## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**Case Study**: - Health Insurance- Linear Regression

**URL** - [https://drive.google.com/file/d/17SmRk7yHW2CMx68pQsxi5EcqoBH0WXlR/view?usp=sharing](https://drive.google.com/file/d/17SmRk7yHW2CMx68pQsxi5EcqoBH0WXlR/view?usp=sharing" \t "https://www.smartwebsolutions.org/text-to-url/_blank)A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Bharadwaj Tanishka Kautish

Roll Number: -2210990894

Group - 13

**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
* Shape of Dataset
* Information of Dataset
* Check Null Values in Dataset
* Get Overall Statistics of Dataset
* Covert Columns From String ['sex' ,'smoker','region' ] To Numerical Values
* Store Feature Matrix In X and Response(Target) In Vector y
* Train/Test split
* Import the models
* Model Training
* Prediction on Test Data
* Compare Performance Visually
* Predict Charges for New Customer

**Library: -**

* Pandas
* Sk.learn

**Methods: -**

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

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

1. **test\_size**:

Description: The proportion of the dataset to include in the test split

1. **random\_state:**

Description: To ensure reproducibility

1. **train\_test\_split:**

Description: function for splitting arrays or matrices into random train and test subsets

1. **.fit():**

Description: Used for training the model on training data.

1. **.score():**

Description: Used to evaluate the performance of the model on the dataset

1. **.predict():**

Description: Used for making predictions on new data using the trained model

1. **describe():**

Description: Used to generate statistics about the dataset

1. **replace():**

Description: Used to replace the data in the dataset

1. **tail():**

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

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.