International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 6365-6372

https://doi.org/10.70517/ijhsa464540

Enhancing Judicial Credibility through Legal Judgment Prediction: Optimized Particle Swarm Algorithms and Big Data Analysis

Ying Chieh Lin[1,*] and Shaojun Liu[1]
[1] School of Civil, Commercial, and Economic Law, China University of Political Science and Law, Beijing, 100000, China
Corresponding authors: (e-mail: inja@asdgac.org).

Abstract The public's confidence in legal institutions is frequently damaged by the unpredictable nature of court rulings, which also creates a divide between the public and judicial elites. Improving the predictability of rulings is essential for promoting overlapping consensus, bolstering judicial authority, and increasing public confidence in legal institutions. In order to increase the precision and interpretability of legal judgment forecasts, this study investigates the use of an optimized particle swarm algorithm. We suggest a hybrid model that combines particle swarm optimization with semantic and decision-element analysis methods using the CAIL2018 dataset, a sizable collection of criminal case records. According to experimental data, the revised algorithm improves accuracy, resilience, and convergence speed by 5% across a range of case circumstances. The suggested framework addresses issues of interpretability and decision-making support while greatly increasing the efficiency of legal judgment prediction by utilizing cutting-edge data mining techniques like matrix-based distributed representations and scaled dot-product attention mechanisms.

Index Terms Legal judgment prediction, optimized particle swarm algorithm, judicial credibility, big data analysis, semantic representation, interpretability, artificial intelligence

I. Introduction

In recent years, the complexity of legal cases and the rapid growth of data have presented unprecedented challenges to judicial systems worldwide [1]. The evolution of crime patterns and the advancement of criminal investigation techniques have contributed to the accumulation of vast and intricate datasets. Judges, lawyers, and other legal professionals face increasing difficulties in managing and interpreting this information, often leading to inconsistencies in judgments and a diminished public trust in judicial processes [2], [3]. The unpredictability of judgments not only undermines judicial credibility but also disrupts social cohesion and weakens the authority of legal systems.

Legal judgment prediction has emerged as a critical technology to address these challenges [4]. By leveraging artificial intelligence (AI) and data mining, this field aims to predict judicial outcomes based on factual case descriptions. Such predictions provide valuable references for legal professionals, enabling them to make more consistent and informed decisions [5]. Moreover, predictive technologies offer a pathway to reduce discrepancies caused by individual biases among judges, thereby promoting the principle of "equal judgment for similar cases."

Current judicial judgment prediction systems have a lot of drawbacks despite their potential. Deep neural networks and other high-accuracy models frequently function as "black boxes," providing little explanation for the logic underlying their predictions [6], [7]. The adoption and real-world implementation of these models in legal systems that demand open decision-making procedures are hampered by their lack of interpretability. Simpler models, like as decision trees and linear regression, on the other hand, are easier to understand but frequently compromise accuracy, which reduces their usefulness in complex situations [8]. For the field to advance, this gap between interpretability and accuracy must be closed.

The era of big data and advancements in computational power provide new opportunities to enhance the predictability of judicial judgments. The availability of extensive legal datasets, such as the China Judgment Document Network and the CAIL2018 dataset, allows researchers to explore novel methodologies for case analysis and prediction [9], [10]. These datasets contain millions of annotated legal cases, offering a rich resource for developing machine learning models capable of identifying patterns and drawing inferences across diverse scenarios. The integration of big data analytics with optimized algorithms promises to revolutionize the field of legal prediction by enabling more accurate, transparent, and efficient decision-making processes [11].

Particle swarm optimization (PSO) algorithms, inspired by the collective behavior of biological systems, have shown significant potential in addressing complex optimization problems [12]. Recent advancements in PSO techniques, including hybrid models and enhanced parameter tuning, have demonstrated improved convergence speed, accuracy, and adaptability. Applying these optimized algorithms to legal judgment prediction offers a promising avenue for overcoming existing limitations in the field [13]. By balancing global exploration and local exploitation, PSO-based models can efficiently analyze large datasets, identify relevant patterns, and provide interpretable predictions.

