iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52003

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

A Semantic Network Encoder for Associated Fact Prediction

Abirami.N[1], Sneha.J[2], Mr.D.Shanmugavel[3]
[1, 2, 3]Department of Computer Science and Engineering Sri Muthukumaran Institute of Technology

Abstract: Semantic network is a network of concepts connected by semantic relations. It contains two forms of binary semantic network and multiplex semantic network. The associated fact prediction is a link prediction task that aims to infer the implicitly connected facts by mining the high-level representation of the network. Previous methods for associated fact prediction put much emphasis on the topological feature of network but not utilize the information of semantic expression. This paper proposes a Semantic Network Encoder (SemNE), which learns a feature mapping function from the binary semantic networks and can be applied to the multiplex semantic networks in a pre-training manner. SemNE is a two-stage framework that contains an embedding encoder and a prediction decoder. It jointly models the semantic information and network topology to enrich the network representation. A word self-organization method based on the factual boundary is proposed to unify the topological feature and the semantic feature representations. Experimental results on binary semantic networks show that SemNE achieves the state-of-the-art results in associated fact prediction and experimental results on multiplex semantic networks show that SemNE is scalable and can effectively improve the performance of existing
Keywords: Semantic networks, link prediction, word self-organization, network embedding, jointly modeling.

I. INTRODUCTION

Tasks such as robot planning [1], cyber threat detection, [2] and text clustering [3] are made more informative through the use of external general knowledge sources, like Pro base [4] and Concept Net [5], etc. which are well- known as Semantic Networks (SN). Typically, an SN is a network of concepts connected by semantic relations, in which each concept contains a group of words while the semantic relation bridges the concepts to compose the whole network [6]. Each concept considering the word- level parse is called the semantic facts.

It can be defined as $fact$ = ($wordm$, $wordm$+1), where m denotes the position of the $word$ in the fact [7]. Each semantic relation can either be the unified binary relationship (e.g., “Is A” relationship) that is indicated by a fact pair or the specific relationship (e.g., “Take Office” relationship) that is explicitly described in the tuple [8]. Therefore, the extended terms of binary semantic network and multiplex semantic network are used to describe the different SNs. For example, “{(Tom, Cruise), [(famous, action), (action, actor)]}” is the basic word component of the semantic fact “famous action actor”.

II. OBJECTIVE

Monitoring is difficult to do in coal mines, particularly underground mines. Even when monitoring is in place, accidents occur on a regular basis because they are unexpected. These accidents inflict catastrophic injuries and costs to the coal mining business. This research primarily examines two types of data: seismic activity using Seismic Bump Analysis and coal mine risk assessment utilizing accident data reports. The risk evaluation is based on the number of casualties documented.

Accidents can occur in a variety of ways; important criteria are studied and assessed for prediction. In the coal mining sector, monitoring is set up to calculate seismic bumps using a geophone. Many geophones are used in underground mining. Geophone data is collected and forwarded for examination.

The analysis is completed, but the record is not instantly updated and submitted to higher authorities since the person of higher authority might be anywhere in the coal mine or outside the coal mine. There is no rapid reaction measure, and phones are not permitted inside coal mines for safety reasons.

This is where the project comes in, since the web application is beneficial for obtaining clearance for a certain data set, and prediction is also performed automatically and in a comprehensible manner. The data may reach the higher authorities quickly by email and can also be viewed within the online application. Authorities can review the data and preventative measures and approve them.

