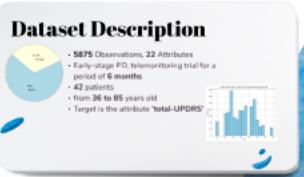


Parkinsons Telemonitoring

Parkinson Disease(PD) is a brain disorder that leads to shaking,stiffness, and difficulty with walking, balance and coordination.



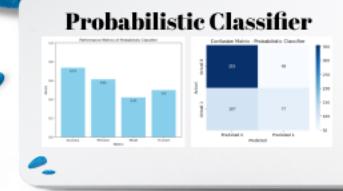
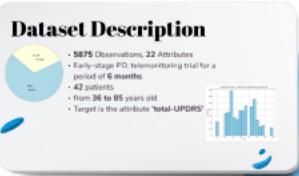
”
“Remembering the Parkinson's. Where your value belongs to the Class 0 or Class 1 based on the prediction.”
Thank you!

Parkinsons Telemonitoring

Parkinson Disease(PD) is a brain disorder that leads to shaking,stiffness, and difficulty with walking, balance and coordination.

Parkinsons Telemonitoring

Parkinson Disease(PD) is a brain disorder that leads to shaking,stiffness, and difficulty with walking, balance and coordination.



IS 669: Big Data & Information Systems

Group 4: KalyanKumar Papani

Ayush VimalKumar Patel

Dhruv Harishkumar Patel

Miliben Vineshkumar Patel

Sirivenni Pati



Problem Statement

- **Requires Patients Physically Appear in Clinic**

To track Unified Parkinson's Disease Rating Scale(UPDRS)

- **Routinely Checking is Diffcult**

Time consuming for immobility patients

Solution

- **Telemonitoring**

Predicts UPDRS scores with Voice measures rapidly
and remotely with clinically useful accuracy

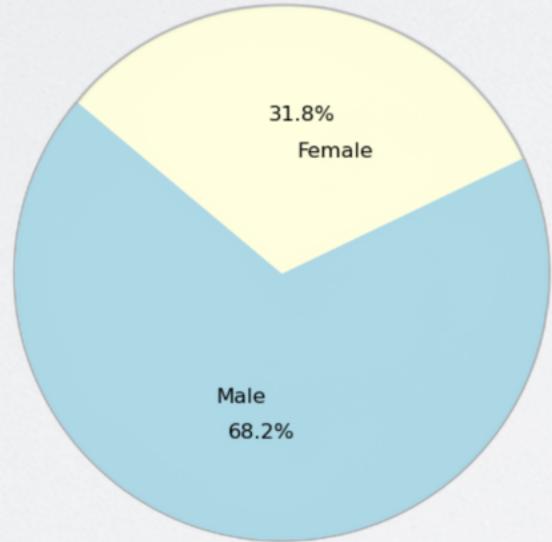
Research Questions

- Which regression model can best predict UPDRS based on voice measures?
- Which attributes has the highest feature importance in the prediction of UPDRS?

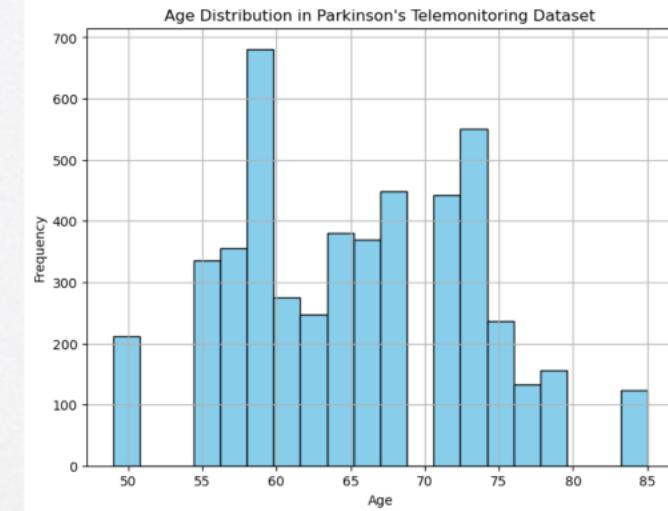
Aim

- The main aim of the data is to build a predictive model that can accurately predict total UPDRS scores ('**total_UPDRS**') from the 16 voice measures.

Dataset Description



- **5875** Observations, **22** Attributes
- Early-stage PD, telemonitoring trial for a period of **6 months**
- **42** patients
- from **36 to 85** years old
- Target is the attribute '**total-UPDRS**'



Data Cleaning

Missing Values

- Dataset has **0** missing values

Outliers

- **22** attributes contain outliers.
- They were removed by Interquartile Rule. Instances with values fall outside the lower and upper bound of boxplots are removed.

