Exploration of Data Science Toolbox and Predictive Models to Detect and Prevent Medicare Fraud, Waste, and Abuse

Adam Canton1, Eli Fuller2, Benjamin Goodwin3, Babatunde Olanipekun 4, Chris Papesh5

1 Master of Science in Data Science, Southern Methodist University,  
Dallas, TX 75275 USA

{eli.fullerin, ovens, chris.papesh}@unlv.edu

{bgoodwin, cantona, bolanipekun}@smu.edu

**Abstract.** The Federal Department of Health and Human Services spends approximately $700 Billion annually on Medicare of which an estimated $30 to $110 Billions some form of fraud, waste, or abuse (FWA). Despite the Federal Government’s ongoing auditing efforts, fraud, waste, and abuse is rampant and requires modern machine learning approaches to generalize and detect such patterns. New and novel machine learning algorithms offer hope to help detect fraud, waste, and abuse. The existence of publicly accessible datasets complied by The Centers for Medicare & Medicaid Services (CMS) contain vast quantities of structured data. This data, coupled with industry standardized billing codes provides ample opportunity for the application of machine learning for fraud, waste, and abuse detection. This research aims to develop a new model utilizing machine learning to generalize the patterns of fraud, waste, and abuse in Medicare. This task is accomplished by linking provider and payment data with the list of excluded individuals and entities to train an algorithm on previously fraudulent behavior. Placeholder sentence for the main result. Placeholder sentence for main conclusion.

1 Introduction

Medicare fraud, waste, and abuse (FWA) is a problem on the national scale, causing an enormous burden (e.g, loss of billions of dollars, Medicare Learning Network, 2021) on public finances and the allocation of resources for some of the most vulnerable populations (Medicare Learning Network, 2021). Actions that constitute FWA are poorly defined, but The Centers for Medicare & Medicaid Services (CMS) provides a list of the kinds of claims that are candidate for FWA (Medicare Learning Network, 2021). Despite a multitude of prior studies concerning the topic of FWA, an exact amount of Medicare FWA is difficult to closely approximate because much of the FWA simply goes undetected. Conservative estimates claim 3-10% ($19 Billion to $65 billion) of all medical claims fit into FWA (Bauder & Khoshgoftaar, 2018, p. 9), while other estimates claim the number is much higher and somewhere closer to $300 billion (Nicholas, Segal, Hanson, Zhang, & Eisenberg, 2019, p. 788). Medicare is a large complex government program and in its current form has few controls in place to detect FWA. Unlike private insurers who use complex algorithms to detect FWA and subsequently deny such claims before they are paid, Medicare operates in an opposite fashion, paying providers first, and then investigating supposed FWA (Pande & Mass, 2013, p. 10). The Centers for Medicare & Medicaid Services (CMS) provide publicly accessible data containing fields such as provider information and payment data. CMS also maintains a list of excluded providers, The List of Excluded Individuals and Entities (LEIE). The LEIE contains the individuals and entities reprimanded for FWA. CMS data provides an ample number of resources and records to adequately train and develop a machine learning model to disseminate between legitimate Medicare claims and those considered FWA. The potential cost and labor savings for the Department of Health and Human Services from FWA detection via machine learning is significant.

Despite the variability of dollar estimates for FWA, one theme is clear, Medicare suffers from rampant FWA, and these costs are absorbed by the Federal Government as well as the taxpayers of the United States. CMS provides publicly available data on Medicare through various applications and databases and with this data this study proposes implementation of machine learning algorithms to generalize and detect FWA in current and future claims.

Machine learning techniques can complement the efforts of the Department of Health and Human Services in the investigation of potentially fraudulent claims. Such machine learning algorithms could serve as a first layer of detection, which human claims inspectors can further analyze to determine if the claim is a potential candidate for FWA. This step in the investigative process has the potential to save hundreds of thousands of man hours per year as the algorithm has the capability to classify claims as “potentially fraudulent” or “not potentially fraudulent,” saving the humans from the task of manually reviewing each Medicare claim. FWA are often difficult to detect and pursue and the costs of pursuing individual cases many not always be worth the expense, especially in the case where the FWA is a legitimate provider making an honest mistake, or even the case when a dishonest provider only slightly overbills, invoking an investigation costing more than the FWA itself. In the scenario where an algorithm can make this first determination has the potential to save tremendous amounts of money and investigative labor.

The subject of FWA is an often researched and discussed problem within academia and beyond (Pande & Mass, 2013, p. 9). The topic has even finally reached the point of Federal Government admission, with President Obama making the elimination of healthcare FWA a top priority of his administration (Pande & Mass, 2013, p. 10). The Federal Government announced in 2011 that it would include predictive data modeling to assist in the fight against FWA prior to the payment of claims (Pande & Mass, 2013, p. 10). Yet despite this public affirmation of the issue of FWA, 12 years later the problem persists and my many accounts, only continues in increases prevalence (Pande & Mass, 2013, p. 9).

Among the various issues associated with solving such a problem revolves around access to data. Despite the relatively large amount of publicly accessible data, this research is tasked with determining which data best pertains to the topic at hand and whether the data answers the question of interest. The CMS offers robust publicly available data and will be utilized throughout the duration of this study.

The research team believes research in the field of Medicare FWA combined with machine learning will yield statistically significant and insightful results. Applying machine learning techniques to Medicare data, implemented according to the methodologies reported above can and will identify situations where fraud, waste, and abuse are present.

2 Literature Review

Much of the previous research on the topic of Medicare FWA have attempted to generalize fraud and train machine learning models to detect providers who commit Medicare FWA (e.g., Musal, 2010, Liu et al., 2016, Herland et al., 2017, Zhang & He 2017, Bauder & Khoshgoftaar 2018, Obodoekwe & van der Haar, 2019). Specifically, each prior paper has targeted a specific aspect of Medicare where the authors believe FWA could potentially be generalized, such as at the provider level (Bauder & Khoshgoftaar 2018, p. 9), or after the services have been rendered, at time of payment to the provider (Pande & Mass, 2013, p. 10). The researchers believe the prior studies have done an adequate job generalizing specific tenants of the problem but contain several shortcomings that need to be addressed. These shortcomings include, using multiple years of available data, different choices of classification algorithms, and linking the excluded providers with the claims data. The bulk of this analysis is centered around using machine learning to determine if a given claim is legitimate or FWA.

This research aims to use machine learning to identify possible fraudulent trends or activity using public Medicare Data. This research plans to use 3 years of Medicare Part B and Part D data along with labels collected from the LEIE to generate some insights on Medicare FWA. Since occurrences of FWA are rare when compared to legitimate claims, the researchers plan to mine these cases for insight and then use unsupervised techniques to identify FWA behavior.

