Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Machine Learning Approaches on India's Health Insurance Market using Data Analytics

Nidhi Prabhakar[1], Suresh Kumar Kashyap[2], Barnita Ghosh[1]

[1]Research Scholar, Mittal School of Business, Lovely Professional University, Punjab, India
[2]Professor and Additional Dean, Mittal School of Business, Lovely Professional University, Punjab, India
Email: nidhi_arnav@yahoo.co.in

Data has grown in significance in the digital realm with the development of computer technology. Information gathering is crucial when it comes to data analytics. Every industry uses data analytics, from banking to retail, and the importance of data analytics in healthcare is on the rise.This research mainly seeks to categorize and predict healthcare data using data mining tools and machine learning (supervised) approaches. We need to build a sophisticated model (using machine learning) to filter through all the data in our databases since there is so much of it. The amount of data stored exceeds the speed at which human beings can analyze it. This categorization becomes even more significant since classified healthcare data may help with patient identification, diagnosis, and treatment. The huge, complicated, and varied nature of healthcare data is driving the need for data mining tools for classification and forecasting. Machine learning has emerged as a crucial tool for healthcare researchers to quickly, accurately, and efficiently resolve complex classification problems. Although researchers are aware of the problem, premature birth (PTB) continues to go unaddressed, despite the fact that it is a huge public health concern with far-reaching negative consequences for families and communities. The first goal of implementing a model based on machine learning is to predict cases of TB. Decision Tree (DT), Logistic Regression (RR), and Support Vector Machine (SVM) are three learner classifiers utilized for this purpose, in addition to Minimum Information Loss (MIL) discretization.

Keywords: Healthcare, Data Analytics, MDDS, Machine Learning, Classification, Feature Selection, PTB, Learning Classifiers, Data Mining.

1. Introduction

The word "healthcare" describes a system that improves health-related services to meet people's clinical needs. Everyone involved in healthcare—from patients to physicians to researchers to medical businesses—is working to save and recover patient records. There has been a dramatic increase in the need for data mining applications in recent years due to the

exponential growth of data across all industries, including healthcare, made possible by the phenomenal advancements in technology. But medical organizations are producing massive amounts of healthcare data as a result of the healthcare system's digitalization [1]. In a nutshell, any and all health records kept digitally are considered healthcare data [2]. There could be clinical reports, doctor's notes, prescription information, and a patient's whole medical history stored there. Large, multi-dimensional, and diverse data sets characterize all of these records. These days, it's hard to make good decisions since healthcare data is so complicated [1,2]. Academic research in machine learning, data mining, and statistical methods has greatly improved people's decision-making abilities, allowing them to achieve better results in any professional setting [3]. The pace at which humans can analyze stored data is disproportionately low in comparison to the total quantity of data [2]. In the healthcare industry, where there is a severe shortage of qualified personnel, this is of the utmost importance. The need for computerized medical disease diagnosis systems (MDDSs) that may aid in the making of better healthcare choices for all persons has led to their widespread adoption. Better illness detection, higher quality treatment, and lower healthcare expenses are all possible outcomes of this. Healthcare data categorization and prediction issues using supervised machine learning methods are the primary emphasis of this research. When combined with data mining methods, a number of learning algorithms may improve the reliability, accuracy, and speed of healthcare diagnostics by solving categorization challenges.[4]

With the growth of computer capabilities, many renovation studies are being conducted online. The digital world would not exist without data. In the realm of data analytics, data collecting plays a key role. Analyzing data with the purpose of drawing conclusions (useful insights) is called data analytics. Data analytics are used by every sector of the economy. The healthcare industry places a greater emphasis on data analytics compared to other commercial sectors, such as finance [2, 3]. Therefore, healthcare quality and, more crucially, the ability to take preventative actions may both be enhanced by the efficient use of data analytics in this field. The total potential of healthcare data is expected to be over $300 billion, according to a recent big data research [4]. A healthcare data analyst's primary responsibility is to learn about healthcare data so that they may draw conclusions and make decisions based on that knowledge. A number of data mining tools have recently been created to aid decision making in the healthcare industry [3, 6]. In order to improve the organization's capacity to generate money, reduce risk, and control costs, the authors of [1] examined healthcare data analysis, visualization, and mining and determined ways in which data might be effectively handled.[6].

