iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.51005

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Parkinson Detection using Machine Learning Algorithms

Sheetal Gupta[1], Saliha Ghanchi[2], Shabista Idrisi[3]
Department of Information Technology Shah & Anchor Kutchhi Engineering College, Chembur, Mumbai

Abstract: ***Parkinson's Disease (PD is an irreversible, chronic, and progressive disorder, estimated to affect about ten million people worldwide affecting patient movement and feelings of motivation and reward. Motor symptoms of PD include tremor or shaking, which usually begins in one arm or hand, muscle rigidity or stiffness, and slowness, also referred to as bradykinesia. Signs and symptoms are unique to each person and the treatment should be tailored to each person's individual needs. A recent survey of 492 people with PD found most believed self-monitoring would improve their understanding of PD, well-being, ability to cope, and communication with their health-care team. The overlap of symptoms with other neurological disorders leads to difficulties, delays, and errors in diagnosis, negatively impacting patients. A recent study of 1775 patients' experiences of PD diagnosis identified the importance of early diagnosis. and found misdiagnosis occurred at first consultation for 15% of participants. This fallibility of clinical diagnosis is recognized and requires periodic diagnosis reviews. Participants are presented with three activities evaluating Patients Call Response, Fine Motor Movements and Memory. To evaluated Patients, Name Call Response, the application uses a Chatbot to communicate with the patient, and tries to ask questions by calling his name. On successful response a series of other questions are asked on patients likes and dislikes. Fine Motor Deficit is evaluated based on a Finger Tapping Task, the patient as to tap on two buttons shown on the phone screen alternatively. The data is collected after 2 minutes on Tapping and compared with threshold set by the Normal Group. The Memory Activity measures the time taken by the patient to solve the puzzle by finding the corresponding images. Patients are shown images of fruits that are illuminated one at a time and then asked to find the same in the grid.***

Keywords: Parkinson, Fine motor deficit, finger tapping activity, puzzles, accuracy, knn, svm, machine learning.

I. INTRODUCTION

1) Background: Parkinson's disease is a chronic, progressive disease that causes involuntary or uncontrollable movements such as shaking, stiffness, and problems with balance and coordination. Symptoms usually start gradually and get worse over time; people have trouble walking and talking. Patients may also experience mental and behavioural changes, sleep problems, depression, memory problems, and fatigue. Although anyone can be at risk, some studies show that the disease affects men more than women. Most people develop Parkin son's disease after age 60, and about 5_10% develop it before age 50. Since there is no known cure for Parkinson's disease, symptoms can be relieved with treatment.

2) Patients should see their doctor every 4 to 6 months to assess presenting symptoms. Early and reliable diagnosis of Parkinson's disease is an important undertaking that can take a long time as patients improve. To reduce the suffering of people with Parkinson's disease, we offer a mobile application that detects symptoms such as attention deficits, poor response to calls and fine motor deficits. These symptoms can be assessed using defined activities performed by normal and patient groups. To predict the outcome of each campaign, we used the KNN algorithm, which will be trained using the normal group data and thresholds will be set for the patient group.

3) Parkinson Disease is a brain neurological disorder. It leads to shaking of the body, hands and provides stiffness to the body. No proper cure or treatment is available yet at the advanced stage. Treatment is possible only when done at the early or onset of the disease. These will not only reduce the cost of the disease but will also possibly save a life. Most methods available can detect Parkinson in an advanced stage; which means loss of approx. 60% dopamine in basal ganglia and is responsible for controlling the movement of the body with a small amount of dopamine. More than 145,000 people have been found alone suffering in the U.K and in India, almost one million population suffers from this disease and it's spreading fast in the entire world.

4) People diagnosed with Parkinson's disease may have other symptoms, including -

Depression Anxiety
Sleep and memory problems
Loss of smell and balance problems.

©IJRASET: All Rights are Reserved

The cause of Parkinson's disease is unknown, but researchers have shown that a variety of factors contribute to the disease. These include –

a) *Genes* – Some very rare mutated genes have been discovered through research. Genetic variants generally increase the risk of Parkinson's disease, but to a lesser extent each genetic marker.
b) *Environment* - due to the presence of certain harmful toxins or chemicals in the environment which can cause disease but have minor effects. Although it develops at age 65, 15% can be found in young people under 50.