This study investigates the application of an optimized particle swarm algorithm for legal judgment prediction. By integrating semantic analysis, decision-element extraction, and hybrid modeling techniques, the proposed framework aims to achieve high accuracy while maintaining interpretability. Using the CAIL2018 dataset as a benchmark, the study evaluates the performance of the optimized algorithm in predicting charges, applicable legal provisions, and sentencing outcomes. Experimental results demonstrate that the optimized PSO model not only improves prediction accuracy but also enhances the transparency and reliability of judicial decision-making processes.

In addition to technical contributions, this research addresses the broader implications of legal judgment predictability. Enhancing the transparency and consistency of judicial decisions fosters public trust and bridges the gap between judicial authorities and society. Predictive technologies help create a more unified and trustworthy legal system by offering a solid basis for behavioral expectations and value direction. The study's conclusions highlight how AI and big data have the power to completely change the legal system and open the door to a time where the rule of law and technology coexist together to preserve justice and equity.

II. Goal of the study

The semantic links between words and contexts are described by matrix-based distributed representations, which create "word-context" matrices. A latent semantic analysis (LSA) model was proposed in the literature [14]. It reduces feature dimensions and noise by projecting words into a low-dimensional semantic space using singular value decomposition. Although the polysemy issue is partially resolved by the left and right singular matrices in LSA, which capture the correlation between words and topics and topics and documents, respectively, it is challenging to decipher the meaning of negative values and cannot distinguish between polysemous words. In this context, literature [15] suggests Probabilistic Latent Semantic Analysis (PLSA), which is able to recognize polysemous words by treating words and texts as random variables and adding a latent topic layer. However, the PLSA model's parameters rise linearly with training samples and are vulnerable to overfitting because there is no assumption about the topic prior distribution. By adding the Delicacy distribution, the latent Delicacy Distribution (LDA) model is further suggested as a solution to this issue, reducing the overfitting of PLSA and optimizing the parameter distribution of subjects and words. Nevertheless, the matrix decomposition method's vector space structure remains suboptimal, and its performance in the word linear analogy test is restricted.

II. A. Theoretical Foundation

The China Academy of Legal Studies Judicial Artificial Intelligence Cup (CAIL2018) dataset, which includes over 2.6 million criminal cases released by the Chinese Supreme People's Court, was used in this study. The CAIL2018 dataset is larger, contains several times as much data, and offers richer and more extensive judgment annotations than existing legal judgment prediction datasets. The judgments cover the relevant legal provisions, charges, and judgment conclusions, and each case provides factual descriptions and matching judgments. CAIL2018 offers substantial significance for legal intelligence system research in addition to being a resource for professionals looking to increase their productivity (see Table 1).

Table 1: CAIL2018 dataset examples

<table>
<tr><td>Factual explanation</td><td colspan="2">On the morning of November 5, 2015, the victim Sun and the defendant Hu got into a fight over unimportant tasks at Jiaxing duoling Jinniu Clothing Co., Ltd.'s workshop in Zhapu Town, Pinghu City. Later, the defendant Hu used a wooden cushion to damage the victim Sun's left abdomen. The sun's left abdomen injury has progressed to the second stage of serious harm, according to the Pinghu Public Security Judicial Expertise Center.</td></tr>
<tr><td rowspan="4">The outcome of the judgment</td><td>Defendant</td><td>Hu Mou</td></tr>
<tr><td>charge</td><td>Intentional harm</td></tr>
<tr><td>Provisions that are pertinent</td><td>Criminal law Article 234</td></tr>
<tr><td>The length of incarceration</td><td>12 months</td></tr>
</table>

II. B.Building Models

Case similarity judgment is fundamentally a type of analogical reasoning, which is the process of determining the significance of previous and pending cases to be expected. This is done by assuming that two things are similar in certain elements of the reasoning and that they may also be similar in other aspects. Such logic leads to a very probabilistic conclusion: objects A and B contain qualities a, b, and c simultaneously, but object B does not necessarily have attribute d. Even if "different judgments for the same case" is a typical occurrence, it is evident that a system of case guidance can improve judgment predictability.

