DETECTION OF PARKINSON'S DISEASE THROUGH A MACHINE LEARNING APPROACH

Ms. Rekha Jadhav, PG Scholar, Dept. Of Computer Engineering, Trinity College of Engineering and Research, Pune.
Prof. Prasad Bhosle, Assistant Professor, Dept. Of Computer Engineering, Trinity College of Engineering and Research, Pune.

Abstract
Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms such as tremors, bradykinesia, rigidity, and postural instability. Early diagnosis and intervention are crucial for managing the disease effectively and improving the quality of life for patients. This review provides a comprehensive overview of the various methods and technologies employed in the early detection of Parkinson's disease. This review underscores the significance of early detection in Parkinson's disease and emphasizes the need for ongoing research and collaboration to enhance diagnostic accuracy and enable timely interventions. Early detection not only improves the quality of life for patients but also opens avenues for targeted therapies and disease-modifying treatments, bringing us closer to the ultimate goal of finding a cure for Parkinson's disease.

Keywords: *Parkinson's Disease, Neuroimaging Techniques, Machine Learning, Convolutional Neural Network, Parkinson's Dataset*.

I. Introduction
Parkinson's illness (PD) is an ongoing neurodegenerative problem that influences a huge number of people around the world. It is characterized by a progressive loss of motor control, leading to symptoms such as tremors, bradykinesia, rigidity, and postural instability [9]. Early and accurate diagnosis of PD is crucial for timely intervention, personalized treatment planning, and improved quality of life for patients. Traditional diagnostic methods often rely on clinical evaluation, which may lack the sensitivity required for early detection. With the advent of advanced technologies, particularly machine learning (ML) and artificial intelligence (AI), there has been a paradigm shift in the way Parkinson's disease is diagnosed and managed [8].
This paper presents an in-depth exploration of the application of machine learning in the detection of Parkinson's disease. It discusses the challenges associated with traditional diagnostic methods, emphasizing the limitations that hinder early detection. The paper highlights the potential of machine learning algorithms in processing diverse datasets and extracting meaningful features associated with PD. By reviewing recent studies and methodologies, the paper aims to provide a comprehensive overview of the advancements in Parkinson's disease detection facilitated by machine learning techniques [1].
The impact of these advancements on clinical practice, emphasizing the integration of machine learning models into routine diagnostic protocols. The potential benefits, including increased accuracy, reduced diagnosis time, and improved patient outcomes, underscore the transformative power of machine learning in the early detection and management of Parkinson's disease. Additionally, ethical considerations, data privacy, and challenges related to the implementation of these technologies in real-world healthcare settings are also addressed, providing a holistic view of the landscape surrounding Parkinson's disease detection using machine learning.
This paper elucidates the pivotal role of machine learning in reshaping the landscape of Parkinson's Disease detection. By harnessing the capabilities of machine learning algorithms, healthcare professionals can make significant strides towards early and accurate diagnosis, ultimately enhancing the lives of individuals affected by this debilitating neurodegenerative disorder.

II. Problem Definition

Parkinson's disease (PD) is a neurodegenerative disorder affecting millions globally. Early and accurate detection is crucial for timely intervention and improved patient outcomes. However, current diagnosis methods are subjective, rely on specialist evaluation, and often miss the disease in its early stages. Current methods for diagnosing PD are subjective and rely heavily on specialist evaluation, leading to potential inconsistencies and delays in diagnosis. Early stage symptoms can be subtle and mimic other conditions, making accurate detection difficult. No single definitive diagnostic test exists for PD. To delayed diagnoses, hindering timely interventions and personalized treatment strategies. Additionally, the absence of standardized criteria and the need for more accurate, non-invasive diagnostic tools pose significant obstacles in effective PD detection. There's potential to overcome the challenges in Parkinson's disease detection, leading to earlier and more accurate diagnoses, personalized treatments, and improved outcomes for individuals affected by PD.

III. Literature Survey

WU WANG , JUNHO LEE [1], "Early Recognition of Parkinson's Illness Utilizing Profound Learning and AI", Precisely recognizing Parkinson's infection (PD) at a beginning phase is absolutely crucial for dialing back its encouraging and giving patients the chance of getting to sickness changing treatment. Towards this end, the premotor stage in PD should be carefully checked. An imaginative profound learning procedure is acquainted with early reveal regardless of whether an individual is impacted with PD in view of premotor highlights. In particular, to uncover PD at a beginning phase, a few pointers have been viewed as in this review, including Fast Eye Development and olfactory misfortune, cerebrospinal liquid information, and dopaminergic imaging markers.