The development of state-of-the-art healthcare has allowed data mining studies to enhance lives.A wide range of disciplines are included by the term "data mining," such as statistical analysis, pattern recognition, machine learning, and many more. There has been an exponential increase in the volume of data across all sectors. Due to its extensive usage in several sectors, including healthcare, data mining applications are highly sought after. A number of acknowledged difficulties exist in the healthcare domain for end-users, data-miners, and business-class computers. These include, but are not limited to, the heterogeneity of health-related data, the gradual expansion of data warehouses, and the need for intelligent machine-based data analytical tools (to better manage massive amounts of data). The amount of data stored is substantially more than the number of experts who can analyze it. Therefore, semi-

automatic analytical approaches are required to get insights from health-related data and make better decisions. In light of these considerations, a number of data mining techniques have been proposed to mine data warehouses for useful insights that might enhance decision-making. [7-9]

Development of the model and evaluation of the model are the two most critical considerations for incorporating an AI model into real-world datasets. The training dataset is essential for model construction and training. Next, the test dataset has to be used to assess how well the model performed. There is no lack of choices when it comes to assessing various learning methodologies. Both are widely used in many fields, but healthcare is one of them.

Predefined classes must be applied to input data (training data) in order for a supervised learning approach to work. It is the particular method via which an educator keeps tabs on student progress.[10] Classification problems, as mentioned before, may be seen as supervised learning using examples. The learning phase and the testing phase are the two stages of supervised machine learning.[11] During the learning phase, the training dataset is used for model creation, while the test dataset is utilized for model performance evaluation. The input data set is split into a training set and a test set according to any data partitioning technique[12]. A major data mining approach (a subset of supervised learning) for making predictions is classification. Decision Tree (DT), Logistic Regression (LR), Artificial Neural Network (ANN), Naïve Bayes (NB), Support Vector Machine (SVM), and others are examples of competent supervised machine learning algorithms.in [13][14]

Predictive issues including regression and classification are both handled by the KNN algorithm, a fundamental supervised learning technique. Among machine learning approaches, KNN is among the most basic. The basic idea of the KNN method is straightforward. The class of an instance's nearest neighbors really determines its classification. Here, it's sometimes beneficial to take more than one immediate neighbor (let's say K), which is why the method is called KNN. However, because to its large memory storage requirements for both the testing and training data sets, KNN is computationally costly to perform. In particular, the distance measure technique is used to determine the distance between the visible and invisible samples. As far as distance measurements go, the Euclidean distance criteria is by far the most used. Selecting the visible instances that produce the smallest distance from the invisible example allows for its classification (using distance measure). In K-Nearest Neighbor, there are essentially two steps. step one involves choosing the best neighbors, and step two involves using the distances between those neighbours to decide on a class.[15]

Support vector machines (SVMs) excel in two domains: classification and prediction. This method works well for both linear and non-linear classification issues. Classifying non-linear data is the usual use of this approach. Even more so, SVMs employ a non-linear kernel function to raise the dimensionality of the initial training data from a two-dimensional to a three-dimensional space. Then, it selects the best linear hyperplane to split the training set in half, which is called a decision boundary.[16]

Similar to a tuning parameter, the kernel function is used in support vector machine models. The kernel function is used to transform a non-linear separable issue into a separable problem. If you want to divide data into two categories, you may always utilize a hyperplane thanks to

the kernel function. This data separation hyperplane is determined by the SVM using a small subset of the training data (support vectors), which are feature vectors.[17]

Support vector machines (SVMs) use a non-probabilistic classifier that is linear and binary in order to classify instances of unseen data. A binary classifier that works well for dealing with the multiclass problem is pair-wise classifications, however it is time-consuming.

