IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

www.ijitce.com

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

VISUAL ANALYTICS-DRIVEN DETECTION OF COLLABORATIVE FRAUD IN HEALTH INSURANCE SYSTEMS

[1]Sajjana Gandla Ravi Teja, MCA Student, Department of MCA

[2] M G K Priyanka, MCA, (Ph.D), Assistant Professor, Department of MCA

[12]Dr KV Subba Reddy Institute of Technology, Dupadu, Kurnool

ABSTRACT

The functioning of the healthcare system is threatened by collusive fraud, in which many con artists band together to steal money from health insurance. However, owing to the absence of labelled data and the great resemblance of fraudulent behaviours to routine medical visits, current statistical and machine learning-based algorithms are restricted in their capacity to identify fraud in the context of health insurance. Expert knowledge must be included into the fraud detection process in order to guarantee the accuracy of the detection findings. We present FraudAuditor, a three-stage visual analytics method for collusive fraud detection in health insurance, developed in close collaboration with audit specialists in the field. In particular, in order to represent the visit connections of various patients holistically, we initially let users build a co-visit network interactively. Second, suspected fraudulent groupings are identified using an enhanced community identification algorithm that takes the level of fraud possibility into account. Lastly, with customised visualisations that accommodate various time scales, users may compare, examine, and validate questionable patient behaviour using our visual interface. In order to identify the true fraud group and rule out the false positive group, we performed case studies in a real-world healthcare setting. The outcomes and professional opinions demonstrated the approach's efficacy and usefulness.

I. INTRODUCTION

Maintaining social stability, improving people's quality of life, and managing healthcare resources are all greatly aided by an efficient health insurance system. China's National Basic Medical Insurance has more than 1.3 billion members1. Nonetheless, the rise in health insurance fraud incidents has turned into a serious societal issue. Nearly half of China's 815,000 health facilities had inappropriate or even unlawful funding expenses in 2020, according to an inspection by the Ministry of Public Security and the National Healthcare Security Administration. This resulted in an economic loss of over 22.3 billion yuan ($3.4 billion). 2. Of these incidents, the emergence of collusive fraud is the most critical and pressing [1]. Fraudsters get together to buy medications with insurance money and then cash them out. There are severe repercussions from the enormous volume of deception. In order to promptly uncover collusive fraud and stop more losses, efficient and effective detection techniques are desperately needed.

There are two obstacles to identifying collusive fraud in health insurance. Initially, it is difficult to differentiate the conduct of fraudulent medical visits from that of legitimate patients. Fraudsters usually purchase big amounts of medications that are readily marketed. However, patients with chronic illnesses and those undergoing treatment with Traditional Chinese medicine (TCM) have comparable buying habits to those of fraudsters since they must continue taking medicines for an extended period of time. Second, auditing by hand is required but time-consuming. Since a patient must take legal responsibility after being identified as a fraudster, misidentification is unacceptable for fraud detection. Auditors must synthesise a lot of contextual information in order to verify fraud,

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

including the amount of payment, the degree of fit between the patient's condition and medications, and the time of visits.

These issues are almost impossible for current collusive fraud detection techniques to address. Current approaches concentrate on using graphs to represent the interactions between fraudsters and statistical or machine learning (ML) techniques to identify fraudulent groupings. To find anomalous substructures (such as fraud groups or incidents), statistical methods use structural and attribute characteristics [2], [3], [4], or spectral analysis [5]. However, since collusive fraud in health insurance is so murky, audit specialists informed us that these techniques are prone to false positives. For auditors, eliminating false positives takes effort and may drastically lower detection effectiveness. Graph neural network (GNN) models are the primary tool used by machine learning techniques to identify collusive fraud [6, 7, 8]. Homogeneous or heterogeneous graphs are used to design fraudsters and their relationships. Fraudsters may be represented by GNNs trained on labelled data, and they can also be used to evaluate unlabelled people. Unfortunately, for high-performance GNNs, a lot of labelled data is required. In our situation, GNN models are useless without enough labelled fraud.

We suggest a unique visual analytics method to assist health insurance audit professionals in identifying suspect groups, looking into suspicious patients' visitation patterns, and validating collusive fraud findings in order to overcome these difficulties. To depict the relationships between patients, we suggest using a co-visit network. By extracting the characteristics of collusive fraudsters, including the time interval and amount of visits, the weights of the edges are determined. A weighted community identification algorithm can then identify suspicious groups that visit the same place repeatedly at the same time. The technique is included into Fraud Auditor, a prototype solution that helps professionals browse interactively and enhances model detection outcomes. By identifying co-visit relationships in visualisations of patient medical behaviour, Fraud Auditor may assist specialists in promptly identifying and analysing fraud. It is possible to confirm and rule out false positive groupings when combined with contextual data like sickness, medication, and charge information. To verify the efficacy of the suggested method, we provide case studies and expert interviews in actual health insurance situations.