According to Kanakaprabha.S and Arulprakash. P [2], titled "Parkinson Disease Detection Using Various Machine Learning Algorithms," explored the neural aspects of Parkinson's disease.It prompts shaking of the hands, trouble to walk, offset with coordination. No clinical treatment is accessible in the significant level stage. X-beam, CT output and blood tests report are not adequately results accessible in the beginning phase. Around two trillion local area are alive in Parkinson's sickness (PD) in the U.K., which is the largest number of individuals impacted are pinpointed to have different sclerosis, strong dystrophy and Lou Gehrig's sickness. This is depended upon to rise to 1.5 million by 2040. Around the 75,000 Americans are analysis PD with each year.

Asmae Ouhmida, Abdelhadi Raihani [3], "Parkinson's illness arrangement utilizing AI calculations: execution examination and correlation", Discovery of Parkinson's sickness remains challenge for doctors, particularly, in the clinical field because of the trouble of fix. Subsequently, calculations of characterization play the primary part in the appraisal of this neurodegenerative issue. In this paper, we center around the examination and the assessment of nine AI Calculations (MLA), specifically Backing Vector Machine (SVM), Strategic Relapse, Discriminant Investigation, K-Closest Neighbors (KNN), Choice tree, Arbitrary Woodland, Packing tree, Gullible Bayes, and AdaBoost.

M.S. Roobini1 , Yaragundla Rajesh Kumar Reddy [4], "Parkinson's Disease Detection Using Machine Learning", Firstly, Parkinson delineates Parkinson's sickness as a neurologic syndrome, it affects the central system, the patients face trouble talking, walking, quake all through the movement. Parkinson's disorder patient by and large envelops a low-volume commotion with a droning quality, this strategy investigates the order of sound signs highlight dataset to analyze Parkinson's infection (PD), the classifiers we will generally use during this framework region unit from AI.

Sandhiya S, Dr.Ashok.S [5], "Parkinson's Disease Prediction Using Machine Learning Algorithm", Parkinson's disease (PD) is a progressive central nervous system disorder characterized by tremors, vocal cord disorders, bradykinesia, and slurred, slow speech. After Alzheimer's illness, it is the second most pervasive neurological problem. Approximately 63 million people worldwide are affected by this disease. It is brought about by the passing of dopamine-creating neurons and fundamentally influences those beyond 60 years old. Parkinson's infection doesn't have a fix, yet early location might have the option to slow the sickness' movement. Utilizing discourse examination and walk investigation

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 3, April : 2024

information, analysts have gone through the beyond couple of years dealing with the discovery and checking of Parkinson's sickness.

IV. Existing System

Prediction of Parkinson disorder is one of the most important problem that has to be detected in the early phases of the commencement of the disease so as to reduce the disease progression rate among the individuals .Various researches have been made to find the basic cause and some have reached to the heights by proposing a system which differentiates the healthy people from those with any ND'S (Neurodegenerative disorders) using various machine learning techniques. Lots of pre-processing feature selection and classification techniques have been implemented and developed in the past decades.

V. Drawbacks

These drawbacks can lead to delayed diagnoses, hindering the initiation of treatment and potentially impacting the long-term prognosis of patients. They also highlight the need for more objective, accessible, and reliable methods for early detection of Parkinson's disease. These are promising new areas, but they are still under development and require further validation to ensure accuracy and widespread adoption. Data collection and analysis methods need refinement to improve their effectiveness.