The data sets are quite large and combining them presents many challenges. The public Medicare data sets are not released at the patient/event level. They are instead aggregated on National Provider Identifier (NPI) (a unique identification number assigned to all covered health care providers) and other characteristics (Bauder & Khoshgoftaar 2018 p. 2). This means the researchers will have to join some large datasets together to examine the relevant features required for this analysis. Anomaly detection, one of the preeminent methods of fraud detection is employed in many different areas such as procurement fraud, credit card fraud, and Medicare fraud detection (Zhang & He 2017 p. 310). The assumption being that anomalous events or activity is likely to be fraudulent when compared with the rest of the body (Bauder & Khoshgoftaar 2018 p. 3). Previous researchers have used Spatial Density using imLOF (Improved Local Outlier Factor) (Zhang & He 2017 p. 311). As well as unsupervised methods such as Isolation Forest and Unsupervised Random Forest (Bauder, Rosa, & Khoshgoftaar 2018 p. 285), Deviation Clustering, Gaussian Mixture Models, and Bayesian Co-clustering (Ekina, Leva, Ruggeri & Soyer 2013 p. 151). Further, past researchers have seen that Local Outlier Factor (non-improved), K-Nearest Neighbors, and autoencoders are suboptimal performers (Bauder, Rosa, & Khoshgoftaar 2018 p. 286). Though there is some discussion over LOF (Bauder, Rosa, & Khoshgoftaar 2018 p. 287), (Zhang & He 2017 p. 310). The researchers involved in the study "An Anomaly Detection Method for Medicare Fraud Detection" designed a new LOF metric designated imLOF for improved Local Outlier Factor. This metric is designed to detect excessive medical treatment and decomposing hospitalization using spatial density information. (Zhang & He 2017 p. 312) The original measure, developed by Breunig et al. (Breunig, Kriegel, Ng, & Sander 2000 p. 94), gives anomaly scores based on the density of observations; they noted that the density of anomalous events would be less than that of its normal neighbors. This metric has issues with small clusters making it suboptimal for healthcare use (Bauder, Rosa & Khoshgoftaar 2018 p. 288), (Zhang & He p. 312). The authors give an example of a small cluster of hypertension patients with a great deal of fraud. Since the cluster is small and the point’s neighbors are also likely to be anomalous the metric scores a low chance of anomalous activity. The authors then suggested an improvement to the LOF score by adding the size of a cluster into consideration instead of only density, with the additional use of the DBSCAN algorithm the improved LOF score performed much better on healthcare data (Zhang & He 2017 p. 312)

Some research has only focused on single years (Bauder, Rosa, & Khoshgoftaar, 2018 p. 288) (Gordon & Siegel 2020 p. 1), (Hancock & Khoshgoftaar 2020 p. 572) and used either a Supervised Learning design or a combination of unsupervised and supervised (Bauder, Rosa & Khoshgoftaar 2018 p. 288) (Meyers 2017 p. 251) Our design will follow the latter. There have also been studies directed at specific portions of Medicare/Medicaid such as dental, otolaryngology (Ekina, Leva, Ruggeri & Soyer 2013 p. 151) and dermatology (Gordon & Siegel 2020 p. 1). The use of a single year is generally due to the size of the data, given that a single year’s Medicare Provider Utilization and Payment Data for Part B is around 10 million records, 29 columns, and about 3 GB of memory by itself.

An issue with the current data is proper class balance in distribution of the target classes. This was an issue in all of the studies that used supervised methods. Such severe imbalance requires careful data adjustments to somehow increase the representation of the minority class, in this case fraudulent activity. This means some type of special sampling method. Bauder et al. (2018) indicated that random under sampling provided the best results followed by SMOTE (Synthetic Minority Oversampling Technique) (Bauder & Khoshgoftaar 2018 p. 3). In random under sampling (RUS), the algorithm throws out random events from the majority class, thereby increasing the representation of the fraudulent cases. In the case of SMOTE, a more advanced algorithm is used to create new minority class instances by first finding a minority class instance and its k nearest neighbors. Then a new instance is created by choosing a random neighbor and combining it with the original instance. In a further study by Bauder and Khoshgoftaar an RUS method was used along with an adjusted cost function (Bauder and Khoshgoftaar 2018 p. 5)

This research will follow heavily in the footsteps of prior research in the field of Medicare FWA. Modern machine learning and data exploration techniques will be exploited for the purposes of better understanding the divers and factors behind Medicare FWA. Several of the techniques and methodologies referenced above will be modified and adapted for the purposes of this research into Medicare FWA.

Using machine learning algorithms, prior research has indicated FWA is possible to detect, but often these prior studies are inconclusive in their test for statistical significance when testing against the hypothesis of a difference between a fraudulent claim and a legitimate claim. This is primarily due to a handful of challenges split between Medicare data and the current limitations of machine learning. The first challenge concerns the balance of legitimate claims against illegitimate ones. For example, if 3-10% ($19 billion to $65 billion) of all medical claims fit into FWA (Bauder & Khoshgoftaar, 2018, p. 9), approximately 90% of claims are legitimate. This indicates that a machine learning model could perform reasonability well by simply classifying all transactions as legitimate. Class imbalance is a significant issue with Medicare data and will require new and novel approaches to overcome. Fortunately, credit card companies and other large organizations that process huge volumes of transaction data have deeply studied such topics and have developed formidable and complex anti-fraud and theft systems. Unfortunately, much of implementation of these systems is proprietary, nonetheless this indicates that large class-imbalances can be overcome. In addition to class-imbalance, the issue of training data presents a significant barrier. Medicare claims data is just that, information on claims and aggregated to the procedure and National Provider Identifier Standard (NPI), there is no comprehensive database of claims data complete with an indicator if the transaction is a legitimate one or categorized as FWA. The proposed solution involves integration of claims data and the excluded provider list. The intention is to examine those who populated the excluded provider list, and then determine the last year they submitted claims, and find those claims, add the claims and provider to the training data to train the model based on the claims behaviors of known excluded providers. Adding the excluded providers in addition to the legitimate claims will provide the model with the ability to differentiate between the behaviors of FWA and real claims data.

Fraud related activities in Medicare claims include:

* claims for appointments that was not attended by the patient;
* claims for more complex services than those performed or required;
* Claims for services that that were not carried out (Johnson & Khoshgoftaar 2019, p. 18).

On the other hand, medicare abuse includes the practice of knowingly providing medically unnecessary services to patients against recognized standards. For example, misusing billing codes for personal gain. There are applicable Federal laws that prohibit Medicare fraud and abuse. These include the False Claims Act (FCA) and Anti-Kickback Statute (Johnson & Khoshgoftaar, 2019, p. 18).

