# Model Development Phase Template

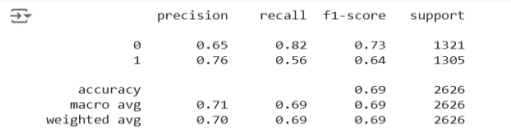| Date | July 2024 |
|---|---|
| Team ID | 739663 |
| Project Title | Estimating the Stock keeping units using Machine Learning |
| Maximum Marks | 10 Marks |

**Initial Model Training Code, Model Validation and Evaluation Report**

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include a summary and training and validation performance metrics for multiple models, presented through respective screenshots.

**Initial Model Training Code (5 marks):**

Paste the screenshot of the model training code

**Model Validation and Evaluation Report (5 marks):**

| Model | Summary | Training and Validation Performance Metrics |
|---|---|---|
| Model 1 | Linear regression model typically include accuracy, precision, recall, R2 score to evaluate its predictive performance and generalization capability. |  |

| | | |
|---|---|---|
| Model 2 | Random forest regression model often encompass accuracy, precision, recall, R2 score to measure its prediction quality and robustness. | **RANDOM FOREST CLASSIFIER**<br><br>```python<br>[45] #random forest classifier<br>from sklearn.ensemble import RandomForestClassifier<br>rfc=RandomForestClassifier()<br>rfc.fit(x_train,y_train)<br>pred=rfc.predict(x_test)<br>from sklearn.metrics import classification_report<br>print(classification_report(y_test,pred))<br>```<br><br>```<br>              precision    recall  f1-score   support<br><br>           0       0.69      0.87      0.77      1321<br>           1       0.82      0.61      0.70      1305<br><br>    accuracy                           0.74      2626<br>   macro avg       0.76      0.74      0.73      2626<br>weighted avg       0.76      0.74      0.73      2626<br>``` |
| Model 3 | Decision tree regression model commonly include accuracy, precision, recall, R2 score which help assess the model's prediction accuracy and generalizability. | ```python<br>#decision tree classifier<br>from sklearn.tree import DecisionTreeClassifier<br>dec=DecisionTreeClassifier()<br>dec.fit(x_train,y_train)<br>predi=dec.predict(x_test)<br>from sklearn.metrics import classification_report<br>print(classification_report(y_test,predi))<br>```<br><br>```<br>              precision    recall  f1-score   support<br><br>           0       0.71      0.69      0.70      1321<br>           1       0.69      0.71      0.70      1305<br><br>    accuracy                           0.70      2626<br>   macro avg       0.70      0.70      0.70      2626<br>weighted avg       0.70      0.70      0.70      2626<br>``` |
| Model 4 | K-nearest neighbors classifier model typically include accuracy, precision, recall, R2 score to evaluate its prediction performance and generalization ability | **K-NEAREST NEIGHBORS**<br><br>```python<br>[48] #knn<br>from sklearn.neighbors import KNeighborsClassifier<br>knn=KNeighborsClassifier()<br>knn.fit(x_train,y_train)<br>p=knn.predict(x_test)<br>from sklearn.metrics import classification_report<br>print(classification_report(y_test,p))<br>```<br><br>```<br>              precision    recall  f1-score   support<br><br>           0       0.69      0.79      0.74      1321<br>           1       0.75      0.65      0.69      1305<br><br>    accuracy                           0.72      2626<br>   macro avg       0.72      0.72      0.72      2626<br>weighted avg       0.72      0.72      0.72      2626<br>``` |
| Model 5 | XGB classifier model typically include accuracy, precision, recall, R2 score used to evaluate the model's predictive performance and ability to generalize | **XGBOOST CLASSIFIER**<br><br>```python<br>[49] #XGBOOST CLASSIFICATION<br>from xgboost import XGBClassifier<br>xg=XGBClassifier()<br>xg.fit(x_train,y_train)<br>p=xg.predict(x_test)<br>from sklearn.metrics import classification_report<br>print(classification_report(y_test,p))<br>```<br><br>```<br>              precision    recall  f1-score   support<br><br>           0       0.71      0.80      0.75      1321<br>           1       0.77      0.66      0.71      1305<br><br>    accuracy                           0.73      2626<br>   macro avg       0.74      0.73      0.73      2626<br>weighted avg       0.74      0.73      0.73      2626<br>``` |

| Model 6 | Ridge classifier model typically include accuracy, precision, recall, R2 score, and mean squared error to evaluate its prediction performance and generalization. | |
|---------|------------------------------------------------------------------------------------------------------------------------------------------------------------|---|

**RIDGE CLASSIFIER**

```
[50] #RIDGE CLASSIFIER
     from sklearn.linear_model import RidgeClassifier
     rg=RidgeClassifier()
     rg.fit(x_train,y_train)
     p=rg.predict(x_test)
     from sklearn.metrics import classification_report
     print(classification_report(y_test,p))
```

```
                 precision    recall  f1-score   support

             0       0.65      0.80      0.72      1321
             1       0.74      0.56      0.64      1305

      accuracy                           0.68      2626
     macro avg       0.69      0.68      0.68      2626
  weighted avg       0.69      0.68      0.68      2626
```