II. LITERATURE REVIEW

A. *Early Detection of Parkinson's Disease using Machine Learning & Image Processing*

Machine Learning and Image Processing, can detect PD in its early stage, The method ensures it separates the detection and prediction phases, so it can accurately predict the disease, better accuracy There is potential chance of improving its capability in terms of handling complex data sets and large size of data. The processing speed of input dataset needs improvement as well.

B. *Parkinson's Detection Using Machine Learning*

Machine Learning, Gives better accuracy by applying decision tree algorithms

C. *Detection of Parkinson's Disease Using Vocal Features: An Eigen Approach*

ML- Eigen approach, Gives accuracy upto 100%, these techniques can be used to detect various other diseases that show variations in the voice patterns, such as dyslexia, asthma, some types of cancer, etc.

D. *Prediction of Parkinson's disease using XGBoost*

XGBoost, The proposed method is being developed to ensure that Parkinson's disease is detected more accurately and that the best treatment options are available, accuracy 92%, This proposed Parkinson data analysis approach can also be extended to include various sorts of variances. The results of these variations experiments can be used to diagnose Parkinson's disease in a variety of ways.

E. *Stratification of Parkinson Disease using python scikit- learn ML library*

ML, python , The results obtained in this paper are promising as they provide enough accuracy in the prediction of the Parkinson disease and can let the patients decide if they want to go for the treatment or not Future work can be extended to classify the datasets of the Parkinson disease with even more number of features that are dependent on one another and to apply the best suitable model that gives the highest accuracy. It can be extended in classifying tele-monitoring dataset.

III. METHODOLOGY

We will use the KNN algorithm and the SVM algorithm, the abbreviation KNN stands for "K_Nearest Neighbor". It is a super vised machine learning algorithm. This algorithm can be used to solve classification and regression problems. The number of nearest neighbors for a new unknown variable to predict or classify is indicated by the symbol "K".

Before we start with this amazing algorithm, let's look at a relevant real-life scenario. We're often told that you have a lot in common with your closest colleagues, whether it's your thought process, work etiquette, philosophy, or other factors. As a result, we form friendships with people we think are similar to us. The KNN algorithm uses t he same principle. Its purpose is to locate all nearest neighbors around a new unknown data point to find out which class it belongs to. This is a distance-based method. Support Vector Machines or SVMs are one of the most popular learning algorithms for classification and regression problems. However, it is mainly used for classification problems in machine learning The goal of the SVM algorithm is to create an optimal line or decision boundary that can divide the n_dimensional space into c lasses so that we can easily assign new data points to the right class at the coming. This optimal decision boundary is called a hyperplane. SVM selects extreme points/vectors which help to create a hyperplane. These extreme cases are called support vectors, so the algorithm is called a support vector machine.

A. *Algorithm Used*

KNN Algorithm

The k_Nearest Neighbor (kNN) algorithm is a simple tool that can be used to solve many realworld problems in finance, healt hcare, recommender systems, and more. This blog post will introduce what kNN is, how it works, and how to implement it in a machine

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

learning project. The k_nearest neighbor (kNN) classifier is a nonparametric supervised machine learning algorithm. It is distancebased: it clas sifies objects according to the category of their nearest neighbors. Although k can be considered an algorithm parameter in a sense, it is actually a hyperparameter. It is manually selected and h eld fixed during formation and transfer time. The k_nearest neighbor algorithm is also nonlinear. Unlike simpler models such as linear regression, it works with data where the relationship between the independent variable (x) and the dependent variable (y) is not a straight line.The parameter k in kNN refers to the number of labeled (neighboring) points considered for classificationThe value of k represents the number of these points used to determine the result. Our task is to calculate distances and determine which classes are closest to our unk nown entities.

The main concept behind k-Nearest Neighbors is as follows.

Given a point whose class we do not know, we can try to figure out which points in the feature space are closest to it. These points are the k nearest neighbors. Since similar things occupy similar positions in feature space, the point is likely to belong to the same class as its neighbors. Based on this, new points can be classified as belonging to either class.

Providing a training set

A training set is a set of labeled data used to train a model.

You can create the training set by manually labeling the data or by using a labeled database available in a public resource, such as this one. A good practice is to use about 75% of the available data for training and 25% for the test set.

B. Finding the k-nearest neighbors

Finding the k_nearest neighbors of a record consists of finding the records that most closely resemble it in common characteristics. This step is also called similarity search or distance calculation.

C. Classification Points

For classification problems, the algorithm assigns class labels based on the majority vote, that is, by applying the most frequent label found in the neighborhood.

