PARKINSON DISEASE PROGNOSIS

Asma Khan[1], Dr. Padmanjali H[2]

[1,2] Dept. of CSE , Guru Nanak Dev Engineering College, Karnataka, India

---

Abstract - *Using the Parkinson's disease dataset, we hope to classify the parametric and nonparametric models. Two models are compared for their ability to classify Parkinson's data. In parametric modelling, Parkinson's data classification is accomplished by Logistic Regression. Using techniques from non-parametric modelling, such as Random Forest and k-menas Algorithms, Parkinson's disease training and test data is classified. Parametric and nonparametric models are used to classify the Parkinson's data. Classification accuracy on parametric and nonparametric models is obtained using the classified value of data. The performance of the Parkinson's dataset is evaluated using a comparison of both parametric and nonparametric models.*

1. INTRODUCTION

Parkinson's disease is a progressive ailment that affects the central nervous system over a lengthy period of time.

The symptoms manifest themselves gradually, with trembling and stiffness being the most noticeable, followed by slowness of movement and difficulties walking.

It is believed that variables in both the genes and the environment contribute to the development of Parkinson's disease.

Each year, India sees the reporting of more than one million new cases.

Even though there is no known cure for this illness, medication may be of some benefit.

Meditation has been shown to be helpful in the treatment of symptoms associated with Parkinson's disease.

The symptoms of the sickness don't appear in the majority of individuals until they've been living with the condition for years.

The following well-known individuals have been diagnosed with Parkinson's disease:

In this work, we make use of two different methods:[1]

The term "data analytics technique" refers to qualitative and quantitative procedures and processes that are used to increase productivity and commercial benefit. This technique is used for the purpose of data analysis.

Extraction and classification of data are required steps in the process of identifying and analysing behavioural data, patterns, and trends.

2. LITERATURE SURVEY

Elbaz A, Bower *et al*. "Survival Study of Parkinson Disease in Olmsted County, Minnesota"

The goal of this study is to evaluate the survival of incident instances of Parkinson's disease (PD) against that of the general population, which is free of PD.

Through the Rochester Epidemiology Project's linked medical records system, we were able to detect new instances of Parkinson's disease in Olmsted County, Minnesota between 1976 and 1995.

An Essay on Shaking Palsy by John Parkinson.

Overall, clinical writers' use of the term "Shaking Palsy" is vague.

Some have used it to label typical cases of Palsy, in which some involuntary shaking of the limbs has occurred; others have used it to describe peculiar acts of kindness that have nothing to do with the disorder.

"Occupation and risk of Parkinsonism: A multicenter case-control study," by Samira A. Factor.

We looked at jobs (agriculture, instruction, medicine, welding, and mining) and exposures (solvents, pesticides) that may increase your risk of developing Parkinson's disease.

The purpose of this study is to investigate the association between parkinsonism risk and certain professions, jobs, or job duties.

RL, Koller WC, editors "Neurologic Principles and Practice, 2nd ed. New York" 2004. p. 177.

Separated into 14 primary categories, each of which is further subdivided into 271 subject-based subcategories.

This massive book is surprisingly readable due to the small and open parts.

All components are meant to have sufficient baseline information to coordinate care.

Therefore, the book is most useful when the discovery is presented and a study of the concerns surrounding a treatment option is necessary — a structure that takes into account practical application.

2012 April. parkinsons disease-statscountry.htm

The degenerative nature of Parkinson's disease (PD) is best understood in terms of its effect on the central nervous system.

In 1817, British physician James Parkinson published a description of the condition he named "the shaking paralysis."

He proposed the major consequences of the disease that would subsequently carry his name in this article.

Autism and PD are both part of a larger group of illnesses known as developmental disorders.

The four primary adverse effects include tremors (in the hands, arms, legs, jaw, or head), rigidity (in the limbs and trunk), slow growth (bradykinesia), and compromised balance (postural precariousness).

Dauer W. "Parkinson's disease:Mechanisms and Models. Neuron" 2003

Parkinson's illness (PD) results basically from the demise of dopaminergic neurons in the substance nigrosubstantial.

