

# AUTO INSURANCE - CUSTOMER VALUE ANALYSIS

## INTRODUCTION

Auto Insurance landscape has been feeling effects of safe car technology with growing popularity of accident-prevention features such as traffic jam assistance, lane departure warnings that partially remove the human element from driving. According to a report by KPMG; within 25 years, the private passenger automobile insurance industry will shrink by as much as by 60 percent.

**Customer Lifetime Value:** In marketing, customer lifetime value (CLV) is a metric that represents the total net profit a company makes from any given customer. CLV is a projection to estimate a customer's monetary worth to a business after factoring in the value of the relationship with a customer over time.

## BUSINESS CASE

We are building a model to predict market behavior, such as if clients would respond yes to our sales using different classification models. Looking for new sales opportunities from legacy data and type of renew policy that they will accept, etc. We would also try and predict the customer lifetime value analysis based on different factors available in data.

## REGRESSION MODELLING

Models: Linear Regression, Random Forest Regression

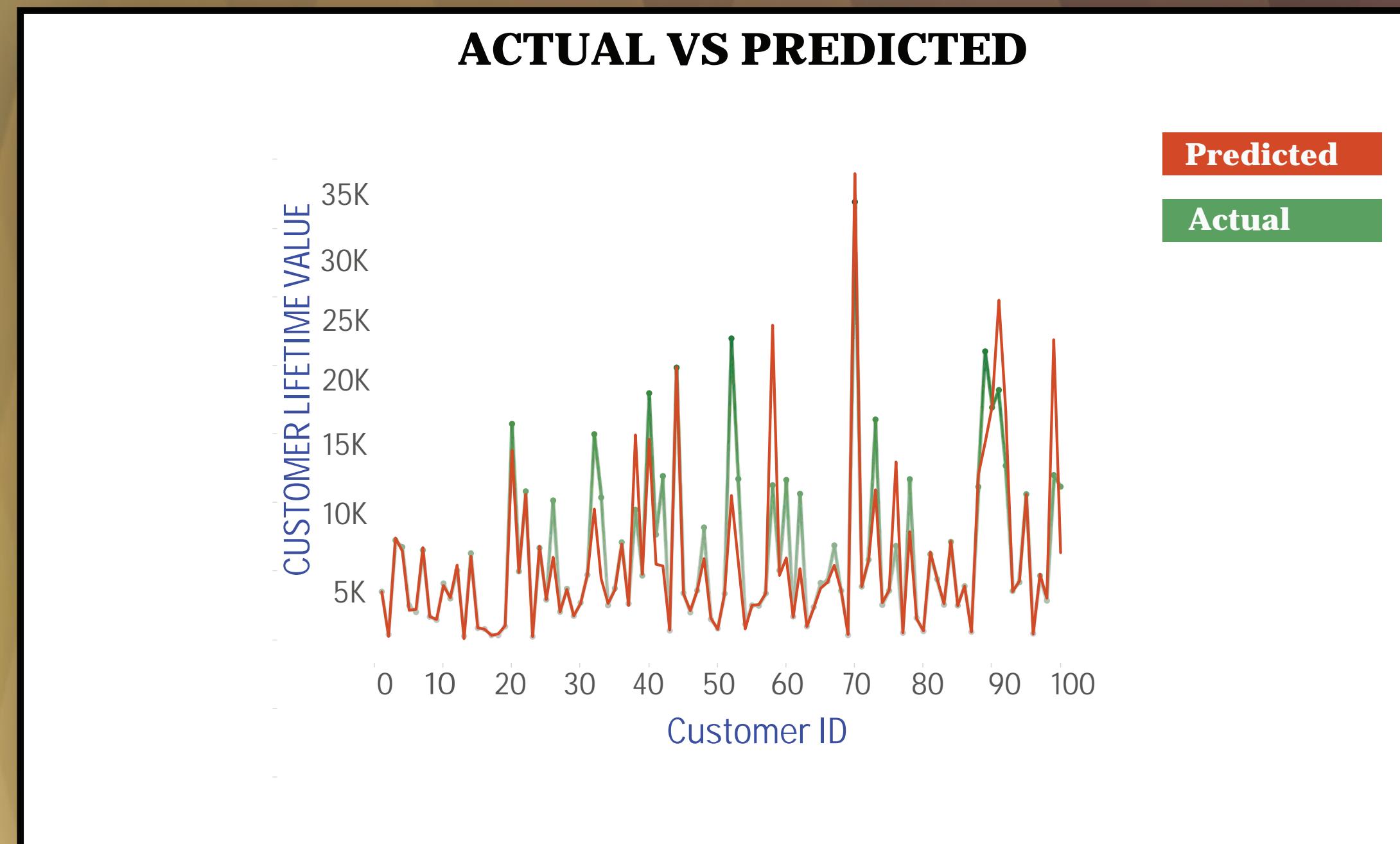
Packages used: Pandas, Numpy, Matplotlib

