**Project Development Phase**

**Model Performance Test**

| Date | 10 November 2022 |
| --- | --- |
| Team ID | PNT2022TMIDxxxxxx |
| Project Name | Project - xxx |
| Maximum Marks | 10 Marks |

**Model Performance Testing:**

Project team shall fill the following information in the model performance testing template.

| **S.No.** | **Parameter** | **Values** | **Screenshot** |
| --- | --- | --- | --- |
|  | Model Summary | Objective: The primary objective of this project is to develop a predictive model for accurate sales forecasting across various product categories in different Walmart stores. This will facilitate optimized inventory management, efficient promotion planning, and overall revenue enhancement.  Dataset: The dataset provided by Walmart encompasses historical sales data, encompassing product details, store locations, promotional events, and relevant features.  Data Preprocessing:  Data Cleaning: We commenced by addressing missing values, outliers, and data inconsistencies for a cleaner dataset.  Feature Engineering: New features were created, including seasonality, holiday indicators, and trend components, to augment the predictive capacity of the model.  Encoding: Categorical variables were encoded into numerical representations, employing one-hot encoding and label encoding techniques.  Scaling: Numerical features were standardized to ensure uniform scaling.  Exploratory Data Analysis (EDA): In-depth analysis of the dataset provided valuable insights into sales trends, correlations, and patterns, steering our feature selection and engineering process.  Model Selection: We experimented with several regression models:  Linear Regression: Utilized as a baseline model for establishing initial performance benchmarks.  Decision Trees and Random Forests: Employed to capture non-linear relationships between features and sales, potentially revealing complex patterns.  Gradient Boosting (e.g., XGBoost, LightGBM): Ensemble methods chosen for their ability to often yield high predictive accuracy.  Model Training and Validation:  Train-Test Split: The dataset was partitioned into training and testing sets to facilitate model evaluation.  Cross-Validation: Implemented k-fold cross-validation to ascertain the model's generalization capability.  Hyperparameter Tuning: Hyperparameters were fine-tuned using grid search and random search methods, optimizing model performance.  Evaluation Metric: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R²) were employed to quantitatively measure the model's accuracy in predicting sales.  Final Model Deployment: The trained model can be deployed using frameworks like Flask. |  |
|  | Accuracy | **Regression Model:** R2 score : 0.9482143725653065 MSE score : 26513619.680080418 RMSE: 5149.137760837286- |  |