For regression problems, the mean is used to identify the k nearest neighbors.

After completing the above steps, the model provides results. Note that we are using discrete values for the classification problem, so the output will be descriptive, like "like bright colors".

For regression problems, apply continuous values, which means the output will be a floating point number.

We can gauge the accuracy of the results by seeing how well the model's predictions and estimates match the known classes i n the test set.

D. How to determine k-value in k-neighbor classifier

The best value of k will help you achieve the maximum accuracy of your model. However, this process is always full of challenges. The simplest solution is to try the k values and find the one that gives the best results on the test set. To do this, we follow the steps below:

Choose a random value of k.

In practice, k is usually chosen at random between 3 and 10, but there is no hard and fast rule. Small values of k lead to unstable decision boundaries. Higher values of k generally result in smoother decision bounds, but do not always result in better metrics. So it's always a matter of trial an d error.

Try different values of k and note their accuracy on the test set. Choose k with the lowest error rate and implement the model.

E. SVM

Support vector machines are a set of supervised learning methods for classification, regression, and outlier detection. All of these tasks are common in machine learning.

You can use them to detect cancer cells based on millions of images, or you can use them to predict future driving routes from a well-fitting regression model.

There are specific types of SVMs that can be used for specific machine learning problems, such as Support Vector Regression (SVR), which is an extension of Support Vector Classification (SVC).

The most important thing to remember here is that these are just mathematical equations designed to give you the most accurate answer as quickly as possible.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

SVMs differ from other classification algorithms because they choose their decision bounds to maximize the distance to the nearest data point for all classes. The decision boundary created by the SVM is called a maximum margin classifier or maximum margin hyperplane. A simple linear SVM classifier works by drawing a straight line between two classes. This means that all data points on one side of the line will represent one category, while data points on the other side of the line will be placed in another category. This means that there can be an infinite number of lines to choose from. The linear SVM algorithm outperforms some other algorithms such as k nearest neighbors because it chooses the best row to classify the data points. It chooses rows that separate the data and are as far away from the cached data points as possible.

2D example to help understand all machine learning terms. Basically you have data points on a grid. You are trying to separate these data points by the category they are supposed to belong to, but you don't want the data to be in the wrong category. This means you are trying to find the line between the two closest points that keeps the other data points apart.

IV. RESULTS

To reduce the suffering of patients with Parkinson's disease, we offer a mobile application that detects symptoms such as attention deficit, poor response to calls and fine motor deficits. These symptoms can be assessed using defined activities performed by normal and patient groups. To predict the outcome of each campaign, we used the KNN algorithm, which will be t rained using the normal group data and thresholds will be set for the patient group.