III. LITERATURE SURVEY

1) Network Analysis Of Coal Mine Hazards Based On Text Mining And Link Prediction

Hazards are a potential source of harm and damage hiding in shadow zones. Without control, they may accumulate and interact with other types of hazards. In the harsh and complicated circumstance, especially, the deep underground space of coal mines, complex and nonlinear interactions among hazards multiply the probabilities that a hazard turns into accidents, more seriously, its effect may trigger more correlated hazards to worsen the accidents and bring huge loss of lives and assets. Therefore, identifying the correlations among hazards and understanding the complexity of interactions among coal mine hazards are significant for ensuring the safety of coal production. From this standpoint, we propose a hybrid method combing text mining and complex network method. First, we abstract the dangerous hazards from a large amount of text data. Then, we establish the coal mine hazard network (CMHN) to capture correlations among hazards. Finally, we analyze and predict the correlations among hazards based on CMHN. Through which, we find the fault-prone hazard and the recurrent hazards, more importantly, we figure out the nonlinear correlations among hazards and reveal the connection preference of hazards. Furthermore, we forecast the unknown correlations among hazards to take precautions of them. This study could be helpful for making prevention strategies for safety management in the coal mine and other highly dangerous industries.

2) Application Of Rule-Based Models For Seismic Hazard Prediction In Coal Mines

The paper presents results of application of a machine learning method, namely the induction of classification and regression rules for seismic hazard prediction in coal mines. The main aim of this research was to verify if machine learning methods would be able to predict seismic hazard more accurately than methods routinely used in Polish coal mines on the basis of data gathered by monitoring systems. In this paper three classification and two regression tasks of prediction of seismic hazards in a longwall were defined. The first part of the paper describes the principles according to which the assessment of seismic hazard in Polish mines is made. These methods are called routine and allow to assess seismic hazard for a particular longwall. The next part of the paper discusses the algorithms of classification and regression rule induction and describes their use for seismic hazard assessment. The input data, which are the basis for rule induction, are: measurement data coming from seismometers and geophones, and the results of routine methods of hazard assessment. Conducted tests showed that automated hazard prediction based on induced rules gives better sensitivity and specificity of predictions than methods currently used in mining practice.

IV. EXISTING SYSTEM

The prediction process is automated, yet the data changes due to continuous monitoring. The data is recorded but not perfectly analyzed, because safety monitoring only indicates the data, the Safety Officer needs to check the data continuously, if hazard prediction occurs, data is recorded and then sent to higher authorities, the analysis takes time, and the risk assessment of accidents is not done extensively.

It may be improved. Risk assessment does not happen automatically with preventative actions. The dataset is not accessible in bulk. The data must be collected in order to do bulk prediction analysis. Annual accident records for total casualties, fatalities, severe and minor casualties and injuries are not gathered. Accident causes are not described in depth, such as roof collapse, how the collapse occurred, measures done, and so on. These kinds of data are not available since accidents happen at random, and the risk environment within underground mining is extremely hostile.