VI. Proposed System

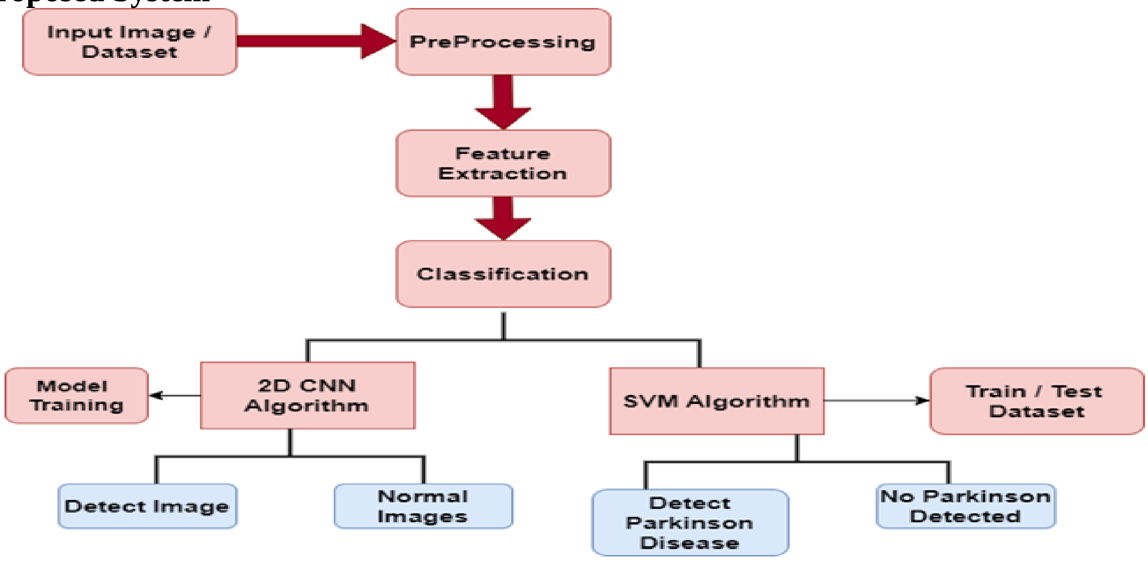


Fig.1- Proposed system

6.1. Input: The initial phase involves data gathering, a crucial step as the quality and quantity of information collected directly influence the scope and accuracy of your prediction model.

6.2. Data pre-processing: Data pre-processing involves visualizing the information to discern relationships among the parameters within the dataset, as well as to address data imbalances. Additionally, data needs to be partitioned into two segments: one for training the model and the other for testing. For instance, in our model, we allocated 70% of the data for training and reserved 30% for testing.

6.3. Feature Selection: In the subsequent phase of our workflow, we engage in feature selection. Researchers and scientists have utilized diverse models tailored for different purposes, including image processing, sequence analysis (such as text, numbers, or patterns), among others. In our scenario, we have delineated samples from Parkinson's disease (PD) patients sourced from various individuals. Consequently, we opt for models capable of distinguishing between unhealthy patients and healthy individuals.

6.4. Training: Training the dataset stands as a pivotal task in machine learning. Through this process, we iteratively enhance the selected model's capability to predict outcomes more accurately, striving for a closer approximation between the predicted and actual results.

6.5. CNN model: Convolutional Neural Networks (CNNs) demonstrate proficiency in processing grid-like data, rendering them well-suited for image analysis tasks. Their capacity to automatically discern patterns relevant to Parkinson's disease within image data is particularly noteworthy. Moreover, CNNs have the potential to identify subtle movement patterns linked to Parkinson's disease through comprehensive analysis.

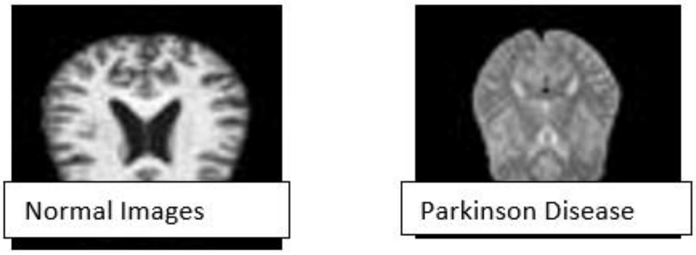
6.6. SVM Algorithm: We have computed various metrics including ROC, Accuracy, Specificity, Precision, and others. These metrics serve to identify the most effective algorithm among the options considered.

6.7. Prediction: During this phase, we prepare the model to accurately detect the presence of Parkinson's disease utilizing the provided dataset.

