Correlation Analysis and Text Classification of Chemical Accident Cases Based on Word Embedding

Sifeng. Jing[a,c], Xiwei. Liu[a,c], Xiaoyan. Gong[a,c], Ying.Tang[c,d], Xiaoyan. Sun[b], Rongshan. Bi[b*]

a. The State Key Laboratory for Management and Control of Complex Systems,Institute of Automation, Chinese Academy of Sciences ,Beijing 100190, China

b. The College of Chemical Engineering, Qingdao University of Science and Technology, Qingdao; 266042, China

c. Institute of Smart Education Systems, Qingdao Academy of Intelligent Industries,Qingdao 266044, China

d. Department of Electrical and Computer Engineering, Rowan University,Glassboro, NJ 08028, USA

* Corresponding author E-mail addresses: birongshan@163.com

***Abstract*—the information of accident precursors and common trends can help people estimate risk and avoid future accident in chemical plants. This information can be obtained by analyzing chemical accident cases. Word embedding and deep learning methods have shown to be very effective in automated text information mining nowadays. In this work, a chemical accident case analysis method is proposed and tested based on word embedding and deep learning method. Firstly, word vectors for the text corpus of chemical accident cases were produced based on *word2vec.* And then Bidirectional LSTM model with attention mechanism for text classification was constructed. Finally, case studies on the correlation analysis of the common trends of chemical accidents and automated text classification of chemical accident cases documents were performed respectively. The results revealed that the proposed method can identify common principles of chemical accidents and classify chemical accident cases. These findings highlight the feasibility of chemical accident case information mining based on word embedding through case studies. Compared with rule-based information mining, our method improves the efficiency, automation, and intelligence of information mining from chemical accident cases documents.**

***Index Terms*—word embedding; text information mining of chemical accident cases; automated text classification; deep learning; chemical safety management;**

I. Introduction

In the petrochemical industry, the chemical processes are complex, the production equipment is diverse and the materials involved in the production process are flammable, explosive, and toxic. Hence, accidents have been a regular occurrence in the petrochemical industry for many years, such as the Bhopal disaster (1984), the accidents that occurred in Jilin, china (2005) [1], the Qingdao 11.2 crude oil leak and explosion (2013), the Tianjin Port 8.12 fire and explosion (2015) [2] and especially serious explosion accident in Xiangshui, China[3]. These disasters resulted in numerous casualties, substantial economic losses and environmental pollution. So it is significant to mine the factors triggered these accidents and understand the likelihood of the cases, which can aid chemical engineers in making better decisions regarding risk prevention and accident control activities, avoiding and reducing recurring accident in petrochemical enterprises.

To avoid accident, projects involving the inherent safety risk assessment of chemical processes [4-7] as well as production safety risk assessment [8, 9] have received increased attentions from chemical engineering researchers and engineers in the chemical industry. The information of accident precursors and common trends can improve efficiency and accuracy of safety risk assessment of chemical process. It can help site managers identify work-areas and instances where the likelihood of an accident may occur [10, 11]. Extracting these valuable information from a huge amount of chemical accident cases is an effective way to enhance safety management in chemical enterprises. Significant advancements in chemical safety management have been achieved [12, 13] by extracting the precursors and causes from these accidents.

Within a chemical accident case, precursors and causes are typically documented in an unstructured or semi-structured free-text data format, which contains information such as the description of the accident cases, its location，time， equipment and people involved. Thus, mining valuable information from a huge amount of chemical accident cases remains a time-consuming, labor-intensive and inefficient process[14－15].

In recent years, new paradigms[17-19] for data processing，representation, mining, and application have emerged in the intelligent era. Facing major obstacles related to safety management in green manufacturing, artificial intelligence (AI) technologies are quite promising [20, 21].Deep neural networks have become an effective way to solving complex problems such as image analysis, information extraction and text mining [22, 23].On the other hand, Word2Vec techniques have been proved efficient words representation approaches learning semantic information from massive unlabeled data[24]. And it was wildly applied to downstream natural language processing tasks [25-27], providing the possibility of efficient and automated information extraction from a large amount of existing cases text. At present, word embedding has been applied to many fields in engineering research. Material related word vectors can be obtained by training large amounts of materials science literature using unsupervised methods. These vectors capture complex materials science concepts and can be used to recommend materials for functional applications several years before their discovery in the materials science field [28]. Pre-trained word vectors are used to analyze the correlations among traffic roads in the transportation field [29]. Safety researchers use pre-trained word vectors to extract the ontology concepts of coal mine accidents in the coal mining industry [30]. Compared with the previous TF-IDF (term frequency–inverse document frequency)method and bag-of-words model, word representation technology has greatly improved precision and recall of Chinese event discovery, and protein interaction recognition has been efficiently and accurately performed [31-33].

