Journal of Artificial Intelligence Practice (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/jaip.2025.080105
ISSN 2371-8412 Vol. 8 Num. 1

Exploration of Formalization Techniques for Geometric Entities in Planar Geometry Proposition Texts

Huiying Tang[1,a], Xijia Wu[2,b], Lian Chen[2,c,*]

[1]School of Modern Information Industry, Guangzhou College of Commerce, Guangzhou, China
[2]School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou, China
[a]thyxxyx@163.com, [b]wuchderz@qq.com, [c]chenlian@gzhu.edu.cn
**Corresponding author*

Keywords: LLMs, NER, BiLSTM, Text Formalization

Abstract: This paper explores the formalization technology of geometric entities in planar geometry proposition texts and proposes a formal definition of planar geometric entities. The study investigates the formalization of planar geometry proposition texts using NER technology and the latest large language models. The research utilizes the classic NER model BiLSTM-CRF for sequence labeling and training to achieve geometric entity recognition, followed by dictionary-based formalization. Additionally, the study employs the PROMPT approach based on Baidu's ERNIE-Bot and iFlytek's Spark large language models for geometric entity recognition and subsequent formalization. The results show that the BiLSTM-CRF model achieves an accuracy of 98% in geometric entity recognition, demonstrating stable performance but limited flexibility. ERNIE-Bot performs exceptionally well, with a recognition accuracy of 99%, although it struggles with certain special geometric entities. The Spark large model performs slightly worse, with issues primarily related to repeated or missed entity recognition. While the classic deep learning model exhibits high stability for recognizing common expressions in planar geometry proposition texts, it has limited adaptability to new and complex expressions not present in the training set. Large language models excel in simple tasks due to their strong language processing and contextual understanding capabilities; however, their accuracy depends heavily on the precision of PROMPT design, and their commercial use entails significant costs. This paper provides technical insights into the formalization of geometric entities in planar geometry proposition texts and offers a valuable reference for applications in intelligent education, particularly in geometry teaching and automated proof generation.

1. Introduction

With the rapid development of artificial intelligence technologies, Natural Language Processing (NLP) has gradually become a crucial research foundation that drives multidisciplinary studies and applications. Named Entity Recognition (NER), as one of the core tasks within the NLP domain, refers to the process of identifying predefined named entities or entity references from unstructured natural language text. These named entities typically include personal names, geographical locations, dates, etc. As an essential tool for information extraction, NER provides critical technical support for numerous downstream tasks. Depending on the specific application scenario, NER technology can be tailored to particular domains. For example, in the field of intelligent education [1], NER technology demonstrates significant potential. By extracting and recognizing core concepts and knowledge points from question texts, it can assist in building intelligent question bank systems [2]. Additionally, extracting keywords from student-submitted answers can support automatic grading and instructional feedback.

Planar geometry proposition texts, as a form of standardized language in the field of mathematics education, typically involve geometric entities such as points, line segments, angles, and shapes, as well as their interrelationships [3]. This unique linguistic characteristic provides an opportunity to apply NER tasks to the domain of geometric proposition texts. The formal study of geometric entities in planar geometry proposition texts can not only enhance the level of intelligence in teaching planar geometry within intelligent education, providing innovative technical solutions for classroom teaching and question bank construction, but also lay the groundwork for the automation of geometric proofs [4, 5]. This, in turn, will promote the deep integration of educational technology and artificial intelligence, offering new perspectives and technical support for the future of geometry education.

1.1. Significance

Geometry, an ancient discipline, is a well-structured system and a prototype of axiomatic theory. Axioms are conclusions repeatedly summarized from practice and are self-evident truths, while theorems are conclusions derived through a finite logical process starting from fundamental propositions and proven to be true. The proof of theorems has always been one of the most significant tasks in mathematical research. However, in practice, proving a theorem may take several generations of scholars to explore. This has led to significant attention from experts in mathematics, computer science, and other fields toward the automation of geometric proofs. The recognition of entities within geometric proposition texts [6] plays a crucial role in realizing geometric proof automation, as the automation process relies on an accurate understanding of key entities and their interrelationships in the proposition texts. By applying Named Entity Recognition (NER) technology to process planar geometry proposition texts, natural language descriptions can be transformed into a formalized [7] input that can be understood by computers, thereby enhancing the reasoning capabilities and application performance of geometric proof machines [8].

