML to describe the operator trees modeling knowledge discovery process.
- It contains more than 100 learning schemes for regression, classification and clustering analysis.
- Rapid Miner has a lot of functionality, is polished and has good connectivity.
- Rapid Miner includes many learning algorithms from WEKA.
- It easily reads and writes Excel files and different databases and it supports twenty two file formats.
- You program by piping components together in a graphic ETL work flows.
- If you set up an illegal work flows Rapid Miner suggest Quick Fixes to make it legal.

4.3 Specializations:

- Rapid Miner provides support for most types of databases, which means that users can import information from a variety of database sources to be examined and analyzed within the application.

- Specialized for Business solutions that include predictive analysis and statistical computing.

4.4 Modelling algorithms

The various modelling available in RapidMiner tool are as follows:

In predictive analysis the models are – decision tree, Naïve Bayes, ID3, random forest, Neural Net, Linear regression, Logistic regression, Support Vector Machine, Discriminant analysis, Ensembles etc.In segmentation the models are- agglomerative clustering, DBSCAN, K-Means etc. In Association the models are- association, FP Growth. Besides these there are different statistical methods like correlation matrix, ANOVA matrix, transition matrix, rainflow matrix. For the optimization process there are parameters optimizations, features selection, feature generations etc.

4.5 Deployment

The main importance of outcome based study is to get the proper results on some tasks in real time environment. RapidMiner is advantageous to find complete solution of the predictive analysis. Deployment branches the gap between analytics and action by supporting better decision making in various areas.

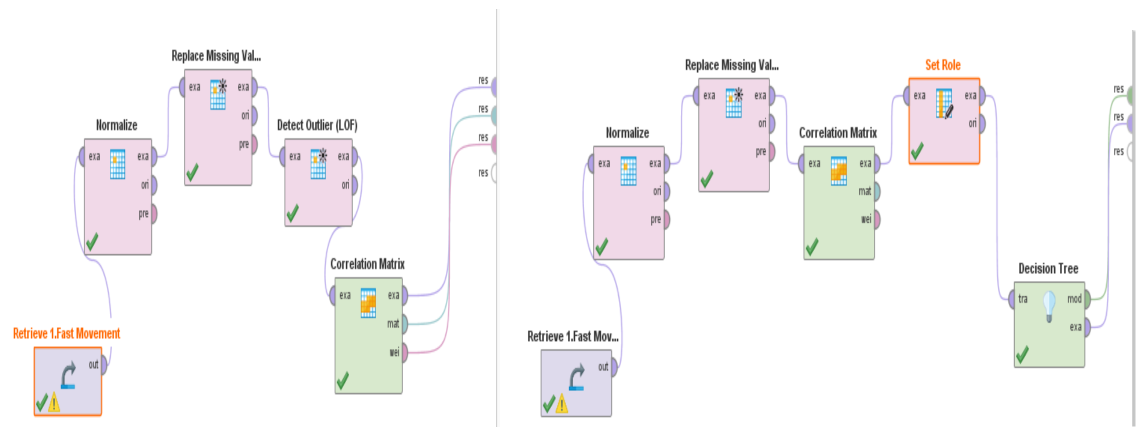

Fig 4.2: Various Deployment scenarios in RapidMiner tool

4.6 Microsoft Kinect:

Microsoft Kinect for Xbox360 is a motion detecting sensor invented by Microsoft. Kinect is a low cost solution for clinical as well as home based assessment. It is an effective tool for tacking body movements that is very promising to investigate tremor and slowness in arm movement.

Brook Galna et al. established that Microsoft Kinect Sensor gives very accurate results for movements in people who are suffering from Parkinson's disease **[22].** The configuration of Kinect camera consists of a RGB camera, two 3D depth sensor cameras, one LED light source, two multi-array microphones and a motorized tilt which is used to move upward or downward.

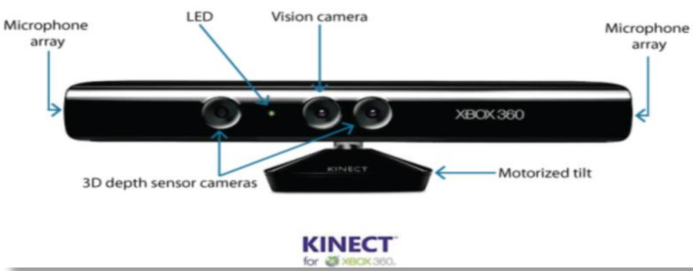

Fig 4.3: Kinect Sensor Camera and its configuration

4.7 Machine Learning tasks

Machine learning a study for data analysis which is used to do knowledge discovery and pattern recognition by iterative learning from the data without being explicitly programmed. The different techniques are used to build models where the class labels are known or not known before the starting of the prediction **[23].** Based on the class label of input and output from the unknown data outcome can be achieved. The concept of machine learning is broadly classified into two parts.

- Predictive Analysis
- Descriptive Analysis

4.7.1 **Predictive analysis** is the use of data, mathematical algorithms and machine learning to identify the likelihood of future events based on historical data. The main goal of predictive analytics is to use the knowledge of what *has* happened to provide the best valuation of what *will* happen. In other words, predictive analytics can offer a complete view of what is going on and the information we need to succeed. This analysis is a huge combination of statistical techniques, models to analyze the pattern of the goal.

4.7.2 **Descriptive analysis** looks at data and analyzes past events for insight as to how to approach the future. Descriptive analytics looks at past performance and understands that performance by mining historical data to look for the reasons behind past success or failure. Almost all management reporting such as sales, marketing, operations, and finance, uses this type of post-mortem analysis. Descriptive models quantify relationships in data in a way that is often used to classify customers or prospects into groups. Unlike predictive models that focus on predicting a single customer behavior (such as credit risk), descriptive models identify many different relationships between customers or products. Descriptive models do not rank-order customers by their likelihood of taking a particular action the way predictive models do.

According to the data the machine learning models are divided into two parts those are given below:

- Supervised learning
- Unsupervised learning

4.7.3 Supervised learning

Key points for supervised learning algorithm:

- The class labels are previously known.
- In the training set input and output both are included.
- This learning technique leads to the certainty.
- There is mapping method which matches with the predefined classes.
- There is a predefined target attribute.
- E.g. Classification(Categorical), regression(numerical)

4.7.4 Unsupervised learning

Key points for unsupervised learning algorithm:

- Class labels are not known previously.
- Prediction undirected and results unknown.
- Main purpose to find the patterns using inputs.
- E.g. Clustering.

4.8 Algorithms categories

There is several machine learning models are there under the algorithms categories like classification, regression, clustering etc. We will mainly focus on classification and rest two will be discussed briefly.

4.8.1 Classification is the machine learning approach that assigns the values to the sub-populations of the specific classes to predict the binary valued, Boolean types or any particular category based results. Classification gives the probabilistic estimation also. The main purpose of the classification is to identify the target labels according to the data with higher efficacy. Let's take an example: In email message it can be identified that a particular message is spam or not based on some attribute of that message. Like email from where the message has been received, content of the message, proper decoration of the whole message etc. Based on the above attributes evaluation will be performed and result can be achieved. Similarly according to our problem it can be revealed that a patient is suffering from PD disease or nor not based on some features. We are providing comparison among some models according to their training speed, accuracy, prediction speed and capability of irrelevant features handling etc.

Table 4.1 Comparison between different classification models on conditions

Model Name	Training Speed	Prediction Speed	Predictive accuracy	Irrelevant features Handle well
KNN	Fast	Depends on inputs	Lower	No
Logistic regression	Fast	Fast	Lower	No
Naïve Bayes	Fast	Fast	Lower	Yes
Decision Tree	Fast	Fast	Lower	No
Neural Networks	Slow	Fast	Higher	Yes
DNN	Fast	Fast	Higher	Yes
Perceptron	Slow	Fast	Average	Yes

4.8.2 Regression analysis is a statistical modeling technique which establishes the relationship between the variables. Basically two types of variable are involved in regression: one is dependent variable, depending on the other variables this variable is calculated and second is independent variables which determine dependent variables. Regression technique is used to predict continuous values (ex: age from 20 to 35) where as classification uses discrete values (ex: Spam/ not spam). This is the main difference between classification and regression.

4.8.3 Clustering is a task for grouping the similar type of data together so that different data cluster will be found and dissimilar data items can be separated those helps to discover any pattern or discover knowledge insights. This technique is used for categorical as wellas numerical data. There is several application of clustering like market research, pattern recognition, data analysis, image processing etc.

Chapter 5
Materials & Methodology

Before going to the actual process of analysis to determine PD the very first step is to collect the raw data from repositories of any websites. Instead of taking the data from website it is also possible to generate the data by own using specific tools. Among two approaches mentioned above we have considered data from the website.

5.1 Data Collection

The Kinect sensor provides numerical data that directly related to hand movement. Captured data which are unidentified numerical data, representing motion vectors in the 3 coordinates(x, y, z), were stored in the database**[24-26].** Some healthy patients were told to perform abnormal arm movement and those data were captured by the researchers at the office of Institute of Media Innovation, Nanyang Technological University, Singapore. We can find the data in any one of the following websites mentioned below-

http://www.elgendi.net/databases.htm.
http://www3.ntu.edu.sg/imi/piconflavien/autres/data-speed-arm.zip

5.2 Data Description

In the dataset, arm movement of 27 healthy patients (21 males and 6 females) was considered. Among of them two were left-handed, with a mean ± SD age of 29.7±4.1, height of 172.9±9.3, arm length of 71.3±5.2.

Table 5.1 Subject Population Characteristics and attributes

Attributes	Description(Mean/Count)
Patients	27
Gender	Male - 21, Female - 6
Age	29.7 ± 4.1
Arm Length	71.3 ± 5.2
Height	172.9 ± 9.3
Handed(R/L)	Right- 25, Left- 2

The motions were measured using the Kinect camera which is located at a distance of 2.7 meters from the subject at a height of 1.2 meters above the floor. The sampling frequency of the Kinect camera is 30 Hz. For the voluntary movement (for certain action) sampling frequency must be above of 7Hz. So here we have taken an arbitrary sampling frequency to 30Hz for better and almost perfect analysis.

During the experiment, the body of the user should be facing the sensor with an angle shift of 45∘ to the right of the Kinect sensor. The reason behind the 45∘ shift to the right is to prevent the arm joints from intersecting with the body joints**[27],** as shown in Fig 5.1. This will generate reliable arm motion in order to study the impact of each joint of the arm on the overall speed of the right arm movement more precisely. These collected arm movements will be used as a benchmark for effective speed detection of an arm movement.

