**Input-Output in data:**

* age - age in years
* sex - (1 = male; 0 = female)
* cp - chest pain type
* trestbps - resting blood pressure (in mm Hg on admission to the hospital)
* chol - serum cholestoral in mg/dl
* fbs - (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
* restecg - resting electrocardiographic results
* thalach - maximum heart rate achieved
* exang - exercise induced angina (1 = yes; 0 = no)
* oldpeak - ST depression induced by exercise relative to rest
* slope - the slope of the peak exercise ST segment
* ca - number of major vessels (0-3) colored by flourosopy
* thal - 3 = normal; 6 = fixed defect; 7 = reversable defect
* target - have disease or not (1=yes, 0=no)

**Input:**

* age - age in years
* sex - (1 = male; 0 = female)
* cp - chest pain type
* trestbps - resting blood pressure (in mm Hg on admission to the hospital)
* chol - serum cholestoral in mg/dl
* fbs - (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
* restecg - resting electrocardiographic results
* thalach - maximum heart rate achieved
* exang - exercise induced angina (1 = yes; 0 = no)
* oldpeak - ST depression induced by exercise relative to rest
* slope - the slope of the peak exercise ST segment
* ca - number of major vessels (0-3) colored by flourosopy
* thal - 3 = normal; 6 = fixed defect; 7 = reversable defect

**Output:**

* target - have disease or not (1=yes, 0=no)

**SLIDE:**

**Heart-Disease Prediction System**

* **Real-World Problem:**
  + **Heart-Disease Prediction**
* **Treated as:**
  + **Supervised Machine Learning Problem**
* **Note:**
  + **Heart-Disease Prediction Problem is treated as a**
    - **Binary Classification Problem because**
      * **The Main AIM is to Distinguish between Two Classes**
        + **Class 01 = Have Disease (1)**
        + **Class 02 = Not Have Disease (0)**
* **Goal:**
  + **Learn an Input-Output Function**
    - **Learn from Input to Predict Output**

**SLIDE:**

**Heart-Disease Prediction System – TASK:**

* **Given:**
  + **A Patient (Represented as Set of Attributes)**
* **Task:**
  + **Automatically Predict whether the Patient have Heart Disease or Not.**

**SLIDE:**

**Heart-Disease Prediction System – TASK:**

* **Input:**
  + **A Patient**
* **Output:**
  + **Have Heart-Disease/Not Have Heart-Disease.**

**SLIDE:**

* **In Kaggle Heart-Disease Dataset, A Patient is represented with many attributes**
* **Kaggle Heart-Disease Dataset:**
  + [**https://www.kaggle.com/code/cdabakoglu/heart-disease-classifications-machine-learning/notebook**](https://www.kaggle.com/code/cdabakoglu/heart-disease-classifications-machine-learning/notebook)
* **For Simplicity and to explain things more clearly** 
  + **In this, Lecture, we have represented a Patient with Five Attributes.**

**SLIDE:**

**Heart-Disease Predication System – Input Attributes:**

* **In this lecture, a Patient is represented with the following Five Attributes**
* **Attribute 01 – Age:**
  + age in years
* **Attribute 02 – Sex:**
  + 1 = male
  + 0 = female
* **Attribute 03 – Cp:**
  + Possible Value 01 = Zero
  + Possible Value 02 = One
  + Possible Value 03 = Two
  + Possible Value 04 = Three
* **Attribute 04 – Chol:**
  + chol - serum cholestoral in mg/dl

**SLIDE:**

**Heart-Disease Prediction System – Output Attributes:**

* **In Heart-Disease Dataset, there is One Output Attribute**
  + **Attribute 05 – Target:**
    - Possible Value 01 = Yes (1)
    - Possible Value 02 = No (0)

**SLIDE:**

**Heart-Disease Prediction System – Summary (Input and Output)**

* **The Following table summarizes the Input and Output Attributes for Heart-Disease Dataset**

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute no.** | **Attribute Names** | **Possible Values** | **Data Types** |
| **1** | **Age** | **Age in years** | **Regression** |
| **2** | **Sex** | **Male(1), Female(0)** | **Categorical** |
| **3** | **CP** | **Zero, One, Two, Three** | **Categorical** |
| **4** | **Col** | **Cholesterol Measures** | **Regression** |
| **5** | **Target** | **Yes, No** | **Categorical** |