Named Entity Recognition (NER), as a classical research topic in the field of Natural Language Processing (NLP), has become one of the foundational technologies for numerous key tasks. Its research paradigm has evolved from early methods based on bag-of-words and rules to traditional machine learning methods, such as Hidden Markov Models (HMM) and Conditional Random Fields (CRF), and more recently, to deep learning-based approaches. Owing to the significant advantages of deep learning in feature extraction, sequence labeling, and other areas, deep learning methods [9] have become the mainstream direction in current NER research. Classic models such as BiLSTM and BERT are widely applied. In recent years, the rapid development of multimodal large models [10], such as ChatGPT and ERNIE-Bot, has injected new vitality into NER research. These large models not only excel in semantic understanding and generation tasks but also provide novel research ideas and technical frameworks for NER, fostering further innovation and application expansion in the field.

2. Related work

2.1. Mainstream models

2.1.1. Deep learning model

Named Entity Recognition (NER) is a classic task in the field of Natural Language Processing (NLP) with numerous well-established models. When applied to geometric entity extraction from planar geometry proposition texts, the goal is to identify predefined geometric elements such as points, lines, angles, and geometric shapes from textual descriptions of geometric propositions. Due to the unique characteristics of geometric proposition texts, which involve domain-specific geometric concepts and properties, NER models require adjustments and enhancements to adapt to the geometry domain.

For instance, the Conditional Random Field (CRF) model, a traditional machine learning model proposed in 2001 [11], employs a sequence labeling approach. It assigns random values to each position in a sequence according to a distribution using a "random field" and models the data based on conditional probabilities to output sequence predictions constrained by labeling rules. In 2015, Huang et al. [12] introduced a BiLSTM-CRF model for NER tasks, achieving enhanced performance. This model leverages the ability of CRF to learn contextual constraints in sequence labeling and applies these constraints to refine the predictions of the BiLSTM model, which captures contextual information in the text. By doing so, it corrects erroneous outputs and ensures the validity of the model's final predictions. In 2018, the Transformer-based pre-trained language model BERT [13] emerged as a breakthrough, achieving state-of-the-art results across various NLP tasks by capturing bidirectional contextual information. Over subsequent years, numerous BERT variants have been developed to address diverse downstream tasks. For example, SpanBERT [14], proposed by Joshi et al. in 2019, introduced span-level refinements to improve the handling of multi-word entities in NER tasks. In 2021, Microsoft introduced DeBERTa [15], which enhanced performance on long texts and complex tasks by incorporating disentangled attention mechanisms and decoding enhancements.

To date, BiLSTM-CRF models and BERT-based models [16, 17] remain classic choices for NER and sequence labeling tasks, each excelling in different scenarios. BiLSTM-CRF models are particularly suited for tasks requiring domain knowledge and feature engineering. They utilize bidirectional sequences to capture contextual information, and the CRF layer further enhances labeling consistency and entity dependencies. Compared to BERT, BiLSTM has lower computational overhead and adapts quickly to specific tasks in smaller training datasets.Therefore, this study adopts the classical BiLSTM-CRF deep learning model to achieve the goal of geometric entity recognition in planar geometry proposition texts, considering its efficiency and effectiveness in scenarios requiring domain-specific adjustments and contextual modeling.

2.1.2. LLMs

In 2022, OpenAI released the ChatGPT large language model [18], marking a revolutionary breakthrough in the field of Natural Language Processing (NLP). It demonstrated unprecedented capabilities in conversational generation, semantic understanding, and multi-task processing. Following the introduction of ChatGPT, the development and application of large language models have advanced rapidly. Numerous technology companies, both domestic and international, have launched their own large language models, such as Baidu's ERNIE-Bot (Wenxin Yiyan), iFlytek's Spark, Anthropic's Claude, Meta's LLaMA, Microsoft's Turing, Google's PaLM, and Tencent's Hunyuan. Notably, given the differences in grammar, vocabulary, and contextual nuance between Chinese and English, domestic language models are often trained on extensive Chinese corpora, resulting in superior performance in Chinese contexts. For instance, ERNIE-Bot, based on ERNIE 4.0, emphasizes semantic understanding and multi-task processing in Chinese, making it particularly well-suited for tasks involving Chinese content generation and understanding. Similarly, iFlytek's Spark is specifically designed for natural language tasks in Chinese contexts. Therefore, for the objectives of this study, domestic large language models are more suitable.