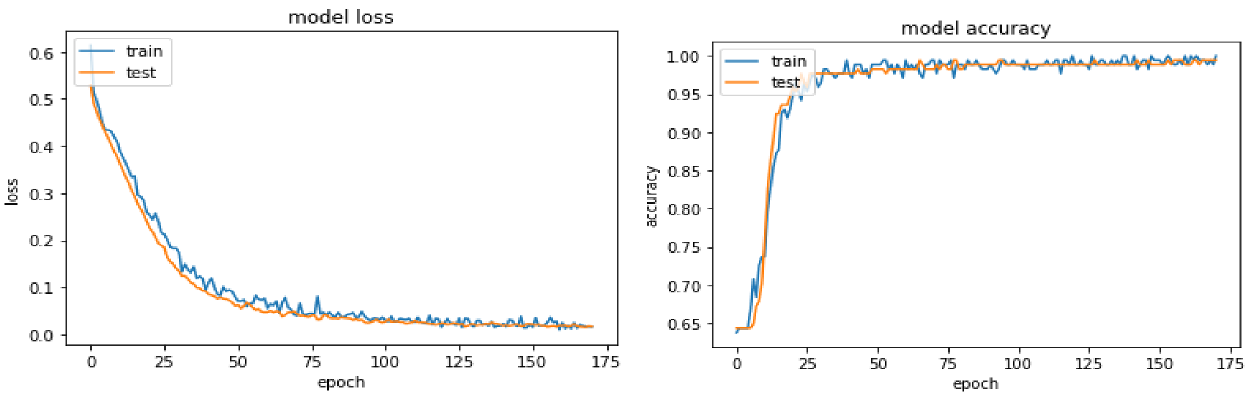
VII. Concept

We outlined essential performance metrics and various datasets employed in diagnosing Parkinson's disease, alongside a comparative analysis of their outcomes. This presentation facilitates the selection of significant datasets, performance assessment metrics, and cross-validation methods by researchers, considering their prevalence and applicability within the field. Additionally, we assessed the experimental results of various Parkinson's disease diagnosis approaches using consistent datasets and performance evaluation criteria. With this contribution, researchers will be able to easily discover proposed models that perform well.

7.1. Images of Parkinson Disease Detection



7.2. Accuracy/Loss in Plot:



Testing Accuracy: 98.83%
Training Accuracy: 99.42%

7.3. Dataset of Parkinson Disease Detection:

Attributes	Purpose
gender	Male/ Female

PPE	Pitch Period entropy
DFA	Detrended Fluctuation Analysis
RPDE	Normalized Root Power of the Decimated Waveform
numPulses	Number of Pulses
numPeriodsPulses	Number of Pitch Periods
meanPeriodPulses	Consecutive Pitch Periods
minIntensity	Minimum Intensity
maxIntensity	Maximum Intensity Level
app_LT_entropy_log_coef	Entropy of Voice Signal
tqwt_kurtosisValue_dec	Discrete Wavelet Transform

7.4. Accuracy of SVM Algorithm:

Metric	**SVM**
Precision	**0.89**
Re-call	**0.88**
F1-score	**0.88**
Support	**40**

Total Accuracy of SVM will be 87.5%

VIII. Technology

Python finds applications in server-side web development, software engineering, mathematical computations, and system scripting. It is notably favoured for Rapid Application Development (RAD) and serves as a scripting or integration language for connecting existing components, owing to its high-level nature, comprehensive built-in data structures, dynamic typing, and dynamic binding. Python's readability and straightforward syntax contribute to reduced program maintenance costs. Furthermore, its support for modules and packages fosters modular programming and code reusability. Being an open-source language with a vibrant community, Python benefits from continuous contributions by independent programmers who develop libraries and expand its functionalities. Professionally, Python is great for backend web development, data analysis, artificial intelligence, and scientific computing. Developers also use Python to build productivity tools, games, and desktop apps.
On the down side, Python isn't easy to maintain. In Python, due to its dynamic typing nature, a single command can hold multiple interpretations based on the context. As a Python application expands in scale and intricacy, the task of maintenance becomes progressively challenging, particularly in identifying and rectifying errors. Users are advised to possess sufficient experience to craft code and devise unit tests that streamline the maintenance process.
Python's speed is often considered a drawback. Its dynamic typing necessitates extensive referencing to determine the correct definition, which can hinder performance. However, this limitation can be alleviated by opting for alternative implementations of Python, such as PyPy.