Analysis, interpretation, and prediction become very difficult when using SVM since it is a black box technique that depends on many unknown characteristics.[18]

The supervised learning process is the foundation of Neural Networks (NN). It has a stellar reputation for fixing categorization issues. Neural networks are a kind of artificial network in machine learning that mimics the way the human nervous system operates. In 1943, McCulloch and Pitts created a computational model called a neural network, which relies on mathematics and algorithms. A mathematical model that mimics the way the brain works is known as a neural network. To illustrate this approach, consider a directed network with three nodes: input, sink, and internal, or hidden. Here, an input layer contains the model's input nodes, and an output layer contains the model's output nodes. The remaining nodes show up in a concealed layer or layers. The edge represents the value of activation. To put it simply, a neural network consists of three distinct layers: input, hidden, and output. There is a connection between every node in every layer and every node in every other tier. To build a deep neural network, one may increase the amount of hidden layers [8]. The three main functions that a neural network typically executes are the error-function, the search-function, and the update-function. The error-function is used to assess the output quality (great or bad) for a certain set of inputs. The search tool finds the areas that need fixing so that the mistake rates are lower. Based on the results of the search, the update function will decide what needs to be changed. This is an iterative procedure that will enhance the algorithm's performance [8]. When faced with situations that defy easy characterization using logical analysis and conventional software, learning methods built on Neural Networks (NNs) prove to be an invaluable resource. Supervised classification is the goal of many neural network models. The backpropagation multilayer feedforward technique is used by the majority of neural networks (NNs).

2. MEDICAL DISEASE DIAGNOSIS SYSTEM (MDDSS)

In an effort to improve healthcare quality while decreasing healthcare costs, researchers are looking into medical disease diagnosis systems (MDDSs) to help doctors make better judgments when it comes to diagnosing and treating patients. Many clinical factors are considered while making these healthcare choices. As a rule, doctors treat patients based on their own expertise and what they've learned in the clinic. Because every doctor has their own unique set of skills and experiences, manual therapy may be time-consuming and prone to incorrect diagnoses. Also, as the population continues to rise, so does the need for doctors and other medical professionals. Therefore, computer-based MDDSs are necessary so that doctors may make better, more timely healthcare choices. To that end, these digital tools are finding widespread usage as a potentially game-changing resource for medical diagnostics, with the dual goals of improving healthcare quality while simultaneously decreasing healthcare

expenditures. MDDSs are becoming an integral aspect of medical treatment. Nevertheless, there are a number of problems and practical limitations linked to health-related data. The most critical point is that medical datasets could become unbalanced due to the fact that certain classes of illnesses may affect a tiny percentage of people. Furthermore, it might be continuously difficult to gather a significant number of individuals representing particular illnesses due to their low frequency. Finally, thirdly, there are contradicting concerns since individuals' clinical appearances vary greatly even if they have the same medical illness. These are the most significant problems that arose along the process of developing effective computerized MDDSs. In most cases, clinical data pertaining to illnesses is heavily used. It is also linked to every single person. New illnesses and their symptoms appear daily, and the search for remedies continues apace. Unfortunately, not everyone in the world has access to the uncommon knowledge that is known about the diagnosis, symptoms, and potential treatments for many illnesses. Model illness specificity, incomprehensibility, decision capacity, and clinical data complexity are some further constraints. For these reasons and more, creating a trustworthy prediction model to aid medical professionals in better diagnosing and treating patients' diseases is an arduous and difficult undertaking. The present study's overarching objective is to help healthcare providers better manage patient data by developing efficient MDDSs.