A problem characterisation that encapsulates the requirements of collusive fraud detection in the context of health insurance is one of the work's contributions. Another is a novel three-stage visual analytics approach that takes expert knowledge and fraud groups' visit patterns into account when detecting collusive fraud in health insurance. Fraud Auditor is an interactive prototype system designed to make it easier to identify, investigate, and validate suspected collusive fraud organisations.

II. LITERATURE SURVEY

"A survey on statistical techniques for detecting health care fraud,"

J. Li, J. Jin, J. Shi, and K.-Y. Huang,

The United States' health care system has seen a considerable increase in costs as a result of fraud and abuse. With an emphasis on classifying fraudulent behaviours, identifying the primary sources and data characteristics that have been used to conduct fraud detection, outlining the crucial steps in data preprocessing, and summarising, classifying, and comparing statistical fraud detection methods, this paper attempts to provide a thorough survey of the statistical methods applied to health care fraud detection. In order to identify potential future study areas, a discussion of the gaps or underaddressed topics in the current research is given based on the results of this survey.

"Oddball: Identifying irregularities in weighted charts,"

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

M. McGlohon, C. Faloutsos, and L. Akoglu,

How can anomalies be found in a huge, weighted graph? Which guidelines must to be broken before a node is classified as abnormal? To locate these nodes, we suggest using the oddball method. The following are the contributions: In order to detect anomalies, we (a) identify a number of new rules (power laws) in density, weights, ranks, and eigenvalues that appear to govern the so-called "neighbourhood sub-graphs" and demonstrate how to apply them; (b) carefully select features and design oddball so that it is scalable and can operate un-supervised (no user-defined constants); and (c) report experiments on numerous real graphs with up to 1.6 million nodes, where oddball does, in fact, identify unusual nodes that make sense.

"Identifying Twitter spammer communities,"

P. S. Thilagam, R. Mishra, and P. Bindu,

In recent years, online social networks have gained enormous popularity and are now the main resources for monitoring the global impact of news and events. However, malevolent individuals are drawn to introduce new types of spam due to the popularity and variety of online social networks. A hazardous practice known as spamming occurs when a phoney user distributes unwanted communications in the form of mass messages, bogus reviews, malware, viruses, hate speech, foul language, or advertisements for marketing scams. Furthermore, it has been shown that spammers often create a network of related spam accounts and utilise them to disseminate spam to a sizable number of genuine users. Therefore, identifying such spammer groups inside social networks is quite desirable. Research on identifying spammer communities and hidden spam accounts is lacking, despite the fact that a great deal of effort has been done in the area of identifying spam messages and accounts. This study suggests using an unsupervised method called SpamCom to identify Twitter spammer groups. In order to detect spammers based on their URL characteristics and structural behaviour, we model the Twitter network as a multilayer social network and take use of the overlapping community-based properties of users represented by hypergraphs. Our method is very reliable and effective due to the utilisation of community-based features, user account graph and URL characteristics, and user content similarity.

"Networked-guarantee loan risk management using visual analytics,"

According to Z. Niu, D. Cheng, L. Zhang, and J. Zhang, groups of businesses may create intricate networks and guarantee one another in an effort to get bank loans. The regulatory commission and banks are quite concerned about keeping an eye on a network's financial health and avoiding or minimising systemic risk in the event of a crisis. Our ultimate objective is to provide a visual analytical method and instrument for risk analysis and decision-making. Four primary analytical activities carried out by financial specialists have been combined by us: i) Multifaceted Default Risk Visualisation, which consists of the development of an interface to visualise important indicators and a hybrid representation to anticipate the default risk; ii) Risk Guarantee Patterns Discovery. iii) Network Evolution and Retrospective, which uses animation to help users understand the guarantee dynamic; iv) Risk Communication Analysis; and iii) Shneiderman mantra guidance for designing interactive visualisation applications, which describes an interactive risk guarantee community detection and a motif detection based risk guarantee pattern discovery approach. Banks and the government may both benefit from using the temporal diffusion route analysis to track the spread of default status. It also offers guidance on how to minimise and eliminate systemic financial risk by taking preventative action. We use case studies using actual bank loan data to put the system into practice. The produced tool is endorsed by two financial experts. This is, as far as we are aware, the only visual analytics tool created to

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

systematically examine networked-guarantee loan risks.

