Automated Adjudication Logic and Denial Pattern Intelligence: Machine Learning Models for Health Insurance Claims Processing Optimisation

Dr. Natalia Borisova, Associate Professor of Artificial Intelligence, ITMO University, Russia

1. Introduction to Health Insurance Claims Processing

Health insurance claims processing partners insurers with healthcare providers to ensure quick and efficient financial management of patient care. Dealing with the complexity and number of claims filed by large hospital systems is important in sustaining those healthcare providers financially. Providers authorize procedures and document patient encounters, then submit claims to payers for reimbursement. Payers review these claims for errors and seek unnecessary claims using automated and human pending processes. Claims processing accuracy varies by payer; providers report rejection rates of up to 5% of claims filed. Claim rejections can occur when providers use the wrong forms, fail to sign forms, or use out-of-date codes. Delays also cause claims to get flagged for suspended or denied status as records are checked and corrected.

Efficiency in claims handling reduces payouts to former care recipients as well as the technical and administrative costs associated with the activities. Unfortunately, the healthcare sector has been slow in developing standardized procedures and information technology to streamline claims handling and billing processes. One challenge to standardization is the wide variety of medical care offered in different categories. Even within a single hospital, care varies between services. One system's billing practices may not align with another's documentation procedures. While considering the unique range of care, the sector must prioritize quick, efficient claim handling that reduces both administrative errors and the potential for fraud. Insurers must avoid the costs of delayed or inaccurate payment as well as the monetary harm to their customers that arises from botched claims processing. A tech tool that uses data to accurately predict a patient's illness and the corresponding treatment will enable automated claims reviews and minimize the risks of fraud or billing errors.

1.1. Overview of the Current Challenges

Health Claim Procedure and the Recent Scenario: Health insurance operations involve complex claim processing, adjudication, and payment operations involving healthcare providers and health insurers. The entire process of claim operations takes a considerable amount of time and effort. More than 30% of healthcare claim costs are attributed to administrative costs. To make the claim process easier, many of the processing requirements were relaxed, and more focus is enforced on post-payment audits. Recent observations have noted errors from 161 million claims submitted each year, and this is approximately 10% of the total claims processed. The current claim service-level agreement with the provider includes clean claim response, notification to third-party payers about secondary liability processing or adjustments, resolution of the full claim within a certain period, and timely adjustments.

Health insurance claims processing is an integral part of healthcare financing in both governmental and private sectors. Evolution in healthcare management works based on changes in regulations, pay-for-performance models, requirements for compliance and standards, and expectations from stakeholders. The regulatory perspective introduces many changes in statutory functions and policies for health insurers, so current health insurers are expected to be swift-moving enterprises. There are many possibilities for manual errors in claim submission, and a model for healthcare providers' operations assists in a significant way. These challenges are numerous, time-consuming, and lack a definite representation suited for developing a model for operations. The above discussion explicates the classic challenges and operational inefficiencies caused by manual claims operations, associating with an organization where the operations are outsourced.

1.2. Importance of Efficiency in Claim Review and Processing

Efficiency in the review and processing of medical insurance claims is a critical function of health insurance administration and payer organizations. From an operations perspective, companies seek efficiencies in their claims process to reduce costs associated with operational administration. From a customer service perspective, individuals and companies that purchase insurance seek that their claims be paid in a timely and accurate manner. Failure to do so can lead to dissatisfaction with their insurance company. For healthcare providers, payment is an indication of the value and

service that they provide. The insurance company becomes an intermediary that should not delay, prevent, or inhibit access to quality of care. Medical and pharmaceutical costs are a significant and increasing portion of the total cost of healthcare, which was approximately 18% of US GDP; it rose and is currently estimated closer to 20-21%.

