International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Fraud Detection and Risk Measurement in Insurance Systems using AI, Blockchain

[1] ABHISHEK HN,[2] NARASIMHA KARANTH,[3]SANJAY V, [4]TARUN R KARKERA, [5] ASHA MS

[1][2][3][4]B.E Students, Department of Computer Science and Engineering
[5] Asst. Professor, Department of Computer Science and Engineering
[1][2][3][4][5] Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

Abstract—

In today's world, the insurance industry is widely recognised as one of the fastest expanding businesses. Insurance companies are on the cusp of adopting cutting-edge technology, processes, and mathematical models in order to optimise profits and minimise losses while handling client claims. Traditional approaches, which rely solely on human models and are both time-consuming and imprecise. In this study, we provide a secure and automated insurance system framework that eliminates human involvement, secures insurance activities, detects fraudulent claims, and decreases monetary loss for the insurance industry. The extreme gradient boosting (XGBoost) machine learning technique is used for the insurance services after presenting the blockchain-based framework to secure data exchange among different interacting agents within the insurance network. Furthermore, an outstanding technique is proposed for dealing with insurance network scams automatically.

Index Terms— Blockchain, Artificial Intelligence,Machine Learning, Insurance , risk measurement, Data Analysis, Fraud detection,Neural Networks**.**

Introduction

Due to claims leakage, insurance companies in today's society are losing a significant amount of money. Insurance companies have a massive and costly problem in the form of fraudulent claims, which cost the business billions of dollars each year. Insurers have always relied on mathematicians to estimate risk, develop premium rates for policy underwriting, and ensure that good levels of payouts are achieved without jeopardising the company's financial stability. Traditional approaches for detecting insurance fraud are extremely complicated and time-consuming. Expert inspection, adjusters, and specifically skilled investigation services are the mainstays of their operations. Added to that, manual detection results in additional expenses and inaccurate results. In this paper, a Smart Insurance System based on Blockchain and Artificial Intelligence - Insure AI is a service that helps businesses codify their operations, automate claim processing, understand their client’s risk profiles, and detect fraudulent claims. The usage of Blockchain is suggested in this situation because it is highly recommended for protecting sensitive and personal information. Insurance firms can use shared Blockchain data to protect themselves from fraudsters, double claim filing, and improve fraud detection efficiency. The fraud detection and risk measurement modules are also built using two machine learning algorithms. The first offline method relies on a batch learning strategy, in which the algorithm trains the entire large dataset at the same time. The second online method is based on an online learning strategy that dynamically trains, updates, and upgrades the learning weights as fresh data enters the system, eliminating the need to retrain the entire model from start as new data continues to arrive.

Related Work

Blockchain has recently sparked a lot of interest since it is a ground-breaking database technology that can help solve complex technical difficulties in a variety of fields. Indeed, Blockchain is no longer solely connected with financial and banking applications. This technology has tremendous potential in a variety of fields, including, but not limited to, information security, health care, logistics, insurance, military, and others. Blockchain is being used to combat Distributed Denial of Service (DDoS) assaults in the world of cybersecurity. The study presents a method to reduce DDoS attacks by implementing a private blockchain that uses decentralized Content Delivery Networks (CDNs) with trusted node participants authorized by military or government agencies.

In 2016, it was recommended to do research and use the random forest model to vehicle insurance fraud. When applied to a big dataset with a large number of important explanatory factors, the random forest technique reduced variable filtering, according to this study. Following that, it is concluded that random forest is appropriate for large datasets and unbalanced data.

Artificial intelligence (AI) and machine learning systems have the potential to be integrated into the insurance industry's claims processing, customer care, and fraud detection sub-sectors.

In 2017, a survey on insurance fraud analysis employing a prediction model was developed. This is a survey of decision trees, Random Forests, SVMs, neural networks, and XG Boost. Different sorts of fraud were investigated, and it was discovered that big data analytics was effective in forecasting claims in both vast and diverse data sets.

Using fuzzy logic membership functions, the latter technique was then utilised to anticipate fraud in large and high-dimensional data sets.

In order to detect fraud, the health care insurance fraudster detection approach was highly recommended.Moreover, detecting fraud in auto insurance based on nearest neighbor models utilized in concert with traditional statistical methods was investigated where distance-based, density based,statistics methods were the used methods to detect fraud occurrence.

