A KNOWLEDGE BASE DATA MINING BASED ON PARKINSON'S DISEASE

BY

Md. Redone Hassan

ID:152-15-547

Md. Aminul Islam

ID:152-15-537

AND

S.K.Obidul Kadir

ID:152-15-528

Bachelor of Science in Computer Science and Engineering

Supervised By

Sheikh Abujar

Lecturer

Department of CSE
Daffodil International University

Co-Supervised By

Md. Tanvir Rahman

Senior Lecturer

Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

APPROVAL

This Project/internship titled **''A KNOWLEDGE BASE DATA MINING BASED ON PARKINSON'S DISEASE''**, submitted by Md. Redone Hassan, ID No: 152-15-547 and Md. Aminul Islam ,ID No:152-15-537 and S.K.Obidul Kadir ,ID No:152-15-528 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on April 6th ,2019

BOARD OF EXAMINERS

_____

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. S M Aminul Haque — **Internal Examiner**
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Saif Mahmud Parvez — **Internal Examiner**
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Sheikh Abujar , Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

_____

Sheikh Abujar
Lecturer
Department of CSE Daffodil International University

Co-Supervised by:

_____

Md. Tanvir Rahman

Senior Lecturer
Department of CSE Daffodil International University

Submitted by:

_____

MD.Redone Hassan
ID: 152-15-547
Department of CSE
Daffodil International University

MD.Aminul Islam
ID:152-15-537
Department of CSE
Daffodil International University

_____

S.K.Obidul Kadir
ID: -152-15-528
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

We are very much thankful to our advisor Sheikh Abujar, Lecturer, Department of CSE Daffodil International University who guide us nicely to do this work. His patience and mentality always give us a spirit that how can we complete this work. He provides all type of resource related to this. So we can get the actual theme to fulfill our work

.

We also give thanks to our Co-Advisor Md. Tanvir Rahman Senior Lecturer, Department of CSE Daffodil International University who always suggest us about our research. Like to thank Dr. S.M. Aminul Haque, Assistant Professor and Associate Head, Department of CSE he is one of the good people and guide us nicely for doing research oriented work. We would like to express our heartiest gratitude to Head, Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to special thanks to our team members always co-operated with this work.

Finally, we thank our family for their guidance and gives us support to complete this work.

ABSTRACT

The approaches of detecting Parkinson's disease (PD) in human body from voice data using Classification techniques apply three different algorithms for finding the growth rate of Parkinson's disease. Support vector machine (SVM), K-nearest neighbor (K-NN) and Decision tree (DT) are applied to detect the progression rate of Parkinson's disease (PD) by using Unified Parkinson's disease rating scale (UPDRS) and Hoehn &Yahr scale (HYS) . Unified Parkinson's disease rating scale (UPDRS) deals with motor fluctuations and change over voice after certain period and that can measure the people affected by Parkinson's disease and healthy people. Hoehn & Yahr scale (HYS) are measures the symptoms which are related to progression of Parkinson's disease (PD) in human body. Classifier algorithms used to detect the factors and symptoms which are related to progression of Parkinson's disease (PD) in human body using voice data. The algorithms can detect Parkinson's disease (PD) by several approaches which criteria are related to progression of Parkinson's disease (PD) in human body. From the distinctions of all algorithms measures which algorithm give the best accuracy for several approaches to diagnosis Parkinson's disease (PD) and risk factors of had Parkinson's disease (PD) in human body.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES **PAGE**

List of abbreviation

SVM	Support Vector Machine
KNN	K-nearest neighbor
PD	Parkinson's Disease
DT	Decision tree
UPDRS	Unified Parkinson’s disease rating scale
HYS	Hoehn&Yahr scale
TCS	Transcranial sonography

CHAPTER 1

INTRODUCTION

1.1 Motivation

Parkinson's disease is a progressive neuro-degenerative disease which affects in our central nervous system. Our brain consists 86 billion neurons. Every neuron has his own functionality. In our mid brain have an area called substantia nigra . In this have a neuron called dopamine. Reducing dopamine neurons is the causes of being Parkinson's disease in human body. Approximately 4-6 million people are affected by Parkinson's disease in every year.

In Bangladesh, Every year 1600 people are die by Parkinson's disease and every year the growth rate of this disease are getting larger. The life expectancy of PD affected people is 5-10 years. But the pooled report on clinical research says that [16] accuracy of diagnosis Parkinson's disease is only 80.6%. Till today has no recovery of [3] Parkinson's disease(PD).In the early stage diagnosis parkinson's disease by using some brain scans like MRI, fMRI, SPECT etc. But it can't give the best accuracy.

So for that we are come up to work with it by different approaches to improving the diagnosis of PD research that [3] in the early stage can diagnosis Parkinson's disease(PD) by [20] Support vector Machine (SVM) , K-nearest neighbor (KNN) and Decision tree (DT). So that for medical research it can very helpful that which approaches is best for diagnosis PD by using classification technique.

1.2 Literature review

[15] Parkinson’s disease is a progressive degenerative disease which was first invented by James Parkinson in 1817 [1]. In the world, 5-6 millions of people years affected by this disease. It has some symptoms like shaking, rigidity, tremor and so on. The [17] term “Parkinson’s disease” was coined in 1865 by William Sanders and later popularized by French neurologist [2].

[12] The accuracy of diagnosis of Parkinson's disease using machine learning technique is limited till now. At the beginning stage when cardinal systems not conclusive then CCT or MRI structures features do not provide the characteristics of Parkinson's disease [1]. Functional neuro imagine using PET and SPECT is very efficient to diagnosing first [8] symptoms or sign of Parkinson's disease. But these processes are highly expensive not board use in work and unavailable. So for this [8] transcranial sonography (TCS) is proved to be very useful and around 90% of patients show [8] hyperechogenicity of substantia nigra (SN) on TCS [1]. It allowing the distinction between mildly affected people with healthy people. Also additionally specific ultrasound features for some forms of secondary Parkinson's disease. So TCS are recommended tools to facilitate the [12] diagnosis of Parkinson's disease. PD diagnosis has been difficult and doctors treated to work with some symptoms to predict a result [1].The advanced technology providing voice tools to diagnosis PD from daily activities. The audio data translate to diagnostic tools to figure out [19] Parkinson's disease-related data to the diagnosis of Parkinson's disease [2].

By using voice data the machine learning classification algorithms and neural network algorithms deal with data to perform to diagnosis [2] Parkinson's disease. From the overall theme try to explore the classification algorithms that the showing interact to different factors to build the [2] symptoms of Parkinson's disease in the human body [3]. Working with [6] Support vector machine(SVM), K-nearest neighbor and decision tree with voice data identify relationship to had a Parkinson's disease in the human body [4]. The system also recommends which way is better to diagnosis Parkinson’s disease.

And in this disease mostly affected people are male. Doing Parkinson’s disease diagnosis several Parkinson’s are used in previously doing research.implement the system using machine learning. This system works on PD symptoms and identifies [2] which are related to occur in Parkinson’s

disease in the human body. It collects voice data that which people are most affected by

Parkinson's disease that can identify by this [3]. After all, building a specific way to diagnosis the factors related to had Parkinson's disease in the human body.

1.3 Thesis Orientation

In the overall thesis, Chapter 2 Background Studies , Chapter 3 Introduced Systems , Chapter 4 Results , Chapter 5 Discussion , Chapter 6 Conclusions the thesis.

CHAPTER 2

BACKGROUND STUDY

2.1 Parkinson's disease

Parkinson's disease is a progressive degenerative disorder which affects the motor systems of the brain. By this disease also impacts on non-motor systems. In the midbrain, an area called substantia nigra has to produce dopamine neuron. Decreasing the protein and Lewy bodies of dopamine neuron be a reason for having Parkinson's disease in the human body.