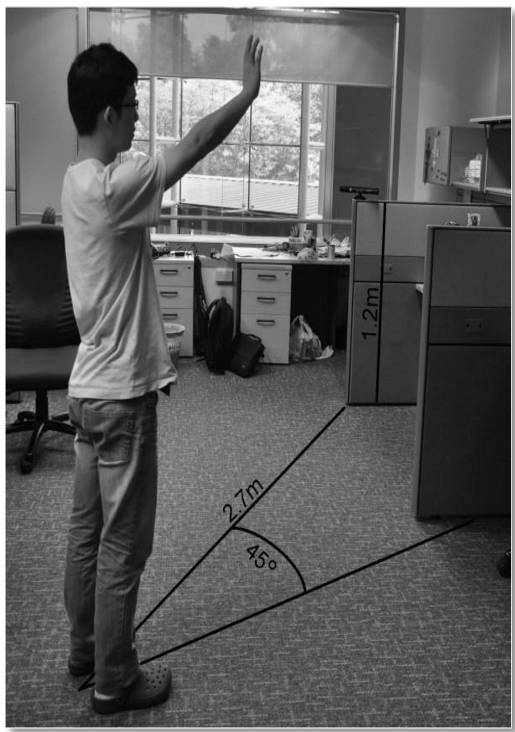

Fig 5.1: Experimental setup of patient and Kinect**[14]**

Data were recorded while each patient is standing up with an initial position both arms extended along the body side. Then, the patient is asked to move arm in upward in curved motion. Patient performs three trials: ‘slow’, ‘normal’, and ‘fast’; with five arm movements for each set. Therefore the number of recorded movements is 405 (27 subject × 5 movements × 3 speeds). The slow arm movement indicates Bradykinesia (central nervous system disorder, especially PD disorder) while the fast movement represents a healthy arm. For the slow action, the patient has been asked to move his arm with the feeling of a heavy object in his arm that does not allow him to move normal. Capturing the arm movement has been done manually. In other words, the subjects wait for a signal from the recording person to start their movement and then they maintain their arm up until they get a signal to come back to the initial position.

5.3 Proposed Approach

The main purpose of this work is to analyze the PD symptoms and differentiate between PD and non-PD patients. The whole analyzing system is divided into 4 parts which are given below.

- Data preparation
- Attributes Selection
- Model simulation
- Results analysis

The first and foremost step of the analysis is to collect the data from different sources like website, any repository and by data generation. Now, the data should be prepared as per our usability. After the data preparation step, next step is attributes selection. This is very important step because the prediction will not be proper if we consider all the attributes present in the data set though our requirement is few among them. The third step is model simulation where the model will be trained and decision will be taken based on training. The simulation processes show the performance and results in terms of accuracy. In the last stage the results will be analyzed based on the classification. The whole system diagram for the prediction is shown in fig5.2.

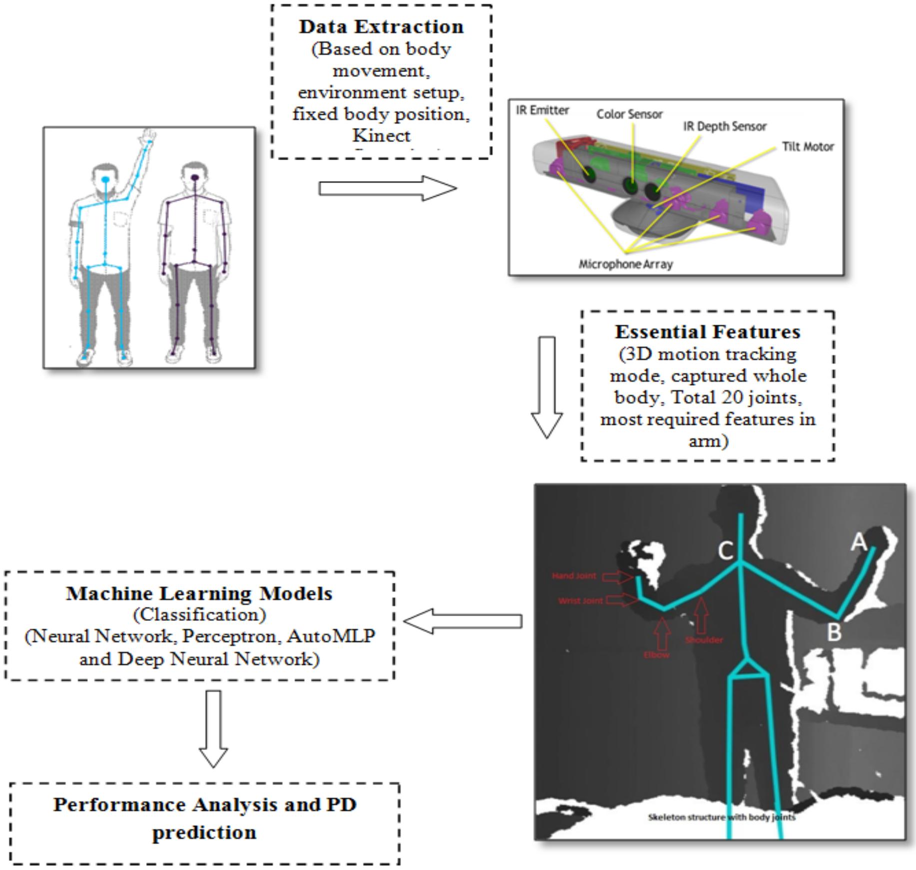

Fig 5.2: PD detected Feed forward Machine Learning Process

In this worktotal 27 subjects have been considered and randomly subjects have been chosen. In one subject,randomly 4 experimentshave been chosen out of 15 experiments to analyze. All the dataset was for postural tremor. In each category, there were three types of movement: *slow*, *medium and fast*.From the various machine learning approach we have learned that for the categorical data we have to go for classification method. This can be justified by the two main reasons:

- We understand that the collected dataset from the Kinect sensor are based on the annotator observations. So it might have possibility that two annotator have the different observations range for different speed. So there will be uncertainty if we consider any annotator is true, which can lead to the wrong prediction of the PD. This problem has been overcome by the classification which gives the correct prediction as category1- *slow movement*, category2- *medium movement* and category3- *First movement*

- The speed gap of shoulder to elbow, elbow to wrist and wrist to hand are overlapping for some experiments. The classifier model shows a clear picture of instantaneous velocity gap for symmetric experiments as well as asymmetric experiments. So we can make difference PD patients from Non-PD patients by analyzing the mean instantaneous velocities which are used in the feature selection process.

Now the aim is to find out whether it is possible to classify the movements and to find out the prediction of the PD patients by using four different classifier methods. To implement this, the instantaneous velocity has been calculated as an intermediate result on each experiment to get the positional changes for 27 patients. The calculated instantaneous velocity was based on the following equation:

$$MIV= \frac{\sqrt{((X_2-X_1)^2+ (Y_2-Y_1)^2 +(Z_2-Z_1)^2)}}{f} \quad [1]$$

Where (x_2-x_1) is the positional change in x coordinates in the change of any time instance, similarly (y_2-y_1) and (z_2-z_1) is for y and z coordinates respectively and f is the sampling frequency. This formula is used on the entire instances in each experiment. After getting the mean instantaneous velocities, Perceptron,AutoMLP, Neural Network, and DNN are used to determine the classification results and performance is evaluated. To get the actual PD affected patients with respect to this prediction we have calculated the sensitivity and specificity on the intermediate results. The intermediate results for all instantaneous velocities have been passed through the above four classifiers to check the correctness of the prediction on the slow movement, medium movement and fast movement separately.

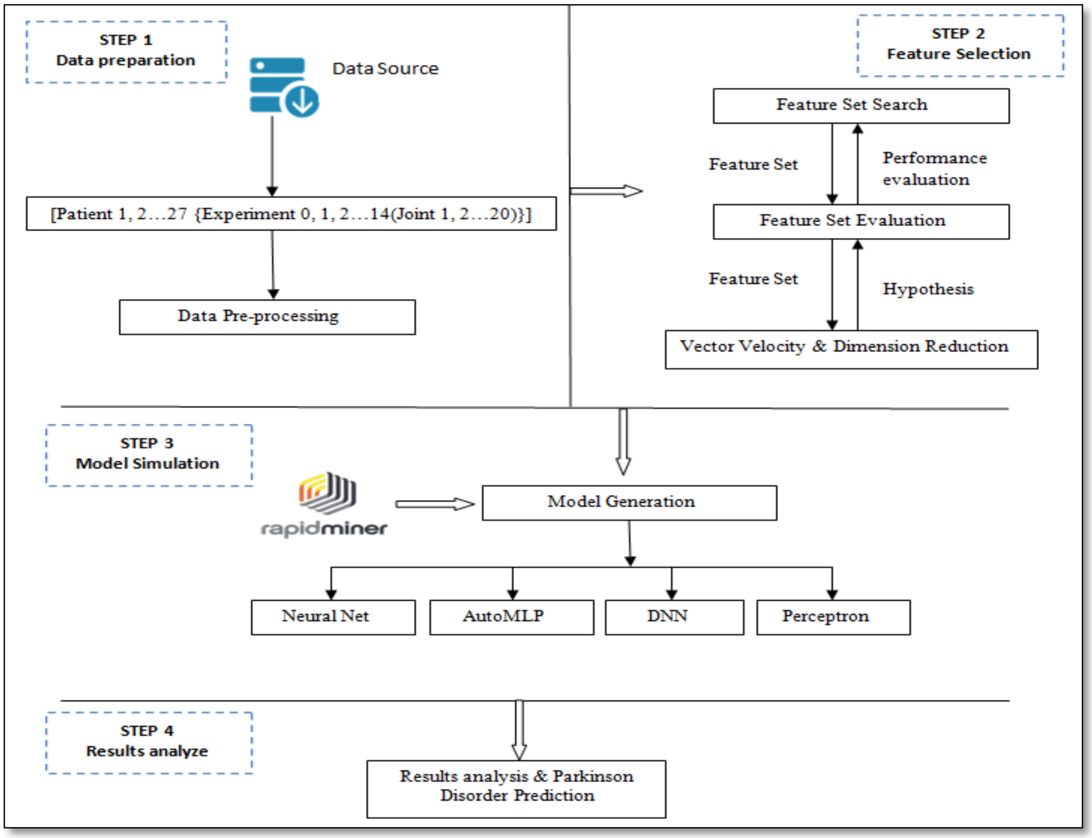

Fig 5.3: Workflow diagram of PD prediction

5.4 Description of Workflow

Before going to the actual analysis for prediction, though it is often neglected but it's very much necessary step to pre-process the dataset. Kinect XBOX360 and API tool has been used to capture the positional data for the arm movement**[28].** Though we have focused on the position but we are interested to record all joints. First of all the whole body joints will be captured by Kinect and we have considered only 4 joints for further analysis.