III. PROPOSED BLOCKCHAIN-BASED AND AI-DRIVEN INSURANCE NETWORK ARCHITECTURE

In a nutshell, we propose two techniques for sharing insurance data. The suggested methods use both homomorphic and aggregate signatures to link information about the legitimacy of insurance firms' data. As a result, in order to validate data, a party must use the correct consent of the entity that owns the data. One method provided allows service providers to double-check the accuracy of client information. Without having to engage with the authority, a person can convey any subset of approved material or test results using a homomorphic signatureIt also assures that when the test results are shared, the individual does not expose any potentially harmful information. The use of aggregate signatures successfully prevents service providers from communicating insurance data in an illegal (or unapproved) manner. In this case, either the receiving company recognises that the data was sent without the data owner's permission, or the data owner can determine which service provider disclosed his data without his permission and hold that party responsible for the breach. The fundamental originality of the presented study is the proposed system, which addresses the coupling of homomorphic and aggregate signatures, as well as the application of the proposed technique to insurance data. The proposed system is based on known cryptographic primitives (especially homomorphic and aggregate signatures), but its implementation is not straightforward. In general, sharing privacy-sensitive data between entities is a novel research topic. The key distinctions between insurance data and other types of sensitive data are as follows: It is typically shared in pieces (as different parts of it or different computations on it are requested by different parties); and (iv) its credibility is critical to the parties who use it. It contains information about family members; it is not revokable (and thus it is critical to ensure that it is not leaked); it is typically shared in pieces (as different parts of it or different computations on it are requested by different parties); and it is typically shared in pieces (as different parts of it or different computations on The suggested method addresses several of the aforementioned unique characteristics of insurance data.

Multiple certified institutions (CIs), clients, and service providers (SPs) are assumed to exist in the system. To keep things simple, we'll discuss the proposed scheme with only one CI, client, and SP. The CI is primarily in charge of sequencing, encrypting, and signing the data that has been sequenced. One of these systems may be able to easily include our proposed strategy to provide a full pipeline. Having such a CI is also unavoidable with today's sequencing technologies. The SP might be a company that provides insurance in practise. The SP is most interested in getting a piece of the data from insurance clients. In each of these cases, the client want to share their information either anonymously (without revealing her genuine identity) or publicly (with her true identity known). They also want to know if the SP can share the client data it receives from the insurance companies with other organisations (either anonymously or with the real identities of clients).

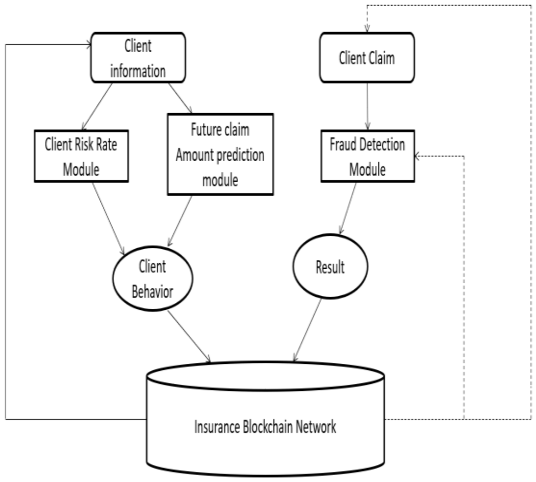

Fig 3.1. Architecture of InsureAI

IV. MACHINE LEARNING-DRIVEN FRAUD DETECTION AND RISK MODELLING

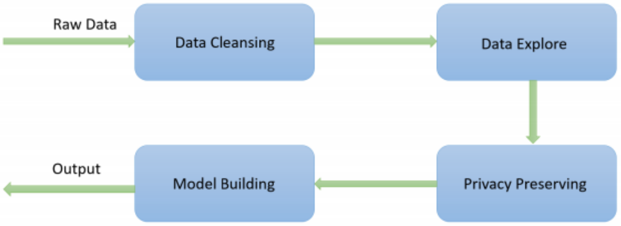

Fig 4.1. Proposed machine learning module pipeline

A. OFFLINE LEARNING: EXTREME GRADIENT BOOSTING ALGORITHM

XGBoost is one of the most efficient gradient boosted decision tree implementations, and it has been selected as one of the best offline machine learning algorithms in Kaggle competitions. XGBoost minimises execution time while enhancing performance when compared to various machine learning methods. It was created with the goal of optimising memory use and maximising the computational power of the device. The primary principle behind boosting is to create sub-trees from an initial tree in such a way that each one minimises the preceding one's faults. In this method, the new subtrees will update the prior residuals, lowering the cost function's inaccuracy.

