



INFLUENCE OF PREDICTIVE ANALYSIS IN HEALTH INSURANCE FRAUD DETECTION

Surya Susan Thomas

Department of Computer Science
Quaid-E-Millath Government College for Women
Anna Salai, Chennai, Tamil Nadu, India

Dr. Ananthi Sheshasaayee

Department of Computer Science
Quaid-E-Millath Government College for Women
Anna Salai, Chennai, Tamil Nadu, India

Abstract— Fraud in insurance sector is a snag faced by both the insurance companies as well as the policy holders. Annually, it costs millions of dollars to the turnover of the insurance sector. Pinpointing fraudulent claims from a vast number of insurance claims is tedious and frenetic. Traditional healthcare fraud detecting analytical solutions turn to be outdated as the inflow of claims are soaring. With the influx of various datamining techniques and tools, a better solution for popping out frauds can be visioned. This paper studies the efficiency of predictive analysis to detect fraudulent claims especially from the health insurance claims.

Keywords— Health insurance, fraud detection, datamining, predictive analysis

I. INTRODUCTION

Insurance claim money is an easy target for fraudsters who obtain financial gain by using fraud methods. Insurance frauds has been there for some time and causes terrible loss to the insurer and policy holders. According to Clarke, insurance fraud is “to appear as normal and to be processed and paid in routine manner” [1]. The loyalty between the ‘insured’ and the ‘insurer’ is key to the success of the insurance business.

Insurance frauds are mainly classified into three categories

- Internal (e.g. employee, agents and managers) vs external (e.g. policyholders, claimant) or deceit of internal and external entities
- Underwriting (e.g. application fraud, premium fraud) vs claims (e.g. consciously false and formulated fraud)
- Soft (e.g. inflated fraud made by otherwise honest policyholder) vs hard (e.g. planned fraud with staged and bogus claims) [2]

According to an Indian Forensic study (2012) “The losses caused to Indian insurance sector due to different kinds of frauds are Rs 30,401 crore (approximately \$50 billion) which is roughly 9 per cent of the total estimated size of insurance industry in the year 2011 “[3]. This is a bogus value when compared to developed countries where frauds are less than 0.6% of industry size. The frauds in non-life insurance

industry such as motor and health sectors are almost six times higher than in life insurance sector.

II. FRAUD ANALYTICS

The health insurance sector must shift to a fraud control system that clubs fraud prevention and detection at the upper level of the payer system, applies analytic reviews throughout the system, and assimilate post-pay checking and recovery processes at the back-end of the system. The claim checking processes should include rules-based data analytics, predictive modeling, and refined technologies to diagnose fraud before the claim amount is paid [4]. Fraud detection is best attained through a layered idea which flags claims analysis, including identity analytics, claims analytics (predictive modeling and rules-based fraud detection), and social network analytics.

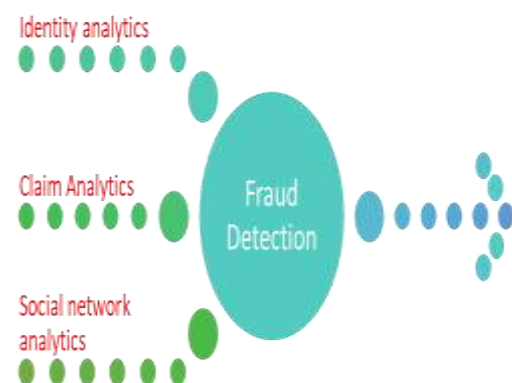


Fig 1: Illustrating role of data analytics in fraud detection

A. Identity Analytics

A valid identity analysis includes verification and authentication of all provider details and approving of the background evaluations. All the new providers should undergo an identity verification step as part of an enrollment qualification process.[5] Besides this, the existing providers’



identity should be periodically reviewed enabling the identification of risky providers from a payer's network.

B.Claim Analytics

Pre-pay claims analytics is always best suited for an efficient fraud detection model. Pre-pay analytics combines rule-based screens and edits and predictive modeling that helps in identifying the claims that sweeps beyond their territory before they are paid. [6] Moving the working of the rule-based fraud detection systems to the initial stage of claim approval and also complementing it with predictive modeling techniques helps to understand the fraudsters easily and efficiently. [7]

C.Social network analytics

The filthy money game that is played among service providers and between providers and patients can be easily identified by social network analysis by the hidden patterns of information sharing and coalition within potentially fraudulent groups like patient relationships with known perpetrators of healthcare fraud, links between recipients, businesses, assets and relatives and associates, links between licensed and non-licensed and sanctioned providers and inappropriate relationships between patients, providers, employees, suppliers and partners. [8]