Now in this process we have collected the data from the Kinect in 30Hz frequency. So, we have raw data with 20 joints and each are having 3 coordinates x, y and z (real numbers) representing positional value with direction by their sign, each single movement was estimated as total 60 vector points (20 joints × 3 dimensions). Now from the unlabeled data we have to found the meaningful relation. This relation has been explained as: In the whole dataset total 27 subjects (patients) 15 experiments for each subject and in each experiment 20 joints. 60 vector points is divided into 20 joints and for each joint 3D data points are available. After getting the relation among them the preprocessing has been applied as explained below:

Data cleansing is the process of detecting and correcting corrupt or inaccurate records from a record set, table or database and refers to identifying incomplete, inaccurate or irrelevant parts of the data and then replacing, modifying or deleting the dirty or coarse data**[29].** In our dataset we have found some missing values which have been replaced by # and then analysis was carried forward and outliers has been removed.

Normalization is the very crucial step in data pre-processing phase where the attribute data is scaled to fit into a specific range. In our work it is not needed to perform.

5.5 Methods used

Our analysis is based on classification problem and there are several classifier methods available for classification. We have to find the appropriate classifier for arm movement with its significance. To choose the classifier we are interested to analyze their internal configuration and mathematical rules for signal passing techniques through one neuron cell to another neuron cell. Because depending on the signal passing and synchronization of the axons and den drones, tremor is appeared **[31].**We have taken 4 classification techniques which all are under the Artificial Neural Network domain. The four models are Perceptron, AutoMLP, Neural Network and Deep Neural Network. Now we are going to study them and try to find out the best techniques among them.

5.5.1 Neural Network

Neural network model is the simplest model which is first invented under artificial neural network. Neural network is defining as the pattern recognition techniques for machine learning approach.

To determine the classifier we are interested to analyze the neural network, where and how signals pass through one neuron cell to another because depending on the signal passing and synchronization of the axons and den drones tremor will appear **[32].**

As illustrated in the following figure (fig 5.4), incoming signals are received by the cell's dendrites through a biochemical process. The process allows the impulse to be weighted according to its relative importance or frequency. As the cell body begins accumulating the incoming signals, a threshold is reached at which the cell fires and the output signal are transmitted via an electrochemical process down the axon. At the axon's terminals, the electric signal is again processed as a chemical signal to be passed to the neighboring neurons across a tiny gap known as a synapse **[33].**

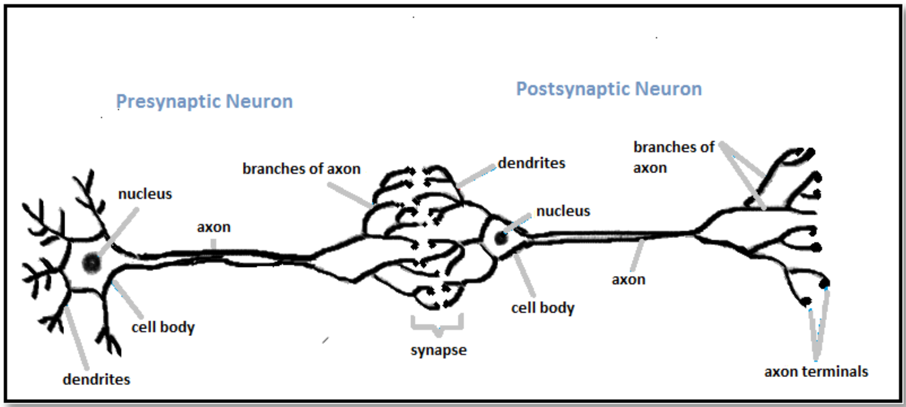

Fig 5.4: Diagram of biological neurons **[15]**

The model of a single artificial neuron can be understood in terms very similar to the biological model**[34].** As depicted in the following figure, a directed network diagram defines a relationship between the input signals received by the dendrites (x variables), and the output signal (z variable). Just as with the biological neuron, each dendrite's signal is weighted (w values) according to its importance—ignore, for now, how these weights are determined. The input signals are summed by the cell body and the signal is passed on according to an activation function denoted by f:

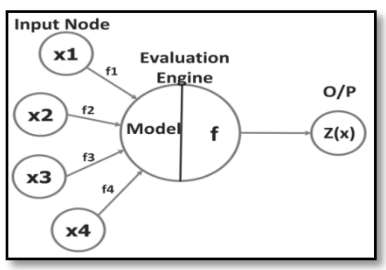

Fig 5.5: Block structure of a single neuron

In artificial neural network there are different types of activation functions are used to propagate the signal from one neuron to another like sigmoid, tanh, ReLU etc.

- **Sigmoid:** takes a real-valued input and squashes it to range between 0 and 1

$$\sigma(x) = 1 / (1 + e^{-x}) \qquad [2]$$

- **tanh:** takes a real-valued input and squashes it to the range [-1, 1]

$$\tanh(x) = 2\sigma(2x) - 1 \qquad [3]$$

- **ReLU**: ReLU stands for Rectified Linear Unit. It takes a real-valued input and thresholds it at zero (replaces negative values with zero)

$$f(x) = \max(0, x) \quad [4]$$

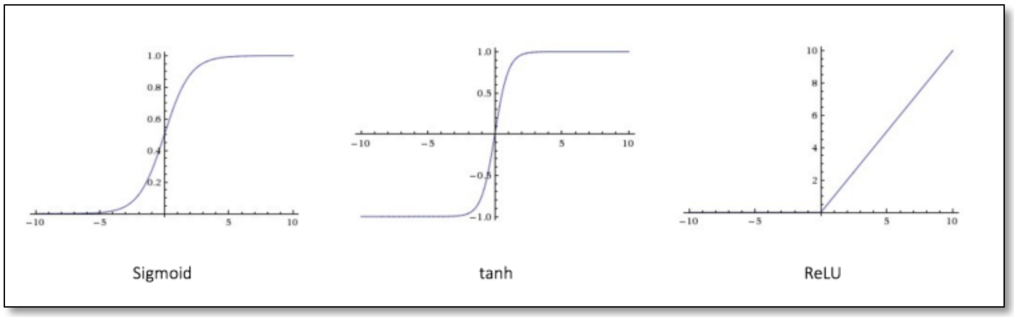

Fig 5.6: Graphical representation of sigmoid, tanh and ReLU activation functions**[35]**

A typical artificial neuron with n input dendrites can be represented by the formula that follows. The w weights allow each of the n inputs (denoted by x(i)) to contribute a greater or lesser amount to the sum of input signals. The net total is used by the activation function f(x), and the resulting signal, y(x), is the output axon **[36]**:

$$y(x) = f(\sum_{i=1}^{n} w(i)x(i)) \quad [5]$$

Now depending on the problem, Artificial Neuron can be divided into two types of models:

(a) Single-layer model (b) Multi-Layer model

Single-layer networks can be used for basic pattern classification, particularly for patterns that are linearly separable, but more sophisticated networks are required for most learning tasks. The details study of the single layer concept will be discussed in neural network.

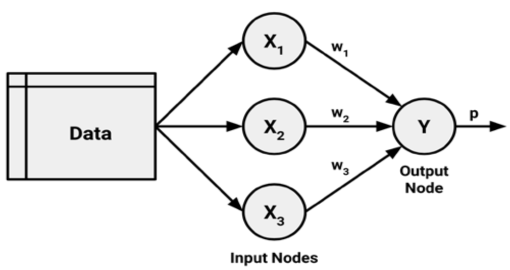

Fig 5.7: Implementation structure for single layer

5.5.2 Deep Neural Network

Deep neural network is often called as multi layer neural network. The main difference between neural network and DNN is the number of hidden layers. A multilayer network adds one or more hidden layers that process the signals from the input nodes prior to it reaching the output node.

Most multilayer networks are fully connected, which means that every node in one layer is connected to every node in the next layer. A neural network with multiple hidden layers is called a Deep Neural Network (DNN) and the practice of training such network is sometimes referred to as deep learning. As tremor occurs due to the abnormality of the neurons, so multi-layer output model (DNN) will give us a most accurate result because actual links of the hidden pattern are analyzed and output will be a PD patient and not a PD patient.

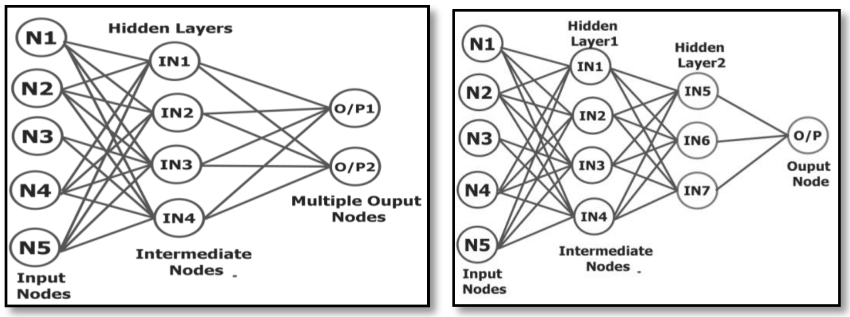

Fig 5.8: Multi layer for multi nodes output and multi layer single node output

According to our problem there are two predictive classes such as: PD patients or non-PD patients. So multi layer multi nodes output will be effective in our analysis.

5.5.3 Perceptron

The technique for the analysis of PD disorder Perceptron plays an important role in terms of performance efficacy. The details description of the Perceptron is given below:

Perceptron is feed forward system where all the direction is about to the output direction. There is a possibility for the existence of loops and if loops exist then this type of network is called recurrent network. It is treated in Perceptron that the arcs which are going to the output direction from layer **i** to layer **i+1** and all the connections (arcs) are present in the network. In this model two intermediate hidden layers exist for the filtering, one input and one output. It is used to classify linearly separable classes. This technique is often used in binary classification.

The Perceptron learning rule was originally developed by Frank Rosenblatt in the late 1950s. Training patterns are presented to the network's inputs; the output is computed. Then the connection weights $\mathbf{w_j}$are modified by an amount that is proportional to the product of

- the difference between the actual output, **y,** and the desired output, **d,** and
- Input pattern, **x**.

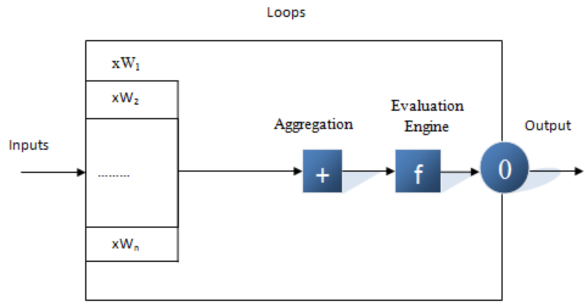

Fig 5.9: Block structure of Perceptron

The algorithm is as follows:

1. Initialize the weights and threshold to small random numbers.
2. Present a vector **x** to the neuron inputs and calculate the output.
3. Update the weights according to:

$$W\ (t+1) = W\ (t) + \eta\ (d-y)\ x \quad [6]$$

Where **d** is the desired output, **t** is the iteration number, and **eta** is the gain or step size, where 0.0 < n < 1.0

4. Repeat steps 2 and 3 until:
   - the iteration error is less than a user-specified error threshold or
   - A predetermined number of iterations have been completed.

