Project Design Phase-
IITechnologyStack(Architecture&Stack)

TeamID	PNT2022TMID01568
ProjectName	*Project – Visualizing and predicting heartdiseaseswithaninteractivedashboar d.*
MaximumMarks	*4 Marks*

TechnicalArchitecture: *TheDeliverableshallincludethearchitecturaldiagramasbelowandtheinformationas perthetable1&table2*

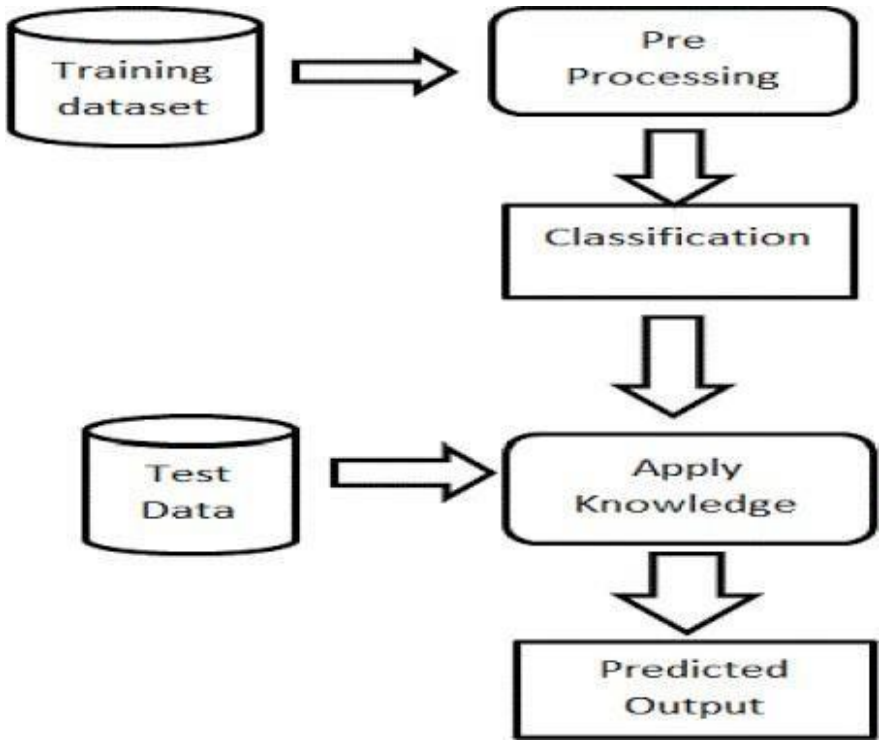
Training dataset
Pre Processing
Classification
Test Data
Apply Knowledge
Predicted Output

Table-1: Components& Technologies:

S.No	*Component*	*Description*	*Technology*
1.	*Importingdata*	*DataImport letsyouuploaddatafromexternalsourcesandcombi neit withdatayoucollect viaAnalytics*	*Python,numpy,pandas.*
2.	*DataCleaning*	*Datacleaningisaprocessbywhichinaccurat e,poorlyformatted, or otherwise messydata isorganizedandcorrected*	*Python,numpy,pandas*
3.	*DataPreprocessing*	*Data preprocessing, a component of data preparation,describesanytypeofprocessingperfor medonrawdatatoprepareitforanotherdata processingprocedure*	*Python,numpy,scipy,pandas*
4.	*Trainingdata*	*Trainingdataisthesubsetoforiginaldatathatisused totrain themachinelearningmodel,*	*Numpy,scipy, pandas*
5.	*Testingdata*	*Testdataisdatawhichhasbeenspecificallyidentifiedfo rusein tests,typically of acomputerprogram.*	*Numpy,scipy, pandas*
6.	*Machinelearningmodel*	*Amachinelearning modelisafilethathasbeen trainedto recognize certain types of patterns. Youtrain a modelover a set of data, providing it an algorithm that it can usetoreason over andlearnfromthosedata*	*Numpy,scipy, pandas,sklearn*
7.	*Improvemodelperformance*	*Accuracyisonemetricforevaluatingclassificationm odels. Informally, accuracy is the fraction ofpredictionsour modelgotright.*	*sklearn*

8.	*Checkingaccuracy*	*A data accuracy check, sometimes called a data sanitycheck,isasetofqualityvalidationsthattakeplac ebeforeusingdata.*	*Sklearn*

Table-2:ApplicationCharacteristics:

S.No	***Characteristics***	***Description***	***Technology***
1.	*Collectionofdata*	*Datacollection istheprocessofgathering,measuring,anda nalyzingaccuratedatafrom avariety of relevant sources to find answers toresearch problems, answer questions, evaluateoutcomes,andforecasttrendsand probabilities*	*IBMCognos,Python.*
2.	*EDAAnalysis*	*Exploratory Data Analysis (EDA) is an approach toanalyze the data using visual techniques. It is used todiscover trends, patterns, or to check assumptions withthe help of statistical summaryand graphicalrepresentations*	*Python, EDAtools*
3.	*Train& Testsplit ofdata*	*Thetrain- testsplitisusedtoestimatetheperformance of machine learning algorithms thatare applicable for prediction-basedAlgorithms/Applications.Thismethodisa fast andeasy procedure to perform such that we cancompare our own machine learning model resultstomachine results.*	*IBMCloud,Python.*
4.	*Modelprediction*	*Predictive modelling is a commonly usedstatisticaltechniquetopredictfuturebe haviour.*	*CreationofDashboardusingIBMCognos.*