# Model Development Phase Template

| Date | 5 May 2024 |
|---|---|
| Team ID | 737906 |
| Project Title | Walmart Sales Analysis for Retail Industry with Machine Learning |
| Maximum Marks | 6 Marks |

## Model Selection Report

The Model Selection Report evaluates various models for predicting Walmart store sales, including Decision Tree, Random Forest, ARIMA, and XgBoost. Each model's performance was assessed using metrics such as MAE, RMSE, accuracy, and F1 score. While Decision Tree provided moderate accuracy, ensemble methods like Random Forest showed higher performance. ARIMA was suitable for time series but less accurate than advanced models. The XgBoost emerged as top performers, offering the highest accuracy and F1 scores, making them the most suitable choices for accurate sales forecasting.

## Model Selection Report:

| Model | Description | Hyperparameters | Performance Metric (e.g., Accuracy, F1 Score) |
|---|---|---|---|
| Random Forest | Combines multiple decision trees to enhance accuracy and reduce overfitting, showing high performance and accuracy. | Max_depth: 30, min_samples: 5 | Performance: High<br><br>Metric: RMSE<br><br>Training Accuracy: 94%<br><br>Testing Accuracy: 96% |

| Decision Tree | Simplistic and interpretable, but prone to overfitting. Performs better than linear regression but is outperformed by ensemble methods. | Max_depth: 30, min_samples: 5 | Performance: Moderate<br><br>Metric: RMSE<br><br>Training Accuracy: 100%<br><br>Testing Accuracy: 94% |
|---|---|---|---|
| XgBoost | Known for its speed and efficiency, providing high accuracy and performance, making it one of the top contenders. | n_estimators: 500<br><br>Max_depth:4<br><br>Learning_rate: 0.5 | Performance: High<br><br>Metric: RMSE<br><br>Training Accuracy: 94%<br><br>Testing Accuracy: 94% |
| ARIMA | Suitable for time series data, offering moderate performance, but can be outpaced by more advanced models in terms of accuracy and adaptability. | Year<br><br>month<br><br>date | Performance: High<br><br>Metric: MAE |