Data Analysis for Vehicle Claims

Reena.D Sathyan.S

UG Scholar, UG Scholar

Dept of CSE, Dept. of CSE

KGiSL institute of technology KGiSL institute of technology

[reenaprajapati1704@gmail.com](mailto:Reenaprajapati1704@gmail.com) [sathyans1998@gmail.com](mailto:sathyans1998@gmail.com)

Vinagayamoorthy.G

UG Scholar,

Dept of CSE,

KGiSL institute of technology

Abstract-

A system that helps in analyze the motor insurance vehicle claims and visualize the data in a form of charts and graphs.It will be useful for future prediction of the data by using those charts and graphs and comparison of that data with a past years vehicle insurance claim data activities. Claims analysis is a technique for examining the consequences of design features that are described in current or future scenarios of use. A "claim" is a statement of the consequences of a specific design feature or artifact on users and other stakeholders. This system is planned to build using the TABLEAU software to visualize the data. This software is a commerce ability tool for visually analyzing the data, and provide the information to improve the insurance claims.

Keywords: Data collection,Analyze and Visualizing the data,Comparison of data, Prediction and Report creation.

# INTRODUCTION

Vehicle insurance (also known as car insurance, motor insurance or auto insurance) is [insurance](https://en.wikipedia.org/wiki/Insurance) for [cars](https://en.wikipedia.org/wiki/Automobile), [trucks](https://en.wikipedia.org/wiki/Truck), [motorcycles](https://en.wikipedia.org/wiki/Motorcycle), and other road vehicles.it will plays a key roll in financial industry and it is also importance for people economical life.Its primary use is to provide financial protection against physical damage or bodily injury resulting from [traffic collisions](https://en.wikipedia.org/wiki/Traffic_collision) and against [liability](https://en.wikipedia.org/wiki/Legal_liability) that could also arise from incidents in a vehicle. Vehicle insurance may additionally offer financial protection against [theft](https://en.wikipedia.org/wiki/Motor_vehicle_theft) of the vehicle, and against damage to the vehicle sustained from events other than traffic collisions, such as [keying](https://en.wikipedia.org/wiki/Keying_(vandalism)), weather or [natural disasters](https://en.wikipedia.org/wiki/Natural_disaster), and damage sustained by colliding with stationary objects. The specific terms of vehicle insurance vary with legal [regulations](https://en.wikipedia.org/wiki/Regulation) in each region. Analyze and visualize the motor insurance vehicle claims automatically in order to make more understandable.

Claim analysis is a technique for examining the positive and negative consequences of design features that are described in current or future scenarios of use. A claim is a statement of the consequences of a specific design feature on user and stakeholder

The main objectives of the project are as follows:

* Database: Relevant information of vehicle insurance claim are collected and stored in the database
* Algorithm: For data cleaning process some sort of algorithms are used in order to retrieve the useful data.

In this our banking bot, our main objectives are as follows:

* Working on large number of data
* Classify the insurance claim details
* Visualizing the data in the form of charts and graphs.
* Has eliminated most of the flaws in the existing applications.

The study seeks to analyze the claims of motor insurance in public and private insurers and their effects. In this particular case, it was believed that the study would have the following specific significances.

* It briefly explains the importance of motor insurance and enhances knowledge of the customers.
* It depicts the extent of customer satisfaction with existing motor insurance operational practices.
* It reveals the effect of centralization in prompt claim settlement and flexibility of premium discounts between the public and private companies.
* It will be used as a lesson and creates opportunity for detail and subsequent research on the same area.

The main goal of this project is to analyse and visualise the insurance claim by using some sample data .The analysing process is done by some criteria.

Some of the applications of the twitter have been described below:-

1. LITERATURE REVIEW

Nowadays, People choose various insurance policy to assist themselves and their properties. There are various insurance type which are followed by people such as property, health, medical and auto insurance. Car insurance is mostly imposed by the authority for those who own and drive vehicles and it was established in the early 1900s, and has been applied since then. It is now compulsory in most residences around the world. Also in African country, it is mandatory when buying a car with a loan to buy a comprehensive insurance which can compensate the potential loss or damage to the vehicle as well as any damage occurred to others. Today, most of the companies in the world are using various marketing strategies to attract new consumer and successfully differentiate from other competitors to compete with their rivals by raising their market shares and increasing profit. They spend in marketing, sales, distribution, research and development, and other functional departments to meet customers needs better.