V. PROPOSED SYSTEM

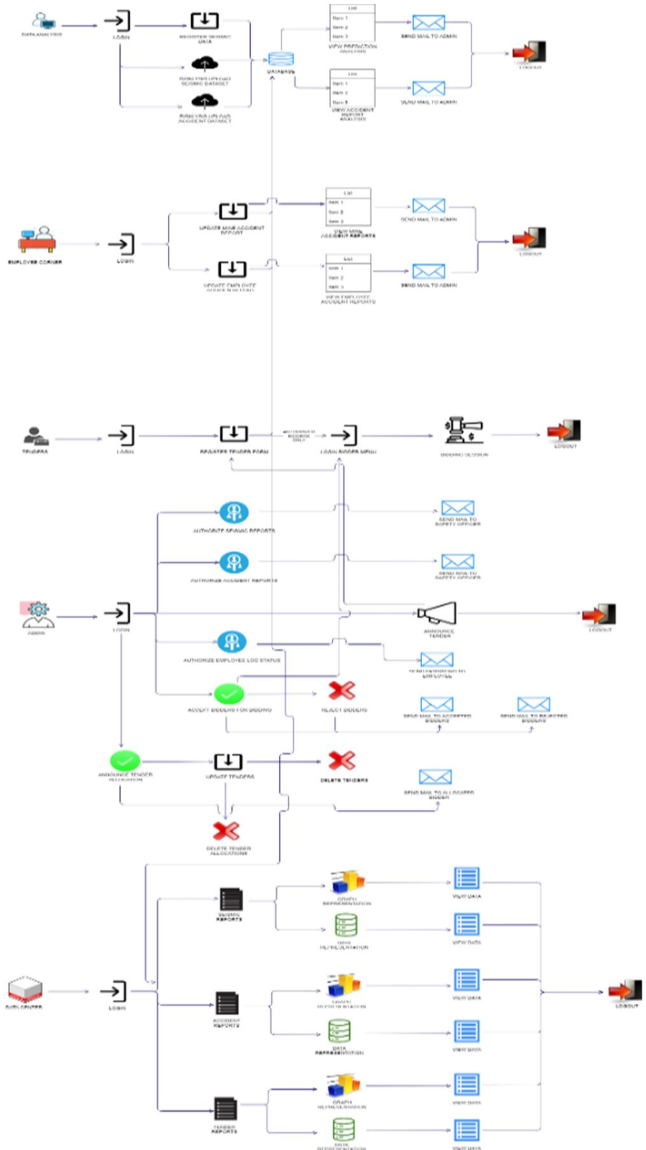
Daily seismic analysis can be done by Safety Officer by easily uploading the data into the web application. Report data can be successfully analyzed for risk assessment. The prediction model uses semantic analysis for dataset we stored in a table, predicts simultaneously and the results are shown immediately.

The data results are stored for other usage too. Immediately we can send email quickly to higher authorities for approval, also immediate action can be taken to reduce losses both financially and loss of lives. The algorithm used in this project is Naïve Bayes algorithm. It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple.

The Naïve Bayes algorithm determines the output of the risk assessment and seismic bump analysis. The individual accident reports the total casualties, fatal casualties, severe injuries, minor injuries are noted from Mine supervisor reports that is updated.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

VI. ARCHITECTURE DIAGRAM



VII. MODULES

The modules are performed using different techniques and they are described.

A. Data Analysis

In this module the Safety Officer can daily update the seismic bump activity that is happening under coal mines, The Safety Officer has to update details like Date, Time of seismic activity, Coal mine name, Seismic activity before, Shift type, G energy, G pulse, GD energy, N bumps, N bumps2, N bumps3, N bumps4, N bumps5, N bumps6, N bumps7, N bumps8, Max Energy. After updating these data the prediction reports values are updated with the Seismic after state and Hazard State, which are key predictions whether the next shift should be continued or not. The Safety Officer need not to take preventive measures as the Naïve Bayes Algorithm will predict the safety measures for the data recorded in the prediction reports. Safety officer sends email to admin in pdf format with all the data of seismic activity along with predicted results and preventive measures. In risk analysis we can upload dataset for bump analysis for multiple seismic data. The dataset from yearly accident reports also needs risk assessment. So yearly accident dataset is also uploaded into the Accident report analysis for predictions like probability and severity. The preventive measures also sent to email by safety officer to admin for approval of safety measures. The probability and severity from yearly accident data determines the safety of coal mines. The seismic reports determine the dangers upcoming before the shift gets started.

B. Employee Corner

In Employee Corner the Mine Supervisors update their accident reports, the data collected from the accident reports like total casualties, fatal casualties, severe casualties, minor injuries are taken into account. Accidents cannot be predicted but can be preventable. The accident reports are randomly occurring; it is divided into two types of reports, Employee error accidents and Natural error accidents. In Report Portal, Mine Supervisor selects the type of report Employee Error or Natural Error, If Employee Error selected, the form has details like Name of the reporting person, Date of accident, Name of Coal Mine, Location of accident, Time of the accident, Shift Type. What error employee done in detail, Total No of casualties, Fatal casualties, Severe injuries Minor injuries, Total amount of loss in damage, Proof of the accident, Action needs to be taken. The data is uploaded into Employee Error reports, where we can send the report to admin for approval for action. We can view the image in the Employee reports. If natural accident like roof fall, side wall fall, poisonous gases, coal burst, Manual handling of objects, etc. happens it is reported in the Natural error. The date and time of accident, Name and location of coal mine where the accident occurred, Type of accident, Shift type, cause of accident, cause of accident in detail, Total casualties, fatal casualties, severe and minor injuries, total amount of loss in damage, Proof of the accident, Action needed to be taken are updated and sent to admin for approval.

