

# Machine Learning and Exploratory Data Analysis in Cross-Sell Insurance



**Anand Jha**

*Rustamji Institute of Technology, BSF Academy, Tekanpur, India*

**Brajkishore Prajapati**

*Rustamji Institute of Technology, BSF Academy, Tekanpur, India*

## INTRODUCTION

A financial shield against a risk or threat is insurance. A business corporation that sells such financial shield and an entity that buys such financial shield are called insurer and insured respectively. A written agreement between insurer and the insured defines this financial shield and called *policy document* or *insurance policy*. But this financial shield is given to insured in lieu of monetary payment called premium (Aswani et al., 2020).

The insurance policies can be of several kinds according to the different kinds of risks. Some popular types of insurance are Life Insurance, Health Insurance, Vehicle Insurance, Travel Insurance and Property Insurance (Bacry et al., 2020). This basic idea of risk or monetary risk reduction is the reason to run entire industry (Batra et al., 2021).

Insurance industry access to volume of data generated with the advent of technologies has tremendously increased (Accenture, 2018). It can be said that more data have been created, collected in the past couple of years than the human society has ever produced due to irresistible explosion in data from a host of sources, like telematics, Internet usage, social media bustle, voice analytics, connected sensors and wearable devices.

Big data can be unstructured, semi-structured and structured (Chakraborty & Kar, 2017; Kar, 2016). It is well known fact that *actuarial formulas* are still in use by several insurance businesses, but now data science and analytics can be utilized to excerpt hidden information that can help for improved strategic as well as administrative decision making than the conventional techniques. (Chowdhury et al., 2020; Karhade et al., 2019). Use of data technologies can have a serious role in different facets of the insurance business like risk assessment, claim analysis, underwriting analysis, customer profiling and fraud detection etc. (Das et al., 2020; Das et al., 2021).

Together with various data analysis and visualization techniques, Machine Learning (ML) as a sub-domain of artificial intelligence (AI) come as a rescue to Insurance Industry to facilitate such datasets. Machine learning teaches computers to think in a similar way as humans, thus learning and improving upon past experiences. Practically almost all tasks that can be accomplished with a data-defined pattern or set of rules can be automated with machine learning (Accenture, 2018). Machine learning (ML) techniques can be effectively used across Structured, Semi Structured or Unstructured datasets (Burri et al., 2019), (Chakraborty & Kar, 2017; Kar, 2016). But most insurers are struggling to maximise the benefits of machine learning and thus not able to unearth analytical insights hidden into datasets. Since machine learning is in use during last few decades, it is not a novel technology. *Supervised learning*, *Unsupervised Learning* and *Reinforcement Learning* are three core classes of machine learning. Majority

DOI: 10.4018/978-1-7998-9220-5.ch039

of insurers are working with supervised learning for risk assessment by means of identified parameters to obtain preferred outcome. But during last few decades, unsupervised learning is also gaining popularity among present age insurers.

Thus we can say that state of affairs in insurance industry is witnessing a slow but steady change due to an environment characterised by increased competition, elastic marketplaces, complex claims, risk appetite and premium leakage, expense administration, subrogation, new business, new product development, policy servicing, distribution fraud behaviour, sales, customer experience and tighter regulation. This changing scenario also forcing insurers to explore new ways to use predictive modelling and machine learning to retain their competitive edge, boost business operations and enhance customer satisfaction. Most insurers process and analyse a small percentage of data they have access to. They process mostly structured data since it is obtainable from their traditional databases. This worsens to unravel worth not only from their structured data, but also ignoring the valuable insights hidden in their unstructured data. Emergent data analytical techniques can reap benefits from this unstructured data by effectively mining it to get business insights.

This chapter discusses relevance of machine learning and application of data visualization techniques for making better strategic decisions in insurance industry, especially for cross-sell insurance through a dataset.

## **BACKGROUND**

Insurance companies have been gradual to embrace technological change compared to other industries because insurance is among one of the old and extensively regulated industry in the world with marine insurance was the first kind of insurance (Rawat et al., 2021). Insurance Industry is still soaked in paper-based processes that are slow, manual and require human intervention. Even in this digital era, customers are asked to go for a time-consuming paperwork when getting a claim reimbursed or signing up for a new insurance policy. In this process customers may be paying more for insurance policies since policies are not custom-made according to their exclusive needs. In a digital era, this may end up with an unhappy customer experience.

