1. Name

-RubelAhmod

2. StudentID(IDthatisrelatedto thecourse)

-2

3. ProjectTitle

- HeartDisease Prediction

4. Provideabriefdescriptionofyourproject(2-3sentences).

- Thisprojectaimstopredictthepresenceof heartdiseaseusing patienthealthdata.Variousmachine learning models are used to analyze medical attributes and determine whether a person is at risk.

5. Whatistheprimaryproblemyourprojectaimstosolve?

- Tohelpinearlydetectionof heartdiseasebasedon patienthealthdata,allowingfortimelymedical intervention.

6. Whoarethestakeholdersorbeneficiariesofyourproject?

- Doctors,Patients,Healthcareinstitutions,Medicalresearchers

7. Whatdataset(s)didyouuse?Providethesource(s).

-Dataset:**HeartDiseaseDataset**
-Source:Kaggle

8. Howdidyoupreprocessandcleanthe data?

-HandlingMissingValuesusingdifferentimputationtechniqueslikeSimpleImputer, KNNImputer, and IterativeImputer.

-FeatureScalingusingStandardScalerandMinMaxScaler.

-LabelEncodingforcategoricaldata.

9. Didyoufaceanymissingdataoroutliers?Howdidyouhandlethem?

-MissingData**:**Handledusingmultipleimputationtechniques.
-Outliers**:**Nospecificdetailsfoundinthenotebookabouthandlingoutliers.10.Whatexploratory

10. dataanalysis(EDA)techniquesdidyouapply?

- Heatmaps(correlationanalysis),Pairplots,Seaborn&Plotlyvisualizations

11. Didyoucreateanynewfeatures?Ifyes,explaintheirsignificance.

-No

12. Whichmachinelearningmodelsortechniquesdidyouuse?Why?

- LogisticRegression
- K-NearestNeighbors(KNN)
- SupportVectorMachine(SVM)
- DecisionTree
- RandomForest
- AdaBoost&GradientBoosting

13. Describeyourmodeltrainingandvalidationapproach(e.g.,train-testsplit,cross-validation).

- Train-test split:80%training,20%testing.
- Cross-validationused forbettergeneralization.

14. Whatperformancemetricsdidyouusetoevaluateyourmodel?

 -AccuracyScorePrecision,Recall,F1-Score, ConfusionMatrix

15. Whatwerethekeychallengesinmodelselectionandtraining?

- ImbalancedData:Somemodelsmaystrugglewithclass imbalance.
- Overfitting inDecisionTrees:The modelmightperformwellontrainingdatabut poorlyon test data.
- HyperparameterTuning:Requiredoptimizationforbettermodelperformance.-IntheDecision tree that the training error is 0, while the test error is 0.58 so it is an overfitted model.

-Inrandomforestmodel ithaslesserrorinthetestdata

16. Whatwereyourmodel'sfinalresults(accuracy,precision,recall,F1-score,RMSE,etc.)?

- Exactresultsarenotextracted,butthemodelswere evaluatedbasedon accuracy**,**precision**,**recall**,** and F1-score.

 17. Didyoudeployyourmodel(e.g.,FlaskAPI,Streamlitapp)?Ifyes,providedetails.

-Yes --> https://heart-disease-prediction-1234.streamlit.app/

18. ProvidealinktoyourGitHubrepositoryInterpretationandInsights

 - https://github.com/rubelahmod/Heart_desease_detection.git

19. Whatkeyinsightsdidyougainfromyourproject?

- Learnedhowdifferenthealthfeaturesrelatetoheartdisease.

- Understoodhowdifferentmodelsperformonthistypeofdata

20. Whatarethestrengths andlimitationsofyourapproach?

- Strength:Usedmultiple modelsto compareresults.

-Limitation:Mightneed betterhandlingofclassimbalance

21. Ifyouhadmoretime,whatwouldyouimproveoraddto thisproject?

 - Trymoremodels likeXGBoost.

-Improvedatasetsizeforbetterpredictions.