2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

Patent Value Prediction Method Based on Bibliographic Item Concatenation, Legal Value Calculation, and Deep Learning Fusion

Xiaoyuan Ma[1], Liang Zhang[2(✉)][0000-0003-3738-0767] and Wei Yan[1(✉)][0000-0003-0054-3657]

[1]School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China
wyaninsa@foxmail.com

[2] Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao 266237, China
zhang.l@sdu.edu.cn

Abstract.Citation analysis in high-value patent identification faces challenges such as regional bias, time lag, and insufficient legal event analysis. This paper adopts a strategy combining multi-source data fusion and deep learning techniques to enhance the accuracy and comprehensiveness of patent value assessment. The dataset is sourced from the IncoPat patent database. Patent validity duration serves as the key metric for categorization. The data is divided into three value classes: "low," "medium," and "high" after value calculation. Oversampling is applied to address imbalanced sample distribution, laying the groundwork for subsequent model research. The study introduces a patent value assessment model built on BERT and BiLSTM. The BERT embedding layer captures word semantics. The BiLSTM encoder deeply encodes the semantic structure of the text. The value prediction layer outputs classification probabilities. The BERT-BiLSTM model is compared with the BERT model. Experimental results on the test set show that the BERT-BiLSTM model achieves a lower test loss of 0.53 and a higher test accuracy of 79.40%, surpassing the BERT model's 77.31%. For the "high" value class, the BERT-BiLSTM model outperforms the BERT model in recall and F1 scores. The results demonstrate the superior performance of the BERT-BiLSTM model in patent text value classification tasks. This method exhibits significant advantages in patent value prediction.

Keywords: BERT, Patent Recommendation Framework, Patent Value Judgment, Deep Learning.

1 Introduction

Patents have long served as strategic resources for enterprise development and core elements of international competitiveness. Identifying high-value patents from a vast pool is a critical task for various patent-related activities. These activities include patent transfers, pledges, financing, auctions, and national technology strategy planning.

Regardless of their scale, such economic and political initiatives necessitate prior patent value evaluations to ensure informed decision-making. These assessments ensure effective resource allocation and the formulation of accurate intellectual property strategies. For investors, high-value patents influence decisions on whether to fund a startup. They also aid in predicting stock price fluctuations when a company files new patents [1]. For consumers, choosing products with high-value patents offers access to advanced technologies. These technologies enhance convenience, innovation, and overall quality of life. Thus, identifying high-value patents can generate greater societal benefits [2].

However, the WIPO 2022 statistical report reveals significant challenges. From the end of 2011 to the end of 2021, the global number of valid invention patents increased by 207.59%, reaching 16.4 million. Figure 1 illustrates the global distribution of valid patents.

Meanwhile, the annual growth rate of global patent applications stands at 2.9%. In 2023 alone, inventors worldwide submitted 3.02 million patent applications. This massive patent volume complicates the identification of high-value patents for policymakers, investors, and consumers.

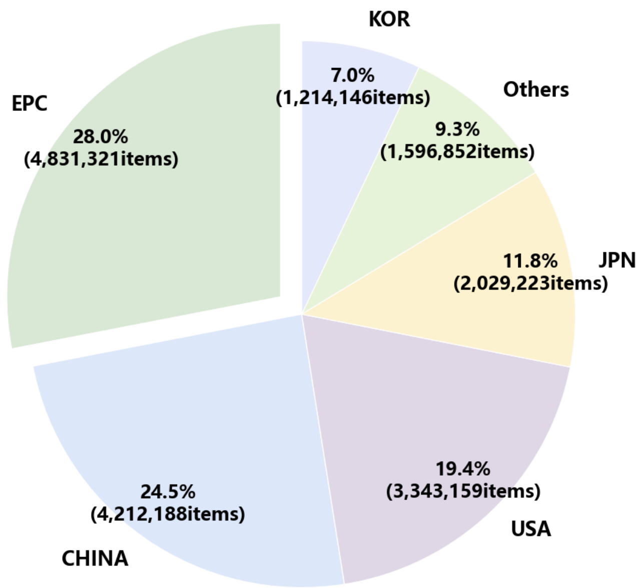

Fig 1 Global Distribution of Valid Patents as of the End of 2022

2 Literature Review

2.1 Patent Value Assessment Metrics

Current methods for identifying high-value patents can be broadly categorized into two types based on the patent information used. The first type relies on discrete data features, such as citation counts and total page numbers from bibliographic items. The second type utilizes textual information, including titles, abstracts, and applicant names.

