A Life Insurance Policy Bundling Recommendation System

PURDUE UNIVERSITY | Krannert School of Management

Mengwei Li, Shashi Pingolia, Tianyi Yang, Matthew A. Lanham

Purdue University, Krannert School of Management

li4003@purdue.edu; shashi@purdue.edu; yang1193@purdue.edu; lanhamm@purdue.edu

ABSTRACT

In this study we have researched and developed a life insurance bundling recommendation system that identifies among current home or auto insurance policy holders who is most likely to add a life insurance policy product to their existing plans as well as when to optimally recommended a life product to the customer. The implementation process of matching customers to the right products is not widely known and likely could be improved using analytical frameworks found in other domains. Our goal is to provide insurance companies a more efficient, analytically driven, and scalable approach to sell additional products that are most suitable to their customers and also increase the business revenue.

INTRODUCTION

Insurance bundling is a win-win trade practice for both Insurance companies and for Customers. For customers, it saves money and efforts as insurance companies offers multipolicy discount as a means to save on overall insurance bills. For insurance companies, it's a source of additional revenue from the same customer. Research has also indicated that customers who avails the benefits of insurance bundling stay more loyal to their insurance companies.

Research questions:

1. What is the most suitable model to use when building the recommendation system?
2. What business impact and revenue does this recommendation system bring?

EXPLORATORY DATA ANALYSIS

Auto and home policies have completely different attributes, so we trained on two different datasets, one with auto policies related information and one with home policies related information.

