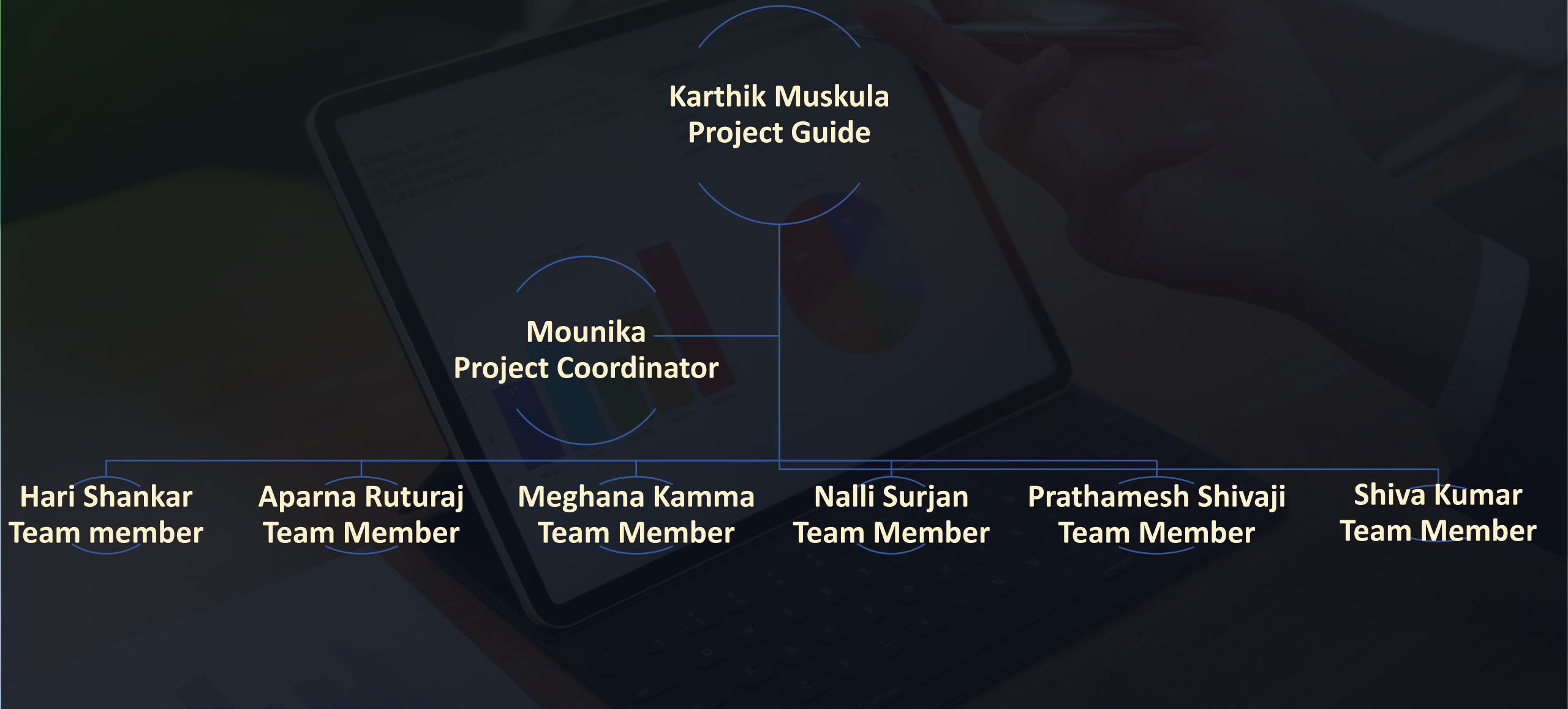
PROJECT PRESENTATION

CLASSIFICATION OF MYOCARDIAL INFARCTION

Project Team Structure:



Business Objective:

- Myocardial Infarction(Commonly known as “Heart attack”) is one of the most challenging problems of modern medicine. Acute myocardial infarction is associated with high mortality in the first year after it. The incidence of MI remains high in all countries. This is especially true for the urban population of highly developed countries, which are exposed to chronic stress factors, irregular and not always balanced nutrition.
- The Given dataset consists of 124 variables with 1700 records of patients. Classify the Lethal outcome (cause) (LET_IS)(Y variable) by using the given dataset.

Roadmap to Complete the Project

Action plan:

Sl.No	Action	Sub Action	Responsibility	Target date	Status	Remarks
1	Frame the problem	Identify the problem	All	26th July	Closed	
2	Collecting and reading the data	Collect the data and understand the data	All	29th July	Closed	
3	Exploratory data analysis	Analyze and summarize the main data sets	All	1st Aug	Closed	
4	Data cleaning or data preprocessing	Remove unnecessary columns/rows like Nulls,NA,NAA, unusual values and etc. and Fill Nulls,NA,NAA with Mean or meadian or mode or others.	All	3rd Aug	Closed	
5	Data Transformation	Converting the data into usable format or change the scaling	All	5th Aug	Closed	PCA selected for analysis
6	Data partition	Splitting the dataset in to known and unknown data sets.	All	7th Aug	Closed	
7	Selecting few models	Select the best suitable models from machine learning models	All	8th Aug	Closed	
8	Cross validation of chosen models	Re verify the selected models before any conclusion on the results	All	11th Aug	Closed	
9	Evaluation of all models and selecting the best model	Compare the accuracy of the results with the actuals, then select the best model and do the evaluation again with diff data sets for performance	All	11th Aug	Closed	
10	Prediction on new data by deployment	Final predictions for the data which is unseen data	All	18th Aug	Closed	

COLLECTING & READING THE DATA

- ✓ **The Given dataset consists of 124 variables with 1700 records of patients. Classify the Lethal outcome (cause) (LET_IS)(Y variable) by using the given dataset.**

