To

IITD-AIA Foundation of Smart Manufacturing

Date:06-08-2023

Subject: ***Weekly Progress Report for Week-9.***

Dear Sir,

Following is the required progress report of this week dated from 31-07-2023 to 06-08-2023.

Weekly Progress:

31 July & 1 August :

Topics covered:

- The dataset is loaded and preprocessed.
- The k-fold cross-validation is performed using k-fold from scikit-learn with n-splits set to the desired number of folds.
- The average mean squared error and average R-squared across all folds are then calculated to provide a more reliable estimate of the model's performance.
- This approach helps to assess how well the model generalizes to unseen data and reduces the dependency on a specific train-test split.
- The model is trained and evaluated on each fold, and the mean squared error and R-squared scores are computed for each fold.
- During the cross-validation loop, we train the model on the training data, make predictions on the test data, and calculate the mean squared error for each fold.
- Finally, we compute the average mean squared error across all folds to get an estimate of the model's performance.
- Model validation is a crucial step in the machine learning workflow that involves assessing the performance and generalization ability of a trained model on unseen data.
- The goal of model validation is to determine how well the model can make accurate predictions on new data and avoid overfitting, where the model performs well on the training data but poorly on unseen data.
- There are several techniques for model validation, and the choice of method depends on the dataset size, available data, and the specific problem being solved.
- During model validation, several performance metrics are used to assess the model's accuracy and generalization ability, such as:
    - Mean Squared Error (MSE)
    - Root Mean Squared Error (RMSE)
    - Mean Absolute Error (MAE)
    - R-squared (R2)
    - Accuracy, Precision, Recall, F1-score

02 August & 03 August:

Topics covered:

- Stratified K-Fold Cross-Validation is a variant of the traditional K-Fold Cross-Validation technique that is used primarily for classification problems with imbalanced class distributions.
- It ensures that each fold's test set contains a proportional distribution of class labels, which helps in obtaining more representative and reliable performance estimates for the model.
- The original dataset is divided into k equally-sized folds while maintaining the proportion of class labels in each fold as close as possible to the original dataset.
- During each iteration of the cross-validation process, one fold is used as the test set, and the remaining (k-1) folds are used as the training set.
- The model is trained on the training set and then evaluated on the test set.
- The performance metric (e.g., accuracy, precision, recall, F1-score) is calculated for each fold. The final performance metric is often the average of the performance metrics from all k iterations.
- Leave-One-Out Cross-Validation (LOOCV) is a special case of cross-validation, where the number of folds (k) is equal to the number of data points in the dataset.
- In LOOCV, each data point is used as the test set once, while the rest of the data points are used for training.
- This means that for a dataset with N data points, the model is trained and evaluated N times, with each iteration leaving out one data point for testing.
- The model is trained on the training set, and then its performance is evaluated on the single data point held out for testing.
- The performance metric (e.g., mean squared error, accuracy, etc.) is calculated for each iteration of LOOCV.
- It provides an unbiased estimate of the model's performance since each data point is used as the test set exactly once.
- It is useful when the dataset is small, and we want to make the most of the available data for evaluation.

04 August:

Topics covered:

- I have started documentation of the project.
- I have written abstract which includes

1. The machining industry plays a pivotal role in manufacturing various products.
2. One critical aspect of machining is tool wear and surface roughness, which directly impacts product quality and production costs.

3. Accurate prediction of tool wear and surface roughness can help optimize cutting parameters, improve product quality, and enhance overall productivity., and other points.

- I have written table of content and introduction of the project.
- how we develop a machine learning model that can predict tool wear and surface roughness of workpieces produced by a lathe machine. By analyzing various process parameters, tool characteristics, and historical data, the model will estimate the tool wear progression and surface roughness for different operating machines.

05 August & 06 August:

Topics covered:

- I have written Methodologies used which includes-
- The methodology involves data preprocessing, feature engineering, model training, hyperparameter tuning, and model evaluation.
- An Inclusive Data Analysis was conducted on the dataset, descriptive statistics, correlation analysis and Data Visualization techniques were employed to explore the relationships between the variables.
- Multiple machine learning algorithms, such as Support Vector Machines (SVM), Random Forest, Neural Networks, and Gradient Boosting, are employed and compared to identify the most accurate and reliable model.
- I have worked on final report and Presentation of the project.
- I have included literature review, results, conclusion, references etc.
- Several existing approaches and methodologies have been explored to address these prediction challenges.
- The results of Data Preprocessing, Data Analysis, Data visualization, Feature Engineering, Performance Evaluation Metrics, Different Model comparisons.
- Results of Evaluation metrices performed on the dataset and different model comparision.
- Future Directions and Potential Applications- Include Real-Time Monitoring and Control Predicting, tool wear and surface roughness in real-time and adjusting cutting parameters automatically. Industry 4.0 Implementation Integrating predictive maintenance with industry 4.0 technologies for better manufacturing efficiency. Online Tool and Surface Roughness Databases Creating reliable databases for tool wear and surface roughness data for better research and development.