Software Requirements Specification

- Coding Language : Python
- Operating System : Windows 10

Hardware Requirements Specification

- Processor : Pentium-IV
- RAM :512 MB(min)
- Hard Disk : 40 GB
- Key Board : Standard Windows Keyboard
- Mouse : Two or Three Button Mouse

- Monitor : LCD/LED

IX. Algorithm

9.1. A Convolutional Neural Network (CNN) is a type of Deep Learning neural network architecture commonly used in Computer Vision. Computer vision, a branch of Artificial Intelligence, empowers computers to comprehend and interpret visual data or images.

In a regular Neural Network there are three types of layers:

1. **Input Layers**: This layer serves as the entry point for data into our model. The number of neurons within this layer corresponds to the total number of features present in our dataset (in the case of an image, this would be the number of pixels).
2. **Hidden Layer**: Following the Input layer, the data is transmitted to the hidden layer. The number of hidden layers can vary based on our model and dataset dimensions. Each hidden layer may contain a different number of neurons, typically exceeding the number of features. Output from each layer is calculated through matrix multiplication of the preceding layer's output with the learnable weights specific to that layer, followed by the addition of learnable biases. Subsequently, an activation function is applied, imparting nonlinearity to the network.
3. **Output Layer**: The output from the hidden layer undergoes processing through a logistic function such as sigmoid or softmax. This transformation converts the output of each class into a probability score, facilitating the classification of each class.

9.2. Support Vector Machine The Support Vector Machine (SVM) represents a modern learning system rooted in recent advancements in statistical learning theory. It operates effectively on both linear and non-linear datasets. SVM achieves this by transforming the original data into a higher-dimensional space, enabling the identification of a hyperplane for data separation using essential training instances termed support vectors. Formally, SVM is a discriminative classifier defined by a separating hyperplane. In essence, given labeled training data, SVM constructs a hyperplane or a set of hyperplanes in a high-dimensional or infinite-dimensional space, applicable for classification, regression, or similar tasks. In essence, optimal separation is achieved by the hyperplane that maintains the maximum distance to the nearest training data point of any class, referred to as the functional margin. A larger margin generally corresponds to lower generalization error for the classifier.

X. Implementation Part: stages, python

CNN coding:

```
def train_model():
    update_label("Model Training Start...............")
    start = time.time()
    X= CNNModelp.main()
    end = time.time()

    ET="Execution Time: {0:.4} seconds \n".format(end-start)
    msg="Model Training Completed.."+'\n'+ X + '\n'+ ET
    print(msg)
   import functools
   import operator


def convert_str_to_tuple(tup):
    s = functools.reduce(operator.add, (tup))
    return s
```

```
    def test_model_proc(fn):
    from keras.models import load_model
    #from keras.optimizers import Adam

#   global fn

    IMAGE_SIZE = 64
    LEARN_RATE = 1.0e-4
    CH=3
    print(fn)
    if fn!="":
        # Model Architecture and Compilation

        model = load_model('parkisons.h5')

        img = Image.open(fn)
        img = img.resize((IMAGE_SIZE,IMAGE_SIZE))
        img = np.array(img)

        img = img.reshape(1,IMAGE_SIZE,IMAGE_SIZE,3)

        img = img.astype('float32')
        img = img / 255.0
        print('img shape:',img)
        prediction = model.predict(img)
        print(np.argmax(prediction))
        plant=np.argmax(prediction)
        print(plant)

        if plant == 0:
            Cd="Parkisons Disease Not Detected"
        elif plant == 1:
            Cd="Parkisons Disease Detected”

        A=Cd
        return A
```

SVM Model:

```
def Model_Training():
data = pd.read_csv("dataset.csv")
data.head()

data = data.dropna()
"""One Hot Encoding"""

"""Feature Selection => Manual"""
x = data.drop(['id','class'], axis=1)
data = data.dropna()

print(type(x))
```

