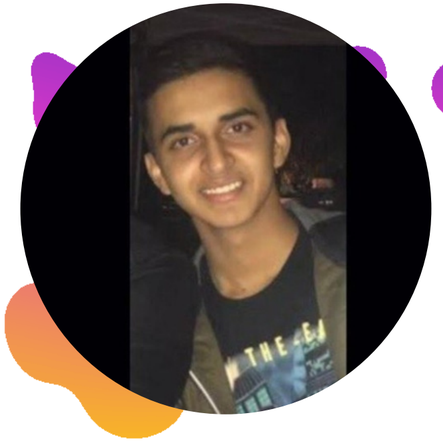
Project Exhibition II
Group 9

MACHINE LEARNING APPLICATION IN MEDICINE

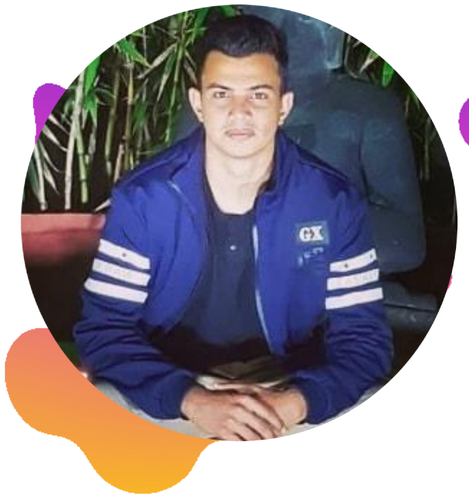
Guide: Dr. Suganya E

Team Members



Atharv Damble

20BCE10181



Shubham Sangwan

20BCE10519

Anubhav De

20BCE10085

Aashish Kumar

20BCE10163

Harshil

20BCE10071

Overview/Problem Statement

- For heart disease identification we need vascular specialists, which are very finite in number in developing nations and expensive medical procedures is not affordable for people in such countries.

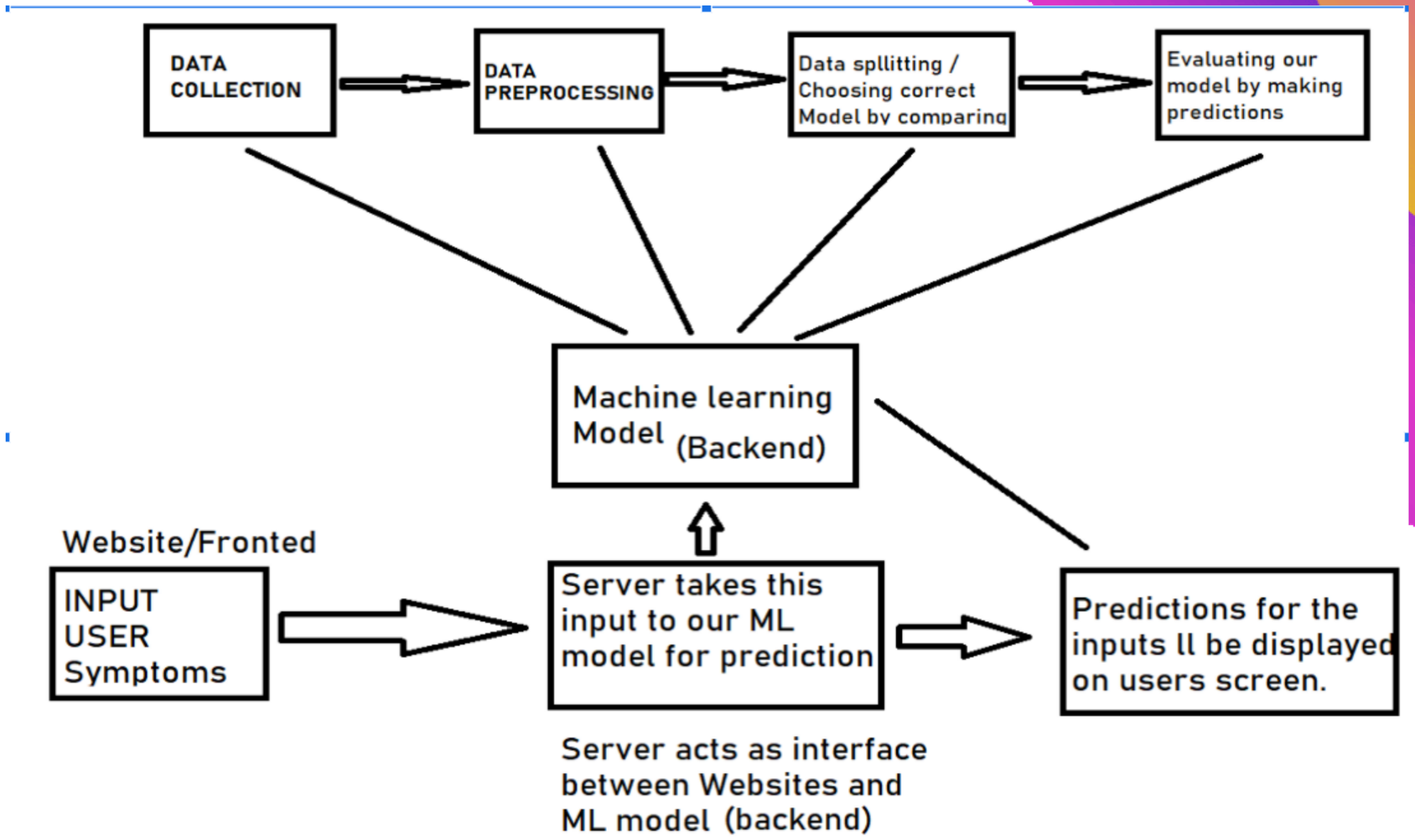
- Early detection is critical in diseases like such and this is where Machine Learning Algorithms can be applied. The ability to identify heart illnesses in a timely manner can aid in making judgments about patients, lowering their risks.