Bauder et al. (2018) applied several anomaly detection techniques to segment medical provider fraud in the 2012 to 2015 Medicare Provider Utilization and Payment Data: Physician and Other Supplier which is publicly available from the Center for Medicare and Medicaid Services. To evaluate the performance of candidate learners, the authors mapped fraud labels dataset using the List of Excluded Individuals and Entities (LEIE). The novelty in their study is the application of Isolation Forest and Unsupervised Random Forest on this big Medicare dataset. Bauder et al., 2018 worked with only half of their preprocessed dataset as the reduced the dataset from 3.7 million to 1.8 million due to hardware limitations.

Prior research on Medicare FWA has studied algorithms such as XG-Boost, CatBoost, and Gradient Boosted Decision Trees (Hancock & Khoshgoftaar, 2019, p. 572). While extremely useful the machine learning space, these algorithms have proven to be inconclusive on Medicare FWA (Hancock & Khoshgoftaar, 2019, p. 578). Other techniques such as regression and clustering analysis have been deployed with comparable results (Musal, 2010, p. 2828). Chief among these algorithms is a basis in frequentist statistics. Looking at Medicare FWA through a Bayesian lens could provide the missing link between sporadic classicization and a clearly defined approach to detecting FWA. The notion of prior probabilities can help better train models, and approach the data with a unique perspective, treating both model parameters and data as random. As the algorithm evaluates an individual claim, the assistance of a prior probability related to its legitimacy has the potential to facilitate updating the likelihood of flagging transactions as FWA or legitimate with more precision. Better performance was reported by Johnson and Khoshgoftaar 2019 who reported AUC score > 85% from similar Medicare dataset. They tackled the class-imbalance with random over-sampling and random under-sampling techniques prior to fitting it on a 3-layer dense neural network. Perhaps Bauder et al., 2018 could have achieved better performance by using random under-sampling and/or random over-sampling to manage the severe class-imbalance in the dataset. Other methods that have been applied to the Medicare dataset for fraud detection. Liu et al., (2016) utilized isolation forest; Bauder et al., 2016 compared supervised learning techniques: Gradient Boosted Machine, Random Forest, Deep Neural Network, and Naive Bayes and a suite of unsupervised learning techniques: autoencoder, Mahalanobis distance, KNN, and local outlier factor, and hybrid (multivariate regression and Bayesian probability) machine learning approaches.

3 Data

Current study will use public Medicare data available from CMS – The Center for Medicare and Medicaid Services. The Medicare Provider Utilization and Payment Data, hereafter referred to as MPUPD, contains Medicare data aggregated to the NPI-procedure level. The clustering and other unsupervised methods are applied to this data set to attempt to find fraudulent patterns. However, the research plans to use a supervised model first to find some guidance in our unsupervised methods. To achieve this, the team will use the List of Excluded Individuals and Entities. This list is updated monthly by CMS and the OIG (Office of the Inspector General) to reflect any actions taken against individuals or entities committing fraud, waste, or abuse towards the system. This can cause issues for supervised learning methods because the LEIE is aggregated only to the NPI level. Thus, to get a one-to-one relationship, the MPUPD data must be aggregated up to the same level (Bauder & Khoshgoftaar 2018 p. 2). The team plans to conduct a multi-year study using the MPUPD part b and part d data ranging from 2016 to 2018. We will then use these insights to evaluate the 2019 data. The team also considered using the Medicare Open General Payments Data set. But this dataset did not contain the correct features to be joined to the List of Excluded Individuals and Entities. These sets were also double and triple the size of the MPUPD. In the future, depending on what the researchers find to be relevant features in the data may be useful to apply our techniques to this dataset and other in the MPUPD that were not included in this study.

4 Methods

Using modern and novel approaches to machine learning, the goal is to analyze publicly accessible Medicare data to determine patterns behind FWA. While the algorithms play an extremely key component in this research, the dissection of claims data and subsequent paring with excluded provider information is a unique approach to the problem of detecting Medicare FWA. Pairing claims data with excluded provider information is a key differentiator between this study and those proceeding and should result in significantly different results.

Modern machine learning often involves a handful of algorithms such as XGBoost, Random Forests, and Boosted Gradient Descent. An often underutilized and poorly understood approach to problems in machine learning is a Bayesian approach. Bayesian statistics take a fundamentally different approach to data and parameters when compared to frequentist statistics. Bayesian statistics consider both the data and the parameters to be drawn from random distributions, while frequentist statistics only consider the data to be random. This is a key distinction because Bayesian statistics include notions of prior probabilities that can help define otherwise unknown distributions. In the case of detecting Medicare FWA, the Bayesian perspective can approach the detection problem with an understanding of the claim being FWA based on an assigned prior probability. As the chosen algorithm evaluates the claim at hand, the prior probability is updated as the algorithm is fed more and more data, with the intention of updating its belief in the legitimacy of the claim as it evaluates a given claim. This approach also impacts the algorithms view of the data, when each claim is determined to be either FWA or legitimate, the algorithm will recall this global prior probability of FWA occurring and being evaluating each claim as FWA or legitimate with this new understanding of its prior probability. The prior probabilities will be determined using both Markov-Chain Monte Carlo and simple statistics to examine prevalence of FWA within the data.

The best model Local Outlier Factor produced a 63% Area Under the Curve, just 13% higher than random. As also averred by the study, such a low score makes it difficult to deploy this study for real world application. Situations like this usually demands scrutiny of the data and its source/s. One explanation for the low AUC scores could be a lack of known fraudulent providers to use as fraud labels for validation, creating a highly imbalanced dataset (Bauder & Khoshgoftaar 2017, p. 860).

This research will use entirely open-source models and publicly accessible data to embrace transparency in claim evaluation criteria. As a final note, in addition to the novel Bayesian approaches to this problem, the research team will use an assortment of calibration models such as random forest, and Boost to determine baseline performance for determining FWA.

Machine learning is a subjective science and criteria for success can be measured on any number of levels. Since this data is extremely imbalanced and an algorithm that simply classifies all claims as legitimate would perform at around 90% accuracy. As a solution to this problem, F1 score will be used as the primary metric for success, F1 is an evaluation metrics that balances precision and recall. F1 is the harmonic mean of both precision and recall and will balance the overabundance of legitimate claim data with the need to detect and accurately classify transactions categorized as FWA. For the purposes of this research, the threshold for a successful algorithm will achieve an F1 score of >0.6. In addition to the F1 metric, the research team will also look at the contextual implications of additional criteria to evaluate performance of different algorithms. As an example, XGBoost and a Bayesian classifier operate under fundamentally different assumptions, and despite using one as a baseline, the research team will need to look at a variety of different metrics across the different classifiers.