Therefore, in this study, we utilized Word2Vec model to obtain word-vectors of chemical accident cases text corpus firstly, then identified the accident common trends by correlation analysis of accident cases based on word vectors of chemical accident cases text. Finally, a deep neural network was constructed to perform automated text classification based on word vectors of chemical accident cases text.

II. METHODOLOGY

Figure 1 show the main framework of correlation analysis and text classification of chemical accident cases based on word embedding in our work. This work is mainly divided into two major tasks: (1) obtaining the word vectors of the text corpus of chemical accident cases and (2) performing a case study of correlation analysis and text classification of chemical accident cases based on word embedding. Strategies to produce word vectors for the cases text corpus and models to classify chemical accident cases text are detailed in this section. The critical steps of obtaining word vectors mainly include case collection, case text prepossessing, case text corpus word2vec model construction and model evaluation. The case study consists of two parts: correlation analysis of the accident common trends and automated text classification of chemical accident cases.

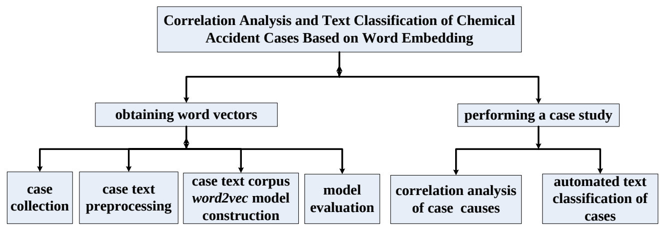

Figure 1. Main framework of word vector-based chemical accident case analysis

1. Strategies to produce word vectors for the cases text corpus

At present, word vectors can be used as a feature in many natural language processing and machine learning applications. Word vectors are derived from the distributional hypothesis, that is, words occurring in similar contexts have similar meanings [34-36]. They are generated during the training of deep language probabilistic models. This weakens the interpretability of statistical machine learning methods and relies heavily on data. The greater of the amount of data is, the better the universality shows. Two distinct neural models can be used in *word2vec*: the continuous bag-of-words (CBOW) model and the continuous skip-gram (skip-gram) model. Generally, obtaining word vectors for a domain-specific text corpus requires the extensive training and fine-tuning of multiple parameters. In this study, a series of experiments will be carried out to select a suitable word2vec model for text corpus of chemical accident cases.

In this study, the *Word2vec model* was called from the *Gensim module in* Python to generate word vectors for the text corpus of chemical accident cases based on *Anaconda Spyde*r. The word vectors generation steps for the corpus of chemical accident cases are shown in Figure 2. Theses mainly consist of information collection of chemical accident cases, text corpus preprocessing, case text corpus *word2vec* model construction, and *word2vec* model evaluation.

1.1 Information collection of chemical accident cases

A total of 10000 accident cases happened from 1988 to present were collected and a vocabulary with a capacity of 1G was obtained after a preliminary screening in this work. These chemical accident reports were compiled by major Chinese chemical companies. 3000 of them were collected from the chemical factories, such as Qilu Petrochemical, Maoming Petrochemical, Petro China Fushun Petrochemical. Another 7000 cases were obtained by crawling the websites ChemMade.com, safehoo.com and the China Chemical Safety Association. The keywords used to crawl these websites were “chemical accidents,” “chemical production process," and "petrochemical industry." Half of them actually occurred, whereas the other half were discovered in a timely fashion and effectively controlled.

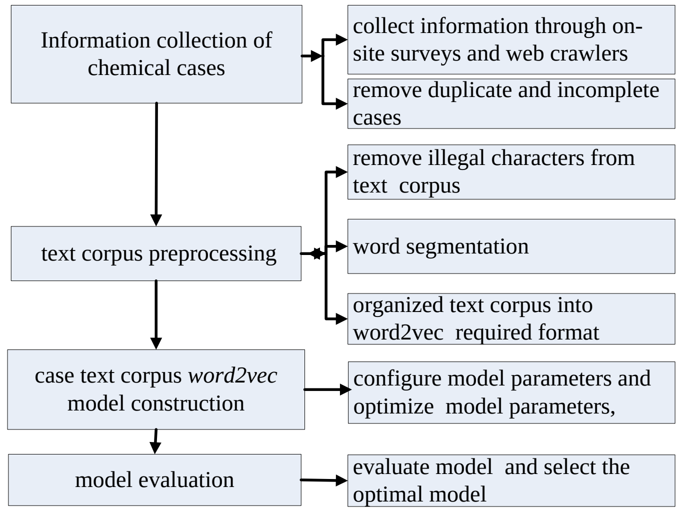

Figure2. Steps in word vector generation from the text corpus of Chemical accident cases

1.2 Text corpus preprocessing

Spaces, line breaks, letters, punctuation, and stop words in documents have no practical meaning and reduce the training efficiency and quality of the word vectors obtained. Therefore, it was vital to remove these characters from the training corpus.

