

Risk assessment for commercial crime and legal protection based on modularity-optimized LM and SCA-optimized CNN

YunZhe Wang and FengLan Su

HeBei Sport University, Shijiazhuang, Hebei, China

ABSTRACT

With the Internet of Things (IoT) making significant strides in recent years, the challenges associated with data collection and analysis have emerged as a pressing concern in public security. When employed to tackle extensive criminal networks, the conventional deep learning model encounters issues such as heightened computational complexity, sluggish operational efficiency, and even system failures. Consequently, this research article introduces an intricately devised framework for detecting commercial offenses, employing a modularity-optimized Louvain-Method (LM) algorithm. Additionally, a convolutional neural networks (CNN)-based model is formulated to determine the feasibility of extending legal aid, wherein feature transformation is facilitated by utilizing TFIDF and Word2vec algorithms aligned with diverse legal text corpora. Furthermore, the hyper-parameter optimization is accomplished using the sine cosine algorithm (SCA), ultimately enabling the classification of relevant legal guidance. The experimental outcomes comprehensively affirm the exceptional training effectiveness of this model. The commercial crime identification model, grounded in modular optimization as proposed in this article, adeptly discerns criminal syndicates within the commercial trading network, achieving an accuracy rate exceeding 90%. This empowers the identification of such syndicates and bestows the judicial sphere with pertinent legal insights.

Submitted 3 August 2023
Accepted 16 October 2023
Published 16 November 2023

Corresponding author
FengLan Su, wyz@hepec.edu.cn

Academic editor
Bilal Alatas

Additional Information and
Declarations can be found on
page 14

DOI 10.7717/peerj-cs.1683

© Copyright
2023 Wang and Su

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Artificial Intelligence, Data Mining and Machine Learning, Text Mining, Neural Networks

Keywords Commercial crime, LM, CNN, SCA, Legal protection

INTRODUCTION

In recent years, with the continuous development of the economy and society, the situation of comprehensive social governance has become increasingly serious. Commercial crime is becoming increasingly prominent, and it is easy to induce major social risks. Leaders at all levels of the state attach great importance to this. General Secretary Xi Jinping made important instructions to crack down on various crimes, stressing the need to strengthen key rectification and strictly curb the spread trend of various crime problems. To improve the people's sense of security and life happiness index, the public security departments continue to promote the governance of various criminal problems and always maintain a high pressure on various criminal activities. Through the strong investigation and control

of public security units at all levels, remarkable results have been achieved in the fight against various crimes (Feng, 2020). To maximize data's leading and early warning role in comprehensive social governance, we should gather massive amounts of information, such as people flow, capital flow, information flow, and call flow, and conduct deep mining to form relevant forecasts (Stevenson & Slobogin, 2018). At the same time, public security organs are facing the challenge of explosive data growth, especially with the increasingly mature development of the Internet of things in recent years, leading to more serious problems of data collection, analysis and processing faced by public security. Faced with such problems and challenges, public security organs are often passive in solving public security data analysis and prediction. How to fully excavate valuable information from public security data and effectively predict criminal behavior is the key to improving the public security system's work efficiency and reducing the risk to public security.

Compared with human decision-making, the automatic decision-making algorithm has relatively objective, fair and efficient characteristics, so its application has gradually spread to all fields of social life (Deng et al., 2021; Goel et al., 2021; Cai et al., 2021). For example, a university in China uses the algorithm to identify students' economic status according to consumption records to help determine the payment of subsidies for poor students; Banks widely use algorithms to evaluate customers' credit to decide whether to make loans; The US Department of Education uses algorithms to determine whether the teacher employment contract is renewed; In the courts of the United States, some of the crime risk assessment algorithms are used by judges. The automatic decision-making of algorithms directly affects the basic rights of human beings in social security, medical care, supervision of public officials and judicial system.

The traditional algorithm faces a significant dilemma when it comes to regulating paths primarily because it tends to overlook the contextual aspects of the algorithm itself. This oversight is a critical issue, as the behavior and performance of an algorithm can vary significantly depending on the specific context in which it operates. Various factors, such as the subject matter it is applied to, the nature of the objects it interacts with, and the unique problems it seeks to address, all play a pivotal role in shaping the algorithm's characteristics and outcomes.

The challenges and complexities have become increasingly pronounced in our ever-evolving landscape of comprehensive social governance. Society now grapples with multifaceted issues spanning various sectors, from public policy and law enforcement to economic and commercial activities. These challenges are further compounded by the growing prominence of commercial crimes, which have become more sophisticated and widespread.

In light of these pressing concerns, it becomes evident that a one-size-fits-all approach to algorithm regulation is no longer sufficient. Regulatory frameworks must adapt to acknowledge the nuanced interplay between algorithms, subjects, objects, and problems to effectively address the intricate and dynamic issues of our modern world. Only by doing so can we hope to harness the potential benefits of algorithms while mitigating their potential risks and ensuring that they serve the best interests of society as a whole.

This article formulates a legal decision-making model leveraging Convolutional Neural Networks (CNN). The model employs the Tf-idf and Word2Vec algorithms to transform features from various legal text corpora and fine-tune hyperparameters using the SCA to classify pertinent legal guidance, reducing the risk to public security.