Although other marketing strategies are also beneficial, research demonstrates that defensive strategies can be more worthwhile because it brings higher profit for companies than strategies focusing on market share expansion or cutting cost.

Thus, companies within the insurance field such as other areas must pay peculiar attention to this matter to reach long lasting relationships with customers. Customers are clearly more pleased with high quality service, and when a company gains the customers’ satisfaction, profitability is ensured hence, assessing their relationship can be beneficial for the managements in the companies. Obviously, one of the basic sources of competitive advantage between companies to attract customers is the price of goods and services, in our insurance industry case, premium. However, unlike price, there are many quality issues which can be improved and stretched over time to attract customers by satisfying them. The service quality theory is considered one of the most debated subjects in service literature (Ooi, et al., 2011) [6] since it lacks agreement when it comes to describing and defining service quality (Gupta, et al., 2005) [7] . Service quality is a focused evaluation reflecting the customer's perception of specific service dimensions including reliability, responsiveness, assurance, empathy, and tangibility (Zeithaml & Bitner, 2003) [8] . Furthermore, Customer satisfaction is one of the main areas of interest in marketing, business and the academic world (Tikkanen & Alajoutsijaui, 2002) [9] . In today’s changing world, developed economies are becoming service oriented, meaning that customer impressions play a critical role in this field. A company delivers services to customers, while overall customer attitude towards the company is defined in the relationship between service quality and customer satisfaction. Different research was done on the positive relationship between service quality and customer satisfaction. In other words, if service quality decreases quickly, customer satisfaction declines dramatically and if service quality grows radically, customer satisfaction rises rapidly too. Some researchers have tested the inspiration of service quality and customer satisfaction. An assessment model suggested by Sharareh et al. (2012) [10] emphasizes on the relationship between perceived service quality, customer satisfaction and interest in purchasing. This research shows that customer satisfaction is an intervening variable of service quality and repurchases interest. In other words, service quality influences customer satisfaction, and customer satisfaction affects repurchase interest. In addition, Zeithaml et al., (1996) [11] , documented that behavior trends demonstrate the effect of service quality on customer satisfaction. Moreover, other research has been carried out on the relationship between service quality and customer satisfaction, a connection which has received a good deal of attention in literature (Bolton & Drew, 1994) [12] . In spite of the results of relationship between service quality and customer satisfaction, I didn’t come across a study which compares public insurers with that of the private insurers in relation to customer satisfaction. In addition, as far as my knowledge is concerned, there was no any research conducted on insurance service in the study area. Therefore, this research is expected to fill this gap in literature.

1. PROPOSED SYSTEM

Our proposed system is automated process which follows from collecting the data from the users using the google forms or an insurance company’s own web or mobile application.

After collecting the necessary data from the subscriber the analyzing process is carried out using the table au software and analyzing the data means organizing or structuring the data.

The visualization will be shown accurately using the same software.the comparision will be carried out simulataneous and the future prediction is analyzed

Proposed System is to implement latest technology for making the system still more beneficial to cover the core areas.

Apart from the previous mentioned techniques some of the other techniques involved in the paper are described below:-

1. Collection

The data is required in order to process visualization.so we should collect data from the user and then process,analyze,visualize and so on.The only process we are carrying manually is the data collection.this is a manual process.this process is done by distributing forms to the subscribed users.The different type of vehicle insurance details of different customers will be collected by using the Google forms and it will be stored as a CSV file in a database.

1. Data cleaning data storing

Data cleaning will be performed in python by using Python’s Pandas and NumPy libraries to use clean data

We will cover the following:

* Dropping unnecessary columns in a Data Frame
* Changing the index of a Data Frame
* Using .str() methods to clean columns
* Using the DataFrame.applymap() function to clean the entire dataset, element-wise
* Renaming columns to a more recognizable set of labels
* Skipping unnecessary rows and columns in a file

After the data cleaning will be completed that data will be stored in database with .CSV extension

1. Analysing

The cleaned data will be retrieved from the database and it will be analysed based on the customer needs and problems. It will be performed by following steps

1 .Define a problem-The main objective will be defined

2 .Determine the problem-

3 .Design the problem-the problem and its solution will be designed

4 .Analyse the problem

5 .Sample

6 .Report- The report will be generated in a form of charts and graphs

Those graphs and charts will be used to perform a multiple operation like comparison and prediction

KMEANS, Linear Regression are the algorithms to use for the analysis process.