- This project presents a prediction model that uses entered symptoms in a webform to forecast whether a patient has a heart illness or not.

- We have developed a system that uses Machine Learning algorithms to forecast illnesses based on information supplied by users. It predicts the patient's ailment based on the data or symptoms submitted into the online system and provides findings based on that data. The health industry can benefit from this technique by simply asking the user for symptoms and entering them into the system, and in a matter of seconds be able to diagnose the patient and advice the needed medical care.

System Architecture Diagram

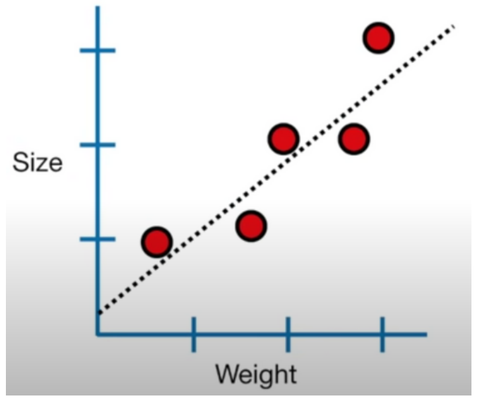
- Disease forecasting using machine learning predicts the existence of the sickness for the user oriented on various symptoms and the information entered by the user.



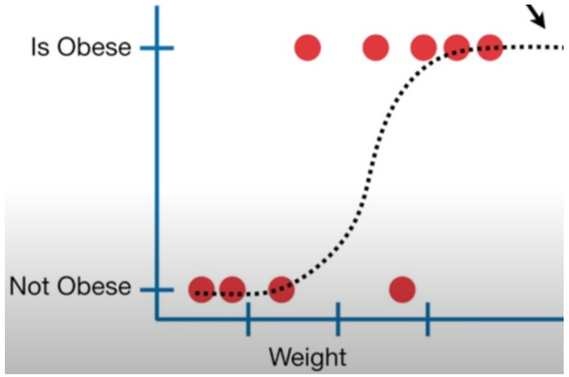
Logistic Regression

- A method of statistical analysis that models the probability of an event

-

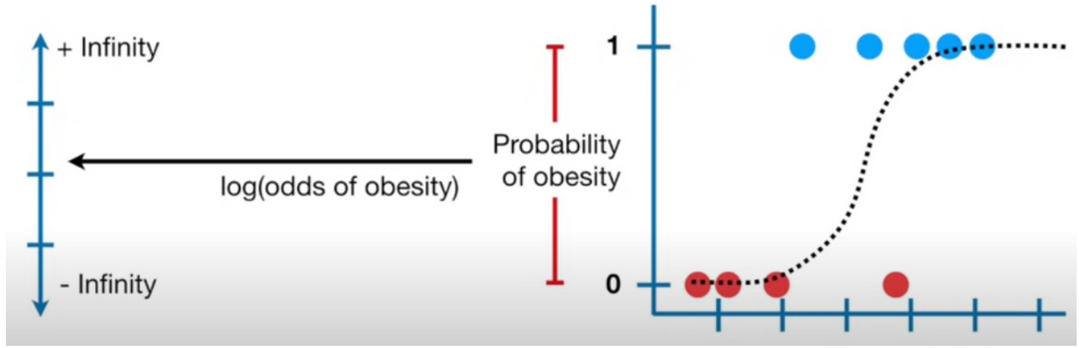


Linear Regression



Logistic Regression

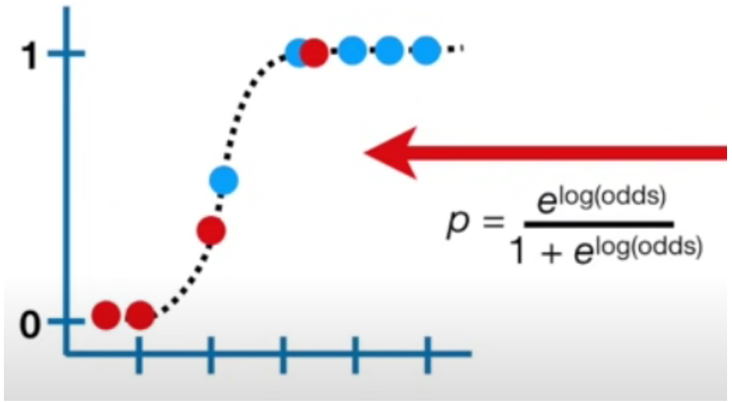
- How do we obtain the Logistic Regression Model from the given data?

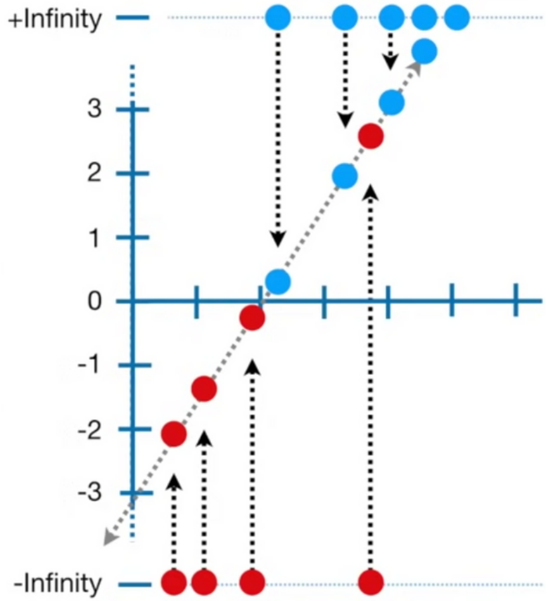


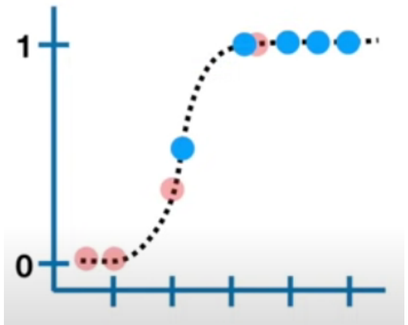
$$\text{logit}(p) = \log(\frac{p}{1-p})$$

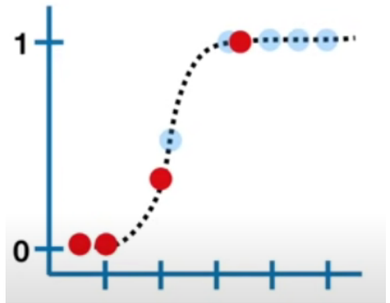
Maximum Likelihood

+ After performing the stated operations, the values tend to positive and negative infinity. So we project the data onto a candidate line.