A. Necessity of Medical Disease Diagnosis Systems (MDDSs)

The present section describes the necessity of MDDSs in the area of healthcare on the basis of classification-prediction models. Pattern is everything around in this digital world. There is a pattern for every real object or an abstract notion. An object' s description would be a pattern when discussing about classification of objects. A new pattern must be classified under its own class. An important part of pattern classification involves selection of attributes and representation of patterns. In computer science, a pattern is represented using vector feature values. In fact, any medical informational dataset (under natural domain) such as disease dataset is a collection of patterns. Such dataset generally comprises of huge data with high dimensionality. Therefore, developing an appropriate and generalized predictive model for disease diagnosis using disease datasets is definitely a difficult task. In context to healthcare domain, data mining techniques are currently being utilized to find interesting (valuable) patterns from medical datasets, which are subsequently used to diagnose diseases. It's no exaggeration to say that without referring to comparative cases, diagnostic decision taken by doctors based on current medical condition of a patient is a complex task. Consequently, a number of Medical Disease Diagnosis Systems (MDDSs) have been developed to assist doctors in their decision making. Such computerized systems are broadly utilized for classification, prediction, and diagnosis of several diseases on the basis of clinical data stored in electronic medical records (EMRs). Interestingly, one noticeable point is that most of the Medical Disease Diagnosis Systems (MDDSs) are focused on Decision Tree (DT) Classifier for healthcare data analysis, disease diagnosis, and prediction purpose because it has the ability to deal with the worst features that arises normally in medical data sets.

B. Health Insurance Data Analytics Trends

When it comes to maintaining their competitive edge in a market that is becoming more cutthroat, health insurers must make it a priority to be current on the most recent innovations

in data analytics within the insurance industry.

One of the most noteworthy changes that has taken place in the health insurance industry in recent times is the trend toward treating members of health plans more like individuals and less like a group. The reasoning for this is straightforward: members do not want to be treated like a faceless mass, but rather they want their health plan personnel to recognize them, understand them, and most importantly, pay attention to them.A combination of data from internal and external sources is being integrated: In addition to the vast volumes of data created internally via member contact and sales, health insurance companies also receive a wealth of data from a wide variety of other sources. One of the most important developments in the health insurance sector is the implementation of analytics-enabled solutions that are able to combine and combine data from a variety of sources inside a single system. The reason for this is because when this information is dispersed throughout a variety of different systems, it becomes very challenging to make use of it.It is possible that health insurers may follow the trend of wearable technology, which is now being investigated by the healthcare industry as a method of encouraging consumers to lead better lifestyles (more on this topic will be discussed later). Because of the information that can be gathered about a person's actions via the Internet of Things, health insurers and medical experts have taken note of wearables such as the Health app on the Apple Watch and the fitness tracker Fitbit.With the right algorithms, artificial intelligence in insurance might evaluate claims data via a Health Risk Assessment, taking into consideration the specific characteristics of each individual who participates in a health care plan. Robots that are capable of learning and remembering information. HRAs are surveys, assessments, and other similar activities that average out members. As an example, members who are considered to be at a high risk may get gift cards as an incentive to quit smoking or participate in programs that promote heart health. This is going to be beneficial not only for the health of the members, but it is also going to be beneficial for the plan.Using chatbots, social media, portals, and other similar channels, artificial intelligence customer service may also be able to give consumer support in the form of automated guided plan selection. It is also possible for members to get the bulk of their questions addressed without ever having to pick up the phone with them. In the event that they need more support, it is possible that someone may phone them directly rather than having someone else contact them.As we have said in the past, big data is a really significant issue. On the other hand, were you aware that machine learning may potentially make it simpler to manage enormous amounts of data? This data analytics trend, which can be thought of as a kind of artificial intelligence, may be of assistance to health insurance companies in the development of algorithms that are capable of automatically evaluating both internal and external data information that is entered into their systems. These machine learning approaches have a wide range of applications, including the ability to lead health plan customers to the proper level of coverage, the ability to observe market trends and product performance, the ability to build predictive analytics models, and a great deal more.When we talk about analytics models that make use of predictive data, we should also mention that predictive modeling is another big data trend that has everyone in the health insurance industry on edge. Actuarial models have been used by health insurers for a considerable amount of time. This is done in order to evaluate the risks that are associated with insuring certain individuals and to calculate the right pricing for health plans.For the purpose of mining huge amounts of data for insights and developing more complicated models, predictive analytics has lately gained a lot of popularity among health insurance companies. It