Large language models, pre-trained on massive datasets, exhibit exceptional contextual understanding capabilities, enabling them to perform tasks without relying on extensive domain-specific labeled data. For example, the task of identifying geometric entities from geometric proposition texts can be approached by inputting planar geometry propositions directly into a large language model, which can output corresponding geometric entities without requiring a complex annotated dataset specific to the geometry domain. In this paper, the classical BiLSTM-CRF deep learning model and the domestic large language models ERNIE-Bot and Spark will be employed to perform geometric entity extraction tasks from planar geometry proposition texts [19]. The experimental results will be compared to provide data support for further formalization of planar geometry proposition texts [20, 21].

2.2. Formal Definition of Planar Geometric Entities

This paper classifies the geometric entities potentially involved in planar geometry propositions into seven categories: points, lines, angles, triangles, quadrilaterals, pentagons, and circles. Each geometric entity can be represented by one or more symbols, with the formalized representation expressed as $\{x_1, x_2, \ldots, x_n\}$. Taking "angles" as an example, in addition to the basic "angle," there are specific types of angles as geometric entities, including right angles, obtuse angles, acute angles, straight angles, and exterior angles. These specific types of angles not only belong to the "angle" category but also carry distinct geometric meanings. For instance, an acute angle has a measure less than 90°, and a right angle has a measure equal to 90°. Consequently, the geometric entity "angle" requires further concrete definition. The specific descriptions are detailed in the Table 1 below.

Table 1: The formal representation corresponding to the geometric entity of an "angle".

Geometric Entity	Formal Representation	Example
Angle	$angle\{[x_1, x_2, x_3]\}$	$angle\{ABC\}, angle\{1\}$
Right Angle	$right\_angle\{[x_1, x_2, x_3]\}$	$right\_angle\{ABC\}$
Obtuse Angle	$obtuse\_angle\{[x_1, x_2, x_3]\}$	$obtuse\_angle\{A_1B_1C\}$
Acute Angle	$acute\_angle\{[x_1, x_2, x_3]\}$	$acute\_angle\{A'B'C'\}$
Straight Angle	$straight\_angle\{[x_1, x_2, x_3]\}$	$straight\_angle\{\beta\}$
Outside Angle	$outside\_angle\{[x_1, x_2, x_3]\}$	$outside\_angle\{C\}$
Inner Angle	$inner\_angle\{[x_1, x_2, x_3]\}$	$inner\_angle\{ACD\}$

2.3. Data preparation

2.3.1. Data pre-processing

A total of 420 pieces of Chinese plane geometry proposition text data were collected in this paper. Due to the presence of numerous geometry-related mathematical special symbols in the plane geometry propositions, such as "▱ABCD", "⊙O", "AB⊥CD", etc., these symbols may be lost or misrepresented in the compilation environment due to encoding issues. Therefore, pre-processing operations are required to convert these special symbols into standardized LaTeX expressions. The following Table 2 shows the possible mathematical special symbols in the plane geometry proposition text and their corresponding LaTeX representations, along with examples.

Table 2: Geometrical mathematical symbols and their corresponding LaTeX representations.

Mathematical symbols	LaTeX representations	Examples
△	$\triangle\{\}$	△ ABC → $\triangle\{$ ABC $\}$
∠	$\angle\{\}$	∠ A → $\angle\{$ A $\}$
⊙	$\odot\{\}$	⊙ O → $\odot\{$ O $\}$
▱	$\parallelogram\{\}$	▱ ABCD → $\parallelogram\{$ ABCD $\}$
⊥	$\{\}\bot\{\}$	AB ⊥ CD → $\{$ AB $\}\bot\{$ CD $\}$
≌	$\{\}\cong\{\}$	△ ABC ≌ △ DEF → $\{\triangle\{$ ABC $\}\}\cong\{\triangle\{$ DEF $\}\}$
∽	$\{\}\sim\{\}$	△ ABC ∽ △ DEF → $\{\triangle\{$ ABC $\}\}\sim\{\triangle\{$ DEF $\}\}$

2.3.2. Sequence labeling

After obtaining the proposition text pre-processed with LaTeX, the next step for the deep learning BiLSTM-CRF model is sequence labeling. In this paper, brat is used as the named entity recognition tool to label the types and index positions of different entities in each proposition text, generating the corresponding .ann files. The entity annotation files help determine the number of entities in the current text, their corresponding categories, and the index positions of each entity within the text. However, the .ann annotation files alone cannot achieve a one-to-one correspondence between the labeled information and the proposition text; they need to be combined with the proposition text for sequence labeling. Named entity sequence labeling methods include BIO, BMES, BIOES, and others [22]. This study adopts the BIO (Begin, Inside, Outside) three-tier labeling method, where "B-X" indicates that the current word is the beginning of the textual description of entity X, "I-X" indicates that the current word still belongs to the textual description of entity X, and "O" represents non-entity portions of the text.