+ 



+ 

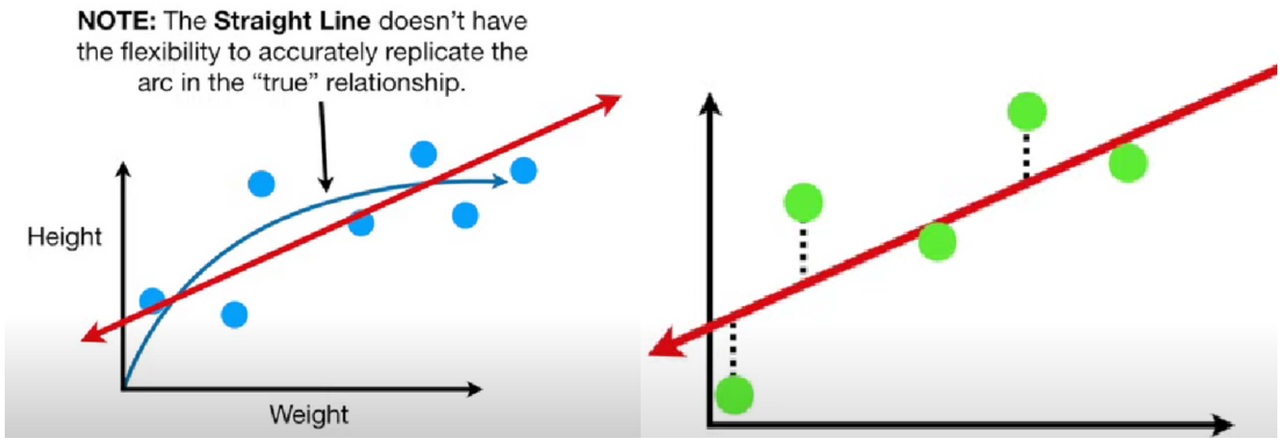


likelihood of data given the squiggle = 0.49 × 0.9 × 0.91 × 0.91 × 0.92 ×
(1 - 0.9) × (1 - 0.3) × (1 - 0.01) × (1 - 0.01)

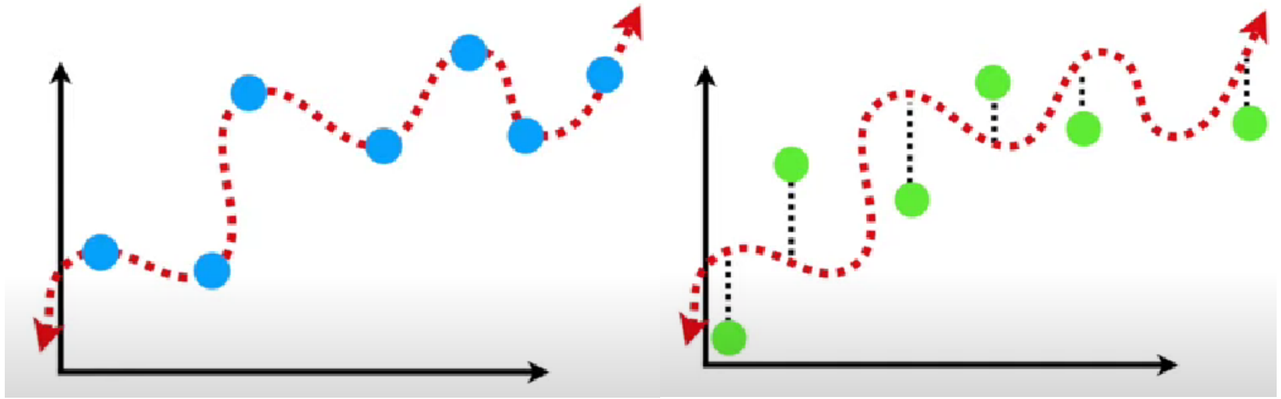
+ We repeat the process taking different candidate lines until we get the value of log(likelihood) = 0 and that equation is the best fit.

Bias-Variance Tradeoff

- Bias–Variance tradeoff is the property of a model that the variance of the parameter estimated across samples can be reduced by increasing the bias in the estimated parameters.



High Bias, Low Variance
(fits the training data better)



Low Bias, High Variance
(fits the testing data better)

KNN Algorithm

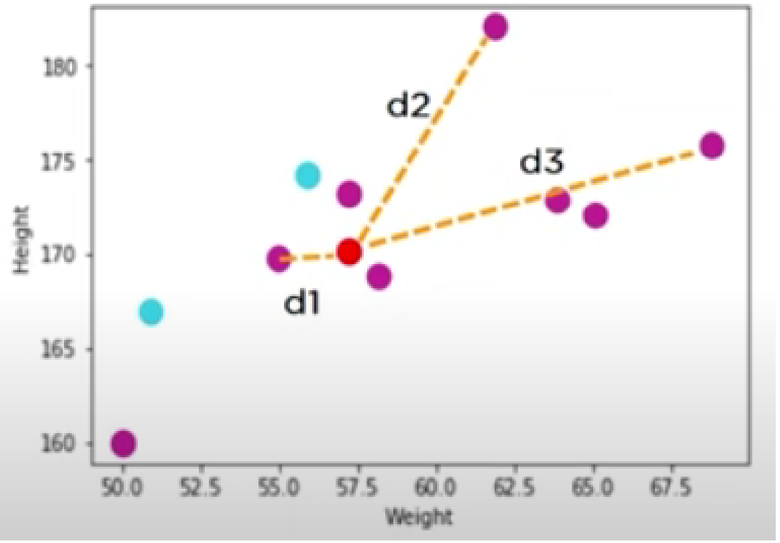
- K Nearest Neighbors is a Machine Learning Algorithm that classifies a data point on how it's neighbors are classified
- KNN is also known as the lazy-learners algorithm since it stores all available data and classifies new cases based on similarity measure
- k in KNN is a parameter that refers to the number of nearest neighbors to be considered
- We find the nearest neighbors using Euclidean distance and classify the data based on that.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Euclidean Distance

