

**Parkinson Disease Detection**

**(BSCS)**

Department of Computer Science Lahore Garrison University

**Submitted By:**

S# Name Registration /Roll /Section

1. Muhammad Mateen Mujahid Sp-2022-BSCS-170/C
2. Muhammad Sohail Shahzad Sp-2022-BSCS-99/C

**Submitted to:**

Dr. Sadaf Hussain

**1. Abstract**

Parkinson's illness is a neurodegenerative problem that influences development, equilibrium, and coordination. Early finding is urgent for powerful administration and working on the personal satisfaction for people with Parkinson's. This study investigates the utilization of AI calculations for the expectation of Parkinson's illness utilizing voice and discourse information. The dataset used in this study contains different biomedical voice estimations from people with and without Parkinson's. some preprocessing method has to be performed on data including dealing with missing values, eliminating unimportant features, balancing the data and normalization of data. The machine learning models which are utilized are: k-nearest Neighbors (k-NN), Logistic Regression, Naive Bayes, Decision Tree , Support Vector Machine and some methods of Ensemble Learning. These models were trained and their performance are measured using performance metrics : accuracy, Precision, Recall and F1-score. The outcomes showed promising execution for all models, with k-NN accomplishing the most noteworthy accuracy. Furthermore , Ensemble learning techniques have also showed better accuracy. This study features the capability of AI in helping the early detection of Parkinson's disease, preparing for further developed medical services results. Future work will be on investigating strategies and consolidating extra clinical information to upgrade the forecast model.

**2. Introduction**

Parkinson's illness is a dynamic neurological problem that fundamentally influences development. Early analysis is pivotal for successful administration and working on the personal satisfaction for people with Parkinson's. AI has arisen as a promising device for illness expectation and determination. This study explores the utilization of different AI calculations to foresee Parkinson's infection utilizing voice and discourse information. By investigating biomedical voice estimations, the point is to foster a model that can precisely recognize people with Parkinson's. The review assesses the presentation of individual models and investigates group learning methods to upgrade expectation precision possibly. The discoveries of this study add to the progression of AI based demonstrative instruments for Parkinson's sickness, at last intending to further develop medical care results.

**3. Literature Review**

| **Model** | **Accuracy (Your Study)** | **Accuracy (Other Studies)** | **Reference** |
| --- | --- | --- | --- |
| k-NN | **98.3%** | **91.5%** | Caliskan et al. (2017) |
| Logistic Regression | **83.1%** | **85.2%** | Speelman et al. (2013) |
| Naive Bayes | **83.1%** | **81.0%** | Bott et al. (2014) |
| Decision Tree | **83.1%** | **86.7%** | Goldman et al. (2015) |
| Ensemble Learning (Hard) | **93.2%** | **92.0%** | Grover et al. (2018) |
| Ensemble Learning (Soft) | **98.3%** | **95.4%** | Grover et al. (2018) |
| Boosting | **96.6%** | **93.8%** | Goldman et al. (2015) |
| Ensemble Learning (Stacking) | **98.3%** | **96.0%** | Grover et al. (2018) |
| SVM | **98.3%** | **85.0%** | Sachin Shetty et al. (2016) |
| Random Forest | - | **92.0%** | Goldman et al. (2015) |
|  |  |  |  |

**4. Methodology**

**4.1. Proposed Model Diagram**

**A computer screen shot of a diagram

Description automatically generated**

**4.2 Dataset Description**

The dataset utilized in this study is the "Parkinson Disease" dataset. It involves 195 rows and 24 columns .It contains different biomedical voice estimations from people with and without Parkinson's. The elements incorporate proportions of vocal key recurrence, jitter, gleam, and other discourse related boundaries. The objective variable is a double marker addressing the presence or nonattendance of Parkinson's infection.

The features in dataset are:

name - ASCII subject name and recording number

MDVP:Fo(Hz) - Average vocal fundamental frequency

MDVP:Fhi(Hz) - Maximum vocal fundamental frequency

MDVP:Flo(Hz) - Minimum vocal fundamental frequency

MDVP:Jitter(%), MDVP:Jitter(Abs), MDVP:RAP, MDVP:PPQ, Jitter:DDP - Several measures of variation in fundamental frequency

MDVP:Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA - Several measures of variation in amplitude

NHR, HNR - Two measures of the ratio of noise to tonal components in the voice

status - The health status of the subject (one) - Parkinson's, (zero) - healthy

RPDE, D2 - Two nonlinear dynamical complexity measures

DFA - Signal fractal scaling exponent

spread1, spread2, PPE - Three nonlinear measures of fundamental frequency variation

**4.3. Description of Models**

• k-Nearest Neighbors (k-NN): A non-parametric algorithm that classifies data points based on the majority class among their nearest neighbors.