First, the *re* module in Python was called to remove spaces, line breaks, letters, punctuation and other illegal characters from the text corpus. Next, to improve the efficiency and accuracy of segmentation, a proprietary dictionary of chemical vocabulary was constructed and "*jieba* segmentation" was utilized to accurately segment the preprocessed text of chemical accident cases into words in our project. Then, Stop words were then removed, and a chemical text corpus of approximately 1.35 million words was finally obtained（see Attachment 1 for details）.

The CBOW and skip-gram models both consider semantic coherence when training word vectors. To this end, in this study the text corpus was processed sentence - by - sentence from text preprocessing, to text segmentation, to the removal of stop words. Ultimately, the text corpus was organized into a list of list format, which was then fed to the *word2vec* model.

1.3 Case text corpus word2vec model construction

Word2vec primarily uses skip-gram and CBOW to learn word representations. CBOW mainly predicts a target word based on its context words, whereas skip-gram mainly predicts context words based on a given target word. Both of these methods use artificial neural networks as classification algorithms. For CBOW, a key step is comparing its output with the word itself in order to correct its representation based on the back propagation of the error gradient. For skip-gram, a key step is comparing its output with each word of the context in order to correct its representation based on the back propagation of the error gradient. These are all related to vocabulary size and the window size of each word. In order to improve the efficiency of learning word representations, the negative sampling algorithm and the hierarchical softmax algorithm were used in the learning process. In *word2vec,* however, CBOW, skip-gram, the negative sampling algorithm and the hierarchical softmax algorithm have their own advantages and disadvantages for different training data. Therefore, evaluations of the performance of different models and algorithms were conducted in order to determine the most efficient *word2vec* model for the text corpus of chemical accident cases.

In our study, *word2vec* model parameters such as the CBOW model, skip-gram model, negative sampling algorithm, hierarchical softmax algorithm, number of iterations, word vector dimension, and window size, were all considered. The performance of the word vectors obtained under different parameter configurations are presented in the model evaluation section.

1.4 Word2vec model evaluation

Word vector training is an unsupervised process based on *word2vec.* Generally, there are two methods for evaluating word vectors: one is to apply word vectors to an existing natural language processing system and then draw conclusions by observing system evaluation indicators such as the precision, recall and F-score of word vector accuracy. The other is to evaluate the quality of word vectors in terms of linguistics. In order to decrease cost and improve efficiency, model evaluation was conducted from the perspective of linguistics in this study. Specifically, the mean average precision (MAP) [37] of the closest 50 words to the target word was used as the evaluation criterion. The following formulas were employed to compute the MAP:

$$\text{MAP} = \frac{\sum_{n=1}^{N} AveP(q_n)}{N}. \quad (1)$$

$$AveP(q) = \frac{\sum_{i=1}^{R} M_i}{2\times 20R}. \quad (2)$$

In these equations, N represents the number of selected target words, $AveP(q_n)$ is the average precision of each target word, R is the number of words closest to the target word, and M is the manual evaluation score of the similarity between each word and the target word. In our experiment, N = 5, R = 50, and M was measured as follows:

1) If a word was completely unrelated to the target word, then M = 0;

2) If a word was roughly related to the target word, then M = 1; and

3) If a word exactly matched the target word, then M = 2. The model evaluation process is illustrated in Figure 3.