The computational and operational ease of supervised keyword extraction techniques makes them popular in fields like search and document classification. It is mostly used to assess a word's significance inside a corpus or dataset and quantify its impact in order to determine the word's weight in information retrieval. The degree of information a phrase in a query provides is reflected in its importance. The technique makes the assumption that a word's information value may be efficiently distinguished by how frequently it appears in a text. The fundamental tenet is that a word's significance is directly correlated with how frequently it appears in the text.

The task summary's frequency shows how frequently terms appear in the text:

$$TF_{w,D_i} = \frac{C(w)}{|D_i|} \tag{1}$$

The converse is Inverse Document Frequency (IDF). The word's IDF value is inversely correlated with the sentences in which it appears. The following are the calculation formulas (2) and (3):

$$IDF_w = \ln \frac{n}{\sum_{i=1}^{n} I(w, D_i)} \tag{2}$$

$$I(w, D_i) = \begin{cases} 1 & w \in D_i \\ 0 & w \overline{\in} D_i \end{cases} \tag{3}$$

Formula (2) is useless when the word "w" is absent from any document in the dataset, hence the IDF value is calculated as Formula (4) illustrates the slight modification in the calculation:

$$IDF_w = \ln \frac{n}{1 + \sum_{i=1}^{n} I(w, D_i)} \tag{4}$$

Therefore, formula (5) can be used to determine the text's TF-IDF value.

$$\mathrm{TFIDF}_{w,D_i} = TF_{w,D_i} \times IDF_w \tag{5}$$

III. Methods

The fundamental concept of optimization is to model the word graph by seeing the text as a collection of various units (words, sentences, etc.) and then classifying these units based on their criticality. The voting mechanism that underpins the sorting rules determines a unit's criticality based on the sum of the criticalities of the other units that point to it. Keyword extraction can be done immediately in the text, which is efficient and concise, as it eliminates the need for corpus learning and training. The logical flow is obvious and intuitive, as seen in Fig. 1:

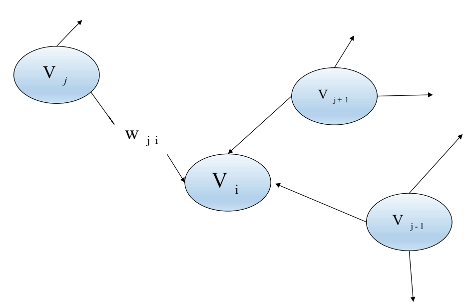

Figure 1: The algorithm's underlying logic diagram

III. A. Classification of Texts

A crucial stage in transforming text into a computer-readable format, text rendering determines the content's quality. The issue of content search is covered in Chapter 2. Text categorization performance can be weakened by complicated multi-level repetition, which affects the majority of text categorization functions (i.e., text groups). Selecting more representative and varied text representations is therefore especially crucial. By removing separate functions, this method not only partially resolves the issue but also lessens the categorization challenge. As illustrated in Figure 2, the selection process is separated into three categories based on the type of selection: filtering, packaging, and synthesising.

