Predicting Risk for New Insurance Using Customer Information

Problem Statement

Risk management is important for insurance industry to identify new authenticated customers. Insurers consider every available quantifiable factor to develop profiles of high and low insurance risk. Level of risk will be able to determine the expected insurance premiums. Generally, insurance policies involving factors with greater risk of claims are charged at a higher rate. With much information at hand, insurers can evaluate risk of insurance policies at much higher accuracy.

For this insurer collect a vast amount of information about policy holders and insured objects. Statistical methods and tools based on data mining techniques can be used to analyze or to determine insurance policy risk levels. In this project, an Insurance risk predictive modeling approach is discuss for predicting risk for new insurance using customer information. This approach is about RME-EP (Rule-based Model Evaluation with Event Processing) which is an amalgam of predictive modeling and (forward chaining) rule engine. It uses Deep Learning approach for improving the accuracy of system. RME-EP can be used to combine a number of predictive models into a single model, producing combined predictions such as maximum, minimum, average, etc. In addition, it can be used to classify combined predictions into classes such as "Very high risk", "High risk", "Medium risk", "Low risk", etc.

Background

Instead of treating all customers equally, managers have come to understand that it is more effective to develop customer-specific strategies. As a result, companies are now adopting customer relationship management (CRM). CRM means that companies manage relationships with individual customers with the aid of customer databases and interactive mass customization technologies. The adoption of CRM has been enhanced by recent developments in Information and Communication Technology. By using customer information contained in databases, companies can invest in the customers that are potentially valuable for the company, but also minimize their investments in non-valuable customers. Researches in this genre do not target at analyzing particular insurance types but discussing issues of risk classification based on dichotomy instead studying cases of both perfect information and information asymmetry and of both competitive and non-competitive markets.

Methodology

Step 1: Data collection and dataset preparation

This will involve collection of customer information containing dataset that is real world dataset so this raw dataset needs preprocessing and feature extraction by noise reduction, feature extraction and data normalization so that this preprocessed data can be used on the proposed model.

Step 2: Developing A RME-EP (Rule-Based Model Evaluation with Event Processing) model for Predicting Risk for New Insurance Using Customer Information

In this step a Deep Learning based approach i.e., RME-EP (Rule-based Model Evaluation with Event Processing) which is an amalgam of predictive modeling and (forward chaining) rule engine is developed for predicting risk for new insurance using customer information.

Step 3: Training and experimentation on datasets

The RME-EP (Rule-based Model Evaluation with Event Processing) predictive model that is Deep Learning based approach will be trained on the dataset to do the predicting risk for new insurance using customer information and notify the insurance company about the risk factors of customer by dividing it into classes such as "Very high risk", "High risk", "Medium risk", "Low risk".

Step 4: Deployment and analysis on real life scenario

INPUT- Customer Information Data	→	**Data Preprocessing**- Feature Extraction	→	**Machine learning Implementation**- RME-EP	→	**Training and Testing**- Categorizing Customers	→	**OUTPUT**- Predicting Risk

Figure 1: The proposed system architecture of Predicting Risk for New Insurance Using Customer Information
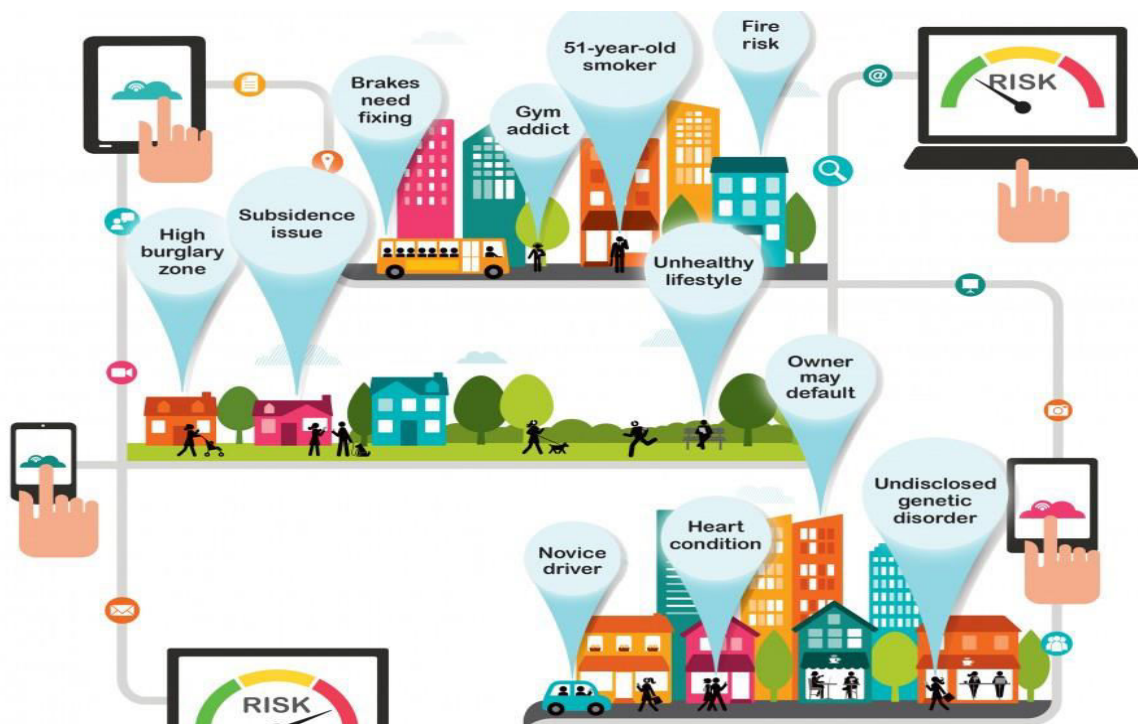
Figure 2: Explaining how Insurance companies analyses risks with customers[1]

---

[1] https://www.fnlondon.com/articles/insurance-industry-gets-digital-20160505.

The trained and tested Predicting Risk for New Insurance Using Customer Information model will be deployed in a real-life scenario made by the human experts & will be leveraged for further improvement in the methodology and will follow the above architecture. It will also follow the proposed architecture shown and its working is explained with the help of diagram.

Experimental Design

Dataset

Insurance Company Benchmark (COIL 2000)[2] data Set is used which contains information on customers of an insurance company. The data consists of 86 variables and includes product usage data and socio-demographic data, which has different type of insurance dataset will be used for experimentation and evaluation**.**

Evaluation Measures

Measures such as Accuracy, risk factor and mean standard deviation will be measured and evaluation is done for predicting risk for new insurance using customer information**.**

Software and Hardware Requirements

Python based Deep Learning libraries will be exploited for the development and experimentation of the project. Tools such as Anaconda Python and libraries such as Tensorflow will be utilized for this process. Training will be conducted on NVIDIA GPUs for training the proposed system for predicting risk for new insurance using customer information.

---

[2] https://archive.ics.uci.edu/ml/datasets/Insurance+Company+Benchmark+(COIL+2000)