One way to help create additional capacity and increase both access and affordability is to reduce operational costs by providing healthcare services more efficiently. Another is to enhance consumer satisfaction and trust in an insurer's ability to provide payments that indicate timely performance and value-added service. Machine learning techniques and predictive models can be used to streamline the review of claims for medical necessity, determine whether treatments or patient standards are appropriate, and shorten the decision-making process along the entire treatment and healthcare path. Pharmaceutical reviews in precisely this area have already been in use for over 25 years and have documented price avoidance savings for every $1 spent in program and service delivery and administration. An efficient medical and hospital claim review process can help reduce costs, streamline capacity obligations, maintain operations, provider satisfaction, and business revenues, and be a proactive tool to further operations by preventing potential future delays, grievances, and bad debt. Inefficiency in the payment process creates not only a payment delay but also, in some cases, decreases the full value and worth of the services and costs of care delivery. Machine learning models can be used effectively to expedite and improve service delivery and validity of care decisions. Health insurance claims volumes have peaked and led to a substantial delay in sliding PDA. It is critical to begin the transition to novel machine learning, AI, and NLP to review medical and hospital claims in detail to improve payer-physician relations, network relations, and to bolster system and institutional compliance in an efficient and expert-on-top fashion. Such artificial intelligence models can help predict, identify, streamline, or process and audit medical health insurance claims as part of overall operational compliance and payment integrity solutions.

2. Fundamentals of Machine Learning in Health Insurance

The chief acts of the process known as the ML cycle have the singular objective of predicting the type of claim in order to redistribute claims to those with low or no fraud scores. Within the realm of machine learning, the most central concept is predictions. The problem of predictions is further decomposed into a number of predictive sub-

problems, of which classification is the most relevant for the health insurance sector. Machine learning classification is the process of automatically sorting discrete data preferences into predefined classes. Any data point can be sorted into these classes based on features that are important in relation to the classification. The fundamental method for any machine learning algorithms to be effective is to be trained on labeled data. Typically, machine learning models can handle supervised, unsupervised, semi-supervised, and reinforcement learning scenarios. For health insurance, based on data characteristics, a number of typical machine learning models, including linear models based on various statistical techniques, such as nearest neighbors, decision trees, support vector machines, and neural networks, are presented. Putting data into machine learning models to train them to automate a task involves writing scripts in a programming environment. The goal is to train a model to optimally make a prediction given its specifications. One of the most vital parts of the process often relates to feature engineering, where new variables are constructed from the initial raw data with the objective of providing more insights than the raw attributes alone. One of the contemporary strengths of machine learning is the creation of models that can handle numerous types of data. This can include time-series data, numerical data, text data, or images such as X-rays or MRI scans. Not least, the so-called supervised learning is a popular technique used for classifying predictive models. Some of the most common applications are in the insurance industry, where machine learning can help reduce the complexity of customers' risk, increase the speed of processing claims, and more accurately assess the value of assets and reserves for insurers. It can also provide significant potential for accurately analyzing the potential of InsurTech companies.

2.1. Basic Concepts and Terminology

This article assumes that readers have a basic understanding of the fundamental health insurance concepts relevant to the claims process and sections of the healthcare business. The following section provides stakeholders with insight into common machine learning concepts and terminology used in health insurance for claims processing. As artificial intelligence, machine learning, and automated techniques become topical in healthcare, it is apparent that having at least a rudimentary understanding of machine learning will be beneficial to managed care professionals. After reading this article, a reader should have a working understanding of machine learning operations, downside risk, upside opportunities, limitations, and productive paths for application in healthcare,

particularly in health insurance. Typically, supervised learning is used for predictive modeling, as indicated in the sections that follow, but a basic understanding of unsupervised learning is represented as well.

At the core of artificial intelligence applications are advanced machine learning techniques. In short, the generic supervised machine learning process involves training a model (or multiple models in the case of ensemble modeling) with data. The "model" is a mathematical description of patterns found in the "training data" that should help the model to accurately predict certain outcomes when given previously unseen "testing data." Other common terms synonymous with "model" being discussed in digital health are "algorithm," when referring to unsupervised learning, and "method" or "procedure," reserved generally for artificial neural networks application. Precision and accuracy increase with model complexity and the quality and quantity of clean data. Together, the data and any computer code used to analyze or function upon the data form a binary relationship, where misuse of one piece results in misapplication of the other. It is important to recognize that in determining the model and training data, the problem is generally defined and constrained. The combination may be referred to as a class of "optimization problems" with various degrees of complexity. Therefore, the model can be used through scoring to support payment integrity in the delivery system by indicating required claims that are more likely to be erroneous or otherwise inappropriate. It is also important to note that, if done improperly, attempts to apply machine learning paradigms and advanced methodologies inappropriately incur further risks. Thus, the principles and important considerations discussed here will help explain such situations, noting the limitations and opening into high upside opportunities in lieu of downside risk.