Working of KNN Algorithm

- Calculate Euclidean distance of the unknown data point from all the data point in scatter plot
-



Weight(x2)	Height(y2)	Class	Euclidean Distance
51	167	Underweight	6.7
62	182	Normal	13
69	176	Normal	13.4
64	173	Normal	7.6
65	172	Normal	8.2
56	174	Underweight	4.1
58	169	Normal	1.4
57	173	Normal	3
55	170	Normal	2

Euclidean distance of unknown data point from other data points

- Depending upon the value of k, we classify the unknown data point
- Choosing a value of k is done through a process called parameter tuning
- Generally, Value of k= sqrt(n); n = total no of data points or k value is estimated using trial-and-error

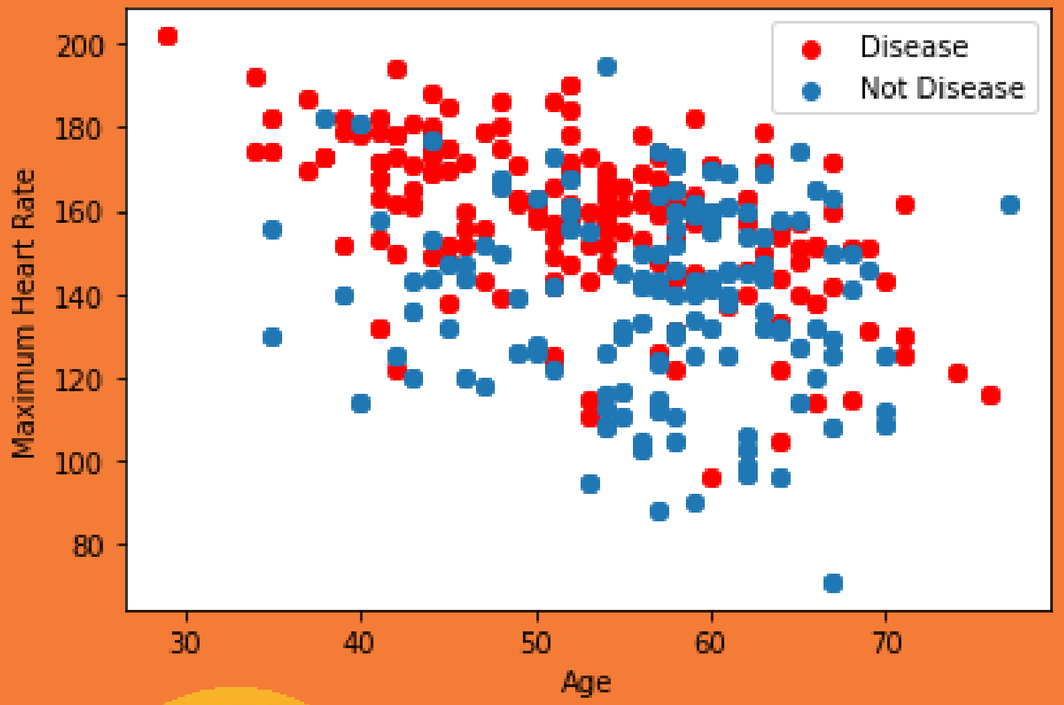
Modules & Explanation

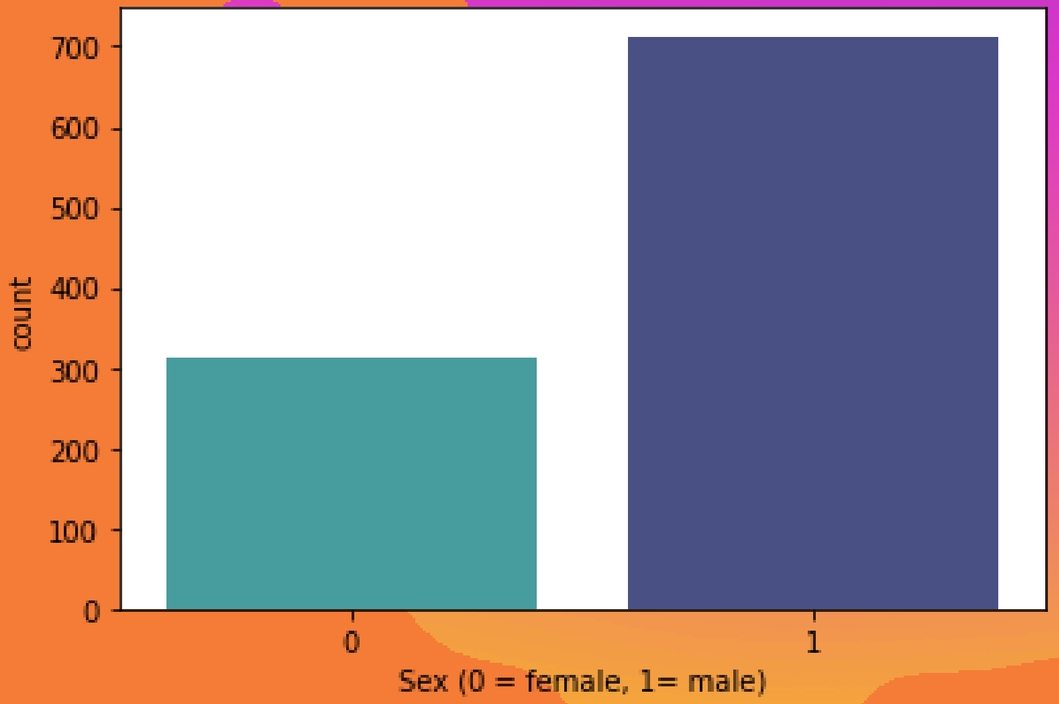
+ The Modules include a variety of designs that we have implemented in our system to predict disease using machine learning.

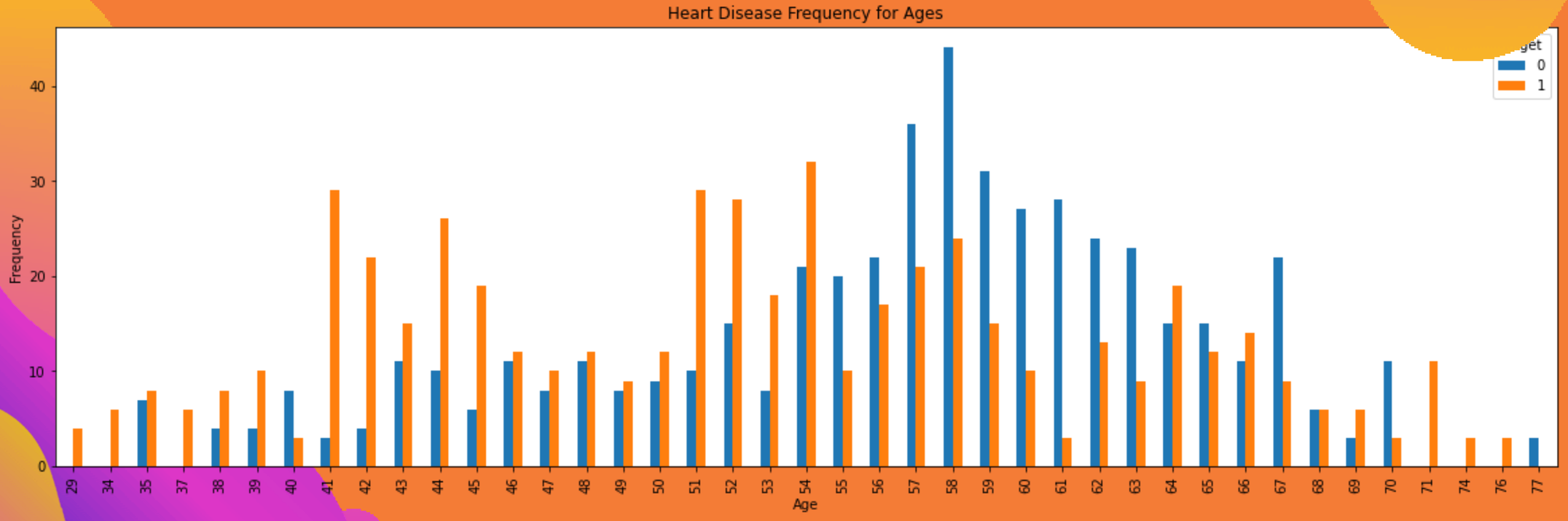
+ Here are some of the things that this system can do.
  1. Disease Prediction
  2. Entering Symptoms/Information

