GlobalLogic®

Role of Big Data Analytics in Insurance

by Dr. Param Jeet, Specialist
and Er. Subodh Pandey, Consultant

Contents

Introduction

The insurance industry runs on data and the success of its business model is based on analyzing data to evaluate information and make appropriate decisions. As we know, this is the era of big data. Companies around the world have access today to more points of information than at any other time in human history. 90% of all the data in the world was created in the last two years. Every day, the planet creates more than 2.5 quintillion bytes of information.

Every interaction you perform with your computer or phone creates data, whether it's your phone tracking your location through its GPS sensors or you posting on social media, shopping with a debit card, or streaming a song, movie or podcast.

What is Big Data?

Data that is very large in size (GBs ,TBs, PBs) is called big data. As previously stated, almost 90% of today's data has been generated in the past 2 years. This data comes from many sources: social networking sites, e-commerce sites, weather stations, telecom companies, and share markets.

The 3 V's of Big Data

Velocity: Data is increasing at a very fast rate. It is estimated that the volume of data will double every 2 years.

Variety: Nowadays, data is not stored in rows and columns. Data is structured as well as unstructured. Log file and CCTV footage is unstructured data. Financial transactions are one example of structured data.

Volume: The very large amount of data we deal with is on the scale of Petabytes.

Challenges in Analyzing Data with Volume, Variety, and Velocity

An edge computing ecosystem consists of a large number of systems like hardware vendors, platform companies, application developers, IoT devices, edge devices, and telecom providers. Apart from these, there are two major key players in the ecosystem, i.e., Hyperscale Cloud Providers (HCPs) and Operations Technology (OT) vendors.

Amazon Web Services (AWS), Microsoft Azure, and Google are a few HCPs that provide core cloud infrastructure and platforms. They have thousands of developers working towards building application ecosystems that can serve multiple enterprises in many sectors globally.

It is a distributed landscape where the computing responsibilities are being shared between edge devices and the central cloud server, thus reducing the load on cloud servers. Edge devices processes some of the data locally and transmits only relevant information to the cloud server for further and complex processing. This way, edge computing helps to achieve faster responses to the end device/user, low bandwidth utilization, etc.

An example of the edge computing landscape is illustrated in the figure below.

What is Big Data Analytics?

Big data analytics is the use of processes and tools to glean insights from huge amounts of data. This data is either characterized as high velocity, large volume, or extreme variety. Analyzing this amount of data was previously beyond our reach using old tools like spreadsheets.

Big data analytics aims to derive conclusions and correlations. This is done with the help of tools like SAS, Hadoop, R, etc., which are more powerful than our columns and rows.

The Benefits of Big Data Analytics

- Using big data analysis leverages valuable data assets such as unstructured information in the form of emails, claim notes, etc., which significantly improve enterprise-wide information flow.
- The aggregation of data is much more important for different and new sources of information. Analytics can be used to effectively leverage partner data, for instance, and allow an enterprise to share processed data to improve operational intelligence.
- Despite the lack of face-to-face customer interactions, it is still possible to create an accurate customer profile using data aggregated from various sources such as demographics data, social media, public records, etc. A personalized product offer and risk-based pricing can be created by big data analytics using different sources of information.
- The most significant benefit of big data analytics is improved fraud detection. Analytics can help put the disparate pieces together to identify patterns at the beginning of the claim life cycle so insurers can protect against fraud.

Big Data Adoption in the Next 3 Years

Big data is quickly being adopted across the industry as a tech stack for data warehousing and analytics. The facts below summarize a survey conducted by SNS Telecom IT.

It is expected that insurance industries will invest $3.6 billion by 2021 in such technologies. This will cause a 30% increase in better data access to insurance services. It will scale up cost savings by 40-70% and there will be a 60% increase in fraud detection rates.

