PARKINSON'S DISEASE DETECTION USING MACHINE LEARNING

Muthangi Bhavana[1], Mirdoddi Richa[2], M. Amarender Reddy[3],

[1] *Student, Department of Information Technology, B V Raju Institute of Technology, Medak, Telangana, India.*
[2] *Student, Department of Information Technology, B V Raju Institute of Technology, Medak, Telangana, India.*
[3] *Assistant Professor, Department of Information Technology, B V Raju Institute of Technology, Medak, Telangana, India.*

ABSTRACT

The Parkinson's disease is progressive neuro degenerative disorder that affects a lot only people significantly affecting their quality of life. It mostly affect the motor functions of human. The main motor symptoms are called "parkinsonism" or "parkinsonian syndrome". The symptoms of Parkinson's disease will occur slowly, the symptoms include shaking, rigidity, slowness of movement and difficulty with walking, Thinking and behavior change, Depression and anxiety are also common. There is a model for detecting Parkinson's using voice. The deflections in the voice will confirm the symptoms of Parkinson's disease. This project showed 73.8% efficiency. In our model, a huge amount of data is collected from the normal person and also previously affected person by Parkinson's disease. These data is trained using machine learning algorithms. From the whole data 60% is used for training and 40% is used for testing. The data of any person can be entered in database to check whether the person is affected by Parkinson's disease or not. There are 24 columns in the data set each column will indicate the symptom values of a patient except the status column. The status column has 0's and 1's.those values will decide the person is effected with Parkinson's disease. I's indicate person is effected, 0's indicate normal conditions.

Keywords: *Parkinson's disease; Machine learning (ML), Support Vector Machine*

---

1. INTRODUCTION

In our model, a huge amount of data is collected from the normal person and also previously affected person by Parkinson's disease. These data is trained using machine learning algorithms. From the whole data 80% is used for training and 20% is used for testing. The data of any person can be entered in database to check whether the person is affected by Parkinson's disease or not. There are 24 columns in the data set each column will indicate the symptom values of a patient except the status column. The status column has 0's and I's. Those values will decide the person is effected with Parkinson's disease. I's indicate person is effected, 0's indicate normal conditions.

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks T, as measured by P, improves with experience.

2. LITERATURE SURVEY

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems Fig 2.1 It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

The data points from each class that lie closest to the classification boundary are known as support vectors. If an SVM is given a data point closer to the classification boundary than the support vectors, the SVM declares that data point to be too close for accurate classification.

• The margin should be as large as possible.

• The support vectors are the most useful data points because they are the ones most likely to be incorrectly classified.

The advantages of support vector machines are: Effective in high dimensional spaces. Still effective in cases where number of dimensions is greater than the number of samples. Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

2.1 EXISTING SYSTEM

In existing system, Parkinson's is detected at the secondary stage only which leads to medical challenges. Also doctor has to manually examine and suggest medical diagnosis in which the symptoms might vary from person to person so suggesting medicine is also a challenge.

♣ Thus the mental disorders are been poorly characterized and have many health complications. Parkinson's is generally diagnosed with the following clinical methods as,

♣ MRI or CT scan - Conventional MRI cannot detect early signs of Parkinson's disease

♣ PET scan - is used to assess activity and function of brain regions involved in movement

♣ SPECT scan - can reveal changes in brain chemistry.

2.2 DISADVANTAGES OF EXISTING SYSTEM

• As the existing systems we need to verify with the doctor, the local doctor may or may not be aware of this disease. Dept. of IT, BVRIT, Narsapur.

• Therefore, this method's main flaw was that we need to search a particular specialist doctor.

2.3 PROPOSED SYSTEM

♣ As we will be using python for this project, we will also be needing a suitable environment for our code to run. We used Google Col-laboratory, as it removes the tedious process of compilation in the computer itself, and any type of code can be run very easily. All we need for using Google Collab is a stable internet connection.

♣ Here we proposed predictive models that are adaptive, flexible, and scalable. The proposed research provides the weather the person have Parkinson's or not.

2.3.1 ADVANTAGES OF PROPOSED SYSTEM

• One drawback we found is, for the greatest and most accurate prediction, in our opinion, both of these criteria need to be considered simultaneously.

• So data from the dataset will be useful to predict the accurate value.

3. DESCRIPTION OF ALGORITHMS:

3.1 Support Vector Machine Algorithm

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithm, which is used for Classification as well as regression problems. However, primarily, it is used for Classification problem in machine Learning

3.2 TYPES OF SVM

SVM can be of two types:
- **Linear SVM**: Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.

- **Non-linear SVM**: Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

4 TESTING & VALIDATION

4.1 Testing

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

4.2 Validation

Validation means observing the behavior of the system. The verification and validation mean, that will ensure that the output of a phase is consistent with its input and that the output of the phase is consistent with the overall requirements of the system. This is done to ensure that it is consistent with the required output. If not, apply certain mechanisms for repairing and there by achieved the requirements.

5. CONCLUSIONS

♣ The average accuracy obtained in the classification of breast cancer is around 0.84 by using Support Vector Machine Algorithm.
♣ This model will provide precise and error-free Predicted data.
♣ The process of Predicting data using big data sets will be streamlined by Support Vector Machine.

6. REFERENCES

Papers Referred: https://ieeexplore.ieee.org/document/9768002
For Implementation:
https://www.cdc.gov/genomics/hugenet/casestudy/parkinson/parkcoffee_view
https://www.kaggle.com/datasets/vikasukani/parkinsons-disease-data-set
https://www.tutorialspoint.com/google_colab/index.htm