None of the current treatments for PD are able to halt or even significantly slow the death of dopaminergic neurons.

A lack of knowledge of the critical molecular events that provoke neurodegeneration is the primary roadblock to developing neuroprotective therapies.

3. PROBLEM DEFINITIONS

Researchers have had a tough time predicting Parkinson's disease in its early stages due to the fact that signs of the condition did not present themselves until middle age or later.

There are many different symptoms that might be caused by Parkinson's disease.

On the other hand, the symptoms of speech articulation problems in Parkinson's disease patients are the primary topic of this paper, in which an effort is made to construct a model utilising three different data mining approaches.

These three data mining approaches are from three distinct data mining domains: tree classifiers, statistical classifiers, and support vector machine classifiers. Tree classifiers are the oldest way of data mining.

There are three performance matrices that are used in order to evaluate the accuracy, sensitivity, and specificity of these three classifiers.

SCOPE:

By making use of the information on Parkinson's disease that was acquired in a dataset, we are able to rank the parametric and nonparametric models.

The Parkinson's information is tested with two different models in order to determine whether model provides the most accurate categorization.

Logistic Regression is used in parametric modelling to organise the Parkinson's information that has been collected.

ML Algorithms are applied to organise the preparation and test information of Parkinson's disease. These algorithms are derived from non-parametric showing.

The order is determined by combining the results of the parametric and non-parametric models with the information acquired about Parkinson's disease.

It is possible to achieve grouping accuracy on parametric and nonparametric models with the information that has been sorted in its worth.

Assessing the Parkinson's dataset involves looking at it through the lens of parametric and nonparametric models respectively.

Methodology:

The focused sensory system is especially vulnerable to the slow degeneration caused by Parkinson's disease.

Manifest adverse consequences manifest over time, the most notable of which are trembling, rigidity, slow growth, and walking difficulties.

Parkinson's disease has an unknown but likely complex genetic and environmental basis.

In India, almost a million new cases are recorded every year.

There is no cure for this illness, although treatment might assist.

Parkinson's disease symptoms may be managed by contemplation.

Most adverse effects take a while to manifest, and they tend to stick around for a very long period when someone is unwell.

Boxer Muhammad Ali was diagnosed with Parkinson's disease after sustaining many severe head wounds, joining the ranks of famous people who have been affected by the disorder.

It was conclusively confirmed that Pope John Paul II suffered from Parkinson's disease.

In the year 1991, Michael J. Fox was diagnosed with Parkinson's disease.

The most obvious symptoms of Adolf Hitler's Parkinson's infection were agitation and tremors.

Existing System:

The symptoms of Parkinson's disease may be classified into two categories: motor and non-motor.

Motor symptoms are well-known since they may be seen by everyone.

Some of the most common signs include a trembling hand at rest, slowness of movement (bradykinesia), trouble with balance, and stiffness [2].

It is now understood that non-motor signs may be seen throughout a window of time.

Dopamine doesn't help these symptoms.

Symptoms include cognitive decline, sleep disturbances, loss of smell, constipation, speech and swallowing difficulties, aches that cannot be pinned down, drooling, and low blood pressure while standing.

Although none of these non-motor symptoms alone is sufficient to diagnose PD, they may be useful when paired with additional indicators such as Cerebrospinal Fluid (CSF) testing and dopamine transporter imaging.

Proposed System:

This investigation takes into account both motor and non-motor symptoms by measuring biomarkers such cerebrospinal fluid and imaging the dopamine transporter.

In this paper, we adopt a similar strategy, though we make an effort to make use of alternative machine learning algorithms that can improve the model's performance and play a crucial role in making early predictions of PD, thereby allowing us to initiate neuroprotective therapies at the optimal time.

Modules:

Dataset information:

A total of 31 people, including 23 people diagnosed with Parkinson's disease, contributed biomedical voice estimates to this dataset (PD).

A total of 195 voice recordings were used to compile this data, with each table cell representing a different voice measurement (name segment).

The major purpose of the data is to distinguish healthy people from those with PD, as shown by the "status" field, which is set to 0 for healthy people and 1 for those with PD.

Data is presented in an ASCII CSV format.

