**Model Development Phase Template**

|  |  |
| --- | --- |
| Date | 09 July 2024 |
| Team ID | SWTID1720023141 |
| Project Title | Prediction and Analysis of Liver Patient  Data 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:**

Paste the screenshot of the model training code

**Model Validation and Evaluation Report:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Classification Report** | **Accuracy** | **Confusion Matrix** |
| Logistic  Regression | Logistic Regression is a linear  model used for binary  classification tasks. It  estimates the probability that  a given input belongs to a  particular class by fitting a  logistic function (sigmoid) to  the linear combination of  input features. It's simple,  interpretable, and works well  for linearly separable data. | random\_state=42 | Accuracy score = 74% |
| K  neighbors  Classifier | The KNeighbors Classifier  (k-NN) is a non-parametric,  instance-based learning  algorithm. It classifies a data  point based on the majority  class among its k-nearest  neighbors in the feature  space. The value of k  determines the number of  neighbors considered. | n\_neighbors=6,  weights='uniform',  algorithm='kd\_tree',  leaf\_size=20 | Accuracy score = 77% |
| Random  Forest  Classifier | Random Forest Classifier is  an ensemble learning method  that combines the predictions  of multiple decision trees to  improve accuracy and prevent  overfitting. Each tree is  trained on a random subset of  the data and features, and  their results are aggregated  for final classification. This  approach leverages the  diversity of the individual  trees to enhance robustness  and accuracy. | n\_estimators=500,cri  terion='entropy',rand  om\_state=18 | Accuracy score = 72% |
| SVC | Support Vector Classifier  (SVC) is a supervised  learning model that constructs  a hyperplane or set of  hyperplanes in a highdimensional space to classify  data points. SVC aims to  maximize the margin between  different classes, making it  robust to outliers and  effective for high-dimensional  data . | kernel="rbf",random  \_state=100,gamma='  auto',verbose=2,deci  sion\_function\_shape  ='ovo' | Accuracy score = 78% |