RELATED WORKS

Algorithm prediction relies on the existing feature data and feature information. It establishes a data model for these data through a classification or clustering algorithm and calculates and predicts according to the model. *Kshatri et al. (2021)* proposed that association rules can be used to find potential association rules from criminal behavior information or criminal records to analyze and predict crimes and also to analyze the relationship between different types of cases and the characteristics of cases, as well as the relationship between different types and criminal methods. Although the crime prediction is successfully realized, there are still some defects. For example, the algorithm based on association rules is FP-growth, which will cost a lot of memory and is not widely available. In addition, the K-Means algorithm is sensitive to the initial centroid, and improper selection of the initial centroid will easily lead to the phenomenon that the convergence centroid process is very long and the execution time of the algorithm is too long. *Safat, Asghar & Gillani (2021)* advocated the utilization of Support Vector Machines (SVM) for crime prediction. They posited the viewpoint that SVM tackles the solution process as a convex quadratic programming problem, thereby capable of providing global solutions. This makes SVM a viable choice for crime prediction.

However, the voluminous data required for crime prediction presents a substantial challenge. The SVM algorithm, when confronted with a sizable training dataset, demands substantial memory and extends the algorithm's execution time considerably. In addition, the classical support vector machine algorithm only furnishes two classification methods, rendering it less adept at tackling multi-classification problems. Neil and his colleagues posit that the current escalation in crime rates necessitates a departure from traditional, sluggish crime resolution techniques. Consequently, they advocate the deployment of machine learning and computer vision algorithms to aid law enforcement in forecasting crimes, alleviating the police force's burden and contributing to crime prevention (*Neil, Nandish & Manan, 2021*). It is worth noting, however, that while they suggest the utility of machine learning and computer vision algorithms for crime prevention, they have yet to implement this proposal.

On the other hand, *Shen et al. (2019)* proposed utilizing the naive Bayesian algorithm to assess and compute the conditional probabilities of each feature attribute amenable to classification. The data set is then categorized based on these attribute-specific conditional probabilities, thereby achieving the classification and crime prediction objective. Nevertheless, it is imperative to recognize that the naive Bayes algorithm assumes attribute independence, a presumption that does not always hold in practical application and real-life scenarios. Furthermore, Lan did not undertake optimizations to enhance the naive Bayes algorithm. Suppose a particular attribute value remains absent in the training

set. In that case, its prior probability value defaults to zero, which can cause the naive Bayes calculation formula to yield zero, impacting the final classification outcome and diminishing the Bayesian algorithm's accuracy. Lastly, [Pu et al. \(2023\)](#) advocated using a spatiotemporal coarse-grained algorithm for crime prediction. Her approach entails extracting diverse time components through time series decomposition, sequential input of these components into the most suitable algorithms, and, ultimately, their effective integration.

To achieve high-performance training data, the GPU used is relatively high-end and expensive, and it also needs to be equipped with better CPU, SSD and RAM. It can be seen that deep learning requires very high machine configuration. However, the Hadoop cluster runs on some cheap commercial hardware.

Hence, executing deep learning algorithms on Hadoop clusters, particularly those with modest configurations, exerts significant computational stress on the hardware. Conversely, machine learning operates proficiently with a standard CPU for effective data training, demonstrating no particular hardware demands. Within this study, we employ a scanning algorithm to scrutinize crime patterns and assess their characteristics. Nonetheless, establishing the parameters for the DBSCAN algorithm poses a challenging task. Inadequate parameter selection can affect clustering outcomes, especially in scenarios with numerous data dimensions.

Furthermore, when confronted with high data dimensionality, DBSCAN exhibits suboptimal performance. Pan has proposed using the KNN algorithm to construct a predictive model for crime classification ([Zhongying, 2019](#)). However, the KNN algorithm entails extensive computations and exhibits limited fault tolerance. In cases where sample data for a particular class is scarce, it can lead to reduced classification accuracy for those specific classes. Kounadi conducted a comprehensive analysis of 32 papers from 2000 to 2018, culminating in synthesizing and appraising cutting-edge spatial crime prediction technologies ([Kounadi, 2020](#)). Gaurav introduced a spatiotemporal crime prediction technology rooted in machine learning and two-dimensional hotspot analysis. They assert that each crime category correlates with time, weather, location, and demographic parameters ([Hajela, Chawla & Rasool, 2020](#)).

COMMERCIAL CRIME IDENTIFICATION MODEL BASED ON MODULARITY-OPTIMIZED LM

Commercial crime cannot be separated from money transactions, so the capital transaction relationship between accounts in a commercial network can be effectively characterized by calculating transaction weight. Suppose the calculated capital transaction weights are applied to the Louvain-Method (LM) algorithm. The suspected subnets with the same fund transaction characteristics and close fund transactions can be effectively divided in that case. LM is a powerful community detection algorithm used in network analysis to identify clusters of nodes with strong internal connections. This algorithm maximizes modularity, a measure of network partition quality. LM starts with an initial partition and then iteratively optimizes modularity by moving nodes between communities. It aggregates the network

into a new representation, repeating the process until convergence. Finally, it assigns nodes to communities in the aggregated network. LM's efficiency and scalability make it ideal for uncovering hidden structures in complex networks across various domains, from social networks to biology and transportation planning.