5 Context Architecture

(Placeholder) This will be filled out as the research team, as data architecture is determined.

6 Results

Through the research of Medicare data, the research team hopes to generalize the common occurrences of fraud, abuse and overbilling throughout the Medicare system using machine learning. The research team will determine “good” results using various statistical tests as well as participating in a continuous feedback loop with stakeholders who regularly handle and identify fraud, abuse, and overbilling.

Exploratory Data Analysis

The dataset used for this project contains ?? Instances and ?? Attributes. These were aggregated to show the distribution of claims among the uniquely identified providers (Figure 1). Most of the claims studied occurred are categorized in the diagnostic radiology and internal medicine. Interestingly podiatry, gastroeterology and urology, respectively, have the least claims (Figure 2). To provide a perspective on the claims submitted, in 2019 (Figure 3), 2.59 billion services were performed on 879 million individuals at a cost of $3.81 billion. Of these charges, medicare made a payment of $816 million. Of particular interest is the distribution of providers in the LEIE. Figure 4 shows that the proportion of these individuals correlates with the population distribution of states in the USA. For instance, the most populous states of the United States (California, New York and Florida) records the highest number of providers found in the Excluded List.

Fig 1: Aggregations Over National Provider Identifier (NPI) 2017-2019

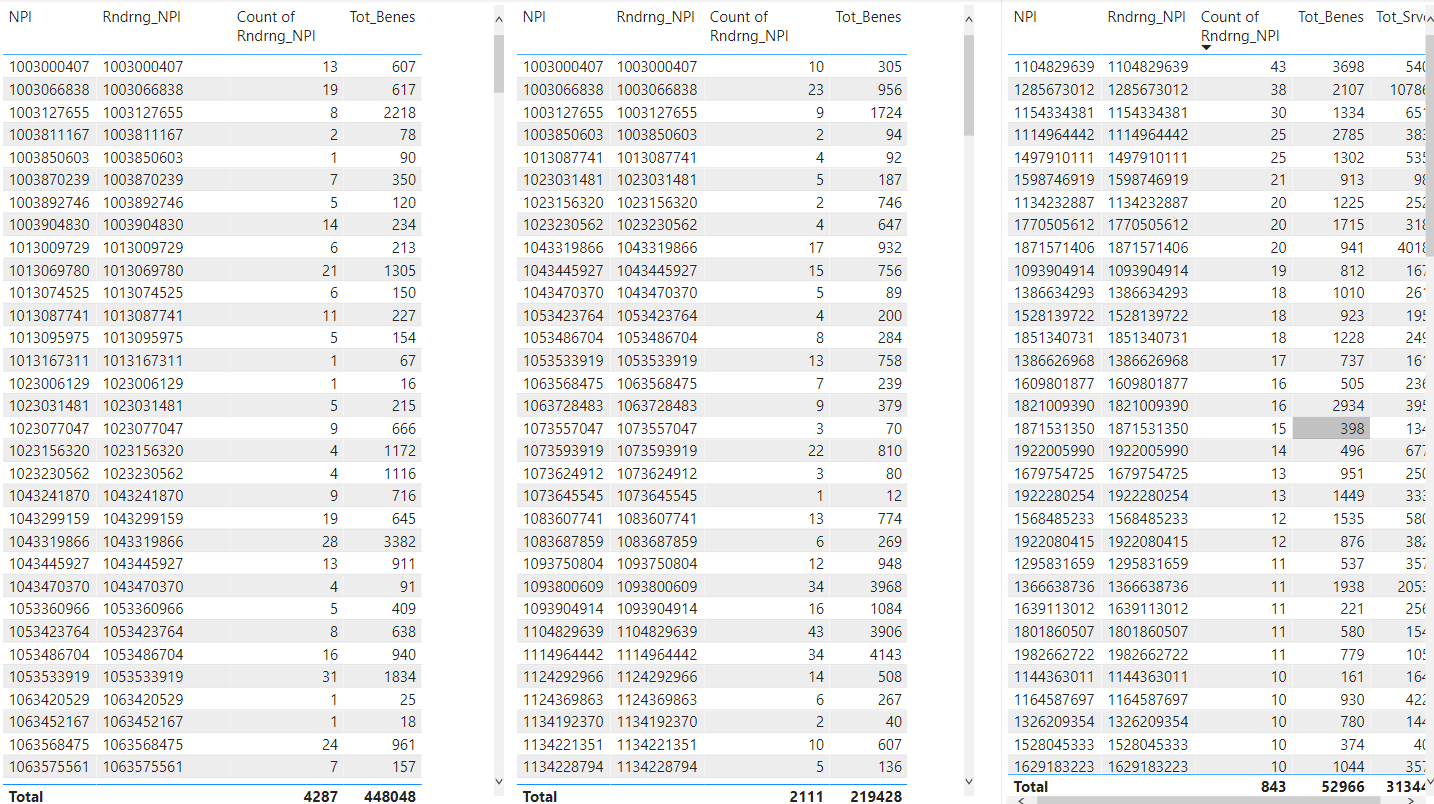


Fig 2: HCPCS Codes by Provider Type – an indication of volume 2019

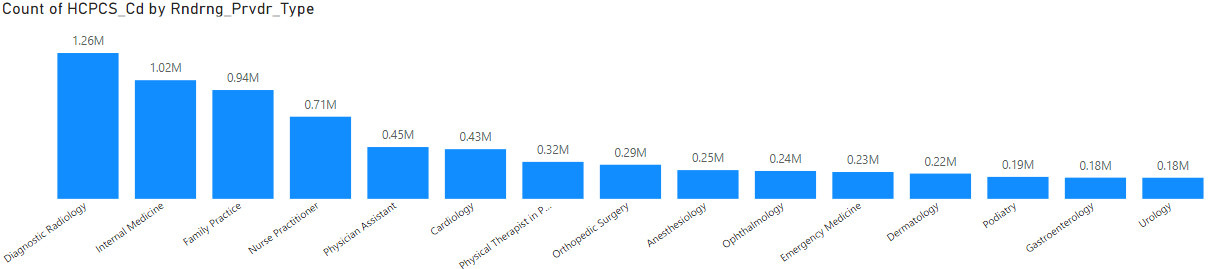


Fig 3: Some Summary Stats from 2019 MPU and PD

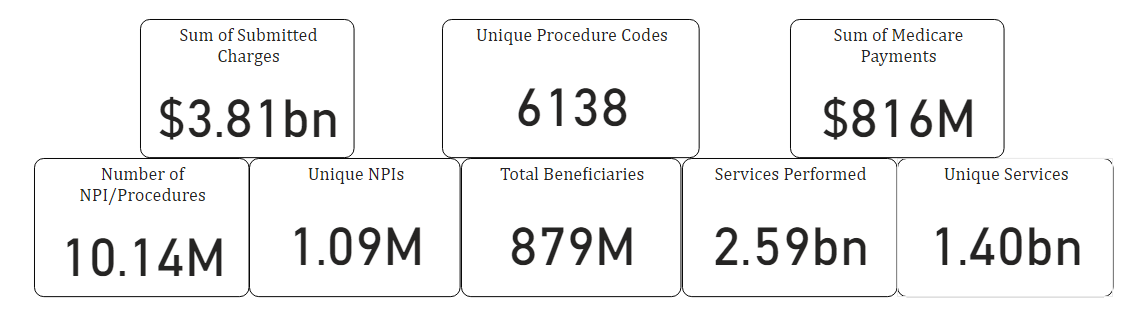


Fig 4: Excluded Individuals and Entities by Year (Compare to NPIs in 2019)

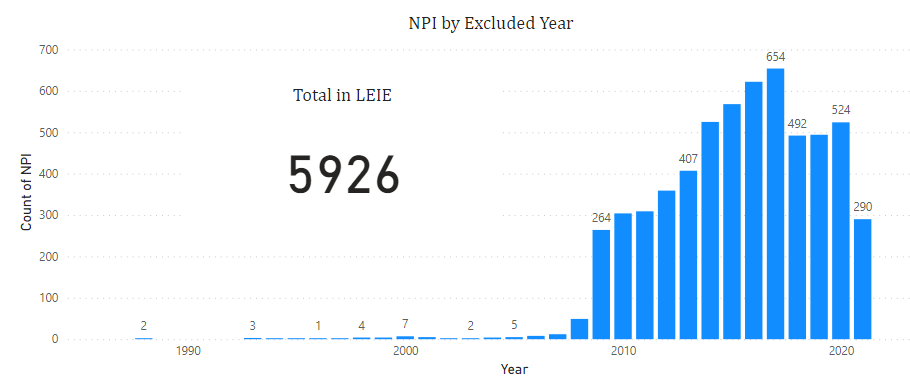


Fig 5: Excluded Individuals and Entities by State

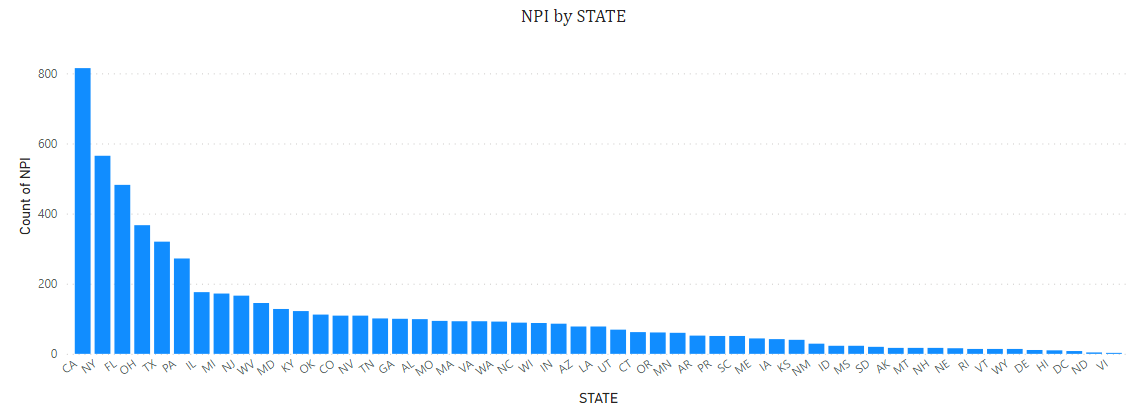


Fig 6: Differences and Errors in Credentials and Specialties MPUPD 2019 v LEIE

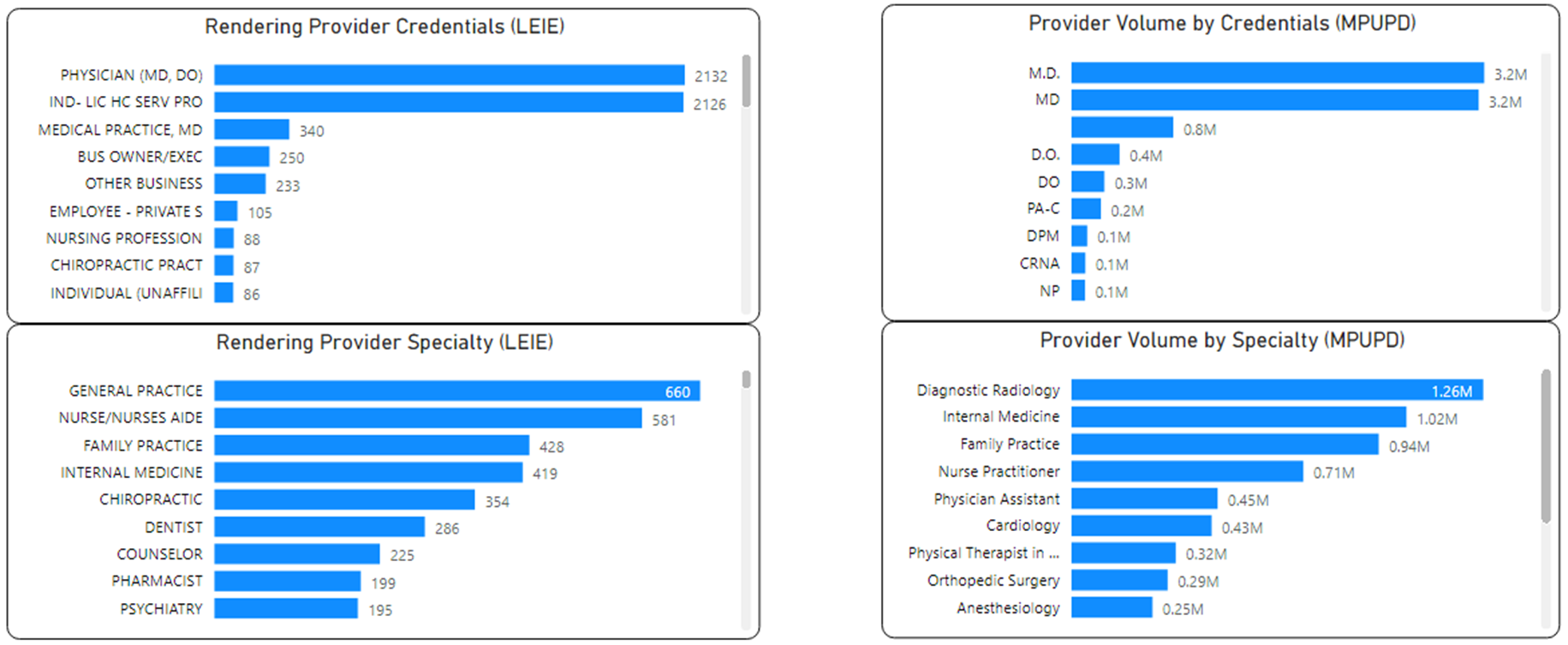


Fig 7: Word Cloud of HCPCS Code Descriptions (Stop Words Removed)

