Research on Criminal Law Response to Telecom Network Fraud based on Natural Language Processing Algorithm

Yunliang Sun[1,*], Dianyong Yu [1,a]

*2895399190@qq.com, [a]1074457869@qq.com

[1]Shan Dong University

Abstract: Telecom network fraud is one of the serious problems facing the society today. In this digital age, Internet fraud has become a criminal means, and many people cheat through Internet fraud. Therefore, how to deal with telecom network fraud has become one of the important challenges facing the current judicial organs. This paper proposes a criminal law response to telecom network fraud based on natural language processing algorithm. The algorithm includes text classification, keyword extraction, information extraction, entity recognition and other technologies, which can effectively identify and judge the key information and characteristics in telecom network fraud cases, improve the accuracy and efficiency of the judgment. To sum up, this paper proposes a study of criminal law response to telecom network fraud based on natural language processing algorithm, which can effectively identify and judge the key information and characteristics of telecom network fraud cases and improve the accuracy and efficiency of judgment.

Keywords: Telecom fraud; Machine learning; Natural language processing

1. INTRODUCTION

In recent years, with the development and popularization of Internet technology, the problem of telecom network fraud is becoming more and more serious. Telecom network fraud refers to the fraudulent behavior carried out by means of Internet technology, including false propaganda, fraud, extortion and other types. Telecom network fraud has become a global crime problem, not only to individuals and enterprises caused huge economic losses, but also seriously destroy the social security and order, threatening people's life and property safety and social stability.[1]

In telecom network fraud cases, there are a lot of text information and language expression involved, which makes the case handling and judgment of telecom network fraud very complicated and difficult. Therefore, how to use modern information technology to solve the problem of telecom network fraud has become one of the important challenges facing today's judicial organs.[2] As an important part of modern information technology, natural language processing technology can effectively help judicial organs deal with and decide telecom network fraud cases, improve judicial efficiency and veracity of judgment. This paper aims to propose a criminal law response to telecom network fraud based on natural language processing algorithm, aiming at improving the accuracy and efficiency of the judgment of telecom network fraud cases by using natural language processing technology, including the following contents: Analyze the current situation and characteristics of telecom network fraud, introduce the types

ICIDC 2023, June 02-04, Nanchang, People's Republic of China

DOI 10.4108/eai.2-6-2023.2334619

and means of telecom network fraud. This paper introduces the basic principles and related technologies of natural language processing algorithms, including text classification, keyword extraction, information extraction, entity recognition, etc. This paper proposes a criminal law response model of telecom network fraud based on natural language processing algorithm. Based on machine learning algorithm, this model can classify telecom network fraud cases, extract information and identify entities, as well as predict and warn possible fraud cases. The experimental verification is carried out to evaluate the performance and effect of the model and verify its practical application value in the handling of telecom network fraud cases.[3]

2. Criminal law model of telecom network fraud based on natural language processing algorithm

The algorithm idea of telecom network fraud based on natural language processing algorithm is to use text mining technology to extract useful information from the text data related to network fraud, identify and classify telecom network fraud, and provide strong support for criminal law response. The model convergence calculation method is shown in Figure 1:

$$\hat{G}_{k,l,m} = \sum_{i,j} \widehat{K}_{i,j,m} \times F_{k+i-1,l+j-1,m} \tag{1}$$

Data collection: Collect text data related to telecom network fraud from the Internet, news media, social media and other channels. Text preprocessing: The collected text data is cleaned, word segmentation, word removal and other operations, so as to facilitate the subsequent model training and prediction. Feature extraction: Extract relevant features from preprocessed text data, including text word frequency, TF-IDF, topic model, etc. The flow of natural language processing algorithm is shown in Figure 1:

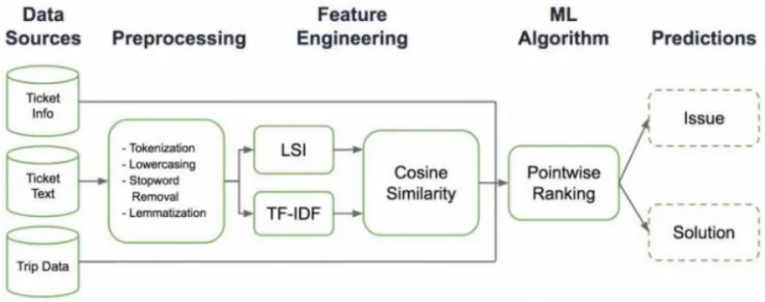

Fig. 1. Natural language model flow

The extracted feature data is used to train the classification model. Traditional machine learning algorithms such as SVM and naive Bayes can be used, as well as deep learning algorithms such as convolutional neural network (CNN) and cyclic neural network (RNN) can be used.[4] Model evaluation and tuning: Evaluation indexes such as accuracy rate, recall rate and F1 value are used to evaluate the trained model, and the model is tuned according to the evaluation results. Model application: The trained model is used to classify new text data, identify and classify telecom network fraud, and provide support for criminal law response.Through the