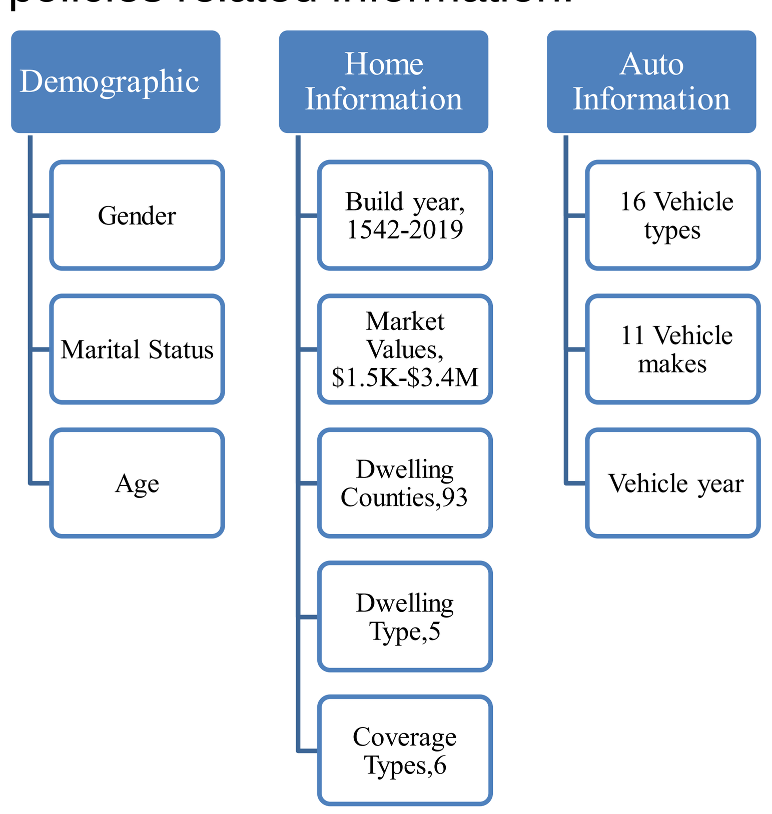

Fig 1. Data Structure

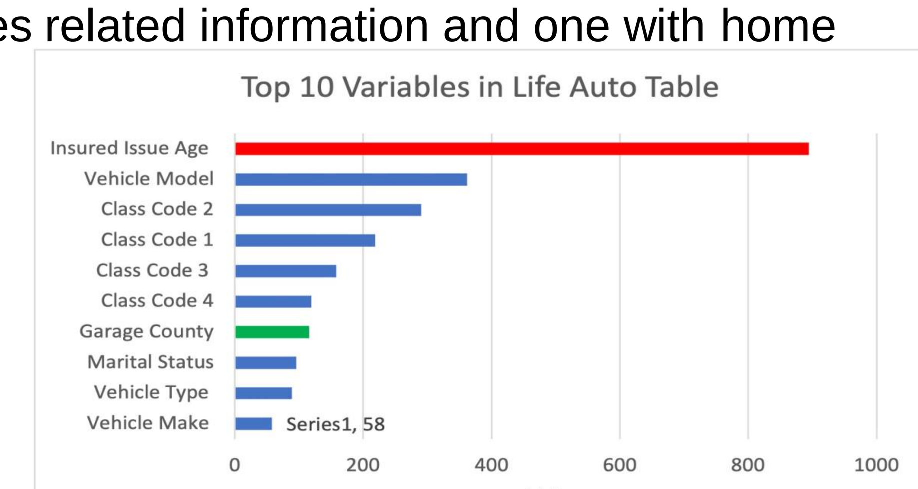

Fig 2. Top Variable-Auto

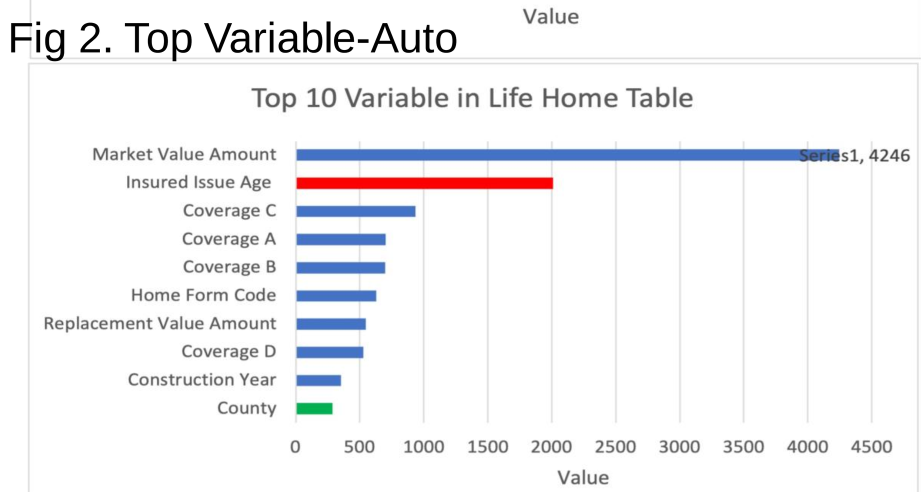

Fig 3. Top Variable-Home

LITERATURE REVIEW

There are not many documented instance of recommender system for insurance domain.

Our research contribution: Provide an empirically justified methodological design that recommends who is likely to want to bundle life insurance with their existing auto or home policies, what product is best for them, and when to offer the bundle.

METHODOLOGY

We're trying to build a recommendation system that first identifies our potential customers. Next is to provide them with the correct insurance type, and finally identify the best time for them to purchase the policy.

