EpiMetrics
FEBRUARY 2018 VOLUME 3 ISSUE 5

PRESCRIPTIONS

POLICY BRIEF

In a Sea of Health Insurance Claims, Machines Find the Outliers

With millions of claims processed each year, the Philippine Health Insurance Corporation Inc. (PHIC) runs the risk of disbursing funds to claims that are either fraudulent or in violation of policy. The company relies on experience to determine if a single claim is for audit, then assigns cases to auditors for investigation. With numerous claims to evaluate, PHIC does not have the manpower to thoroughly audit each claim. To help increase the speed and accuracy of flagging unusual claims ("outliers"), this study created an outlier detection model using machine learning. The study used machine learning methods to select outliers for auditing. These outliers must then be audited or validated if they are truly fraudulent or perhaps invalid in some way, which will pave the way for PHIC to create better decision support tools for auditors, and lay the groundwork for supervised machine learning for fraud cases.

POLICY LESSONS

Machine Learning detects outliers from millions of claims

Using machine learning or computer programming, outliers can easily be sifted through millions of claims without creating a burden on human resources and time. PHIC must then validate these results in order to train future machine learning models to recognize fraud, bad form entry, and other true causes for the outlier behavior.

Build a database of verified fraudulent claims

Despite the use of Unsupervised Machine Learning for outlier detection, Supervised Machine Learning is still the standard method of doing this form of artificial intelligence. A training dataset with a large number of cases verified as fraud or not fraud is required for machine learning programs to learn and detect future cases of fraud. As more examples of claims fraud are presented to the model, the more it can improve its capacity to detect fraud.

Unsupervised methods are particularly apt in situations like these, where cases of true fraud are far fewer than normal claims. If PHIC attempted to create a training dataset by randomly sampling then evaluating cases, the process would be extremely time consuming as most of the cases evaluated would be not fraud, and one would need a minimum number of fraud cases in order to train the model.

But first, define fraud

One of the reasons why a training data set has yet to be provided is the lack of successful convictions filed by PHIC. Given the legal definitions of fraud, no samples can be categorized as such. PHIC must establish an operational definition of fraud.

Invest in the necessary Information Technology infrastructure

Since PHIC has a large database of claims and other insurance data, it is necessary to invest more in building their technology infrastructure for data storage and processing. This may be executed via increasing on-premise installations or increasing utilization of cloud platforms. Increased cloud specifications in terms of processing and memory will allow for more claims to be processed with more robust machine learning models.

No need to reinvent the wheel

The Medical Post-Audit Module (MPAM) is a potentially robust data set that allows medical auditors to classify if a claim is possibly fraudulent or in violation. Unfortunately, data fields have been disabled to make claims processing more efficient. The reactivation of the MPAMs full capacity will allow for the creation of a data set that can at least be used for the detection of possible fraud.

INTRODUCTION

The Philippine Health Insurance Corporation (PHIC) is the country's foremost agency when it comes to financing universal health care (UHC) in the Philippines. It has steadily improved population coverage and financial risk protection for all Filipinos.

With billions of pesos being lost by PHIC annually to fraudulent claims, machine learning models and a decision support system can help PHIC identify and detect questionable transactions for investigation and subsequent action. Money recovered from prevention of fraud can then be invested in other efforts to achieve UHC, such as increasing the range of services covered, including all outpatient consultations, and increasing the percentage of costs covered per claim.

METHODS

This study is a proof of concept that Machine Learning can be used to detect outlier claims. Outliers detected can then be audited to determine if they are fraudulent or not.

Claims on Pneumonia and Cataract from 2014-2015 were analyzed using five (5) Unsupervised Machine Learning models: Isolation Forest, Principal Component Analysis Outlier Detection, Local Outlier Factor, HDBSCAN Clustering, One-Class Support Vector Machine for Outliers. These were then made into a majority ensemble for classifying claims as an outlier or inlier. Microsoft Azure was the cloud computing service utilized for this study, a service aligned to what is currently implemented by PHIC. After the data was processed, analysis for outlier detection took a total of 10 weeks.

RESULTS

Out of 2,228,800 claims for Moderate Risk Pneumonia, 2,127 (0.095%) were detected as outliers. Out of 171.517 claims for High Risk Pneumonia, 1,335 (0.78%) were determined as outliers. Out of 351,500 Cataract (requiring CPSA) Claims, 695 (0.198%) were determined as outliers. These outliers are far more likely to contain cases of fraud and other irregularities compared to a randomly sampled subset of claims. This means improved speed and efficiency in auditing enough cases to create a subset of 1,000 cases of confirmed fraud, which is a prerequisite for fraud minimization and flagging work.

To maximize the output of this project, it is recommended that these claims are investigated to determine if they are in fact candidates for audit.