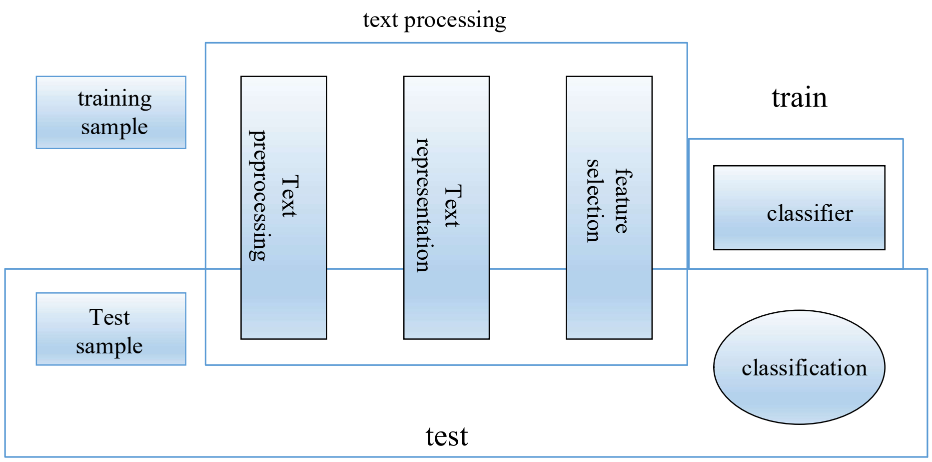

Figure 2: The fundamental text classification logic diagram

III. B. Decision-element extraction

This chapter adds "0" labels to non-elementary sentences, divides the dataset's factual descriptive paragraphs into numerous factual descriptive sentences, and combines many labels into a single label to make model processing easier. Table 2 displays the processed experimental data.

Table 2: The quantity of samples in each kind of dataset

Data	Set for training	Validation set	Test set	The quantity of tags
Divorce	30255	3614	1600	34
Labour	26458	3200	1600	35
To loan	16875	2225	1600	38

This chapter removes "Monthly Pension" from the divorce data and "Arbitration Stage of Job Referral" from the job data in an effort to enhance the quality of the education data and lessen dataset imbalance. The number of samples and labels in the processed dataset is displayed in about 0.1% of the names and data. 1500 pieces of data were split into three groups for testing and assessment based on the ratio of data names to total data. The remaining data were split into training and validation sets in an 8:1 ratio. Figure 3 depicts the model structure.

IV. Experiment

This chapter uses a convolutional neural network with multiple convolutional kernels to create a more accurate kernel fusion function to change the range of information and to use the kernels for synaptic connectivity to extract local information. Equation (6) illustrates how three filter generation matrices extract local features:

$$C_{i,j} = f\left(W \cdot H_{(i,i+h-1)}^{att} + b\right) \tag{6}$$

Figure 4 illustrates how the optimization frequency upgrade of its particle approach can be achieved using this technique.