is becoming more common for insurers to make use of predictive modeling in order to match members with the coverage that is most suited to meet their specific needs. Previously, insurers would exclude customers from certain health plan choices.In the past, health plan providers have always been entrusted with the responsibility of protecting the confidentiality of their policyholders' medical information. However, the exponential rise of data in the current digital age has made this obligation far more urgent. Both the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) were established by federal and international agencies in response to the need for data privacy. HIPAA mandates that organizations must adhere to particular criteria when dealing with the personal health information of individuals, while GDPR grants individuals the right to be forgotten. In order for health insurers to be in compliance with the laws and regulations that they are required to follow, they are required to increase data security measures. These methods include data encryption and rigorous privacy policies.Value-based insurance: An increasing number of health insurance companies are beginning to embrace the concept of value-based insurance as a means of distinguishing themselves from the competition and putting an emphasis on the member experience. Value-based care in the health care sector pays physicians for proactive treatment and positive health outcomes, and it sets a priority on the well-being of individuals. This is in contrast to the previous fee-for-service model, which placed an emphasis on the number of treatments and procedures that were provided.When it comes to health insurance, value-based plans are all about lowering costs without compromising the quality of coverage. According to the National Conference of State Legislatures, value-based insurance has the potential to assist in the removal of barriers that prevent improved health care from being provided. Coverage of wellness checks, treatments, and preventive care may allow health plans to decrease costs by reducing the need for expensive medical procedures in the future. This might be accomplished by reducing the need for operations.In the absence of a systematic format: When it comes to describing their members and the outcomes of their products, health insurance firms have traditionally relied on data that is already stored in a database or in a standard file format. On the other hand, as a result of the explosion of social media, health insurers have come across a whole new type of data, which is known as unstructured data.

C. Benefits of Health Insurance Data Analytics

Health insurance firms who are ready to engage in data analytics technologies might reap many advantages from big data.Provide an individualized service to each member. Exciting changes have occurred as the health insurance sector has moved its focus from products to members. Insurers may build comprehensive member profiles that provide health insurance agents and reps a bird's-eye perspective of each member by analyzing big data using a customer relationship management (CRM) system. Customer service insights gained from this data include a better grasp of members' identities, values, problems, and lifetime value as customers, among other things, allowing for more tailored, member-centric care. Also, with the help of AI, health insurance companies may tailor plans to each customer via chatbots and on-demand insurance.Prevent fraud by identifying it in advance. An estimated 3%-10% of the $2.26 trillion spent on health care in the US goes toward fraud, which costs the country an additional $68– $230 billion annually. Even a single case of fraud may greatly raise health plan prices for members, so it's not only the healthcare and insurance businesses that suffer

from this kind of deception. As a result, health insurers should put money into stronger fraud detection systems or, even better, put measures in place to stop fraud from occurring.

With the use of predictive analytics, claims investigators may sift through social media postings and other unstructured data for signs of possible fraud, and then mark certain claims for further examination. Machine learning allows insurers to track this activity over time and apply new rules in response to suspicious trends; this takes the guesswork out of preventing and detecting fraud.

Insurance AI may also detect false claims for services that were never ever provided. A male patient seeking an obstetrician's opinion, for instance, would trigger more scrutiny. Everything comes full circle when we consider the concept of "auto-adjudication," which means that with the ability to automatically accept claims comes the possibility of automatically rejecting them or marking them as fraudulent.