**SLIDE:**

**Heart-Disease Prediction System:**

* **Task**
  + **Develop a Heart-Disease Prediction System to Predict the Disease of a Patient.**
* **Input**
  + **Four Attributes**

|  |
| --- |
| 1. **Age** 2. **Sex** 3. **CP** 4. **Col** |

* **Output**
  + **One Attribute**

|  |
| --- |
| 1. **Yes** |

* **Treated as a**
  + **Supervised Machine Learning Problem**
* **Goal**
  + **Learn an Input-Output Function**
    - **Learn form Input to predict the Output**

**SLIDE:**

**Heart-Disease Prediction System is a Classification Problem**

* **Heart-Disease Prediction System is a Classification Problem because**
  + **Output is Categorical**

**SLIDE:**

**Heart-Disease Prediction System – Input and Output**

* **Input**
  + **Categorical**
  + **Regression**
* **Output**
  + **Categorical**

**SLIDE:**

**Project Focus**

|  |
| --- |
| **Heart-Disease Prediction System** |

**SLIDE:**

**Steps – Treating Heart-Disease Prediction Problem as a Classification Problem**

* **In Sha Allah (انشاء اللہ), I will follow the following steps to treat the Titanic Passenger Survival Prediction Problem as a Classification Problem** 
  + **Step 1: Decide the Learning Settings**
  + **Step 2: Obtain Sample Data**
  + **Step 3: Understand and Pre-process Sample Data**
  + **Step 4: Represent Sample Data in Machine Understandable Format**
  + **Step 5: Select Suitable Machine Learning Algorithms**
  + **Step 6: Split Sample Data into Training Data and Testing Data**
  + **Step 7: Select Suitable Evaluation Measure(s)**
  + **Step 8: Execute First Two Phases of Machine Learning Cycle**
    - **Training Phase**
    - **Testing Phase**
  + **Step 9: Analyze Results**

|  |
| --- |
| **If (Results are Good)**  **Then**  **Move to the Next Step**  **Else**  **Go to Step 1** |

* + **Step 10: Execute 3rd and 4th Phases of Machine Learning Cycle**
    - **Application Phase**
    - **Feedback Phase**
  + **Step 11: Based on Feedback**
    - **Go to Step 1 and Repeat all the Steps**

|  |
| --- |
| **Step 1: Decide the Learning Setting** |

**SLIDE**

**Step 1: Decide the Learning Setting**

* **In Sha Allah (انشاء اللہ), I will treat the Heart-Disease Prediction Problem as a** 
  + **Supervised Machine Learning Problem**
* **Since Output is Categorical, it will be treated as a**
  + **Classification Problem**

|  |
| --- |
| **Step 2: Obtain Sample Data** |

**SLIDE**

**Step 2: Obtain Sample Data**

* **Since I am Treating Titanic Heart-Disease Prediction Problem as a Supervised Machine Learning Problem, I will need**
  + **Annotated Data**
* **For more accurate learning, I need**
  1. **Large amount of Annotated Data**
  2. **High-quality Annotated Data**
  3. **Balanced Data**
* **Note**
  + **For simplicity, In Sha Allah (انشاء اللہ) I will use a toy Corpus / Dataset of 100 instances**

**SLIDE**

**Obtain Sample Data Cont…**

* **Total Instances in Sample Data = 100**
  + **Survived = 50**
  + **Not Survived = 50**

