**Insurance Premium Prediction-By Rishil.B and G.SaiTeja**

**ABSTRACT:**

The goal of this project to give people an estimate of how much they need based on their individual health situation. After that, customers can work with any health insurance carrier and its plans and perks while we keeping the projected cost from our study in mind. This can assist a person in concentrating on the health side of an insurance policy rather than the ineffective part.

**Dataset :**

The dataset is taken from a Kaggle. You can download the dataset from [here](https://www.kaggle.com/noordeen/insurance-premium-prediction)

**Approach :**

Applying machine learning tasks like Data Exploration, Data Cleaning, Feature Engineering, Model Building and model testing to build a solution that should able to predict the premium of the personal for health insurance.

* **Data Exploration :** Exploring the dataset using pandas, numpy, matplotlib, plotly and seaborn.
* **Exploratory Data Analysis :** Plotted different graphs to get more insights about dependent and independent variables/features.
* **Feature Engineering :** Numerical features scaled down and Categorical features encoded.
* **Model Building :** In this step, first dataset Splitting is done. After that model is trained on different Machine Learning Algorithms such as:
  + Linear Regression
  + Decision Tree Regressor
  + Random Forest Regressor
  + Gradient Boosting Regressor
  + XGBoost Regressor
  + KNN
* **Model Selection :** Tested all the models to check the RMSE & R-squared.
* **Pickle File** : Selected model as per best RMSE score & R-squared and created pickle file using pickle library.
* **Webpage &Deployment :** Created a web application that takes all the necessary inputs from the user & shows the output. Then deployed project on the Heroku Platform.