"A new method for using spectral analysis to find health care frauds,"

A. Gangopadhyay and S. Chen,

Fraudulent actions have been the subject of several study publications and are found in many facets of both our everyday lives and enterprises. Credit card transactions, telecommunications, network intrusions, banking and insurance, and scientific applications are the industries where these kinds of operations are most common. However, there has been relatively little study in this field since healthcare fraud detection has not received much attention. The dearth of study isn't because health care fraud causes little losses; rather, it's because the losses are enormous. Tens of billions of dollars are lost annually as a result of health care fraud, according to 2011 NHCAA estimates. Health care data are seldom made available to research communities due to privacy issues. In this paper, we will present and evaluate a new fraud detection method based on a community identification algorithm using spectrum analysis using a de-identified health claims dataset. In terms of spotting possibly fraudulent patterns in prospective physician collusions, our data demonstrate strong performance and encouraging outcomes. We have examined possible fraudulent situations in order to assess our results. additional areas of fraud detection issues might be addressed by this community detection method and a list of additional comparable algorithms.

III. SYSTEM ANALYSIS AND DESIGN

EXISTING SYSTEM

In order to identify egonets, Akoglu et al. [2] used structural characteristics from the network, such as node degree or centrality. SpamCom [3] used structural and attribute characteristics, including user topology, user profile, and Twitter content similarity, to identify spammer communities on Twitter. Chen et al. [5] used a spectrum analysis-based community detection technique in healthcare contexts to identify instances of patient referral fraud from a bipartite network of specialists and doctors. A dynamic heterogeneous information network including patients, hospitals, and illnesses was created by Zhao et al. [13]. Then, over set or variable durations, they found anomalies that matched established fraud patterns (such the expensive single treatment). Although statistics-based techniques may provide first fraud candidates, they may yield inaccurate findings that need expert confirmation.
By using attention processes and conditional random fields, Xu et al. [7] describe GRC, a new GNN model that learns representations of various person kinds and identifies loan fraud. But since these machine learning techniques are supervised or semi-supervised, they need data that has been labelled with fraud, which is not available in our health insurance case.
The loan guarantee network is shown by Niu et al. [4] using a node-link diagram, in which each node is associated with a community that is determined by a random walk technique and is coloured accordingly. Tao et al. [25] suggested a high-order correlation graph to enable analytic procedures beginning with an aberrant node in order to detect collective abnormalities.The high-order correlation graph makes it simple to identify the corresponding nodes that contribute to the abnormality. Both graph and sequence visualisation are included into our system. Our method offers more contextual information, including sickness, medication, and visit frequency, to concentrate on collusive fraud in health insurance situations..

Disadvantages

- The system did not include the detection of suspicious groupings in the current work.
- This system's absence of a graph neural network results in lower performance.

PROPOSED SYSTEM

In order to assist health insurance audit professionals in identifying suspect groups, looking into the visitation patterns of suspicious

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

patients, and validating collusive fraud findings, the system suggests a revolutionary visual analytics technique. To depict the relationships between patients, we suggest using a co-visit network. By extracting the characteristics of collusive fraudsters, including the time interval and amount of visits, the weights of the edges are determined. A weighted community identification algorithm can then identify suspicious groups that visit the same place repeatedly at the same time.

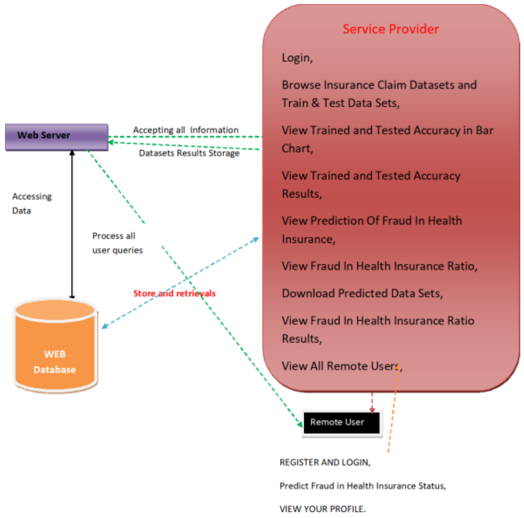
The technique is included into Fraud Auditor, a prototype solution that helps professionals browse interactively and enhances model detection outcomes. By looking for co-visit ties in visualisations of patient medical behaviour, FraudAuditor may assist professionals in promptly identifying and investigating fraud.