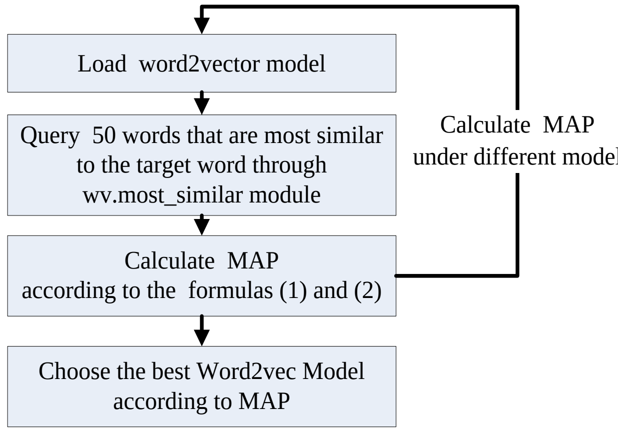

Figure 3. Flowchart of model evaluation

In our study, the CBOW model, skip-gram model, number of iterations, hierarchical softmax and negative sampling optimization tricks, vector dimension, window size and min_count were the key model parameters used to learn the 300-dimensional word embedding of the vocabularies related to chemical accidents. Based on the actual demand for chemical accident case analysis, the selected evaluation target words were set to “accident,” “fire,” “heat exchanger,” “air respirator,” and “ethylene”; the model evaluation results are presented in Figure 4-6. Figure 4 and Figure 5 shows changes of MAP with number of iterations under the CBOW model and the skip-gram model respectively. Figure 6 shows changes of MAP with different optimization tricks under the CBOW model. From these results, it can be seen that for the text corpus of chemical accident cases, the CBOW model performed better than the skip-gram model, 30 was the optimal number of iterations and the negative sampling trick was better than hierarchical softmax trick for training. Therefore, in our study, the parameters of the *word2vec* model used to learn the word embedding for the text corpus of chemical accident cases were those listed in Table Ⅰ.

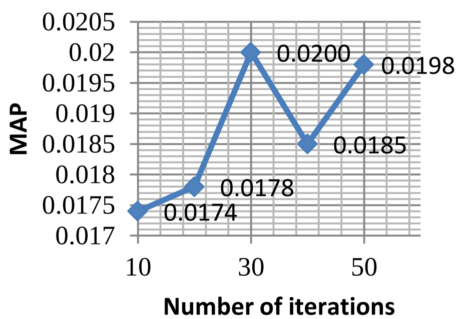

Figure 4. Number of iterations versus MAP under the CBOW model

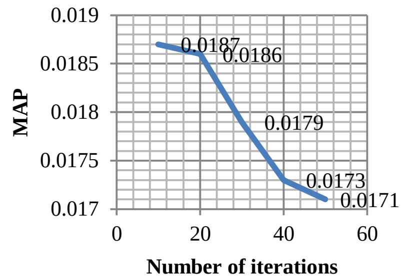

Figure 5. Number of iterations versus MAP under the skip-gram model

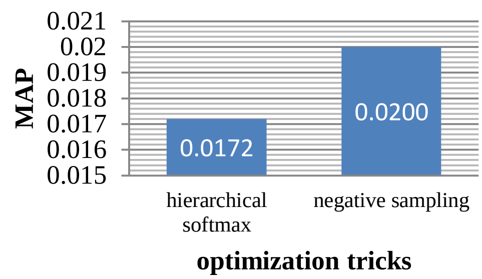

Figure6. Optimization tricks versus MAP under CBOW model

Table Ⅰ SETUP OF THE *WORD2VEC* MODEL PARAMETERS

Parameter	Set-up
sg	0(CBOW)
Vector dimension	300
Number of iterations	30
hg	0(negative sampling)
Window size	10
Negative	5
Sample	1*10-3
min_count	5

2. Model of automated text classification of chemical accident cases

Text classification of Chemical accident cases is essential for risk assessment of chemical process. It can improve the efficiency of information mining and provide valuable information for the supervision and early warning of chemical accident in chemical safety management. Recently, deep learning techniques have shown impressive results in text classification. Convolutional neural networks (CNN)[38],Recurrent neural network (RNN)[39] and Long Short-Term memory (LSTM) [40] are the most commonly used deep learning techniques in automated text classification.

However, CNN cannot obtain the semantic relationship between non-adjacent words in a sentence because CNN's filter has a limited word capacity, while RNN may cause gradient explosion or gradient disappearance for long data sequence. In addition, LSTM can capture long-distance and close-range dependencies between word sequences[41].

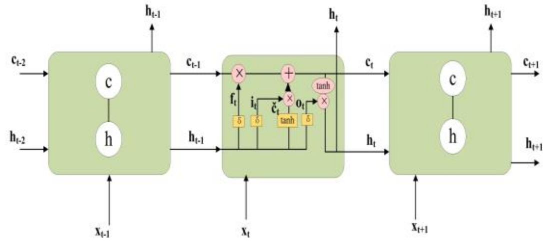

Figure7.LSTM topology

A topology of LSTM is shown in Figure 7. *h* and *c* are two state of LSTM cell. *c* denotes cell output state of previous time step. *h* denotes output of hidden layer. At each iteration *t*, the current cell input x_t,output of previous time step h_{t-1} and cell output state of previous time step c_{t-1} are incorporated by LSTM cell to update network parameters during the training. Gates mechanisms are introduced to control cell states of LSTM by allowing information to pass through optionally. i_t, f_t, o_t, $\check{c}_t$ denote the input gate, forget gate ,output gate and output state of current cell respectively. Values of gates and the input state of cell are calculated by Equations(3-6).