Descriptive Statistics

- **Descriptive Statistics for total_UPDRS:**

Minimum Value: 7.00

Maximum Value: 54.99

Mean Value: 29.02

Median Value: 27.58

Mode Value: 32.00

Range Value: 47.99

Variance Value: 114.50

Standard Deviation Value: 10.70

1st Quartiles Value: 21.37

2nd Quartiles Value: 27.58

3rd Quartiles Value: 36.40

Data Shaping

Tables:

- **Patients Table**

Columns: patient_id, subject_id, age, sex

- **Test Table**

Columns: test_id, patient_id, test_time, total_UPDRS

- **Results Table**

Columns: result_id, test_id, Jitter_percent, Jitter_Abs, Jitter_RAP, Jitter_PPQ5, Jitter_DDP, Shimmer, Shimmer_dB, Shimmer_APQ3, Shimmer_APQ5, Shimmer_APQ11, Shimmer_DDA

Relationships:

- **Patients to Tests:** One-to-many (Patients -> Tests)

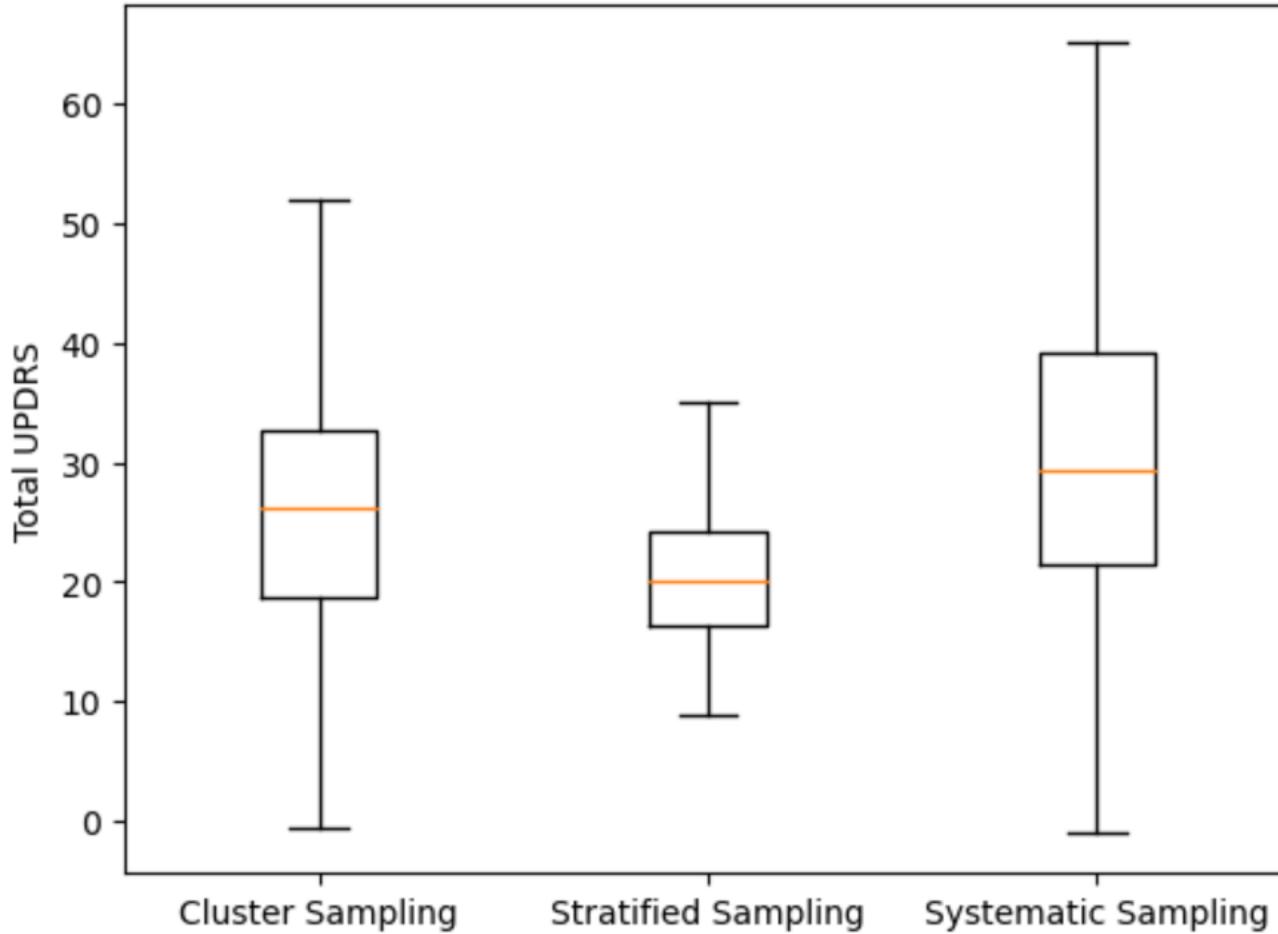
- **Tests to Results:** One-to-many (Tests -> Results)