**SLIDE**

**Sample Data**

* **We obtained a Sample Data of 100 instances**
  + **See heart-disease-sample-data.csv File in Supporting Material**
* **The following Table shows the Sample Data**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Instance No.** | **Input** | | | | **Output** |
| **Age** | **Sex** | **CP** | **CHOL** | **Target** |
| **x1** | **63** | **1** | **3** | **233** | **1** |
| **x2** | **37** | **1** | **2** | **250** | **1** |
| **x3** | **41** | **0** | **1** | **204** | **1** |
| **x4** | **56** | **1** | **1** | **236** | **1** |
| **x5** | **57** | **0** | **0** | **354** | **1** |
| **x6** | **57** | **1** | **0** | **192** | **1** |
| **x7** | **56** | **0** | **1** | **294** | **1** |
| **x8** | **44** | **1** | **1** | **263** | **1** |
| **x9** | **52** | **1** | **2** | **199** | **1** |
| **x10** | **57** | **1** | **2** | **168** | **1** |
| **x11** | **54** | **1** | **0** | **239** | **1** |
| **x12** | **48** | **0** | **2** | **275** | **1** |
| **x13** | **49** | **1** | **1** | **266** | **1** |
| **x14** | **64** | **1** | **3** | **211** | **1** |
| **x15** | **58** | **0** | **3** | **283** | **1** |
| **x16** | **50** | **0** | **2** | **219** | **1** |
| **x17** | **58** | **0** | **2** | **340** | **1** |
| **x18** | **66** | **0** | **3** | **226** | **1** |
| **x19** | **43** | **1** | **0** | **247** | **1** |
| **x20** | **69** | **0** | **3** | **239** | **1** |
| **x21** | **59** | **1** | **0** | **234** | **1** |
| **x22** | **44** | **1** | **2** | **233** | **1** |
| **x23** | **42** | **1** | **0** | **226** | **1** |
| **x24** | **61** | **1** | **2** | **243** | **1** |
| **x25** | **40** | **1** | **3** | **199** | **1** |
| **x26** | **71** | **0** | **1** | **302** | **1** |
| **x27** | **59** | **1** | **2** | **212** | **1** |
| **x28** | **51** | **1** | **2** | **175** | **1** |
| **x29** | **65** | **0** | **2** | **417** | **1** |
| **x30** | **53** | **1** | **2** | **197** | **1** |
| **x31** | **41** | **0** | **1** | **198** | **1** |
| **x32** | **65** | **1** | **0** | **177** | **1** |
| **x33** | **44** | **1** | **1** | **219** | **1** |
| **x34** | **54** | **1** | **2** | **273** | **1** |
| **x35** | **51** | **1** | **3** | **213** | **1** |
| **x36** | **46** | **0** | **2** | **177** | **1** |
| **x37** | **54** | **0** | **2** | **304** | **1** |
| **x38** | **54** | **1** | **2** | **232** | **1** |
| **x39** | **65** | **0** | **2** | **269** | **1** |
| **x40** | **65** | **0** | **2** | **360** | **1** |
| **x41** | **51** | **0** | **2** | **308** | **1** |
| **x42** | **48** | **1** | **1** | **245** | **1** |
| **x43** | **45** | **1** | **0** | **208** | **1** |
| **x44** | **53** | **0** | **0** | **264** | **1** |
| **x45** | **39** | **1** | **2** | **321** | **1** |
| **x46** | **52** | **1** | **1** | **325** | **1** |
| **x47** | **44** | **1** | **2** | **235** | **1** |
| **x48** | **47** | **1** | **2** | **257** | **1** |
| **x49** | **53** | **0** | **2** | **216** | **1** |
| **x50** | **53** | **0** | **0** | **234** | **1** |
| **x51** | **67** | **1** | **0** | **286** | **0** |
| **x52** | **67** | **1** | **0** | **229** | **0** |
| **x53** | **62** | **0** | **0** | **268** | **0** |
| **x54** | **63** | **1** | **0** | **254** | **0** |
| **x55** | **53** | **1** | **0** | **203** | **0** |
| **x56** | **56** | **1** | **2** | **256** | **0** |
| **x57** | **48** | **1** | **1** | **229** | **0** |
| **x58** | **58** | **1** | **1** | **284** | **0** |
| **x59** | **58** | **1** | **2** | **224** | **0** |
| **x60** | **60** | **1** | **0** | **206** | **0** |
| **x61** | **40** | **1** | **0** | **167** | **0** |
| **x62** | **60** | **1** | **0** | **230** | **0** |
| **x63** | **64** | **1** | **2** | **335** | **0** |
| **x64** | **43** | **1** | **0** | **177** | **0** |
| **x65** | **57** | **1** | **0** | **276** | **0** |
| **x66** | **55** | **1** | **0** | **353** | **0** |
| **x67** | **65** | **0** | **0** | **225** | **0** |
| **x68** | **61** | **0** | **0** | **330** | **0** |
| **x69** | **58** | **1** | **2** | **230** | **0** |
| **x70** | **50** | **1** | **0** | **243** | **0** |
| **x71** | **44** | **1** | **0** | **290** | **0** |
| **x72** | **60** | **1** | **0** | **253** | **0** |
| **x73** | **54** | **1** | **0** | **266** | **0** |
| **x74** | **50** | **1** | **2** | **233** | **0** |
| **x75** | **41** | **1** | **0** | **172** | **0** |
| **x76** | **51** | **0** | **0** | **305** | **0** |
| **x77** | **58** | **1** | **0** | **216** | **0** |
| **x78** | **54** | **1** | **0** | **188** | **0** |
| **x79** | **60** | **1** | **0** | **282** | **0** |
| **x80** | **60** | **1** | **2** | **185** | **0** |
| **x81** | **59** | **1** | **0** | **326** | **0** |
| **x82** | **46** | **1** | **2** | **231** | **0** |
| **x83** | **67** | **1** | **0** | **254** | **0** |
| **x84** | **62** | **1** | **0** | **267** | **0** |
| **x85** | **65** | **1** | **0** | **248** | **0** |
| **x86** | **44** | **1** | **0** | **197** | **0** |
| **x87** | **60** | **1** | **0** | **258** | **0** |
| **x88** | **58** | **1** | **0** | **270** | **0** |
| **x89** | **68** | **1** | **2** | **274** | **0** |
| **x90** | **62** | **0** | **0** | **164** | **0** |
| **x91** | **52** | **1** | **0** | **255** | **0** |
| **x92** | **59** | **1** | **0** | **239** | **0** |
| **x93** | **60** | **0** | **0** | **258** | **0** |
| **x94** | **49** | **1** | **2** | **188** | **0** |
| **x95** | **59** | **1** | **0** | **177** | **0** |
| **x96** | **57** | **1** | **2** | **229** | **0** |
| **x97** | **61** | **1** | **0** | **260** | **0** |
| **x98** | **39** | **1** | **0** | **219** | **0** |
| **x99** | **61** | **0** | **0** | **307** | **0** |
| **x100** | **56** | **1** | **0** | **249** | **0** |