Each line of the CSV file represents one case compared to one audio file.

For each individual patient, there are around six separate accounts, with the principal one serving as the primary means of identification.

Performance matrices:

In this study, we utilise three performance matrices to assess the classifiers and algorithms we've covered.

Classification accuracy, sensitivity, and specificity are standard performance metrics used to quantify a diagnostic method's efficacy.

• Accuracy: Accuracy characterises the degree to which a projected value is similar to the actual value.

It is possible to calculate accuracy as (TP1 + TN1)/(TP1 + TN1) * (FP1 + FN1).

• Sensitivity is the degree to which a change in one of the model's inputs affects the value of the output.

It prioritises the attributes that matter most for calculating the right result.[62]

A new formula has been established for it: Sensitivity = TP1/(TP1 + FN1).

Machine learning algorithms in disease :

Medical diagnosis and prognosis are two areas where AI calculations have a long and storied history.

AI calculations have been widely used in the medical industry, with many published examples demonstrating their usefulness in areas such as diagnosis, prognosis, and survival and identification of disease evidence.

Initially, AI developed in three distinct areas: symbolic learning, quantifiable procedures, and brain structures.

Hunt illustrated representative learning, Nilsson demonstrated quantifiable techniques, and Rosenblatt described neural networks.

An abundance of artificial intelligence (AI) tools have emerged in the machine learning community, and these tools are now often utilised to obtain classification models, such as clinical prognostic models.[16,17]

Choice tree classifiers and simulated neural networks have been used successfully in illness diagnosis and analysis.

Pendharker used A-NN, a Decision Tree, and a Logistic Regression to predict the likelihood that breast cancer patients would survive their disease.

Predictions of pneumonia mortality were made using a relapse and K nearest neighbour model and six alternative artificial intelligence computations.

4. SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS

CPU: Any processor with a clock speed of at least 500 MHz.

Ram : 4 GB

Disc Space: 4GB

Keyboard and mouse, the workhorses of input devices.

Video Graphics Array with High-Resolution Display as the Output Device.

REQUIREMENTS FOR SOFTWARE

OS X 10.7 or later

VS2010 is the latest version of the Visual Studio.

2008 SQL Server.

python2.7.x and later is supported by the IDE.

Programming Language for Developing Applications using Python

For versions 3.6 and above, you'll need to install the setup tools and pip.

5. FEASIBILITY REPORT

FEASIBILITY STUDY:

Examining a proposed endeavour to see whether it makes sense, can be carried out within the estimated budget, and will provide desirable results.

Studies of feasibility are often aimed at situations involving very large numbers of objects.

Achievability analysis is another name for this.

The main issue here is

You have probably noticed that the same online page might seem different depending on the application and even the version of the programme you are using to view it.

Sometimes a page on a site won't function properly until you've upgraded to the latest version of a required application.

In addition, a website page may look fine in an older application but poorly in a newer one.

Now think about the company that sends out a handful of software-based apps.

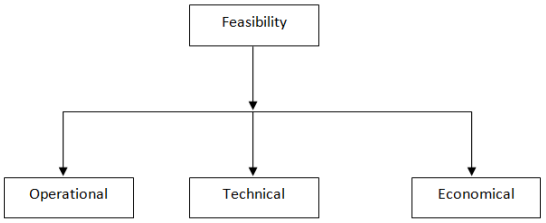
Then, all of a sudden, one of these apps gets an upgrade that necessitates switching to the currently deployed standard organisation software.

Assuming the company decides to restructure the program, it's probable that certain features of at least one of the program-based apps won't be compatible with the new program. This has effectively put an end to the company.

Do they upgrade the programme and maybe interfere with other software that relies on it?

Before delivering the new programme, do they devote significant resources to checking the sent program-based apps to ensure compatibility?

Instead, do they continue to rely on tried-and-true methods of progress?



Economic Feasibility:

The effective element linked with the building life cycle is the primary subject of this study's investigation.

This is very necessary in order to validate the reliability of the structure.

It is possible for the design team to identify the extent to which the development of the project on which they are

presently working is related to the practical use of the product.