Discrete features refer to statistical metrics like the number of independent claims and IPC classifications. Among these, citation counts (backward citations) and reference counts (forward citations) are the most commonly used. For backward citations in patent literature, Reitzig et al. [3] directly use total citation counts. In contrast, Harhoff et al. [4] and Yang et al. [5] differentiate citations from patent and non-patent literature, counting them separately. For forward citations, a straightforward approach is to count the number of times a patent is cited. Yang et al. [5] distinguish citations within five years and ten years after patent publication. Fisch et al. [6] address the time lag in cumulative citations by using the time to first citation as an indicator of patent value.

Despite their prevalence, methods based on surface features have significant limitations. These include regional and temporal biases, which hinder cross-country patent value comparisons under a unified framework. The drawbacks of using citation metrics are as follows:

First, patent citations originate from scientific literature citations. Eugene Garfield [7] introduced the Science Citation Index in 1955, aiming to reflect important literature based on researchers' informed judgments. Scientific citations are autonomous and self-regulated by scholars. Patent citations, however, differ significantly.

Second, regional differences lead to variations in examiner practices and systems. Some citations are made by patent examiners rather than inventors. This results in discrepancies in citation counts due to differing habits of examiners across regions and varying patent policies. For example, the U.S. patent system mandates applicants to disclose all relevant technical information via Information Disclosure Statements (IDS). This leads to over-citation in U.S. patents, unlike in other countries without such requirements.

Third, language barriers arise from global regional differences. Applicants and examiners often use familiar languages for patent searches, which may result in non-English or non-major language patents being overlooked internationally. This reduces their visibility and citation counts.

Fourth, patent citations are limited by time constraints. Scientific literature is published upon acceptance by conferences or journals. In contrast, there is a significant gap between patent application and publication dates, which varies based on filing strategies. For instance, an important patent A may be filed early for priority but published 18 months later, while patent B is published six months after filing to expedite authorization. This one-year difference can cause patent A to miss its citation peak in fast-evolving industries like telecommunications and aerospace.

Additionally, differences in technological fields, market sizes, and the completeness and accuracy of patent databases across countries contribute to citation imbalances.

These factors collectively undermine the reliability of citation data in reflecting true patent value and influence.

2.2 Patent Recommendation Methods

With the rapid advancement of information technology, the development of systems capable of intelligently analyzing and recommending patents using artificial intelligence and big data has become a significant research direction. Leveraging technologies such as Natural Language Processing (NLP) [8] and Machine Learning (ML) [9], these systems can automatically understand the technical content of patent documents, identify patents highly relevant to users' areas of interest, and recommend similar patent documents.

Currently, commonly used patent recommendation methods can be categorized into three types. First, Keyword-based methods [10, 11]: These methods automatically extract keywords from patents, using keyword clusters to represent the content of a patent. Recommendations are made by matching keywords with user queries during searches. However, the results of such methods often exhibit significant variability and lack accuracy. Second, Machine learning-based text processing methods: These methods extract features from patent texts and employ traditional machine learning algorithms (e.g., Support Vector Machines (SVM), Random Forests, Naive Bayes) for patent classification, clustering, or recommendation [12, 13]. While these methods can capture local features of patent texts to some extent and optimize model performance through feature engineering, they face limitations in handling long texts and cross-domain semantics due to the dense technical terminology and complex semantics of patent texts. Additionally, feature engineering relies on manual design, which is time-consuming and difficult to scale for large patent datasets. Last, Deep learning-based methods: These methods use word embeddings to convert text into numerical representations, mapping large vocabularies into low-dimensional dense vector spaces to capture semantic relationships between words. Numerous transformer-based neural network methods have been applied to patent text analysis and recommendation tasks [14-16]. Compared to keyword-based and machine learning-based methods, deep learning-based approaches offer superior performance, efficiency, robustness, and cost-effectiveness, making them more suitable for large-scale patent data analysis and recommendation.

Despite their excellent performance in intelligent patent text analysis and recommendation tasks, deep learning-based methods still face challenges. Recommendation algorithm architectures are often complex, requiring users to process multiple intermediate results during patent recommendation [17], making end-to-end recommendation difficult. While these algorithms achieve high recommendation accuracy, they often lack sufficient judgment of patent value, making it challenging to prioritize high-value patents at the top of recommendation lists. This limitation hinders the practical application of patents.

3 Construction of Patent Value Assessment Methodology

3.1 Framework

This paper proposes a patent value assessment methodology based on the integration of bibliographic item concatenation, patent validity calculation, and deep learning models. Figure 2 illustrates the detailed flowchart of the patent value assessment process, which is divided into three parts: value assessment methodology, data preprocessing, and model training and optimization.

Part 1: Value Assessment Methodology

The process begins with the patent database. INID (Internationally agreed Numbers for the Identification of (bibliographic) Data) are extracted and concatenated to form patent data texts. Historical legal statuses are then incorporated to calculate patent lifespan. Finally, patent value labels are derived based on these calculations.