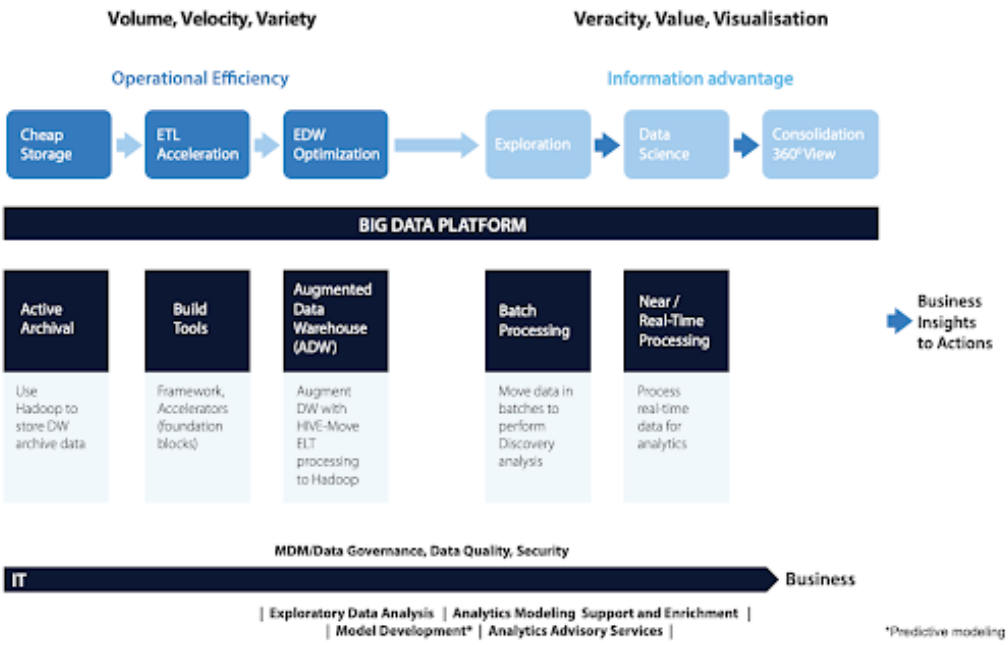
How to Build, Store, Analyze, and Use Unstructured Data

Unstructured data represents around 80% of data. It often includes text and multimedia content. Examples include e-mail messages, word processing documents, videos, photos, audio files, presentations, webpages, and many other kinds of business documents.

A strategy for using unstructured data should include the following:

- A clear view of the data that is being aggregated— different sources and locations, and different formats.
- A framework to integrate structured and unstructured data to gain insight
- A detailed description of the kind of data required—not all available data needs to be collected and used. Aligning business objectives to data collection and its sources to define the goal and the kind of data needed to achieve it will help identify the appropriate data sources and decide what must be archived or discarded.
- Data cleansing to ensure that data is accurate and complete.
- The tools that will be used to extract value from the aggregated data.
- The presentation of information in a way that people can understand and use.
- Privacy concerns and policies around the use of personal data.

Implementing Big Data Analytics Solutions



Big data technologies can be deployed to target either operational efficiencies in IT or an 'information advantage' for business. In the evolutionary stage, they are best suited to pursuing IT operational efficiency using the following:

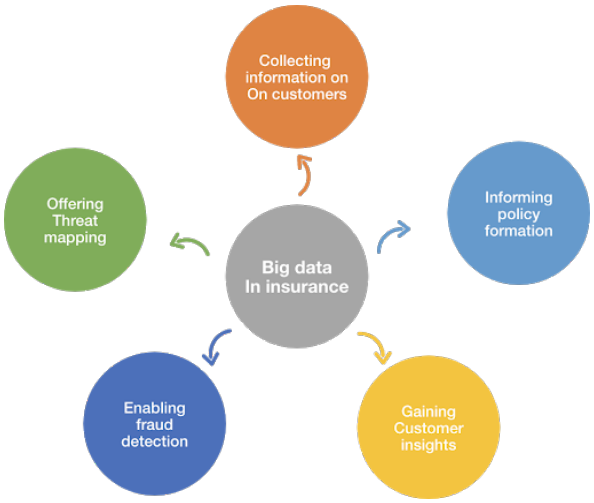
- Hadoop for 'Cheap Storage' and to create an 'Active Archive'.
- To build data warehouses that offer ETL acceleration and enterprise data warehouse optimization, resulting in efficiencies and cost savings.
- Data lakes that use big data/Hadoop to provide a strong data foundation and create a 'data factory'.

As the use of big data technology matures and organizations move towards business transformation, we will see the following results:

- Big data being used to manage and process huge volumes and varieties of data.
- Hadoop 2.0 and NOSQL big data technologies will facilitate real-time data processing and big data analytics through self-service BI.
- Emerging data discovery and data visualization tools will help create business insights and enable timely actions and interventions. However, to truly leverage the power of big data analytics, insurers must optimize the triad of technology, processes, and people.

Big Data Use Cases in the Insurance Industry

The insurance industry is important for both individuals and companies. Insurance provides stability to people and their dependents during uncertain times in life. With technology evolving quickly, how we can apply it across industries is an exciting possibility.



Pricing and Underwriting

The insurance business is all about estimating premiums for insurance policies, which are based on the law of large numbers and risk assessment procedures.