If the most cost-effective use can be determined, the organisation may continue in the manner that is recommended by this information.

The possibility of carrying out the plan:

After it has been finished, this assessment examines the application in its entirety to determine how well it functions.

The operational functionality of the application has been fully implemented.

This is accomplished by making use of a stage and mix of some kind; more often than not, this involves the creative minute.

Technical Feasibility:

The exact data required for the engineer to realise the whole of the framework's functioning is understood thanks to this study's Technical Feasibility section.

The investigation is essential since it paves the path for future advancement.

Organizational Feasibility: The challenge will be to make certain that the structure can successfully validate the information with the client.

This establishes a connection between the framework and the client's location, so enabling the customer to make efficient use of the action plan.

The customer shouldn't let the framework be a barrier for them, but they should keep in mind that it is a must.

Feasibility in Terms of Time:

Clients will be able to change the riddle articulations of the record, and we may add the variables to reestablish the passphrase if the client forgets it. This will allow clients to change the riddle articulations in the event that they need to update the login credentials of the foundation while they are travelling.

6. LANGUAGE OVERVIEW

PYTHON

Python is an interpreted high-level programming language. The Python used by GuidoanRossum has a focus on whitespace and readability. Python's internal memory management and type system are completely independent of any other software.

In addition to a large standard library, it supports imperative, functional, procedural, and object-oriented programming paradigms.

PC-LEARNING

To me, it's dishonest to classify machine learning under artificial intelligence.

Rather than seeing machine learning as an end in itself, it is more fruitful to view it through the lens of data science's emphasis on model construction. When it comes to understanding data, machine learning relies on mathematics.

The term "learning" is introduced when we provide these models with a wide range of parameters, allowing the programme to "learn" from the information provided. These models gain the ability to foresee and make sense of new information when they are coupled with preexisting data.

Sizes M-L

Labeled training data is used for both classification and regression model training. In order to reach the desired standard, the learning process will be repeated.

Unsupervised learning uses factor and cluster analysis techniques to gain insight into data that has not been labelled.

To learn in a semi-supervised manner, labelled and unstructured data are combined.

Improved accuracy at a lower cost than supervised learning is achieved by using labelled data.

It is via trial and error that reinforcement learning achieves its results.

Success at this stage will be measured by your ability to establish routines of study that will maximise your potential for future gain.

DJANGO\sPython-based

Django is a robust web framework that allows for fast iteration and smart, aesthetically pleasing design. Capable software experts created it to solve a wide range of problems encountered while making a website. So you can stop worrying about wasting time and go straight to work on your app.

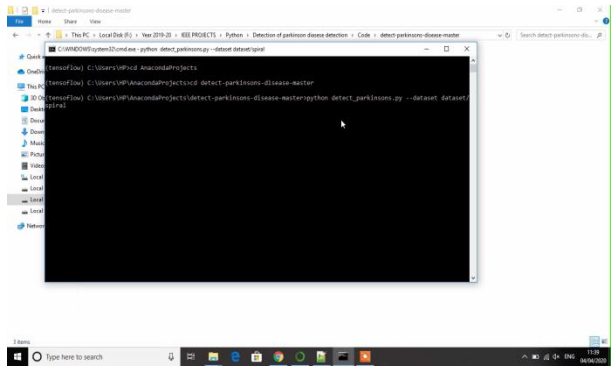
7. SYSTEM DESIGNS

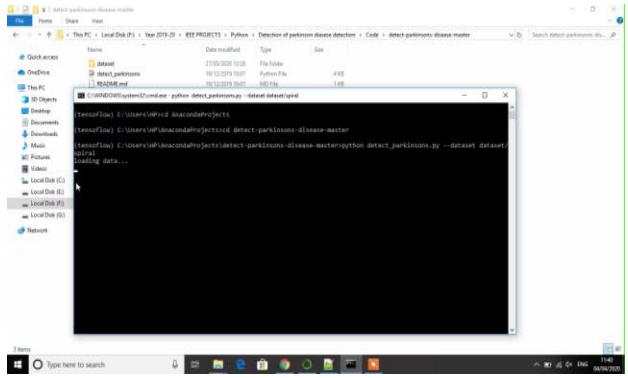
The next step in the process is System Design, which is where the system's big-picture structure is established.