Table 1. Outliers Detected Per Case

Disease / Facility level	Number of claims	Number of outliers (%)
Moderate risk pneumonia	2,228,800	2,127 **(0.095%)**
High risk pneumonia	171,517	1,335 **(0.78%)**
Cataract requiring CPSA	351,500	695 **(0.198%)**

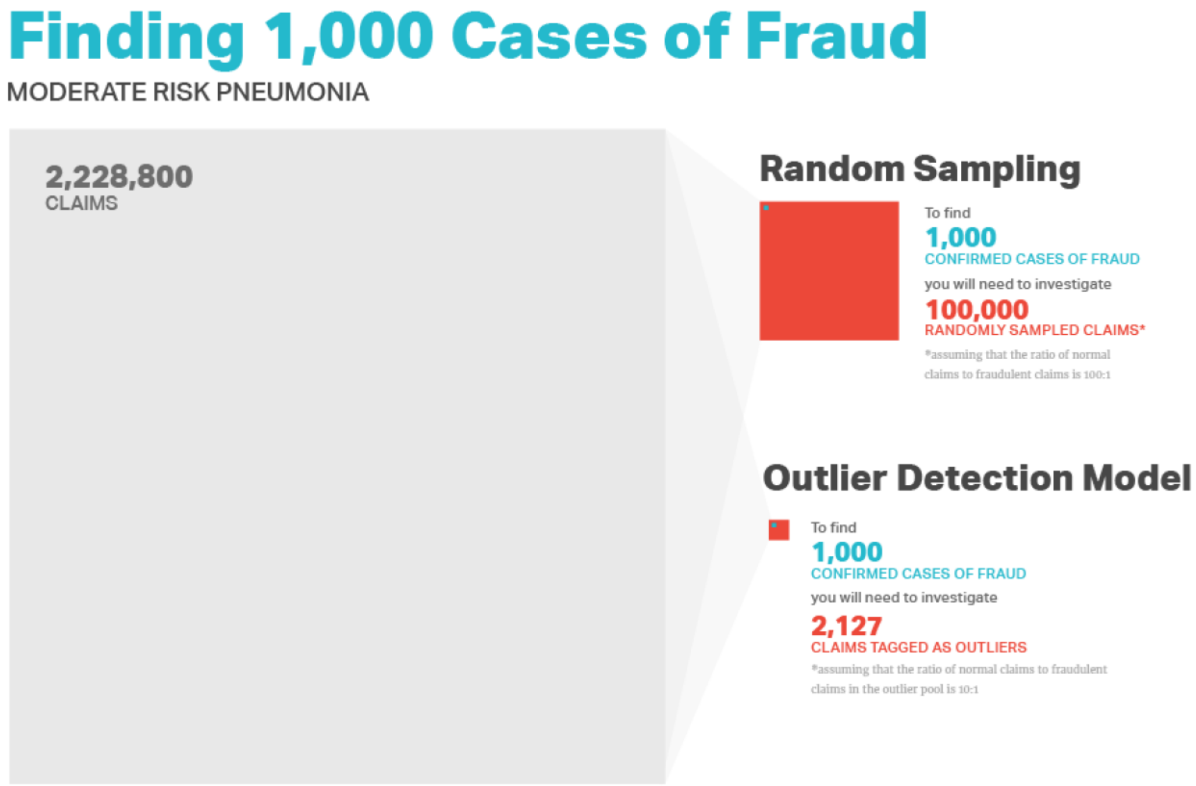

Figure 1. Sample Outlier Detection for Moderate Risk of Pneumonia

CONCLUSION

Given the results of the study, there is now a high-potential dataset of 4,157 outliers that contains proportionally far more cases of irregularities compared to a randomly sampled subset of all claims. By manually auditing these claims, PHIC can quickly generate a robust number of cases of true irregularities, which can be used for decision trees, fraud detection models, and so on.

The models generated as part of the study means that it is technically feasible for claims to undergo an unsupervised machine learning analysis to detect outliers for possible audit. Higher cloud computing specifications would allow for larger volumes of claims to be processed in a shorter duration of time.

RECOMMENDATIONS

As this is a proof-of-concept fraud detection model, the following steps are recommended for PHIC to undertake:

1. Work on a viable training data set. This entails:

- Developing a set of objective criteria, if possible, on how PHIC auditors and supervisors can determine if an outstanding insurance claim is fraudulent or not.
- Conducting audits and inspections on the outliers to confirm the accuracy of the model predictions.
- Building a database of verified fraudulent claims and flag the disease-specific features that contributed to a specific claim being tagged as fraud.
- Utilizing the Medical Post Audit Module (MPAM), to serve as a rich data source for the detection of probable Fraud. Improve policy on its implementation.
- Validating the outliers (current results) internally.

2. Fine-tune current processes for Unsupervised Machine Learning to Detect Outliers and future fraud detection efforts. This entails:

- Doing a rerun on the models and tuning the hyperparameters of the outlier detection model to increase its accuracy in predicting whether a new claim will be fraudulent or not.
- Encoding Item 7 for all claims within Claim Form 2. As a stopgap, the Item 7 of a representative sample of claims can be encoded by going back to the paper Claim Form 2.
- Developing standard text descriptions for entries in Item 7. For electronically submitted claims, this can be implemented using text input suggestions.

3. Invest in the necessary Information Technology Infrastructure both in terms of hardware, software and personnel. This entails:

- Having the necessary memory and processing capacity to handle millions of claims.
- Implementing an IT infrastructure with a varied ecosystem to handle a variety of supervised and unsupervised machine learning models.
- Training the necessary personnel to maintain this system. This may also entail procuring a 3rd party cloud-based ecosystem with comprehensive post sales service and support.

Authors:
Carlo Emmanuel L. Yao, MD-MBA
Stephanie Sy

Original Research:
Wong, J., and Yao, C. (2018) *Three Use Cases for Big Data in Universal Health Care: Social Determinants of Health, Service Delivery Network, and Fraud Detection*. Makati City: EpiMetrics, Inc.

Recommended Citation:
Yao, C. (2018) 'In a Sea of Health Insurance Claims, Machines Find the Outliers', *Prescriptions*, 3(5). Makati City: EpiMetrics, Inc.