• Logistic Regression: A linear model that predicts the probability of a binary outcome using a logistic function.

• Naive Bayes: A probabilistic classifier based on Bayes' theorem, assuming independence between features.

• Decision Tree: A tree-based model that makes predictions by recursively partitioning the data based on feature values.

• Ensemble Learning (Voting Classifier): Combines the predictions of multiple base learners using a voting mechanism to improve overall performance.

• Support Vector Machine (SVM): A supervised learning algorithm that finds an optimal hyperplane to separate data points into distinct classes by maximizing the margin between the closest points (support vectors) of each class.

**5. Results and Discussion**

**5.1. Environment Setting**

The study was conducted using Python and various machine learning libraries, including scikit-learn, in a Google Collab environment.

**5.2. All Performance Metrics Obtained**

The performance of each model was evaluated using accuracy, precision, recall, and F1-score.

**5.3. Discussion of Results of each ML Model**

**1. KNN (K-Nearest Neighbors)**

- Accuracy: 0.983

- Precision: 0.967

- Recall: 1.0

- F1-score: 0.983

**Discussion:**

KNN performs very well in terms of overall accuracy (98.3%), indicating that it correctly classifies most instances. Its precision is also high (96.7%), meaning that the positive predictions made by the model are generally correct. The perfect recall of 1.0 shows that KNN is able to identify all relevant instances of the positive class without missing any. This is ideal for applications where it’s critical not to miss any positives, such as medical diagnoses. The F1-score is also very high at 0.983, suggesting a good balance between precision and recall.

**2. Logistic Regression**

- Accuracy: 0.831

- Precision: 0.852

- Recall: 0.793

- F1-score: 0.821

**Discussion:**

Logistic Regression performs moderately well, with an accuracy of 83.1%. However, its precision (85.2%) is higher than recall (79.3%), indicating that while it is good at identifying positive instances when it makes a prediction, it misses a fair number of actual positive cases. The relatively lower recall suggests the model might not be sensitive enough to detect all relevant positive instances. The F1-score of 0.821 indicates that there is a moderate trade-off between precision and recall.

**3. Naive Bayes**

- Accuracy: 0.831

- Precision: 0.852

- Recall: 0.793

- F1-score: 0.821

**Discussion:**

The performance of Naive Bayes is identical to Logistic Regression in all metrics. This means it exhibits the same trade-off between precision and recall, with a moderate F1-score. It also has similar limitations, as the recall is not as high as desired for sensitive applications, leading to missed positive instances.

**4. Decision Tree**

- Accuracy: 0.831

- Precision: 0.852

- Recall: 0.793

- F1-score: 0.821

**Discussion:**

Decision Tree, like Logistic Regression and Naive Bayes, performs similarly with respect to accuracy, precision, recall, and F1-score. Its performance suggests it is not highly tuned or optimized for this particular task. Decision Trees are often prone to overfitting, which may lead to suboptimal performance on unseen data.

**5. Ensemble Learning (Hard)**

- Accuracy: 0.932

- Precision: 1.0

- Recall: 0.862

- F1-score: 0.926

**Discussion:**

Ensemble Learning (Hard) performs well with an accuracy of 93.2%. The precision is perfect (1.0), meaning all positive predictions made by the model are correct. However, the recall is somewhat lower (86.2%), which indicates that while the model is very confident about positive predictions, it does miss some positive cases. Despite this, the high F1-score of 0.926 suggests a good balance between precision and recall.

**6. Ensemble Learning (Soft)**

- Accuracy: 0.983

- Precision: 1.0

- Recall: 0.966

- F1-score: 0.982

**Discussion:**

Ensemble Learning (Soft) performs nearly as well as KNN with an accuracy of 98.3%. It achieves perfect precision (1.0) and a very high recall of 96.6%, making it almost as ideal as KNN for applications where both precision and recall are critical. The F1-score of 0.982 reflects an excellent balance and performance across all metrics.

**7. Boosting**

- Accuracy: 0.966

- Precision: 0.966

- Recall: 0.966

- F1-score: 0.966

**Discussion:**

Boosting performs evenly across all metrics with an accuracy of 96.6%. Its precision, recall, and F1-score are all equal (96.6%), indicating that the model is well balanced, without favoring one metric over the other. While this balance is generally desirable, it may not reach the high recall or perfect precision seen in other models like Ensemble Learning (Soft).