**SLIDE**

**Step 3: Understand and Pre-process Sample Data**

* **Understanding Data**
  + **The Sample Data contains Five Attributes** 
    - **Age**
    - **Sex**
    - **Cp**
    - **Chol**
    - **Target**
  + **Separating Input from Output**
    - **Input comprises of Four Attributes** 
      * **Age**
      * **Sex**
      * **Cp**
      * **Chol**
    - **The Output comprises of a Single Attribute**
      * **Target**
* **Pre-processing Data**
  + **Corpus is already pre-processed**
    - **Therefore, no pre-processing is needed 😊**

|  |
| --- |
| **Step 04: Represent Data in Machine Understandable Format** |

**SLIDE**

**Step 4: Represent Sample Data in Machine Understandable Format**

* **Feature-based Classification Algorithms (implemented in Scikit-learn) can understand data in** 
  + **Attribute-Value Pair** 
    - **Values of Attributes / Features must be Numeric**
* **Problem**
  + **Our Sample Data is not in Attribute-Value Pair form**
    - **We need to transform our Sample Data into Machine Understandable Format**
* **Solution**
  + **There are many approaches to transform Sample Data into Machine Understandable Format**

**SLIDE**

**Transforming Sample Data in Machine Understandable Format**

* **In our Sample Data**
  + **Input is Categorical and Regression**
  + **Output is Categorical**
* **Considering Input (Age, Chol) we will need to** 
  + **Transform Input (Regression) into Categorical than in Numerical Representation.**

**SLIDE**

**Converting Age and Chol attributes in categorical form by using ranges in the data:**

* **Age:**
  + **In Our Age attributes the ranges will be as follows:**
    - **30-50 = Mature**
    - **51- 80 = Old**
* **Chol:**
  + **In our Chol attributes the ranges will be as follows:**
    - **170-250 = Normal**
    - **250-above = High**
* **See heart-disease-sample-data-Ranges.csv File in Supporting Material**