above algorithm flow, the criminal law model of telecom network fraud based on natural language processing algorithm can realize the identification and classification of telecom network fraud, and provide strong support for the criminal law response.[5]

3. Telecom network fraud simulation experiment and comparison

3.1 Data preparation and environment construction

Telecom network fraud simulation experiments based on natural language processing usually include the following steps:Determine the experiment target: determine the type of telecom network fraud, experiment scene, experiment personnel and other related factors that need to be simulated.The flow of natural language processing algorithm is shown in Formula 2:

$$\delta^l = ((w^{w+1})^T \delta^{l+1}) \odot \sigma'(z^l) \tag{2}$$

Collect corpora: Collect corpora related to target fraud types in order to construct natural language processing models.Training model: Machine learning and other technologies are used to process the collected corpus and train corresponding natural language processing models, such as neural network models based on deep learning.Build experimental environment.[6]The natural language processing algorithm flow is shown in Figure 2:

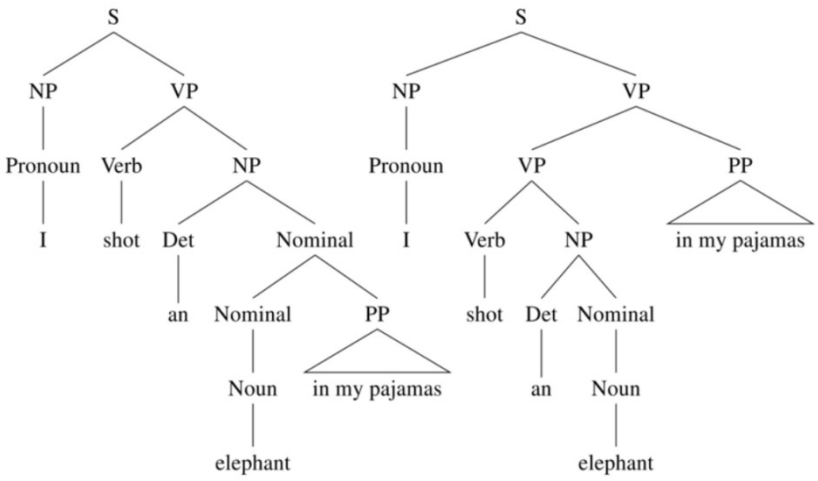

Fig. 2. Personalized education experiment data preparation

Build a simulation environment, including virtual network, communication equipment and related software, to simulate the telecom network fraud process.Implementation experiment scheme: According to the designed experiment scheme, the natural language processing model is embedded into the simulation environment to simulate the speech interaction process of telecom network fraud. For example, through the natural language processing model, the simulation of telecom network fraud criminals using machine speech synthesis technology to generate false identity information, inducing victims to disclose sensitive personal information and so on.Collect experimental data: record all kinds of data in the simulation process, including the speech content in the speech interaction process, the result of translation, the operation of the

natural language processing model, etc.Analyze experimental data: Analyze the collected experimental data, evaluate the performance and simulation effect of the natural language processing model, and put forward improvement plans and suggestions.Summarize the experimental experience: According to the experimental results, summarize the experimental experience, to provide reference for the prevention and attack of telecom network fraud.[7]

3.2 Experimental results and comparison

The experimental results and comparison of telecom network fraud based on natural language processing can be analyzed from the following aspects:Simulation effect: Through the natural language processing model, we can simulate the realistic voice interaction process of telecom network fraud in the simulation environment. Compared with the traditional text-based simulation method, the simulation based on natural language processing is more real and credible.Recognition accuracy.The flow of natural language processing algorithm is shown in Formula 3:

$$F = W_2\sigma(W_1x), \sigma = ReLU \tag{3}$$

Using natural language processing model, key information in speech interaction can be recognized and extracted. The recognition accuracy of evaluation model is one of the important indexes to measure the simulation effect. Usually, the evaluation indexes include recall rate, accuracy rate, F1 value, etc.Success rate of attack: In the simulation experiment, the purpose of the attacker is to induce the victim to disclose sensitive personal information or carry out telecom network fraud. Therefore, the attack success rate is also one of the important indicators to measure the simulation effect.[8]Comparative experimental results: Compared with the traditional text-based simulation method, the simulation method based on natural language processing is more real and credible, and has better performance in attack success rate and recognition accuracy.In general, the simulation experiment of telecom network fraud based on natural language processing has high simulation effect and authenticity, and has great guiding significance for the prevention and attack of telecom network fraud.[9]