+ Entering Symptoms: Once user successfully open the system then he/she has to enter the necessary inputs.
  Disease prediction: The predictive model predicts the disease of a person and gives output to the user in maybe tabular form that have disease, not disease, depends on the symptoms entered by the user.

VISUALIZATION


Disease
Not Disease
Maximum Heart Rate
Age
200
180
160
140
120
100
80
30
40
50
60
70


count
Sex (0 = female, 1= male)
700
600
500
400
300
200
100
0
0
1


Heart Disease Frequency for Ages
Frequency
Age
0
1
40
30
20
10
0
29
34
35
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
74
76
77

MACHINE LEARNING MODEL

FRONT END & INTEGRATION

Webpage Design (Input)

Heart Disease Predictor

A Machine Learning Web Application that predicits chances of having heart disease or not, with an accuracy of 96.58%.

Age:	Enter your age
Sex:	Select
Chest Pain Type:	Select
Resting Blood Pressure:	Enter a number in range (94-200) mmHg
Serum Cholesterol:	Enter a number in range (126-564) mg/dl
Fasting Blood Sugar:	Select
Resting ECG Results:	Select
Max Heart Rate:	Enter a number in range (71-202) bpm

Webpage Design (Input)

Max Heart Rate: Enter a number in range (71-202) bpm

Exercise-induced Angina: Select

ST Depression: Typically in range (0.0 - 6.0)

Slope Of Peak Exercise ST Segment: Select

Number of Major vessels: Typically in range (0-4)

Thalassemia: Select

Predict

Webpage Design (Results)

Heart Disease Predictor

A Machine Learning Web Application that predicits chances of having heart disease or not, with an accuracy of 96.58%.

Prediction: Great! You DON'T have chances of Heart Disease.

User Input - Entering Data

- Patient’s Symptoms and other information is entered in this webpage and the result depends on the symptoms entered by the user.

- In the webpage, user has to enter personal details such as Chest Pain Type, Resting Blood Pressure, Serum Cholesterol, Fasting Blood Sugar, Resting ECG Results, Heart Rate and other such inputs.

- These inputs are processed onto the ML model using Flask (Python Framework) and a prediction is made which is then displayed on the website.

Creating the Flask Application

- Flask is a generic web framework that can be used to build almost anything using HTML and CSS.
- Step 1: Creating an HTML web page for the application
- As a first step in creating a web application using Flask, we should first design an HTML webpage.
- Step 2: Creating the Flask application using python
- The next step in this process is to create a python file (app.py).

Deploying the Flask application on Heroku

- Step 1: Download Heroku CLI and Git on your computer.
- Step 2: Creating a requirement file
- Step 3: Creating a Procfile
- Step 4: Open browser, log in to Heroku, click on the create new app option.
- Step 5: Now, we can deploy the application on Heroku

Conclusion

+ The developed Machine Learning model provided a platform to solve a large problem of administering heart diseases. This Machine Learning model made use of available Python based libraries and mathematical algorithms.
  For linking the front-end part of our application with the model, we made use of the python-based framework Flask.

+ During the development process we found that KNN algorithm worked better and provided higher accuracy rate compared to Logistic Regression (as observed in figure 3&4)and hence was used in this model. The value of K in our model is 2 and thus, the classification is done based on the two nearest neighbours of a new data.