1. VISUALIZATION

This is the graphical representation of our analyzed data.visualization will be better if the analyzation is better.this process is automated by jupyter followed by python.The more we see the more we understand.this is basic idea of visualization and this is what we carry out.

Visualization will be a form of graphs and charts

1. Comparison

The analyzed data will be visualized in many form of charts and graphs. These charts and graphs will be used to perform a comparison operation.Comparision is carried out for better understanding of data.we can compare it with our previous data and can predict our growth in which business call YOY called as year on year growth.data must be compaed for better understanding of ourselves.The data activity will be compared year by year and find the percentage rate of claim, deduction, income rate and other details.

1. Prediction

We should be able to predict the future responses.however the prediction may not be accurate.data prediction is best when carried out with data comparision.sometimes prediction can save us from huge disastrous for instance,when yoy starts decreasing ,it can be a warning to change our strategy.

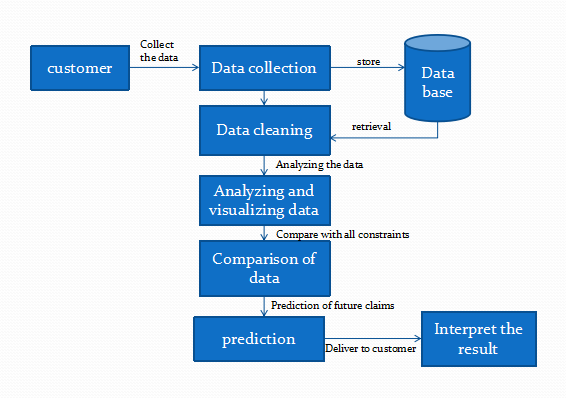


Figure 2: Architecture Diagram

1. CONCLUSION

The main aim of this paper is to provide the essential information to the insurance companies for increase their income and claim rate in a future. The key idea we applied here is the automation . the automation we applied can save us from time ,labour and much more.the simpler is now gone much simpler and the application can be managed by a few people.

The process we done above is collected data is transferred into visualized data and are compared and finally predicted.this will be very useful in vehicle insurance as the insurance customer can detaily view the progress such as subscription etc.

We would like to conclude that our

REFERENCES

[1]. A. Sil and A. Yates, “Re-ranking for joint named-entity recognition and linking,” in Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage.,2013, pp. 2369–2374

[2]. B. Han and T. Baldwin, “Lexical normalisation of short text messages: Makn sens a #twitter,” in Proc.49th Annu. Meeting. Assoc.Comput. Linguistics: Human Language Technol., 2011, pp. 368–378.

3]. C. Li, J. Weng, Q. He, Y. Yao, A. Datta, A. Sun, and B.-S. Lee, “Twiner: Named entity recognition in targeted twitter stream,” in Proc. 35th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2012, pp. 721–730.

[4]. Chenliang Li, Aixin Sun, Jianshu Weng, and Qi He, Member, “Tweet Segmentation and Its Application to Named Entity Recognition”, IEEE transaction, VOL. 27, NO. 2, FEBRUARY 2015.

[5]. K.-L. Liu, W.-J. Li, and M. Guo, “Emoticon smoothed language models for twitter sentiment analysis,” in Proc. AAAI Conf. Artif.Intell., 2012, pp. 1678–1684.

[6]. X. Meng, F. Wei, X. Liu, M. Zhou, S. Li, and H. Wang, “Entity centric topic-oriented opinion summarization in twitter,” in Proc.18th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining, 2012, pp. 379–387.

[7]. X. Wang, A. McCallum, and X. Wei, “Topical n-grams: Phrase and topic discovery, with an application to information retrieval,” in Proc. IEEE 7th Int. Conf. Data Mining, 2007, pp. 697–702.

[8]. Zhenhua Wang, Lidan Shou, Ke Chen, Gang Chen, and Sharad Mehrotra, “On Summarization and Timeline Generation for Evolutionary Tweet Streams”, IEEE transaction, VOL. 27, NO. 5, MAY 2015.