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (3)$$

$$\mathrm{f_t} = \sigma(\mathrm{W_f x_t} + \mathrm{U_f h_{t-1}} + \mathrm{b_f}) \quad (4)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

$$\check{c}_t = \tanh(W_c x_t + U_o h_{t-1} + b_c) \quad (6)$$

$$c_t = f_t * c_{t-1} + \check{c}_t * i_t \quad (7)$$

$$\mathrm{h_t} = \mathrm{o_t} * \tanh(\mathrm{c_t}) \quad (8)$$

where w_i ,w_f, w_o,w_c denote the weight matrices between the input of hidden layer, input gate, forget gate, output gate and input cell state. U_i,U_f,U_o,U_c denote the weight matrices between previous cell output state, input gate, forget gate, output gate and input cell state.b_i ,b_f, b_o ,b_c denote the corresponding bias vectors. σ is the sigmoid function. And then,at each iteration t, cell output state c_t and LSTM layer output h_t are calculated by Equations (7 and 8).

To improve the accuracy of text classification of chemical accident cases, two key factors were considered in this study. The first one is how to represent the contextual semantic information of a sentence, and the other is how to express the important part of a sentence for a given chemical accident cases text.

Bi-directional Long Short-Term Memory (BiLSTM) can process sequence data in both direction with forward LSTM and backward LSTM layer and these two hidden layers are connected to the same output layer. Attention mechanism has gained great success in many text classification tasks [42,43].The core idea of attention mechanism is to assign a weight to each position at a lower level of a neural network when computing a higher level of representation [44]. To extract the dependencies feature between word sequences and the contextual semantic information of the chemical accident cases text, a BiLSTM text classification model with attention mechanism was constructed to classify document-level chemical accident cases text as shown in Figure 8.

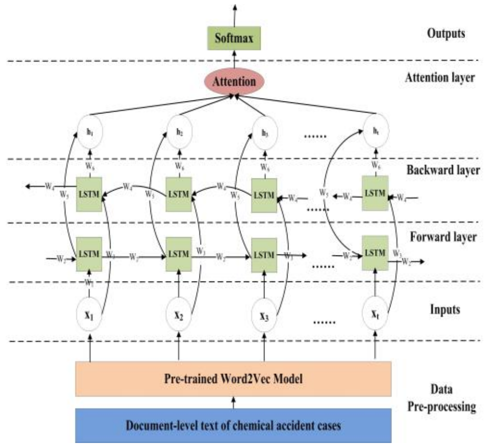

Figure8. Model architecture of chemical accident cases text classification

In this model, firstly, each text of chemical accident cases was reprocessed and represented sentence by sentence based on pre-trained Word2Vec model. Secondly, each text vector was feed to the forward LSTM cell and backward LSTM cell sentence by sentence. Thirdly, the output vectors h_t of BiLSTM are connected and transformed according Equations (9). An attention weight vector α_t was obtained by Equation (10). The final representation of a text d is calculated by Equation (11) .

$$u_t = \tanh(W_u u_t + b_u) \qquad (9)$$

$$\alpha_t = \mathrm{softmax}\left(\frac{u_t u_{t-1}}{\sqrt{(u_t - u_{t-1})(u_t - u_{t-1})^T}}\right) \qquad (10)$$

$$d = \sum_t \alpha_t u_t \qquad (11)$$

III. Case study

In this section, pre-trained word vectors for a chemical accident case text corpus were applied to chemical accident case analysis. Correlation analysis of the common trends of chemical accident cases and automated text classification of chemical accident cases were then carried out respectively.

A. Correlation analysis of common trends of chemical accident cases

Accident-related factors, such as root causes, accident time, accident-related equipment, accident type, accident-related materials, and accident characteristics can help chemical safety management personnel to quickly understand the basic situations of accidents and grasp the accident common trends. This helps to enhance the safety awareness of the relevant chemical production personnel, and promotes chemical companies to strengthen their own safety management strategies. In this case, the correlation between the word vectors was used to obtain the accident correlation information, that is, “accident” was selected as the target word and 50 associated words were then found. The word cloud diagram is illustrated in Figure 9.