2.2. Applications of Machine Learning in the Healthcare Industry

An extensive body of research illustrates the potential of machine learning to optimize processes in a variety of healthcare settings. This portion will focus on machine learning applications, categorizing them as process optimizations, predictive analytics, personalized treatment plans, or accuracy improvements. Machine learning applications include improving diagnostic processes, developing personalized treatment plans, and streamlining back-office operations. Successful machine learning applications are

transforming processes, focusing on developing and integrating resources to improve efficiency, patient retention, and encourage retention.

Admission processes have applications from predicting the number of subjects likely to enroll in a clinical trial to understanding sub-segments and likely behavior to optimize marketing spend and patient outcomes. Operational support applications involve specific examples in Medicare and Medicaid, white-collar insurance funds, and blue-collar carrier states. There are several machine learning applications that support the back-office billing operation, which makes them more efficient. Results show how to use machine learning and data science to reduce operating costs in insurance. Applications can include, but are not limited to, improving labor efficiency in personal medicalization and the quality of treatment and patient services. The impact of machine learning has yet to develop patient retention and reduce time to match patients and subjects. In the general segment of the personal health industry, the approval and speed of patient recruitment are described. Other services include the development of personalized treatment plans. In the segment controlling pharmaceutical trials for depression and schizophrenia, different populations demonstrate that they can be used in a personalized way to improve decision-making about strategic alternatives. Some success has been documented using unlabeled machine learning internally and extending the use of labeling machines with larger internal databases and tracking that provide patient profiles. There is evidence that multimodal technology can extend the use of the machine learning module. Machine learning applications also translate across the back-office insurance segment. Providing blue-collar carriers spread across Texas and Oklahoma with a decision-support tool has triggered and sped up their ability to return employees to work, reducing the number of open claims. White-collar carriers reduced the number of claims required for policyholders they settled without difficulty. Overall, application reducing potential pools creates a reduction in contested claims and increases labor and revenue.

3. Streamlining Claim Review with Machine Learning

One of the key stages for health insurance claims is the claim review process, in which insurer representatives ensure that the claim complies with the insurance policy and the stated loss is covered under the policy. Most of the process today is manual and, consequently, prone to error, suboptimal in many instances, and very time-consuming.

However, machine learning and predictive modeling techniques have the potential to automate and substantially enhance this process. We discuss some of these machine learning models for health insurance claim review.

One aspect of claim review for health insurance and other types of insurance is the automated claim adjudication system, in which a computer-based claims processing system makes the decision to pay the claim and, if so, the amount. The insurer can use rules to filter claims that are to be passed straight through the system. Predictive models can be used to assess the probability that a claim is correctly approved; claims that are unlikely to be approved can be passed to the manual adjudication system before the approval decision is made. Machine learning models can estimate the likelihood that a claim will be approved. If a claim is denied, predictive models can further estimate the likelihood of an appeal being successful. Conducting some approximation of these calculations over large claims volumes is extremely valuable, especially because any small percentage error in this process can really add up. This can result in underpaid claims and potentially a negative experience for the insured and their healthcare provider. Not having these effective capabilities in these realms can result in a large backlog of appeals and re-adjudications that need to be resolved. Machine learning can streamline this process.