Data Information:

- ✓ **Range Index: 1700 entries, 0 to 1699**
- ✓ **Columns: 123 entries, AGE to LET_IS**
- ✓ **Data Types: float64(110), int64(13)**
- ✓ **Memory Usage: 1.6 MB**

Duplicated Rows:

- ✓ **Number of Duplicated Rows: 0**

Additional Note:

- ✓ **The "ID" Column: It can be neglected as it is a reference to patients.neglected because It is a reference to patient**

EXPLORATORY DATA ANALYSIS

- Box plot for overall data set: To know the variable wise data range & Outliers through visually
- Data set description: Variable wise statistical summary table
- Density plot: Show the specific visualization details such as column names, data range, etc.
- Variable wise unique value count visually
- Variable wise percentage of Null Values
- Data set visualization(Null vs Values)
- Data set visualization(Variable wise Nulls vs Values)
- Correlation between Variables visually
- Pie chart for gender
- Bar graph for age vs LET_IS(outcome)

DATA CLEANING OR DATA PRE-PROCESSING

- Dropping top columns, which having high null values
- Filling missing values of columns having null count less than 112 with the help of interpolation technique
- Filling remaining missing values of remaining columns with the help of fillna() function and median central tendency
- Correlation after data preprocessing to know the relationship and visualizations

Note: **from above correlation graph we did not find any high correlated variables with outcome and having more independent variables(High dimensional data set), so we will go with dimensional reduction techniques**

DATA TRANSFORMATION

Applying Principal Component Analysis(Dimensional reduction technique) to reduce the number of independent(X) variables.

PCA and Feature Selection:

- We utilize PCA to determine the optimal number of columns for 'k' in SelectKBest().
- The Select Best() method aids in feature selection and engineering tasks.

Accuracy Assessment with KNN:

- We employ the K-Nearest Neighbors (KNN) algorithm.
- It helps us understand how varying the number of columns impacts the model's accuracy.

Shape of the data before PCA (1700, 119)
Shape of the data after PCA (1700, 11)

DATA PARTITION

Splitting the dataset in to known and unknown data sets.

In the ratio of 70%(Known) and 30%(Unknown)
Shapes of the data train(known) and test(unknown)

FINAL MODEL BUILDINGS & EVALUATION

Model buildings with for loops and their Accuracy scores:

Sl.No	Model Name	Accuracy%
1	**XGBOOST**	**98%**
2	KNN	93%
3	SVC	84%
4	Random Forest	97%
6	Naive_Bayes	41%

XGBOOST (Extreme Gradient Boosting) preforming very good compare to other classification models with 98% accuracy.

MODEL DEPLOYEMENT

Integrating ML Model into Django and Deploying to Heroku

- 4: progress of congestive heart failure
- 5: thromboembolism
- 6: asystole
- 7: ventricular fibrillation

Classifying complication level of Myocardial Infarction

AGE

STENOK_AN

ZSN_AN

S_AD_ORIT

D_AD_ORIT

K_SH_POST

ant_im

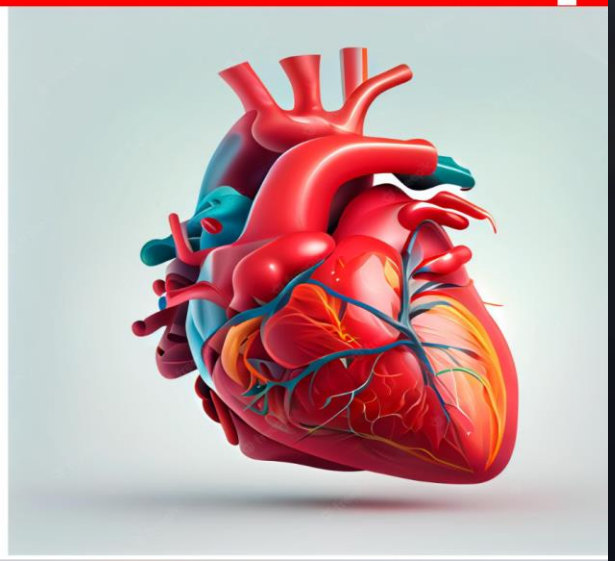
Screenshots from web

Myocardial Infarction

Myocardial infarction (MI), colloquially known as "heart attack". A heart attack usually occurs when a blood clot blocks blood flow to the heart. Without blood, tissue loses oxygen and dies. Symptoms include tightness or pain in the chest, neck, back or arms, as well as fatigue, lightheadedness, abnormal heartbeat and anxiety. MI is one of the most challenging problems of modern medicine. Acute myocardial infarction is associated with high mortality in the first year after it. The incidence of MI remains high in all countries. This is especially true for the urban population of highly developed countries, which is exposed to chronic stress factors, irregular and not always balanced nutrition. In the United States, for example, more than a million people suffer from MI every year, and 200-300 thousand of them die from acute MI before arriving at the hospital. The course of the disease in patients with MI is different. MI can occur without complications or with complications that do not worsen the long-term prognosis. At the same time, about half of patients in the acute and subacute periods have complications that lead to worsening of the disease and even death. Even an experienced specialist can not always foresee the development of these complications. In this regard, the risk of complications of myocardial infarction in order to timely carry out the necessary preventive measures is an important task.

Lethal outcome : Complication level

- 0: unknown (alive)
- 1: cardiogenic shock
- 2: pulmonary edema
- 3: myocardial rupture
- 4: progress of congestive heart failure
- 5: thromboembolism



3 : Myocardial rupture

Web address:

https://myocardial-infarction-bae0828dfd3d.herokuapp.com/predict


Thank You