Big Data Analytics

Homework 01

APPLICATION OF BIG DATA IN PROPERTY & CASUALTY INSURANCE

GUGHAPRIYAA ELANGO

SU ID : 585992381

HOME WORK - 1

Big Data Analytics

Homework 01

APPLICATION OF BIG DATA IN PROPERTY & CASUALTY INSURANCE

ABSTRACT

This study explores the impact of big data in the Property & Casualty (P&C) Insurance sector, emphasizing its important role in automating processes, enhancing customer experiences, and improving cost-efficiency. It highlights the potential of data mining applications and discusses the transformative impact of AI in predicting bodily injuries and detecting fraud. Additionally, the paper addresses the ethical dimensions of handling sensitive data, emphasizing the importance of data security, transparency, and responsible use. Overall, it sheds light on how big data is revolutionizing the insurance industry, by offering insights to remain competitive and ethically responsible at the same time.

INTRODUCTION

Having some data engineering experience in the Property & Casualty Insurance sector, I was interested in diving deeper into the application of big data in this niche. Automotive property damage bills or casualty health bills can no longer fit in spreadsheets. They are produced every minute with big data streaming in real time at a large scale. This brought a need to process this huge amount of data in an optimal way, thus entering the big data era.

There is major impact of big data in the insurance sector. Not just in the United States but all over the world, the revenue generated by the insurance industry is in billions of dollars (‘Leveraging Text Mining in Property and Casualty Insurance’, 2022). The industry is ideal for data mining or text mining applications where the business of P&C insurance has very diverse and complicated uses. They also contain a range of varieties of insurance like commercial or personal insurance for home, auto or other possessions where big data is generated and used. (Huang and Meng, 2019)

The insurance sector provides a promise to its customer to restitute an insured subject, be it in an event of a possible future covered loss. Accurate predictions are critical in this industry, creating a high demand for quantitative analysis in the P&C sector. With highly growing competitive markets, there is a need to identify and differentiate themselves from their competitors, control and optimize costs and allocated resources in increasingly smart and meaningful ways.

Operations of customers around their claims is the most important customer experience the insurance company must provide. This claim process needs to be provided in the most seamless way possible so that the company can ensure customer retention and satisfaction. There is a lot of data in the Petabytes which are generated around these insurance claims daily.

Big Data Analytics

Homework 01

APPLICATION OF BIG DATA

Big data in this industry can bring about many advantages in the insurance industry like claim management, faster recovery, risk assessments and customer satisfaction, to name a few. With good quality data, that is up to date and accurate, insurance companies can improve their services by automating many processes.

With insurance claims available in online applications to the users to file, insurance companies are increasingly leaning towards automated Straight-through-processing without any human interventions. Still an insurance adjuster's role is at large in many cases of anomalies to negotiate better with clients. Claim automation can be done up to a certain extent using NLP and Machine learning increasing the operational efficiency. Research shows that with the use of big data technologies companies save 40-70% of their costs, better access to insurance packages and 60% more efficient in identifying fraud. (Huang and Meng, 2019). They could also help save processing and operational costs.

The second important application of big data in Insurance sector is risk assessment and underwriting. With vast amount of information from various sources, insurers can gather and get predictive analytics that help assess risk(Koutsomitropoulos and Kalou, 2017). This is possible by identifying patterns, trends and correlations in the data that can be tailored towards the policies and pricing. With this precise and fairer premium rate can be achieved. Flagging unusual claim patterns or inconsistencies are very important to determine liabilities between the parties more accurately.

While big data streamlines the process of claim processing, we can also leverage it to enhance customer experiences by providing personal recommendations and services. They can help insurers better understand the policyholders' needs, preferences, and behaviors. Using Virtual Assistants using NLP or text mining, customers can get their inquiries promptly responded to.

Not just customer retention, customer segmentation can be more precisely done using big data and machine learning. Customers can be segmented into focus groups on various factors of their demographics, preferences and behaviors.(Ellili et al., 2023)

With AI and advanced car sensing technology, we can predict accurately the bodily injuries possible in the seated positions when two cars hit each other in any way. The physics and dynamics of the car collision can be used to predict the injuries people can sustain. This is particularly useful in fraud detection of swoop and squat cases. The perpetrators fit tightly well into a car and target a new luxury vehicle to stop suddenly in front to rear end the vehicle. This could cause the target vehicle to be liable for the property and casualty losses.

Tracking of telematics of the vehicles, we can more precisely predict driver behavior. This data of speed, braking pattern and locations can help determine premiums and assess risks better.(Ellili et al., 2023)

Big Data Analytics

Homework 01

In terms of Pricing and Product development, big data can help drive insights to develop innovative insurance products that cater to highly evolving customer trends and demands. With fine tuning algorithms, pricing strategies can be maintained with competitiveness while maintaining profitability. We can also use big data to identify the focus groups of people among the different insurance policies, which can be used as opportunity for better marketing and sales.