It is possible to confirm and rule out false positive groupings when combined with contextual data like sickness, medication, and charge information. To confirm that the suggested method works, we provide case studies and expert interviews in actual health insurance situations.

Advantages

- A problem description that enumerates the prerequisites for detecting collusive fraud in the context of health insurance.
- A brand-new, three-stage visual analytics method that takes into account expert knowledge and the visitation patterns of fraud groups to identify collusive fraud in health insurance.
- FraudAuditor, an interactive prototype system designed to make it easier to find, investigate, and validate questionable collusive fraud organisations.

SYSTEM ARCHITECTURE


Service Provider
Login,
Browse Insurance Claim Datasets and Train & Test Data Sets,
View Trained and Tested Accuracy in Bar Chart,
View Trained and Tested Accuracy Results,
View Prediction Of Fraud In Health Insurance,
View Fraud In Health Insurance Ratio,
Download Predicted Data Sets,
View Fraud In Health Insurance Ratio Results,
View All Remote Users,
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
Predict Fraud in Health Insurance Status,
VIEW YOUR PROFILE.

IV. IMPLEMENTATION

Modules

Service Provider

The Service Provider must use a working user name and password to log in to this module. He may browse insurance claim datasets and train and test data sets after successfully logging in. See the results of the trained and tested accuracy, the bar chart showing the accuracy, the prediction of health insurance fraud, the ratio of health insurance fraud, and the predicted data sets that can be downloaded. View All Remote Users and Fraud in Health Insurance Ratio Results.

View and Authorize Users

The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use his password and authorised user name to log in. Following a successful login, the user will do tasks such as registering and logging in,

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

predicting health insurance fraud, Examine your profile.

ALGORITHMS

Logistic regression Classifiers

The relationship between a collection of independent (explanatory) factors and a categorical dependent variable is examined using logistic regression analysis. When the dependent variable simply has two values, like 0 and 1 or Yes and No, the term logistic regression is used. When the dependent variable contains three or more distinct values, such as married, single, divorced, or widowed, the technique is sometimes referred to as multinomial logistic regression. While the dependent variable's data type differs from multiple regression's, the procedure's practical application is comparable.

When it comes to categorical-response variable analysis, logistic regression and discriminant analysis are competitors. Compared to discriminant analysis, many statisticians believe that logistic regression is more flexible and appropriate for modelling the majority of scenarios. This is due to the fact that, unlike discriminant analysis, logistic regression does not presume that the independent variables are regularly distributed.

Both binary and multinomial logistic regression are calculated by this software for both category and numerical independent variables. Along with the regression equation, it provides information on likelihood, deviance, odds ratios, confidence limits, and quality of fit. It does a thorough residual analysis that includes diagnostic residual plots and reports. In order to find the optimal regression model with the fewest independent variables, it might conduct an independent variable subset selection search. It offers ROC curves and confidence intervals on expected values to assist in identifying the optimal classification cutoff point. By automatically identifying rows that are not utilised throughout the study, it enables you to confirm your findings.

Naïve Bayes

The supervised learning technique known as the "naive bayes approach" is predicated on the straightforward premise that the existence or lack of a certain class characteristic has no bearing on the existence or nonexistence of any other feature. However, it seems sturdy and effective in spite of this. It performs similarly to other methods of guided learning. Numerous explanations have been put forward in the literature. We emphasise a representation bias-based explanation in this lesson. Along with logistic regression, linear discriminant analysis, and linear SVM (support vector machine), the naive bayes classifier is a linear classifier. The technique used to estimate the classifier's parameters (the learning bias) makes a difference.

Although the Naive Bayes classifier is commonly used in research, practitioners who want to get findings that are useful do not utilise it as often. On the one hand, the researchers discovered that it is very simple to build and apply, that estimating its parameters is simple, that learning occurs quickly even on extremely big datasets, and that, when compared to other methods, its accuracy is rather excellent. The end users, however, do not comprehend the value of such a strategy and do not get a model that is simple to read and implement.

As a consequence, we display the learning process's outcomes in a fresh way. Both the deployment and comprehension of the classifier are simplified. We discuss several theoretical facets of the naive bayes classifier in the first section of this lesson. Next, we use Tanagra to apply the method on a dataset. We contrast the outcomes (the model's parameters) with those from other linear techniques including logistic regression, linear discriminant analysis, and linear support vector machines. We see that the

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

outcomes are quite reliable. This helps to explain why the strategy performs well when compared to others. We employ a variety of tools (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b, and RapidMiner 4.6.0) on the same dataset in the second section. Above all, we make an effort to comprehend the outcomes.