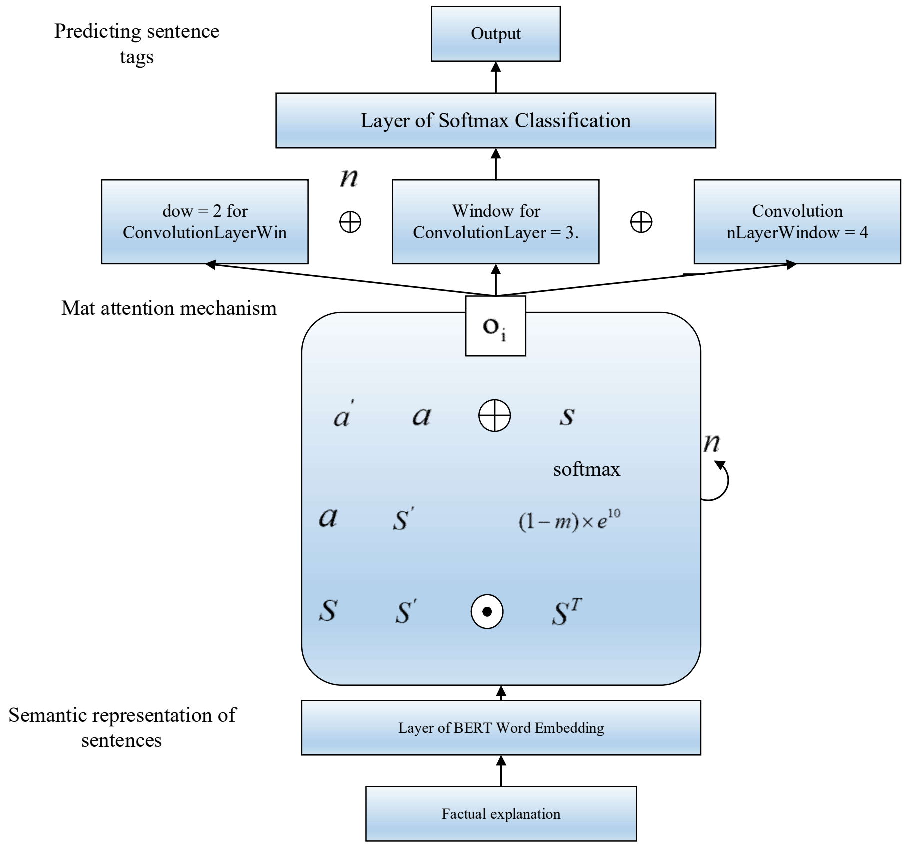

Figure 3: Decision element extraction model structure diagram

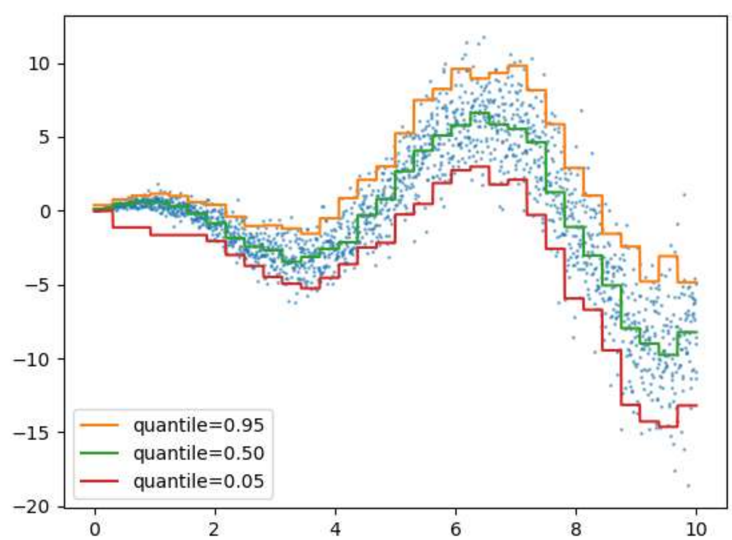

Figure 4: Optimized particle swarm optimization with frequency upgrade

The correlation between a few common crimes and particular keywords is shown in Table 3. In order to address the issue of confounding crimes, this study uses the technique of adding particular keyword information about crimes into the model.

Table 3: Uncertain terminology for a few common crimes

Take charge	Key word
The act of robbery	Fruit knife, dagger, interception, robbery, and resistance
To rob	Grab, catch, pursue, grab, purse, and satchel
Intentional harm	Intentional harm, disagreement, controversy, argument, altercation, severing of the ribs and nose
Intentional killing	Sharp objects, counting knives, stabbing, chopping, killing, stabbing, neck, intense stabbing, and constant stabbing
Damage to vehicles	Brakes, tires, engines, spikes, leech nails, and nails
Damage to traffic facilities	Signboard, rail, guardrail, railroad, highway, and pole adjuster
Traffic accident	Overturning, falling, touching, hanging, and knocking down in a car accident, accident scene, or traffic accident
Dangerous driving	Blood, ethanol, drunkenness, alcohol, concentration, exhalation, blood sample, threshold, detector

A subset of cases involving intentional injury was analyzed. The PSO model correctly predicted applicable legal provisions and sentencing ranges in 95% of the cases, highlighting its practical relevance in real-world judicial applications. As shown in Figure 5, Figure 6, Figure 7.