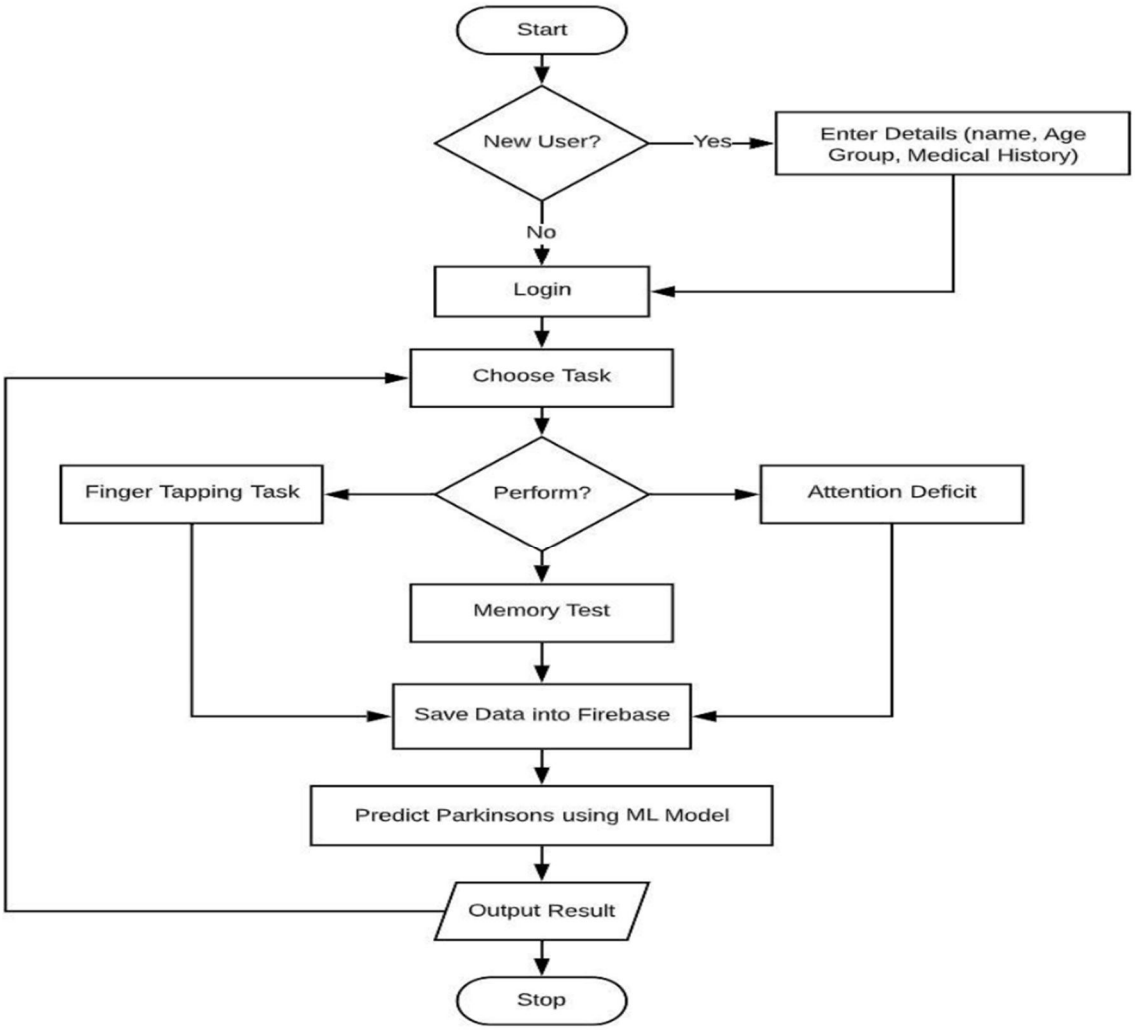


Fig 1: Flow Diagram

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

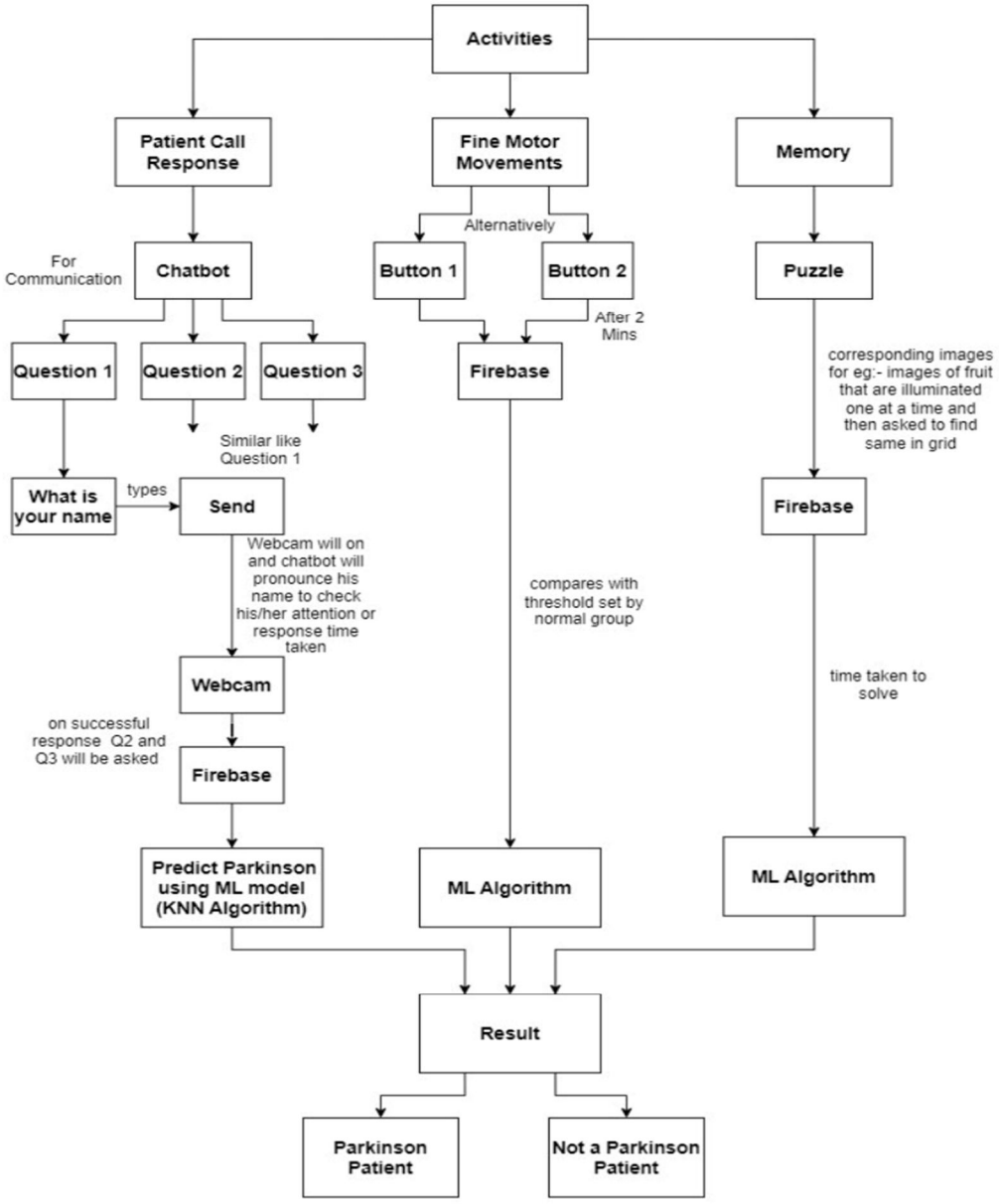


Fig 2: Activity Diagram

V. DISCUSSION

We will use the KNN algorithm and the SVM algorithm, the abbreviation KNN stands for "K_Nearest Neighbor". It is a supervised machine learning algorithm. This algorithm can be used to solve classification and regression problems. The number of nearest neighbors for a new unknown variable to predict or classify is indicated by the symbol "K".

Before we start with this amazing algorithm, let's look at a relevant real life scenario. We're often told that you have a lot in common with your closest colleagues, whether it's your thought process, work etiquette, philosophy, or other factors. As a result, we form friendships with people we think are similar to us. The KNN algorithm uses the same principle. Its purpose is to locate all nearest neighbors around a new unknown data pointto find out which class it belongs to. This is a distance-based method.

Support Vector Machines or SVMs are one of the most popular learning algorithms for classification and regression problems. However, it is mainly used for classification problems in machine learning.

The goal of the SVM algorithm is to create an optimal line or decision boundary that can divide the n_dimensional space int o classes so that we can easily assign new data points to the right class at the coming. This optimal decision boundary is call ed a hyperplane SVM selects extreme points/vectors which help to create a hyperplane. These extreme cases are called support vectors, so the algorithm is called a support vector machine.