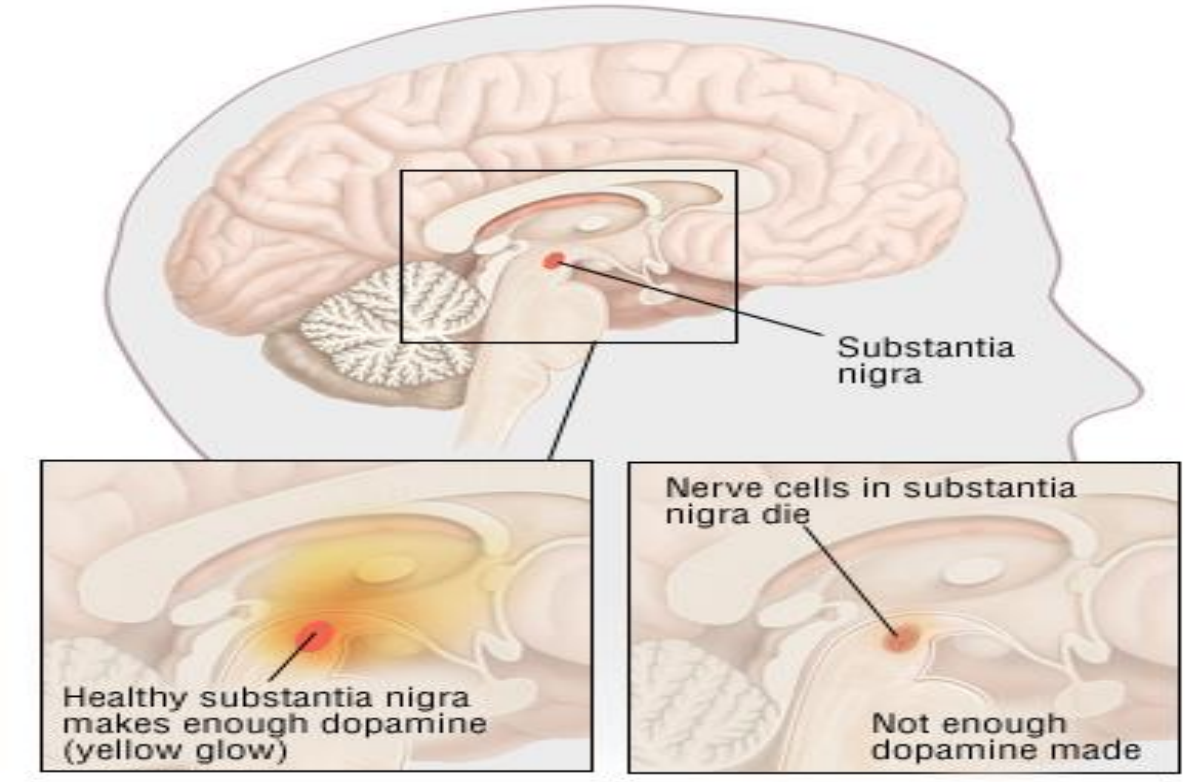

Figure 2.1: Parkinson's disease (PD) [1]

2.2 Symptoms and factors of PD

[11] Parkinson's disease is a progressive degenerative disease. It sustains in the human brain for

many times. So that it had some symptoms that very slowly change in terms of time. At the very

first stage Shaking , rigidity, tremor, slowness of movement and reckless movement in the sleep, difficulty with walking are occurs. At the advanced stage, dementia [3] is the most common

symptoms in Parkinson's disease. sleeping, emotional problems, thinking and behavioral problems may occur to have Parkinson's disease in the human body. Its process is very slow so it can grow up in the human body for many years. So some symptoms can exist in the body for many years.

2.3 Impact of PD

Parkinson's disease gives the so many bad impact to the people. People can't memorize any things clearly and can't make sense about the responsibility like in family or any work place. Another bad impact is always want to be lonely so that the relationship with other people will fall down day by day. The detected people suffer emotional problems and can't be energetic doing any kinds of works.

2.4 Data mining in Neurological field

Data mining is the technique of finding patterns from large datasets and extract knowledge. Using Data mining techniques can predict any probable decision. In the neurological field, Data mining playing vital role because neurological disease is progressive degenerative disease which impact on neurons into the brains. Data mining works several symptoms which are related and make a prediction from the data. So a vast use of Data mining algorithms in this field working qualitive research and diagnosing neurological disease.

2.4.1 Uses of data mining algorithms

Data mining algorithms doing calculations from large data and creating a model based on data.The algorithms finding parameter from the data and these parameter are predict on overall data to find the model.

The Data mining algorithms are dealing with different approaches to solve the problems.Every algorithms had the own working procedure to extract knowledge from data.Classification algorithm Anticipate at least one discrete factors, in view of different characteristics in the dataset. Regression algorithm predict at least one nonstop numeric factors from the dataset.Segmentation algorithm partition information into gatherings, or bunches, of things that have comparative

properties. Association algorithm discover relationships between's various qualities in a dataset.The most widely recognized utilization of this sort of calculation is for making affiliation rules, which can be utilized in a market container examination.Sequence Algorithm calculations outline visit successions or scenes in information, for example, a progression of snaps in a site, or a progression of log occasions going before machine support. So every algorithms had different approaches to find patterns from the data.

2.4.2 Knowledge gathering procedure

In the research to work parkinson's disease had to learn many resoures that the procedure of diagnosis parkinson's disease by clinically. But seen that working with it is very hard because there are many symptoms and factors associated to build [14] parkinson's disease in human body. In machine learning research many data mining algorithms use to figure out the symptoms finding. So for Knowledge gathering we try make questions about parkinson's disease collecting data and work with it. Then making discussion about the working procedure and finding problems and try to find out the way to solve it.

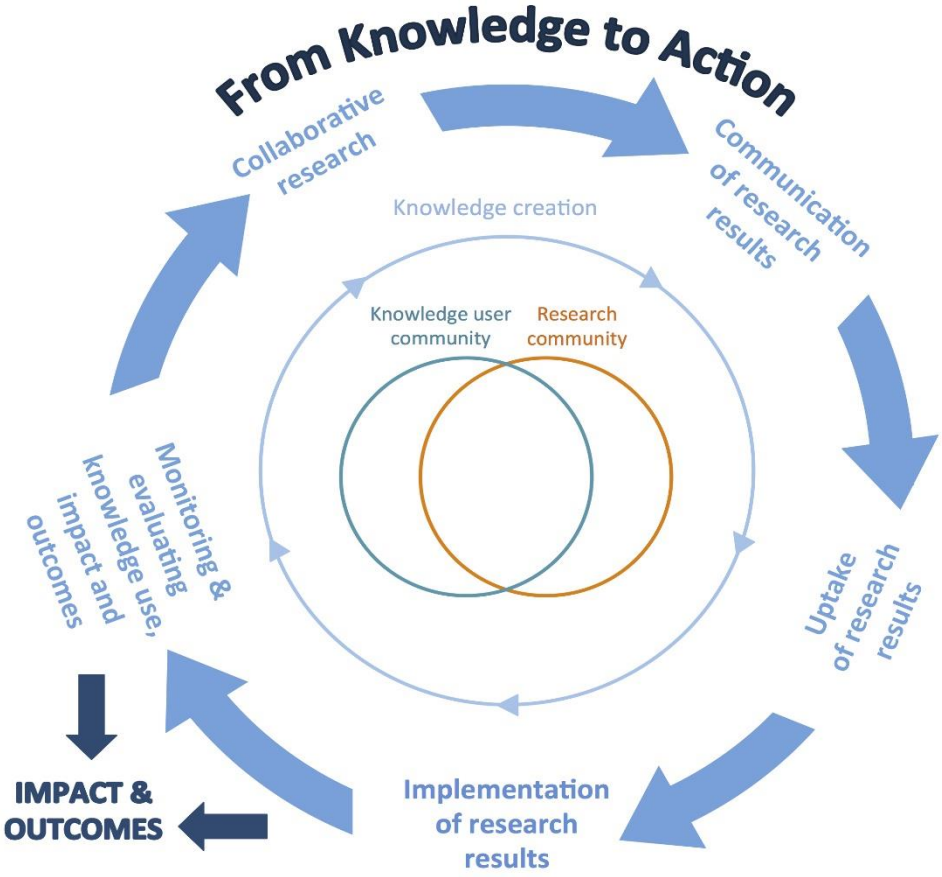

Figure 2.4.2: Knowledge exploration and exchanges [2]