Novelty and Future scope

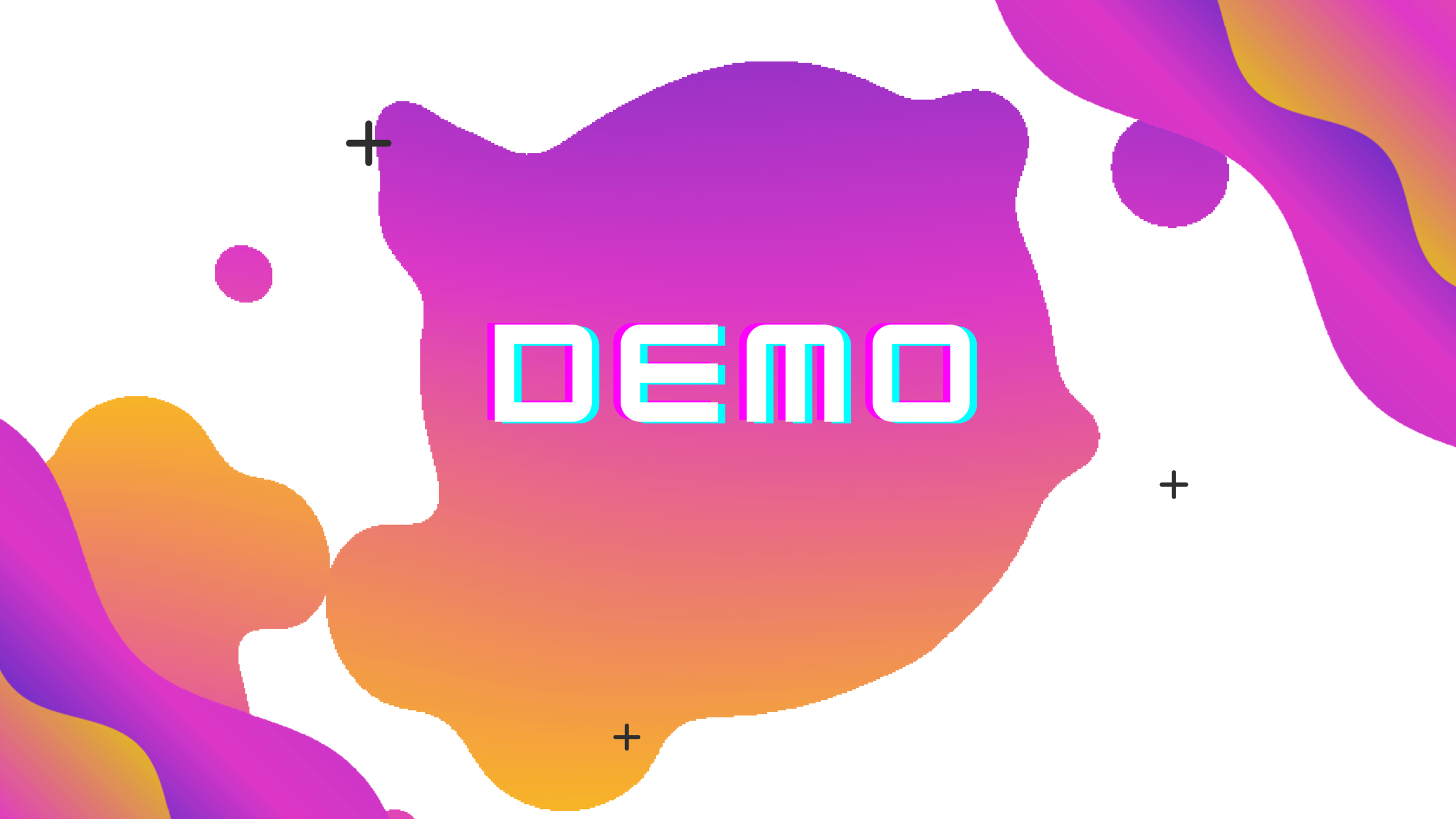
+ The model address the scarcity of vascular specialists in poor countries, as well as the fact that expensive medical operations are out of reach for most people.

+ As heart disease detection is a very intricate It acts as a platfrom for speacialists to get a second person opinion.

+ With our project being dynamic in nature, the scope for future enhancements is vast . (Using NLP we can develop a more user interactive website, User login/signup, Prediction of preventive measures, etc)

References

- Santhana Krishnan.J, Geetha.S, “Prediction of Heart Disease Using Machine Learning Algorithms”, IEEE 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019.
- https://www.researchgate.net/profile/Nikhil-Kumar-Mutyala/publication/324162326_Prediction_of_Heart_Diseases_Using_Data_Mining_and_Machine_Learning_Algorithms_and_Tools/links/5ac251eba6fdcccda65f7753/Prediction-of-Heart-Diseases-Using-Data-Mining-and-Machine-Learning-Algorithms-and-Tools.pdf
- https://jpinfotech.org/prediction-of-heart-disease-using-machine-learning-algorithms/
- https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0244-x
- https://www.sciencedirect.com/science/article/abs/pii/S1386505614001907


DEMO

Thank you!