Model Development Phase

Date	26th June 2024
Team ID	SWTID1720080161
Project Title	Revolutionizing Liver Care : Predicting Liver Cirrhosis Using Advanced Machine Learning Techniques
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:

Using SVM to test the model

Splitting the data into Train and Test

```
[695] from sklearn.model_selection import train_test_split, cross_val_score
```

```
[696] X_train, X_test, y_train, y_test = train_test_split(X, y_encoded, test_size=0.2, random_state=42)
```

Since the outcome is highly skewed we oversample the data

```
[697] from imblearn.over_sampling import RandomOverSampler
      os=RandomOverSampler(random_state=0)
      X_resampled, y_resampled = os.fit_resample(X_train, y_train)
```

```
[88] model = svm.SVC()
     model.fit(X_resampled, y_resampled)
     y_pred = model.predict(X_test)
     print("Test Accuracy:", accuracy_score(y_test, y_pred))

     from sklearn.metrics import confusion_matrix, classification_report

     confusion_matrix = confusion_matrix(y_test, y_pred)

     print("Confusion Matrix:")
     print(confusion_matrix)

     classification_report = classification_report(y_test, y_pred)

     print("Classification Report:")
     print(classification_report)
```

```
Test Accuracy: 0.902834008097166
```

Using Logistic Regression to test the model

```
from sklearn.metrics import confusion_matrix, classification_report

model = LogisticRegression(penalty="l1",C=0.01,solver="liblinear")
model.fit(X_resampled, y_resampled)


y_pred = model.predict(X_test)

print("Test Accuracy:", accuracy_score(y_test, y_pred))
```

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy
SUPPORT VECTOR MACHINE	Classification Report: precision recall f1-score support 0 0.72 0.97 0.82 58 1 0.99 0.88 0.93 189 accuracy 0.90 247 macro avg 0.85 0.92 0.88 247 weighted avg 0.92 0.90 0.91 247	Test Accuracy: 0.902834008097166
Model 2	Screenshot of the classification report	Accuracy Value
LOGISTIC REGRESSION	Classification Report: precision recall f1-score support 0 0.85 0.97 0.90 58 1 0.99 0.95 0.97 189 accuracy 0.95 247 macro avg 0.92 0.96 0.94 247 weighted avg 0.96 0.95 0.95 247	Test Accuracy: 0.951417004048583 Confusion Matrix:

MODEL 1 CONFUSION MATRIX	MODEL 2 CONFUSION MATRIX
Confusion Matrix: [[56 2] [18 171]]	Confusion Matrix: [[56 2] [10 179]]