Notice that the learning only occurs when an error is made otherwise the weights are left unchanged. This rule is thus a modified form of Hebb learning.

During training, it is often useful to measure the performance of the network as it attempts to find the optimal weight set. A common error measure or cost function used is sum-squared error. It is computed over all of the input vector/output vector pairs in the training set and is given by the equation below:

$$E = \frac{1}{2}\sum_{i=1}^{p} ||\, y^{(i)} - d^{(i)} ||^2 \quad [7]$$

Where p is the number of input/output vector pairs in the training set.

5.5.4 AutoMLP

AutoMLP is a well known disease prediction model in artificial intelligence model and AutoMLP stands for Automatic Multi Layer Perceptron. This technique is almost same like simple Perceptron model. The only difference of AutoMLP and Perceptron is in the numbers of hidden layers. In simple Perceptron there is only one hidden layer and in AutoMLP there may be one or may be more than one hidden layers exist. Like Perceptron there is one input system and one output system in AutoMLP. The mechanism is of the model is when input signals arrive to any node in the network all the signals are aggregated and form a single signal and evaluation is happened. Similarly when multi outputs are form just before the final output then at the last stage all the signals are combined and form a single output. This model can be used for the complex network whereas Perceptron is used for quite simpler network.

A MLP can be formulated for the one hidden layer as f: $R^D \rightarrow R^L$where D is the size of input vector **x** and **L** is the size of the output vector **f(x),** such that, in matrix notation:

$$f(x) = G(b^{(2)} + w^{(2)} + (s(b^{(1)} + w^{(1)}))) \quad [8]$$

with a bias vector $b^{(1)}$ and activation functions G and s

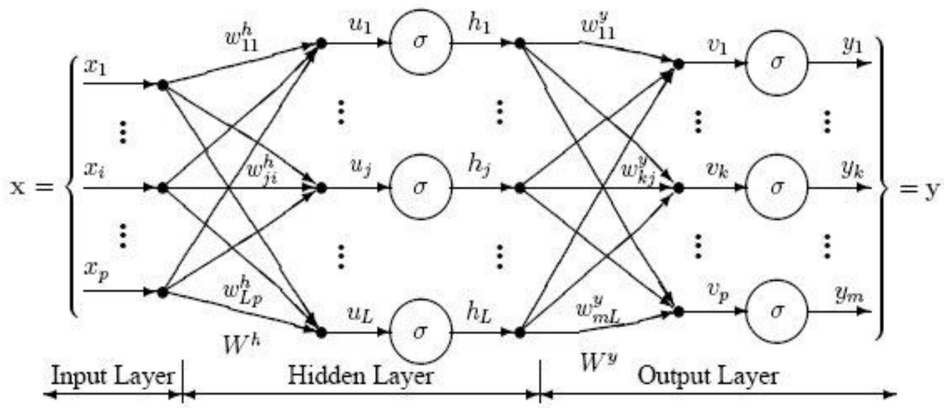

Fig 5.10: Multi layer Perceptron structure

For the multi class classification softmax activation function is used for the better response rate. But for the faster training speed tanh can be used. The output vector of the AutoMLP method is as follows:

$$f(x) = G(b^{(2)} + w^{(2)}h(x)) \quad [9]$$

Multi layer Perceptron shows better prediction than Perceptron model for the complex network. But for simpler network Perceptron is better. We will discuss the performance analysis later according to our problem.

5.6Performance Measure

The main objective of any research work is to measure the performance of the analysis. According the problem of PD we have considered artificial neural network domain for evaluation where we trained four models those are illustrated above in details. Now the efficiency of those methods and their comparison is important to get the most accurate model from those four. The most popular technique which is used to obtain the comparison is confusion matrix. The detail of the technique is given below.

Confusion Matrix: -The Deep learning model performance is estimated on the sensitivity and specificity. Sensitivity (S_N) is the percentage of true positive rate (TPR) means actually a patient have a Parkinson's disease, how correct the prediction is to determine the disease. Specificity (S_P) is percentage of true negative rate(TNR) means a patient does not have PD and how correct the prediction is determine that patient is not suffering from PD.

Sensitivity is the ratio of the correct prediction of PD suffering patient (TP) to the total of true positive (TP) and false negative (TN) value as defined below:

$$\textbf{Sensitivity } (\mathbf{S_N}) \textbf{ or Recall} = \frac{\mathbf{TP}}{\mathbf{TP+TN}}$$

Specificity is the ratio of the correct prediction of not detected PD (TN) to the total of true negative (TN) and false positive (FP) value as defined below:

$$\textbf{Specificity: } (\mathbf{S_P}) \textbf{ or Precision} = \frac{\mathbf{TN}}{\mathbf{TN+FP}}$$

A true positive is a true PD patient that is also classified as true PD suffering patient. A true negative PD patient means actually the patient is not suffering from PD and our prediction also says the same thing. False Positive is not a PD patient but prediction indicates Patient is suffering from PD and last terminology is False Negative which is a true PD patient that is wrongly classified as not a PD patient.

The accuracy (AC) is the proportion of the total number of predictions that were correct. It is determined using the equation:

$$\textbf{Accuracy} = \frac{\mathbf{TP+TN}}{\mathbf{TOTAL(TP+TN+FP+FN)}}$$

Chapter 6
Implementation & Results

In this chapter, we will explain the prediction process of PD step by step in which computational techniques are described in details. We will proveslow movement shows accurately PD disorder but there might be a chance for medium movements which can also indicate the abnormality. So there was an importance to analyze medium movements also. Since fast action performance indicates about the healthy patients so there will be no chance for the abnormality. That's why fast movements can be ignored. From the consideration of slow and medium we would be able to distinguish between the PD and non-PD patients.The whole process is divided into two stages. In first step the necessary attributes are extracted from the experiments by two methods and in the second stage, all the required attributes are taken and Mean Instantaneous Velocity (MIV) are calculated and the results are passed to the models to train them by four ANN algorithms.

According to our dataset each patient has 20 joints. All the features are not necessary for this analysis because our prediction is on the arm movement. The most effective 4 joints are hand; wrist, elbow and shoulderare indicated in the following figure 6.1:

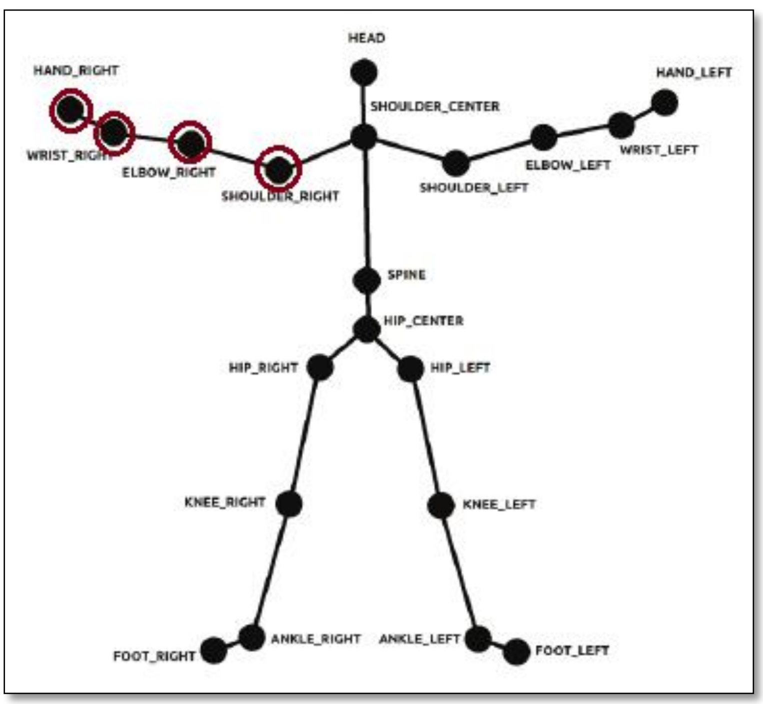

Fig 6.1: body joints captured by Kinect sensor

6.1 Attributes Selection

Before going to the prediction of the joint signals, it's very much important to find that which features can be extracted from the all body joints for the arm movement. Parkinson disorder was predicted through the upper-limb movement with different time intervals using instantaneous velocity **[37]**. So we have considered velocity as a key factor to evaluate essential features. In these work two processes has been performed for the features selections: (a) **Mean Vector Velocity (MVV):** In this process, vector velocity has been taken as the feature evaluating parameter and then calculating the mean (b) **Rank of Correlation:** it indicates strength between attributes.

6.1.1 Mean Vector Velocity

Mean vector velocity is the concept of physics where the velocity differences with small intervals of time are calculated for three dimensional vector values. Patient's data those are captured in small intervals are in three dimensions. So MVV will be calculated to get velocities of the arm movement. Lastly absolute velocities will be compared to get the essential attributes.Among 27 subjects one subject has been chosen (Here we haveconsidered subject1) for feature selection. In this subject, 4 experiments have been considered arbitrarily (considered exper1, exper2, exper5, exper11) to calculate the velocity by the equation as follows: The vector displacement of the body is given by the equation of the positional change of the vector for the patient is given by-

$$d(t) = [p(t), q(t), r(t)] \qquad [10]$$

The vector velocity equation for the one dimensional space is given by is simply the derivative of r with respect to t. In other words,

$$V(t) = \frac{\lim_{\Delta t -> 0} [d(t+\Delta t) - d(t)]}{\Delta t} = \frac{dr}{dt} \qquad [11]$$

Calculating the difference between each data point by the formula as follows:

$$A' = (a_2 - a_1), (a_3 - a_2) \ldots (a_n - a_{n-1}) \text{ as } A_1, A_2 \ldots A_{n-1} \qquad [12]$$

$$B' = (b_2 - b_1), (b_3 - b_2) \ldots (b_n - b_{n-1}) \text{ as } B_1, B_2 \ldots B_{n-1} \qquad [13]$$

$$C' = (c_2 - c_1), (c_3 - c_2) \ldots (c_n - c_{n-1}) \text{ as } C_1, C_2 \ldots C_{n-1} \qquad [14]$$

After calculating the velocity for each coordinates of all points mean of the velocities are performed to get average velocity for each joint.The below equation is used to minimize the three coordinates to one.