Hierarchical Model:

- Patients -> Tests -> Results

Data Sampling

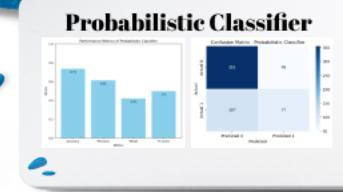
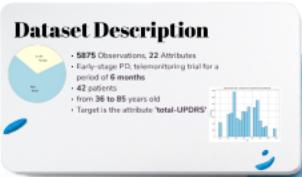
Comparison of Total UPDRS across Sampling Techniques



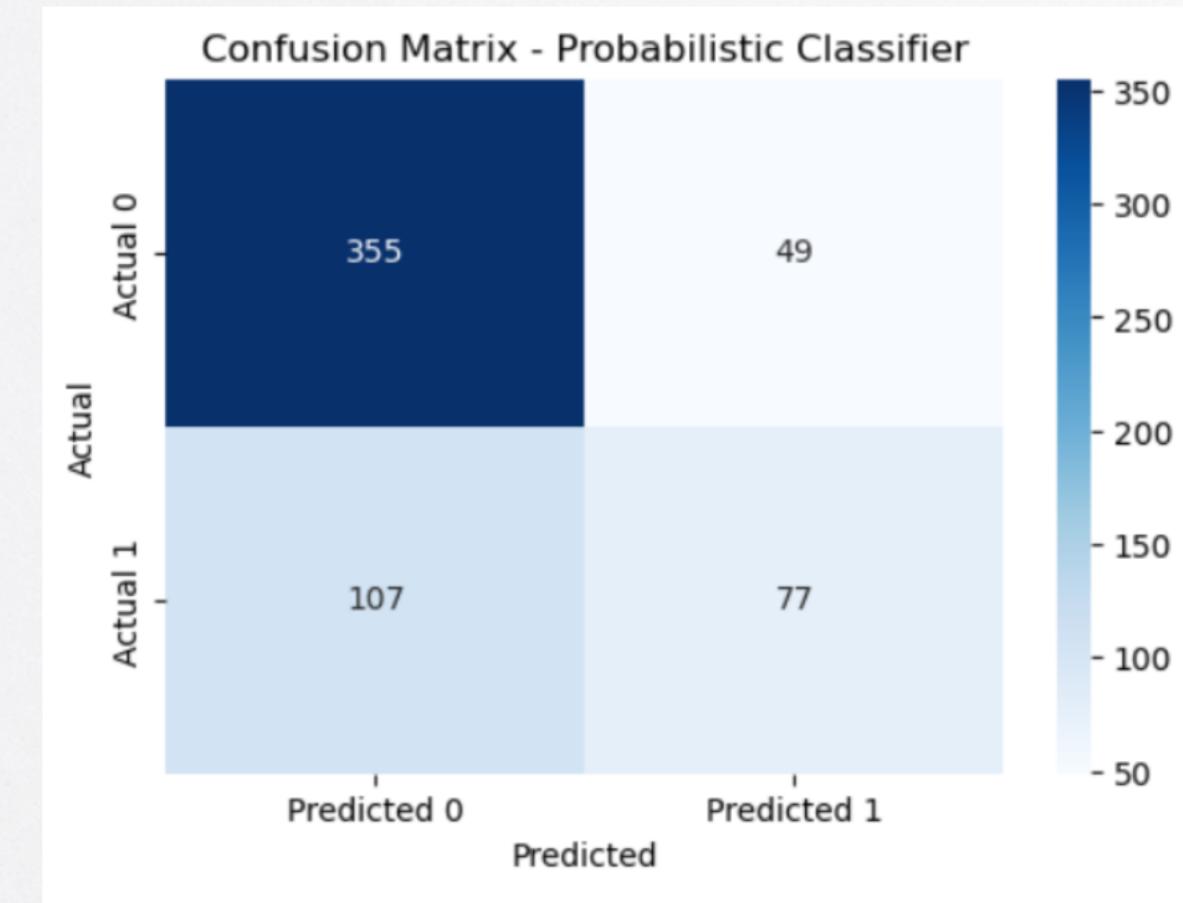
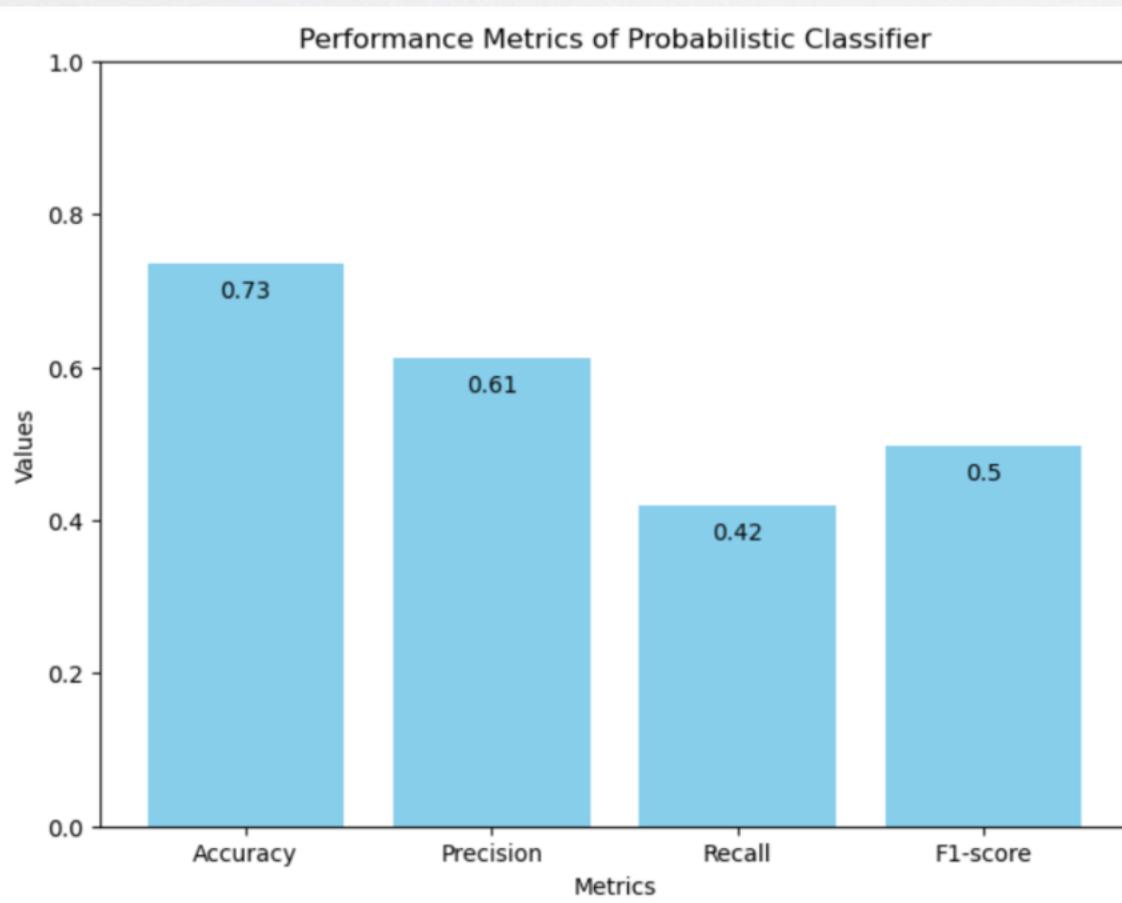
- If the target variable 'total UPDRS' needs to represent a wide range of patient experiences, a broader distribution (**Cluster or Systematic Sampling**) might be necessary.
- Conversely, if the goal is to minimize variability and focus on a more typical patient experience, **Stratified Sampling** would be appropriate.