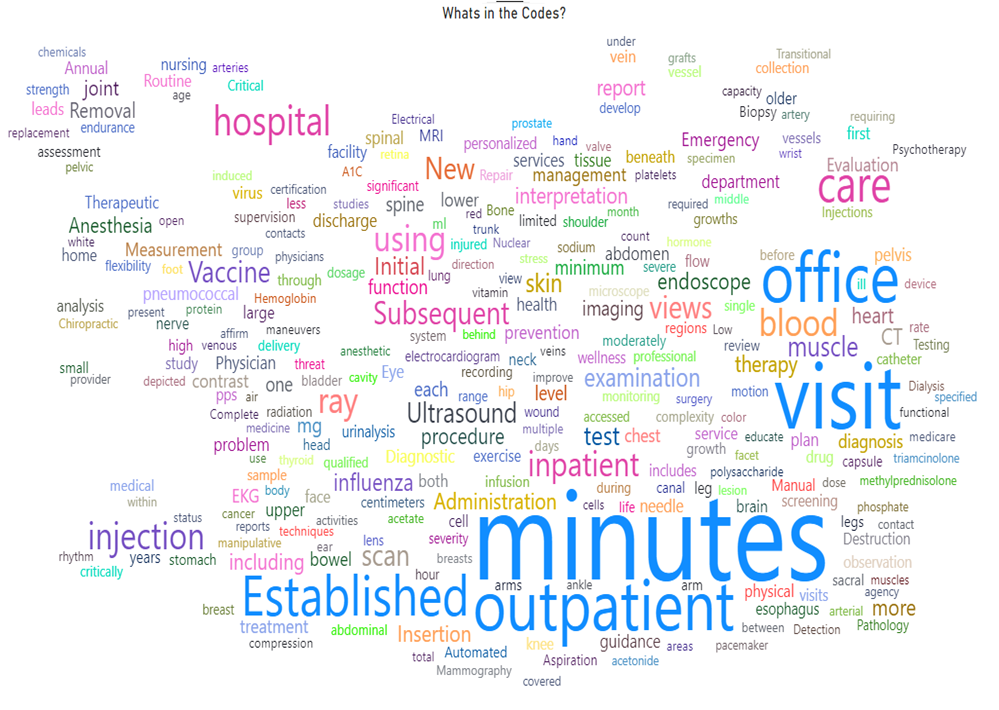


Fig 8: Median Unique Services Performed by NPI YoY 2013-14 (Top 50)

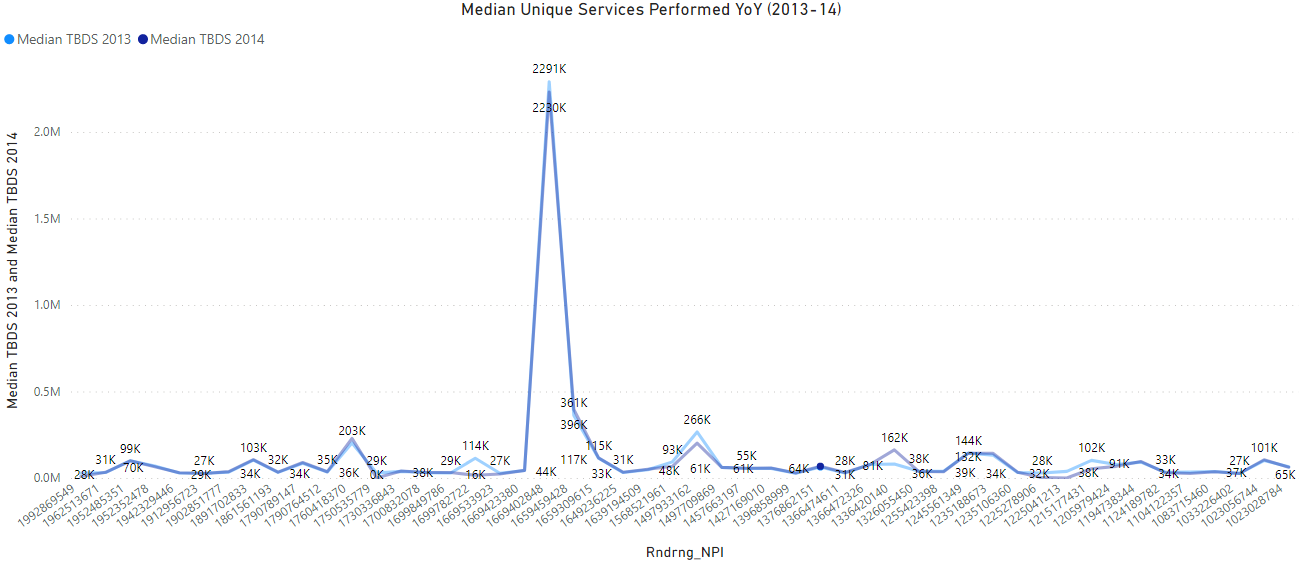


Fig 9: Average Medicare Payments by NPI YoY 2013-14 (Top 50)

