# CAPSTONE PROJECT SUBMISSION

## Health Insurance Cross Sell Prediction

An insurance policy is an arrangement by which a company undertakes to provide a guarantee of compensation for specified loss, damage, illness, or death in return for the payment of a specified premium. A premium is a sum of money that the customer needs to pay regularly to an insurance company for this guarantee.

Just like medical insurance, there is vehicle insurance where every year customer needs to pay a premium of a certain amount to the insurance provider company so that in case of an unfortunate accident by the vehicle, the insurance provider company will provide compensation (called 'sum assured') to the customer.

This rapid increase in automobiles has caused automobile insurance to emerging as an essential business target for insurance companies. Therefore, if car insurance sales are predicted and sold using the information of existing health insurance customers, it can generate continuous profits in the insurance company's operating performance. Therefore, this study aims to analyze existing customer characteristics and implement a predictive model to activate advertisements for customers interested in such auto insurance. The goal of this study is to maximize the profits of insurance companies by devising communication strategies that can optimize business models and profits for customers.

Cross-Selling is a new marketing strategy based on data analysis, which found that different needs exist as well, who can become customers and meet their needs through sales of various related services or products.

 As the first step, performing data-wrangling over the raw data further, we divided the complete project data observation and cleaning, exploratory data analysis, checking multicollinearity, and data modeling. In the first step, we collected the raw data from Almabetter. Then we did a basic data cleaning check and removed zero NaN values because there were no null values or missing values in the dataset.

 the exploratory data analysis we briefly study on variables after that we found our dependent and independent variables. After separating the  variables we did some statistical analysis with variables and did some feature engineering, the correlation has been checked with the help of heatmap, distribution plots, and scatterplots and there is a very high correlation among all variables. So we check multicollinearity with . Then we drop some features to prevent wrong predictions.

We are utilizing the  Logistic Regression, Random Forest, and Xgb classifier for building different prediction models.
It was possible to derive a predictive model with more accurate results. It is expected that building a car insurance purchase prediction model based on the car insurance buyer's personal purchase information can be used in the future when an insurance company provides personalized advertising. Suppose an insurance company collects various pieces of personal information from customers who typically use insurance other than car insurance or health insurance. In this case,

you can predict and analyze insurance in related fields by predicting the interests of your customers based on the data. In this case, even if the number of customers in the insurance company does not increase significantly, it can be expected to have a positive impact, guaranteeing much higher profitability based on its customers.

Additionally, being able to make more confident predictions with a single data point is expected to have the benefit of increased data usage. In this way, in the future, we are currently building a machine learning model that can predict, recommend and suggest insurance products in areas where customers feel their usual need besides car insurance by collecting and using various information besides the buyer's personal insurance purchase.

## Contribution Roles:

i). **Sunanda Debnath**
Performed Data Preprocessing
General Analysis  Perform EDA(Exploratory Data Analysis)
Build Model Such As
Logistic Regression

ii). **Ajay Tiwari**
Feature engineering
Build Model Such As
Random Forest
XGB classifier

| Please paste the GitHub Repo link. |
| --- |
| GitHub Link:            https://github.com/AjayUnique/Health-Insurance-Cross-Sell-Prediction |

**Drive Link:**
https://drive.google.com/drive/folders/1iSTGTkcx0IqBAQpF5fbhly-C0UcxDQww