

Model Development Phase Template

Date	03 June 2024
Team ID	739676
Project Title	Harvesting Brilliance : A Taxanomic Tale of Pumpkin Seeds Varieties
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	<pre>random_forest_classifier() 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)</pre> <p>Accuracy Score: 0.8751580543299</p> <table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>cerisevolik</td><td>0.86</td><td>0.91</td><td>0.88</td><td>257</td></tr><tr><td>drdp SlvrIsi</td><td>0.90</td><td>0.84</td><td>0.87</td><td>248</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.88</td><td>497</td></tr><tr><td>macro avg</td><td>0.88</td><td>0.87</td><td>0.87</td><td>497</td></tr><tr><td>weighted avg</td><td>0.88</td><td>0.88</td><td>0.87</td><td>497</td></tr></tbody></table>		precision	recall	f1-score	support	cerisevolik	0.86	0.91	0.88	257	drdp SlvrIsi	0.90	0.84	0.87	248	accuracy			0.88	497	macro avg	0.88	0.87	0.87	497	weighted avg	0.88	0.88	0.87	497	0.875
	precision	recall	f1-score	support																												
cerisevolik	0.86	0.91	0.88	257																												
drdp SlvrIsi	0.90	0.84	0.87	248																												
accuracy			0.88	497																												
macro avg	0.88	0.87	0.87	497																												
weighted avg	0.88	0.88	0.87	497																												

Logistic regression	<pre>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)</pre> <table><tr><th colspan="5">Accuracy Score: 0.860753917661</th></tr><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr><tr><td>Cerebellik</td><td>0.85</td><td>0.93</td><td>0.88</td><td>257</td></tr><tr><td>Urgip Sivrisi</td><td>0.85</td><td>0.83</td><td>0.86</td><td>240</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.87</td><td>497</td></tr><tr><td>macro avg</td><td>0.87</td><td>0.87</td><td>0.87</td><td>497</td></tr><tr><td>weighted avg</td><td>0.87</td><td>0.87</td><td>0.87</td><td>497</td></tr></table>	Accuracy Score: 0.860753917661						precision	recall	f1-score	support	Cerebellik	0.85	0.93	0.88	257	Urgip Sivrisi	0.85	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	0.86
Accuracy Score: 0.860753917661																																					
	precision	recall	f1-score	support																																	
Cerebellik	0.85	0.93	0.88	257																																	
Urgip Sivrisi	0.85	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	<pre>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)</pre> <table><tr><th colspan="5">Accuracy Score: 0.829738436587581</th></tr><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr><tr><td>Cerebellik</td><td>0.81</td><td>0.87</td><td>0.84</td><td>257</td></tr><tr><td>Urgip Sivrisi</td><td>0.85</td><td>0.78</td><td>0.82</td><td>240</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.83</td><td>497</td></tr><tr><td>macro avg</td><td>0.83</td><td>0.83</td><td>0.83</td><td>497</td></tr><tr><td>weighted avg</td><td>0.83</td><td>0.83</td><td>0.83</td><td>497</td></tr></table>	Accuracy Score: 0.829738436587581						precision	recall	f1-score	support	Cerebellik	0.81	0.87	0.84	257	Urgip 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	0.82
Accuracy Score: 0.829738436587581																																					
	precision	recall	f1-score	support																																	
Cerebellik	0.81	0.87	0.84	257																																	
Urgip 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	<pre>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)</pre> <table><tr><th colspan="5">Accuracy Score: 0.81466936458906</th></tr><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr><tr><td>Cerebellik</td><td>0.75</td><td>0.95</td><td>0.84</td><td>257</td></tr><tr><td>Urgip Sivrisi</td><td>0.93</td><td>0.67</td><td>0.78</td><td>240</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.81</td><td>497</td></tr><tr><td>macro avg</td><td>0.84</td><td>0.81</td><td>0.81</td><td>497</td></tr><tr><td>weighted avg</td><td>0.86</td><td>0.81</td><td>0.83</td><td>497</td></tr></table>	Accuracy Score: 0.81466936458906						precision	recall	f1-score	support	Cerebellik	0.75	0.95	0.84	257	Urgip Sivrisi	0.93	0.67	0.78	240	accuracy			0.81	497	macro avg	0.84	0.81	0.81	497	weighted avg	0.86	0.81	0.83	497	0.81
Accuracy Score: 0.81466936458906																																					
	precision	recall	f1-score	support																																	
Cerebellik	0.75	0.95	0.84	257																																	
Urgip Sivrisi	0.93	0.67	0.78	240																																	
accuracy			0.81	497																																	
macro avg	0.84	0.81	0.81	497																																	
weighted avg	0.86	0.81	0.83	497																																	
Gradient Boosting Classifier	<pre>support_vector_classifier() support_vector_classifier.fit(X_train,Y_train) Y_pred=support_vector_classifier.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)</pre> <table><tr><th colspan="5">Accuracy Score: 0.660888822807385</th></tr><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr><tr><td>Cerebellik</td><td>0.64</td><td>0.89</td><td>0.71</td><td>257</td></tr><tr><td>Urgip Sivrisi</td><td>0.71</td><td>0.53</td><td>0.60</td><td>240</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.67</td><td>497</td></tr><tr><td>macro avg</td><td>0.68</td><td>0.66</td><td>0.66</td><td>497</td></tr><tr><td>weighted avg</td><td>0.68</td><td>0.67</td><td>0.66</td><td>497</td></tr></table>	Accuracy Score: 0.660888822807385						precision	recall	f1-score	support	Cerebellik	0.64	0.89	0.71	257	Urgip 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	0.66
Accuracy Score: 0.660888822807385																																					
	precision	recall	f1-score	support																																	
Cerebellik	0.64	0.89	0.71	257																																	
Urgip 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																																	
Support vector classifier	<pre>svm=svm.SVC(kernel='rbf') svm.fit(X_train,Y_train) Y_pred=svm.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)</pre> <table><tr><th colspan="5">Accuracy Score: 0.883792982927566</th></tr><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr><tr><td>Cerebellik</td><td>0.86</td><td>0.92</td><td>0.89</td><td>257</td></tr><tr><td>Urgip Sivrisi</td><td>0.71</td><td>0.85</td><td>0.80</td><td>240</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.88</td><td>497</td></tr><tr><td>macro avg</td><td>0.88</td><td>0.88</td><td>0.88</td><td>497</td></tr><tr><td>weighted avg</td><td>0.88</td><td>0.88</td><td>0.88</td><td>497</td></tr></table>	Accuracy Score: 0.883792982927566						precision	recall	f1-score	support	Cerebellik	0.86	0.92	0.89	257	Urgip Sivrisi	0.71	0.85	0.80	240	accuracy			0.88	497	macro avg	0.88	0.88	0.88	497	weighted avg	0.88	0.88	0.88	497	0.88
Accuracy Score: 0.883792982927566																																					
	precision	recall	f1-score	support																																	
Cerebellik	0.86	0.92	0.89	257																																	
Urgip Sivrisi	0.71	0.85	0.80	240																																	
accuracy			0.88	497																																	
macro avg	0.88	0.88	0.88	497																																	
weighted avg	0.88	0.88	0.88	497																																	