3. Formalization of geometric entities

3.1. Based on the BiLSTM-CRF model

This paper employs the classical BiLSTM-CRF deep learning model [23, 24] to perform geometric entity recognition and formalization of Chinese plane geometry proposition texts. The overall architecture is shown in Figure 1. The input layer receives the proposition text, which is first pre-processed through LaTeX standardization to handle special characters specific to the geometry domain. Then, the entity annotation files of the proposition text are combined with the text data to perform a one-to-one BIO sequence labeling task, resulting in the input sequence. The core model consists of a BiLSTM (Bidirectional Long Short-Term Memory) network that captures the contextual features of the text, followed by a CRF (Conditional Random Field) layer to further optimize the global consistency of the sequence labeling and correct the entity annotations. Finally, a formalization module converts the sequence prediction results into a formal format (such as triangle{ABC}, line{AB}, angle{C}) to achieve the formal output of geometric entities in the geometry proposition text. BiLSTM (Bidirectional Long Short-Term Memory) is a bidirectional recurrent neural network composed of a forward LSTM and a backward LSTM. The LSTM unit is controlled by three gating mechanisms: the forget gate, input gate, and output gate, which manage the information selection at each time step.

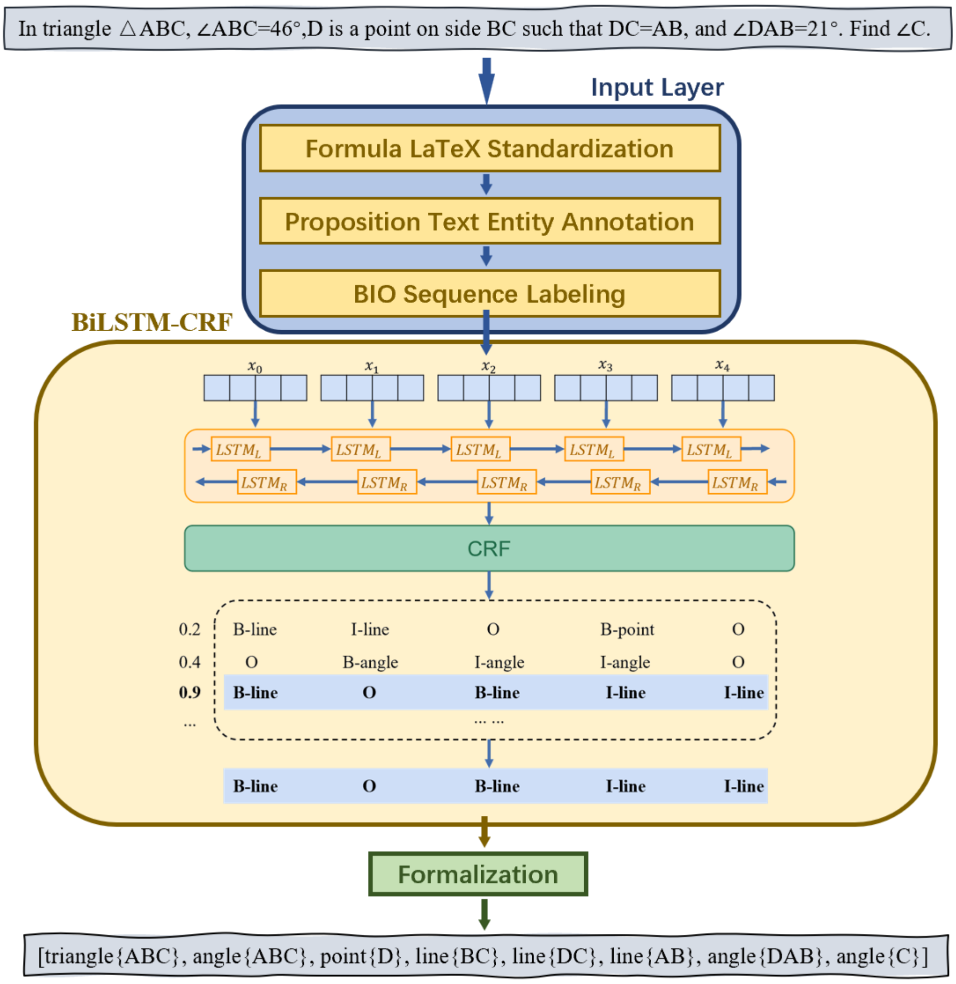

Figure 1: The overall architecture of geometric entity formalization based on BiLSTM-CRF

