Model Optimization and Tuning Phase Template

Date	5 May 2024
Team ID	737906
Project Title	Walmart Sales Analysis for Retail Industry with Machine Learning
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values

Random Forest	`n_estimators,`
	`max_depth,`
	`min_samples_split,`
	`min_samples_leaf`
		`n_estimators= 150,`
		`max_depth=30,`
		`min_samples_split=5,`
		`min_samples_leaf=5`

Decision Tree	`_estimators,`
	`max_depth,`
	`min_samples_split,`
	`min_samples_leaf`
		`n_estimators= 150,`
		`max_depth=30,`
		`min_samples_split=5,`
		`min_samples_leaf=5`

XgBoost	`learning_rate,`
	`nthread,`
	`n_estimators,`
	`max_depth,`
		`learning_rate = 0.5,`
		`nthread=4,`
		`n_estimators= 500,`
		`max_depth = 4,`

	alpha, random_state	alpha = 10, random_state=0
ARIMA	Year, Month, Date	Year, Month, Date

Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metric
Random Forest	85%	88%
Decision Tree	72%	78%
XgBoost	87%	92%
ARIMA	78%	82%

Final Model Selection Justification (2 Marks):

Final Model	Reasoning

XgBoost	XgBoost was chosen as the final optimized model due to its superior performance in terms of accuracy and generalization ability. It consistently outperformed other models after hyperparameter tuning and optimization, demonstrating robustness and efficiency. XgBoost's ensemble method effectively handles non-linear relationships in the data, making it suitable for complex datasets like sales prediction at Walmart stores. Additionally, its computational efficiency allows for faster training and prediction times, making it a practical choice for real-world applications. Overall, XgBoost offers the best balance of accuracy, generalization, and efficiency, making it the optimal choice for sales forecasting at Walmart.