Model Development Phase Template

Date	11th July 2024
Team ID	SWTID1720435231
Project Title	Walmart Sales Analysis For Retail Industry With Machine Learning
Maximum Marks	6 Marks

Model Selection Report

In the Model Selection Report, the various models that have been tested will be given a brief, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
Linear Regression	Models sales based on linear relationships with predictors like CPI and unemployment, offering simplicity and interpretability but limited in capturing complex interactions.	-	Accuracy score = 89%
Random Forest	Constructs multiple decision trees to predict sales, effectively handling complex feature interactions like store, department, and economic	-	Accuracy score = 96%

	indicators such as CPI and unemployment.		
Decision Tree	Divides data into subsets based on key features to predict sales, effective in capturing interactions between variables such as store, department, and seasonal factors, but may overfit without ensemble methods.	-	Accuracy score = 94%
XGBoost	Optimizes weak learners sequentially to predict sales with high accuracy, especially beneficial for capturing non-linear relationships among diverse features like temperature, markdowns, and sales history.	-	Accuracy score = 94%
ARIMA	Forecasts sales based on historical patterns and autocorrelation in time series data, suitable when predicting sales trends over time without explicit consideration of external factors.	-	Accuracy score = 97%