CHAPTER 3

INTRODUCED SYSTEM

3.1 Flow of Work

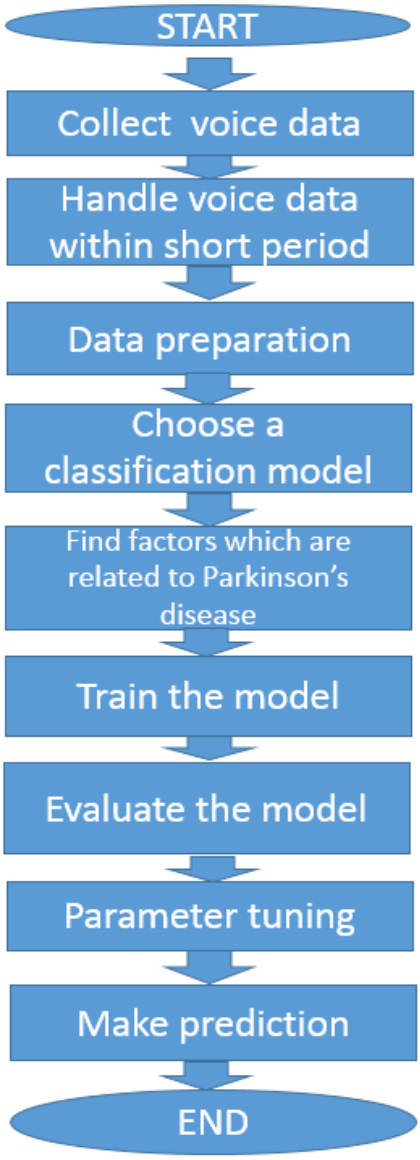

Figure 3.1: Process workflow of the work

Collect voice data:

In this research collect voice data from UCI dataset repository. The voice data consist Demographical information,clinical information e.t.c which is used find out the scaling of UPDRS scoring.

Handle voice data within short period:

For very short time handle the dataset because of in short period can't measure some time to detect parkinson's symptoms.

Data preparation:

Preparing data from large dataset is too much hard to analyze.The inconsistent data had to prepare for mining or calculating the result.

Choose the model:

The classification model is use for appling to predict the result.Three datamining model is to diagnosis parkinson's disease.

Find the factors [2] which are related to parkinson's disease:

To [2] apply the algorithms finding the related symptoms and factors that the directly and indirectly related to had parkinson's disease in the human body.Because every algorithm gives the different types of results.

Train the model:

In applying every algorithms should had to train up. Because inconsistent data and missing data gives bad output result.So for the all criteria had to train up each and every model to get the better result.

Parameter tuning:

Parameter tuning is the way of finding the best parameter that will gives the best output from dataset.The way to diagnosis parkinson's disease from the overall parameter.

Make prediction:

The fulfill of criteria finally making predict the model that the accuracy rate of each algorithms and finally generating the output.

3.2 Analyzing and processing data of PD and non-PD dependent:

The apply of classification algorithms had to analyze the PD and non-PD dependent data from datasets. Because many data can't find the result.The dataset consist the behavior of both data so the inconsistent data can't find the actual output or making error to prediction the result.So find appropriate data to getting output from PD and non-PD dependent data.In every operation had to filter out the dataset to getting the result.

3.3 Data acquirements

In the dataset had both types of data like numerical data and categorical data. So for different types of operation had to prepare the dataset. Because every algorithm can't handle both types of data.

There are some samples of dataset is given below:

	Demographic information	Demographic information	Clinical information	Clinical information
Participant code	Age (years)	Gender	Positive history of Parkinson disease in family	Age of disease onset (years)
PD01	58	F	No	56
PD02	68	F	No	67
PD03	68	M	No	67
PD04	75	M	No	73
PD05	61	M	Yes	60
PD06	58	M	No	58
PD07	79	M	No	78
PD08	59	F	No	57
PD09	73	M	No	72
PD10	66	M	No	65
PD11	71	M	No	70
PD12	37	F	No	32
PD13	43	M	No	37
PD14	70	M	No	69
PD15	75	F	No	73

Figure 3.3: [11] Demographic and clinical information of Parkinson's disease

Speech examination: speaking task of reading passage	Speech examination: speaking task of reading passage	Speech examination: speaking task of reading passage	Speech examination: speaking task of reading passage	Speech examination: speaking task of reading passage	Speech examination: speaking task of reading passage
Entropy of speech timing (-)	Rate of speech timing (-/min)	Acceleration of speech timing (-/min2)	Duration of pause intervals (ms)	Duration of voiced intervals (ms)	Gaping in-between voiced intervals (-/min)
1.564	354	6.05	146	264	58.65
1.564	340	27.52	173	253	48.26
1.55	211	11.97	377	322	47.54
1.519	140	-2.49	360	663	13.72
1.543	269	6.72	211	328	42.9
1.553	317	24.19	186	286	43.83
1.563	269	2.47	214	310	60.3
1.568	338	-4.36	145	272	61.31
1.552	374	11.1	117	265	52.5
1.549	281	16.09	213	323	28.8
1.537	234	21.82	198	414	26.7

Figure 3.3.1: Speech examination of reading passage

Here are some samples of dataset. The dataset also contains the symptoms and two scaling measurement scale and by this is use to find out the Parkinson's disease in the human body.

3.4 Data Pre-processing

Data preprocessing is used to removing the inconsistent and noisy data from dataset. In the data preprocessing data wasn't meaningful but to find the patterns from dataset had to prepare and customize data in the dataset. So it is must necessary to do without preprocessing can't find the appropriate output.

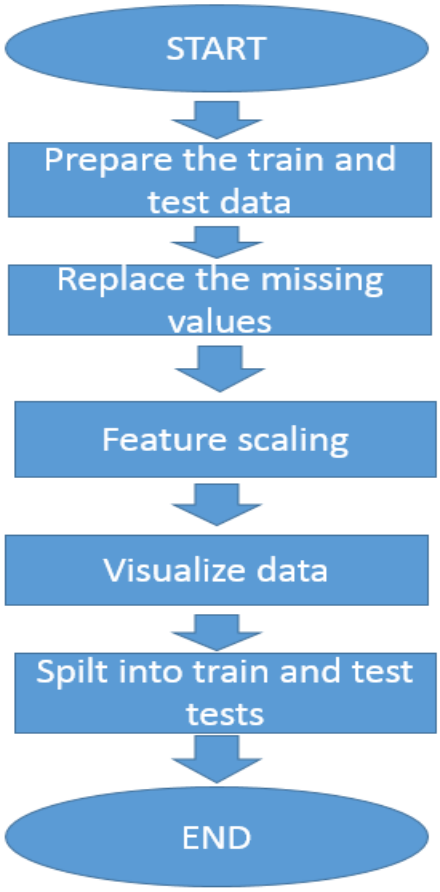

Figure 3.4: Data preprocessing

So uses some procedure like feature scaling,replaceing missing data and then finally visualize the data to making prediction from the dataset.

3.5 Analyzing Data

From the overall operation analyzing the data that the relevant to build parkinson's disease in the human body. Then adding the factors that are related to progression of parkinson's disease.

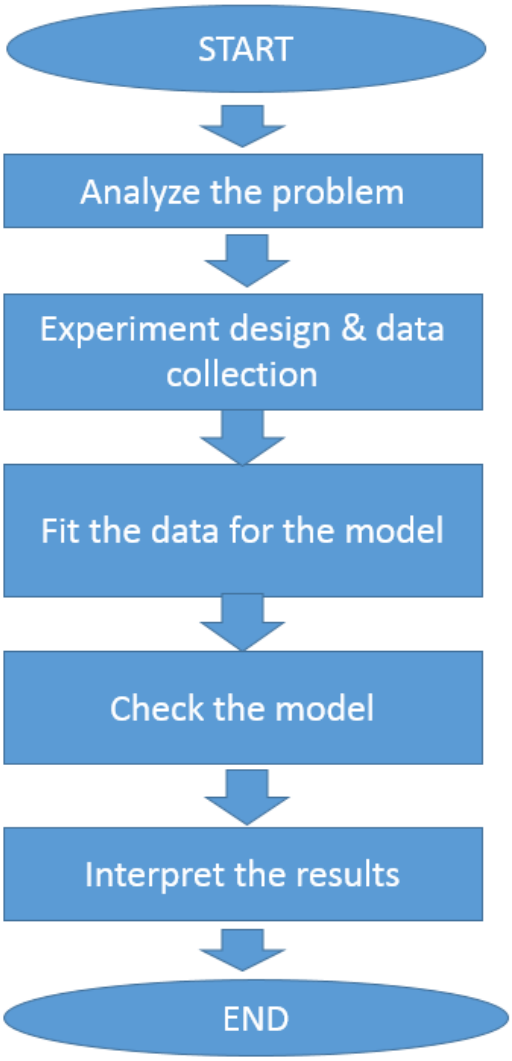

Figure 3.5: Analyze the data

First analyze the problem from the data build a experimental design and collect the data then fit the data for the model to predict the result. Then finally check the model and interpret the result from the data.