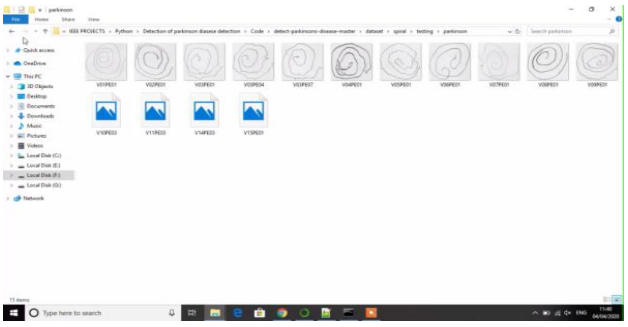
This system is structured as a network of interacting modules.

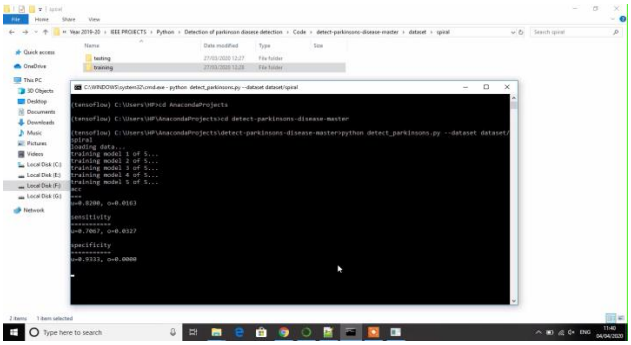
When designing the system as a series of interacting subsystems, the analyst takes into account both the requirements discovered in system analysis and the expectations of the end user.

