## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**Case Study**: - Heart Disease Prediction Using Logistic Regression

https://drive.google.com/file/d/14WFdqCxXsoiXkPBnYuQHQ6IQ4y2\_6aOp/view?usp=drive\_link

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajiv Bhardwaj Taranpreet Kaur

2210990908

G – 13(b)

**Department of Computer Science and Engineering**

## Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

**Description about Case Study: -**

* Read dataset Insurance
* Display Top 5 rows of the dataset
* Display the Last 5 rows of the dataset
* Information of Dataset
* Check Null Values in Dataset
* Data Visualization with correlation matrix, pairplot, countplot
* Machine Learning- logistic regression
* Model Training
* Testing the data
* Predicting the Data
* Plotting Confusion Matrix
* Compare Performance Visually

**Library: -**

* Pandas
* Numpy
* Matplotlib.pyplot
* seaborn

**Methods: -**

1. **read\_csv():**

Description: Reads a CSV file and converts it into a data frame.

1. **dropna()**:

Description: removes missing values (None, NaN, or NaT) from a pandas DataFrame.

1. **fillna():**

Description: fills missing values (None, NaN, or NaT) in a pandas DataFrame with a specified value or a method of imputation.

1. **figure():**

Description: used in matplotlib library to create a new figure object.

1. **heatmap():**

Description: creates a heatmap visualization of a matrix

1. **show():**

Description: This will display the plot of the given data as a line chart.

1. **pairplot():**

Description: The pairplot() function in seaborn library creates a grid of scatter plots for pairwise relationships between variables in a dataset.

1. **countplot():**

Description: type of bar plot in seaborn library that shows count of occurrences of each unique value in a variable on the x-axis.

1. **legend():**

Description: The **legend()** method in matplotlib library adds a legend to the current plot, which shows a label for each unique value in the data, making it easier to understand the plot.

1. **head():**

Description: Displays the first few rows of the data frame.

1. **shape():**

Description: Returns the shape (number of rows, number of columns) of the data frame.

1. **info():**

Description: Provides basic information about the data frame, such as column types and missing values.

1. **isnull():**

Description: Returns True/False for each value in the data frame, indicating whether the value is missing (NaN) or not.

1. **subplot():**

Description: The **subplot()** function is a part of the Matplotlib library, which is used to create a grid of subplots within a single figure. It can take three integer arguments: the number of rows, the number of columns, and the index of the plot to be created in this scheme.

1. **title():**

Description: To add a title to a Seaborn plot

1. **fit():**

Description: The **fit()** method in scikit-learn library is used to find the optimal parameters of a given model by minimizing the cost function, for example, finding the best weights for a linear regression model.

1. **predict():**

Description: The **predict()** method in Seaborn is used to generate predictions based on the relationship learned from a previous regression plot

1. **plot():**

Description: The **plot()** method in matplotlib library is a function used to create a line plot, scatter plot, or bar chart.