The LM algorithm can be considered a clustering algorithm to partition a large-scale network into distinct communities. Nodes within a community exhibit similar characteristics and maintain strong connections, whereas nodes bridging different communities exhibit distinct characteristics and sparse connections (Tian et al., 2022). Due to its good classification accuracy and excellent adaptability to large-scale networks, the LM algorithm is slowly applied to the scene of commercial network crime identification based on capital transactions (Bahulkar et al., 2018). However, the original LM algorithm is serial. When applied in large-scale financial networks, it is easy to lead to slow operation efficiency and even system crashes. To solve this problem, we need to optimize the modularity of the original LM algorithm.

Modularity optimization process

As an important index to measure the reasonable degree of community division, the main idea is in a random business network, the probability of capital transaction relationship between two accounts (that is, there are edges between two nodes) is equal. It has the same number of nodes as the actual network, but the edges in the random network are connected randomly. By comparing the proportion of the number of edges of each community node with the total number of edges in the actual network, subtracting the expected difference of the total number of edges in the community of the corresponding random network. If the difference is larger, the randomness is smaller, the relationship between nodes in the network is closer, and the situation of community division is better (Seifikar, Farzi & Barati, 2020). Therefore, we can define modularity as follows:

$$Q = \frac{1}{2m} \sum_{i,j} \left| A_{ij} - \frac{k_i k_j}{2m} \right| \delta(c_i, c_j) \quad (1)$$

where m represents the number of edges in the network, A_{ij} represents the weight of edges between node i and node j , indicating the closeness between node i and node j . k_i is the sum of the values of all edges connected to node i , and $\frac{k_i k_j}{2m}$ represents the closeness of i and j when the network is random. c_i is the cluster number of node i . If c_j is the same, then $\delta(c_i, c_j) = 1$, indicating that node i and node j are in the same cluster. Otherwise, $\delta(c_i, c_j) = 0$, indicating that node i and node j are not in the same cluster. $\frac{1}{2m}$ is the introduced inertia factor, which is to control the value of modularity within the interval (0,1).

According to the idea of modularity, the definition of community modularity gain is given as follows:

$$\delta Q = \left[\frac{\sum in + k_{i,in}}{2m} - \left(\frac{\sum tot + k_i}{2m} \right)^2 \right] - \left[\frac{\sum in}{2m} - \left(\frac{\sum tot}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \quad (2)$$

The above formula can be simplified as

$$\delta Q' = k_{i,in} - \frac{\sum tot \times k_i}{m} \quad (3)$$

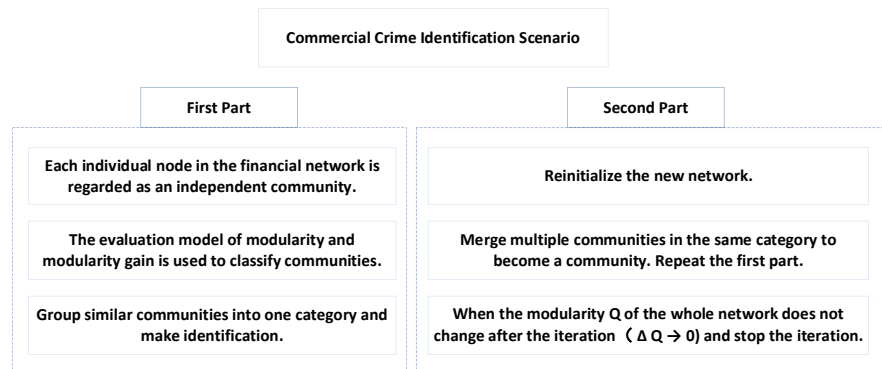


Figure 1 Commercial crime identification scene.

Full-size DOI: 10.7717/peerjcs.1683/fig-1

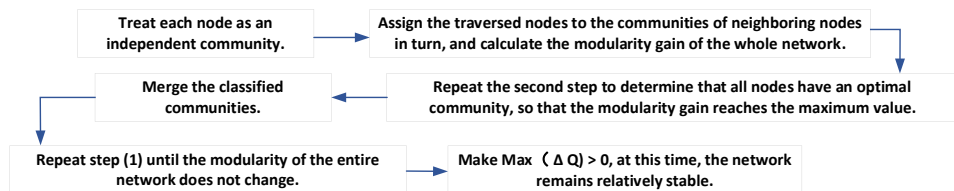


Figure 2 Algorithm steps.

Full-size DOI: 10.7717/peerjcs.1683/fig-2

whereas, $k_{i,in}$ signifies the summation of weights about cluster C linked to node i , and \sum_{tot} symbolizes the cumulative sum of total weights connected to cluster C . Additionally, k_i denotes the cumulative sum of total weights linked to node i . It is noteworthy that the simplified calculation yields a relative gain rather than an absolute gain. This transformation significantly mitigates programming complexity.

As shown in Fig. 1, the application of the LM algorithm based on modularity in the scene of commercial crime identification is divided into two parts.

In the first part, each node in the financial network is regarded as an independent community, and these communities are classified by the above evaluation model of modularity and modularity gain, and similar communities are classified into one group and identified.

In the second part, we reinitialize the new network, merge multiple communities into a community, and then repeat the first part to complete a round of iteration. When the modularity Q of the whole network no longer changes after iteration ($\delta Q \rightarrow 0$), the iteration stops.