Part 2: Data Preprocessing

The data is first cleaned to remove invalid or erroneous entries. It is then split into three subsets: 80% for training, 10% for testing, and 10% for validation. This prepares the data for subsequent model training.

Part 3: Model Training and Optimization

The BERT and BiLSTM models are combined to construct the BERT-BiLSTM model. The model is trained using the training set and evaluated using the testing and validation sets. This process ultimately enables the assessment of patent value.

By integrating data processing and model training, this workflow provides a scientifically robust approach to patent value assessment.

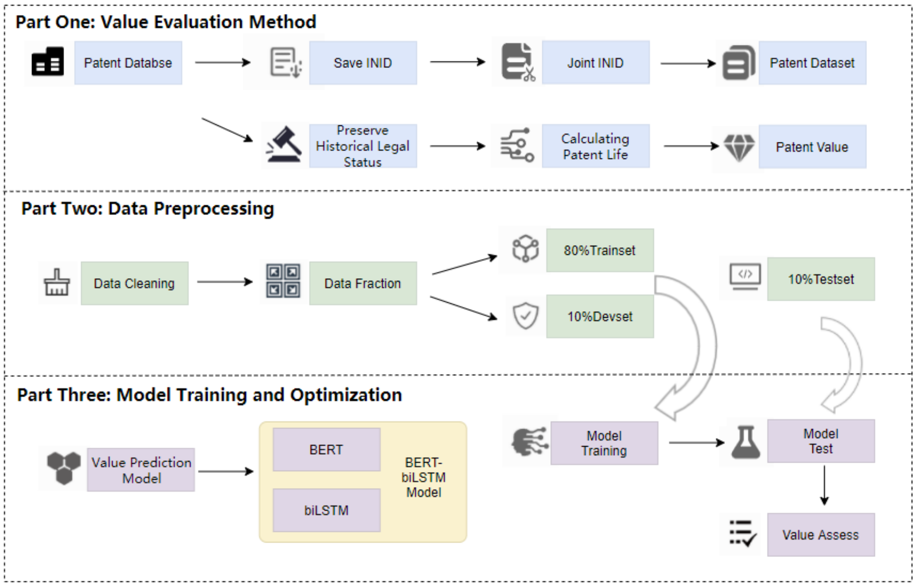

Fig 2 Construction Process of the Patent Value Prediction Methodology

3.2 Patent Lifecycle and Patent Legal Status

This section introduces two critical concepts closely related to patent value: the patent lifecycle and patent legal status. It systematically explains the dynamic mechanism of patent value formation within the framework of patent legal systems. It focuses on the cumulative effects of legal events at each stage of the patent lifecycle on the market value of technological achievements.

The patent lifecycle encompasses the entire process from patent application, authorization, implementation, to expiration. Each stage contains rich information that influences patent value assessment to varying degrees. Patent legal status, on the other hand, clarifies the legal protection status of a patent, such as its validity and the existence of infringement disputes. This directly affects the enforceability and market stability of the patent.

From the date of application submission, a patent begins its lifecycle, which can last up to 20 years. As illustrated in Figure 3, "The Full Cycle of Invention Patents," the patent undergoes numerous rigorous and complex legal events during this period.