Parkinsons Telemonitoring

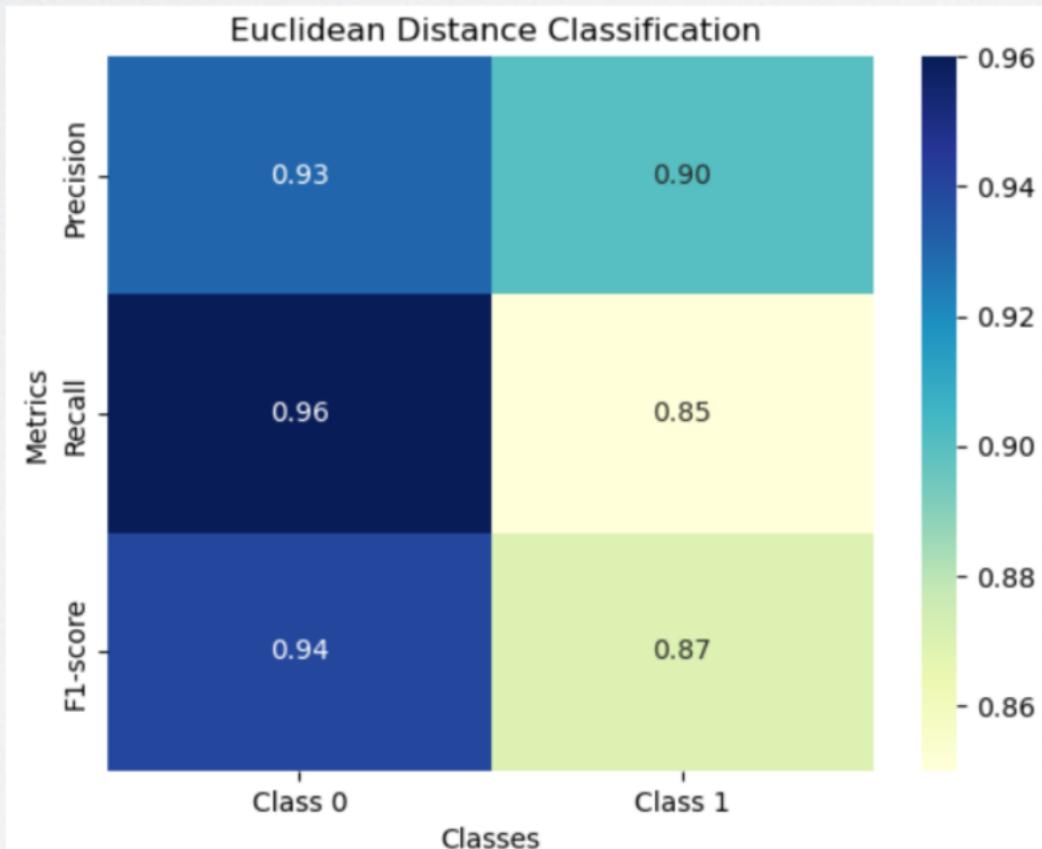
Parkinson Disease(PD) is a brain disorder that leads to shaking,stiffness, and difficulty with walking, balance and coordination.