Random Forest

Random forests, also known as random decision forests, are ensemble learning techniques that build a large number of decision trees during training for tasks like regression and classification. The class chosen by the majority of trees is the random forest's output for classification problems. The mean or average forecast of each individual tree is given back for regression tasks. The tendency of decision trees to overfit to their training set is compensated for by random decision forests. Although random forests are less accurate than gradient enhanced trees, they often perform better than choice trees. However, their performance may be impacted by data peculiarities.

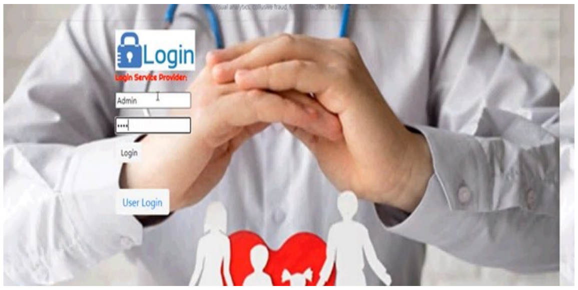
Tin Kam Ho[1] developed the first algorithm for random decision forests in 1995 by using the random subspace technique, which in Ho's definition is a means of putting Eugene Kleinberg's "stochastic discrimination" approach to classification into practice.

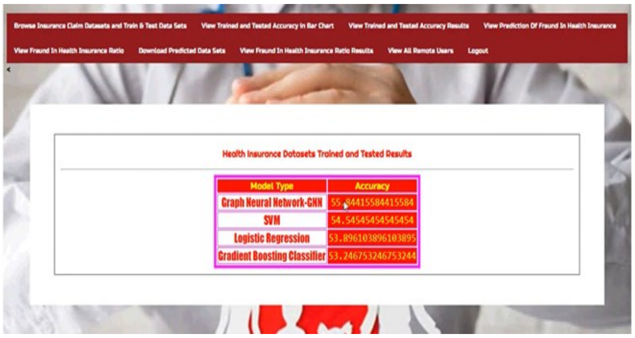
Leo Breiman and Adele Cutler created an algorithm extension and filed for a trademark in 2006 for "Random Forests" (owned by Minitab, Inc. as of 2019).The extension builds a set of decision trees with controlled variance by combining Breiman's "bagging" concept with random feature selection, which was initially proposed by Ho[1] and then separately by Amit and Geman[13].

Businesses often employ random forests as "blackbox" models since they need minimal setup and provide accurate forecasts across a variety of inputs.

V. SCREEN SHOTS

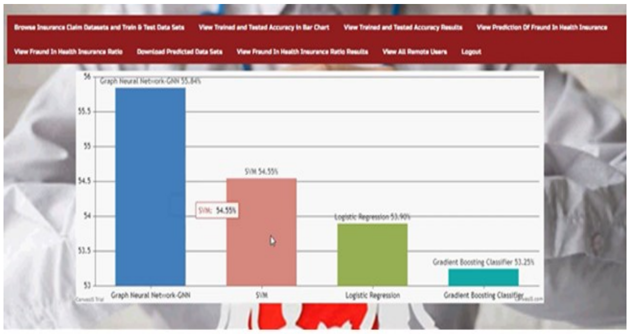
Visual analytics, collusive fraud, fraud detection, health insurance.

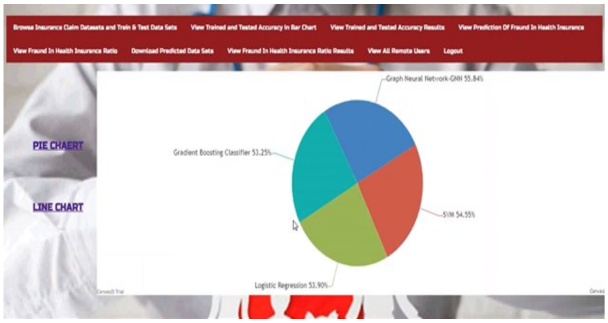


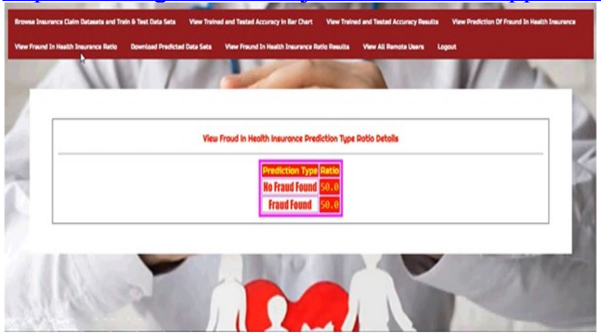

Browse Insurance Claim Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Fraund In Health Insurance
View Fraund In Health Insurance Ratio
Download Predicted Data Sets
View Fraund In Health Insurance Ratio Results
View All Remote Users
Logout