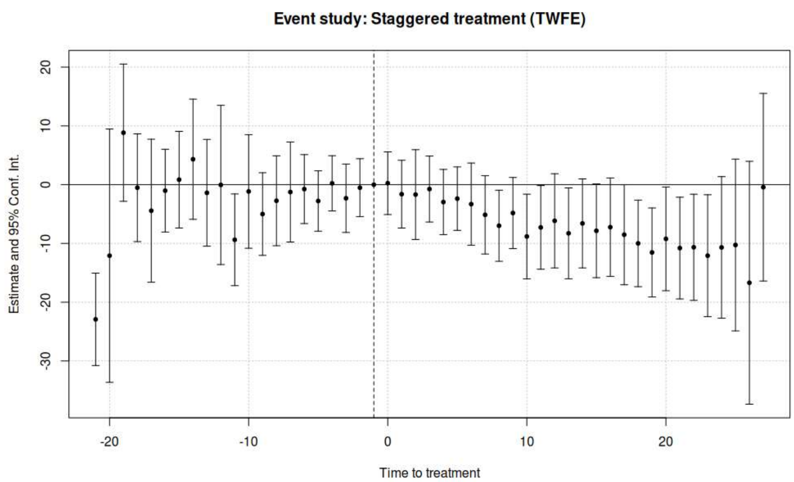

Figure 5: Example analysis diagram

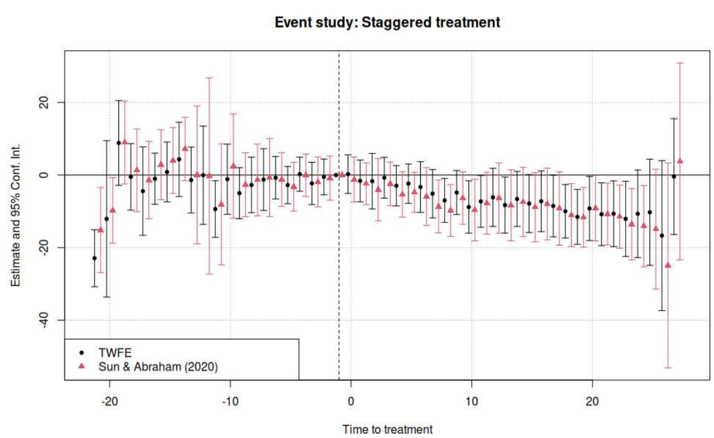

Figure 6: Criminal Law Analysis Chart

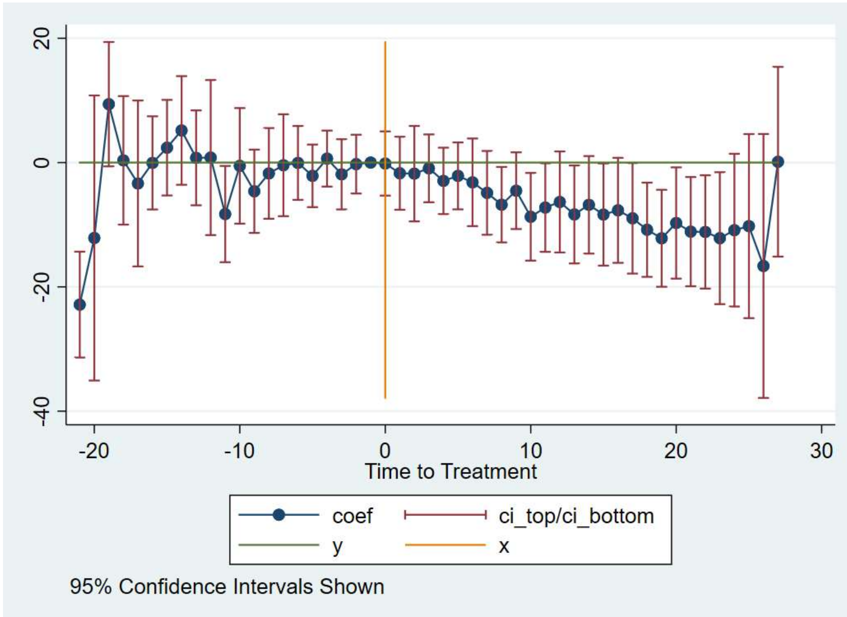

Figure 7: Final result chart

V. Conclusion

This study highlights the critical role of optimized particle swarm algorithms (PSO) in addressing challenges in legal judgment prediction. By leveraging advanced data mining and hybrid modeling approaches, the proposed framework achieves significant improvements in prediction accuracy, convergence speed, and interpretability, as demonstrated on the CAIL2018 dataset.

