Model Development Phase

Date	10 July 2024
Team ID	SWTID1720078683
Project Title	Anemia Sense: Leveraging Machine Learning for Precise Anemia Recognitions
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:

```
log = LogisticRegression()

log.fit(x_train,y_train)

▼ LogisticRegression
LogisticRegression()
```

```
rf = RandomForestClassifier()

rf.fit(x_train,y_train)

▼ RandomForestClassifier
RandomForestClassifier()
```

```python
dec = DecisionTreeClassifier()

dec.fit(x_train,y_train)
```

▼ DecisionTreeClassifier
DecisionTreeClassifier()

```python
NB = GaussianNB()

NB.fit(x_train,y_train)
```

▼ GaussianNB
GaussianNB()

```python
SVM = SVC()

SVM.fit(x_train,y_train)
```

▼ SVC
SVC()

```python
GB = GradientBoostingClassifier()

GB.fit(x_train,y_train)
```

▼ GradientBoostingClassifier
GradientBoostingClassifier()

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix
Logistic Regression		0.9798	
Random Forest Classifier		1.00	
Decision Tree Classifier		1.00	

Logistic Regression classification report:

```
acc_lr = accuracy_score(y_test,y_predict)

acc_lr

0.9798387096774194

rep_lr = classification_report(y_test,y_predict)

print(rep_lr)
              precision    recall  f1-score   support

           0       0.99      0.97      0.98       123
           1       0.97      0.99      0.98       125

    accuracy                           0.98       248
   macro avg       0.98      0.98      0.98       248
weighted avg       0.98      0.98      0.98       248
```

Logistic Regression confusion matrix:

```
confusion_matrix(y_test,y_predict)

array([[119,   4],
       [  1, 124]], dtype=int64)
```

Random Forest Classifier classification report:

```
acc_rf = accuracy_score(y_test,y_predict)
acc_rf

1.0

rep_rf = classification_report(y_test,y_predict)
print(rep_rf)

              precision    recall  f1-score   support

           0       1.00      1.00      1.00       123
           1       1.00      1.00      1.00       125

    accuracy                           1.00       248
   macro avg       1.00      1.00      1.00       248
weighted avg       1.00      1.00      1.00       248
```

Random Forest Classifier confusion matrix:

```
confusion_matrix(y_test,y_predict)

array([[123,   0],
       [  0, 125]], dtype=int64)
```

Decision Tree Classifier classification report:

```
acc_dc = accuracy_score(y_test,y_predict)
acc_dc

1.0

rep_dc = classification_report(y_test,y_predict)
print(rep_dc)

              precision    recall  f1-score   support

           0       1.00      1.00      1.00       123
           1       1.00      1.00      1.00       125

    accuracy                           1.00       248
   macro avg       1.00      1.00      1.00       248
weighted avg       1.00      1.00      1.00       248
```

Decision Tree Classifier confusion matrix:

```
confusion_matrix(y_test,y_predict)

array([[123,   0],
       [  0, 125]], dtype=int64)
```

| Gaussian Naïve Bayes | ```
acc_NB = accuracy_score(y_test,y_predict)
acc_NB

0.9516129032258065

rep_NB = classification_report(y_test,y_predict)
print(rep_NB)

              precision    recall  f1-score   support

           0       0.97      0.93      0.95       123
           1       0.93      0.98      0.95       125

    accuracy                           0.95       248
   macro avg       0.95      0.95      0.95       248
weighted avg       0.95      0.95      0.95       248
``` | 0.9516 | ```
confusion_matrix(y_test,y_predict)

array([[113,  10],
       [  2, 123]], dtype=int64)
``` |
| Support Vector Machine | ```
acc_svm = accuracy_score(y_test,y_predict)
acc_svm

0.9032258064516129

rep_svm = classification_report(y_test,y_predict)
print(rep_svm)

              precision    recall  f1-score   support

           0       0.98      0.82      0.89       123
           1       0.85      0.98      0.91       125

    accuracy                           0.90       248
   macro avg       0.91      0.90      0.90       248
weighted avg       0.91      0.90      0.90       248
``` | 0.9032 | ```
confusion_matrix(y_test,y_predict)

array([[101,  22],
       [  2, 123]], dtype=int64)
``` |
| Gradient Boosting Classifier | ```
acc_GB = accuracy_score(y_test,y_predict)
acc_GB

1.0

rep_GB = classification_report(y_test,y_predict)
print(rep_GB)

              precision    recall  f1-score   support

           0       1.00      1.00      1.00       123
           1       1.00      1.00      1.00       125

    accuracy                           1.00       248
   macro avg       1.00      1.00      1.00       248
weighted avg       1.00      1.00      1.00       248
``` | 1.00 | ```
confusion_matrix(y_test,y_predict)

array([[119,   4],
       [  1, 124]], dtype=int64)
``` |