Health Insurance Datasets Trained and Tested Results

Model Type	Accuracy
Graph Neural Network-GNN	55.84415584415584
SVM	54.54545454545454
Logistic Regression	53.896103896103895
Gradient Boosting Classifier	53.246753246753244

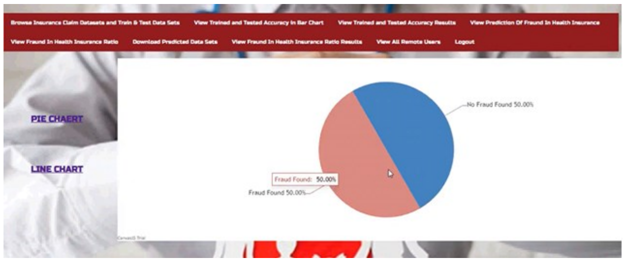

Browse Insurance Claim Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Fraund In Health Insurance
View Fraund In Health Insurance Ratio
Download Predicted Data Sets
View Fraund In Health Insurance Ratio Results
View All Remote Users
Logout
56
55.5
55
54.5
54
53.5
53
Graph Neural Network-GNN 55.84%
SVM 54.55%
SVM: 54.55%
Logistic Regression 53.90%
Gradient Boosting Classifier 53.25%
Graph Neural Network-GNN
SVM
Logistic Regression
Gradient Boosting Classifier
CanvasJS Trial

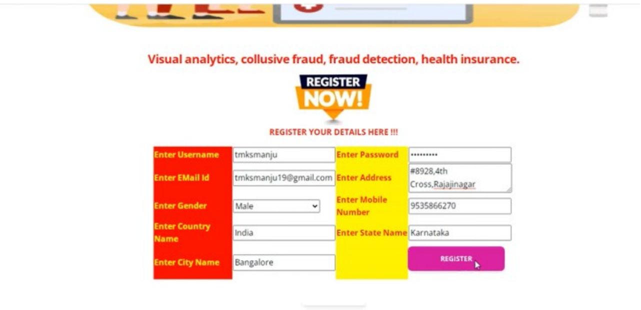

Browse Insurance Claim Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Fraund In Health Insurance
View Fraund In Health Insurance Ratio
Download Predicted Data Sets
View Fraund In Health Insurance Ratio Results
View All Remote Users
Logout
PIE CHAERT
LINE CHART
Graph Neural Network-GNN 55.84%
Gradient Boosting Classifier 53.25%
SVM 54.55%
Logistic Regression 53.90%
CanvasJS Trial


Browse Insurance Claim Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Fraund In Health Insurance
View Fraund In Health Insurance Ratio
Download Predicted Data Sets
View Fraund In Health Insurance Ratio Results
View All Remote Users
Logout

View Froud In Health Insuronce Prediction Type Rotio Detoils

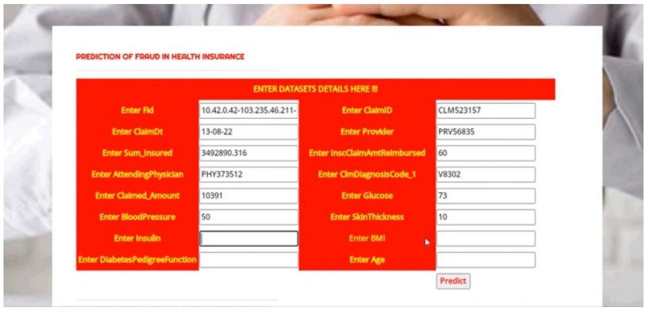
Prediction Type	Ratio
No Fraud Found	50.0
Fraud Found	50.0


Browse Insurance Claim Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Fraund In Health Insurance
View Fraund In Health Insurance Ratio
Download Predicted Data Sets
View Fraund In Health Insurance Ratio Results
View All Remote Users
Logout
PIE CHAERT
LINE CHART
No Fraud Found 50.00%
Fraud Found: 50.00%
Fraud Found 50.00%
CanvasJS Trial


Visual analytics, collusive fraud, fraud detection, health insurance.

REGISTER YOUR DETAILS HERE !!!