Implementation process

The specific steps of the above algorithm are shown in Fig. 2.

- (1) Each node in the transaction network is regarded as an independent community.
- (2) After traversing each node of the transaction network, the nodes traversed are allocated to the community of adjacent nodes in turn, and the modularity gain of the

whole network δQ is calculated if $\max(\delta Q) > 0$, that is, there is an adjacent node of the current node. After merging the current node into a community, the module degree of the whole network increases, while if it does not exist $\max(\delta Q) > 0$. The current node does not join the community where any node is located, and the original state remains unchanged (Zhang et al., 2021).

(3) Repeat Step (2) to determine that all nodes in the network have an optimal community so that the network modularity gain reaches the maximum value.

(4) In this way, all the nodes belonging to the same community are merged into one node. The weights of the edges between the nodes in the community are transformed into the weights of the new nodes, and the weights of the edges between the communities are transformed into the weights of the edges between the new communities.

(5) Repeat Step (1) until the modularity of the whole network is no longer changed; that is, there is no merging of any two communities so that $\max(\delta Q) > 0$. At this time, the network remains relatively stable.

In traversing the intricate web of transaction nodes within the network, a systematic procedure is followed to determine the allocation of these nodes to their respective adjacent communities (Niu et al., 2023). This allocation step plays a pivotal role in assessing the community structure's robustness and optimizing it for enhanced modularity, a key measure of network organization. At each juncture of node traversal, the algorithm assesses the potential benefit of reallocating the current node to a different community. This evaluation hinges on the concept of "modularity gain", which quantifies the improvement in the network's modularity when a node is moved to a new community. In this context, modularity measures how densely connected the nodes are within a community compared to what would be expected by chance.

In the implementation of the community discovery algorithm, the most complex and time-consuming task is the community division in step (2) because there are a large number of nodes participating in the calculation, and the modularity gain of the whole network needs to be calculated when the current node joins each neighbor node. In principle, the community division of each node can be optimized with the idea of parallelization because the modularity change is only related to the community of the current node and the neighbor node of the current node and has nothing to do with other communities. When step (2) is executed, the order of traversing nodes has little influence on the final community clustering effect (Linhares et al., 2020).

DESIGN OF LEGAL ASSISTANT DECISION MODEL

Overall design

Artificial intelligence technology can meet the needs of many judicial aspects, such as intelligent conviction, classification of legal advisory issues, document recommendation, etc. The task of intelligent conviction is to use the computer to read the case description and automatically give the appropriate charges. The classification task of legal consultation is to predict which legal level the problem may belong to based on the description of the user problem. The task of document recommendation is to search for similar documents

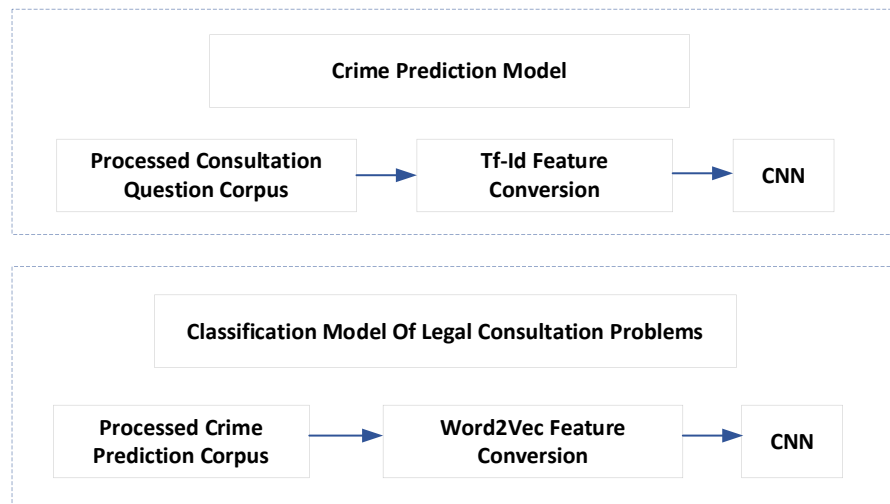


Figure 3 Legal assistant decision model.

Full-size DOI: [10.7717/peerjcs.1683/fig-3](https://doi.org/10.7717/peerjcs.1683/fig-3)

from massive data to assist judges and lawyers in judging cases. Given the identification of commercial crimes, this article proposes a corresponding legal assistant decision-making model, as shown in Fig. 3. The model mainly solves two problems: one is crime prediction, and the other is the classification of legal advisory issues.

CNN algorithm design

Input layer

In the input layer of the convolutional neural network, to realize the representation from single word feature to whole document feature, a more intuitive feature organization method is adopted. For a document, a combination of several words is formed after processing. After each word is transformed into a word vector, and then each word vector is arranged vertically, the text data can be converted into matrix data.