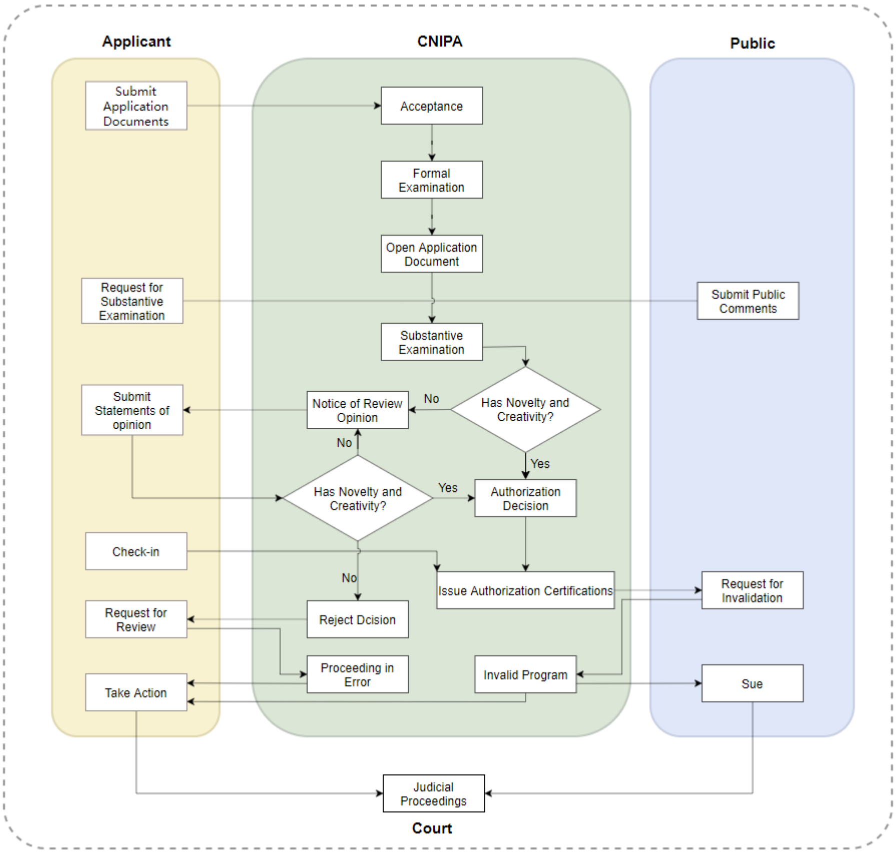

Fig 3 Lifecycle of Invention Patents and Legal Events

As shown in Figure 3, the full-cycle management system for invention patents follows a typical three-stage evolutionary path: the front-end examination procedure (including formal examination, substantive examination, and authorization announcement), the mid-term rights maintenance phase (including invalidation procedures and administrative litigation procedures), and the end-term rights termination phase. This institutional design establishes a complete closed loop from legal confirmation of rights to market realization through strict procedural requirements.

From the perspective of the examination procedure, invention patent applications undergo sequential stages, including formal examination (20 days), preliminary examination (approximately 3-4 months), and substantive examination (typically 2-4 years). Among these, the substantive examination stage, with its triple criteria of inventiveness, novelty, and utility, constitutes the core mechanism for screening technical value. Statistical studies [18] indicate that only about 40%-58% of patent applications pass the substantive examination, significantly raising the technical quality baseline of authorized patents. Notably, on November 29, 2024, Heng Fuguang, spokesperson for the China National Intellectual Property Administration, announced that the average authorization period for invention patents in China has been shortened to 15.6 months. However, the examination period for complex technologies may still extend beyond three years. This positive correlation between time cost and technical complexity indirectly confirms the role of examination rigor in enhancing value.

The post-authorization rights maintenance phase consists of three key subsystems: First, the registration and effectiveness procedure in administrative confirmation (within three months from the authorization announcement date). Second, the rights stability guarantee mechanism (invalidation request examination period typically 1-2 years). And third, the judicial relief pathway (first-instance administrative litigation cycle of approximately six months). This institutional safeguard not only extends the statutory protection period but, more importantly, reinforces the legal certainty of patent technologies through repeated verification.

Particular attention should be paid to the procedural characteristics of the rights termination phase. When a patent is terminated due to non-payment of annual fees, voluntary abandonment, or invalidation, its remaining protection period incurs "institutional loss." Invention patents, due to their higher technical complexity, require greater maintenance costs, and their rights duration exhibits a significant positive correlation with the technical lifecycle. Furthermore, the extension of rights stability allows patent holders to more fully implement technology commercialization strategies, forming barriers of technical advantage.

Through the analysis of the above institutional framework, it becomes evident that patent legal status essentially constitutes a time function of technical value. The passage or failure of each procedural node not only affects the remaining length of the statutory protection period but, more importantly, converts the market potential of technological achievements into quantifiable economic value through the layer-by-layer confirmation of legal effectiveness. This value conversion mechanism is particularly pronounced in the later stages of the lifecycle: when a patent enters the litigation phase, its market valuation often increases by 2-3 times compared to the stable period, reflecting the strengthening effect of legal dispute resolution on technical value signals. Therefore, the extension of patent validity is essentially the result of technical value being certified by the legal system, rather than a simple linear accumulation of time.

3.3 Bibliographic Item Concatenation

The logical coherence of patent texts is primarily reflected in the intrinsic relationships among their bibliographic items. These items are not only integral components of patent applications but also embody the rigor of patent law. By conducting an in-depth analysis of the interconnections among these items, a multidimensional framework for patent value assessment can be constructed. Table 1 presents the classification and examples of patent bibliographic items.

Table 1 Patent description item classification and examples

Type	No.	Patent Information	Examples
Indicator Data	1	Application Date	20240611
	2	Number of Applicants	1
	3	Citation	7
	4	Number of Document Pages	30
Discrete Data	1	Patent Type	Invention, Utility Model
	2	IPC	H02K3/28；H02K3/12
Continuous Data	1	Applicant	Xiaomi Automobile Technology Co., Ltd.
	2	Title of Invention	Stator, Electric Motor, Power Assembly and Vehicle
	3	Abstract	This disclosure relates to a stator, an electric motor, a power assembly and a vehicle. The stator includes a stator core and a flat - wire winding. A plurality of stator slots are arranged circumferentially on the stator core. The flat - wire winding comprises a multi - phase winding passing through the plurality of stator slots.