From the term frequency of associated words with “accident” in Figure 9, it can be seen that the main characteristics of chemical accidents were similarities, recurrences, and emergencies. Mistakes, illegal operations, out of control, poor management, and negligence were closely related to chemical accidents. In the case of accident type, explosions, fires and poisoning accidents were high-frequency chemical accident. In addition, chemical accidents often led to environmental pollution. These are in line with people’s understanding of chemical accidents [44] and theories of accident causation [45]. Hence, these results highlight the possibility of using word embedding to obtain the common trends of chemical accidents.

Figure 9. Word cloud for accident- related information

B. Automated text classification of chemical accident cases

In this case, *tensorflow2.3.0* module was imported to construct a BiLSTM model with attention mechanism to classify case text of chemical accident based on *Anaconda Spyder.*

A series of experiments were conducted on a labeled dataset in a supervised learning setting. Each text of chemical accident cases was manually labeled. The labels consist of Fires, Explosions, Poisoning, Fall and Others according to correlation analysis in Case A.

The input of the model is a document-level text of chemical accident cases. Each sentence in the text is represented by the vector sum of each word in this sentence based on pre-trained word vectors. And then we merged sentence-level features from each time step into a document-level features vector through attention layer. Finally, a document-level vector feature vector is used for case text classification of chemical accident cases. Maximum number of sentences of the text is 60 in our datasets.Some chemical accident cases were attached in Attachment 2 for details.

The model parameters is shown as Table Ⅱ.The result of text classification is shown in Table Ⅲ.

Table Ⅱ BILSTM MODEL PARAMETERS

Parameter	Value
Input tensor	[60,60,300]
n_classes	5
Hidden layer	2
Hidden size	128
Embedding size	300
Learning rate	0.001
Opertimizer	Adam
Batch size	60
epoch	30
Dropout rate	0.5
rate	0.8

Table Ⅲ RESULTS

Model	Precision(P)	Recall(R)	F(F_score)
BiLSTM with Attention	0.9134	0.9026	0.9079

The performance of the model is evaluated by Precision(P)， recall(R) and F_score(F). The calculation formulas of P, R and F are as equation（12-14）:

$$P=c1/c2\times 100\% \quad (12)$$

$$R=c1/c3\times 100\% \quad (13)$$

$$F=(\beta^2+1)PR/(\beta^2 P+R)\times 100\% \quad (14)$$

And here c1 means the number of returned correct results, c2 means the number of all returned results, c3 means the number of results that should be returned. Here P represents the correct rate of the classification results, and its value is calculated by formula (12), while R is the recall rate of the classification results, and its value is calculated by formula (13). The F value is used to measure the importance of R and P. The value of β represents the contribution of R and P to the measurement standard. In this paper, R and P are considered to be equally important for evaluating the extraction effect. The value of β is 1, and F value is calculated by formula (14).

It can be observed from Table III that the BiLSTM model with Attention mechanism can classify the chemical accident cases precisely based on pre-trained word vetors of chemical accident cases text.It was noted that pre-trained word vectors ensure the stability of the training process,BiLSTM model with attention mechanism can learn more hidden feature information and taking sentence-level features vectors as input improved the computational efficiency of the model.

IV. CONCLUSIONS

In this study, word vectors were obtained and applied to examine the correlation analysis of common characteristics of chemical accident cases, as well as the automated text classification of chemical accident cases. Our results revealed the following: (1) Using the correlation among word vectors can effectively identify the key causes of chemical accidents, common trends and other information. (2) The BiLSTM model with Attention can precisely classify the chemical accident cases based on pre-trained word vetors of chemical accident cases text. (3) Taking sentence-level features vectors as input of BiLSTM model with attention mechanism improved computational efficiency.(4) Our results highlight the possibility of analyzing chemical accident cases based on word embedding.

ACKNOWLEDGMENT

This work was supported in part by Natural Science Foundation of China Projects (61533019) and Science and Technology Innovation Project of Hunan Province (2018GK1040) .Engineers Yan Feng, Chuanzhong Xu and Mr. Li Yulin for their excellent co-operation and assistance during the work are acknowledged.

REFERENCES

[1] Aimin, WANG, Chun, et al. Phototransformation of nitrobenzene in the Songhua River: Kinetics and photoproduct analysi[J]. Journal of Environmental Sciences, 2008.

[2] Bing Wang, Chao Wu, Genserik Reniers, et,al. The future of hazardous chemical safety in China: Opportunities, problems, challenges and tasks[J], Science of The Total Environment, 2018,643():1-11.