After the text is processed, suppose that a text contains m words, and a corresponding word vector represents each word; then, the text can be expressed as a vertical arrangement of m f -dimension vectors, and f is the dimension set by the word vector during the training process. Afterword segmentation and stop word filtering, the longest legal text in the training set contains 935 words. The word embedding length of each word is set to 300 dimensions, that is, $m = 935$, $k = 300$. Therefore, the corresponding input layer organization is a $935 * 300$ network structure. Text data often has no fixed length; some may be shorter. 935 is just the longest text length in all data sets. The text that cannot reach the length needs to be processed separately. In this article, the method of zero padding is used. For a text with n words ($n < 935$), the first n words are arranged vertically by word embedding, and the insufficient parts are filled with zeros, *i.e.*, $935-n$ zeros.

Convolution layer

The input of the convolution layer is a text matrix; each row of the matrix represents a word, and the matrix size is 935×300 (Lin et al., 2020). Convolution kernel extracts

features by sliding on the text matrix. When the width of the convolution kernel remains unchanged, the height of the convolution kernel represents the size of the convolution kernel. If the text matrix is defined as $A \in R^{s*d}$ and the size of the convolution kernel is h , then the weight vector of the convolution kernel is $w \in R^{h*d}$, then there are $h*d$ parameters to be calculated. Line i to line $i+h-1$ of the text matrix can be expressed as $A[i:i+h-1]$. When the convolution window acts on the text submatrix $A[i:i+h-1]$, the following formula can be obtained:

$$o_i = \omega \cdot A[i:i+h-1]. \quad (4)$$

Among them, $i = 1 \dots, s-h+1$, the symbol \cdot represents the dot product operation. If the convolution kernel operates on the entire sentence matrix, then the length of the output O is $s-h+1$. Then, the deviation term $b \in R$ and the activation function f are used to work on each O_i , then C_i can be obtained. The semantic feature map $c \in \mathbb{R}^{s-h+1}$ can be obtained by all C_i . The calculation formula of C_i is as follows:

$$c_i = f(O_i + b). \quad (5)$$

In order to improve the training speed, The Relu function is used as the activation function in this article:

$$f(x) = \max(0, x). \quad (6)$$

Full connection layer

The full connection layer is usually at the tail of the convolutional neural network, and all neurons have a weight connection. According to the K features with the greatest influence, the feature data is divided into a one-dimensional vector and sent to the full connection layer with the Softmax classifier (Ma, He & Peng, 2023) so that all the local features extracted can be considered comprehensively to complete the classification task of the model.

Super parameter optimization

In this article, SCA is proposed to optimize the parameters, that is, to find the most suitable set of parameters in the range of optional parameters so that the model's accuracy is the highest. In the operation of SCA, it is necessary to determine the fitness function. In this experiment, the accuracy of the model on the test set is selected as the fitness function, as shown in Eq. (7)

$$\text{acc}(f; D) = \frac{1}{m} \sum_{i=1}^m \Pi(f(x_i) = y_i) \quad (7)$$

To balance the exploration and development phases and the range of sine and cosine in the equation. The following equation is used for adaptive changes:

$$r_1 = a - t \frac{a}{T} \quad (8)$$

where t is the current iteration, T is the maximum number of iterations, and a is a constant. The specific optimization steps are as follows:

(1) The processed and transformed data are divided into a test set and a training set, which are used for model training and testing. The hyper-parameter range of the prediction model was determined.

(2) Initialize SCA and set the parameters of SCA.

(3) Execute the first iteration of SCA, where the fitness value is calculated according to the fitness function, and the solution with the highest fitness in the current initial solution set is selected and saved.

(4) When t is 2, the parameter $r1$ is calculated and the position of the candidate solution is updated.

(5) After the new candidate solutions are obtained, the fitness of each candidate solution is calculated. The optimal fitness in this iteration is selected for comparison with the candidate solutions saved in the last round. If the current candidate solution is superior to the previous one, the current candidate solution is selected. Otherwise, the solution from the previous iteration is selected.

(6) Judge whether the termination conditions are met, and if so, proceed to the next step; Otherwise, make $t = t + 1$ and repeat Step (5) and Step (6).

(7) When the optimal fitness solution and the location of the solution are obtained, the hyper-parameters are output, and the parameters are input into the model for training to get the final model.

EXPERIMENT AND ANALYSIS

Experimental data

In this research endeavor, the experimentation process relied on authentic bank production data, underscoring the validity and practicality of the findings. The experiment was conducted meticulously and precisely, leveraging a genuine data set sourced from a banking institution. This data set, accessible at the following link (<https://doi.org/10.7272/Q61G0JHJ>), comprises a comprehensive and extensive collection of financial records.

To provide a snapshot of the data's magnitude, it encompasses a staggering 78,760 distinct accounts, embodying a diverse cross-section of financial activity. Within the confines of a single month, an astonishing 568,978 transaction records were meticulously logged and analyzed. Each transaction record is replete with a rich tapestry of information, containing 70 distinct fields, each offering valuable insights into the intricate world of financial transactions.

It is imperative to note that the bank, in a conscientious effort to safeguard customer privacy and protect sensitive financial information, has taken stringent measures to ensure data security. Consequently, the real production data set has been transformed into a secure and confidential format. This transformation includes desensitization, which involves removing or obfuscating personally identifiable information, and encryption, which fortifies the data against unauthorized access. These measures collectively ensure that the integrity and confidentiality of the data are upheld, complying with the highest standards of data protection and privacy regulations.

