**HEALTH INSURANCE COST PREDICTION**

**Project Synopsis**

**INTRODUCTION**

**Problem Statement:**

Building a machine learning model which helps the health insurance companies in predicting premiums for their customers based on certain important factors.

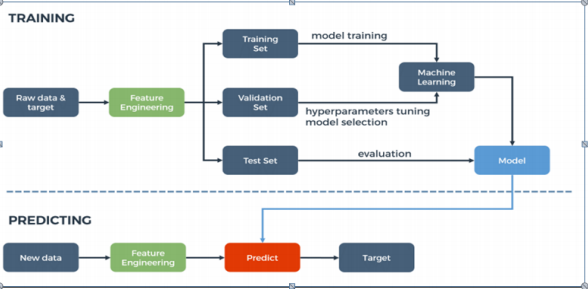
* **Overview**:

The aim of the project is to build a machine learning model which helps the health insurance companies in predicting the premium health insurance offers for their customers by following the health laws in the country.

* **Existing Problem:** Health insurance companies face a problem in determining the premium insurances for their customers. They need a trained model which predicts what kind of premiums are suitable to a customer. The companies should also follow certain rules set by the health care law in that country.
* **Proposed Solution:** The main motto of the project is to provide the health insurance companies with a machine learning model using the machine learning algorithms and python programming that can predict the customer's eligibility for different premium insurance services based on certain features and the model is tested based on the accuracy and performance of the model.

**THEORETICAL ANALYSIS**

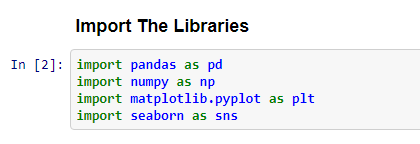
* **Flow chart:**

****

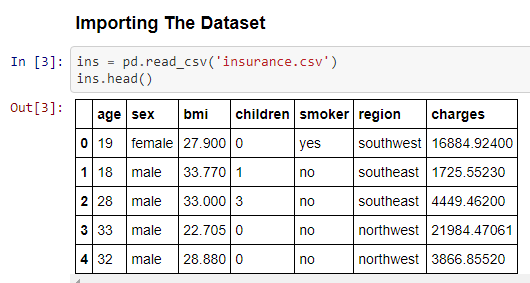
**EXPERIMENTAL INVESTIGATIONS**

1. Collect the data set. For this I have downloaded the dataset from Kaggle.
2. Before building the machine learning model the data collected should be preprocessed. In the data preprocessing the following steps are to be followed:

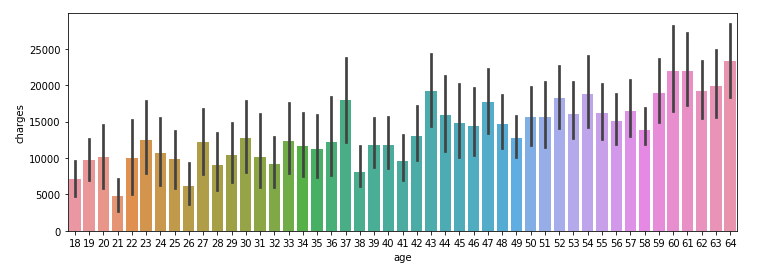
* Importing the required libraries.

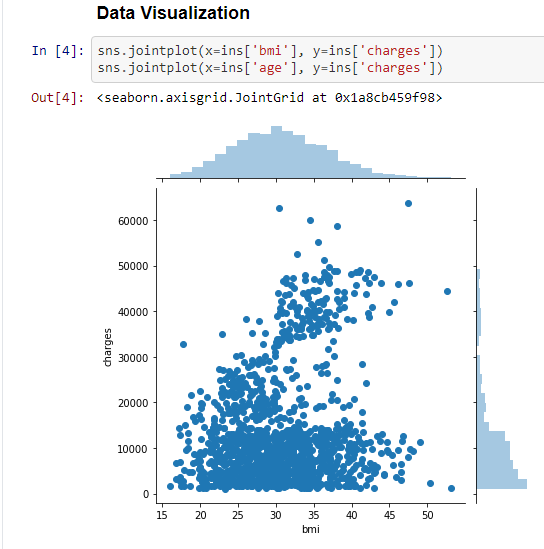


* Importing the dataset.

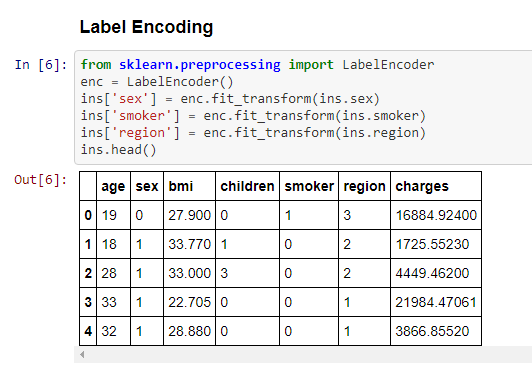


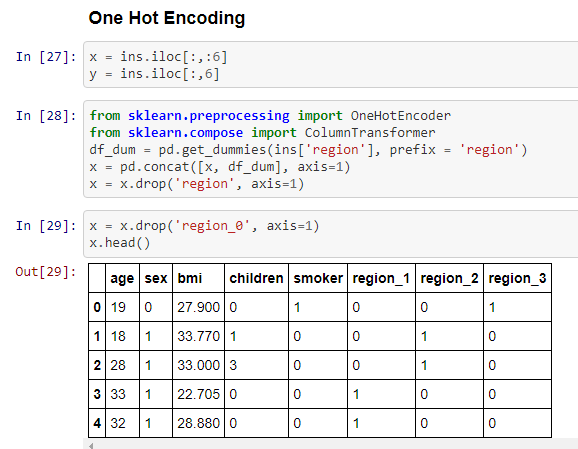
* Visualizing the data. By this we can understand the columns in the dataset and also know about the missing data and null values.