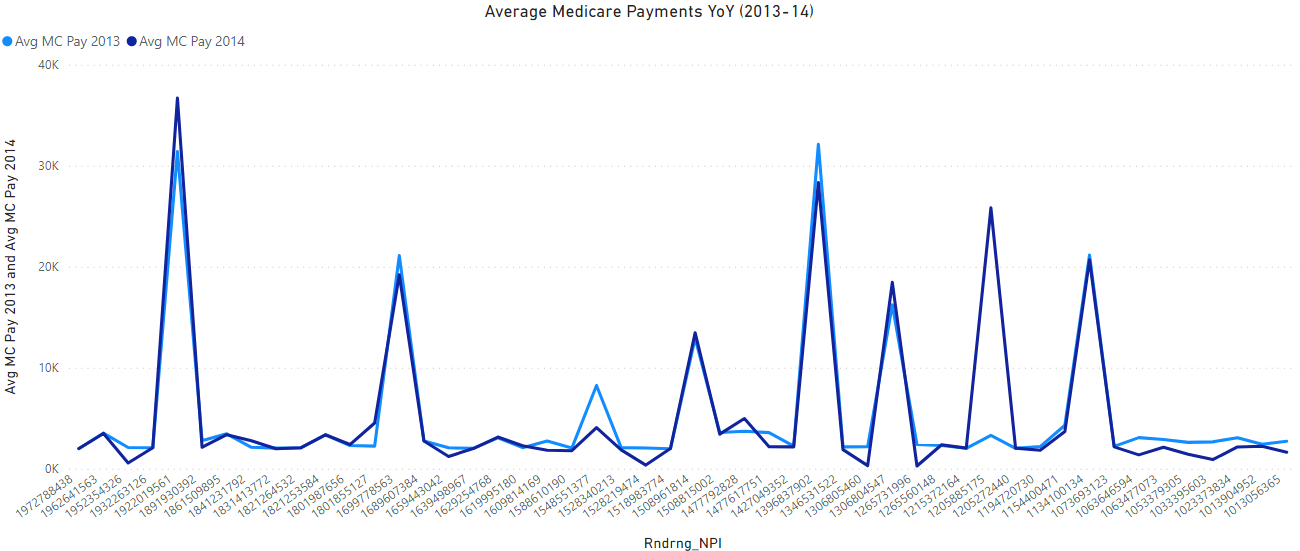


Fig 10: Average Submitted Amount from Provider YoY 2013-14

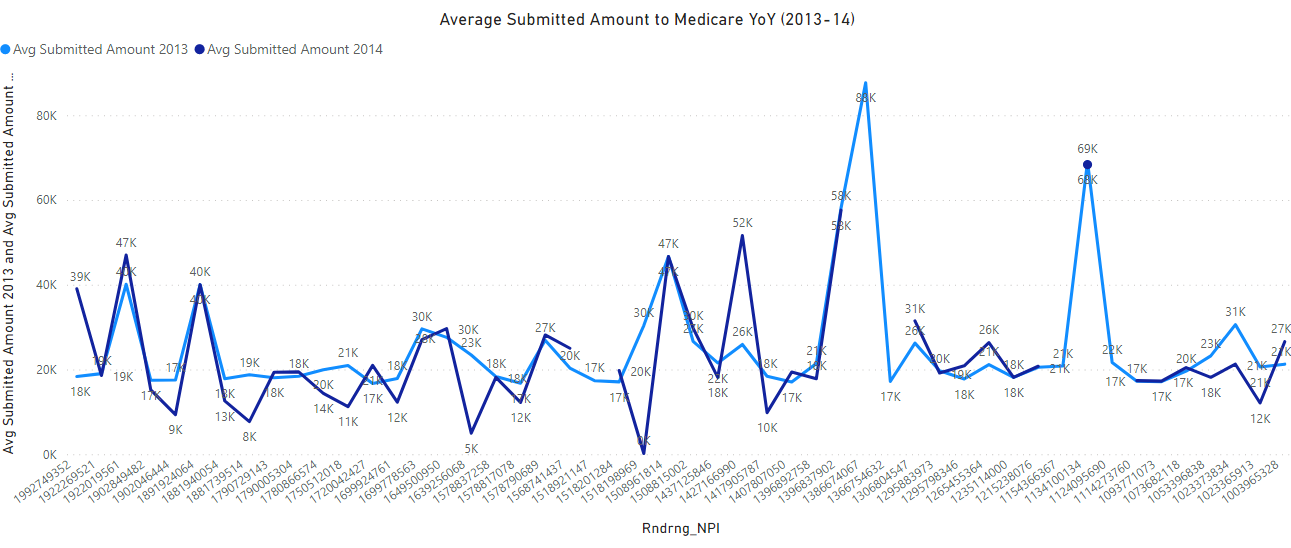
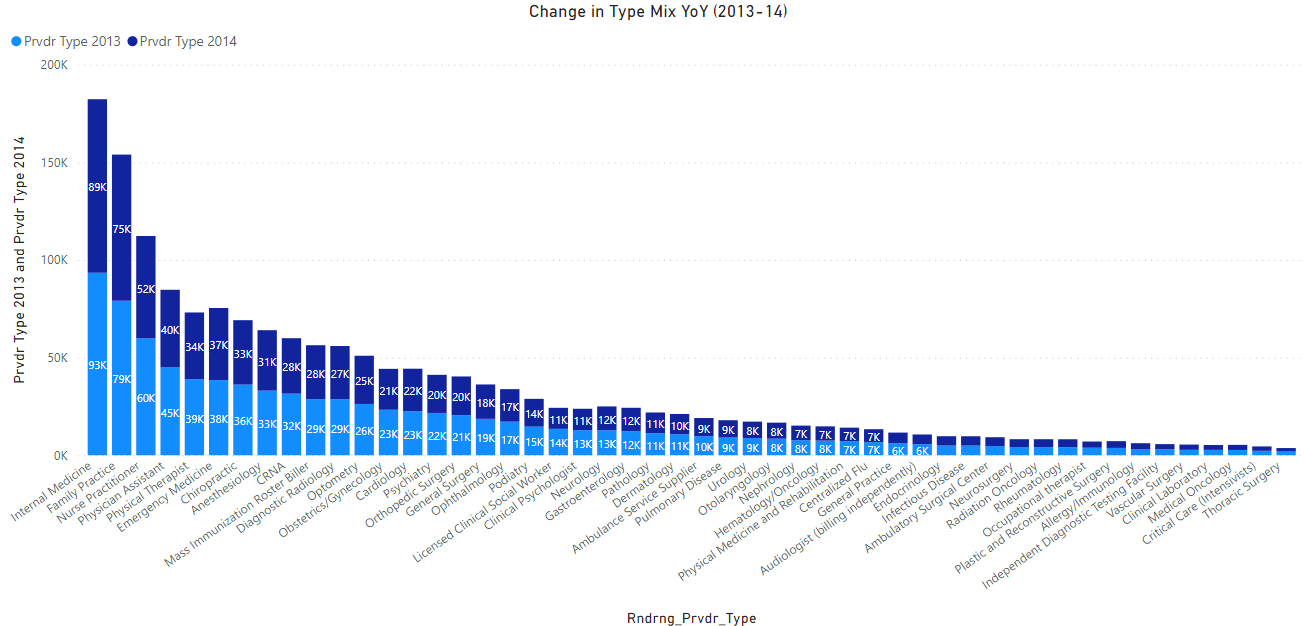


Fig 11: Change in Overall Provider Specialty Mix YoY 2013-14



7 Discussion

(Placeholder Section)

Interpretations: Are the results sufficiently convincing in their ability to predict and detect fraud and theft? As a secondary objective, do the results encompass all the vast supply chain data?