**SLIDE:**

**Converting Output into Numerical Representation Cont…**

* **Step 01: Define an Encoding Scheme**
* **Encoding Scheme for Survived Attribute**
  + **Age:**
    - **Mature = 0**
    - **Old = 1**
  + **Chol:**
    - **Normal = 0**
    - **High = 1**
* **Step 02: Use Encoding Scheme defined in Step 01, to convert Categorical Output Values to Numerical Input Values for all instances in the Sample Data**
* **The Table below shows Sample Data after Encoding Categorical Input Values to Numerical Output Values**
  + **See heart-disease-sample-data-encoding.csv File in Supporting Material**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Instance No.** | **Input** | | | | **Output** |
| **Age** | **Sex** | **CP** | **CHOL** | **Target** |
| **x1** | **1** | **1** | **3** | **0** | **1** |
| **x2** | **0** | **1** | **2** | **0** | **1** |
| **x3** | **0** | **0** | **1** | **0** | **1** |
| **x4** | **1** | **1** | **1** | **0** | **1** |
| **x5** | **1** | **0** | **0** | **1** | **1** |
| **x6** | **1** | **1** | **0** | **0** | **1** |
| **x7** | **1** | **0** | **1** | **1** | **1** |
| **x8** | **0** | **1** | **1** | **1** | **1** |
| **x9** | **1** | **1** | **2** | **0** | **1** |
| **x10** | **1** | **1** | **2** | **0** | **1** |
| **x11** | **1** | **1** | **0** | **0** | **1** |
| **x12** | **0** | **0** | **2** | **1** | **1** |
| **x13** | **0** | **1** | **1** | **1** | **1** |
| **x14** | **1** | **1** | **3** | **0** | **1** |
| **x15** | **1** | **0** | **3** | **1** | **1** |
| **x16** | **0** | **0** | **2** | **0** | **1** |
| **x17** | **1** | **0** | **2** | **1** | **1** |
| **x18** | **1** | **0** | **3** | **0** | **1** |
| **x19** | **0** | **1** | **0** | **0** | **1** |
| **x20** | **1** | **0** | **3** | **0** | **1** |
| **x21** | **1** | **1** | **0** | **0** | **1** |
| **x22** | **0** | **1** | **2** | **0** | **1** |
| **x23** | **0** | **1** | **0** | **0** | **1** |
| **x24** | **1** | **1** | **2** | **0** | **1** |
| **x25** | **0** | **1** | **3** | **0** | **1** |
| **x26** | **1** | **0** | **1** | **1** | **1** |
| **x27** | **1** | **1** | **2** | **0** | **1** |
| **x28** | **1** | **1** | **2** | **0** | **1** |
| **x29** | **1** | **0** | **2** | **1** | **1** |
| **x30** | **1** | **1** | **2** | **0** | **1** |
| **x31** | **0** | **0** | **1** | **0** | **1** |
| **x32** | **1** | **1** | **0** | **0** | **1** |
| **x33** | **0** | **1** | **1** | **0** | **1** |
| **x34** | **1** | **1** | **2** | **1** | **1** |
| **x35** | **1** | **1** | **3** | **0** | **1** |
| **x36** | **0** | **0** | **2** | **0** | **1** |
| **x37** | **1** | **0** | **2** | **1** | **1** |
| **x38** | **1** | **1** | **2** | **0** | **1** |
| **x39** | **1** | **0** | **2** | **1** | **1** |
| **x40** | **1** | **0** | **2** | **1** | **1** |
| **x41** | **1** | **0** | **2** | **1** | **1** |
| **x42** | **0** | **1** | **1** | **0** | **1** |
| **x43** | **0** | **1** | **0** | **0** | **1** |