In quantitative data, numerical values are the primary form of representation. In discrete feature information, categories are the main form of representation, such as the number of claims and application types. The specific content of claims is often drafted by patent agents, primarily to provide legal protection for the technology rather than directly reflect the degree of innovation. For example, the patent with publication number USRE041156E, titled "Notched Brush and Makeup Device Including the Same," contains 339 claims. However, these claims only represent minor improvements or simple extensions of existing technologies rather than substantive innovations. Moreover, if patent value assessment systems widely adopt the number of claims as a key metric, it may incentivize patent agents to maximize the number of claims when drafting patent applications to enhance the assessed value of the patent. This behavior could lead to a "quantity over quality" phenomenon, thereby undermining the scientific rigor and accuracy of the patent assessment system.

From the perspective of information complementarity, the title constructs the technical orientation of the patent through the chain enumeration of technical components. The abstract, as the core description of the patent's technical solution, syntactically highlights the substantive innovation points of the technical solution, making it the peak region of information entropy for technical features. The applicant information, as the identity marker of the patent subject, holds unique value in technical characterization. Empirical analysis shows that patents from different applicants (e.g., enterprises, research institutions, individuals) exhibit significant differences in the distribution of technical features. Enterprise applications often have a stronger commercialization orientation, research institution patents tend to reflect technological frontiers, and individual applications may showcase unique innovative perspectives. These differences not only reflect the sources of technological innovation but also provide important references for in-depth assessment of patent technologies.

Through in-depth analysis of patent bibliographic items, it becomes evident that patent technical characterization is a multidimensional systematic project. The title, abstract, and applicant information, as the core elements of patent documents, each serve distinct technical characterization functions while complementing one another, collectively forming the foundational framework for in-depth assessment of patent technologies. This multidimensional feature analysis method offers a new research pathway for patent information mining and patent value assessment, holding significant theoretical and practical value.

In summary, the title provides technical orientation, the abstract presents technical details, and the applicant information reflects the source of innovation. These three elements complement each other, forming a comprehensive technical characterization system. This multidimensional feature analysis framework not only effectively supports the assessment of patent technical depth but also provides a new research perspective for patent information mining and patent value analysis.

3.4 Value Prediction Model

This paper primarily constructs a patent value assessment model based on BERT and BiLSTM. This model is a BERT-BiLSTM hybrid designed for text tri-classification tasks. In terms of model architecture, the input text sequence passes through the BERT embedding layer. Each word is transformed into a word embedding vector enriched with contextual semantic information, effectively capturing the meaning of words in different contexts. Subsequently, the word embedding vectors are fed into the BiLSTM encoder. BiLSTM utilizes the coordinated operation of input, forget, and output gates. It processes the sequence simultaneously through both forward and backward LSTM units, thoroughly capturing the contextual dependencies of words. This enables deep encoding of the semantic structure of the text. Finally, the output of the BiLSTM encoder is passed to the value prediction layer. Through a fully connected layer and the softmax function, the model outputs the probability values of the text belonging to each of the three categories. During model training, the cross-entropy loss function is employed to measure the discrepancy between predicted results and true labels, quantifying the classification error of the model. The optimization process leverages the AdamW optimizer. It effectively updates model parameters while preventing overfitting through a weight decay mechanism, enhancing the model's generalization capability and robustness.

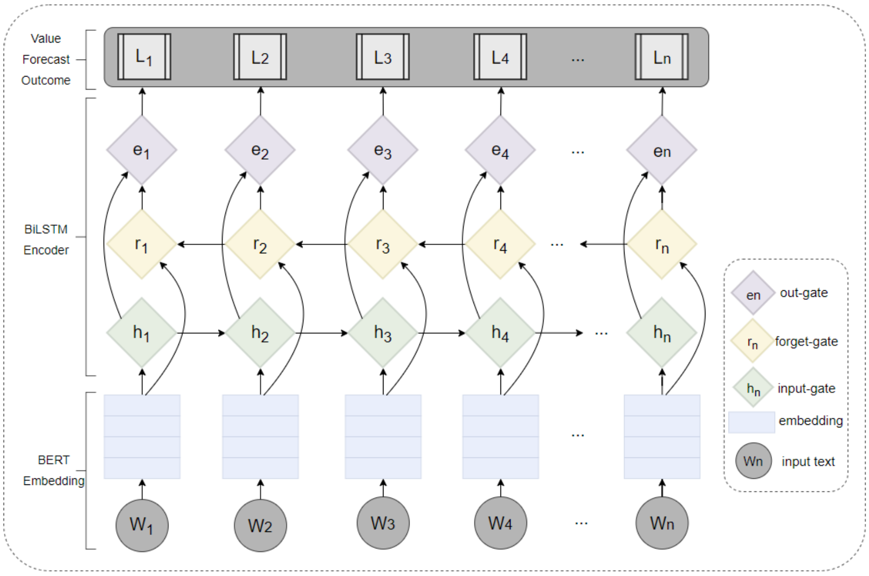

Fig 4 The architecture of the patent recommendation algorithm based on the BERT-BiLSTM hybrid model.