[3] https://www.chinanews.com/gn/2019/11-15/9008586.shtml

[4] Xiaolong Yu, Yin H , Ye J S , et al. Degradation of tris-(2-chloroisopropyl) phosphate via UV/TiO2 photocatalysis: kinetic, pathway, and security risk assessment of degradation intermediates using proteomic analyses[J]. Chemical Engineering Journal, 2019

[5] Li K , Cui W , Li J , et al. Tuning the reaction pathway of photocatalytic NO oxidation process to control the secondary pollution on monodisperse Au nanoparticles@g-C3N4[J]. Chemical Engineering Journal, 2019, 378:122184.

[6] Song J , Wang X , Ma J , et al. Removal of Microcystis aeruginosa and Microcystin-LR using a graphitic-C 3 N 4 /TiO 2 floating photocatalyst under visible light irradiation[J]. Chemical Engineering Journal, 2018:S1385894718307551.

[7] Chen J , Zhang D , Li G , et al. The health risk attenuation by simultaneous elimination of atmospheric VOCs and POPs from an e-waste dismantling workshop by an integrated de-dusting with decontamination technique[J]. Chemical Engineering Journal, 2016:299-305.

[8] Wang B , Qian F , Zhong W . Wind field reconstruction for the dispersion modeling of accidental chemical spills on complex geometry[J]. Chinese Journal of Chemical Engineering, 2019, 27(11).

[9] Kwanghee Lee, Hyuck-myun Kwon, Seungsik Cho,et,al. Improvements of safety management system in Korean chemical industry after a large chemical accident[J]. Journal of Loss Prevention in the Process Industries, 2016, 42:6-13

[10] Jinsong Zhao, Johanna Suikkanen, Maurren Wood. Lessons learned for process safety management in china[J].Journal of loss prevention in the process industries, 2014,29():170-176.

[11] Kamarizan Kidam, Haslinda A. Sahak, Mimi H. Hassim, Haslenda Hashim, Markku Hurme. Method for identifying errors in chemical process development and design base on accidents knowledge[J]. Process Safety and Environmental Protection,2015,.97:49-60.

[12] Valeria Casson Moreno, Genserik Reniers, Ernesto Salzano, Valerio Cozzani. Analysis of physical and cyber security-related events in the chemical and process industry[J]. Process Safety and Environmental Protection, 2018,116() :pp 621-631.

[13] Ching-Wu Cheng, Hong-Qing Yao, Tsung-Chih Wu. Applying data mining techniques to analyze the causes of major occupational accidents in the petrochemical industry[J]. Journal of Loss Prevention in the Process Industries, 2013,26(6):1269-1278.

[14] S.F.Jing,X.W.Liu,C.J.Cheng,X.Q.Shang,G.Xiong.Study on a Process Safety Management System -- Design of a Chemical Accident Database, in proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics, Qingdao,china,October 2014, pp.282-285.

[15] Sifeng.Jing, Xiwei.Liu,Chuanzhong.Xu. A Simple and Effective Method for the Use of Chemical Accident Cases, in proceedings of 2016 IEEE International Conference on Service Operations and Logistics, and Informatics, Beijing,china,July 2016, pp.235-239.

[16] JING Sifeng, XIONG Gang, LIU Xiwei, GONG Xiaoyan, HU Bin.Study on key information extraction of chemical accident cases.[J] Industrial safety and environmental protection.2019,45(08):61-65.

[17] Wang F Y， Zhang J J， Zheng X H， et al． Where Does Alphago Go: From Church-Turing Thesis to Alphago Thesis and Beyond． IEEE /CAA Journal of Automatica Sinica， 2016， 3(2) : 113－120．

[18] Liu X， Wang X， Zhang W S， et al． Parallel Data: From Big Data to Data Intelligence． Pattern Recognition and Artificial Intelligence， 2017， 30(8) : 673－681．

[19] Wang F Y， Yuan Y， Li J J， et al． From Intelligent Vehicles to Smart Societies: A Parallel Driving Approach． IEEE Transactions on Computational Social Systems， 2018， 5(3) : 594－604．

[20] Mao S , Wang B , Tang Y , Qian F. Opportunities and Challenges of Artificial Intelligence for Green Manufacturing in the Process Industry[J]. Engineering(English), 2019, 005(006):P.995-1002.

[21] Xu B , Cheng W , Qian F , et al. Self-adaptive differential evolution with multiple strategies for dynamic optimization of chemical processes[J]. Neural Computing and Applications, 2019.