In automotive insurance, the premium calculation for each policyholder is performed through big data analytics by developing complex predictive analytics models. These models are very accurate and quick because they use big data technologies to process large pools of data. Some large IT companies like IBM are already supporting such systems for insurance companies.

Big data also fetches value from external data (social media, newspapers, analyst comments) in order to optimize their finances and instant payouts.

Settling Claims

Settling policy claims is a time-consuming process. It requires manual effort from an adjuster to assess the loss and recommend whether a claim should be paid or not.

In this current era of big data analytics, the claim settlement process is being improved through automation and analyzing historical data to provide valuable information. This is achieved by machine learning algorithms that use historical claims data as training data by identifying dependent and independent variables.

With big data, all data is stored in an efficient manner, and analytics are used to score the claims in an order of priority such that claims with higher levels of complexity get assigned to the appropriate adjuster. These processes save lots of time and help prevent insurance companies from issuing the wrong settlement amounts.

Using these kinds of analytics models, insurance companies can automate decisions such as whether a policy holder will be given a claim amount or not.

Gaining Insight on Healthcare for Customers

The health insurance sector is a very big business not only in Western countries but increasingly in Asia-Pacific regions as well.

Oscar, a US-based health insurance company, is utilizing data from medical records aggregators and linking it with patient medical history. By doing this, it's helping patients by making permission-based recommendations on its user platform. Oscar has achieved this via implementation of prescriptive analytics.

Basically, they are applying mathematical models to big data in order to provide recommendations to patients and thus help them with preventive care.

For example, suppose a patient living with asthma normally orders an inhaler drug but forgot to do so before it ran out. An Oscar nurse can be notified of this situation by Oscar's telemedicine service, an appointment can be arranged with a physician, and the patient can have the prescription filled in a timely manner.

This is one example of how health insurers can intervene when patients are at risk.

Fraud Detection

Similar to the automation of insurance claim settlements, a rapid increase in fraud detection can be achieved in big data by predictive analytics. The independent variables of a newly filed claim are checked against the variables present in the training dataset of previous fraud claims. If a match is found with high accuracy, a new claim is filed and sent for further investigation.

The detection of complex fraudulent claims is dependent on many factors that can be missed when using manual human effort, but big data analytics can perform these tasks successfully.

Threat Mapping

Setting policy premiums also becomes easy when big data provides organizations with ample information to analyze. When a customer wants car insurance, for instance, the insurance company can analyze information about the areas where they travel the most, determining the likelihood of the vehicle being damaged or discovering how prone that area is to road accidents. When an insurance agency has this sort of information available at their disposal, they can easily decide how much premium they should charge to avoid any losses.

Conclusion

Data in the insurance industry is increasing daily, with a 90% increase in the last two years. Insurance businesses run on data analysis so they can better decide the price of insurance policies, settle claims, analyze customer behavior to facilitate benefit payouts, detect fraud, and map threats, etc. Big data analytics process volume, velocity, and variety in datasets in ways that cannot be done using older technologies. Big data analytics provide many benefits, such as lower processing costs, improving your pricing, competing with big businesses, focusing on local preferences, and increasing your efficiency, sales and customer loyalty.

About the Authors

Param Jeet is a Ph.D. in Mathematics and has more than 10 years of hands-on industry experience. He has published many research papers in international journals and authored a book. Currently, he is leading AI/ML practice in GlobalLogic, India, and collaborating with universities to enhance the Industry-Academic research collaboration program.

Subodh Pandey - Subodh Pandey is an Engineering Graduate in Information Technology with more than 8 years of industry experience. He has done Executive Program in Data and Decision Sciences from the Indian Institute of Technology Delhi. Currently, he is working as a data scientist with one of GlobalLogic's Insurance domain Client. Also, he is a core member of AI/ML practice in GlobalLogic, India.

References

How Big Data is Changing the Insurance Industry

Top 7 Big Data Use Cases in Insurance Industry — Exastax

4 Big Data use cases in the insurance industry you should know about

GlobalLogic®

GlobalLogic, a Hitachi Group Company, is a leader in digital product engineering. We help our clients design and build innovative products, platforms, and digital experiences for the modern world. By integrating our strategic design, complex engineering, and vertical industry expertise with Hitachi's Operating Technology and Information Technology capabilities, we help our clients imagine what's possible and accelerate their transition into tomorrow's digital businesses. Headquartered in Silicon Valley, GlobalLogic operates design studios and engineering centers around the world, extending our deep expertise to customers in the automotive, communications, financial services, healthcare & life sciences, media and entertainment, manufacturing, semiconductor, and technology industries.



www.globallogic.com