**Model Optimization and Tuning Phase Template**

| Date | July 2024 |
| --- | --- |
| Team ID | 740053 |
| Project Title | Estimating the stock keeping units using Machine Learning |
| Maximum Marks | 10 Marks |

**Model Optimization and Tuning Phase**

The Model Optimization and Tuning Phase involves refining neural network 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 (8 Marks):

| **Model** | **Tuned Hyperparameters** |
| --- | --- |
| Linear Regression | #importing the library for grid search  from sklearn.model\_selection import GridSearchCV  The ‘lr\_param\_grid’ specifies different values for regularization strength (C), solvers (solver), and penalty types (penalty). GridSearchCV (lr\_cv) is employed with 5-fold cross-validation (cv=5), evaluating model performance based on accuracy (scoring="r2 score").  C:\Users\home\AppData\Local\Packages\Microsoft.Windows.Photos_8wekyb3d8bbwe\TempState\ShareServiceTempFolder\screenshot16.jpeg |
| Random Forest | The parameter grid (make\_regression) for hyperparameter tuning. It specifies different values for the number of trees (n\_estimators), splitting criterion (criterion), maximum depth of trees (max\_depth), and maximum number of features considered for splitting (max\_features). GridSearchCV (rfc\_cv) is employed with 3-fold cross-validation (cv=3), evaluating model performance based on accuracy (scoring="r2 score").  C:\Users\home\AppData\Local\Packages\Microsoft.Windows.Photos_8wekyb3d8bbwe\TempState\ShareServiceTempFolder\screenshot17.jpeg |
| Decision Tree | The parameters (params) define a randomized search for hyperparameter tuning of the Decision Tree Regressor (DecisionTreeRegressor), including max\_depth, min\_samples\_leaf, min\_samples\_split and max\_features . RandomizedSearchCV is used to evaluating model performance based on r2 score(scoring="r2 score")  C:\Users\home\AppData\Local\Packages\Microsoft.Windows.Photos_8wekyb3d8bbwe\TempState\ShareServiceTempFolder\screenshot18.jpeg |

### 

### 

### 

### 

### Final Model Selection Justification (2 Marks):

| **Final Model** | **Reasoning** |
| --- | --- |
| **Random Forest** | Random Forest model is chosen for its robustness in handling complex datasets and its ability to mitigate overfitting while providing high predictive r2 score.  C:\Users\home\AppData\Local\Packages\Microsoft.Windows.Photos_8wekyb3d8bbwe\TempState\ShareServiceTempFolder\screenshot12.jpeg  Above all the models Random Forest model have the highest r2 score among all the models.  A higher r2 score is generally considered better as it indicates a more accurate and reliable model. |