$$MVV = \sqrt{(Avg(x_1\ to\ x_n))^2 + (Avg(y_1 to\ y_n))^2 + (Avg(z_1\ to\ z_n))^2} \qquad [15]$$

The result which has been calculated is being represented in table as follows:

Table 6.1 MVV Analysis of Patient1

Subject1					
	Exp1	Exp2	Exp5	Exp11	
Joint No	MVV	MVV	MVV	MVV	Remarks
J1	0.016314	0.03133	0.01258	0.14682	
J2	0.014472	0.02586	0.01084	0.12928	
J3	0.009458	0.008921	0.00574	0.09769	
J4	0.008874	0.005849	0.00854	0.10509	
J5	0.004557	0.00593	0.00321	0.1032	
J6	0.002421	0.004871	0.00307	0.09262	
J7	0.004062	0.005523	0.00464	0.10119	
J8	**0.606815**	0.009669	0.00421	0.10569	
J9	0.02631	**0.041878**	**0.02336**	**0.15622**	**Shoulder**
J10	**0.120695**	**0.180256**	**0.12879**	**0.67047**	**Elbow**
J11	**0.234467**	**0.309899**	**0.21641**	**1.13042**	**Wrist**
J12	**0.262412**	**0.349702**	**0.23799**	**1.26534**	**Hand**
J13	0.014176	0.03097	0.01046	0.13244	
J14	0.013828	0.012028	0.00781	0.10672	
J15	0.006383	0.007126	0.00453	0.10453	
J16	0.080255	0.018467	0.0088	0.15954	
J17	0.017188	0.030311	0.01368	0.13817	
J18	0.011780	0.025944	0.01538	0.12585	
J19	0.00327	0.00543	0.00472	0.08742	
J20	0.002977	0.008214	0.0274	0.0746	

The Mean Vector Velocities of 20 joints have been calculated as shown in table 6.1 for four joints.These four joints are mainly responsible for hand, wrist, elbow and shoulder from all 20 joints and rest will not be considered. As we are considering the data based on the arm movement while rest of the body is in rest thenit’s quite obvious that the velocity of the arm will be higher with respect to the whole body. The above table shows that the mean vector velocity in Exp1 of J12 is highest, J11 is at 2nd highest position, J10 obtains the 3rd highest position and J8 is the lowest among four joints. On the other hand, Exp2, Exp5 and Exp11 show similar type of results where J12 shows the highest MVV, then J11, J10 and the lowest velocity is of J9. Based on the above interpretation it is confirmed that J12, J11 and J10are indicating the 3 joints of arm joints. J9 will be considered as the 4th highest velocity based on the majority.

The significant point is that shoulder is closed to body and hand is at far distance so the velocity of hand is obviously shows greater value compares than shoulder, wrist and elbow. Wrist and Elbow will be after hand, shoulder would carry least value because loosely coupled object shows higher velocity than tightly coupled if they are enforced to move. Finally we are able to conclude that J12 is assigned for hand, J11 is for wrist, J10 is for elbow and J9 is for shoulder according to their MVV values. So, with respect to the arm all the other joints will not be considered. From this analysis these 4 joints will be selected for further prediction as main features.

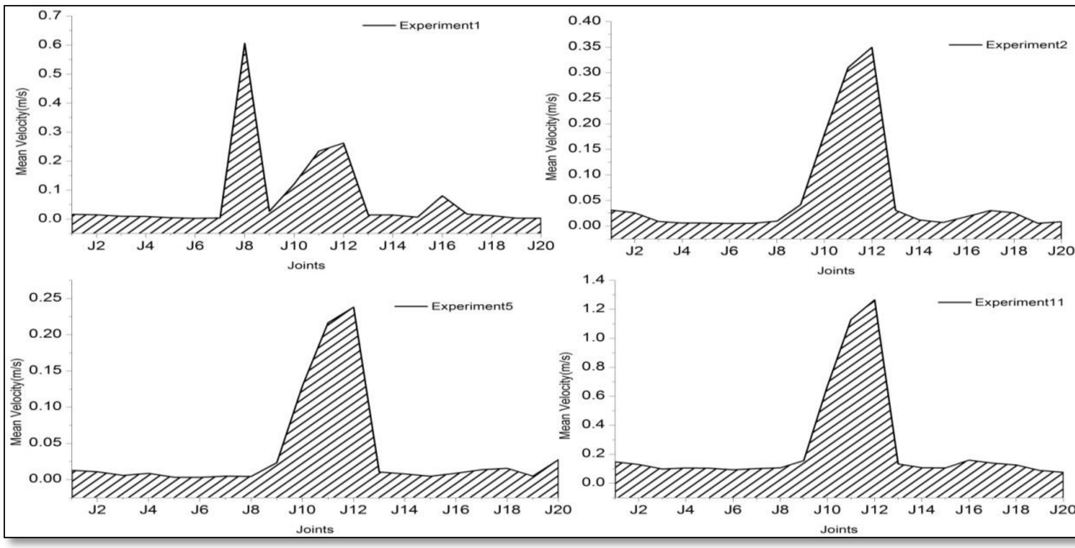

Fig.6.2. Graphical representation of Mean Vector Velocity of Patient1

From the above graph of Mean velocities with time, it is clear that out of 4 experiments there are three experiments (experiment2, experiment5 and experiment11) are showing almost symmetrical behaviors but experiment1 shows difference. So we are considering three experiments as ideal to draw a conclusion on the essential features that J9, J10, J11 and J12 these 4 joints are key features for shoulder, elbow, wrist and hand respectively.

6.1.2 Rank of Correlation

For further analysis of attribute selection, rank of correlation has been taken into the consideration and it is the well approach to identify features for machine learning problems **[38, 39]**. Correlation indicates the strength of the relationship between joints. Rank of correlation returns the vector of integers based on the strength of the correlation. This rank of correlation will be significant above the cutoff of 0.75. The cutoff indicates that the less strength of correlation will be reduced if it is greater than 0.75. We are only interested to evaluate the ranks in significant position and cutoff value is taken as 0.8. Then by summing up the ranks we get the sum of highest rank for each experiment. For calculating the ranks of the joints we have taken 4 joints together as a set (like J1-J4, J5-J8, J9-J12, J13-J16 and J17-J20). The R code is given below:

```
# ensure the results are repeatable
install.packages ("mlbench")
install.packages ("caret")
# load the library
Library (mlbench)
Library (caret)
# load the data
Dataset <- read.csv ("give your path here\\exp1.csv")
# calculate correlation matrix.
CorrelationMatrix<- cor (dataset [, 2:13]) # Change the column
number as per your dataset.
# summarize the correlation matrix
Print (CorrelationMatrix)
# find attributes that are highly corrected (ideally >0.75)
HighlyCorrelated <- findCorrelation (CorrelationMatrix, cutoff=0.5)
# Print indexes of highly correlated attributes
Print (HighlyCorrelated)
```

First of all “mlbench” and “caret” these two packages have been installed and included the corresponding libraries to get inbuilt function findcorrelation(). By reading the data file correlations have been evaluated and at last highly correlated values have been found. The results of finding the rank of correlations are given below in the table:-

Table 6.2 Rank of correlation of Patient1

Experiment1									Total Rank	Remarks
J1-J4	1	4	8	3	2	12			30	
J5-J8	10	11	8	3	6				38	
J9-J12	10	7	2	12	9	5	8	6	59	**Highest Rank**
J13-J16	2	6	11	12					31	
J17-J20	10	2	5	11					28	
Experiment2										
J1-J4	8	2	5	4	7	3	6	9	44	
J5-J8	11	1	5	7	9	3			36	
J9-J12	11	10	6	3	9	5	2		46	**Highest Rank**
J13-J16	1	2	6	10	12				31	
J17-J20	5	11	1						17	
Experiment5										
J1-J4	7	5	2	4	1	8	6	3	36	
J5-J8	4	7	2	9	8				30	
J9-J12	11	8	5	3	2	10	6	9	54	**HighestRank**
J13-J16	5	6	2	10	12				35	
J17-J20	5	2	7	8	3	11			57	

Experiment11											
J1-J4	4	1	2	5	11	8	6	7	10	54	
J5-J8	11	8	9	6	10	7	5	3	2	61	
J9-J12	11	8	5	7	10	6	4	9		60	**Highest Rank**
J13-J16	9	1	6	11	10					37	
J17-J20	4	7	5	6	1	9	10	11		53	

In the rank of correlations finding process from the above result every time we got the highest rank from the joint9 to Joint12. There is no any abnormality in the result set. Though for experiment5 J1-J4 and J9-J12 both are showing the highest rank, as per other experiments we can take J9-J12 as the highest ranker. So from the feature selection we can conclude that Joint9, Joint10, Joint11 and Joint 12 are the essential features corresponding to shoulder, elbow, wrist and hand respectively for the analysis of tremor arm movement.

6.2 Results & Analysis

The four essential attributes were found by using two methods (i) Mean Vector Velocity (ii) Rank of Correlation. In this portion we will show how the MVV and Rank of Correlation are being calculated on the real data.

In the process of MVV there are few steps are involved. In the first step the difference of displacement with each interval of time is calculated to get the small positional change. The below figure (fig 6.3) is of patient1. A small part of the dataset is visible here.

A	B	C	D	E	F	G	H	I
Time		j_1x	j_1y	j_1z		j_2x	j_2y	j_2z
0.32342		-0.33065	-0.26131	3.14731		-0.34059	-0.20054	3.20016
0.35561		-0.33059	-0.26137	3.1474		-0.34046	-0.20062	3.20027
0.39445		-0.33054	-0.26142	3.14747		-0.34034	-0.20069	3.20037
0.43532		-0.3305	-0.26147	3.14754		-0.34024	-0.20074	3.20046
0.46628		-0.33047	-0.26151	3.14759		-0.34014	-0.2008	3.20054
0.49507		-0.33047	-0.26157	3.14762		-0.34008	-0.20086	3.20059
0.52643		-0.33047	-0.26163	3.14764		-0.34003	-0.20092	3.20063
0.55691		-0.33047	-0.26168	3.14766		-0.33997	-0.20097	3.20067
0.59102		-0.33047	-0.26171	3.14767		-0.3399	-0.20101	3.2007
0.62323		-0.33047	-0.26174	3.14767		-0.33985	-0.20104	3.20072
0.65722		-0.33048	-0.26176	3.14767		-0.33979	-0.20107	3.20073
0.69034		-0.33049	-0.26178	3.14766		-0.33976	-0.20108	3.20073
0.72382		-0.33053	-0.2618	3.14764		-0.33975	-0.20109	3.20073
0.75806		-0.33056	-0.26182	3.14761		-0.33973	-0.20111	3.20072
0.79438		-0.3306	-0.26185	3.14758		-0.33971	-0.20112	3.20071

Fig: 6.3 Experiment5 dataset ofpatient1

The first step of the MVV calculation on each coordinate is given below.

$$A_1 = \frac{(-.33059 + 0.33065)}{(.35561 - .32342)} = 0.0018 \quad [16]$$

$$B_1 = \frac{(-.26137 + .26131)}{(.35561 - .32342)} = -0.00196 \quad [17]$$

$$C_1 = \frac{(3.1474 - 3.14731)}{(.35561 - .32342)} = 0.0028 \qquad [18]$$

We have shown here only for the first time interval. This process will continue on all the instances. After this step we will get n-1 (n is the number of instances) numbers of positional changes in each experiment. The next step is of this process is to calculate the mean values of vector velocities to minimize the dimensions. The evaluation of MVV is given below:

$$MVV_{(p1)} = \sqrt{(Avg(.0018\ to - .05973))^2 + (Avg(-.00196\ to - .03663))^2 + (Avg(.0028\ to - .57988))^2} \qquad [19]$$

$$= 0.01258$$

This similar way will be followed on each joint to calculate mean vector velocity. After calculating the MVV on 15 experiments, velocities are compared to get the essential attributes. On the basis of the attribute selection process we finally predict the J12 as hand feature, J11 as wrist, J10 as elbow and J9 as shoulder (described in the previous section 6.1).

