|  |  |  |  |
| --- | --- | --- | --- |
| ​​Machine Failure Prediction​ | | | |
| **​​SUBJECT​** | **CODE** | **TEAM.NO** | **​​DATE​** |
| Machine Learning | 19CSE305 | 7 | 10/12/2023 |

**​​**

**OVERVIEW​**

This project explores the use of machine learning algorithms to predict machine failure.

|  |  |  |
| --- | --- | --- |
| **ROLL NO.** | **NAME** | **CONTRIBUTION** |
| CB.EN.U4CSE21027 | Karan Singhania | Under Sampling, Oversampling, PCA (Dimensionality Reduction), Random Forest (Ensemble),  K-Means,  Agglomerative Clustering |
| CB.EN.U4CSE21037 | Muhilan H | Under Sampling, Oversampling, PCA (Dimensionality Reduction), Naïve Bayes with Bagging (Probabilistic Model),  Cross Validation and Hyperparameter Tuning |
| CB.EN.U4CSE21041 | Nithyleshwaran PS | Under Sampling, Oversampling, PCA (Dimensionality Reduction), Catboost(Ensemble),  Cross Validation and Hyperparameter Tuning |
| CB.EN.U4CSE21056 | Shri Venkatakrishnan | Under Sampling, Oversampling, PCA (Dimensionality Reduction), Gradient Boosting (Ensemble),  Cross Validation and Hyperparameter Tuning |
| CB.EN.U4CSE21061 | Syed Fareed | Under Sampling, Oversampling, PCA (Dimensionality Reduction), AdaBoost (Ensemble),  Cross Validation and Hyperparameter Tuning |

**Implementation:**

|  |  |
| --- | --- |
| **APPROACH** | **TOPICS COVERED** |
| ***PROBABILISTIC MODEL*** | * Naive Bayes       -Gaussian      -Multinomial |
| ***UNSUPERVISED MODEL*** | **Clustering:**   * K-Means clustering. * Hierarchical clustering. |
| ***ENSEMBLE MODEL*** | 1. **Bagging:**  * Base Model   -Naïve Bayes  -Random Forest   1. **Boosting:**  * AdaBoost. * Gradient Boosting * Catboost |
| ***DIMENSIONALITY REDUCTION*** | **Linear Techniques:**   * Principal Component Analysis (PCA). |

**ANALYSIS:**

Note: Each member in the team has done the following Analysis for their parts.

|  |  |
| --- | --- |
| **CONCEPT** | **TOPICS COVERED** |
| ***VISUALISATION*** | -Created visualizations (e.g., heatmaps, scatter plots) to illustrate the relationships between properties of the machine like Torque, Rotational Speed and with Target Variable |
| ***Cross VALIDATION*** | Employed K-Fold Cross Validation to get a more generalized value of performance metrics |
| ***HYPERPARAMETER TUNING*** | - Employed grid search to vary different parameters of different models and to obtain the best performing model |
| ***ROC CURVES*** | Plotted ROC-AUC curves to judge the performance of each model |
| ***PERFORMANCE METRICS*** | - Calculated and compared metrics such as accuracy, precision, recall, and F1 score to evaluate the overall performance of each model. |