3.1. Automated Claim Adjudication

Automated Claim Adjudication

Health insurance policies typically have coverage agreements that specify the treatment conditions and exclusions, the required qualifications of medical personnel, length and coverage of hospital stays, discharge planning, etc. The process of determining eligibility and services provided are covered by the patient's health insurance coverage. A high level of benefits, coverage, and limitations are, by design, machine understandable so that the adjudication process can occur in real time. This insight illustrates the applicability of such a process for the Medicaid and Medicare segments where authorized plans are more restrictive in allowable service than the menus selected. Machine learning helps untangle the complexity of managing thousands of providers and tens of contracts, considers literal volumes of data spanning from hospital procedures to cost structure compositions, even zip code driven cost and usage figures,

and allows denial rates, cost, and quality competition to create multiple operational profiles differentiated by product or market.

The AI adjudication engine is claimed to deliver a decision rate that is exponentially higher than is typically possible with a human touch system. The core business value lies in the combination of claim adjudication and service delivery, and top-line growth potential offering true differentiation and competitive edge. A casual review of the decision results did not note any variances that could be construed as detrimental to the patient and could only describe them as 'exceptional'. The intentional blending of technology advantages and human touch reflects that the value of a turned down claim is not based entirely on its DRG value. A thousand dollar decision and a thousand dollar bad debt both represent a loss. Swift technology makes other performance differentiators a potential point of branding, such as the technology to take all attributes of a patient's profile to create a message to the patient's doctor. By sending a denial letter drafting message to a doctor, reasons for why a procedure is denied can be communicated to the hospital's key programmatic decision makers. The results of this could have financial implications in reduced lawsuits, reimbursements, and settlements, plus ancillary revenues generated in the areas of education and training. We are presently analyzing the potential market to define market segments and quantify potential revenues and expect to be implementing the application shortly.

3.2. Predictive Modeling for Claim Approval

Predictive modeling can also be applied to health insurance in the context of claim approvals. This technique is based on analyzing historical data to quantify the probability of the claim’s approval. The insurer examines the distribution of historic approvals and denials and creates a predictive model of what factors in a claim are predictive for a particular outcome (approval or denial). This technique offers several dominant features compared to other techniques such as rules-based systems that are derived from experts’ knowledge. For example, there are five advantages of predictive modeling when utilized in decision support software by insurers including: (a) constructing the optimal claim acceptance cutoff for acceptance or denial that minimizes total processing cost while taking into account the cost of error; (b) addressing missing data through imputation using a random split sample; (c) providing “concurrent validity” since the model includes data up to the day the claim is evaluated, allowing

prediction of approval for a specific claim at that time; (d) updating the model over time as new cases appear; and (e) conducting "prediction validation" in order to estimate the additional benefit of employing predictive modeling over an established approach.

There are typically three main steps involved in the process of predictive modeling including (1) the analysis of past data, the development of an outcome, the identification of accurate variables in the model, and the generation of the model, (2) the utilization of the model to predict outcomes for new claims, assessing whether the claims the model approves would in fact be paid correctly, and utilizing accumulated "evidence" to adjust the decision boundary, and (3) the comparison of the performance of predictive modeling against standard desk/telephone adjudication decisions. In creating a predictive model, one must first determine what relation between the outcome and the selection criteria is needed—be it a linear or nonlinear association. Next, in building the model, the appropriate level of expression must be chosen: either the source variables and the raw data values, or potentially an index formed from the source values, or a weight sum of the source variables. Once the model has been constructed, it must be "trained" in two samples of data before performing its final application. The need for variable selection often perturbs the development of the model. Once the model is fully trained, it enters production system maintenance and this may involve regularly checking and updating the results that can be based on recent data. Several case studies indicate the benefits resulting from the utilization of predictive modeling for claim approval in health insurance by increasing the "hit rate" and/or the provider payment accuracy.

While predictive modeling aims to increase the number of approved clean claims, several potential challenges might limit its applicability. These include the inflexibility of models in addressing the varied or non-standard cases, fraudulent claims, or new procedures. Consequently, different cases need to be handled differently based on "flexible" rules, which might not be addressed by predictive models alone. Additional pitfalls include the potentially high cost of developing models if the volume of claims is not sufficient to cover development costs, the risk of over- or under-training the model if the model is subjected to a poor sample, and the difficulty in updating the model if the trend changes. Finally, there are many factors that impact the predictability of what will happen in the future based on past data or, conversely, when the current time is

relatively short in a “history.” Insurers need to address these factors if they opt to implement predictive modeling.