CHAPTER 4

RESULTS

In the result section showed the results and output of [6] support vector machine (SVM),K-nearest neighbor and Decision tree algorithms output figure that are achieving from the train data.From the above of these algorithms the decision tree provides the highest accuracy of 55%. On the other hand, Support vector machine and K-nearest neighbor giving the accuracy up to 50 %.The result are finding from different factors which are [14] related to build parkinson's disease in the human body.In the paper working with 30 early affected parkinson's disease people, 50 are eye rapid disorder people and 50 healthy people.

4.1 Tabular Representation of area of PD dependent

In this paper , [23] Support Vector Machine(SVM),k-nearest neighbors and Decision Tree are applied on the dataset.

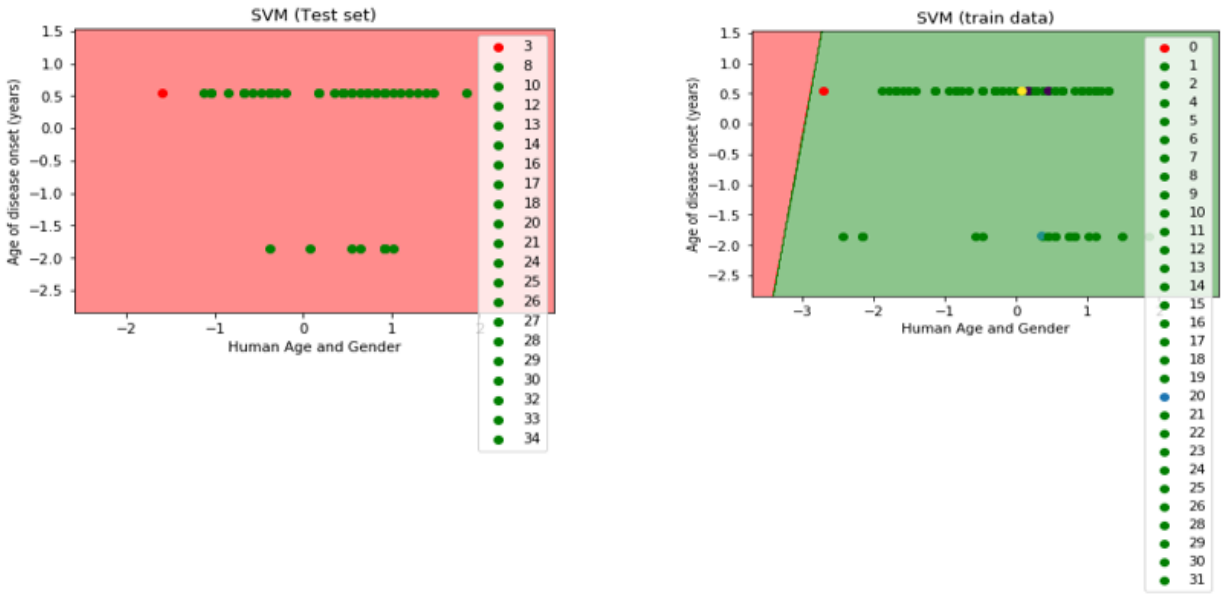

Figure 4.1: Comparing Human age and Gender with Age of disease onset(years) using SVM

On the dataset of age of disease are predicted from human age and gender. In this section which gender and when the disease is first affected are classifying Using Support Vector Machine(SVM).

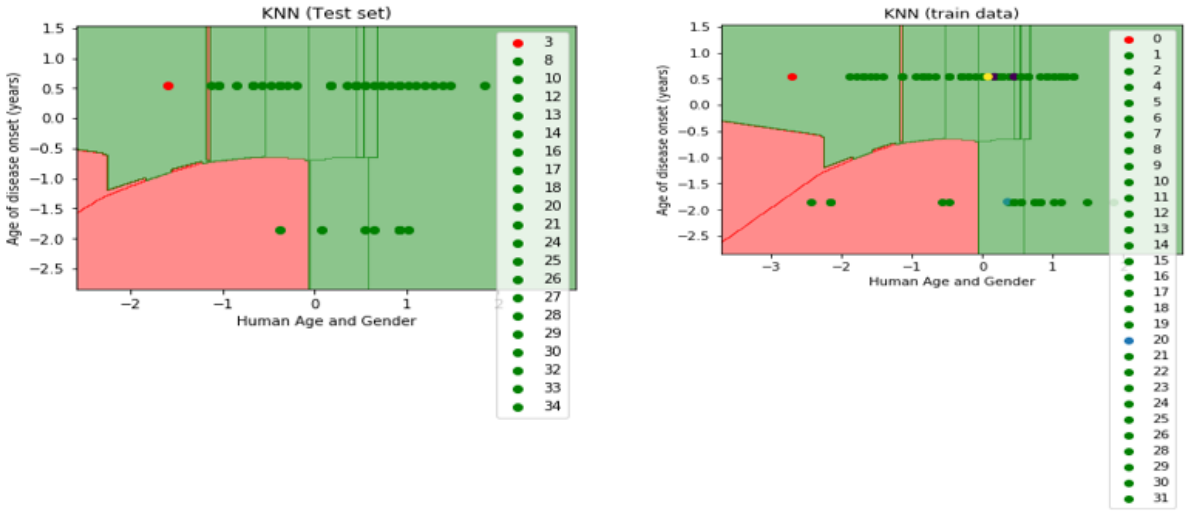

Figure 4.1.1: Comparing Human age and Gender with Age of disease onset(years) using KNN

Using k-nearest neighbor are classifying the age of disease onset(years) based on human age and gender. age of disease onset(years) can easily predicted from the human age and gender. Which type of gender are affected in the disease can classified.

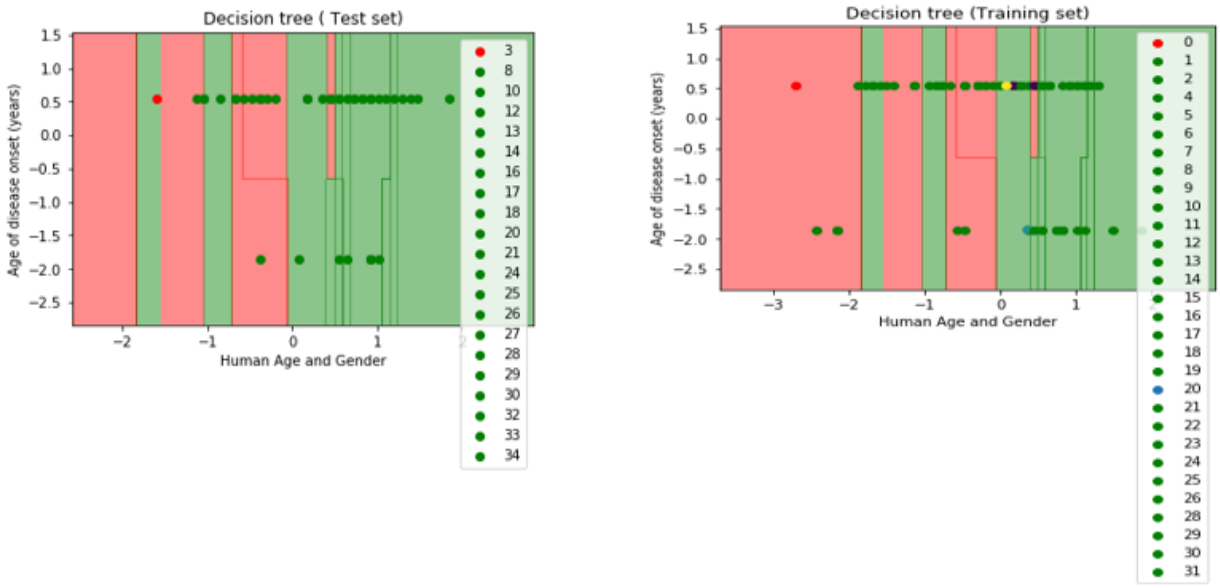

Figure 4.1.2: Comparing Human age and Gender

with Age of disease onset(years) using Decision Tree

The comparison in between age and gender with age of disease onset (years) using decision tree algorithm find the relation which classified the gender determine increasing the age be the probability of had parkinson's disease in the human body.