4 Experimental Results and Analysis of Patent Value Assessment

To evaluate the effectiveness of the proposed model, we compared it with the BERT model. The experiments were conducted on an AutoDL cloud computing server equipped with an RTX 3080x2 (20GB) GPU. The runtime environment included Python 3.8 (Ubuntu 20.04), CUDA 11.3, and PyTorch version 1.11.0.

4.1 Data Description and Parameter Settings

The dataset was sourced from the IncoPat patent database, covering patents filed with the China Intellectual Property Office from 2000 to 2020. A sample of the processed patent dataset is shown in Table 2. Table 3 summarizes the distribution information of the dataset. The dataset focuses on patent value assessment, using the length of patent validity [0 - 20] as the evaluation dimension. The data was discretized into bins with cut points at [0, 0.333, 0.8, 1], corresponding to three value categories: "low," "medium," and "high," labeled as 0, 1, and 2, respectively. In the original data, the sample distribution across value categories was imbalanced. The "medium" category had 85,306 samples, accounting for 0.474856 of the total. The "low" category had 64,736 samples, representing 0.360353, while the "high" category had only 29,604 samples, constituting 0.164791. The actual bin boundaries were: low level [0.00, 8.00), medium level [8.00, 14.00), and high level [14.00, 20.00]. To optimize sample balance, the dataset underwent oversampling, ensuring that each category reached 85,306 samples. This provided a more balanced data foundation for subsequent model training and analysis. The hyperparameter settings for model training are detailed in Table 4.

Table 2 The partial bibliographic items and legal status information of patents to be concatenated.

Title	Applicant	Abstract	Historical legal state	Value
Software protecting method and device	Beijing Thinking Rock Software Technology Co., LTD	The invention discloses a software protecting method. The method comprises the following steps : a plurality of code segments in N code segments of protected software	2011Publication|2013Reject	0
Diagnostic information providing system for construction machine	Hitachi Construction Machinery Corporation	Disclosed is an information providing system for a construction machine, which can enhance the accuracy of a warning/abnormality occurring factor diagnosis. The information providing system comprises a user-side personal computer	2009Publication|2013Grant|2016Lapse	7
Electric rheologic liquid electrode plate for surface modification	The Institute of Physics, Chinese Academy of Sciences	The invention relates to a surface-modified electrorheological fluid electrode plate. A rough, wear-resistant and low-conductivity modified layer is added to the surface of a metal electrode plate.	2006Publication|2010Grant	18

Table 3 The distribution information of the dataset after random oversampling processing.

Tags	Class name	Number of original samples	Raw sample distribution	Actual box boundary	Number of samples after oversampling
0	low	64736	0.360353	[0.00, 8.00)	85306
1	medium	85306	0.474856	[8.00, 14.00)	85306
2	high	29604	0.164791	[14.00, 20.00]	85306

Table 4 The hyperparameter settings for model training.

Hyperparameter	BERT	BERT - BiLSTM
Epochs	6	6
Bach_size	128	128
Number_class	3	3
Pad_size	64	64
Learning_rate	5e-5	5e-5
Hidden_size	768	768
LSTM_hidden_size	-	256

4.2 Experimental Results

Figure 5 records the changes in the Loss values during the training of the BERT-BiLSTM model. This figure illustrates the variations in training loss (Train Loss) and validation loss (Val Loss) throughout the training process. The model demonstrates progressive learning dynamics over six training epochs, with the final validation accuracy reaching 80.98% and the test accuracy achieving 79.40%. The specific analysis is as follows:

From an overall trend perspective, the fitted lines (smooth red and blue curves) for both training loss (light blue fluctuating curve) and validation loss (orange fluctuating curve) show a downward trend. This indicates that as the number of iterations increases, the model's performance on both the training and validation sets gradually improves, with the loss values steadily decreasing.