C. Tenders

In Tenders module, once admin announces tender, bidders can apply for tender application in the announcements menu, Bidders will fill details like their company name company's license holder's name, category of the company like Micro Unit, Small Unit, Medium Unit, as per MSME(Ministry of Micro Small and Medium Enterprises),Ancillary Unit, Project affected person of the company, Company Address, City, State, Legal Status of the company like Limited company, Undertaking, Joint Venture, Partnership, etc. Contact person details also taken for registration for contacting details about bid via email like email, name and phone number of the contact person. Admin checks the details and bidders can see the status of their application in bidder status. In bidding option, bidders can enter the bid by admin allocated password after admin accepted their applications. The bidding can be logged in at certain date and time and it's a session only. Two types of bidding is happening, Service Provider and Business process. Service provider is for tenders which are based on service for a certain amount of time. Service bid process has a bid limit at which is least required for doing the service process at coal mine industry. In Business process, there is no value limit for bidding but a tender base value is there. Tender winners can be seen in tender allocations after the bidding is finished. It is declared by the admin.

D. Admin

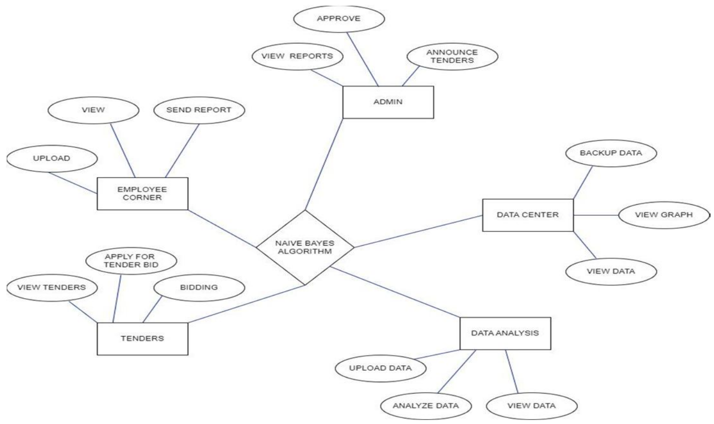
Admin has to authorize all the actions that needs to be taken from Data Analysis, Employee Corner, Tenders modules, In S/O Damage reports, the Seismic reports he needs to give authorization for immediately stopping the work shift in coal mine. The Accident reports from yearly analysis and a preventive measure is also authorized by admin by sending mail to the corresponsive person. The Employee accident reports and Natural accident reports also needs to be authorized by admin, The log status of employees needs to be authorized by admin by allocating a password that can be used to see confidential information, employees can access information only by admin allocated password and cannot be done by others. In Tender Auctions, the bidder status of whether to accept the application for bid or not is decided by admin, Tender announcements is updated in tender status, Tender allocations is done in manage tender allocations option by clicking the allocate button and automatically announces the winner of the bid. In Announce tender admin can announce tender for bidders to apply for it. In Tender Allocation manager admin can delete the bidder allocated for the tender, if needed