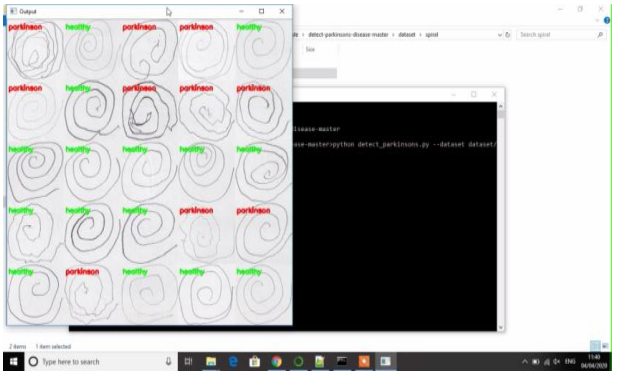
8. OUTPUT

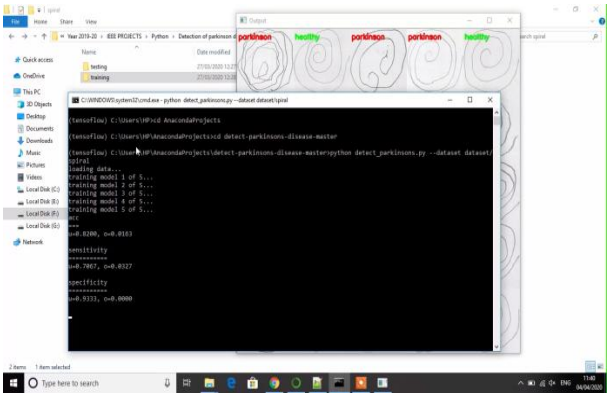


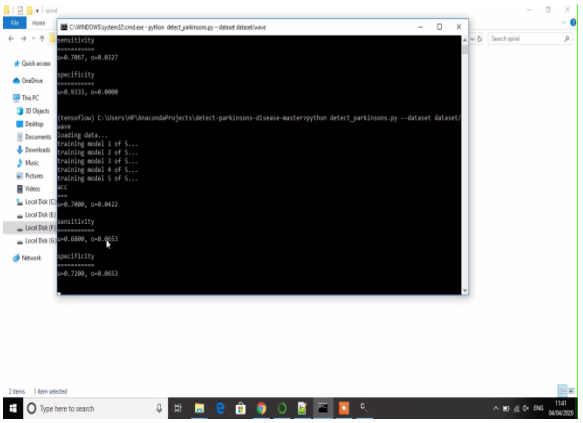


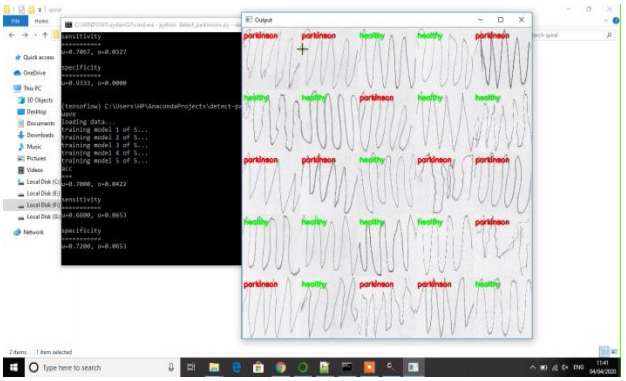












9. SYSTEM TESTING & IMPLEMENTATIONS

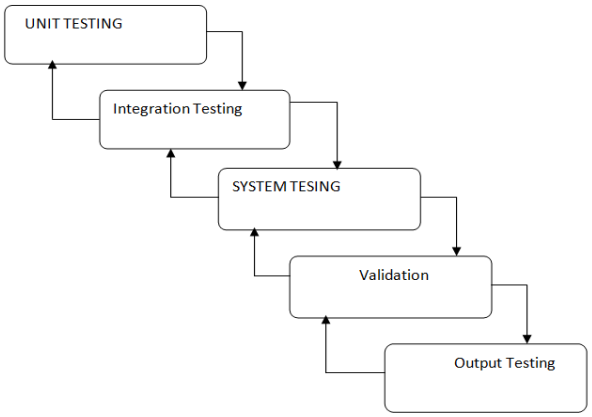


Figure: Level of Test-ing

1. Unit Testing

The excitement lashing test is designed to expose weaknesses. A workaholic's "test" is an exhaustive examination of their potential shortcomings. It's a step toward making sure that everything from individual components to subgroups to social events and the finished product itself is easily accessible. Second, Tests of Integration A joining test will be needed to see whether the merged piece of code is really functioning as one. The main issue with point drives is the delay in the primary outcome on screen or field, which compounds the difficulty. The mistake in the fragment is accurate and stable, as shown by consolidate tests, independent of the part's public finish. Finding the problem you can lean against in a jumble of possibilities no longer requires a joint appointment. Error detection is the goal of system testing. Testing is the practise of looking for flaws or weak spots in a product. Multiple kinds of examinations may be taken. Depending on the kind of test being conducted, a variety of methods may be used to satisfy this criterion.

Testing for functionality:

All technical and business criteria, as well as those defined in the system documentation and user manuals, are satisfied and more throughout functional testing. Areas where functional testing is most prevalent include: Acceptable Input Classes All valid input classes must be accepted.

White Box Testing: This kind of testing is performed by software testers who have some understanding of the program's design and implementation (or at least its intended purpose). It serves a useful function. As such, it is used for the purpose of evaluating black-box-level features. The logic of each module is mapped out in a flowchart, and then test cases are written and run to ensure that all logical choices are tested. It has been put to use in generating test cases in the following scenarios:

Checking the "Black Box": The irrational or nonexistent performance of expected functions Mistakes in the user interface (ii) Inconsistencies in the data structures (iii) Use of a third-party database (or lack thereof) Errors in Performance, Part iv Mistakes during startup or shut down (v). In this check, we look just at the output to make sure it's correct. There is no attempt to follow the data's logic.

Tests for Acceptance: Getting user buy-in during the testing phase of a project is critical and requires significant input from the target audience. The system's functional needs are also verified.

Results from the tests:

All of the aforementioned tests came out positive. No problems were found. Test Methodology: Top-Down

Approach: Testing may be performed sequentially, beginning with the simplest and most fundamental components. Bottom-up testing involves running each module as part of a smaller programme that provides it with the input data it needs to simulate its eventual role in a larger system

10. CODE EFFICIENCY

MEASURES OF CODE EFFICIENCY

The code is planned in light of the accompanying attributes.