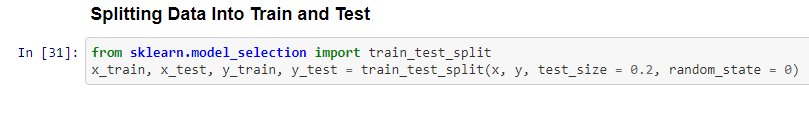


* After this we should take care about the missing data.
* Then we should do label encoding followed by one hot encoding for converting textual data.

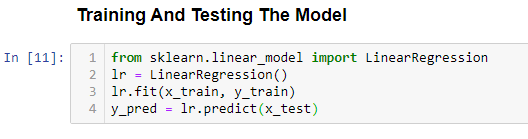


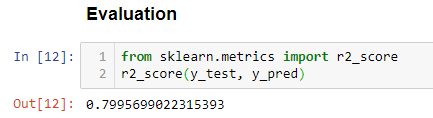


* Then feature scaling of the data should be done.
* Then we split the data into training data and test data



1. After data preprocessing, we need to choose the best machine learning algorithm that can be used for building the model.

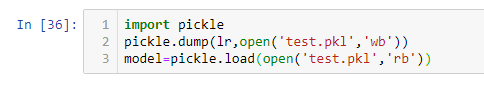




1. After building the prediction model we need to build the application. This involves the following steps:

* Creating an HTML page along with CSS design for the model for linking the script to flask backend code for creating a web API for using the model.



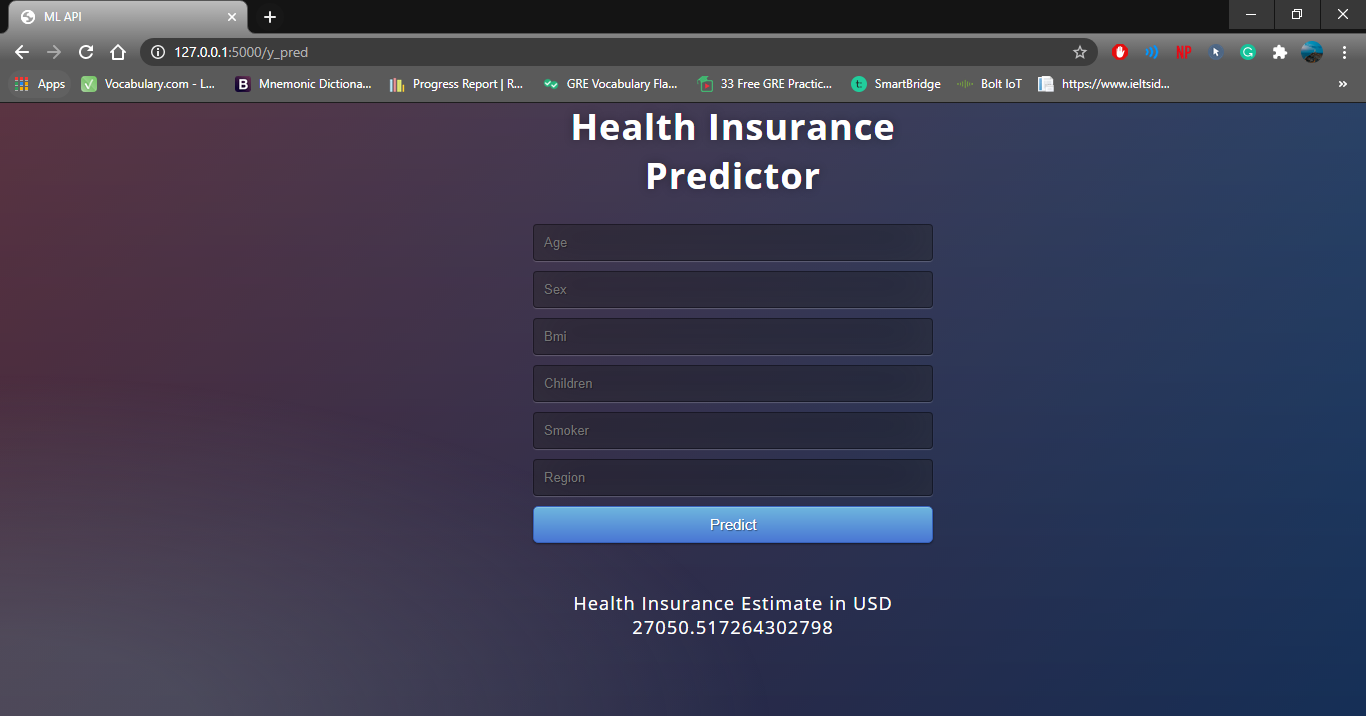


* Then we build the required python code for Flask API linking the HTML file and the model.

1. After this the model is to be deployed.
2. Then we test the model against the test dataset and check the accuracy of the prediction model.
3. Once the model is deployed it can be seen on the user interface.

**RESULT**

The machine learning model that works on the regression algorithm predicts the cost of the health insurance package that can be given to a customer based on the details of the customers depending upon the various given input features of the dataset.



**APPLICATIONS**

This model predicts the various health insurance premiums applicable to a given customer and can be used by the health insurance companies in order to improvise their services and can also bring new changes to their policies based on the statistics.

**CONCLUSION**

This was a great experience with Smart bridge learning new and interesting things and also applying them in real time. Coming to my project I can say that the machine learning model that is created to predict the cost of health insurance has a wide range of applications and makes the work of health insurance companies simpler. This gives all the predictions just by giving basic details of the customer without actually going through all the manual calculations.