Comparing training loss and validation loss, during the initial iterations [0, 45], the training loss is higher than the validation loss. This is likely because the model is in the early stages of training and has not yet fully learned the data features. As the iterations progress, within the range of [45, 80], the training loss decreases rapidly and falls below the validation loss. At this point, the model's performance on the training set surpasses that on the validation set, indicating a tendency toward overfitting. Toward the end of the iterations, the training loss exhibits significant fluctuations and an upward trend, while the validation loss remains relatively stable. This further suggests that the model overfits the training set in the later stages, while its generalization ability on the validation set remains relatively stable. The training loss decreases from 1.1 to 0.2, and the validation loss decreases from 1.1 to 0.53, demonstrating the effectiveness of gradientbased optimization. Due to the application of early stopping, severe overfitting is avoided.

The fitted line for training loss (red curve) shows a significant decline and continues to decrease in the later stages, reflecting the model's continuous learning and improving fitting ability on the training set. The fitted line for validation loss (blue curve) declines more gently, indicating that the model's performance improvement on the validation set is relatively slower. It stabilizes in the later stages, suggesting that the model's generalization ability does not improve correspondingly with excessive fitting on the training set.

Fig 5 The changes in loss values during the training process of the hybrid model.

Based on the confusion matrix heatmap in Figure 6, it can be concluded that the hybrid model performs best for the "high" category (F1=0.91). This is likely due to the unique linguistic patterns in long-lived patents, such as broader technical claims and standardized abstract structures. The confusion in the "medium" category (F1=0.69) suggests inherent ambiguity in determining intermediate lifespans. Additionally, there is a precision-recall imbalance in the low and medium categories (low: P=0.74/R=0.82; medium: P=0.74/R=0.65), reflecting differences in the salience of features across categories. The asymmetry of the confusion matrix reveals that "medium" category patents are frequently misclassified as low (2,313 cases) or high (693 cases), possibly indicating a nonlinear relationship between textual features and medium lifespan outcomes. The experiments demonstrate that the BERT component effectively captures domain-specific semantics, while the BiLSTM layer likely models sequential dependencies in inventor networks and technical descriptions.

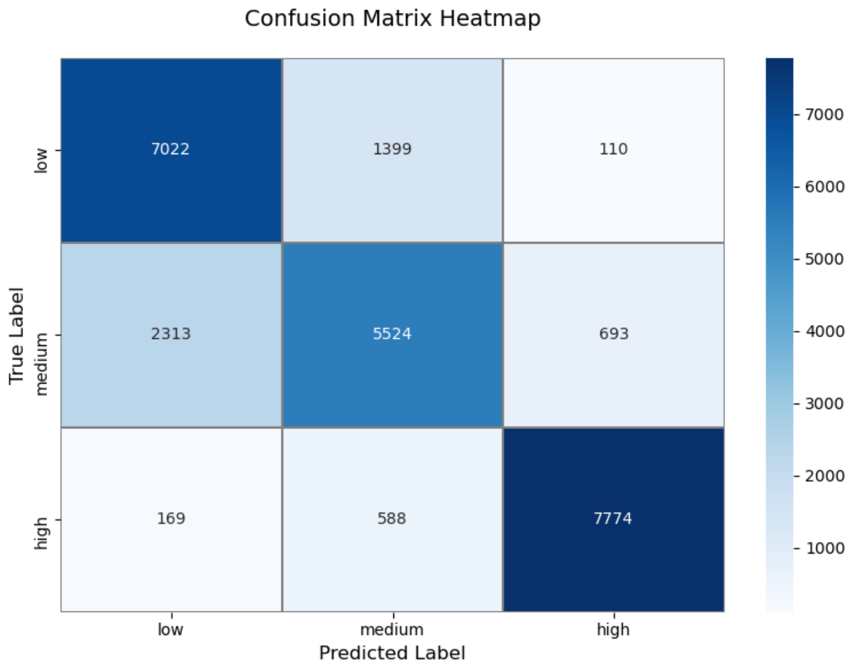

Fig 6 The confusion matrix heatmap of the classification results for the hybrid model.

Table 5 Comparison of the training experimental results between BERT-BiLSTM and BERT.

	Test set loss value	Test set accuracy	High-class recall rate	High-class f1-score
Bert-Bilstm	0.53	79.40%	91.13%	0.9088
Bert	0.54	77.31%	86.97%	0.8995

Table 5 demonstrates that in the experiment of patent text value tri-classification, the BERT-BiLSTM model exhibits superior performance compared to the BERT model. From the test results, the test loss of BERT-BiLSTM is 0.53, lower than BERT's 0.54, indicating a smaller discrepancy between its predictions and the true labels. Additionally, the test accuracy of BERT-BiLSTM reaches 79.40%, higher than BERT's 77.31%, suggesting its overall classification correctness is improved.

Table 6 The classification results report for the BERT-BiLSTM model.