Language used: Python

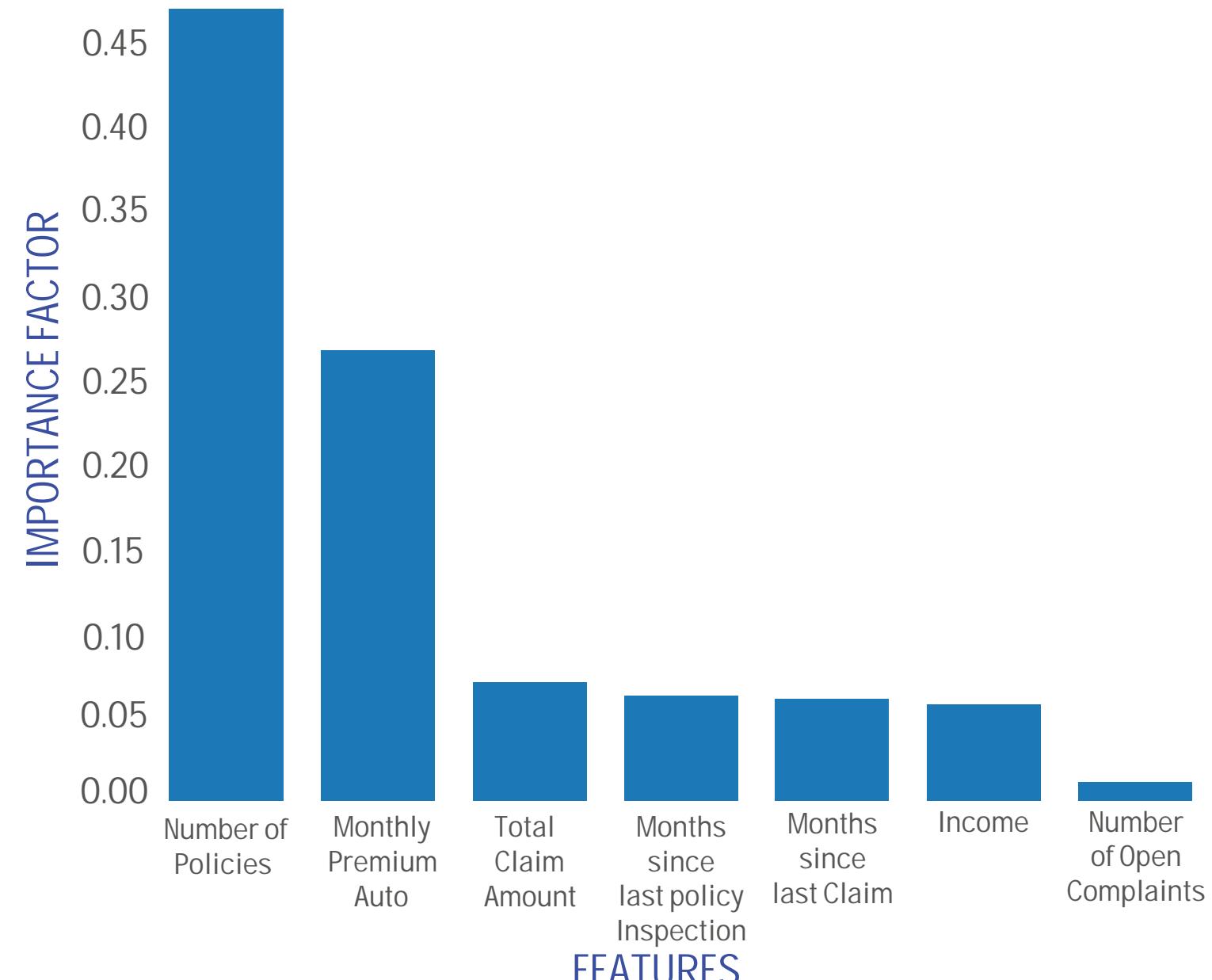
MSE: Linear Regression > Random Forest Regression

We applied linear regression and random forest regression models to predict Customer Lifetime value. After training both the models, better results were found using random forest using test data.

Actual vs Predicted customer lifetime values has been depicted using the line chart plotted.