In the process of training the models the very first step that we have considered is Mean Instantaneous Velocity. This MIV gives the mean positional changes of Joint9 (shoulder), Joint10 (elbow), Joint11 (wrist) and Joint12 (Hand). The only condition for normal and healthy patient is that the mean instantaneous velocities of shoulder, elbow, wrist and hand will be increasingorder and will show almost symmetrical behaviors. The velocity meanfor elbow is almost the half of the highest velocity mean (i.e. hand velocity); the velocity of shoulder is lesser than hand, wrist and elbow; elbow's velocity should be less than wrist and hand; wrist's velocity will be less than the hand velocity. According to this discussion we get a specific order for different velocities as Shoulder<Elbow<Wrist<Hand and whereverwe find this order is hampered; we can say that this condition is of abnormal patients.

Two steps are involved to calculate the MIV of *slow, medium and fast* movements: in first step we have calculated the frequency of the all instances. In the second step MIV has been calculated by the formula (eq.1) in fewer sub steps. All the steps are discussed below.

A	B	C	D	E	F	G
Time	J_9x	J_9y	J_9z	J_10x	J_10y	J_10z
0.65301	-0.19529	0.027	3.1926	-0.14886	-0.24705	3.2196
0.68486	-0.19523	0.02708	3.19241	-0.15033	-0.24929	3.21052
0.72131	-0.19515	0.02717	3.19225	-0.15078	-0.24963	3.20743
0.7519	-0.19504	0.02726	3.19209	-0.15102	-0.24932	3.20326
0.78415	-0.19496	0.02748	3.19164	-0.15128	-0.24892	3.20003
0.81975	-0.19494	0.02782	3.1909	-0.15182	-0.24757	3.19512
0.85175	-0.19493	0.02804	3.19032	-0.1519	-0.24643	3.18614
0.884	-0.19491	0.02817	3.18995	-0.1526	-0.2459	3.17526
0.92293	-0.19489	0.02829	3.18962	-0.15278	-0.24404	3.16579
0.95604	-0.19488	0.02839	3.18932	-0.15055	-0.23452	3.14484
0.9895	-0.19489	0.02857	3.18912	-0.15112	-0.22727	3.13485
1.02481	-0.19488	0.0287	3.18895	-0.14199	-0.19941	3.09118
1.0541	-0.19491	0.02882	3.18777	-0.14213	-0.19898	3.09095
1.08413	-0.19497	0.02908	3.18565	-0.1451	-0.18845	3.07613
1.12074	-0.19458	0.02948	3.18273	-0.14625	-0.1818	3.07025

Fig: 6.4features (Joints) of experiment5 in patient1

We have considered experiment5 of patient1 for the calculation of MIV. The frequency is calculated from the above dataset (fig 6.4) by following way for each instance:

$$F1= (-0.19523+0.19529)/44 = 0.0007238 \qquad [20]$$

Here 44 is the total number of instances. For all other instances, this same process will be performed to get the frequency. After getting the frequencies MIV is being found on each instance by following way:

$$MIV_{(exp5)} = \frac{\sqrt{(-.19523+.19529)^2 + (.02708-.027)^2 + (3.19241-3.1926)^2}}{0.0007238} = 5.6906 \qquad [21]$$

The above result of expeiment5 indicates the MIV of shoulder of medium movement in patient1. This above step will be applicable for other three joints as well as all the other experiments in patient1. The whole result of patient1 is given below.

Table 6.3 MIV table for Patient1

		J9(Shld)	**J10(Elb)**	**J11(Wrst)**	**J12(Hnd)**
Slow	**Exp0**	6.294	37.748	79.066	76.922
	Exp1	6.511	38.869	86.27	68.711
	Exp2	6.028	39.3179	71.5741	79.445
	Exp3	6.233	52.177	84.782	85.821
	Exp4	6.0335	43.889	82.896	88.845
Medium	**Exp5**	5.6906	29.676	61.59	70.296
	Exp6	6.076	30.674	63.666	70.696
	Exp7	14.401	43.619	67.605	80.582
	Exp8	14.073	40.493	63.111	68.122
	Exp9	6.658	43.434	78.442	89.33
Fast	**Exp10**	11.434	35.641	59.614	67.472
	Exp11	7.572	29.322	48.614	58.3
	Exp12	5.468	26.659	52.486	63.669
	Exp13	7.223	30.661	49.415	59.204
	Exp14	10.687	35.519	57.424	63.81

In the Table 6.3 the average of different experiments from Exp0 to Exp4 (for J9, J10, J11 and J12 joints) for slow movement of arm is calculated. Similarly average is calculated for medium and fast movements of arm. The average is calculated by the following way:

$$\text{Mean}_{p1,\ slow,\ J9} = (6.294+6.511+6.028+6.233+6.0335)/5 = 6.219 \quad [21]$$

All the other means will be calculated similar way as above (eq. 21). We have computed mean instantaneous velocities for each and individual experiments of the patients and compared the results according to the above velocity approach. The results of all the patients are given below (table 6.4) in the table format.

Table 6.4 Mean Instantaneous Velocity of all patients with three types of movements.

	Slow Movement				Medium Movement				Fast Movement						
Patient No.	**J9(Shld)**	**J10(Elb)**	**J11(Wrst)**	**J12(Hnd)**	**J9(Shld)**	**J10(Elb)**	**J11(Wrst)**	**J12(Hnd)**	**J9(Shld)**	**J10(Elb)**	**J11(Wrst)**	**J12(Hnd)**	**Remarks1**	**Remarks2**	**Remarks3**
P1	6.219	42.4	80.917	79.948	9.379	37.579	66.882	75.80	8.476	37.579	66.882	75.805	**PD YES**	PD NO	PD NO
P2	3.404	2.023	2.434	2.167	1.684	1.036	0.833	1.001	0.995	0.746	0.803	1.009	**PD YES**	**PD YES**	**PD YES**
P3	7.763	39.578	69.618	70.701	6.129	37.535	62.751	66.62	8.583	34.396	54.276	64.531	PD NO	PD NO	PD NO
P4	6.315	33.492	65.622	66.964	4.831	34.654	63.696	71.38	5.378	30.112	57.337	71.381	PD NO	PD NO	PD NO
P5	3.604	49.543	82.073	62.964	3.411	30.913	45.787	52.58	4.368	27.859	46.287	57.552	**PD YES**	PD NO	PD NO
P6	2.506	25.469	49.232	50.015	1.9	23.253	38.810	46.40	1.89	22.009	33.652	45.83	PD NO	PD NO	PD NO
P7	4.201	27.673	40.823	48.094	3.741	28.586	43.512	51.96	6.124	29.965	49.988	58.954	PD NO	PD NO	PD NO
P8	3.362	24.609	48.127	52.805	2.711	25.410	46.393	52.74	4.628	33.232	59.618	69.932	PD NO	PD NO	PD NO
P9	7.486	37.59	61.028	56.256	4.646	25.431	45.742	47.31	5.716	29.157	47.309	53.40	**PD YES**	PD NO	PD NO
P10	9.008	30.334	53.209	58.868	9.848	32.007	53.274	57.73	8.319	33.225	59.372	66.709	PD NO	PD NO	PD NO
P11	4.574	27.265	50.409	55.744	4.795	29.446	48.538	54.69	10.90	32.401	54.955	65.068	PD NO	PD NO	PD NO
P12	3.962	30.318	58.802	67.108	5.369	32.873	59.942	64.98	9.757	34.215	61.063	70.276	PD NO	PD NO	PD NO
P13	9.099	34.64	51.595	77.466	7.473	33.603	50.611	62.60	8.802	32.8	55.183	62.878	PD NO	PD NO	PD NO
P14	5.124	38.99	61.211	64.326	4.057	31.815	53.687	57.48	6.743	33.271	53.572	60.924	PD NO	PD NO	PD NO
P15	10.278	53.05	64.092	62.777	14.087	38.120	62.049	68.05	8.591	34.347	62.071	69.225	**PD YES**	PD NO	PD NO
P16	3.495	30.199	68.33	62.305	4.013	29.611	55.565	58.31	8.727	30.268	51.481	57.054	**PD YES**	PD NO	PD NO
P17	3.324	26.173	46.534	56,159	3.478	27.135	48.852	58.00	3.425	33.399	52.388	66.00	PD NO	PD NO	PD NO
P18	4.117	33.545	60.731	60.797	3.175	27.394	52.057	56.66	3.956	32.917	60.885	69.742	PD NO	PD NO	PD NO
P19	2.653	35.738	60.479	54.493	3.918	28.220	49.613	56.11	6.401	30.542	55.703	63.905	**PD YES**	PD NO	PD NO
P20	8.382	42.151	74.214	65.611	7.6388	34.466	62.599	62.59	13.14	35.428	52.636	67.859	**PD YES**	**PD YES**	PD NO
P21	1.601	7.937	13.665	4.993	1.179	2.9270	2.9090	2.812	1.831	3.55	7.728	4.461	**PD YES**	PD NO	**PD YES**
P22	4.494	30.95	54.57	58.451	4.118	29.402	53.036	62.13	10.11	36.153	65.153	73.135	PD NO	PD NO	PD NO
P23	5.879	31.842	61.614	65.713	4.433	22.845	49.076	55.98	15.25	37.3	70.422	80.012	PD NO	PD NO	PD NO
P24	7.091	47.691	65.672	66.109	6.289	38.791	61.525	66.80	4.754	34.357	62.484	69.415	PD NO	PD NO	PD NO
P25	8.307	41.342	86.457	66.411	8.775	33.941	59.360	63.67	8.855	32.712	56.831	64.143	**PD YES**	PD NO	PD NO
P26	7.008	32.647	59.264	66.299	6.851	29.401	50.580	57.23	7.999	30.891	54.376	64.348	PD NO	PD NO	PD NO
P27	12.048	40.531	62.079	74.694	15.471	41.417	64.417	73.48	20.36	41.06	68.382	75.603	PD NO	PD NO	PD NO

From the above table, let's take an example of slow movement of Patient16 where the MIV value for shoulder is 3.495, for elbow is 30.199, for wrist is 68.33 and for hand that is 62.305. The data statistics shows there is a problem of PD. But there is not any abnormality for medium and fast movement because from shoulder to hand MIV is in increasing order. So we can conclude from the above case, this patient is only suffering for the SM (Slow Movement). As the slow movement indicates the abnormality that's why this patient is having PD. Similar way PD detection is performed manually and we have got the results as remark1 (SM), remark2 (MM) and remark3 (FM).