Implications: If the models can accurately detect fraud or theft, will organizations utilize and trust this new tool?

What stood out as interesting/unique/unexpected?

Limitations

* 1. What challenges occurred during analysis?

Ethics, questions to consider here such as, by detecting fraud, does the research team too often target providers making honest mistakes?

Future Research

* 1. Are there areas of research where others can pick up and go deeper?

8 Conclusion

Within this section the research team will discuss findings, interpretations, and practical implications from the study. (Placeholder)

References

1. R. Bauder and T. Khoshgoftaar, "A Survey of Medicare Data Processing and Integration for Fraud Detection," 2018 IEEE International Conference on Information Reuse and Integration (IRI), 2018, pp. 9-14, doi: 10.1109/IRI.2018.00010.
2. R. Bauder, R. da Rosa and T. Khoshgoftaar, "Identifying Medicare Provider Fraud with Unsupervised Machine Learning," 2018 IEEE International Conference on Information Reuse and Integration (IRI), 2018, pp. 285-292, doi: 10.1109/IRI.2018.00051.
3. Bauder, R. A., & Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for Medicare fraud detection with imbalanced big data. Health Information Science and Systems, 6(1), 1–14. <https://doi.org/10.1007/s13755-018-0051-3>
4. Bauder, R.A., & Khoshgoftaar, T.M. (2017). Medicare Fraud Detection Using Machine Learning Methods. 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 858-865.
5. Buppert, C. (2001). Avoiding Medicare Fraud part 2. The Nurse Practitioner, 26(2), 34–41. <https://doi.org/10.1097/00006205-200102000-00005>
6. Ekina, T., Leva, F., Ruggeri, F., & Soyer, R. (2013). Application of Bayesian Methods in Detection of Healthcare Fraud.
7. Gad, B., Warren, C., & Paskhover, B. (2020). Why Otolaryngologists Get Excluded From Medicare and Medicaid. Ear, Nose, & Throat Journal, 14556132093304–145561320933040. <https://doi.org/10.1177/0145561320933040>
8. Gordon, D., & Siegel, D. M. (2020). Machine learning and the future of Medicare fraud detection. Journal of the American Academy of Dermatology, 83(2), e133. https://doi.org/10.1016/j.jaad.2020.03.059
9. J. Hancock and T. M. Khoshgoftaar, "Performance of CatBoost and XGBoost in Medicare Fraud Detection," 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), 2020, pp. 572-579, doi: 10.1109/ICMLA51294.2020.00095.
10. M. Johnson and Taghi M. Khoshgoftaar. Medicare fraud detection using neural networks. Journal of big data. <https://doi.org/10.1186/s40537-019-0225-0>. 2019. pp. 6–63.
11. Kose, Ilker & Gokturk, Mehmet & Kilic, Kemal. (2015). An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Applied Soft Computing. 36. 283–299. 10.1016/j.asoc.2015.07.018.
12. Markus M Breunig, Hans Peter Kriegel, Raymond T Ng, and Jrg Sander. Lof: identifying density-based local outliers in ACM SIGMOD International Conference on Management of Data May 16-18 2000 Dallas, Texas, Usa, pages 93-104, 2000.
13. Meyers, T. J. (2017). Examining the network components of a Medicare fraud scheme: the Mirzoyan-Terdjanian organization. Crime, Law, and Social Change, 68(1), 251–279. <https://doi.org/10.1007/s10611-017-9689-z>
14. Musal, R. M. (2010). Two models to investigate Medicare fraud within unsupervised databases. Expert Systems with Applications, 37(12), 8628–8633. https://doi.org/10.1016/j.eswa.2010.06.095
15. Nicholas, L. H., Segal, J., Hanson, C., Zhang, K., & Eisenberg, M. D. (2019). Medicare beneficiaries' exposure to fraud and abuse perpetrators. Health Affairs, 38(5), 788-793,793A-793C. doi:http://dx.doi.org/10.1377/hlthaff.2018.05149
16. Pande, V., & Maas, W. (2013). Physician Medicare fraud: Characteristics and consequences. International Journal of Pharmaceutical and Healthcare Marketing, 7(1), 8-33.doi:http://dx.doi.org/10.1108/17506121311315391
17. Westerski, Kanagasabai, R., Shaham, E., Narayanan, A., Wong, J., & Singh, M. (2021). Explainable anomaly detection for procurement fraud identification—lessons from practical deployments. International Transactions in Operational Research, 28(6), 3276–3302.<https://doi.org/10.1111/itor.12968>
18. W. Zhang and X. He, "An Anomaly Detection Method for Medicare Fraud Detection," 2017 IEEE International Conference on Big Knowledge (ICBK), 2017, pp. 309-314, doi: 10.1109/ICBK.2017.47.
19. Medicare Learning Network (2021). “Medicare Fraud & Abuse: Prevent, Detect, Report”. ICN MLN4649244 January 2021, <https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/Fraud-Abuse-MLN4649244.pdf>. Accessed on: Dec. 4, 2021.
20. Musal, R. M. (2010). Two models to investigate Medicare fraud within unsupervised databases. *Expert Systems with Applications*, *37*(12), 8628–8633. <https://doi.org/10.1016/j.eswa.2010.06.095>
21. Liu, J., Bier, E., Wilson, A., Guerra-Gomez, J. A., Honda, T., Sricharan, K., Gilpin, L., & Davies, D. (2016). Graph analysis for detecting fraud, waste, and abuse in healthcare data. *AI Magazine*, *37*(2), 33–46. <https://doi.org/10.1609/aimag.v37i2.2630>
22. Obodoekwe, N., & van der Haar, D. T. (2019). A comparison of machine learning methods applicable to healthcare claims fraud detection. *Advances in Intelligent Systems and Computing*, 548–557. <https://doi.org/10.1007/978-3-030-11890-7_53>
23. Herland, M., Bauder, R. A., & Khoshgoftaar, T. M. (2017). Medical provider specialty predictions for the detection of Anomalous Medicare Insurance claims. *2017 IEEE International Conference on Information Reuse and Integration (IRI)*. <https://doi.org/10.1109/iri.2017.29>
24. CMS. (n.d.). *Medicare provider utilization and payment data: Physician and other supplier*. CMS. Retrieved December 7, 2021, from <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Physician-and-Other-Supplier>.