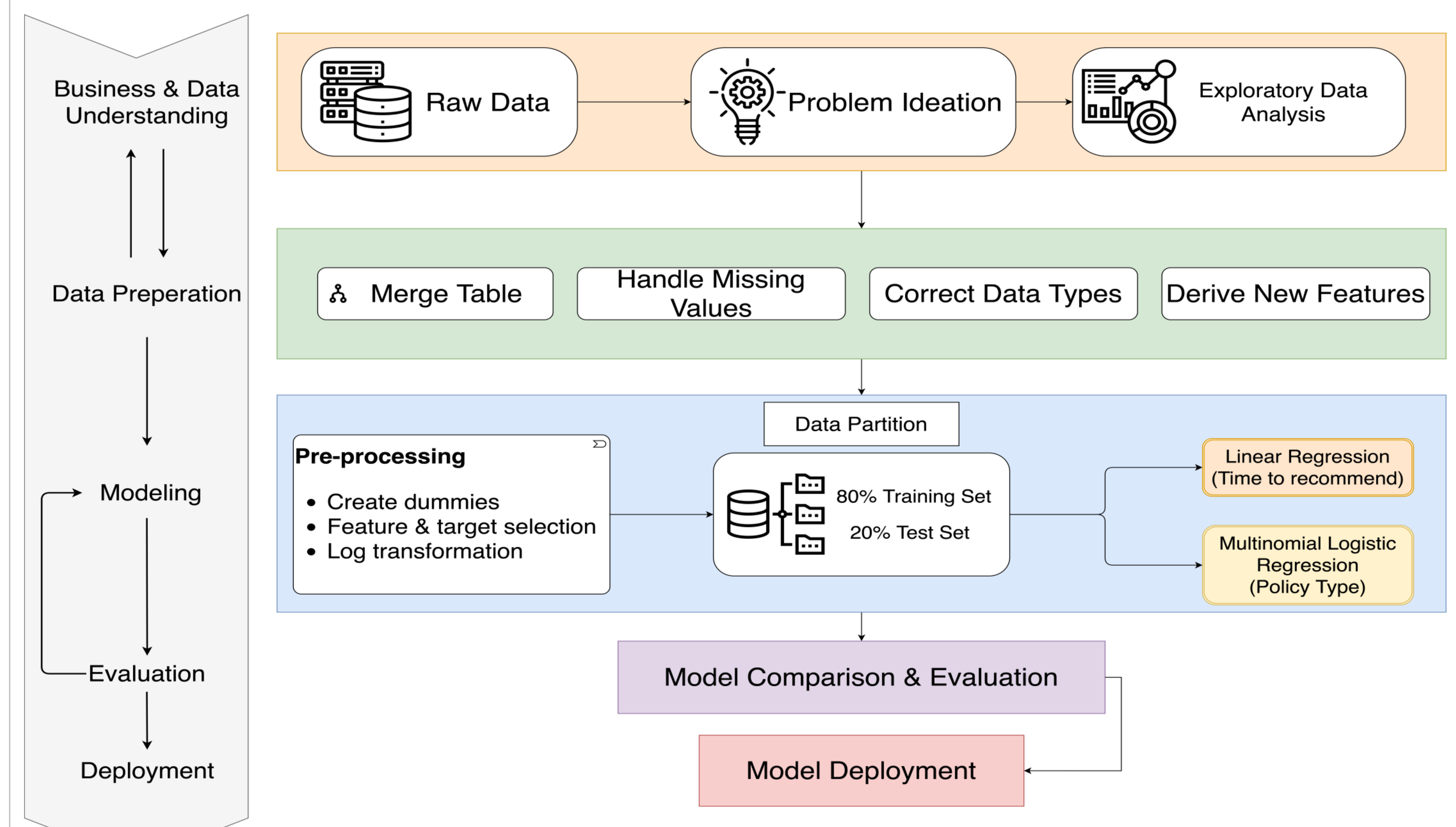

Fig 4. Methodology Process Workflow

Models Used

- We used random forest and neural network for the numeric prediction on timespan and random forest for the multi-classification prediction on policy type.
- Other supervised learning models such as Neural Network and Random Forest are also used. These models were not as effective so we stick with the linear and logistic regression given that the coefficients are more interpretable thus can aid HR decision-making.

Tools

- Caret library in R and Scikit Learn package in Python to train and identify the best candidate model.