[22] Syed Umar Amin, Mansour Alsulaiman, Ghulam Muhammad, Mohamed Amine Mekhtiche, M. Shamim Hossain.Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion.Future Generation Computer Systems,2019,101(): 542-554.

[23] Ozal Yildirim, Ulas Baran Baloglu, Ru-San Tan, Edward J. Ciaccio, U. Rajendra Acharya.A new approach for arrhythmia classification using deep coded features and LSTM networks.Computer Methods and Programs in Biomedicine,2019,176():121-133.

[24] Mikolov T,Chen K,Corrado G,et al. Efficient Estimation of Word Representations in Vector Space[J]. arXiv preprint:1301.3781,2013.

[25] Hinton G E. Learning Distributed Representations of Concepts.[C]. Eighth Conference of the Cognitive Science Society,1989.

[26] Y. Bengio,R. Ducharme,P. Vincent. A Neural Probabilistic Language Model[J] Journal of Machine Learning Research,2003,3:1137-1155.

[27] Derry Jatnika, Moch Arif Bijaksana, Arie Ardiyanti Suryani.Word2Vec Model Analysis for Semantic Similarities in English Words.Procedia Computer Science,2019,157():160-167,

[28] Tshitoyan, Vahe; Dagdelen, John; Weston, Leigh; et al. Unsupervised word embeddings capture latent knowledge from materials science literature. NATURE ,2019,571(7763), Pages: 95-106.

[29] LIU Kang，QIU Peiyuan，LIU Xiliang，et a1．Measuring Traffic Correlations in Urban Road System Using Word Embedding ModelEJ]．Aeta Geodaetica et Cartographica Sinica，2017，46(12)：2032—2040．DOI：10．11947／j．AGCS．2017．20170166．

[30] LIU Si-Han, LIU Xu-hong, LIU Xiu-lei et,al. Concept extraction of coal mine safety accident ontology based on word embeddings and CRFs[J].Coal

Technology, 2018,37(09):178-181.

[31] Bin Z , Linmei H U , Lei H , et al. Word Embedding Based Chinese News Event Detection and Representation[J]. Pattern Recognition and Artificial Intelligence, 2018.

[32] Yun-Lei P , Yun N . Feature Words Selection Based on Word Embedding[J]. Computer Technology and Development, 2018.

[33] Yang Xiaodong, Gao Lutao, Yang Linnan et,al. Research on data crawling and storage system of agricultural product price based on Hadoop platform[J].Computer Applications and Software, 2017,34(03):76-80.

[34] Peter D. Turney and Patrick Pantel. 2010. From Frequency to Meaning: vector space models of semantics. Journal of Artifical Intelligence Research, 37(1):141–188, January.

[35] Zelig Harris. 1954. Distributional Structure. Word, 10(23):146—162.

[36] Stephen Clark. 2015. Vector Space Models of Lexical Meaning. In Shalom Lappin and Chris Fox, editors,Handbook of Contemporary Semantics, chapter 16. Wiley-Blackwell, Oxford.

[37] ZHANG Kejun,SHI Taimeng,LI Weinan et,al.Word2Vec optimization strategy based on an improved statistical language model[J].Journal of Chinese information processing,2019,33(7):11-19.

[38] A I B , B Y L , C M C C , et al. Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification[J]. Artificial Intelligence in Medicine, 2019, 97:79-88.

[39] Errattahi R , Hannani A E , Salmam F Z , et al. Incorporating label dependency for ASR error detection via RNN[J]. Procedia Computer ence, 2019, 148:266-272.

[40] Messina R , Je´ roˆ me Louradour. Segmentation-free handwritten Chinese text recognition with LSTM-RNN[C]// International Conference on Document Analysis & Recognition. IEEE, 2015.

[41] A S Z , A Z C , A H C , et al. Adversarial training based lattice LSTM for Chinese clinical named entity recognition[J]. Journal of Biomedical Informatics, 99.

[42] Zhou P , Shi W , Tian J , et al. Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2016.

[43] Li L , Nie Y , Han W , et al. A Multi-attention-Based Bidirectional Long Short-Term Memory Network for Relation Extraction[C]// International Conference on Neural Information Processing. Springer, Cham, 2017.

[44] Zhou P , Shi W , Tian J , et al. Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2016.

[45] ZHOU Lin, FU Gui, LIU Xiyang. statistics and analysis on chemical accidents based on behavior –based safety theory[J].Journal of safety science and technology,2016,12(01):148-153

[46] FU Gui, WANG Xiu-ming, LI Ya. On the 24Model and the application of its causative codes to the analysis of the related accidents[J].Journal of safety and environment,2017,17(03):1003-1008.