E. Data Center

In Data center, all the data from all the modules are stored here, Data Analyst uses these data for analysis, if any data that is not present can be checked here, the seismic reports has both representations of data, Graph representation and Data representation, In Graph representation Data Analyst enters the date and can view the graph by clicking on view. All the results of seismic data can be seen in data center, we can individually view data in graph too. Accident reports can also be seen in both data representation and graph representation. In Graph representation enter a year to view the data of total, fatal, severe, minor injuries and casualties. In Tender reports, bidder applying for bid details can be viewed. Tender Allocation reports can also be seen. Data Center purpose is to store all data from all modules which can be useful later when needed in this module, Data reports are classify into graph representation and data representation to view the to view the data. The data center collects and maintains all the bulk of data in the manner way.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

VIII. ENTITY RELATIONSHIP DIAGRAM



IX. CONCLUSION

Inspired by the network embeddings, we propose a semantic network encoder, namely SemNE, for associated fact prediction. SemNE can be generalized to either binary semantic networks or multiplex semantic networks. It explores the word-level features of a fact to enrich the network representation. Also, a word self-organization method is proposed to fuse the word embeddings into the fact representations. SemNE obtains the best performance in all semantic networks we used. But the sampling strategy and using BERT model. We will continue to improve the performance of SemNE in multiplex semantic networks and more complicated semantic networks.

REFERENCES

[1] Y Li, L. Lu, Q Yin, and B Zhang, “Robot Task Planning Based on State Semantic Network,” International Conference on Intelligent Computation Technology and Automation (ICICTA ’17), pp. 420–424Oct. 2017, doi: 10.1109/ICICTA.2017.101.

[2] P He, and G Karabatis, “Using semantic networks counter cyber threats,” IEEE International Conference on Intelligence and Security Informatics, pp. 184–184, 2012, doi: 10.1109/ISI.2012.628429.

[3] I Al, and A Melton, “Semantic-Based Text Document Clustering Using Cognitive Semantic Learning and Graph Theory,” IEEE 12th International Conference on Semantic Computing (ICSC18), pp. 243–247, Jan. 2018, doi: 10.1109/ICSC.2018.00042.

[4] “Microsoft Concept Graph for Short Text Understanding”, available at https://concept.research.microsoft.com/, Microsoft website, 2010.

[5] “Concept Net”, available at https://github.com/commonsense/conceptnet5/, Git hub, 2018.

[6] M Li, D Hu, Z Li, and A Lei, “Study on a Vocabulary Learning System Based on Semantic Network,” IEEE International Conference on Service Operations and Logistics, and Informatics, pp. 1004–1008, June. 2006, doi: 10.1109/SOLI.2006.328888.

[7] R J. Brachman, “What's in a concept: structural foundations for semantic networks”, International journal of man-machine studies,pp. 127-152, 1977.

[8] W A. Woods, “What's in a link: Foundations for semantic networks”, Representation and understanding, pp. 35-82, 1975.

[9] A Rossi, D Firmani, A Matinata, P Merialdo, and D Barbosa,“Knowledge Graph Embedding for Link Prediction: A Comparative Analysis”, arXiv preprint arXiv: 2002.00819, 2020.

[10] H Cai, V W. Zheng, and K C. Chang, “A Comprehensive Survey of Graph Embeddings: Problems, Techniques, and Applications”, IEEE Transactions on Knowledge and Data Engineering, pp. 1616-1637, Sept. 2018, doi: 10.1109/TKDE.2018.2807452.

[11] Q Wang, Z Mao, B Wang, and L Guo, “Knowledge graph embedding: A survey of approaches and applications”, IEEE Transactions on Knowledge and Data Engineering, pp. 2724-2743, Dec. 2017, doi: 10.1109/TKDE.2017.2754499

[12] P Cui, X Wang, J Pei, and W Zhu, “A Survey on Network Embedding”, IEEE Transactions on Knowledge and Data Engineering, pp. 833-852, May. 2019, doi: 10.1109/TKDE.2018.2849727.

[13] C Yang, Z Liu, D Zhao, M Sun, and E Y. Chang, “Network representation learning with rich text information”, Proc. 24th international Joint Conference on Artificial Intelligence (IJCAI 2015), pp. 2111-2117, Jan. 2015.