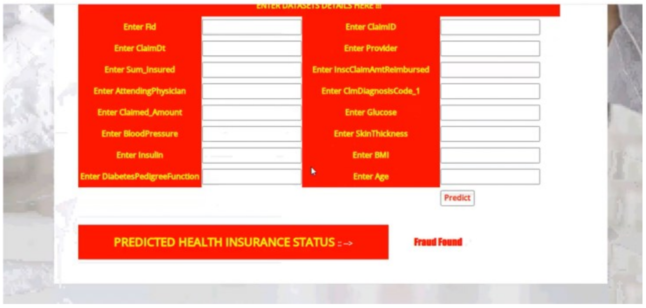
Enter Username	tmksmanju	Enter Password	••••••••••
Enter EMail Id	tmksmanju19@gmail.com	Enter Address	#8928,4th Cross,Rajajinagar
Enter Gender	Male	Enter Mobile Number	9535866270
Enter Country Name	India	Enter State Name	Karnataka
Enter City Name	Bangalore		REGISTER


PREDICTION OF FRAUD IN HEALTH INSURANCE

ENTER DATASETS DETAILS HERE !!!

Enter Fid	10.42.0.42-103.235.46.211-	Enter ClaimID	CLM523157
Enter ClaimDt	13-08-22	Enter Provider	PRV56835
Enter Sum_Insured	3492890.316	Enter InscClaimAmtReimbursed	60
Enter AttendingPhysician	PHY373512	Enter ClmDiagnosisCode_1	V8302
Enter Claimed_Amount	10391	Enter Glucose	73
Enter BloodPressure	50	Enter SkinThickness	10
Enter Insulin		Enter BMI	
Enter DiabetesPedigreeFunction		Enter Age	

Predict


ENTER DATASETS DETAILS HERE !!!

Enter Fid		Enter ClaimID	
Enter ClaimDt		Enter Provider	
Enter Sum_Insured		Enter InscClaimAmtReimbursed	
Enter AttendingPhysician		Enter ClmDiagnosisCode_1	
Enter Claimed_Amount		Enter Glucose	
Enter BloodPressure		Enter SkinThickness	
Enter Insulin		Enter BMI	
Enter DiabetesPedigreeFunction		Enter Age	

Predict

PREDICTED HEALTH INSURANCE STATUS :: --> Fraud Found

VI. CONCLUSION

In this work, we suggested a visual analytics method that facilitates the detection, analysis, and annotation of health insurance collusive fraud. Close cooperation with subject matter specialists served as the foundation for the development and deployment of Fraud Auditor. Fraud Auditor enables a multi-level fraud analysis, including the co-visit network overview, suspicious group identification, and suspicious patient assessment, by using both automated algorithms and human expertise. The identification and investigation of fraud groups is aided by a collection of visualisation designs. Interviews with health insurance audit specialists and case studies demonstrated the prototype system's usability and the efficacy of our methodology.

REFERENCES

[1] J. Li, K.-Y. Huang, J. Jin, and J. Shi, "A survey on statistical methods for health care fraud detection," Health Care Management Science, vol. 11, no. 3, pp. 275–287, 2008.

[2] L. Akoglu, M. McGlohon, and C. Faloutsos, "Oddball: Spotting anomalies in weighted graphs," in Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2010, pp. 410–421.

[3] P. Bindu, R. Mishra, and P. S. Thilagam, "Discovering spammer communities in twitter," Journal of Intelligent Information Systems,vol. 51, no. 3, pp. 503–527, 2018.

[4] Z. Niu, D. Cheng, L. Zhang, and J. Zhang, "Visual analytics for networked-guarantee loans risk management," in Proceedings of Pacific Visualization Symposium, 2018, pp. 160–169.

[5] S. Chen and A. Gangopadhyay, "A novel approach to uncover health care frauds through spectral analysis," in Proceedings of International Conference on Healthcare Informatics, 2013, pp. 499–504.

[6] J.Wang, R.Wen, C.Wu, Y. Huang, and J. Xiong, "Fdgars: Fraudster detection via graph convolutional networks in online app review

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

system," in Companion proceedings of the World Wide Web conference,2019, pp. 310–316.

[7] B. Xu, H. Shen, B. Sun, R. An, Q. Cao, and X. Cheng, "Towards consumer loan fraud detection: Graph neural networks with roleconstrained conditional random field," in Proceedings of AAAI Conference on Artificial Intelligence, 2021, pp. 4537–4545.