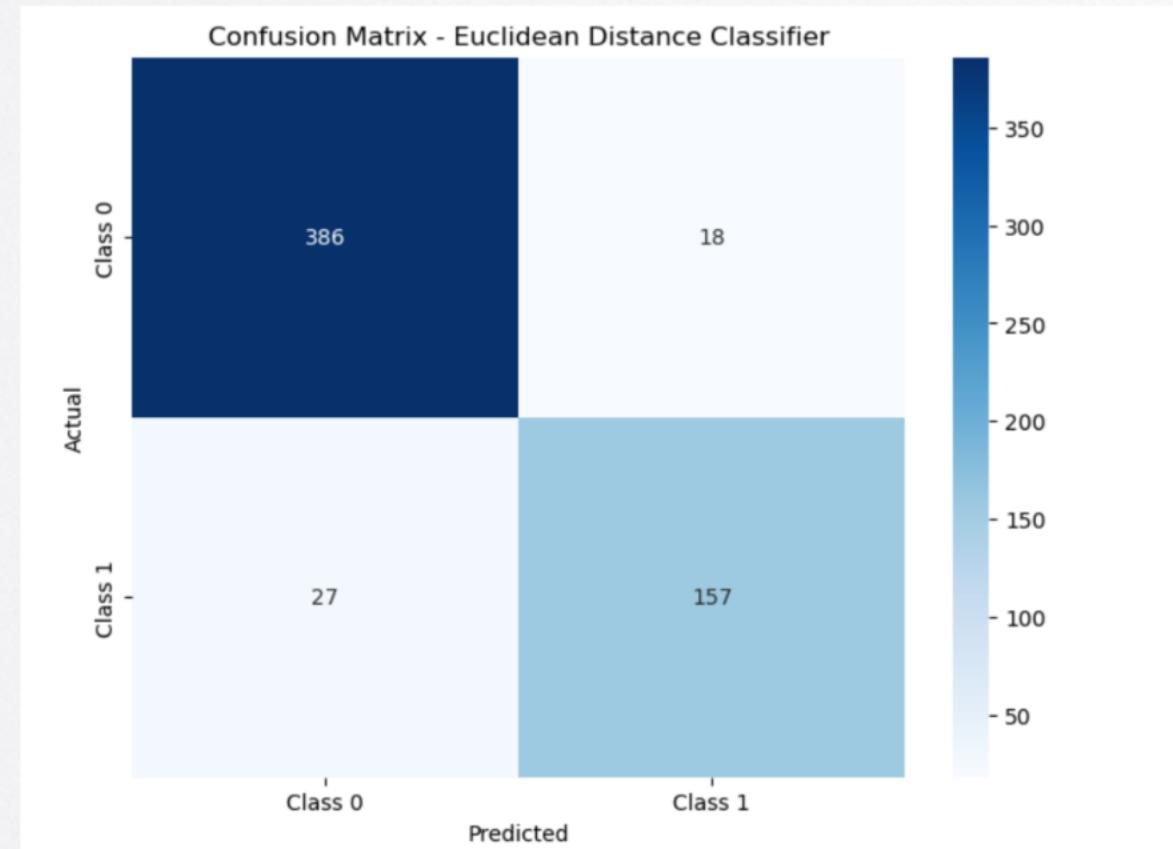
Probabilistic Classifier



Euclidean Distance Classification

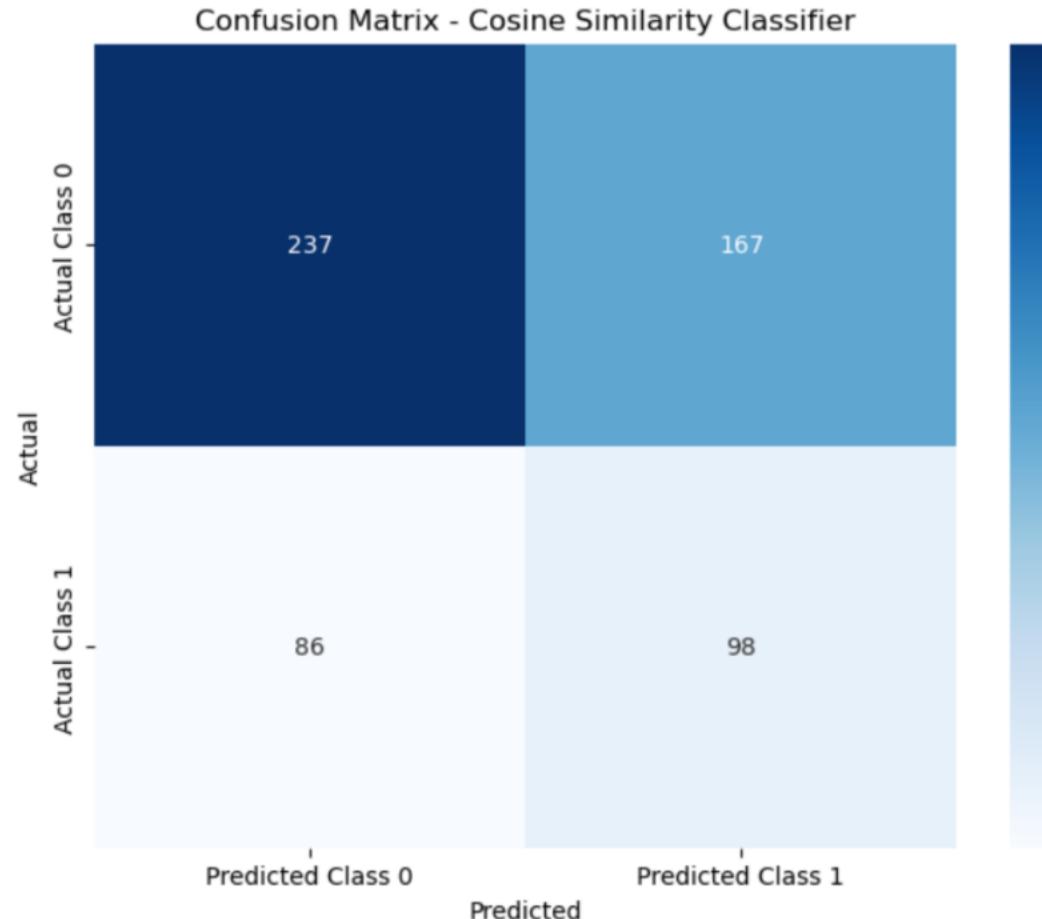


- For this particular model, Class 0 has slightly better performance across all three metrics compared to Class 1, indicating the model may be more effective at correctly identifying Class 0 instances.