Additionally, AI and ML may be used to identify instances of fraudulent invoicing, such as many office visits in a single day, and to identify accounts or providers that repeatedly provide inaccurate information.

Give the appropriate treatment when it is needed. Choosing the best health insurance plan is a daunting and, at times, annoying task for many individuals. In the absence of adequate information, members run the risk of selecting health insurance that is inadequate for their requirements. To assess whether a member's present coverage is sufficient in comparison to their requirements, an agent may access the member's policy and policy usage using the company's CRM system and use health insurance analytics to big data.

An agent may advise a member to upgrade to a more complete plan if, for instance, the member is heavily using their current plan while having the least comprehensive option. Health insurers may get on the value-based insurance bandwagon with this data-driven member service strategy, which also opens up new revenue streams.

Take client service to a whole new level. Products like Dynamics 365 Customer support Insights provide real-time, AI-driven insights that help health insurers' support reps work smarter, close open cases quicker, and delight and retain customers. With this, insurance companies may improve efficiency, see where their money is going, and identify trends in client involvement, all of which can lead to problems being fixed before they affect policyholders.

In customer relationship management, "AI is opening new frontiers in customer experience and success by applying [natural language processing], sentiment analysis, automation, and personalization," as stated by Forbes. "AI is transforming customer journeys, improving interactions, and delivering more compelling experiences for 90% of organizations."

Promote healthy lifestyle choices to lower payments. Health insurers are ecstatic that it appears like everyone is sporting an Internet of Things (IoT)-enabled gadget these days. Insurance companies may get a clearer picture of a policyholder's health, risks, and behaviors by collecting and analyzing data from wearable health monitors. Insurers may also create individualized incentive programs using wearables to cut rates for health plan customers who participate in these programs and engage in healthy behaviors.

Utilize predictive analytics to expedite claims. Due to the large number of claims, claims adjusters aren't always able to meet the expectations of health plan members for prompt and efficient handling of their claims. In contrast, insurers may use predictive analytics to sift through claimants' member profiles and claim histories for trends in behavior, and then apply predictive modeling to ascertain potential outcomes. Instead of spending time going over every little detail, adjusters can utilize this data to react to claims more quickly, which improves efficiency and makes members happier overall.

While advances in AI and ML have the potential to automate claims processing, this is far from a finished product. When members contact their insurers with inquiries, they are often given a certain amount of time to get a response. For instance, because to the complexity of claims data, insurers may be given 45 days to respond to a claim inquiry.

3. Conclusion

The use of machine learning is particularly well-suited to the activities that are often carried out by individuals at a slower pace in the area of health insurance. Through the use of artificial intelligence and machine learning, it is possible to analyze and evaluate huge amounts of data in order to simplify and streamline the operations of health insurance. Both policyholders and insurers will see time and financial savings as a result of the influence that machine learning will have on health insurance. Through the use of artificial intelligence, repetitive tasks will be handled, freeing up insurance professionals to concentrate on procedures that will enhance the experience of policyholders. Patients, hospitals, doctors, and insurance providers will all profit from the potential of machine learning to complete tasks that are now carried out by humans but may be successfully completed by ML in a much shorter amount of time and at a lower cost. Machine learning is one component of cognitive computing that has the potential to handle a variety of difficulties in a wide range of applications and systems. Another component of cognitive computing is the use of historical data. When it comes to the healthcare industry, the problem of forecasting health insurance prices is still one that has to be explored and addressed.

