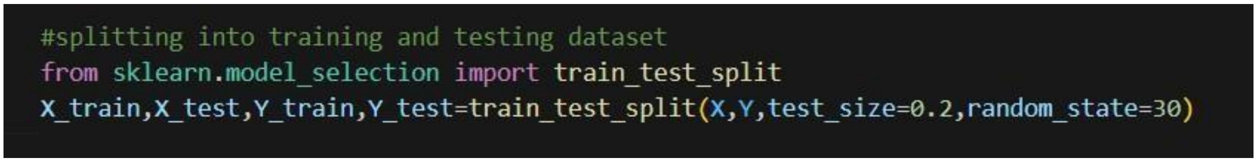
Model Development Phase Template

Date	15 July 2024
Team ID	740682
Project Title	Polycystic Ovary Syndrome Classification Using Machine Learning
Maximum Marks	4 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 classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

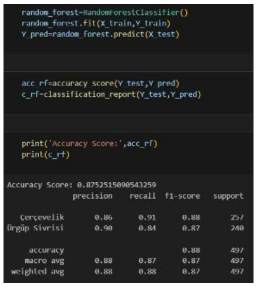
Initial Model Training Code:



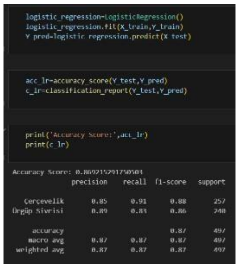
```
#splitting into training and testing dataset
from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=30)
```

Model Validation and Evaluation Report:

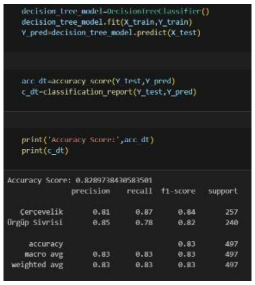
Model	Classification Report& Accuracy	Accuracy

Random forest classifier		0.875

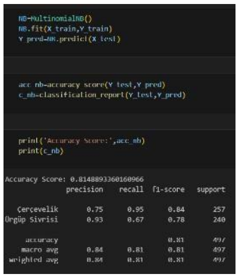
```
random_forest-RandomForestClassifier()
random_forest.fit(X_train,Y_train)
Y_pred=random_forest.predict(X_test)

acc rf=accuracy score(Y test,Y pred)
c_rf-classification_report(Y_test,Y_pred)

print('Accuracy Score:',acc_rf)
print(c_rf)

Accuracy Score: 0.8752515090543259
               precision    recall  f1-score   support

  Çerçevelik       0.86      0.91      0.88       257
Ürgüp Sivrisi       0.90      0.84      0.87       240

    accuracy                           0.88       497
   macro avg       0.88      0.87      0.87       497
weighted avg       0.88      0.88      0.87       497
```

Logistic regression		0.86

```
logistic_regression-LogisticRegression()
logistic_regression.fit(X_train,Y_train)
Y pred-logistic regression.predict(X test)

acc_lr-accuracy_score(Y_test,Y_pred)
c_lr=classification_report(Y_test,Y_pred)

print('Accuracy Score:',acc_lr)
print(c lr)

Accuracy Score: 0.869215291750503
               precision    recall  f1-score   support

  Çerçevelik       0.85      0.91      0.88       257
Ürgüp Sivrisi       0.89      0.83      0.86       240

    accuracy                           0.87       497
   macro avg       0.87      0.87      0.87       497
weighted avg       0.87      0.87      0.87       497
```

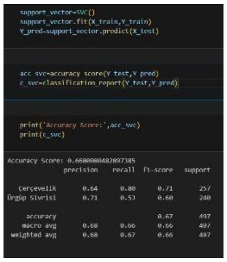
Decision Tree classifier		0.82

```
decision_tree_model-DecisionTreeClassifier()
decision_tree_model.fit(X_train,Y_train)
Y_pred=decision_tree_model.predict(X_test)

acc dt=accuracy score(Y test,Y pred)
c_dt-classification_report(Y_test,Y_pred)

print('Accuracy Score:',acc dt)
print(c_dt)

Accuracy Score: 0.8289738430583501
               precision    recall  f1-score   support

  Çerçevelik       0.81      0.87      0.84       257
Ürgüp Sivrisi       0.85      0.78      0.82       240

    accuracy                           0.83       497
   macro avg       0.83      0.83      0.83       497
weighted avg       0.83      0.83      0.83       497
```

Naïve Bayes		0.81

```
NB-MultinomialNB()
NB.fit(X_train,Y_train)
Y pred-NB.predict(X test)

acc nb-accuracy score(Y test,Y pred)
c_nb-classification_report(Y_test,Y_pred)

print('Accuracy Score:',acc_nb)
print(c_nb)

Accuracy Score: 0.8148893360160966
               precision    recall  f1-score   support

  Çerçevelik       0.75      0.95      0.84       257
Ürgüp Sivrisi       0.93      0.67      0.78       240

    accuracy                           0.81       497
   macro avg       0.84      0.81      0.81       497
weighted avg       0.84      0.81      0.81       497
```

<table>
<tr>
<td>Gradient Boosting Classifier</td>
<td>



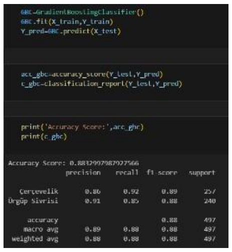
```
support_vector=SVC()
support_vector.fit(X_train,Y_train)
Y_pred=support_vector.predict(X_test)
```

```
acc_svc=accuracy_score(Y_test,Y_pred)
c_svc=classification_report(Y_test,Y_pred)
```

```
print('Accuracy Score:',acc_svc)
print(c_svc)
```

```
Accuracy Score: 0.6680080482897385
              precision    recall  f1-score   support

  Çerçevelik       0.64      0.80      0.71       257
Ürgüp Sivrisi      0.71      0.53      0.60       240

    accuracy                           0.67       497
   macro avg       0.68      0.66      0.66       497
weighted avg       0.68      0.67      0.66       497
```

</td>
<td>0.66</td>
</tr>
<tr>
<td>Support vector classifier</td>
<td>



```
GBC=GradientBoostingClassifier()
GBC.fit(X_train,Y_train)
Y_pred=GBC.predict(X_test)
```

```
acc_gbc=accuracy_score(Y_test,Y_pred)
c_gbc=classification_report(Y_test,Y_pred)
```

```
print('Accuracy Score:',acc_gbc)
print(c_gbc)
```

```
Accuracy Score: 0.8832997987927566
              precision    recall  f1-score   support

  Çerçevelik       0.86      0.92      0.89       257
Ürgüp Sivrisi      0.91      0.85      0.88       240

    accuracy                           0.88       497
   macro avg       0.89      0.88      0.88       497
weighted avg       0.88      0.88      0.88       497
```

</td>
<td>0.88</td>
</tr>
</table>