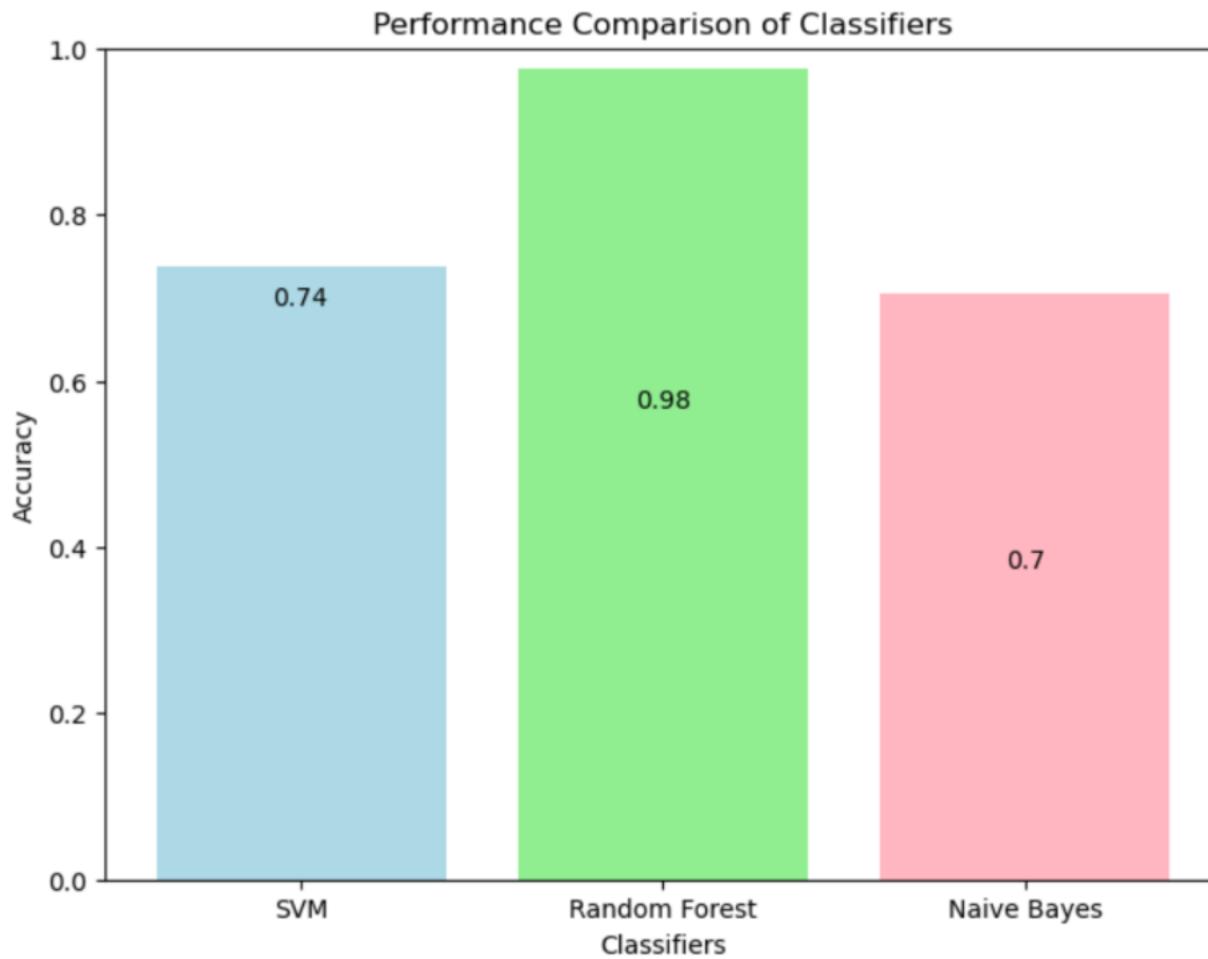
- The color intensity is proportional to the number of instances. Darker colors indicate a higher number of instances. Here, the darkest shade indicates the highest number of true positives for Class 0 (386).