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 3, April : 2024

```
y = data['class']
print(type(y))
x.shape
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.20, random_state=123456789)

from sklearn.svm import SVC
svcclassifier = SVC(kernel='linear')
svcclassifier.fit(x_train, y_train)

y_pred = svcclassifier.predict(x_test)
print(y_pred)


print("=" * 40)
print("==========")
print("Classification Report : ",(classification_report(y_test, y_pred)))
print("Accuracy : ",accuracy_score(y_test,y_pred)*100)
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
ACC = (accuracy_score(y_test, y_pred) * 100)
repo = (classification_report(y_test, y_pred))

label4 = tk.Label(root,text =str(repo),width=45,height=15,bg='seashell2',fg='black',font=("Tempus
Sanc ITC",14))
label4.place(x=250,y=200)

label5 = tk.Label(root,text ="Accuracy : "+str(ACC)+"%\nModel saved as
PD_model.joblib",width=45,height=2,bg='black',fg='white',font=("Tempus Sanc ITC",14))
label5.place(x=250,y=490)
from joblib import dump
dump (svcclassifier,"PD_model.joblib")
print("Model saved as PD_model.joblib")
```

XI. Result O/P

```
Classification Report :              precision    recall  f1-score   support

           0       0.40      0.50      0.44         4
           1       0.94      0.92      0.93        36

    accuracy                           0.88        40
   macro avg       0.67      0.71      0.69        40
weighted avg       0.89      0.88      0.88        40

Accuracy :  87.5
Accuracy: 87.50%
Model saved as PD_model.joblib
```

Fig.2- SVM Result.