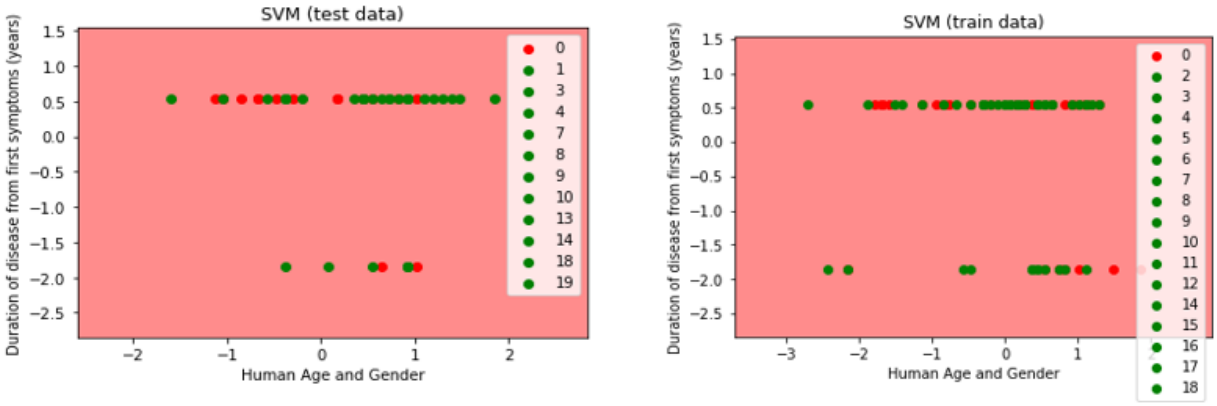

Figure 4.1.3: Comparing Human age and Gender with Duration of disease first symptoms prediction using SVM

The human age and gender also be the factors of parkinson's disease because male people are more suffer from female people that are easily understood from graph that symptoms for male people are first shown the sign from female people in the body.

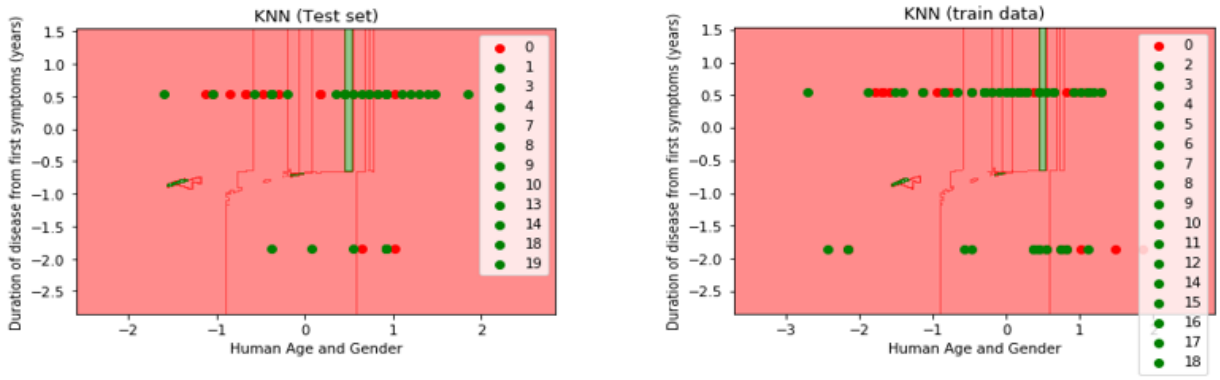

Figure 4.1.4: Comparing Human age and Gender with Duration of disease first symptoms prediction using KNN

The human age and gender also be the factors of parkinson's disease because male people are more suffer from female people that are easily understood from graph that symptoms for male people are first shown the sign from female people in the body.But the fact is K-nearest neighbor algorithm can't easily handle the categorical data.

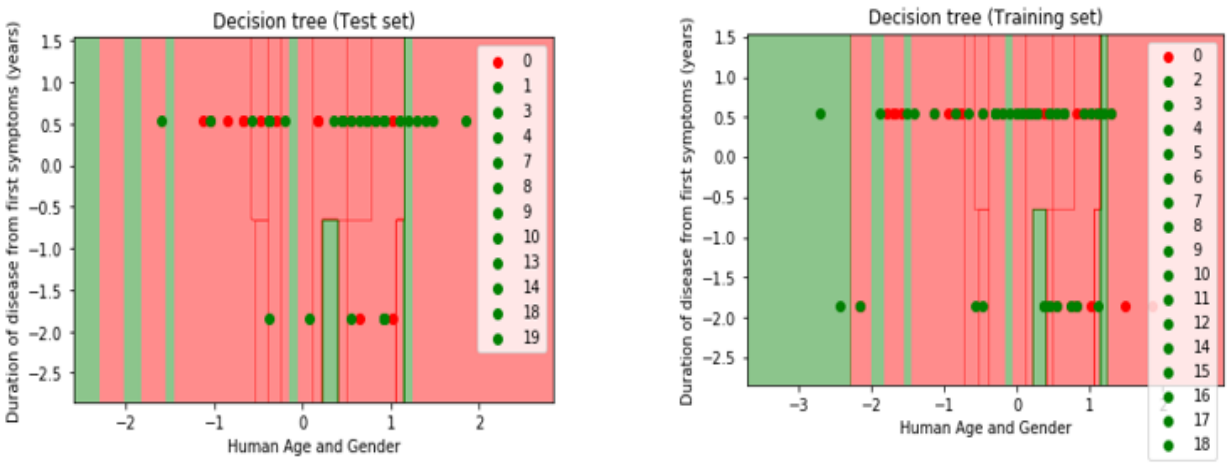

Figure 4.1.5: Comparing Human age and Gender with Duration of disease first symptoms prediction using Decision Tree

The human age and gender also be the factors of parkinson's disease because male people are more suffer from female people that are easily understood from graph that symptoms for male people are first shown the sign from female people in the body.But the fact is Decision tree algorithm can't easily handle the inconsistent data.

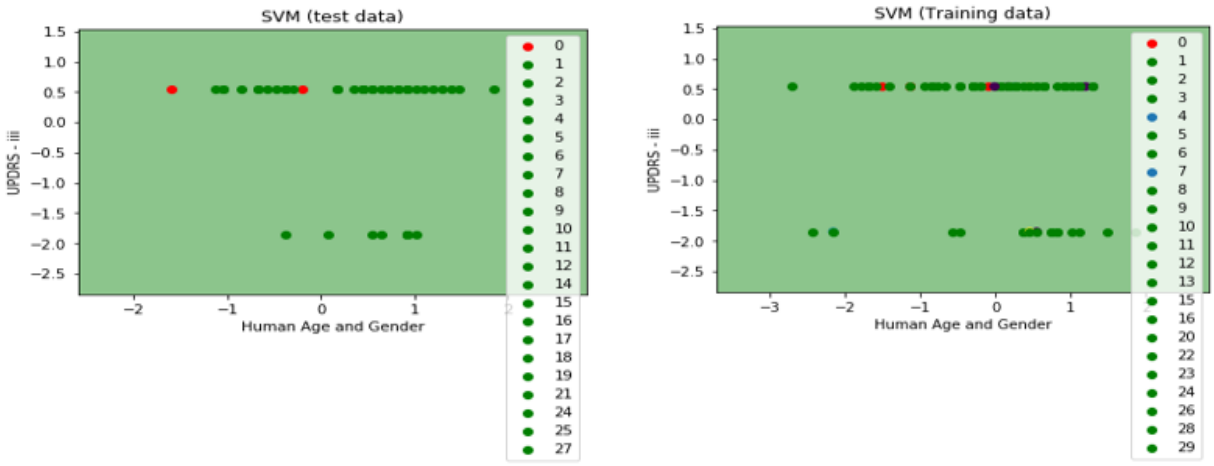

Figure 4.1.6: Comparing Human age and Gender with UPDRS-iii prediction using SVM

Comparing Human age and Gender with UPDRS prediction using SVM shows the effect of changing age is the factors of growing up parkinson's disease in the human body.