Cosine Similarity Classification



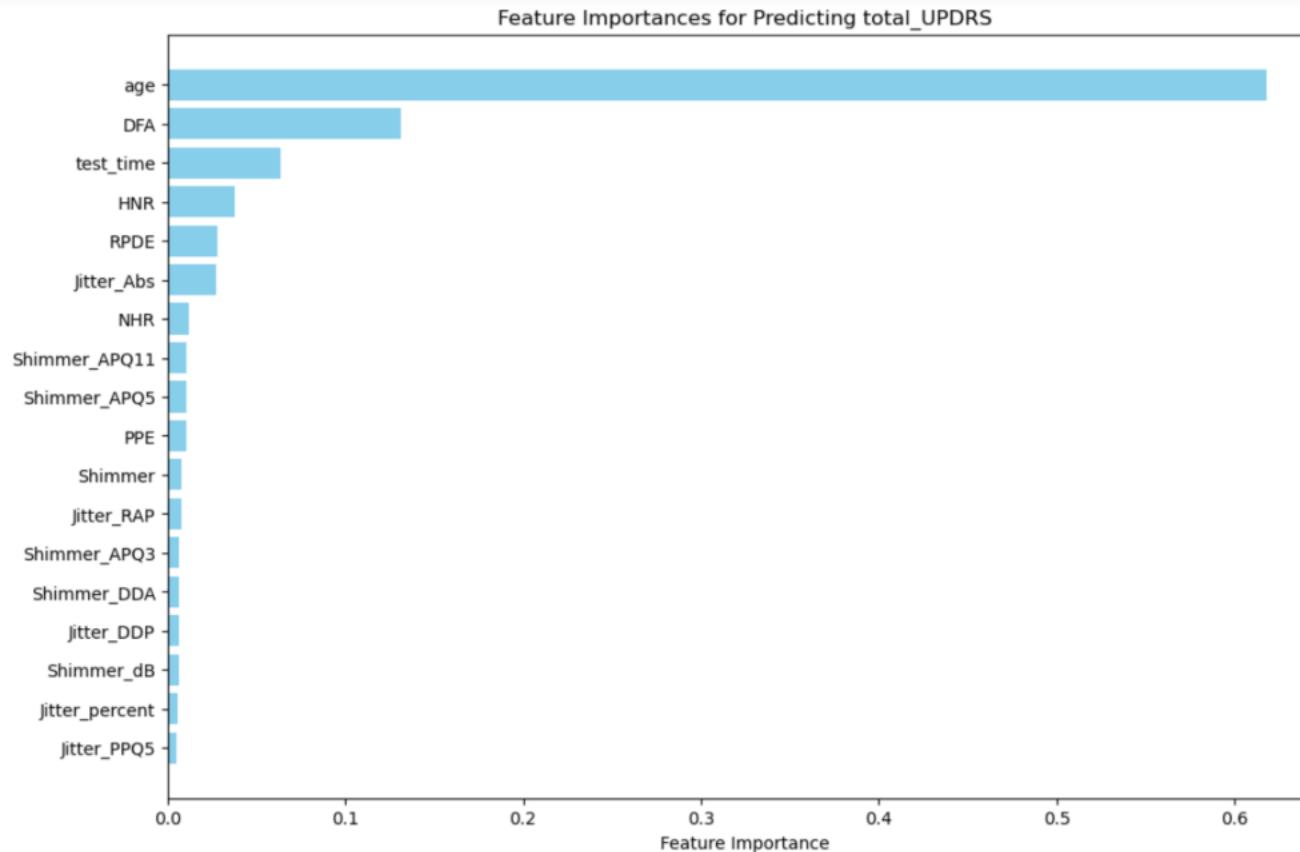
- **Top-Left Cell (Actual Class 0, Predicted Class 0):** Represents the true positives for Class 0 (237 instances correctly identified as Class 0).
- **Top-Right Cell (Actual Class 0, Predicted Class 1):** Represents false negatives for Class 0 (167 instances incorrectly identified as Class 1).
- **Bottom-Left Cell (Actual Class 1, Predicted Class 0):** Represents false positives for Class 1 (86 instances incorrectly identified as Class 0).
- **Bottom-Right Cell (Actual Class 1, Predicted Class 1):** Represents true positives for Class 1 (98 instances correctly identified as Class 1).

Classifier Performance



- **Random Forest:**
Achieves the highest accuracy of 98.0% among the three classifiers used.
- **Support Vector Machine (SVM):**
Provides competitive accuracy at 74.8%.
- **Naive Bayes:**
Shows a moderate accuracy of 70.4%.

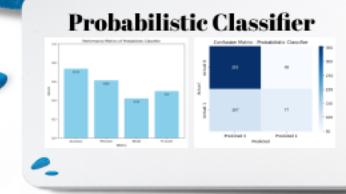
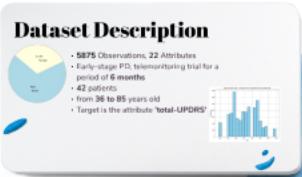
Highest Importance Feature in the prediction of UPDRS



- **Age** holds the highest feature importance (**61.8%**) in predicting UPDRS, signifying its significant role in the model's predictive capacity.
- **DFA (Detrended Fluctuation Analysis)** follows with substantial importance at **13.1%**, contributing significantly to UPDRS predictions.

Parkinsons Telemonitoring

Parkinson Disease(PD) is a brain disorder that leads to shaking, stiffness, and difficulty with walking, balance and coordination.



”

**"Telemonitoring for Parkinson's: Where your voice becomes the
doctor's favorite podcast!"**

"Thank you"

i

Parkinsons Telemonitoring

Parkinson Disease(PD) is a brain disorder that leads to shaking,stiffness, and difficulty with walking, balance and coordination.