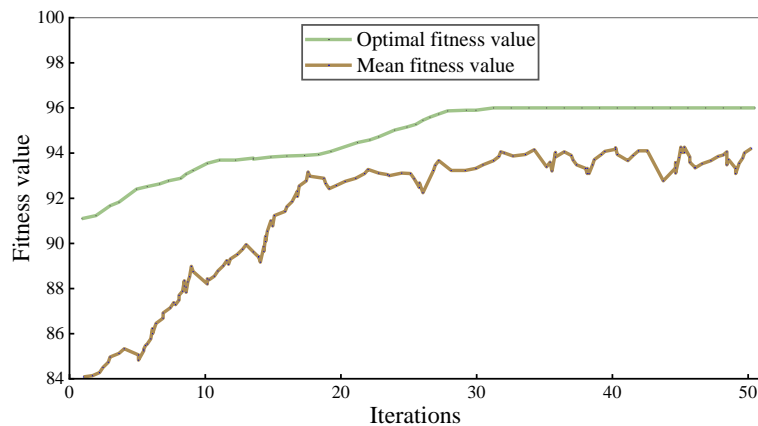


Figure 4 Training results of commercial crime identification model.

Full-size  DOI: [10.7717/peerjcs.1683/fig-4](https://doi.org/10.7717/peerjcs.1683/fig-4)

By employing this authentic and rigorously prepared data set, the study aimed to provide a robust foundation for its analyses, insights, and conclusions. The reliance on actual production data enhances the relevance and applicability of the research findings, ultimately contributing to a deeper understanding of the intricate dynamics surrounding commercial network criminal gangs and their activities within the realm of capital transactions.

Model training effect

All the corpus was divided into 7:3 parts, 70% of which were used to train the model and 30% to verify the performance of the model. The fitness curves in the model training process are shown in Figs. 4 and 5, and accuracy is used to evaluate them.

As can be seen from the figure, the accuracy rate of the legal assistant decision model is maintained at 94%–95%, and that of the commercial crime identification model is maintained at 92%–93%. As the number of iterations increases, the fitness value will also increase. When the number of iterations reaches 25–30, the fitness value tends to be stable. After 50 iterations, the algorithm stops. Finally, the highest accuracy rate of the legal aid decision model is 95.93% and 93.66% of the commercial crime identification model.

Comparison of different models

The commonly used algorithm models are selected for comparison, and the comparison results are shown in Fig. 6.

The research outcomes unequivocally demonstrate the superiority of deep learning methodologies over traditional machine learning algorithms in commercial crime identification. This empirical evidence underscores the potential of advanced neural network-based techniques in tackling complex real-world challenges. Specifically, the deep learning approach exhibited remarkable performance gains compared to its traditional counterparts.

Notably, within the realm of deep learning, a significant enhancement in accuracy was observed with the utilization of an optimized CNN as opposed to a CNN with manually configured hyperparameters. This optimization process underscores the importance of

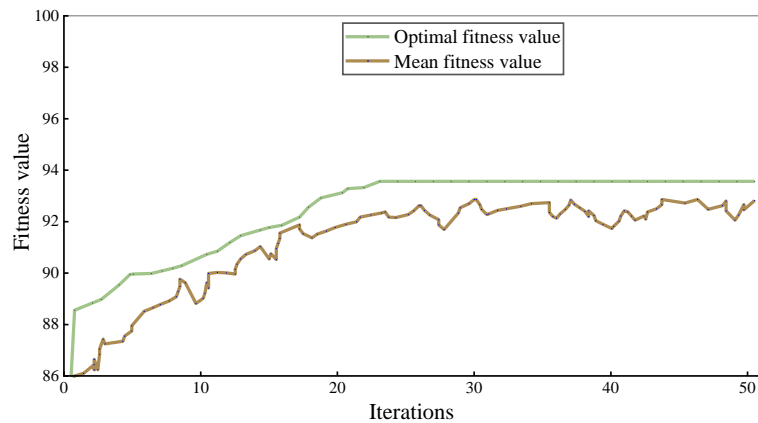


Figure 5 Training results of legal assistant decision model.

Full-size DOI: 10.7717/peerjcs.1683/fig-5

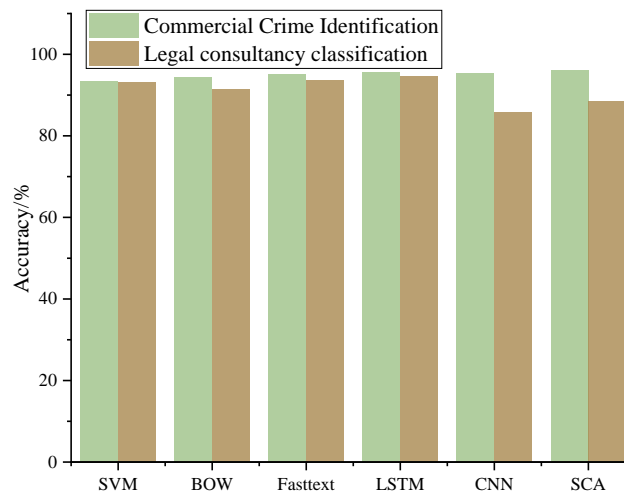


Figure 6 Model comparison results.

Full-size DOI: 10.7717/peerjcs.1683/fig-6

fine-tuning neural network architectures to suit the specific task. By carefully tuning the hyperparameters, the algorithm achieved higher accuracy, further underlining the capacity for deep learning to adapt and excel in intricate problem domains.