4. Fraud Detection Techniques in Health Insurance Claims

One Health Catalyst study found that in the United States, up to $935 billion is lost to fraud every year, with approximately 34% of fraud attributed to insurance fraud. There are various ways health care fraud can be conducted, and they can either involve patients, providers, or in some cases, both. Some specific examples of abuse and fraud by providers include ordering more tests than necessary for patients, filing claims for care that has not been provided, charging insurers for more expensive treatments than were administered, or—so-called "upcoding"—billing for care that is more complex than has been provided. In the United States, millions of dollars' worth of fraudulent claims are filed every year, as insurance payment to the provider has less patient involvement. With losses throughout the economy amounting to more than $80 billion a year, the overall impact of health care insurance fraud on the United States economy is significant if benefits and costs are included.

One of the most effective techniques to combat false and risky claims is to use machine learning and artificial intelligence tools to detect wrong and unusual patterns in large datasets—something frequently performed by health care real-time analysis. Machine learning algorithms like K-means, Support Vector Machine, Self-Organizing Maps, Hierarchical Agglomerative Clustering, and other anomaly or outlier detection algorithms offer a potential solution to the problem. There is no shortage of literature documenting classifiers or clustering algorithms for various fraud detection systems; however, much of the literature focuses on the detection of fraud in credit card abuse. There is much less literature that discusses using these detection systems for the health care industry. As we have explored, health care fraud leads to huge direct, indirect, and sociocultural/psychological costs. Since insurance fraud involves patients, the emotional cost of being a victim of medical identity theft leads to feelings of violation and mistrust. Medical identity theft victims can suffer enormous financial consequences in addition to the emotional distress. Our literature review indicates that this problem is prevalent and requires real-time analytics to catch the mistakes early in the process. Consequently, real-time health care AI may be cost-effective, as it might catch irregular claims early in the processing cycle, making it easier to reverse them before the payment is made. We

provide two instances of AI being employed to detect health care fraud below: one from the market and the other from academia.

4.1. Types of Fraud in Health Insurance

There are many different ways to categorize the types of fraud seen in health insurance claims. Commonly, activities are broken down into those perpetrated by providers against insurance companies, those perpetrated by patients against their own insurance companies, and those perpetrated by patients against insurance companies other than their own. The first category typically encompasses billing for medical services not rendered, also known as "phantom services"; unnecessary surgeries, diagnostic tests, or procedures; improperly bundling services or upcoding them; negotiating and receiving kickbacks for referring patients to another medical provider or facility; and other claims submitted for treatments that fall under false pretenses. These frauds are inspired by the desire either for immediate and often significant monetary reward or career advancement or by optimism that many fraudsters retain regarding the likelihood that their misdoings will not be uncovered.

The second and third categories are both motivated by the desire to save money and hence keep insurance premiums manageable. The second type consists of more passive behaviors that rarely threaten the financial security of the health care system as a whole, such as ignoring the terms of the policy, misrepresenting smoking status; misrepresenting or failing to disclose medical history when the policy is underwritten; letting someone else use the policy; or not informing the insurance company when an insured member is no longer covered by other insurance that might have taken over responsibility for the treatment. The third type is much more aggressive and often contributes to or arises out of abusive behaviors and addictions. Examples include selling or abusing prescription drugs, doctor shopping, and using two different insurance plans to cover services simultaneously while avoiding the promise not to bill the second. More serious forms of this type of fraud involve stolen or faked identities or signatures and forgeries, selling valid insurance cards, and collateral assignment of benefits. Each of these varieties represents potentially serious risks to the health care system.

