JOBATHON JANUARY 2023

Problem Statement

VahanBima is one of the leading insurance companies in India. It provides motor vehicle insurances at best prices with 24/7 claim settlement. It offers different types of policies for both personal and commercial vehicles. It has established its brand across different regions in India. Around 90% of the businesses today use personalized services. The company wants to launch different personalized experience programs for customers of VahanBima. The personalized experience can be dedicated resources for claim settlement, different kinds of services at doorstep, etc.

In order to do so, they would like to segment the customers into different tiers based on their customer lifetime value (CLTV).

In order to do it, they would like to predict the customer lifetime value based on the activity and interaction of the customer with the platform. So, as a part of this challenge, your task at hand is to build a high performance and interpretable machine learning model to predict the CLTV based on the user and policy data. Approach

• After importing data, I did some EDA (check null values, checked unique values etc)

• Following that I made baselines of two models Linear Regression , Gradientboost regression and checked which model is performing best.

• I made observation that Gradientboost regression is performing better than Linear Regression.

• Gradient boost regression is used when large number of environment variables compared to number of obeservations