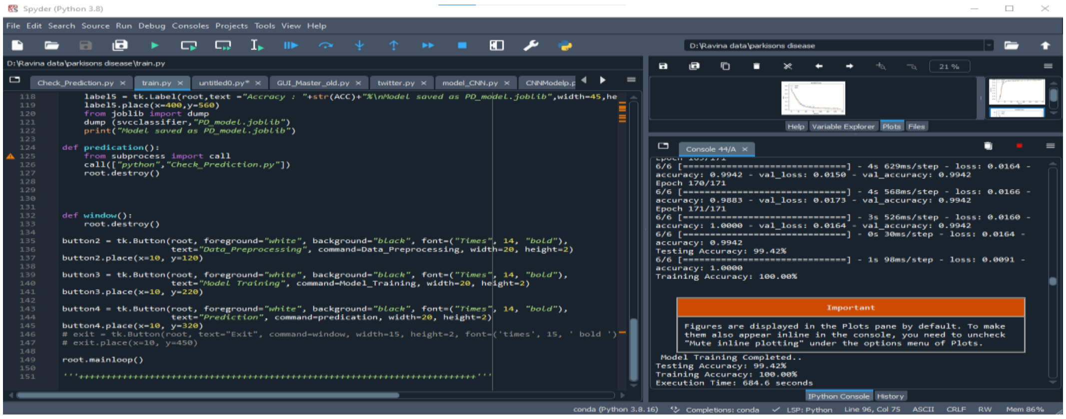


Fig.3- CNN Result

XII. Screenshots of O/P

Fig.4- Login Page

Fig.5- Trained Data

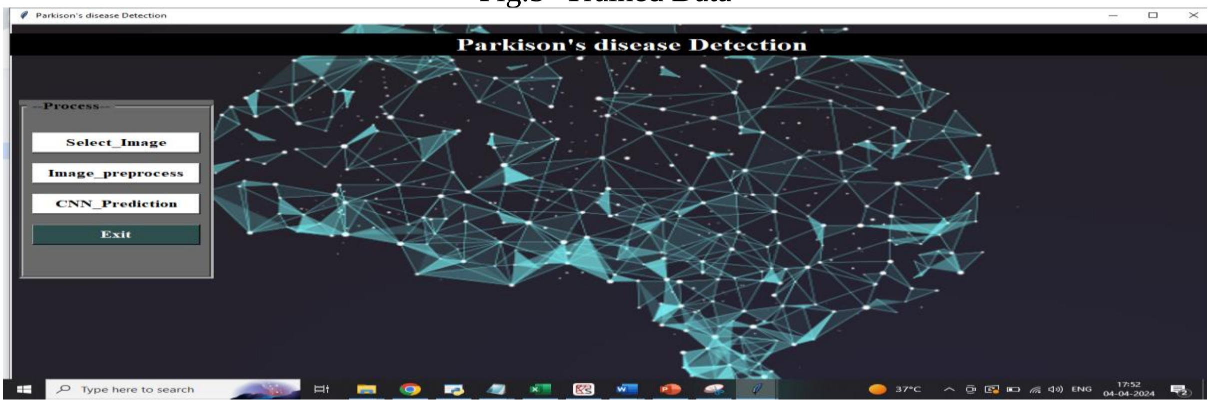


Fig.6- Input Screen

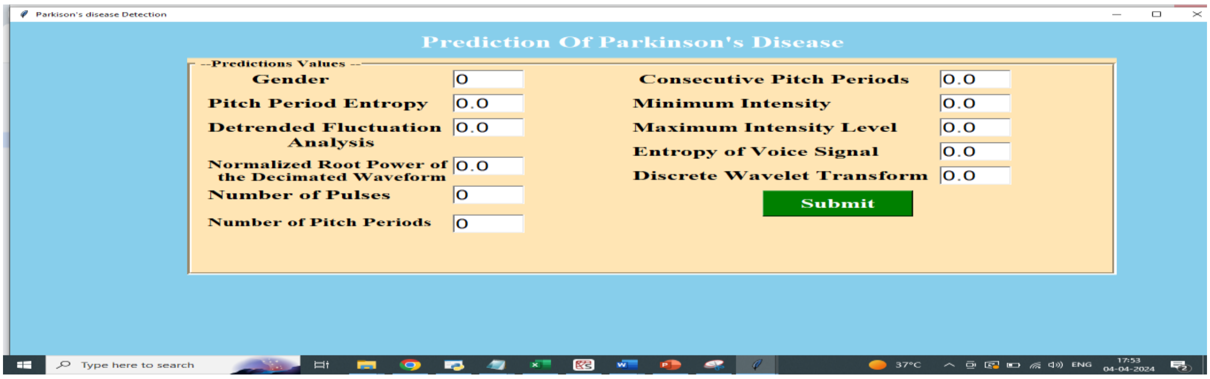


Fig.7- Prediction of Disease

XIII. Conclusion

The model is organized with the dataset of Parkinson's disease by machine learning algorithms. Parkinson's disease dataset is taken and analysed to predict the asperity of the disease. A Decision tree approach is used to predict the disease. The data in the dataset is pre-processed to make it suitable for classification. The CNN and SVM approach to generate efficient classification rules is proposed. Machine learning technique is a multilayer perceptron that is the special design for identification of Parkinson's disease parameter information.

References

[1]Wu Wang , Junho Lee , Fouzi Harrou , (Member, Ieee), And Ying Sun, Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning 2020 IEEE Access Volume: 8, doi: 10.1109/Access.2020.3016062.

[2]Kanakaprabha. S. Arulprakash P,Srikanth R,Parkinson Disease Detection Using Various Machine Learning Algorithms, IEEE 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA),2022, doi: 10.1109/ICACTA54488.2022.9752925.

[3]Asmae Ouhmida; Abdelhadi Raihani; Bouchaib Cherradi; Yasser Lamalem. Parkinson's disease classification using machine learning algorithms: performance analysis and comparison, IEEE 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), 2022,DOI:10.1109/IRASET52964.2022.9738264.

[4]M.S. Roobini, Yaragundla Rajesh Kumar Reddy, Udayagiri Sushmanth Girish Royal, Amandeep Singh K, Babu.K, "Parkinson's Disease Detection Using Machine Learning", 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), IEEE Conference, 2022.

[5]Sandhiya S and Dr. Ashok.S," Parkinson's Disease Prediction Using Machine Learning Algorithm", December 2022, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), DOI:10.1109/ICPECTS56089.2022.10047447.

[6]Dr. C k gomathy, mr. B. Dheeraj kumar reddy, ms. B. Varsha, ms. B. Varshini , the parkinson's disease detection using machine learning techniques,2021, International Research Journal of Engineering and Technology, Volume: 08 Issue: 10.

[7]Dickson, D.W. Neuropathology of Parkinson disease. Parkinsonism Relat. Disord. 2018, 46 (Suppl. 1), S30–S33.

[8]M. Shaban, "Deep Convolutional Neural Network for Parkinson's Disease Based Handwriting Screening," 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), 2020, pp. 1-4, doi: 10.1109/ISBIWorkshops50223.2020.9153407.

[9] J. Goyal, P. Khandnor and T. C. Aseri, "Analysis of Parkinson's disease diagnosis using a combination of Genetic Algorithm and Recursive Feature Elimination," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 268-272, doi: 10.1109/WorldS450073.2020.9210415.