The above result also shows that which portion of the arm is mostly affected by observing the values. Patient20 is suffering from the medium and slow movement of Parkinson. In these two speeds it is observed that the first contradiction occurred to the wrist and continued up to hand for both the cases. So Patient20's MAR (Most Affected Region) is wrist to Hand.

Most of the cases the problem was found in wrist to hand portion. In some cases for medium speed, problems have been detected but intensity is very low. When the problem occurs for slow speed then it obvious that the patient has PD. There may be a situation when the abnormalities may occur in medium speed only then there is a chance of primary PD. But if the problem is in slow and medium movement both then the PD intensity is high in the patient. For the first speed we neglect the chance because fast speed is only possible by healthy patients. Let's take an example of patient e.g. P20, where the effected speed SM and MM both. This is to be concluded that 20th patient is suffering badly by PD.

Depending on the instantaneous velocities values, we categories the tremor problem into some subs categories as per zone: WristToHand, ElbowToWrist, ShoulderToElbow and All Arm. It will be confirmed from these categories that which portion is affected by PD disorder. The result of categories is given below:

Table 6.5 Most Affected Region in arm

Patient No	Slow Movement	Medium Movement	Fast Movement
P1	WTH	Healthy	Healthy
P2	All Arm	All Arm	STH
P3	Healthy	Healthy	Healthy
P4	Healthy	Healthy	Healthy
P5	WTH	Healthy	Healthy
P6	Healthy	Healthy	Healthy

P7	Healthy	Healthy	Healthy
P8	Healthy	Healthy	Healthy
P9	WTH	Healthy	Healthy
P10	Healthy	Healthy	Healthy
P11	Healthy	Healthy	Healthy
P12	Healthy	Healthy	Healthy
P13	Healthy	Healthy	Healthy
P14	Healthy	Healthy	Healthy
P15	WTH	Healthy	Healthy
P16	WTH	Healthy	Healthy
P17	Healthy	Healthy	Healthy
P18	Healthy	Healthy	Healthy
P19	WTH	Healthy	Healthy
P20	WTH	WTH	Healthy
P21	WTH	WTH	WTH
P22	Healthy	Healthy	Healthy
P23	Healthy	Healthy	Healthy
P24	Healthy	Healthy	Healthy
P25	WTH	Healthy	Healthy
P26	Healthy	Healthy	Healthy
P27	Healthy	Healthy	Healthy

This result (table 6.5) shows that wrist to hand (WTH) portion affected most of the cases at the slow movement maximum times. At all the movements patient21 is suffering and most affected region is wrist to hand. This whole analysis process is carried out in two steps: (a) First step is the calculation of the Mean Instantaneous Velocity and (b) Model training.

The first step is already discussed with the calculated values (manual calculation) and conclusion was that which patient has PD disease and which patient is healthy. To justify our results the performance measurement with model training is important where we will get the accuracy of the decision and we will also able to find which model is best among fours so that we will be used to predict the PD in future. To do this all the models (Neural Network, AutoMLP, Perceptron and DNN) were trained on each of the dataset (Slow, medium and fast) which were taken based on the instantaneous values. So this is feed forward system where the output of the MIV is taken as input to the model. That's why it's a hybrid approach design to justify out prediction. After being trained the models we have analyzed the confusion matrix and the table (table 6.6) is given below:

Table 6.6 Performance measurement table of four models

Model Name		TP Instances	TN Instances	FP Instances	FN Instances	Overall Accuracy	Remarks
Neural Network	**Slow**	12	5	5	5	65%	
	Medium	0	24	1	2	88.33%	

	Fast	2	25	0	0	100%	
AutoMLP	**Slow**	14	1	9	3	55%	
	Medium	0	24	1	2	88.33%	
	Fast	2	25	0	0	100%	
Perceptron	**Slow**	17	0	10	0	63.33%	
	Medium	0	25	0	2	93.33%	Highest accuracy
	Fast	0	25	0	2	93.33%	
Deep Neural Network	**Slow**	14	7	3	3	79.60%	Highest accuracy
	Medium	0	24	1	2	88.33%	
	Fast	1	25	0	1	93.33%	

The training of the models on the PD dataset shows satisfactory results in terms of sensitivity, accuracy etc. given in the table 6.5. Our manual prediction the slow movement indicates that the worst model's accuracy (here AutoMLP) is greater than 50% and for the best it's goes up to 88.89% (here DNN). We have taken 5 to 6 trails for each of the movements under each model. As the fast movements indicates the healthy patients that we assumed manually and the same results show in the above table. So we can ignore the fast movements of the models. From the above table it is also observed that DNN shows better accuracy than all other models and the accuracy of medium movement is better than the slow though it's accuracy is higher compare than other's slow movement. For the medium movement Perceptron is the best model compare than other because it has achieved 93.33% of accuracy where as others are 88.33%. Let's take a model, slow movement of Neural Network where 17 instances (TP =12 & TN = 5) are successfully classified and 10 are unsuccessfully classified. Now the accuracy is found by the following way:

$$\text{Accuracy}_{\text{neural Net}} = \frac{12+5}{(12+5+5+5)} \times 100 = 62.9\%$$

Based on the tool the accuracy may vary ±5%. Here classifier indicates that the accuracy of the neural network for the SM is 65%. Let's take another model, slow movement of DNN shows the accuracy is 79.6%. Now accuracy of DNN is = $\frac{14+7}{(14+7+3+3)} \times 100 = 74.07\%$

So we can conclude here that to identify the slow and medium movement the best approaches are DNN and Perceptron respectively. Finally the techniques of prediction show a better throughput and efficacy for the identification of PD.

Chapter 7
Conclusion

In this work we have presented positional changes of the arm movement using Kinect data. The result shows that (1)Mean Vector velocity is well enough for the attribute selection to identify the 4 joints (Hand, Wrist, Elbow and Shoulder) of the arm movement which is cross checked with the highest rank of correlation for its perfectness, (2) Mean of Instantaneous velocity produces better prediction as the intermediate results which has been passed through the four neural network models and shows better classification rate and comparison has been made and DNN is found as the best model to predict PD,(3) Identifying the particular portion (WristToHand, ElbowToWrist, ShoulderToElbow or All arms) of the arm is affected by the PD. MIV is such an evaluating characteristic in which each instant of time changes leads to change in instantaneous velocity. Prediction indicates sensitivity and specificity in high rate where sensitivity is 70.00% and specificity is 82.35% and the accuracy of the prediction analysis is 79.63%. The results are promising, and this approach can be implemented in a human-computer interaction system for inter-active tremor diagnosis, especially for hand related disability and improvement. In this analysis, it has been asked to the healthy patients to mimic abnormality by moving slowly, however testing this approach on the original patient's dataset for hand tremor of PD disorder remains a task for future work.

Chapter 8
Limitation & Future Scope

In this study, dataset of healthy patients has been analyzed as motion detecting system and validated. The necessary step is to do assessment on the actually PD affected patients data collected from the hospital. So we will recommend that the analysis on the actual PD patients who suffer from the arm tremor as a future work. Here we have considered only 27 patients for the prediction but this is not sufficient to perfectly evaluate the classification of the patients which lead the results to uncertainty. We need more data on it. As per our knowledge, there is no such real life data are available from the Kinect sensor except this dataset which we have used here. We have just used mean vector velocity for feature selection and predicting the PD disorder respectively. Linear arm movement is not done always in real life scenario. So it might be possible that angular velocity will show more perfect result with respect to the real life. In the Kinect, 30Hz sampling frequency was used to detect voluntary action. But we know that more than 6Hz frequency shows the postural tremor. Not only 30 Hz frequency is well enough to identify the tremor. So at least more 2 frequency (may be 10Hz, 20Hz etc.) data should be considered. Theses above issues will be major concern for the future work.

References

[1] Barlow, Horace B. *"Single units and sensation: a neuron doctrine for perceptual psychology?."* Perception 1, no. 4 (1972): 371-394..https://doi.org/10.1068/p010371

[2] Bartos, Marlene, Imre Vida, and Peter Jonas. *"Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks."* Nature reviews neuroscience 8, no. 1 (2007): 45-56. http://www.nature.com/nrn/journal/v5/n4/abs/nrn1397.html

[3] Womelsdorf, Thilo, Jan-MathijsSchoffelen, Robert Oostenveld, Wolf Singer, Robert Desimone, Andreas K. Engel, and Pascal Fries. *"Modulation of neuronal interactions through neuronal synchronization."* science 316, no. 5831 (2007): 1609-1612. http://science.sciencemag.org/content/316/5831/1609/F1

[4] Wang, Gang, Ying Wan, Qi Cheng, Qin Xiao, Ying Wang, Jing Zhang, Jian-Fang Ma, Xi-Jin Wang, Hai-Yan Zhou, and Sheng-Di Chen. *"Malnutrition and associated factors in Chinese patients with Parkinson's disease: Results from a pilot investigation."* Parkinsonism & related disorders 16, no. 2 (2010): 119-123. http://doi.org/10.1016/j.parkreldis.2009.08.009

[5] Stern, Yaakov, Richard Mayeux, Jeffery Rosen, and Joyce Ilson. *"Perceptual motor dysfunction in Parkinson's disease: a deficit in sequential and predictive voluntary movement."* Journal of Neurology, Neurosurgery & Psychiatry 46, no. 2 (1983):145-151. http://dx.doi.org/10.1136/jnnp.46.2.145

[6] Hirsch, Etienne, Ann M. Graybiel, and Yves A. Agid. *"Melanized dopaminergic neurons are differentially susceptible to degeneration in Parkinson's disease."* Nature 334, no. 6180 (1988): 345-348. http://www.nature.com/nature/journal/v334/n6180/abs/334345a0.html

[7]http://www.orionpharma.co.uk/Products-and-Services-Orion/Parkinsons-disease/10-facts-about-Parkinsons-disease/

[8]Berardelli, Alfredo, J. C. Rothwell, P. D. Thompson, and M. Hallett. *"Pathophysiology of bradykinesia in Parkinson's disease."* Brain 124, no. 11 (2001): 2131-2146.DOI: https://doi.org/10.1093/brain/124.11.2131

[9]Marjama-Lyons, Jill, and William Koller. *"Tremor-predominant Parkinson's disease.Approaches to treatment."* Drugs & aging 16, no. 4 (2000): 273-278. **DOI:** 10.2165/00002512-200016040-00003

[10]Rahimi, Fariborz, Quincy J. Almeida, David Wang, and FarrokhJanabi-Sharifi. *"Tremor suppression orthoses for parkinson's patients: a frequency range perspective."*In *Engineering in Medicine and Biology Society, 2009.EMBC 2009. Annual International Conference of the IEEE*, pp. 1565-1568. IEEE, 2009.