One of the buzzword of digital decade is Artificial Intelligence and this means the use of machines to pretend human intelligence (Kose et al., 2015; Kraus et al., 2020). AI or ML can help in practically all domains of insurance business beginning from pricing strategies (Larson & Sinclair, 2021; Maehashi & Shintani, 2020), risk selection, personalization of policies to document underwriting etc. (McGlade & Scott-Hayward, 2019; Mita et al., 2021).

Some of the potential use cases (Burri et al., 2019; Accenture, 2018; Shroff R., 2019) of machine learning in health insurance are:

**Insurance Advice:** Machine learning can take a noteworthy role in customer service, from the initial interaction to helping a customer requires to select most relevant and cost effective health insurance policy. Recent surveys show that customers are happy to receive computer-generated insurance advice because they are looking for personalised solutions as made promising by machine learning algorithms that review their profiles and recommend tailor-made insurance policies. Other than using machine learning in back end, insurers are also using artificial intelligence by using chat-bots on messaging apps to resolve claims queries and answer simple questions. (Khan et al., 2014; Knighton et al., 2020) focused on the great scope of betterment in conventionally used methods in insurance business for policy advice and enrolment.

33 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/machine-learning-and-exploratory-data-analysis-in-cross-sell-insurance/317477](http://www.igi-global.com/chapter/machine-learning-and-exploratory-data-analysis-in-cross-sell-insurance/317477)

## Related Content

---

### MCDM-Based Optimization of Performance Characteristics During $\mu$ EDMing of SS 304

Premangshu Mukhopadhyay, Goutam Kumar Bose and Pritam Pain (2021). *Machine Learning Applications in Non-Conventional Machining Processes* (pp. 18-32).

[www.irma-international.org/chapter/mcdm-based-optimization-of-performance-characteristics-during-edming-of-ss-304/271495](http://www.irma-international.org/chapter/mcdm-based-optimization-of-performance-characteristics-during-edming-of-ss-304/271495)

### An Integrated Process for Verifying Deep Learning Classifiers Using Dataset Dissimilarity Measures

Darryl Hond, Hamid Asgari, Daniel Jeffery and Mike Newman (2021). *International Journal of Artificial Intelligence and Machine Learning* (pp. 1-21).

[www.irma-international.org/article/an-integrated-process-for-verifying-deep-learning-classifiers-using-dataset-dissimilarity-measures/289536](http://www.irma-international.org/article/an-integrated-process-for-verifying-deep-learning-classifiers-using-dataset-dissimilarity-measures/289536)

### Cyber Security for Secured Smart Home Applications Using Internet of Things, Dark Web, and Blockchain Technology in the Future

Vinod Mahor, Sujit Kumar Badodia, Anil Kumar, Sadhna Bijrothiya and Ankit Temurnikar (2022). *Dark Web Pattern Recognition and Crime Analysis Using Machine Intelligence* (pp. 208-219).

[www.irma-international.org/chapter/cyber-security-for-secured-smart-home-applications-using-internet-of-things-dark-web-and-blockchain-technology-in-the-future/304212](http://www.irma-international.org/chapter/cyber-security-for-secured-smart-home-applications-using-internet-of-things-dark-web-and-blockchain-technology-in-the-future/304212)

### Ant Miner: A Hybrid Pittsburgh Style Classification Rule Mining Algorithm

Bijaya Kumar Nanda and Satchidananda Dehuri (2020). *International Journal of Artificial Intelligence and Machine Learning* (pp. 45-59).

[www.irma-international.org/article/ant-miner/249252](http://www.irma-international.org/article/ant-miner/249252)

### Analyzing U.S. Maritime Trade and COVID-19 Impact Using Machine Learning

Peter R. Abroades, James Rotella, Partha Mukherjee and Youakim Badr (2023). *Encyclopedia of Data Science and Machine Learning* (pp. 323-341).

[www.irma-international.org/chapter/analyzing-us-maritime-trade-and-covid-19-impact-using-machine-learning/317457](http://www.irma-international.org/chapter/analyzing-us-maritime-trade-and-covid-19-impact-using-machine-learning/317457)