| **x44** | **1** | **0** | **0** | **1** | **1** |
| **x45** | **0** | **1** | **2** | **1** | **1** |
| **x46** | **1** | **1** | **1** | **1** | **1** |
| **x47** | **0** | **1** | **2** | **0** | **1** |
| **x48** | **0** | **1** | **2** | **1** | **1** |
| **x49** | **1** | **0** | **2** | **0** | **1** |
| **x50** | **1** | **0** | **0** | **0** | **1** |
| **x51** | **1** | **1** | **0** | **1** | **0** |
| **x52** | **1** | **1** | **0** | **0** | **0** |
| **x53** | **1** | **0** | **0** | **1** | **0** |
| **x54** | **1** | **1** | **0** | **1** | **0** |
| **x55** | **1** | **1** | **0** | **0** | **0** |
| **x56** | **1** | **1** | **2** | **1** | **0** |
| **x57** | **0** | **1** | **1** | **0** | **0** |
| **x58** | **1** | **1** | **1** | **1** | **0** |
| **x59** | **1** | **1** | **2** | **0** | **0** |
| **x60** | **1** | **1** | **0** | **0** | **0** |
| **x61** | **0** | **1** | **0** | **0** | **0** |
| **x62** | **1** | **1** | **0** | **0** | **0** |
| **x63** | **1** | **1** | **2** | **1** | **0** |
| **x64** | **0** | **1** | **0** | **0** | **0** |
| **x65** | **1** | **1** | **0** | **1** | **0** |
| **x66** | **1** | **1** | **0** | **1** | **0** |
| **x67** | **1** | **0** | **0** | **0** | **0** |
| **x68** | **1** | **0** | **0** | **1** | **0** |
| **x69** | **1** | **1** | **2** | **0** | **0** |
| **x70** | **1** | **1** | **0** | **0** | **0** |
| **x71** | **0** | **1** | **0** | **1** | **0** |
| **x72** | **1** | **1** | **0** | **1** | **0** |
| **x73** | **1** | **1** | **0** | **1** | **0** |
| **x74** | **1** | **1** | **2** | **0** | **0** |
| **x75** | **0** | **1** | **0** | **0** | **0** |
| **x76** | **1** | **0** | **0** | **1** | **0** |
| **x77** | **1** | **1** | **0** | **0** | **0** |
| **x78** | **1** | **1** | **0** | **0** | **0** |
| **x79** | **1** | **1** | **0** | **1** | **0** |
| **x80** | **1** | **1** | **2** | **0** | **0** |
| **x81** | **1** | **1** | **0** | **1** | **0** |
| **x82** | **0** | **1** | **2** | **0** | **0** |
| **x83** | **1** | **1** | **0** | **1** | **0** |
| **x84** | **1** | **1** | **0** | **1** | **0** |
| **x85** | **1** | **1** | **0** | **0** | **0** |
| **x86** | **0** | **1** | **0** | **0** | **0** |
| **x87** | **1** | **1** | **0** | **1** | **0** |
| **x88** | **1** | **1** | **0** | **1** | **0** |
| **x89** | **1** | **1** | **2** | **1** | **0** |
| **x90** | **1** | **0** | **0** | **0** | **0** |
| **x91** | **1** | **1** | **0** | **1** | **0** |
| **x92** | **1** | **1** | **0** | **0** | **0** |
| **x93** | **1** | **0** | **0** | **1** | **0** |
| **x94** | **0** | **1** | **2** | **0** | **0** |
| **x95** | **1** | **1** | **0** | **0** | **0** |
| **x96** | **1** | **1** | **2** | **0** | **0** |
| **x97** | **1** | **1** | **0** | **1** | **0** |
| **x98** | **0** | **1** | **0** | **0** | **0** |
| **x99** | **1** | **0** | **0** | **1** | **0** |
| **x100** | **1** | **1** | **0** | **0** | **0** |