Insurers face some difficult circumstances with respect to detecting and deterring fraud within the health insurance industry. For one, they risk alienating honest policyholders

if they take drastic measures to deter fraud. Privacy reigns supreme. Patients' records cannot be shared between providers or payers without patient consent, making it challenging to use the data profiles of either entity to identify a potential response to fraud. To help protect records from improper use and disclosure, two major regulations were established and put into force, the Security Rule and the Privacy Rule. The Security Rule provides guidance on what constitutes a security risk and security policy procedures that protect data. The Privacy Rule restricts the access of patient information to only certain people in a covered entity's organization, limiting the health insurance industry's ability to share the profiles of claimants or providers between organizations. Both the Security Rule and the Privacy Rule were enacted by the U.S. Department of Health and Human Services. These rules are aimed mainly at health care providers, but the effects also create challenges for insurers. Furthermore, technology is constantly evolving, and fraudsters are always seeking new ways to take advantage of the health insurance industry. It is no easy task to keep up with the ever-changing methods of fraud within the industry to make sure they are addressing all areas.

4.2. Machine Learning Approaches for Fraud Detection

Maliciously submitted claims or claim-related submissions are among the most common forms of health insurance fraud. Insurers use machine learning approaches to prevent a few abusive individuals from driving up costs for honest policyholders, while also keeping customer data and decision-making secure. One approach for fraud detection, supervised learning, trains models using historical claims data to predict the likelihood of fraud based on certain 'fraud indicators' (e.g., service codes, claim costs, etc.). The task is viewed by models as a classification problem: a binary (fraud/not fraud) or multi-class (e.g., low fraud risk, medium fraud risk, high fraud risk, etc.) classification task. A second approach, unsupervised learning, could be used to run anomaly detection, which is used to identify which claims look unusual or different from the rest. Unique algorithms are better suited to isolation forests or one-class support vector machines. Unsupervised methods often have great potential for flagging new fraud or abuse, since such examples may be novel and thus not appear frequently in historical data.

Two important potential areas for future exploration are personalized requests and adversely selected consumers for general insurance business. Hyperparameter settings, trained model complexity (usually more powerful models such as ensemble methods,

decision tree forests, boosted models, and neural networks are likely to be used to create many fraud detection models), more computing or data storage, and higher amounts of training data: models educated on larger, more varied data that may outperform models educated on smaller data. High-frequency updates are necessary because some fraud detection systems need instantaneous insights, but custom-designed real-time algorithms and systems can only be found in cutting-edge, proprietary solutions.

Many fraud detection AI models are currently in operation, with machine learning techniques flourishing. Such tools may derive useful insights from financial data to aid in fraud detection. It is important to understand that the models used by insurance companies are vulnerable to making mistakes during learning. For example, models may yield false positives and become overly suspicious of young and otherwise low-risk people who engage in adventurous activities, leading to higher requests for more information and higher premiums. A model could also introduce biases based on what data employees 'teach' it if that data tends to be biased. But as the technology advances, health insurance companies are likely to devote more resources to understand model biases in data more comprehensively and tweak them to align more precisely with societal values on these and many other questions.

5. Enhancing Payment Accuracy with AI

While most of the conversation about health insurance claims is related to claim review and fraud detection by payers, another dimension of improving the process is to enhance payment accuracy. There are many disconnects that occur in the claims process that lead to unwarranted inaccuracies in the payment amount. Coding errors, data entry mistakes, and insufficient communication between hospitals, clinicians, and other entities are all reasons for inaccurate payments. Artificial intelligence can be used to automate the correction of these issues, not just review the problems identified by people later. AI can be leveraged to manifest consistent results through data validation from multiple, already compared sources. One innovation yet to be adopted in the wider industry is the use of machine learning models to compare and quickly find differences in months' worth of historical data to separate valid rogue claims from statistically inaccurate data.

Attaining more accurate payments has multiple benefits. First, payers ultimately save money, which may manifest as lower customer premiums and cost sharing for patients.

Second, patients have a reduced risk of overpaying and seeing their claim potentially denied, sending them back to the beginning of the claims review process. It is very important to commit the correct amount at the first attempt to pay. Many successful PFSOs have implemented AI in order to enhance processes such as remittance automation validation and inventory relabeling. As in most cases of AI implementation, eliminating manual steps and redundancy in data can improve working capital and work hours. It requires minimal oversight of a data analyst to monitor an AI-based process designed specifically to optimize payment during the first few months of implementation.