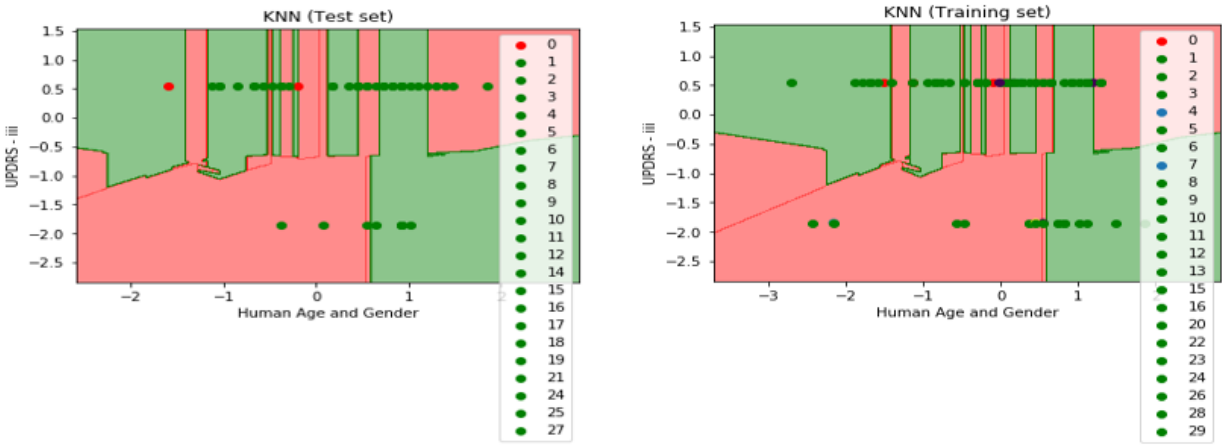

Figure 4.1.7: Comparing Human age and Gender with UPDRS-iii prediction using KNN

Comparing Human age and Gender with UPDRS prediction using KNN shows the effect of changing age is the factors of growing up parkinson's disease in the human body.

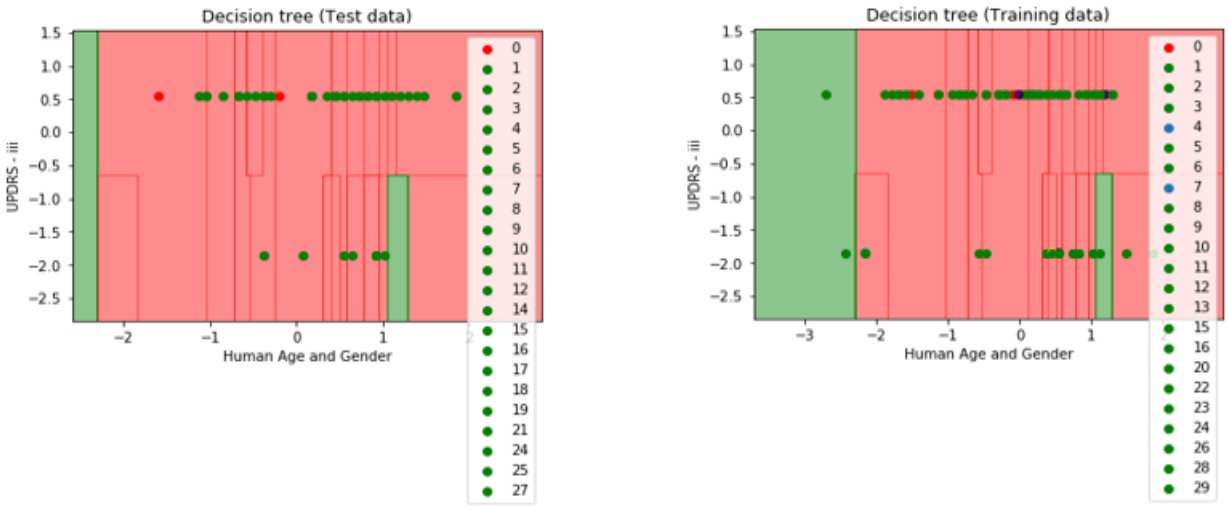

Figure 4.1.8: Comparing Human age and Gender with UPDRS-iii prediction using Decision Tree

Comparing Human age and Gender with UPDRS prediction using DT shows the effect of changing age is the factors of growing up parkinson's disease in the human body.

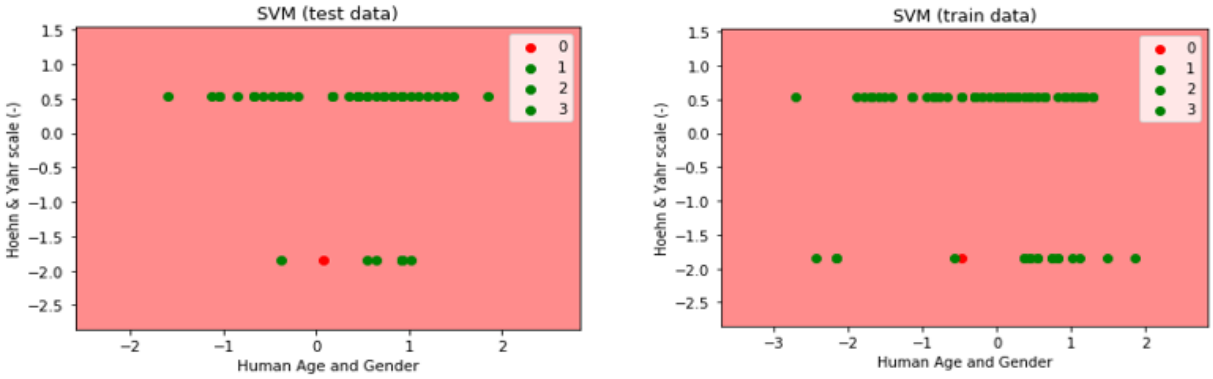

Figure 4.1.9: Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using SVM

Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using SVM can't show the good result because Hoehn & Yahr scale is not directly depend on this. So inconsistent result was shown.

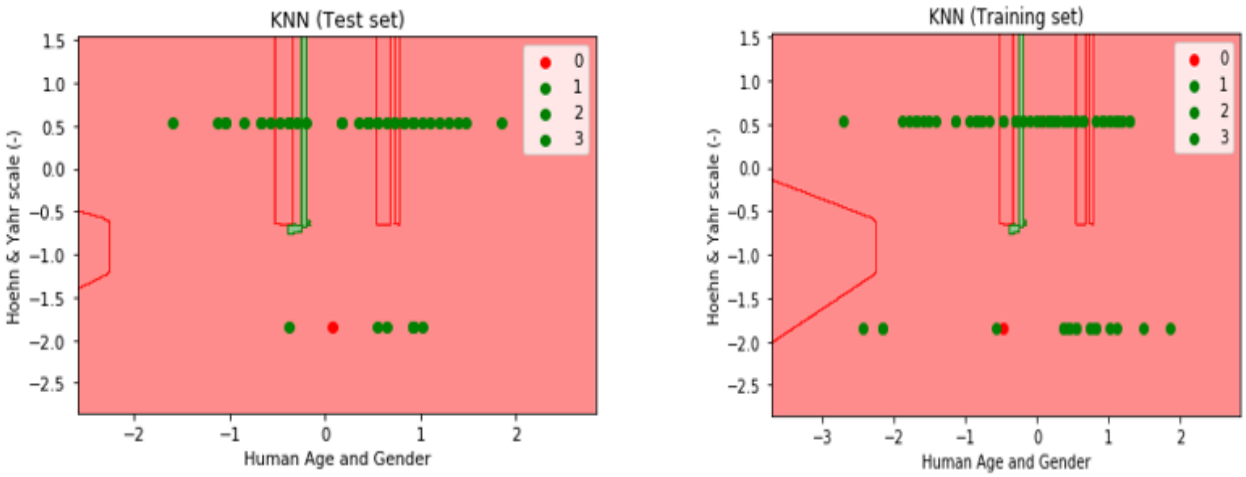

Figure 4.1.10: Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using KNN

Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using SVM can't show the good result because Hoehn & Yahr scale is not directly depend on this. So inconsistent result was shown.