Conversely, a stark contrast emerged when comparing deep learning, represented by CNN, with the traditional random forest algorithm. While the deep learning approach exhibited substantial advantages in terms of accuracy, it is essential to acknowledge that deep learning algorithms may not always yield ideal results in all scenarios. One key reason for this disparity is the nature of the legal consultation questions. These questions tend to be concise and lack substantial linguistic connections. Consequently, the deep learning algorithms face the challenge of discerning and classifying these short sentences, often relying on just a handful of crucial keywords.

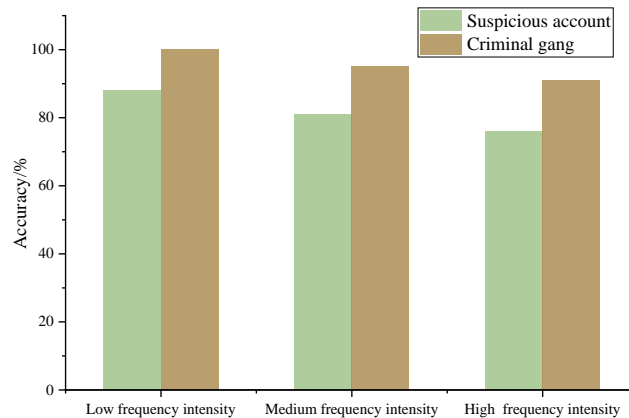


Figure 7 Commercial crime identification effect under different sample frequency intensity.

Full-size  DOI: [10.7717/peerjcs.1683/fig-7](https://doi.org/10.7717/peerjcs.1683/fig-7)

In summary, these findings emphasize the evolving landscape of machine learning, where deep learning techniques are proving their mettle in specialized tasks like commercial crime identification. However, the data's intricacies and the problem's specific characteristics must be considered when selecting the appropriate algorithm. The unique challenge posed by succinct legal consultation questions highlights the importance of continuous refinement and adaptation in achieving optimal algorithmic performance.

Commercial crime identification effect

According to the suspicious report and the later investigation certificate, the commercial criminal account provided by the bank is taken as the real criminal account. The suspicious account in the scope of the real criminal account obtained by the algorithm in this article is regarded as the experimental suspicious account, and the account not within the scope of the real criminal account is regarded as the experimental false alarm account. In the commercial transaction network, the Adjusted Rand Index (ARI) is used to measure the accuracy of the results. The experimental results are shown in [Fig. 7](#).

[Figure 7](#) shows that the commercial crime identification model based on modular optimization proposed in this article can identify the criminal gangs in the commercial trading network, and the accuracy rate is more than 90%. When the fund transaction between accounts is characterized by low amount frequency, the model performs better because the intimacy weight between accounts becomes smaller and the number of suspicious accounts decreases, which indicates that the method meets the expectation and the experimental results are reasonable.

CONCLUSION

This study develops a legal assistant decision-making model based on CNN and a commercial crime identification model based on modular optimization, realizing the precise identification of criminal gangs in the commercial transaction network and through legal informatization. The test data in this research use genuine bank production data sets to

complete the experiment. The findings demonstrate that the model has excellent accuracy in identifying commercial crime and classifying legal consultation problems, which can assist managers in analyzing commercial security data and enhancing the productivity of the legal profession. This study has successfully laid the foundation for advanced applications in legal assistance and commercial crime identification. Moving forward, several promising avenues for further research and development can build upon these achievements.

Firstly, the legal assistant decision-making model based on CNNs shows great potential for broader legal informatization initiatives. Future investigations could delve into expanding the scope of this model to encompass a wider array of legal tasks and domains. This could involve fine-tuning the model to handle more complex legal scenarios, such as contract analysis, legal document classification, or predictive legal decision-making.

Similarly, the commercial crime identification model based on modular optimization presents an exciting opportunity for refinement and expansion. Further research could explore the application of this model in diverse contexts beyond commercial crime, such as fraud detection in financial systems, network security, or even cybersecurity. Additionally, the optimization techniques used in this model could be adapted to improve the efficiency and effectiveness of other machine learning algorithms, potentially leading to advancements in various domains.

Moreover, using genuine bank production data sets in this research opens avenues for developing more comprehensive and nuanced datasets tailored to legal and financial domains. Creating such datasets with carefully curated and labeled examples could facilitate training even more accurate and specialized models.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no funding for this work.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- YunZhe Wang conceived and designed the experiments, performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- FengLan Su conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The code is available in the [Supplementary File](#).

The dataset is available at Dryad: - Rotering, Thomas; Apollonio, Dorie (2022). Cannabis industry businesses, organizations, and individuals that lobby in Colorado [Dataset]. Dryad. <https://doi.org/10.7272/Q61G0JHJ>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.1683#supplemental-information>.