**SLIDE**

* **Alhamdulillah (الحمدللہ), both Input and Output are transformed into Numerical Representation**

|  |
| --- |
| **Step 06: Split Sample Data into Training Data and Testing Data** |

**SLIDE**

**Step 6: Split Sample Data into Training and Testing**

* **We Split the Sample Data using**
  + **Train-Test Split Ratio of**
    - **80% - 20%**
* **Training Data** 
  + **Total Instances = 80**
    - **Having Disease = 40**
    - **Not Having Disease = 40**
* **Testing Data** 
  + **Total Instances = 20**
    - **Having Disease = 10**
    - **Not Having Disease = 10**

**SLIDE**

**Training Data**

* **The following Table shows the Training Data**
  + **See training-sample-data.csv File in Supporting Material**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Instance No.** | **Input** | | | | **Output** |
| **Age** | **Sex** | **CP** | **Col** | **Target** |
| **x1** | **1** | **1** | **3** | **0** | **1** |
| **x2** | **0** | **1** | **2** | **0** | **1** |
| **x3** | **0** | **0** | **1** | **0** | **1** |
| **x4** | **1** | **1** | **1** | **0** | **1** |
| **x5** | **1** | **0** | **0** | **1** | **1** |
| **x6** | **1** | **1** | **0** | **0** | **1** |
| **x7** | **1** | **0** | **1** | **1** | **1** |
| **x8** | **0** | **1** | **1** | **1** | **1** |
| **x9** | **1** | **1** | **2** | **0** | **1** |
| **x10** | **1** | **1** | **2** | **0** | **1** |
| **x11** | **1** | **1** | **0** | **0** | **1** |
| **x12** | **0** | **0** | **2** | **1** | **1** |
| **x13** | **0** | **1** | **1** | **1** | **1** |
| **x14** | **1** | **1** | **3** | **0** | **1** |
| **x15** | **1** | **0** | **3** | **1** | **1** |
| **x16** | **0** | **0** | **2** | **0** | **1** |
| **x17** | **1** | **0** | **2** | **1** | **1** |
| **x18** | **1** | **0** | **3** | **0** | **1** |
| **x19** | **0** | **1** | **0** | **0** | **1** |
| **x20** | **1** | **0** | **3** | **0** | **1** |
| **x21** | **1** | **1** | **0** | **0** | **1** |
| **x22** | **0** | **1** | **2** | **0** | **1** |
| **x23** | **0** | **1** | **0** | **0** | **1** |
| **x24** | **1** | **1** | **2** | **0** | **1** |
| **x25** | **0** | **1** | **3** | **0** | **1** |
| **x26** | **1** | **0** | **1** | **1** | **1** |
| **x27** | **1** | **1** | **2** | **0** | **1** |
| **x28** | **1** | **1** | **2** | **0** | **1** |
| **x29** | **1** | **0** | **2** | **1** | **1** |
| **x30** | **1** | **1** | **2** | **0** | **1** |
| **x31** | **0** | **0** | **1** | **0** | **1** |
| **x32** | **1** | **1** | **0** | **0** | **1** |
| **x33** | **0** | **1** | **1** | **0** | **1** |
| **x34** | **1** | **1** | **2** | **1** | **1** |
| **x35** | **1** | **1** | **3** | **0** | **1** |
| **x36** | **0** | **0** | **2** | **0** | **1** |
| **x37** | **1** | **0** | **2** | **1** | **1** |
| **x38** | **1** | **1** | **2** | **0** | **1** |
| **x39** | **1** | **0** | **2** | **1** | **1** |
| **x40** | **1** | **0** | **2** | **1** | **1** |
| **x41** | **1** | **1** | **0** | **1** | **0** |
| **x42** | **1** | **1** | **0** | **0** | **0** |
| **x43** | **1** | **0** | **0** | **1** | **0** |
| **x44** | **1** | **1** | **0** | **1** | **0** |
| **x45** | **1** | **1** | **0** | **0** | **0** |
| **x46** | **1** | **1** | **2** | **1** | **0** |
| **x47** | **0** | **1** | **1** | **0** | **0** |
| **x48** | **1** | **1** | **1** | **1** | **0** |
| **x49** | **1** | **1** | **2** | **0** | **0** |
| **x50** | **1** | **1** | **0** | **0** | **0** |
| **x51** | **0** | **1** | **0** | **0** | **0** |
| **x52** | **1** | **1** | **0** | **0** | **0** |
| **x53** | **1** | **1** | **2** | **1** | **0** |
| **x54** | **0** | **1** | **0** | **0** | **0** |
| **x55** | **1** | **1** | **0** | **1** | **0** |
| **x56** | **1** | **1** | **0** | **1** | **0** |
| **x57** | **1** | **0** | **0** | **0** | **0** |
| **x58** | **1** | **0** | **0** | **1** | **0** |
| **x59** | **1** | **1** | **2** | **0** | **0** |
| **x60** | **1** | **1** | **0** | **0** | **0** |
| **x61** | **0** | **1** | **0** | **1** | **0** |
| **x62** | **1** | **1** | **0** | **1** | **0** |
| **x63** | **1** | **1** | **0** | **1** | **0** |
| **x64** | **1** | **1** | **2** | **0** | **0** |
| **x65** | **0** | **1** | **0** | **0** | **0** |
| **x66** | **1** | **0** | **0** | **1** | **0** |
| **x67** | **1** | **1** | **0** | **0** | **0** |
| **x68** | **1** | **1** | **0** | **0** | **0** |
| **x69** | **1** | **1** | **0** | **1** | **0** |
| **x70** | **1** | **1** | **2** | **0** | **0** |
| **x71** | **1** | **1** | **0** | **1** | **0** |
| **x72** | **0** | **1** | **2** | **0** | **0** |
| **x73** | **1** | **1** | **0** | **1** | **0** |
| **x74** | **1** | **1** | **0** | **1** | **0** |
| **x75** | **1** | **1** | **0** | **0** | **0** |
| **x76** | **0** | **1** | **0** | **0** | **0** |
| **x77** | **1** | **1** | **0** | **1** | **0** |
| **x78** | **1** | **1** | **0** | **1** | **0** |
| **x79** | **1** | **1** | **2** | **1** | **0** |
| **x80** | **1** | **0** | **0** | **0** | **0** |