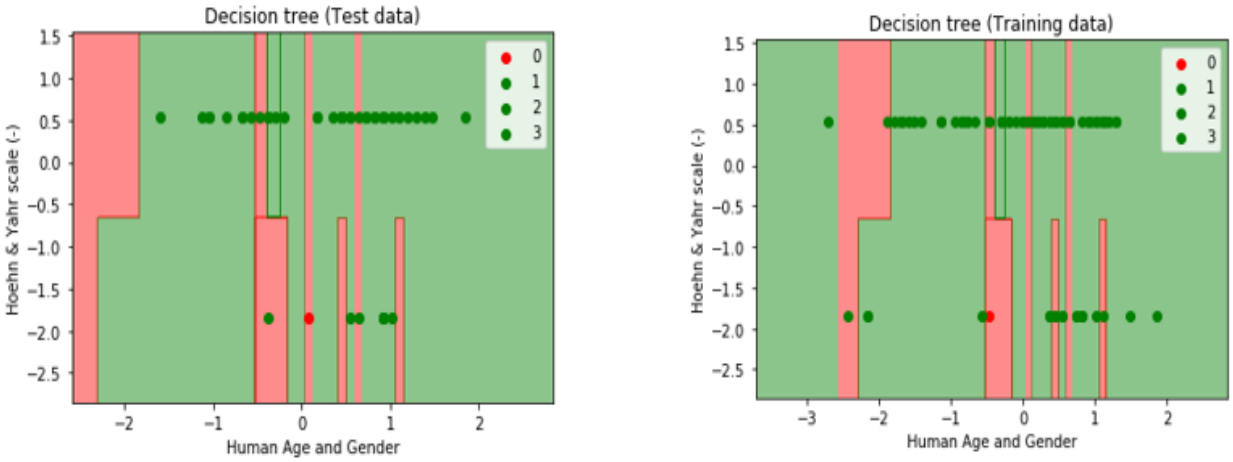

Figure 4.1.11: Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using Decision tree

Comparing Human [4] age and Gender with Hoehn & Yahr scale(-) prediction using SVM can't show the good result because Hoehn & Yahr scale is not directly depend on this. So inconsistent result was shown.

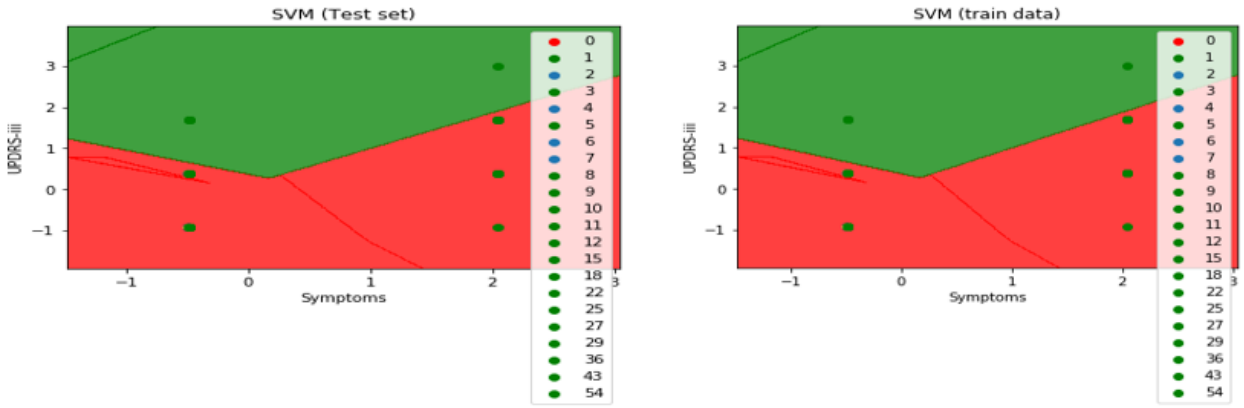

Figure 4.1.12: Comparing UPDRS-iii with Symptom's prediction using SVM

The comparison of UPDRS with the symptoms the result of [2] Unified parkinson's disease rating scale is changing so it was related with this directly. That also can changing the overall result and that is main bad factors of detecting parkinson's disease in the human body.

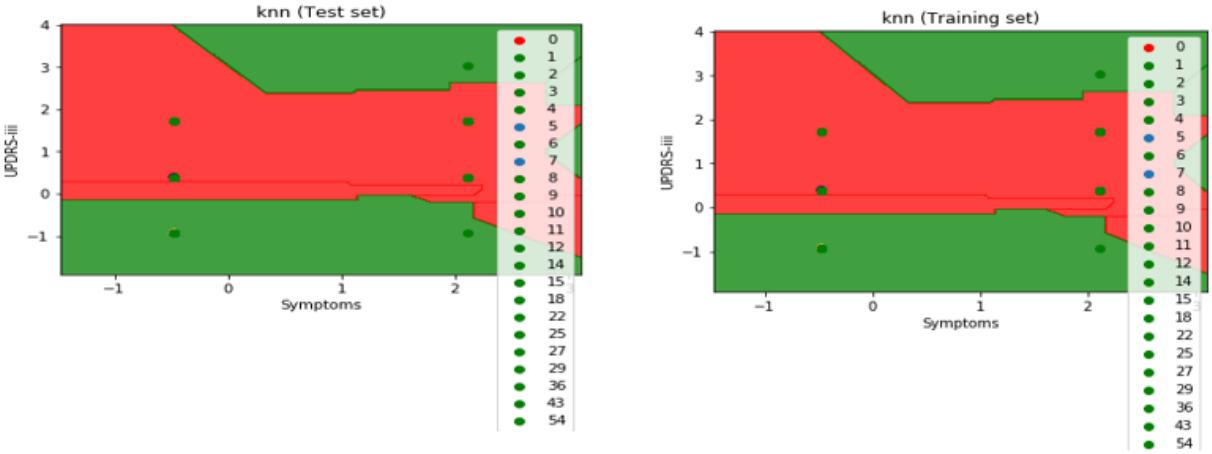

Figure 4.1.13: Comparing UPDRS-iii with Symptom's prediction using KNN

The comparison of UPDRS with the symptoms the result of [2] Unified parkinson's disease rating scale is changing so it was related with this directly. That also can changing the overall result and that is main bad factors of detecting parkinson's disease in the human body.

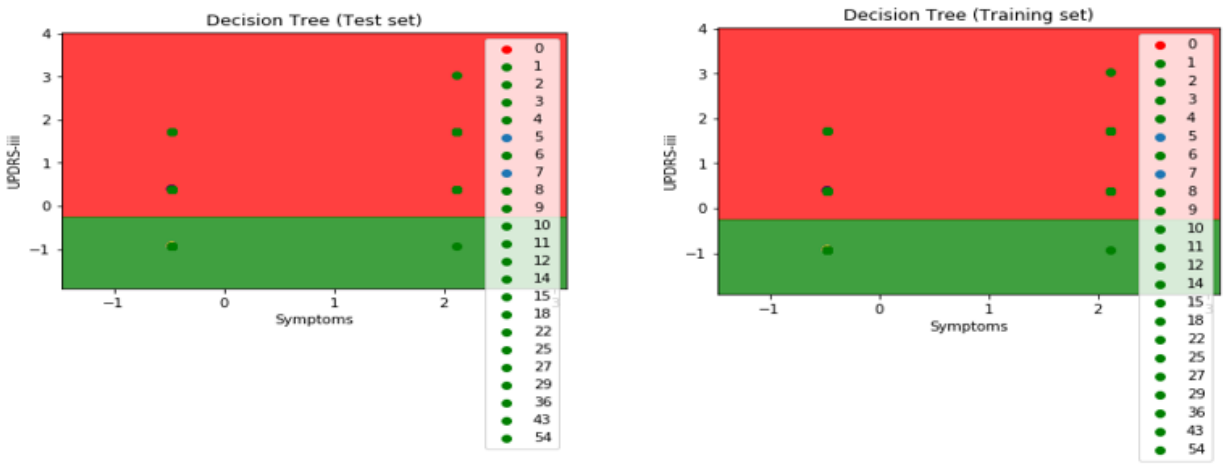

Figure 4.1.14: Comparing UPDRS-iii with Symptom's prediction using Decision Tree

The comparison of UPDRS with the symptoms the result of [2] Unified Parkinson's disease rating scale is changing so it was related with this directly. That also can be changing the overall result and that is main bad factors of detecting [2] Parkinson's disease in the human body.