STATISTICAL RESULTS

1. Confusion Matrices : Prediction evaluation on the Life Policy Type based on random forest model with balanced class-weight

	Precision Rate	Recall rate
Term	74.42%	94.62%
Whole	57.64%	21.09%
Single	16.67%	3.31%

Table 1. Precision and Recall rates

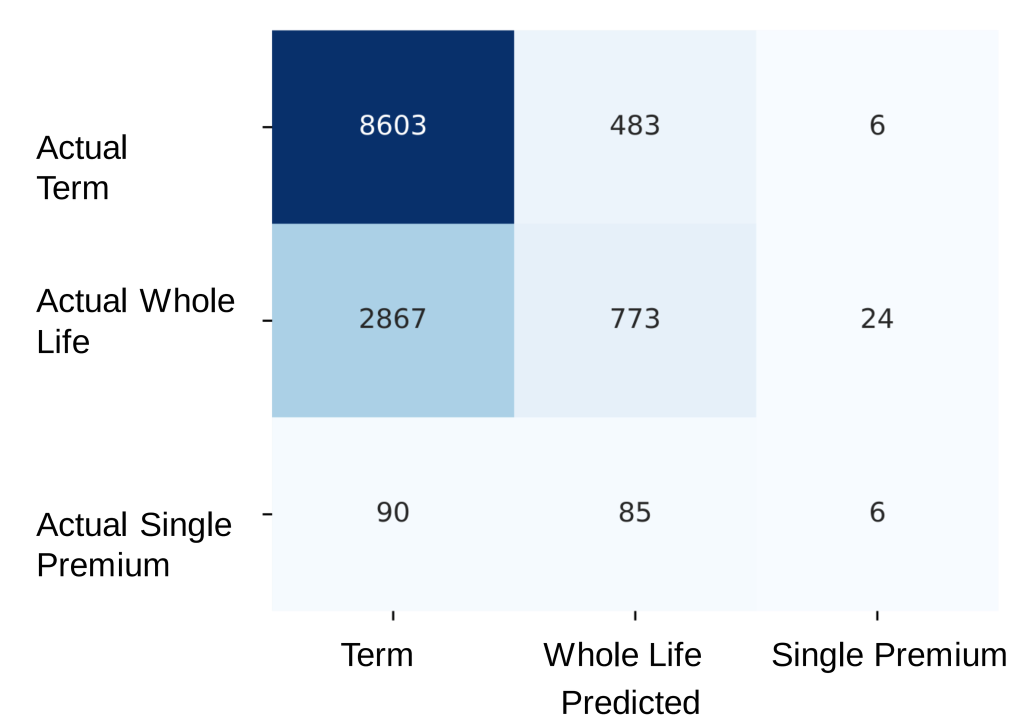

Fig 5. Life policy type prediction based on Home and Auto Insurance

2. Calibration Plots-Random Forest

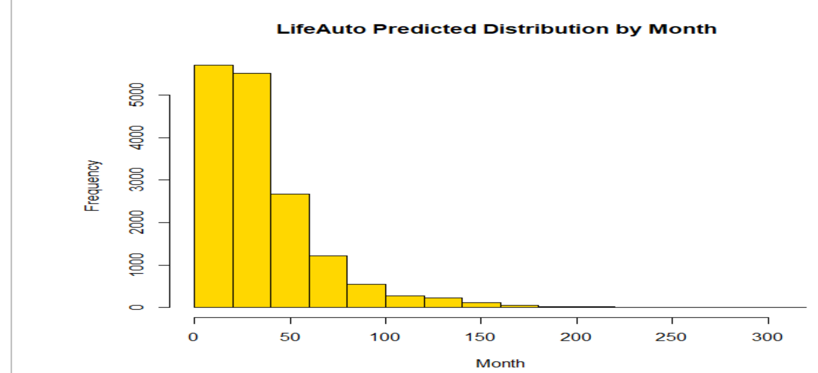

Fig 6. Life-Auto Predicted Distribution

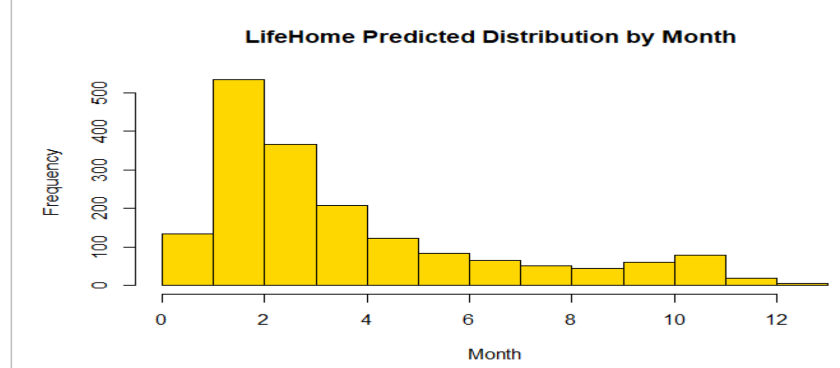

Fig 7. Life-Home Predicted Distribution

3. Calibration Plots-Neural Network

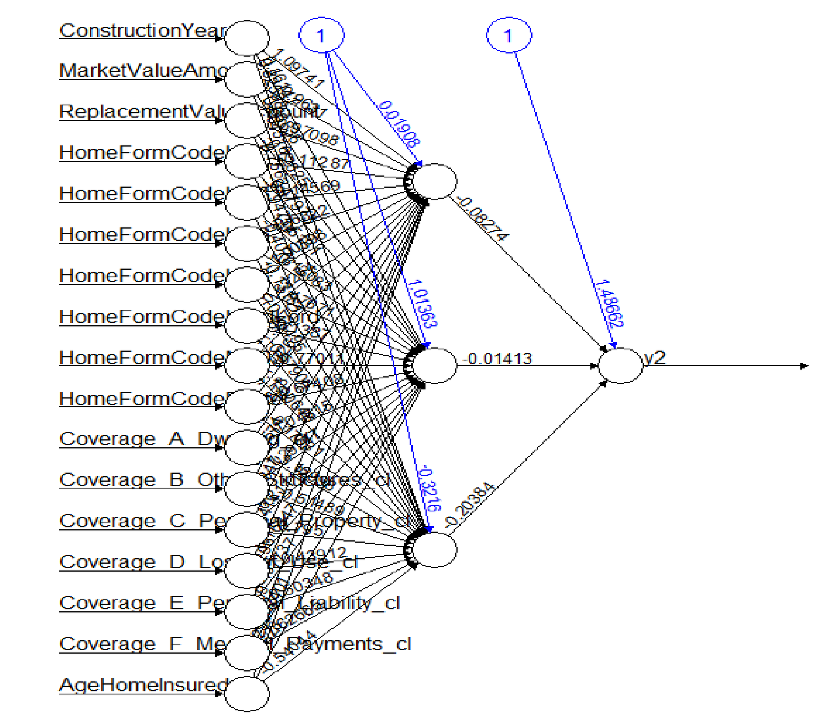

Fig 8. Neural Network

EXPECTED BUSINESS IMPACT

Our solution has provided the insurance company a more efficient, analytically driven, and scalable approach to sell additional products that their customers really want and increase their business revenue. We believe our methodology connects the recommendation system literature to the insurance industry and can be easily adapted by practitioners in this field.

It Is expected that our model will help our client in:

1. Developing business strategy to better target customers for life insurance products.
2. Identify the best time to recommend a life policy to existing P&C policy holders.

CONCLUSIONS

1. As per our analysis, Random Forest and Neural Network performed better than other models and had least average squared errors.
2. Our model can identify the correct life insurance type with an overall accuracy of around 70% on all three policy types.

In addition, our research also identifies that:

- For existing auto insurance customer, the best time to recommend life insurance are 16 to 55 months later after they purchased auto insurance.
- For existing home insurance customer, the best time to recommend life insurance are 1 to 5 months later after they purchased home insurance.
- The differences between groups also have critical influence on the willingness to purchase.

ACKNOWLEDGEMENTS

We would like to thank our faculty and our data scientists from the Insurance firm, for their guidance and support on this project.