References

1. Rakesh Raja, Indrajit Mukherjee, Bikash Kanti Sarkar, " A Machine Learning-Based Prediction Model for Preterm Birth in Rural India" , Journal of Healthcare Engineering, vol. 2021, Article Id 6665573, 11 pages, 2021
2. ul Hassan C.A., Iqbal J., Hussain S., AlSalman H., Mosleh M.A.A., Sajid Ullah S. A Computational Intelligence Approach for Predicting Medical Insurance Cost. Math. Probl. Eng. 2021;2021:1162553. doi: 10.1155/2021/1162553
3. Sailaja N.V., Karakavalasa M., Katkam M., Devipriya M., Sreeja M., Vasundhara D.N. Hybrid Regression Model for Medical Insurance Cost Prediction and Recommendation; Proceedings of the 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT); Visakhapatnam, India. 13– 14 November 2021; pp. 93– 98.
4. Dutta K., Chandra S., Gourisaria M.K., GM H. A Data Mining Based Target Regression-Oriented Approach to Modelling of Health Insurance Claims; Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC); Erode,

India. 8– 10 April 2021; pp. 1168– 1175
5. Shyamala Devi M., Swathi P., Purushotham Reddy M., Deepak Varma V., Praveen Kumar Reddy A., Vivekanandan S., Moorthy P. Linear and Ensembling Regression Based Health Cost Insurance Prediction Using Machine Learning. Smart Innov. Syst. Technol. 2021
6. R. Raja, I. Mukherjee and B. K. Sarkar, " A Systematic Review of Healthcare Big Data" , Scientific Programming, vol. 2020, Article ID 5471849, 15 pages, 2020
7. Kumar Sharma D., Sharma A. Prediction of Health Insurance Emergency Using Multiple Linear Regression Technique. Eur. J. Mol. Clin. Med. 2020;7:98– 105.
8. Van den Broek-Altenburg E.M., Atherly A.J. Using Social Media to Identify Consumers' Sentiments towards Attributes of Health Insurance during Enrollment Season. Appl. Sci. 2019;9:2035. doi: 10.3390/app9102035.
9. S. Debnath, D. P. Barnaby, K. Coppa et al., " Machine learning to assist clinical decision-making during the COVID-19 pandemic," BioMed Central, vol. 6, pp. 2332– 8886, 2020Twin, A. (2021). Stock Volume of Trading. Retrieved from https://www.investopedia.com/terms/v/volumeoftrade.asp#:~:text=Volume of trade is the,and all types of commodities
10. Bhardwaj N., Anand R. Health Insurance Amount Prediction. Int. J. Eng. Res. 2020;9:1008– 1011. doi: 10.17577/IJERTV9IS050700
11. Goundar S., Prakash S., Sadal P., Bhardwaj A. Health Insurance Claim Prediction Using Artificial Neural Networks. Int. J. Syst. Dyn. Appl. 2020;9:40– 57. doi: 10.4018/IJSDA.2020070103
12. Sun J.J. Master' s Thesis. National College of Ireland; Dublin, Ireland: 2020. Identification and Prediction of Factors Impact America Health Insurance Premium
13. Kumar Sharma D., Sharma A. Prediction of Health Insurance Emergency Using Multiple Linear Regression Technique. Eur. J. Mol. Clin. Med. 2020;7:98– 105
14. Fauzan M.A., Murfi H. The Accuracy of XGBoost for Insurance Claim Prediction.;Int. J. Adv. Soft Comput. Appl. 2018 10:159– 171
15. Rustam Z., Yaurita F. Insolvency Prediction in Insurance Companies Using Support Vector Machines and Fuzzy Kernel C-Means. J. Phys. Conf. Ser. 2018;1028:012118. doi: 10.1088/1742-6596/1028/1/012118
16. Boodhun N., Jayabalan M. Risk Prediction in Life Insurance Industry Using Supervised Learning Algorithms. Complex Intell. Syst. 2018;4:145– 154. doi: 10.1007/s40747-018-0072-1
17. Yang C., Delcher C., Shenkman E., Ranka S. Machine Learning Approaches for Predicting High Cost High Need Patient Expenditures in Health Care. Biomed. Eng. Online. 2018;17:131. doi: 10.1186/s12938-018-0568-3.
18. Fauzan M.A., Murfi H. The Accuracy of XGBoost for Insurance Claim Prediction. Int. J. Adv. Soft Comput. Appl. 2018 10:159– 171