In the overall result section try to figure out each and every possibility of had [2] Parkinson's disease in the human body. The working symptoms,age gender and onset of Parkinson's disease deals with [3] Unified Parkinson's disease rating scale (UPDRS) and Hoehn &Yahr scale (HYS) because the first one working with voice data and [2] symptoms which are related with Parkinson's disease on the other hand Hoehn &Yahr scale (HYS) working with all [8] symptoms of Parkinson's disease and finding the output measure from many approaches that depends with voice changing form. So finally the purpose the finding [11] Parkinson's disease in the human body can achieve from different prospective.

CHAPTER 5

DISCUSSION

Parkinson’s disease is a progressive disease so that diagnosing Parkinson's disease in a specific way is not possible. So had to find the factors that are related to build Parkinson's disease in the human brain. There are many factors to had Parkinson's disease in the human body like age, emotions, symptoms etc. is directly related to Parkinson's disease. Using two scaling Unified Parkinson's disease rating scale (UPDRS) and Hoehn & Yahr scale (HYS) used to find the Parkinson's disease rating that defines the growth rate of Parkinson's disease in the human body. So working with many criteria to get the result. The algorithms are working efficiently but each algorithm not giving the best accuracy for all time. The decision tree giving the best accuracy from overall criteria. But for a particular algorithm can’t make a good result for all the criteria. From the working with classification find a thing that uses the best algorithm for each criterion. otherwise build up a hybrid classifier algorithm which can deal with every criterion of had Parkinson's disease in the human body. But that is too much difficult to do. Because forgetting one output it will give multiple outputs for a prediction so that it was very hard to filter out from this.

CHAPTER 6

CONCLUSION

6.1 Future work

In the neurological field, as a progressive degenerative disease Parkinson's disease are bad disease and diagnosis of Parkinson's disease is very hard. At early stage is very difficult to find the symptoms of Parkinson's disease. The life expectancy of affected is 7 -14 years. So diagnosis Parkinson's disease is very necessary to find out. Because that types disease doing very bad impact in the world.

The machine learning research is very efficient to work with these field. So finding the specific way to diagnosis Parkinson's disease, why male is affected mostly affected from women and find the genetic factors related with Parkinson's disease using machine learning technique can be possible by the use of Data mining. Because data mining technique is very efficient to work with complex data.

So in future try build up a hybrid data mining algorithms which can be easily diagnosis Parkinson's disease in a single way

REFERENCES

[1] "Parkinson's Disease," Harvard medical school, Cambridge, Massachusetts,USA.

[2] 2. January 16th, "The Health, Environment, and Indigenous Communities Research Group," 16th January 2014. [Online]. Available: http://heicresearch.com/?p=1514. [Accessed 16th January 2014].

[3] J. Parkinson, "An Essay on the Shaking Palsy," *The journal of Neoropsychiatry and clinical Neurosciences,* vol. 14, no. 2, pp. 223-236, 2002.

[4] F. Andrew J. Lees MD, "Unresolved issues relating to the Shaking Palsy on the celebration of James Parkinson's 250th birthday," *Wiley Online Library,* 2007.

[5] B. D. Gaenslen A1, Early diagnosis of Parkinson's disease., Germany: Department of Neurodegeneration, Hertie Institute of Clinical Brain Research and German, 2010.

[6] Y. O. ¨. 2. D. D. S. C. A. a. R. H. G. Timothy J. Wroge1, Parkinson's Disease Diagnosis Using Machine Learning and Voice, Seattle, Washington, USA: 1. Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania, USA,2. Department of Engineering, Ozye ¨ gin University, Istanbul, Turkey,3. Division of Computing and Software Systems, University of Washington, Seattle, Washington,USA, 2018.

[7] T. V. S. S. V. R. V. S. N. S. V. G. K. Kaladhar, "Diagnosis of Parkinson Disease Using Machine Learning and Data Mining Systems from Voice Dataset," in *Advances in Intelligent Systems and Computing,* Visakhapatnam, India , Dept. of Bioinformatics, GITAM University, Visakhapatnam, India , 2014, pp. AISC, volume 327.

[8] M. B. Ipsita Bhattacharya1, SVM Classification to Distinguish Parkinson Disease, New Delhi ,India: NSIT, University Of Delhi (DU) , 2010.

[9] P. W. J. J. Hao Deng, The genetics of Parkinson disease, Changsha, China: Department of Neurology, the Third Xiangya Hospital, Central South University,, 2017.

[10] J. F. Cavanagha, Diminished EEG Habituation to Novel Events Effectively Classifies Parkinson's Patients, New Maxico: University of New Mexico, Department of Computer Science, 2017.

[11] S. S. G. Sateesh Babu, "Parkinson's disease prediction using gene expression – A projection based learning meta-cognitive neural classifier approach," *ScienceDirect,* 2013.

[12] R. Č. T. T. K. Š. E. R. &. J. R. Jan Hlavnička1, "Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder," www.nature.com/scientificreports, 2017.

[13] "Combined analysis of sensor data from hand and gait motor function improves automatic recognition of Parkinson's disease," *34th Annual International Conference of the IEEE EMBS,* 2012.

[14] "Using Human Factor Cepstral Coefficient on Multiple Types of Voice Recordings for Detecting Patients with Parkinson's Disease," *ElsevierMasson,* 2017.

[15] M. B. Ipsita Bhattacharya1, "SVM Classification to Distinguish Parkinson Disease Patients," Netaji Subhash Institute of technology NSIT, University Of Delhi (DU) , New Delhi ,India.

[16] G. S. ,. G. J. Dr.R.Geetha Ramani, "Feature Relevance Analysis and Classification of Parkinson Disease Tele-Monitoring Data Through Data Mining Techniques," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 2, no. 3, 2012.

[17] H. Chen, "The Changing Landscape of Parkinson Epidemiologic Research," *Journal of Parkinson's Disease* , 2017.

[18] Z. A. F. M. Ross Coomber, " Hip fractures and Parkinson's disease: a case series," 2017.

[19] M. Y. Y. W. R. S. C. S. Filip Simunovic1, "Evidence for Gender-Specific Transcriptional Profiles of Nigral Dopamine Neurons in Parkinson Disease," Mark R. Cookson, National Institutes of Health, United States of America, 2009.

[20] D. T. F. E. T. Mohammad Shahbakhi1*, "Speech Analysis for Diagnosis of Parkinson's Disease Using Genetic Algorithm and Support Vector Machine," *http://www.scirp.org/journal/jbise,* pp. 147-156, 2014.

[21] "A performance comparison of accelerometry based step detection algorithms on a large, non laboratory sample of healthy and mobility impaired persons," *30th Annual International IEEE EMBS Conference,* 2008.

[22] N. M. L. JOSEPH W. HOGAN, "MODEL-BASED APPROACHES TO ANALYSING INCOMPLETE LONGITUDINAL AND FAILURE TIME DATA," in *STATISTICS IN MEDICINE,* 1997, pp. 259-272.

[23] "Determination of Minimal Clinically Important Change in Early and Advanced Parkinson's Disease," *Parkinson Study Group,* 2010.

[24] T. P. E. a, Using partial decision trees to predict Parkinson's symptoms: A new approach for diagnosis and therapy in patients suffering from Parkinson's disease, Unit of Medical Technology and Intelligent Information System, Department of Materials Science and Engineering, University of Ioannina, GR 45110, Ioannina,Greece, 2011.

[25] E. P. W. S. G. W. H. Teasdale, "Clinical Trials and Therapy in Movement Disorders," 3 october 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/mds.116.

[26] "Tau and α-synuclein in susceptibility to, and dementia in, Parkinson's disease," 19 september 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.21192.