Anomaly analysis and detection in health insurance

Kendall Carolissen
Mehmet Kacal

Milliman

In this article we describe a research study developed for VGZ, one of the largest health insurance companies in the Netherlands.

We analyse potential fraudulent claims in a portfolio of healthcare provider invoices for neighbourhood care. Healthcare fraud is considered a material risk in the Netherlands. The excessive number of healthcare fraud cases in the Netherlands has put undue pressure on the Dutch prosecution system, who approximate an annual loss of €100 million due to healthcare fraud.[1]

There is a growing effort by insurers to tackle the issue of health insurance fraud given its materiality. With the use of the expertise and supervision of Milliman, we were able to research and produce a Master's thesis on the application of statistical methods to detect fraud. The analyses were conducted on real-world invoice data from VGZ. Based on interviews with fraud experts, key invoice data features were selected to potentially detect insurance fraud within VGZ. This article briefly describes the approach of the research and the statistical methods applied to more effectively detect fraud.

Our research began with interviews of several domain experts within VGZ. These domain experts detect suspicious behaviour related to fraud by both manual inspection of invoices and with the use of a model. We concluded that the application of two specific statistical methods could potentially complement the existing methods for detecting irregular invoices.

The first statistical method is based on Benford's law. Benford's law is a naturally occurring phenomenon, which makes a statement about the frequency distribution of leading digits in real datasets (e.g., accounting, macro-economic and election data). The method based on Benford's law aims to separate naturally occurring numbers from humanly fabricated numbers. We found that this can therefore also be used to target potential fraud in invoice cost amounts, where these are fabricated in order to invoice treatments that never took place.

Applying Benford's law on the actual sets of invoices of a large number of healthcare providers, we were able to identify healthcare providers which, after further research, were considered to have potentially made fraudulent claims.

The second statistical algorithm applied in this work is the Isolation Forest (iForest). The iForest is a machine learning algorithm for anomaly detection based on the principle of isolating anomalies. Compared to the majority of statistical models, the iForest algorithm has a fundamentally different approach. Statistical models for fraud detection often model legitimate instances in order to identify instances different from those considered legitimate. The iForest method directly detects anomalous instances without modelling legitimate instances. The advantage of the iForest method is that it is likely to complement existing statistical methods that model legitimate instances. Furthermore, the iForest method has proven to be a great tool to detect irregular healthcare providers by using a small number of features as input, using only features in which deviations are likely correlated with fraud.

The Netherlands healthcare system provides access to care for all inhabitants through compulsory medical insurance. The services are provided through health insurers that make specific arrangements with healthcare providers. In 2018, annual healthcare costs in the Netherlands totalled €100 billion, with almost half of this amount accounted for by health insurers.[2] Health insurance companies have therefore extended the governance in place to manage the financial transactions related to these healthcare costs. To manage and reduce the risk of unlawful payments, insurers perform checks and controls on invoices. This allows insurers to identify mistakes or fraudulent payments, helping to reduce unnecessary costs, as ultimately the costs of fraud are paid for by the population through higher premiums.

This research is based on the fact that fraudulent behaviour oftentimes manifests itself by different patterns than what would actually be observed in valid behaviour. With the use of different methods that each tackle a particular type of irregular behaviour, more irregularities can be detected earlier in automatized ways. The application of Benford's law and the iForest provide more insight into the types of potential fraud to detect and categorise each type of potential fraud more effectively. Unlawful payments detected by Benford's law are

[1] NOS. (2020, 2 January). Controle zorgfraude faalt, OM en verzekeraars luiden noodklok. Accessed at https://nos.nl/nieuwsuur/artikel/2317099-controle-zorgfraude-faalt-om-en-verzekeraars-luiden-noodklok.html

[2] CBS. (2019, 2 June). Zorguitgaven; kerncijfers. CBS (StatLine). Accessed at https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84047NED/table?dl=1D372.

likely due to fabricated invoices. The iForest implementation works on detecting deviations in a mix of 17 attribute values. Irregular healthcare providers (i.e., healthcare providers with irregular invoice behaviour) are expected to have a different mix of attribute values from the majority of the data. The attribute values are statistical values of attributes that are often related to fraud (e.g., deviations in age of patients, billing dates and billing amounts). Therefore, healthcare providers flagged as irregular by the iForest are likely a result from deviations in the attribute values.

In this research, the detection of irregular healthcare providers using Benford's law and iForest models are compared with a baseline model that detects irregular healthcare providers at random. The p-value is approximated for the null hypothesis that the number of irregular healthcare providers detected by each model is less than or equal to that detected by the baseline model. It is found that Benford's law and the iForest can detect significantly better than randomly selecting irregular healthcare providers, with p-values of 0.123% and $6.84 \cdot 10^{-5}$%, respectively.

Out of the approximately 5000 healthcare providers, each model selected and ranked 50 healthcare providers that were most likely to be fraudulent according to the two methodologies. Fraud inspectors then further investigated these flagged healthcare providers. Based on the occurrence of irregularities within the total population (estimated ~2%), a random model would be expected to have one true positive flag and 49 false positive ones. A true positive is a healthcare provider that is indeed flagged to have irregularities after further investigations by fraud inspectors. Our models were able to detect 11 true positives: two by Benford's law and nine by the iForest.

Interestingly, the two models detected different irregular healthcare providers and hence complement each other. Moreover, it is worth mentioning that the two true positives in the Benford's law model were ranked in its top three, whereas

none of the nine true positive ones in the iForest model occurred in its top three. Therefore, we conclude that Benford's law is especially powerful where the sample size is limited and the number of false positive needs to be very low, while the iForest model is more powerful if the sample size is larger and false positive ones are less of a problem.

VGZ is currently adding the two aforementioned models to a dashboard, so that fraud investigators can visually observe potentially irregular healthcare providers according to these models. An implementation of these models could result in a more automated and effective way to tackle irregularities oftentimes related to insurance fraud, reducing manual effort. Ultimately, an efficient and effective system that reduces insurance fraud will ultimately lead to lower premiums and the securement of quality healthcare.

Analysis of results: Benford's law and the iForest model

- Population size about 5000 healthcare providers
- Sample size 50
- Based on the assumed fraud occurrence in the population a random model expected to have one true positive flag based on an expected 2% irregular healthcare providers
- A true positive is a healthcare provider that is indeed flagged to have irregularities after further investigation by fraud inspectors
- Benford's law three true positive flags out of 50, of which two occurring in its top three, whereas iForest model nine true positives ones out of 50, but none occurring in its top three
- No overlap in true positives between the models

Milliman

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

Kendall Carolissen
kendall.carolissen@milliman.com

Judith Houtepen
judith.houtepen@milliman.com