	Precision	Recall	F1-score
Low	73.88%	82.31%	0.7787
Medium	73.55%	64.76%	0.6887
High	90.64%	91.13%	0.9088

Table 6 illustrates the performance metrics of the BERT-BiLSTM model in terms of precision, recall, and F1 score. The BERT-BiLSTM model excels in the "medium" and "high" categories, particularly in the "high" category, where the recall rate reaches 91.13% and the F1 score is 0.9088. This demonstrates its enhanced capability in capturing and accurately classifying these categories. Overall, the BERT-BiLSTM model exhibits outstanding performance in the tri-classification task of patent texts.

5 Conclusions

This chapter presents a novel patent value prediction method that integrates bibliographic information, patent validity period, and deep learning techniques. The proposed BERT-BiLSTM model demonstrates superior performance in three-class classification tasks, achieving a test accuracy of 79.40% and outperforming the baseline BERT model. The experimental results reveal that the BERT-BiLSTM model particularly excels in identifying high-value patents, as evidenced by its improved recall and F1 scores. The methodology's effectiveness is further validated through comprehensive analysis of training and validation loss curves, establishing its potential for patent value assessment applications.

However, there are still several issues that need to be addressed in future research. The quantification of legal value remains limited, as the construction of patent validity indicators still relies on manual rules without fully capturing complex legal event correlations. Additionally, the model's capability to process long texts is insufficient for comprehensive analysis of patent specifications and claim documents. Future work should focus on developing more sophisticated legal value quantification methods and enhancing the model's capacity for handling extensive patent text content. These improvements will contribute to more accurate and comprehensive patent value prediction systems.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (62002207), the Natural Science Foundation of Shandong Province (ZR2022YQ05).

References

1. Wenjun Wang. Valuation of Patents Using Stock Market Responses. University of California, Berkeley, 2023.
2. Pu-yan Nie, Hong-xing Wen, and WANG Chan. R & d competition and patent values. Journal of Innovation & Knowledge, 8(1):100308, 2023.
3. Reitzig M. Improving patent valuations for management purposes—validating new indicators by analyzing application rationales[J]. Research policy, 2004, 33(6-7): 939-957.
4. Harhoff D, Scherer F M, Vopel K. Citations, family size, opposition and the value of patent rights[J]. Research policy, 2003, 32(8): 1343-1363.
5. Yang G, Lu G, Xu S, et al. Which type of dynamic indicators should be preferred to predict patent commercial potential?[J]. Technological Forecasting and Social Change, 2023, 193: 122637.

6. Fisch C, Sandner P, Regner L. The value of Chinese patents: An empirical investigation of citation lags[J]. China Economic Review, 2017, 45: 22-34.
7. Garfield E. Citation indexes for science: A new dimension in documentation through association of ideas[J]. Science, 1955, 122(3159): 108-111.
8. Bekamiri H, Hain D S, Jurowetzki R. Patentsberta: A deep nlp based hybrid model for patent distance and classification using augmented sbert[J]. Technological Forecasting and Social Change, 2024, 206: 123536.
9. Miric M, Jia N, Huang K G. Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents[J]. Strategic Management Journal, 2023, 44(2): 491-519.
10. Hu J, Li S, Yao Y, et al. Patent keyword extraction algorithm based on distributed representation for patent classification[J]. Entropy, 2018, 20(2): 104.
11. Choi J, Hwang Y S. Patent keyword network analysis for improving technology development efficiency[J]. Technological Forecasting and Social Change, 2014, 83: 170-182.
12. Wu C H, Ken Y, Huang T. Patent classification system using a new hybrid genetic algorithm support vector machine[J]. Applied Soft Computing, 2010, 10(4): 1164-1177.
13. Wu J L, Chang P C, Tsao C C, et al. A patent quality analysis and classification system using self-organizing maps with support vector machine[J]. Applied soft computing, 2016, 41: 305-316.
14. Choi S, Lee H, Park E, et al. Deep learning for patent landscaping using transformer and graph embedding[J]. Technological Forecasting and Social Change, 2022, 175: 121413.
15. Choi S, Lee H, Park E L, et al. Deep patent landscaping model using transformer and graph embedding[J]. arXiv preprint arXiv:1903.05823, 2019.
16. Pujari S C, Friedrich A, Strötgen J. A multi-task approach to neural multi-label hierarchical patent classification using transformers[C]//Advances in Information Retrieval: 43rd European Conference on IR Research, ECIR 2021, Virtual Event, March 28–April 1, 2021, Proceedings, Part I 43. Springer International Publishing, 2021: 513-528.
17. Ni X, Samet A, Cavallucci D. Similarity-based approach for inventive design solutions assistance[J]. Journal of Intelligent Manufacturing, 2022, 33(6): 1681-1698.
18. Graham S J H, Marco A C, Miller R. The USPTO patent examination research dataset: A window on the process of patent examination[J]. Georgia Tech Scheller College of Business Research Paper No. WP, 2015, 43.