Uniqueness: The code structure should guarantee that only one worth of the code with a solitary significance is accurately applied to a give element or trait.

Expandability: The code structure are intended for such that it should consider development of it's arrangement of elements or characteristics, accordingly giving adequate room to the section of new things with in every grouping.

Succinctness: The code requires the least conceivable number of positions to incorporate and characterize every thing.

Uniform size and organization: Uniform size and arrangement is profoundly positive in automated information handling framework. Effortlessness: The codes are planned in a straightforward way to comprehend and easy to apply.

Flexibility: The code permits altering effectively to reflect fundamental changes in conditions, qualities and

relationship of the encoded substances. Sortability: Reports are generally significant for client proficiency when arranged and introduced in a foreordained configuration or request.

Soundness: Codes that don't need to be often refreshed likewise advance use productivity. Individual code tasks for a given element ought to be made with a negligible probability of progress either in the particular code or in the whole coding structure.

Significance: Code is significant. Code worth ought to mirror the qualities of the coded substances, for example, memory helper highlights except if such a techniques brings about irregularity and rigidity.

11. CONCLUSION

In this study, we make an effort to develop a diagnostic model for Parkinson's disease.

For this, we use sequential minimization optimization, a kind of data mining, as well as decision stump (tree classifiers) and logistic regression (statistical classifiers) (support vector machine).

Data for this article was retrieved from the UCI repository.

From 31 participants, 23 are diagnosed with Parkinson's disease and have had their voices measured using a range of biological instruments (PD).

There are 195 voice recordings here, and each row in the table corresponds to one of those measurements.

A number of 1 indicates that the individual is afflicted with Parkinson's disease, whereas a value of 0 indicates that they are healthy.

Three factors are utilised to evaluate the effectiveness of the classifiers under discussion, and the 10 cross fold approach is used to get the required result.

12. REFERENCES

1. Elbaz A, Bower JH, Peterson BJ, Maraganore DM, McDonnell SK, Ahlskog JE, et al. Endurance Study of Parkinson Disease in Olmsted County, Minnesota. Curve Neurol 2003;60:91‐6.

2. Parkinson J. An exposition on the shaking palsy.1817. J Neuropsychiatry Clin Neurosci 2002; 14:223‐36.

3. Leather expert CM, Ross GW, Jewell SA, Hauser RA, Jankovic J, Factor SA. Occupation and hazard of Parkinsonism: A multicenter case‐ control study. Curve Neurol 2009;66:1106‐13.

4. Marras C, Tanner C. The study of disease transmission of Parkinson's Disease. Development Disorders. In: Watts RL, Koller WC, editors Neurologic Principles and Practice, second ed. New York: The McGraw‐Hill Companies; 2004. p. 177.

5. US Census Bureau. US break projections by age, sex, race, and Hispanic beginning: 2000‐2050. Accessible from: http://www.census.gov/populace/www/projections/usinterimproj. [Last Accessed on 2012 Apr. 7].

6. Accessible from: http://www.rightdiagnosis.com/p/parkinsons_disease/stats‐country.htm. [Last Accessed on 2012 Apr.

7. Dauer W, Przedborski S. Parkinson's illness: Mechanisms and Models. Neuron 2003;39:889‐909.

8. Alonso JB, de Leon J, Alonso I, Ferrer MA. Programmed location of pathologies in the voice by HOS based boundaries. EURASIP J Appl Sig process 2001;14:275‐84.

9. Cnockaert L, Schoentgen J, Auzou P, Ozsancak C, Defebvre L, Grenez F. Low‐ recurrence vocal balances in vowels created by Parkinsonian subjects. Speec Commun 2008;50:288‐300.

10. Revett K, Gorunescu F, Mohamed Salem AB. Highlight Selection in Parkinson's sickness: A harsh Sets approach. Procedures of the International Multi meeting on Computer Science and Information Technology; Oct. 12‐14 Margowo, Polond 2009;4: ps. 425‐8.