**8. Ensemble Learning (Stacking)**

- Accuracy: 0.983

- Precision: 1.0

- Recall: 0.966

- F1-score: 0.982

**Discussion:**

Stacking Ensemble Learning also performs excellently, similar to the Soft ensemble, with an accuracy of 98.3%. It achieves perfect precision (1.0) and a recall of 96.6%, yielding an F1-score of 0.982. This is another highly balanced model with almost ideal performance for tasks requiring both high precision and recall.

**8. Support Vector Machine(SVM)**

- Accuracy: 0.983

- Precision: 1.0

- Recall: 0.966

- F1-score: 0.982

**Discussion:**

SVM achieves excellent performance with an accuracy of 98.3%, indicating that it correctly classifies the vast majority of instances. The high precision of 96.7% reflects that most of the positive predictions made by the model are accurate, minimizing false positives. The recall of 1.0 demonstrates that SVM successfully identifies all positive instances without missing any, making it suitable for tasks where it’s critical to capture all positives. The F1-score of 0.983 confirms a strong balance between precision and recall, highlighting the model's reliability in diverse scenarios.

**Summary:**

- KNN and Ensemble Learning (Soft/Stacking) are the top performers in terms of precision, recall, and F1-score, with KNN achieving the highest recall.

- Ensemble Learning (Hard) provides perfect precision but at the cost of slightly lower recall, still offering very good overall performance.

- Logistic Regression, Naive Bayes, and Decision Tree show similar moderate performance with an accuracy around 83%, but they suffer from lower recall, meaning they miss many positives.

- Boosting shows a good balance across metrics but doesn’t excel in any particular one.

Overall, ensemble methods (Soft and Stacking) provide the best overall performance, followed by KNN.

**5.4. Overall Result in Tabular Form**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| KNN | 0.983 | 0.967 | 1.0 | 0.983 |
| Logistic Regression | 0.831 | 0.852 | 0.793 | 0.821 |
| Naive Bayes | 0.831 | 0.852 | 0.793 | 0.821 |
| Decision Tree | 0.831 | 0.852 | 0.793 | 0.821 |
| Ensemble Learning (Hard) | 0.932 | 1.0 | 0.862 | 0.926 |
| Ensemble Learning (Soft) | 0.983 | 1.0 | 0.966 | 0.982 |
| Boosting | 0.966 | 0.966 | 0.966 | 0.966 |
| Ensemble Learning (Stacking) | 0.983 | 1.0 | 0.966 | 0.982 |
| Support Vector Machine(SVM) | 0.983 | 0.967 | 1.0 | 0.983 |

**6. Conclusion and Future Work**

This study demonstrated the potential of machine learning algorithms, particularly k-NN and ensemble methods, for predicting Parkinson's disease using voice and speech data. Future work will focus on exploring more advanced ensemble techniques, incorporating additional clinical data, and validating the model on a larger and more diverse dataset. The development of accurate and reliable prediction models can aid in the early diagnosis of Parkinson's disease, leading to improved treatment strategies and better healthcare outcomes.

**7. References**

* Sachin Shetty and Y.S. Rao “SVM based machine learning approach to identify Parkinson's disease using gait analysis” on 26-27 August 2016
* Speelman, J.D.; Muslimovic, D.; Schmand, B. Evolution of mild cognitive impairment in Parkinson disease. Neurology 2013, 81, 346–352.
* Goldman, J.G.; Aggarwal, N.T.; Schroeder, C.D. Mild cognitive impairment: An update in Parkinson’s disease and lessons learned from Alzheimer’s disease. Neurodegener. Dis. Manag. 2015, 5, 425–443
* Bott, N.T.; Johnson, E.T.; Schuff, N.; Galifianakis, N.; Subas, T.; Pollock, J.; Pressman, P.; Kramer, J.H.; Possin, K.L. Sensitive measures of executive dysfunction in non-demented Parkinson’s disease. Parkinsonism Relat. Disord. 2014, 20, 1430–1433.
* Caliskan Abdullah, Badem Hasan, Alper Basturk and Mehmet Emin Yuksel, "Diagnosis of the Parkinson Disease by using Deep Neural Network Classifier", IU-JEEE, vol. 17, no. 2, pp. 3311-3318, 2017.
* Grover Srishti, Bhartia Saloni, Akshama, Yadav Abhilasha and K. R. Seeja, "Predicting Severity Of Parkinsons Disease Using Deep Learning", International Conference on Computational Intelligence and Data Science (ICCIDS 2018). Procedia Computer Science, vol. 132, pp. 1788-1794, 2018