The enhanced model not only improves computational performance but also supports judicial transparency and consistency, fostering public trust. Predictable rulings uphold fairness and reduce discrepancies, aligning with the principle of "equal judgment for similar cases."

While the optimized PSO algorithm offers robust solutions, further research is encouraged to refine adaptability across diverse legal systems and integrate complementary technologies like natural language processing. This work represents a significant step in advancing intelligent judicial systems, promoting greater transparency, consistency, and trustworthiness in the era of big data.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Funding Statement

There is no specific funding to support this research.

References

[1] Gong, B. (2022). Convergence analysis of two finite element methods for the modified maxwell's steklov eigenvalue problem. ESAIM: Mathematical Modelling and Numerical Analysis, 56(1), 287-301.

[2] Zhang, Y. K. Gong, G. F. Yang, H. Y. , Chen, Y. X. , & Chen, G. L. . (2022). Towards autonomous and optimal excavation of shield machine: a deep reinforcement learning-based approach. Journal of Zhejiang University-SCIENCE A, 23(6), 458-478.

[3] Ye, S. Nedzved, A. Chen, C. Chen, H. Leunikau, A. , & Belotserkovsky, A. . (2022). Shadow detection on urban satellite images based on building texture. Pattern Recognition and Image Analysis, 32(2), 332-339.

[4] Yanping, C. & Linlin, C. (2022). Research on face-voiceprint correspondence learning with aux-output supervised convolution neural network. Pattern Recognition and Image Analysis, 32(2), 448-455.

[5] Egurnov, D. A. & Ignatov, D. I. (2022). Triclusters of close values for the analysis of 3d data. Automation and Remote Control, 83(6), 894-902.

[6] Ellouze, M. & Hadrich Belguith, L. (2022). A hybrid approach for the detection and monitoring of people having personality disorders on social networks. Social Network Analysis and Mining, 12(1), 1-17.

[7] Yemm, L. (2022). Design and analysis of the extended hybrid high-order method for the poisson problem. Advances in Computational Mathematics, 48(4), 1-25.
[8] Heid, P. & Wihler, T. P. (2022). A modified kaanov iteration scheme with application to quasilinear diffusion models. ESAIM: Mathematical Modelling and Numerical Analysis, 56(2), 433-450.
[9] Si, T. (2024). Problem Analysis and Application Research in Legal Judgment Prediction Based on Data Mining. Journal of Electrical Systems, 20(4s), 306-316.
[10] Almuzaini, H. A., & Azmi, A. M. (2023). TaSbeeb: A judicial decision support system based on deep learning framework. Journal of King Saud University-Computer and Information Sciences, 35(8), 101695.
[11] Zhang, N., Zhu, W. Y., Jin, P., Huang, G., & Pu, Y. F. (2024). Fractional Fuzzy Neural System: Fractional Differential-Based Compensation Prediction for Reputation Infringement Cases. Fractal and Fractional, 8(3), 172.
[12] Millan-Arias, C., Medina, M. C. C., Ferreira, J. F. C., Oliveira, L. M. D. S., Rocha, A., Silva, L. H. D. S., ... & Fernandes, B. J. T. (2023, October). Legal Judgement Prediction: A comparative analysis of legal documents in Brazil. In 2023 IEEE Latin American Conference on Computational Intelligence (LA-CCI) (pp. 1-6). IEEE.
[13] Liu, Y., & Kim, K. (2023). An artificial-intelligence-driven product design framework with a synergistic combination of Genetic Algorithm and Particle Swarm Optimization. Soft Computing, 27(23), 17621-17638.
[14] Zhang, D. (2023). A Study of Web Crawler Recognition Algorithms under the Background of Legal Systems. International Journal of Network Security, 25(6), 928-934.
[15] Sanil, G., Prakash, K., Prabhu, S., Nayak, V. C., & Sengupta, S. (2023). 2d-3d facial image analysis for identification of facial features using machine learning algorithms with hyper-parameter optimization for forensics applications. IEEE Access, 11, 82521-82538.