[11]Anouti, Ahmad, and William C. Koller. *"Tremor disorders.Diagnosis and management." Western journal of medicine* 162, no. 6 (1995): 510.

[12]Yadav, Geeta, Yugal Kumar, and GadadharSahoo. *"Predication of Parkinson's disease using data mining methods: A comparative analysis of tree, statistical and support vector machine classifiers."* In Computing and Communication Systems (NCCCS), 2012 National Conference on, pp. 1-8. IEEE, 2012.

[13]Sanchez-Ramos, Juan, Dacy Reimer, Theresa Zesiewicz, Kelly Sullivan, and Paul A. Nausieda. *"Quantitative analysis of tremors in welders."* International journal of environmental research and public health 8, no. 5 (2011): 1478-1490.

[14]Elgendi, Mohamed, FlavienPicon, Nadia Magnenat-Thalmann, and Derek Abbott. *"Arm movement speed assessment via a Kinect camera: a preliminary study in healthy subjects."* Biomedical engineering online 13, no. 1 (2014): 88.

[15]Elgendi, Mohamed, FlavienPicon, and Nadia Thalmann. *"Towards arm tremor diagnosis using Kinect."*

[16]Ben-Pazi, H., H. Bergman, J. A. Goldberg, N. Giladi, D. Hansel, A. Reches, and E. S. Simon. *"Synchrony of rest tremor in multiple limbs in Parkinson's disease: evidence for multiple oscillators."* Journal of neural transmission108, no. 3 (2001): 287-296.

[17]Ťupa, Ondřej, AlešProcházka, OldřichVyšata, Martin Schätz, Jan Mareš, Martin Vališ, and VladimírMařík. *"Motion tracking and gait feature estimation for recognising Parkinson's disease using MS Kinect."* Biomedical engineering online 14, no. 1 (2015): 97.

[18]Alam, MdNafiul, Benjamin Johnson, Jeffrey Gendreau, KouhyarTavakolian, Colin Combs, and Reza Fazel-Rezai. *"Tremor quantification of Parkinson's disease-a pilot study."*In Electro Information Technology (EIT), 2016 IEEE International Conference on, pp. 0755-0759.IEEE, 2016.

[19]Morris, Meg E. "*Movement disorders in people with Parkinson disease: a model for physical therapy." Physical therapy* 80, no. 6 (2000): 578-597. https://doi.org/10.1093/ptj/80.6.578

[20]Geetha, A., and G. M. Nasira. "*Artificial Neural Networks' Application in Weather Forecasting—Using RapidMiner." International Journal of Computational Intelligence and Informatics* 4, no. 2 (2014): 177-182. http://www.periyaruniversity.ac.in/ijcii/issue/Vol4No3December2014/IJCII%204-1-152.pdf

[21]https://rapidminer.com/wp-content/uploads/2014/10/RapidMiner-5-Operator-Reference.pdf
[22] MousaviHondori, Hossein, and Maryam Khademi.*"A review on technical and clinical impact of microsoftkinect on physical therapy and rehabilitation."* Journal of Medical Engineering 2014 (2014).

[23] W. Klosgen, J. Zytkow, *"Handbook of data mining and knowledge discovery",* Oxford University Press, New York, 2002.

[24] Zhang, Zhengyou. *"Microsoft kinect sensor and its effect."* IEEE multimedia19, no. 2 (2012): 4-10. 10.1109/MMUL.2012.24

[25] Oikonomidis, Iason, NikolaosKyriazis, and Antonis A. Argyros. *"Efficient model-based 3D tracking of hand articulations using Kinect."*In BmVC, vol. 1, no. 2, p. 3. 2011. http://users.ics.forth.gr/~argyros/res_kinecthandtracking.html

[26] Han, Jungong, Ling Shao, Dong Xu, and Jamie Shotton. *"Enhanced computer vision with microsoftkinect sensor: A review."* IEEE transactions on cybernetics 43, no. 5 (2013): 1318-1334. Doi:10.1109/TCYB.2013.2265378

[27] Elgendi, Mohamed, FlavienPicon, Nadia Magnenat-Thalmann, and Derek Abbott. *"Arm movement speed assessment via a Kinect camera: a preliminary study in healthy subjects."* Biomedical engineering online 13, no. 1 (2014): 88; https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-13-88

[28] Catuhe, David. Programming with the Kinect for Windows software development kit.Pearson Education, 2012.https://books.google.co.in/books?hl=en&lr=&id=tKlCAwAAQBAJ&oi=fnd&pg=PT10&dq=Catuhe,+David.+Programming+with+the+Kinect+for+Windows+software+development+kit.+Pearson+Education,+2012.&ots=ynoa0tBLBp&sig=FD8JWOIgL0Eo0NG6kGxCDFm6ty8#v=onepage&q&f=false

[29] Van den Broeck, Jan, SolveigArgeseanu Cunningham, Roger Eeckels, and KobusHerbst.*"Data cleaning: detecting, diagnosing, and editing data abnormalities."*PLoS Med 2, no. 10 (2005): e267. https://doi.org/10.1371/journal.pmed.0020267

[30] Craig, Andrew, Olivier Cloarec, Elaine Holmes, Jeremy K. Nicholson, and John C. Lindon. *"Scaling and normalization effects in NMR spectroscopic metabonomic data sets."* Analytical chemistry 78, no. 7 (2006): 2262-2267 doi:10.1021/ac0519312

[31] Schnitzler, Alfons, Christian Münks, Markus Butz, Lars Timmermann, and Joachim Gross. *"Synchronized brain network associated with essential tremor as revealed by magnetoencephalography."* Movement Disorders 24, no. 11 (2009):1629-1635. https://www.researchgate.net/publication/227797652_Synchronized_Brain_Network_Associated_with_Essential_Tremor_as_Revealed_by_Magnetoencephalography

[32] Schnitzler, Alfons, Christian Münks, Markus Butz, Lars Timmermann, and Joachim Gross. *"Synchronized brain network associated with essential tremor as revealed by magnetoencephalography."* Movement Disorders 24, no. 11 (2009):1629-1635. https://www.researchgate.net/publication/227797652_Synchronized_Brain_Network_Associated_with_Essential_Tremor_as_Revealed_by_Magnetoencephalography

[33] Palay, Sanford L., and George E. Palade. *"The fine structure of neurons."* The Journal of biophysical and biochemical cytology 1, no. 1 (1955): 69. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2223597/pdf/69.pdf

[34] Agatonovic-Kustrin, S., and R. Beresford. *"Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research."*Journal of pharmaceutical and biomedical analysis 22, no. 5 (2000): 717-727. http://doi.org/10.1016/S0731-7085(99)00272-1

[35] Nam, Jinseok, Jungi Kim, EneldoLozaMencía, IrynaGurevych, and Johannes Fürnkranz. "Large-scale multi-label text classification—revisiting neural networks." In *Joint european conference on machine learning and knowledge discovery in databases*, pp. 437-452. Springer, Berlin, Heidelberg, 2014.

[36]Ito, Yoshifusa. *"Representation of functions by superpositions of a step or sigmoid function and their applications to neural network theory."* Neural Networks 4, no. 3 (1991): 385-394.

[37] Almeida, Quincy J., Laurie R. Wishart, and Timothy D. Lee. *"Bimanual coordination deficits with Parkinson's disease: the influence of movement speed and external cueing."* Movement Disorders 17, no. 1 (2002): 30-37. https://www.researchgate.net/publication/11526239_Bimanual_coordination_deficits_with_Parkinson's_disease_The_influence_of_movement_speed_and_external_cueing

[38] 1 Hall, Mark A. *"Correlation-based feature selection of discrete and numeric class machine learning."* (2000). http://hdl.handle.net/10289/1024

[39] Yu, Lei, and Huan Liu. *"Feature selection for high-dimensional data: A fast correlation-based filter solution."*In ICML, vol. 3, pp. 856-863. 2003. http://www.aaai.org/Papers/ICML/2003/ICML03-111.pdf

List of Publications

1. Sagarjit Dash and Ashutosh Mishra, "Detection of Parkinson Disorder using Kinect arm movement - A computational based approach" ***2nd International SCI Index Conference on Big Data, Cloud Computing, and Data Science Engineering (BCD 2017)*** Hamamatsu, Japan [Accepted but not attended].

2. Sagarjit Dash and Ashutosh Mishra, "Detection of Parkinson Disorder using Kinect arm movement - A computational based approach", ***1st International Conference of Advanced Computational and Communication Paradigms 2017.*** [Communicated]

3. Sagarjit Dash and Ashutosh Mishra, "A comparative analysis for detecting Parkinson Disease using Kinect Sensor", ***3rd International Conference on Next Generation Computing Technologies 2017*** [Communicated].

MTech dissertation

ORIGINALITY REPORT

%5	%	%5	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ťupa, Ondřej, Aleš Procházka, Oldřich Vyšata, Martin Schätz, Jan Mareš, Martin Vališ, and Vladimír Mařík. "Motion tracking and gait feature estimation for recognising Parkinson's disease using MS Kinect", BioMedical Engineering OnLine, 2015. Publication	<%1
2	H. Ben-Pazi. "Synchrony of rest tremor in multiple limbs in Parkinson's disease: evidence for multiple oscillators", Journal of Neural Transmission, 03/20/2001 Publication	<%1
3	"Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems", Springer Nature, 2009 Publication	<%1
4	European Business Review, Volume 19, Issue 4 (2007-06-27) Publication	<%1
5	Ahmed, Mansoor, Anirudh Sriram, and Sanjay	<%1