B. ONLINE LEARNING: CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks have been used to make revolutionary discoveries in a variety of pattern recognition areas, from image processing to voice recognition, throughout the previous decade. The most significant benefit of CNNs is that they reduce the number of parameters in ANNs. This achievement has prompted scholars and developers to examine bigger models to address difficult problems that were previously impossible to solve with regular ANNs. The most important assumption regarding the problems that CNN solves is that they should not have spatially dependent features. In other words, in a face detection application, the position of the faces in the photographs is irrelevant.

Conclusions

InsureAI, a unique insurance fraud detection system based on permissioned blockchain and machine learning algorithms, is presented in this study. Based on experimental performance on data from a genuine insurance firm, two learning strategies for detecting and classifying fraudulent claims submissions were chosen from a pool of learning techniques. As an offline learning approach, XGBoost, an extreme gradient boosting machine learning algorithm, will be used, while Convolutional Neural Networks will be used in the creation of online learning methods.

REFERENCES

[1] N. Dhieb, H. Ghazzai, H. Besbes, and Y. Massoud, ‘‘Extreme gradient boosting machine learning algorithm for safe auto insurance operations,’’ in Proc. IEEE Int. Conf. Veh. Electron. Saf. (ICVES), Cairo, Egypt, Sep. 2019, pp. 1–5.

[2] A. Pieroni, N. Scarpato, L. Di Nunzio, F. Fallucchi, and M. Raso, ‘‘Smarter city: Smart energy grid based on blockchain technology,’’ Int. J. Adv. Sci., Eng. Inf. Technol., vol. 8, no. 1, pp. 298–306, 2018.

[3] D. Corum, ‘‘Insurance research council finds that fraud and buildup add up to $7.7 billion in excess payments for auto injury claims,’’ Insurance Res. Council, Malvern, PA, USA, Tech. Rep., Feb. 2015. [Online]. Available: https://www.insurance-research.org/sites/default/files/downloads/IRC%20Fraud%20News%20Release.pdf

[4] S. Wang, J. Wang, X. Wang, T. Qiu, Y. Yuan, L. Ouyang, Y. Guo, and F.-Y. Wang, ‘‘Blockchain-powered parallel healthcare systems based on the ACP approach,’’ IEEE Trans. Comput. Social Syst., vol. 5, no. 4, pp. 942–950, Dec. 2018.

[5] M. Samaniego and R. Deters, ‘‘Internet of smart Things–IoST: Using blockchain and clips to make things autonomous,’’ in Proc. IEEE Int. Conf. Cognit. Comput. (ICCC), Honolulu, HI, USA, Jun. 2017, pp. 9–16.

[6] T. Chen and C. Guestrin, ‘‘XGBoost: A scalable tree boosting system,’’ in Proc. ACM Int. Conf. Knowl. Discovery Data Mining, (SIGKDD), San Francisco, CA, USA, Aug. 2016, pp. 785–794.

[7]] G. Kowshalya and M. Nandhini, ‘‘Predicting fraudulent claims in automobile insurance,’’ in Proc. 2nd Int. Conf. Inventive Commun. Comput. Technol. (ICICCT), Coimbatore, India, Apr. 2018, pp. 1338–1343.

[8]] K. Supraja and S. J. Saritha, ‘‘Robust fuzzy rule based technique to detect frauds in vehicle insurance,’’ in Proc. Int. Conf. Energy, Commun., Data Anal. Soft Comput. (ICECDS), Chennai, India, Aug. 2017, pp. 3734–3739

[9] T. Badriyah, L. Rahmaniah, and I. Syarif, ‘‘Nearest neighbour and statistics method based for detecting fraud in auto insurance,’’ in Proc. Int. Conf. Appl. Eng. (ICAE), Batam, Indonesia, Oct. 2018, pp. 1–5.

[10]] J.-M. Long, Z.-F. Yan, Y.-L. Shen, W.-J. Liu, and Q.-Y. Wei, ‘‘Detection of epilepsy using MFCC-based feature and XGBoost,’’ in Proc. 11th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI), Beijing, China, Oct. 2018, pp. 1–4