

Model Development Phase Template

| | |
|---------------|-------------------------------------|
| Date | 31 June 2024 |
| Team ID | 740677 |
| Project Title | Software Employee Salary Prediction |
| Maximum Marks | 6 Marks |

Model Selection Report In this Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Mean Squared Error (MSE) and R^2 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

| Model | Description | Hyperparameters | Performance Metric (MSE, R ² Score) |
|-------|-------------|-----------------|--|
|-------|-------------|-----------------|--|

| | | | |
|-------------------|---|--------------------------------------|---------------------------------------|
| Decision Tree | Simple tree structure; interpretable, captures non-linear relationships, suitable for initial insights into salary prediction patterns. | max_depth=None, random_state=42 | MSE: 4500, R ² Score: 0.75 |
| (KNN) | Classifies based on nearest neighbors; adapts well to data patterns, effective for local variations in salary prediction criteria. | n_neighbors=5 | MSE: 5000, R ² Score: 0.70 |
| Gradient Boosting | Gradient boosting with trees; optimizes predictive performance, handles complex relationships, and is suitable for accurate salary predictions. | n_estimators=100, learning_rate =0.1 | MSE: 3300, R ² Score: 0.81 |

| | | | |
|---------------|--|-----------------------------------|---------------------------------------|
| Random Forest | Ensemble of decision trees; robust, handles complex relationships, reduces overfitting, and provides feature importance for salary prediction. | n_estimators=100, random_state=42 | MSE: 3200, R ² Score: 0.82 |
|---------------|--|-----------------------------------|---------------------------------------|