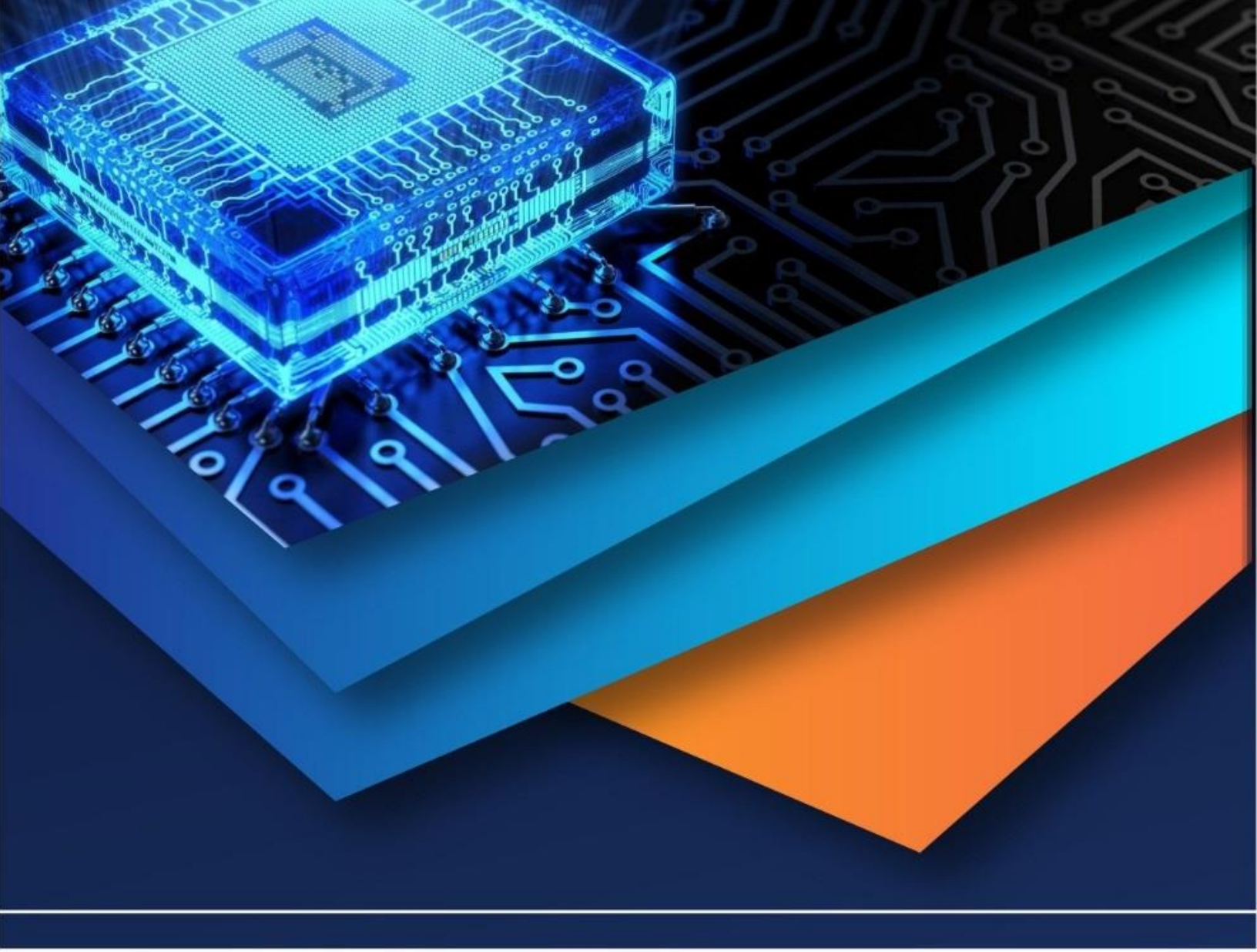
VI. CONCLUSION

We have successfully completed the project with the help of machine learning algorithms which are svm and knn which help us to built a model with providing a accurate results. We created some activities like games which are fun to do and we can understand if the patient is Parkinson patient or not. The UI designs with login registration pages and some activities like finger tapping, puzzle, chatbot etc. Through SVM and KNN we got accuracy of 90%.

REFRENCES

[1] Prashanth R, “A Study on Parkinson’s Disease andSWEDD”, IEEE Access, vol.7, 20, Sep 2017
[2] Rostom Mabrouk, Belkacem Chikhaoui, Layachi Bentabet,” Machine Learning Based Classification Using Clinical and DaTSCAN SPECT Imaging Features: A Study on Parkinson’s Disease and SWEDD” IEEE Access, vol.3, 23, Oct 2018
[3] Mosarrat Rumman, Abu Nayeem Tasneem, Sadia Farzana, Monirul Islam Pavel and Dr. Md. Ashraful Alam” Early detection of Parkinson's disease using image processing and artificial neural network” IEEE International Conference on Computing, 2018.
[4] “Parkinson Disease Prediction Using Machine Learning Algorithm”. Mathur, R., Pathak, V., &Bandil, D. In Emerging Trends in Expert Applications and Security”pp.357-363.Springer, Singapore. 2019.
[5] Timothy J. Wroge et al “Parkinson’s disease diagnosis using machine learning and voice”,2018
[6] M Venkateshwar rao “Diagnosis of Parkinson disease using machine learning and data mining system from voice dataset”,2015
[7] Arvind kumar Tiwari “Machine learning based approaches for prediction of Parkinson disease”,2016
[8] H. Gunduz, “Deep learning-based parkinson’s disease classification using vocal feature sets,” IEEE Access, vol. 7, pp. 115 540– 115 551, 2019.
[9] O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, M. Tutuncu, T. Aydin, M. E. Isenkul, and H. Apaydin, “A comparative analysis of speech signal processing algorithms for parkinson’s disease classification and the use of the tunable q-factor wavelet transform,” Applied Soft Computing, vol. 74, pp. 255–263, 2019.
[10] H. Gunduz, “Deep learning-based parkinson’s disease classification using vocal feature sets,” IEEE Access, vol. 7, pp. 115 540–115 551, 2019

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)