**SLIDE**

**Testing Data**

* **The following Table shows the Testing Data**
  + **See testing-sample-data.csv File in Supporting Material**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Instance No.** | **Input** | | | | **Output** |
| **Age** | **Sex** | **CP** | **Col** | **Target** |
| **x1** | **1** | **0** | **2** | **1** | **1** |
| **x2** | **0** | **1** | **1** | **0** | **1** |
| **x3** | **0** | **1** | **0** | **0** | **1** |
| **x4** | **1** | **0** | **0** | **1** | **1** |
| **x5** | **0** | **1** | **2** | **1** | **1** |
| **x6** | **1** | **1** | **1** | **1** | **1** |
| **x7** | **0** | **1** | **2** | **0** | **1** |
| **x8** | **0** | **1** | **2** | **1** | **1** |
| **x9** | **1** | **0** | **2** | **0** | **1** |
| **x10** | **1** | **0** | **0** | **0** | **1** |
| **x11** | **1** | **1** | **0** | **1** | **0** |
| **x12** | **1** | **1** | **0** | **0** | **0** |
| **x13** | **1** | **0** | **0** | **1** | **0** |
| **x14** | **0** | **1** | **2** | **0** | **0** |
| **x15** | **1** | **1** | **0** | **0** | **0** |
| **x16** | **1** | **1** | **2** | **0** | **0** |
| **x17** | **1** | **1** | **0** | **1** | **0** |
| **x18** | **0** | **1** | **0** | **0** | **0** |
| **x19** | **1** | **0** | **0** | **1** | **0** |
| **x20** | **1** | **1** | **0** | **0** | **0** |

|  |
| --- |
| **Step 07: Select Suitable Evaluation Measure(s)** |

**SLIDE**

**Step 07: Select Suitable Evaluation Measure(s)**

* **I will use the Accuracy Evaluation Measure to evaluate the performance of the Model**
* **Accuracy**
  + **Accuracy is defined as the proportion of correctly classified Test Instances**

|  |
| --- |
|  |

* **Note**
  + **Error = 1 - Accuracy**

|  |
| --- |
| **Step 08: Execute First Two Phases of Machine Learning Cycle** |

**SLIDE**

**Step 8: Execute First Two Phases of Machine Learning Cycle**

* **Recall the Equation**

|  |
| --- |
|  |

* **Training Phase**
  + **Use Training Data to build the Model**
* **Testing Phase**
  + **Use Testing Data to evaluate the performance of the Model**
* **Note that we aim to**
  + **Learn an Input-Output Function**

**SLIDE**

**General Settings - Learning Input-Output Function**

* **Recall – Our goal is to** 
  + **Learn an Input-Output Function**