## FEATURES VS IMPORTANCE FACTOR

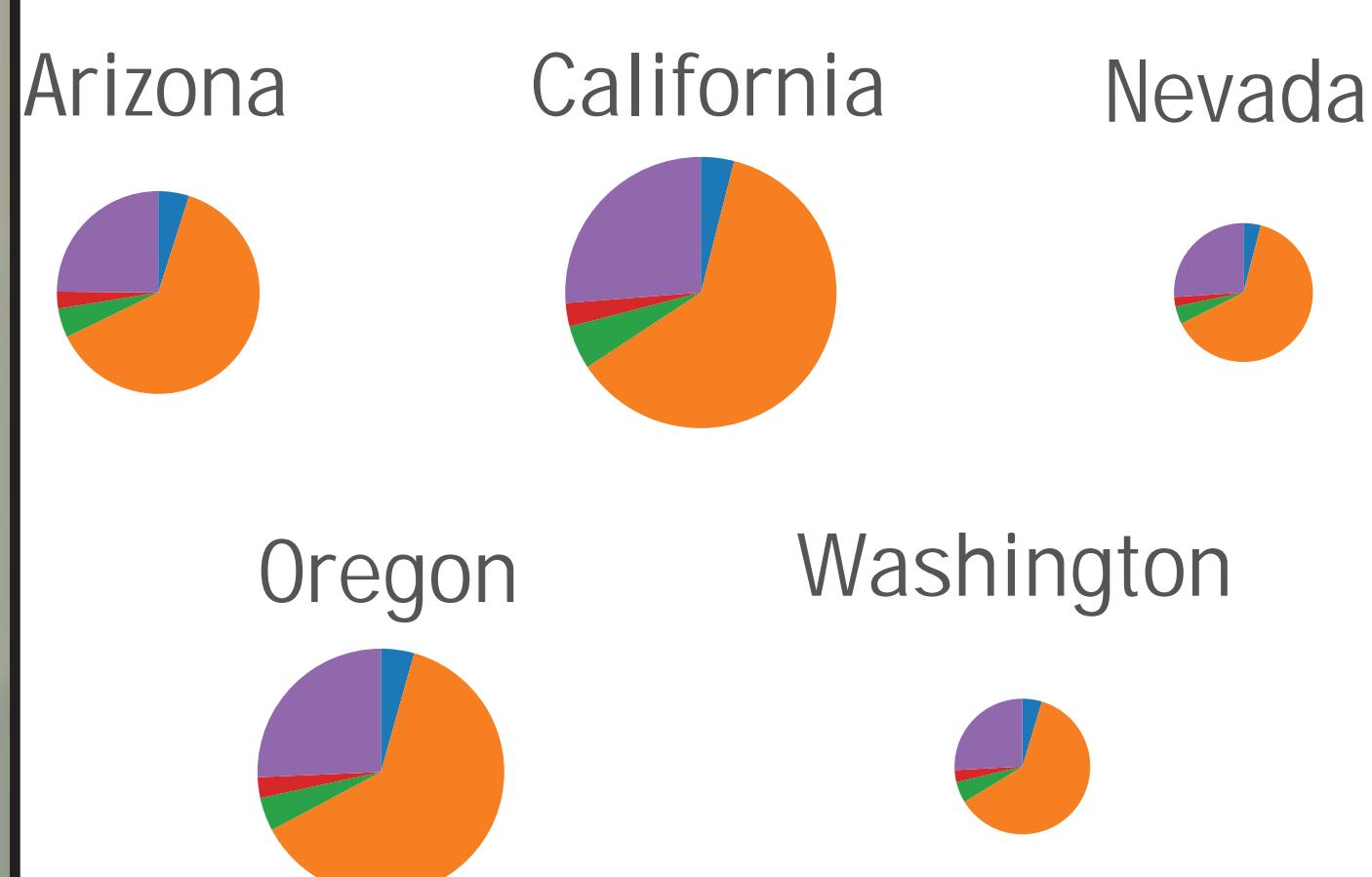


## IMPORTANCE FACTOR

Random forest model gives an option of calculating feature importance. The feature importance helps identify which features are helpful to predict the outcome. In our case we had 7 features that we used in our testing model and tried to identify how much importance each feature has. Below is the outcome of our feature importance:

Feature	Importance
Number of Policies	0.470281
Monthly Premium Auto	0.267772
Total Claim Amount	0.070736
Months Since Last Policy Inspection	0.062404
Months Since Last Claim	0.060539
Income	0.05727
Number of Open Complaints	0.010998

## STATEWISE NUMBER OF POLICIES FOR DIFFERENT EMPLOYMENT STATUS



## RANDOM FOREST CLASSIFICATION EVALUATION