5.1. Challenges in Payment Accuracy

The process of ensuring accurate payment to providers involves a host of challenges, which can be primarily attributed to claims inconsistency and complexity. The first challenge pertains to coding practices being inadequate to describe services fully and unambiguously, which might lead to variability in interpretations of services. Simply put, there exist multiple codes, descriptions, and interpretations for specific services. A related issue is that a significant proportion of services furnished are miscellaneous services and are never described in the coding systems of the insurer. The second source of coding discrepancies arises while recording patient services: many times, entries abstracted make their way into an electronic system, but errors occur in the tool. These errors produce false coding and can come from a variety of sources, including logical flaws in the software, misunderstanding of or ignoring documentation coding rules, or simply a poor quality of programming. The impact of this source of inaccuracy is not negligible because, as most data entry is now performed in batch mode, the overall quality of the services tends to be affected in a cascading effect. While every single error may not count much, the overall aggregate impact is very significant. The challenge is compounded when the claim is held for over ninety days for review. At this point, the software is run again to apply a different policy update and vintaged coding logic, and even if only a very small percentage change occurs on a large number of records, these can be very significant.

Actual human error occurs in three key places in the claim process: data entry, data certifying, and data processing. Inaccuracies might also occur due to either intentional or unintentional fraud. Erroneous payments result in substantial financial losses for

companies. It was reported that the healthcare industry incurred a cumulative annual loss of approximately $21 billion as a result of inaccurate claims. Such inaccuracies further pose risks of compromising care quality, as denial of services that are not compensable by a patient's health plan may concomitantly result in delayed or deferred medical treatment. As things stand, healthcare claims represent a widespread and ever-expanding challenge. The United States alone has seen a steep increase in claims volume, from 1.1 billion in 2014 to 1.2 billion in 2015. Given the dramatic rise in the sheer number of claims being filed, many decision support services are being considered and initiated by stakeholders. Regulators, insurers, and policymakers have become intent on at least attempting to curb inaccurate claims. Regulatory changes and compliance guidelines require adherence to specified standards. A survey indicated that, in 2007, only 20 percent of medical physicians felt that they had access to clear and unequivocal health-related insurance policies. The large number of health insurance programs in the US adds to the complexity of the problem. Currently, there exist over 1,500 health insurance programs for only 47 million uninsured Americans. It becomes evident that the long-term solution for the problem is the dissemination of best practices and aligning interests on billing practices and wants. The need for accuracy does not imply that we have encountered a losing battle. Rather, as we will explore, the error landscape is gradually evolving, and strategies that could at one time have been unproductive should be revisited.

5.2. Machine Learning Solutions for Payment Accuracy

Understanding what medical care factors drive appropriate payment is another area where machine learning can make an impact, as it is something they can likewise accomplish through supervised learning model training. An algorithm could analyze a year’s worth of historical payment data for doctor visits and determine that, in fact, 60% of office visits are coded at a level 3 care visit, 30% as level 4, and 10% as level 5. The algorithm can then recommend appropriate payments for each of these leveled codes, rather than a blanket amount for all office visits. This can be overlayed with other more traditional statistical predictive models to determine which doctors had levels that do not align with the actual services they provided to their patients.

Machine learning’s adaptability to new data also shows promise for more accurately processing payments. Automated audit solutions, enhanced by machine learning, are

able to predict based on historical data which vendors are overpaid, underpaid, and correctly paid. This allowed for the movement from processing audits randomly or being forced to conduct the most audits in particular zip codes to conducting an audit on the vendors predicted to be overpaid based on historical data. This same concept can be extended to real time - a method for checking right the first time. Machine learning solutions could predict in real-time whether or not the invoice being submitted for payment needs to be audited prior to payment processing. This can improve overall processing time and customer experience, as well as reduce vendors calling to check on the status of payment. This type of solution has been successfully tackled in other industries. It helps reduce delayed payments and administrative overhead, where it is urgent to ensure that every supplier mentioned on every invoice actually has an active contract. Importantly, they encode the system with a thirst for new data, so that it becomes more accurate over time.