[8] Q. Zhong, Y. Liu, X. Ao, B. Hu, J. Feng, J. Tang, and Q. He, "Financial defaulter detection on online credit payment via multi-view attributed heterogeneous information network," in Proceedings of The Web Conference, 2020, pp. 785–795.

[9] I. Molloy, S. Chari, U. Finkler, M.Wiggerman, C. Jonker, T. Habeck,Y. Park, F. Jordens, and R. v. Schaik, "Graph analytics for realtime scoring of cross-channel transactional fraud," in Proceedings of International Conference on Financial Cryptography and Data Security,2016, pp. 22–40.

[10] Z. Li, H. Xiong, and Y. Liu, "Mining blackhole and volcano patterns in directed graphs: a general approach," Data Mining and Knowledge Discovery, vol. 25, no. 3, pp. 577–602, 2012.

[11] H. Joudaki, A. Rashidian, B. Minaei-Bidgoli, M. Mahmoodi,B. Geraili, M. Nasiri, and M. Arab, "Using data mining to detect health care fraud and abuse: a review of literature," Global Journal of Health Science, vol. 7, no. 1, pp. 194–202, 2015.

[12] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: a survey," Data Mining and Knowledge Discovery, vol. 29, no. 3, pp. 626–688, 2015.

[13] B. Zhao, Y. Shi, K. Zhang, and Z. Yan, "Health insurance anomaly detection based on dynamic heterogeneous information network," in Proceedings of IEEE International Conference on Bioinformatics and Biomedicine, 2019, pp. 1118–1122.

[14] K. Ding, J. Li, R. Bhanushali, and H. Liu, "Deep anomaly detection on attributed networks," in Proceedings of SIAM International Conference on Data Mining. SIAM, 2019, pp. 594–602.

[15] D. Wang, J. Lin, P. Cui, Q. Jia, Z. Wang, Y. Fang, Q. Yu, J. Zhou, S. Yang, and Y. Qi, "A semi-supervised graph attentive network for financial fraud detection," in Proceedings of the International Conference on Data Mining, 2019, pp. 598–607.

[16] S. Ko, S. Afzal, S. Walton, Y. Yang, J. Chae, A. Malik, Y. Jang, M. Chen, and D. Ebert, "Analyzing high-dimensional multivariate network links with integrated anomaly detection, highlighting and exploration," in Proceedings of IEEE Conference on Visual Analytics Science and Technology, 2014, pp. 83–92.

[17] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin, "Targetvue: Visual analysis of anomalous user behaviors in online communication systems," IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, pp. 280–289, 2015.

[18] C. Mac¸ ˜as, E. Polisciuc, and P. Machado, "Vabank: visual analytics for banking transactions," in Proceedings of International Conference Information Visualisation, 2020, pp. 336–343.

[19] C. Mac¸ ˜as, E. Polisciuc, and P. Machado, "Atovis - A visualization tool for the detection of financial fraud," Information Visualization, vol. 21, no. 4, pp. 371–392, 2022.

[20] J. Zhao, N. Cao, Z. Wen, Y. Song, Y.-R. Lin, and C. Collins, "#fluxflow: Visual analysis of anomalous information spreading on social media," IEEE Transactions on Visualization and Computer Graphics, vol. 20, no. 12, pp. 1773–1782, 2014.

[21] E. Bertini, P. Hertzog, and D. Lalanne, "Spiralview: towards security policies assessment through visual correlation of network resources with evolution of alarms," in Proceedings of IEEE symposium on visual analytics science and technology, 2007, pp. 139–146.

[22] P. Silva, C. Mac¸ ˜as, E. Polisciuc, and P. Machado, "Visualisation tool to support fraud detection," in Proceedings of the International

https://doi.org/10.62647/ijitce.2025.v13.i2.pp552-561

Conference Information Visualisation, 2021, pp. 77–87.

[23] Y. Lin, K.Wong, Y.Wang, R. Zhang, B. Dong, H. Qu, and Q. Zheng, “TaxThemis: Interactive mining and exploration of suspicious tax evasion groups,” IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 849–859, 2021.

[24] W. Didimo, G. Liotta, F. Montecchiani, and P. Palladino, “An advanced network visualization system for financial crime detection,” in Proceedings of the Pacific Visualization Symposium, 2011, pp.203–210.

[25] J. Tao, L. Shi, Z. Zhuang, C. Huang, R. Yu, P. Su, C. Wang, and Y. Chen, “Visual analysis of collective anomalies through highorder correlation graph,” in Proceedings of the Pacific Visualization Symposium, 2018, pp. 150–159.