REFERENCES

- Bahulkar A, Szymanski BK, Baycik NO, Sharkey TC. 2018.** Community detection with edge augmentation in criminal networks. In: *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. Piscataway: IEEE, 1168–1175.
- Cai X, Zhao H, Shang S, Zhou Y, Deng W, Chen H, Deng W. 2021.** An improved quantum-inspired cooperative co-evolution algorithm with multi-strategy and its application. *Expert Systems with Applications* **171**:114629 DOI [10.1016/j.eswa.2021.114629](https://doi.org/10.1016/j.eswa.2021.114629).
- Deng W, Xu J, Song Y, Zhao H. 2021.** Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem. *Applied Soft Computing* **100**:106724 DOI [10.1016/j.asoc.2020.106724](https://doi.org/10.1016/j.asoc.2020.106724).
- Feng H. 2020.** Investigation dilemma and Countermeasures of telecom network fraud crime. *Forest Public Security* **2020(03)**:12–14.
- Goel S, Shroff R, Skeem J, Christopher S. 2021.** The accuracy, equity, and jurisprudence of criminal risk assessment. In: *Research handbook on big data law*. Edward Elgar Publishing, 9–28.
- Hajela G, Chawla M, Rasool A. 2020.** A clustering based hotspot identification approach for crime prediction. *Procedia Computer Science* **167**:1462–1470 DOI [10.1016/j.procs.2020.03.357](https://doi.org/10.1016/j.procs.2020.03.357).
- Kounadi O. 2020.** A systematic review on spatial crime forecasting. *Crime Science* **9(1)**:1–22 DOI [10.1186/s40163-019-0110-3](https://doi.org/10.1186/s40163-019-0110-3).
- Kshatri SS, Singh D, Narain B, Bhatia S, Quasim MT Sinha, GR. 2021.** An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: an ensemble approach. *IEEE Access* **9**:67488–67500 DOI [10.1109/ACCESS.2021.3075140](https://doi.org/10.1109/ACCESS.2021.3075140).
- Lin X, Wang H, Chen Z, Zhang H, Li Y. 2020.** Measurement of the flow rate of oil and water using microwave and Venturi sensors with end-to-end dual convolutional neural network. *Measurement: Sensors* **10**:100018 DOI [10.1016/j.measen.2020.100018](https://doi.org/10.1016/j.measen.2020.100018).
- Linhares CDG, Ponciano JR, Pereira FSF, Rocha LEC, Paiva JGS, Travencolo BAN. 2020.** Visual analysis for evaluation of community detection algorithms. *Multimedia Tools and Applications* **79(25)**:17645–17667 DOI [10.1007/s11042-020-08700-4](https://doi.org/10.1007/s11042-020-08700-4).
- Ma S, He K, Peng X. 2023.** Online coal identification based on one-dimensional convolution and its industrial applications. *Applied Sciences* **13(17)**:9867 DOI [10.3390/app13179867](https://doi.org/10.3390/app13179867).

- Neil S, Nandish B, Manan S. 2021.** Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. *Visual Computing for Industry, Biomedicine, and Art* **4**(1):1–14.
- Niu Y, Kong D, Liu L, Wen R, Xiao J. 2023.** Overlapping community detection with adaptive density peaks clustering and iterative partition strategy. *Expert Systems with Applications* **213**:119213 DOI [10.1016/j.eswa.2022.119213](https://doi.org/10.1016/j.eswa.2022.119213).
- Pu X, Zhu J, Wu Y, Leng C, Bo Z, Wang H. 2023.** Dynamic adaptive spatio-temporal graph network for COVID-19 forecasting. *CAAI Transactions on Intelligence Technology* Epub ahead of print 2023 24 June DOI [10.1049/cit2.12238](https://doi.org/10.1049/cit2.12238).
- Safat W, Asghar S, Gillani SA. 2021.** Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques. *IEEE Access* **9**:70080–70094 DOI [10.1109/ACCESS.2021.3078117](https://doi.org/10.1109/ACCESS.2021.3078117).
- Seifikar M, Farzi S, Barati M. 2020.** C-blondel: an efficient louvain-based dynamic community detection algorithm. *IEEE Transactions on Computational Social Systems* **7**(2):308–318 DOI [10.1109/TCSS.2020.2964197](https://doi.org/10.1109/TCSS.2020.2964197).
- Shen Y, Li Y, Zheng HT, Tang B, Yang M. 2019.** Enhancing ontology-driven diagnostic reasoning with a symptom-dependency-aware Naïve Bayes classifier. *BMC Bioinformatics* **20**(1):1–14 DOI [10.1186/s12859-018-2565-8](https://doi.org/10.1186/s12859-018-2565-8).
- Stevenson MT, Slobogin C. 2018.** Algorithmic risk assessments and the double-edged sword of youth. *Behavioral Sciences & the Law* **36**(5):638–656 DOI [10.1002/bsl.2384](https://doi.org/10.1002/bsl.2384).
- Tian Y, Feng Y, Zhang X, Sun C. 2022.** A fast clustering based evolutionary algorithm for super-large-scale sparse multi-objective optimization. *IEEE/CAA Journal of Automatica Sinica* **10**(4):1048–1063 DOI [10.1109/JAS.2022.105437](https://doi.org/10.1109/JAS.2022.105437).
- Zhang J, Fei J, Song X, Feng J. 2021.** An improved Louvain algorithm for community detection. *Mathematical Problems in Engineering* **2021**:1485592.
- Zhongying P. 2019.** *Design and implementation of crime warning system based on machine learning algorithm*. Sichuan, China: University of Electronic Science and Technology (in Chinese).