Machine learning can directly affect payment accuracy by suggesting and even making the correct payment based on diagnosis and treatment interventions. A well-known use case of this includes a solution that recommended the correct inpatient payments for a large insurance company. After implementation, the insurer saw a significant increase in recoveries within one year. Inputs to the model can also be broadened to include things from claims specifically, cost reports, and EMRs. Improving payment accuracy through harnessing machine learning ultimately means making fewer errors in the first place. When this is applied to health insurance, claims processing errors can be decreased by looking at past payment accuracy in analytics.

6. Future Direction

Given the rapid developments in machine learning and artificial intelligence, the mode of processing health insurance claims might evolve rapidly in the near future. Ever-improving algorithms with AI approaches would conduct a substantial part of health claims processing at a significantly reduced cost and much faster speed. One of the key issues in claims processing is the visibility and auditability of the data since all the data are handled by the claims adjudicator as well as the insurance payer. Blockchain technology offers an opportunity to bring substantial transparency and auditability to health claims processing. An integration of blockchain with the AI model for health claim processing is an important future research direction. As the mode of health claim

processing is primarily led by the requirements of the industry and the changing expectations of the big players, a major design consideration in the architecture would be the adaptability within the AI model, considering the strict regulations and rules set down by the insurance regulatory bodies and other health watchdogs. As we advance into the future, insurance companies must evolve their way of thinking by embracing the talent and potential that AI has to offer in revolutionizing their claims processing strategies. A key cornerstone for such a revolution will be continued collaboration and sharing of information and resources between these insurers and hospitals in order to coherently tackle systematic problems in claims handling processes. It is likely that as technology evolves, all existing regulatory bodies will need to adapt and/or create a separate set of regulations for AI/ML-enabled tools and service providers in the health insurance niche. It is important, however, to reflect on the broader implications of this discussion. The transition of AI/ML into processing in health insurance claims might potentially lead to deeper scrutiny by consumer protection and privacy enforcement stakeholders regarding the use of such technologies. AI ethical concerns such as confidentiality and security, validity, reliability, and fairness of AI decision-making, accountability and transparency, and bias should be important considerations when AI is used in claims processing. Therefore, regulatory considerations will need to be a key driver of innovation in health claims processing. To effectively harness the benefits and evolving capabilities of machine learning models in health insurance scenarios, laws and regulations must be established that strictly govern the use of these AI tools as part of the underwriting and treatment classification process in health insurance.

7. Conclusion

In conclusion, across the worldwide health insurance landscape, innovation focused on technological change increases the value of our health insurance claims processing systems through the development of intelligent and prescriptive machine learning models. Ultimately, an efficient, accurate, and fraud-averse claims system is a fundamental boon for both patient experience and operational efficiency. This is a transformative moment where collaboration, industry investment, and emerging primary AI technologies and tools are aligned to deliver the low-lift, high-impact change most insurance companies desire in claims operations. However, reaching scale will require a concerted effort among insurers and practitioners alike to build platforms, not only models, that adapt to the space and fuel the machine learning innovations of the

future. Doing so responsibly and with the patient in mind is the shared commitment of every segment of the insurance value chain.

While incredible advancements in the management and processing of claims have been made as a result of cutting-edge technologies and more sophisticated analytics, the industry continues to face challenges associated with inherent inefficiencies and costly reporting oversights. The prevention and automation of these errors are more achievable now than ever, given the dramatic improvements that have been made in machine learning and emotional intelligence technology recently. The development and widespread adoption of new data inputs, tools, and technology will undoubtedly lead to the optimization of insurance claim review and fraud detection. Injury and consequence scores are more accurate and comprehensive when considering additional predictive inputs. With the maturation of these technologies, insurance claim oversight and adjudication is likely to decrease in volume and operational expense while driving improvements in integrity and payment accuracy.