An important aspect of any insurance company is the regulatory requirements that need to be maintained for detailed records and for providing auditing trails. Big data and data analytics can help assess and manage the risks associated with any regulatory standards or changes that are brought about.

To summarize, big data in the P&C insurance sector can be leveraged to identify any high-impact opportunities, control and optimize cost and pricing, strategically place themselves apart from competition and work towards resource allocation in an efficient way.

DATA

The main information resource for insurers is the big data of Property and Casualty bills in the event of an automobile accident. The data can be highly complex in nature as they contain multiple information, sometimes in multiple formats. The data can vary in structure like Structured data in traditional databases like claims history or financial records, semi-structured data from sources like IoT devices or PDF documents, unstructured data like emails or customer feedback texts. (Koutsomitropoulos and Kalou, 2017)

The data can be collected from many different input ways, the most common being using an internal application to collect data from users like policy or claim data, financial records, or customer interactions. They can also be generated from external third-party providers like market data, credit scores or regulatory information.

Some of the claim notes can be highly unstructured with abbreviations and jargons, but they are incredibly valuable source of information. Documenting them digitally is a much needed task from insurance adjusters. The sheer scale of data from P&C insurance companies can vary based on the size of the company and its customer base. However, on average the data usually goes in scale of Petabytes.

In recent trends, traditional methods or even batch processing of data is not suitable for applications in the insurance industry. Real-time or Near-real-time processing of the data is required to get fresh and up to date data. The frequency of claims that are filed and processed has increased in interval, providing continuous updates. It is an industry that needs high data availability and make data-driven decisions.

Data collection, storage and analysis are important parts of data processing in the P&C industry. Different technologies have all become inevitably hand in hand for processing such huge amounts of data. Some platform technologies used commonly include data warehouse solutions like Snowflake or Amazon Redshift, Big data technologies like Hadoop, Spark are used for large

data handling, Apache Kafka, Flink are used for Change Data Capture (CDC) and ingestion, business intelligence tools like Tableau, PowerBI and QlikView are used, and cloud platforms like AWS and Azure are being more commonly used.

ETHICAL DIMENSIONS

This disruptive force of AI and big data has brought about changes to society so fast, the society's laws and regulations need some time to get updated. Faster analysis of large volumes of data also makes us vulnerable to easily have a data breach. Despite the benefits of big data, the industry also faces a range of ethical issues.

From harvesting PII or Personal Identifiable Information from millions of people, data companies are also vulnerable to data leakage(Natalija, 2019). Not just PII, but many sensitive information of medical bills is in line. Unfortunately, very less ethical practices are available to the insurance industry, and specific laws are not formed into policy yet.

To start with, informed consent from the user needs to be taken to collect their Personal Identifiable Information. The data being collected needs to transparently declared about its collection and the way it is going to be used. The data collected should be used responsibly with minimal usage and governance. For security reasons, authentication and private networks are very important to users who handle this sensitive information.

Apart from data security, the industry also faces bias and reliability due to any inaccurate, missing, corrupted or inconsistent data. Due to a lack of standard vocabulary, there can be a lot of heterogeneity in data (Natalija, 2019). Bad efficacy in training of the data can lead to skewed models and skewed results. This can also lead to any kind of discrimination and unfair use of the data.

Continuous use of automated systems without a regard to ethical issues can lead to a loss of trust from users on the system. There needs to be an additional initiative of responsibility from companies on this regard. There is an immediate need to have governance about the accuracy of the data and the recommendations it is making. The accessibility of data in the highly sensitive levels needs to be restricted to employees and contractors.

REFERENCES

Ellili, N. et al. (2023) 'The applications of big data in the insurance industry: A bibliometric and systematic review of relevant literature', The Journal of Finance and Data Science, 9, p. 100102. Available at: https://doi.org/10.1016/j.jfds.2023.100102.

Huang, Y. and Meng, S. (2019) 'Automobile insurance classification ratemaking based on telematics driving data', Decision Support Systems, 127, p. 113156. Available at: https://doi.org/10.1016/j.dss.2019.113156.

Big Data Analytics

Homework 01

Koutsomitropoulos, D.A. and Kalou, A.K. (2017) ‘A standards-based ontology and support for Big Data Analytics in the insurance industry’, ICT Express, 3(2), pp. 57–61. Available at: https://doi.org/10.1016/j.icte.2017.05.007.

‘Leveraging Text Mining in Property and Casualty Insurance’ (2012) in Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. Elsevier, pp. 967–982. Available at: https://doi.org/10.1016/B978-0-12-386979-1.00042-6.

Natalija, K. (2019) ‘Insurance, Smart Information Systems and Ethics’, The ORBIT Journal, 2(2), pp. 1–27. Available at: https://doi.org/10.29297/orbit.v2i2.106.