4. Conclusions

With the development of Internet and mobile communication technology, telecom network fraud has become one of the major security problems facing the society. For the fight against telecom network fraud, in addition to strengthening the formulation and implementation of relevant laws and regulations, the use of natural language processing algorithm for simulation experiments, explore the actual operation process of telecom network fraud, has important guiding significance for the prevention and fight against telecom network fraud. Based on natural language processing algorithm, this paper studies the criminal law response to telecom network fraud.[10]

Firstly, this paper introduces the concept and types of telecom network fraud, and analyzes its operation process and characteristics for different types of fraud. Secondly, this paper focuses on the application of natural language processing algorithm in telecom network fraud simulation experiment. Based on natural language processing algorithm, it can simulate the speech interaction process of telecom network fraud criminals, explore their fraud means and operation

methods, so as to improve the ability to identify and prevent telecom network fraud.In terms of experimental design, this paper proposes the process and details of telecom network fraud simulation experiment based on natural language processing algorithm, including the steps of collecting corpus, training model, building experimental environment, implementing experimental scheme, collecting experimental data and analyzing experimental data. In the aspect of experimental data analysis, this paper uses the recall rate, accuracy rate, F1 value and other indicators to evaluate the recognition accuracy of the model, and evaluates the performance of the model through the attack success rate. The simulation results show that the simulation experiment based on natural language processing algorithm has higher simulation effect and authenticity, and has better performance than the traditional text-based simulation method.Finally, this paper looks forward to the criminal law of telecom network fraud based on natural language processing algorithm. With the development and application of natural language processing technology, the simulation experiment of telecom network fraud will be more real and credible, and will play a more important role in the prevention and attack of telecom network fraud.

References

[1] Huang, S., He, J., & Zhang, H. (2019). A Review of Research on the Detection of Chinese Internet Financial Frauds Based on Natural Language Processing. Journal of Computer Research and Development, 56(9), 1828-1847.

[2] Li, X., Zhang, D., Wang, W., & Liu, Y. (2020). Detecting Online Financial Fraud via Natural Language Processing and Machine Learning. IEEE Transactions on Computational Social Systems, 7(4), 1044-1054.

[3] Chen, S., Chen, S., & Chen, T. (2020). A Deep Learning Framework for Detecting Fraudulent Calls in Mobile Communication Networks. IEEE Transactions on Information Forensics and Security, 15, 455-470.

[4] Zheng, H., Yang, X., Wu, H., & Zhang, M. (2019). A Hybrid Approach for Online Fraud Detection Based on Deep Learning and Network Analysis. IEEE Transactions on Computational Social Systems, 6(6), 1266-1278.

[5] Wang, T., Chen, Y., Chen, S., & Hsu, C. (2019). FraudCall: Automatic Detection of Fraudulent Phone Calls Using Deep Neural Networks. IEEE Transactions on Information Forensics and Security, 14, 1783-1795.

[6] Zhou, L., Chen, H., & Zhu, X. (2018). Detecting Online Frauds Using Deep Learning-Based Models. ACM Transactions on Intelligent Systems and Technology, 10(4), 1-21.

[7] Li, Y., Chen, S., Li, Y., & Li, S. (2019). Detecting Fraudulent Transactions Based on Long Short-Term Memory Neural Network. IEEE Access, 7, 43775-43786.

[8] Wang, Z., Li, S., Zhang, C., & Lu, J. (2021). Fraud Detection in E-commerce using Attention-based Convolutional Neural Network. Neurocomputing, 448, 346-357.

[9] Wu, W., Zhang, Y., Chen, J., & Chen, J. (2019). A Novel Method of Financial Fraud Detection Based on Deep Learning. IEEE Access, 7, 109547-109557.

[10] Cui, Y., Zhu, T., Xu, Y., & Chen, W. (2020). An Improved LSTM Model for Financial Fraud Detection in Online Transactions. Journal of Ambient Intelligence and Humanized Computing, 11(5), 2019-2029.