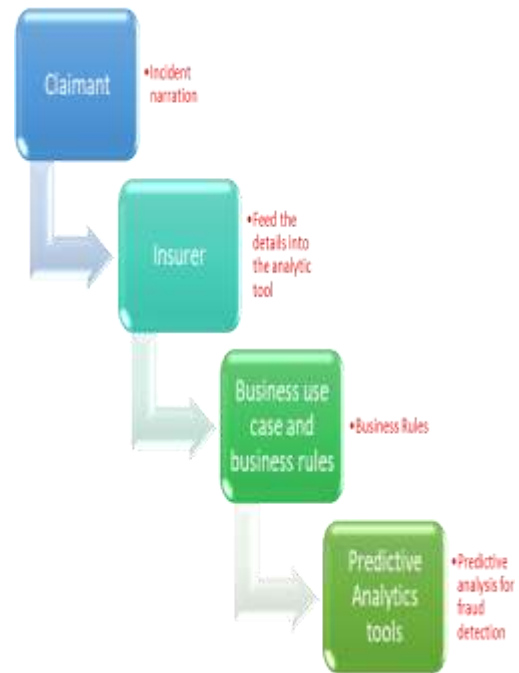


Fig 2: Steps invoking fraud detection through predictive analytics

III. PREDICTIVE ANALYSIS IN FRAUD DETECTION

The insurance sector employs a system that centralizes on effective forecasting. Policy approvals, renewals, risk projections, claim settlement and underwriting are all based from the insurance provider's talent to make factual predictions depending on the data. [9] Defective forecasting pave way to misaligned premiums, unexpected number of claims and inaccurate risk projections—each entity can hamper an insurance provider's foundation and ultimately risking the customer's financial role. [10] Hence being "wrong" in this sector would be a big burden for everyone.

The present day predictive analytics include the use of text analytics and sentiment analysis to detect fraud in the claims. The claim reports usually cover multiple pages leaving a very little chance for text analytics in fraud detection. Predictive analytics technology is deployed in this case to spot potentially fraudulent claims and thereby enhance the payment of legitimate ones. [11] Traditional predictive analytics were mostly used to examine statistical information stored in the structured databases but now the scenario has branched into big data realm too. [12]

IV. ADVANTAGES OF PREDICTIVE ANALYTICS

Predictive Analytics in fraud detection enhances the power of predictive modelling to diagnose fraudulent involvement as it is being processed. This approach has numerous advantages as

- ❖ Employs an adaptive system for pre-payment fraud detection that visualizes the given data precisely. [13]
- ❖ Suspicious claim settlements are automatically detected using predictive models. [14]
- ❖ Suggests a better mode of interaction for every settlement.
- ❖ Drastic reduction in fraud and abuse costs. [15]
- ❖ Settles legitimate and genuine claimants faster and with greater certainty.
- ❖ Use of efficient investigative mechanisms, thus reducing maintenance cost.
- ❖ Decrease in expenditure for recovery measures.
- ❖ Provides stringent investigation against fraudulent cases.

V. CONCLUSION

Fraud detection using predictive analytics allows data analysts to embed advanced fraud detection models into the system thus increasing the predictive capabilities of the transactional systems to the next level. A large number of fraudulent claims



can be brought to light at the processing level and far before payment is sanctioned. Henceforth, a significant reduction in financial costs is observed and dwindle the burden of both the customer and the insurance provider.

VI. REFERENCES

- [1] Clarke, Michael. "The control of insurance fraud: A comparative view." *The British Journal of Criminology* 30.1 (1990): 1-23.
- [2] Viaene, Stijn, and Guido Dedene. "Insurance fraud: issues and challenges." *The Geneva Papers on Risk and Insurance-Issues and Practice* 29.2 (2004): 313-333.
- [3] <http://indiaforensic.com/certifications/india-loses-6-25-billion-to-insurance-frauds-an-indiaforensic-research/>
- [4] <https://risk.lexisnexis.com/insights-resources/article/preventing-healthcare-fraud-with-big-data-and-analytics>
- [5] Siegel, Eric. *Predictive analytics: The power to predict who will click, buy, lie, or die*. Hoboken (NJ): Wiley, 2016.
- [6] Kirlidog, Melih, and Cuneyt Asuk. "A fraud detection approach with data mining in health insurance." *Procedia-Social and Behavioral Sciences* 62 (2012): 989-994.
- [7] Gopinathan, Krishna M., et al. "Fraud detection using predictive modeling." U.S. Patent No. 5,819,226. 6 Oct. 1998.
- [8] Šubelj, Lovro, Štefan Furlan, and Marko Bajec. "An expert system for detecting automobile insurance fraud using social network analysis." *Expert Systems with Applications* 38.1 (2011): 1039-1052.
- [9] Baesens, Bart, Véronique Van Vlasselaer, and Wouter Verbeke. "Predictive Analytics for Fraud Detection." *Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques: A Guide to Data Science for Fraud Detection*: 121-206.
- [10] Marquis, M. Susan, and Martin R. Holmer. "Alternative models of choice under uncertainty and demand for health insurance." *The Review of Economics and Statistics* (1996): 421-427.
- [11] Shmueli, Galit, and Otto R. Koppius. "Predictive analytics in information systems research." *Mis Quarterly* (2011): 553-572.
- [12] <https://www.the-digital-insurer.com/wp-content/uploads/2013/12/53-insurance-fraud-detection.pdf>
- [13] https://www.elpasohealth.com/articles/prepayment_fraud.pdf
- [14] <https://www.casact.org/education/specsem/f2008/handouts/bhattacharyya.pdf>
- [15] <https://www.lexisnexis.com/risk/downloads/.../The-Real-Cost-Health-Care-Fraud.pdf>