The input at time step t is processed through the σ function to obtain the output of the forget gate, denoted as f_t . The calculation formula is shown in Equation (1):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

Where W_f is the connection matrix, b_f is the bias vector, x_t is the input vector, and σ is the sigmoid activation function.

The input gate receives the current state of the unit along with the previous hidden layer state. The output is divided into two parts: one part consists of new information brought in at the current time step, which is normalized using the $tanh$ activation function to obtain the temporary state $\tilde{C}_t$, as shown in Equation (2). The other part is the output i_t, which is the result of applying the σ activation function to the input received by the input gate, as shown in Equation (3):

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

Where W_C is the connection matrix, h_{t-1} is the hidden layer state output from the previous time step, x_t is the input at the current time step, b_C is the bias vector, $tanh$ is the hyperbolic tangent function, which outputs values in the range of -1 to 1, W_i is the connection matrix, and b_i is the bias vector.

The output gate receives the input x_t from the previous time step and the hidden layer state h_{t-1} from the prior time step, which is then processed through the σ function to obtain the output o_t. The calculation formula is shown in Equation (4):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

Where W_o is the connection matrix and b_o is the bias vector.

Conditional Random Field (CRF) is a discriminative probabilistic model that is model based on the conditional probability $P(Y|X)$, where X is the input word sequence $(x_1, x_2, ..., x_n)$, and Y is the output label sequence $(y_1, y_2, ..., y_n)$. Assuming $P(Y|X)$ follows a linear-chain CRF, the output state sequence Y, given the input word sequence X with the value x, is a conditional random field under the input sequence X. The parameterized form is shown in Equation (5):

$$p(y|x) = \frac{1}{z(x)} exp(\sum_{i,k} \lambda_k t_k(y_{i-1}, y_i, x, i) + \sum_{i,l} u_l s_l(y_i, x, i)) \tag{5}$$

Where λ_k and u_l are the corresponding weights, and t_k and s_l are binary feature functions with values of 0 or 1.

During the BiLSTM training process, the model can only learn the correspondence between the input text sequence and the label sequence, but it cannot capture the dependency constraints between the sequence labels. As a result, invalid geometric entity predictions may occur. By adding a CRF layer, the model learns the contextual constraints between sequence labels, incorporates dependency conditions into the BiLSTM predictions, and corrects erroneous outputs, ensuring the validity of the model's output.

3.2. Based on large language models (LLMs)

Large language models (LLMs) represent a significant breakthrough in the field of natural language processing (NLP), with powerful language understanding and generation capabilities. Compared to traditional deep learning models, LLMs are more flexible and versatile, supporting direct input of natural language without the need for complex feature engineering or strict structural preprocessing [25]. This study utilizes Baidu's ERNIE Bot and iFLYTEK's Spark large language models to complete the geometric entity recognition task and formalization of plane geometry proposition texts. The overall architecture of the model is shown in Figure 2, which includes the input layer, large language model layer, formalization processing layer, and output layer.

The input layer consists of plane geometry proposition texts, such as " Given, in right triangle △ABC, ∠C=90°, the length of the median on side AB is 2, and AC+BC=6. Find the area of △ABC. " With the powerful contextual understanding capabilities of large language models, the input corpus does not require the extensive data preprocessing steps typical of traditional deep learning models, nor does it require the construction of one-to-one sequence labeling relationships. Instead, it can extract information and infer semantic relationships from unstructured data [26], thereby achieving the task objectives. The large language model processing layer compares the performance of the ERNIE Bot and Spark models. The formalization processing layer further processes the geometric entities recognized by the large language model into a standardized formal format, such as converting "Triangle ABC " into "right_triangle{ABC}". Finally, the output layer provides the formal representation of the geometric entities corresponding to the proposition text.

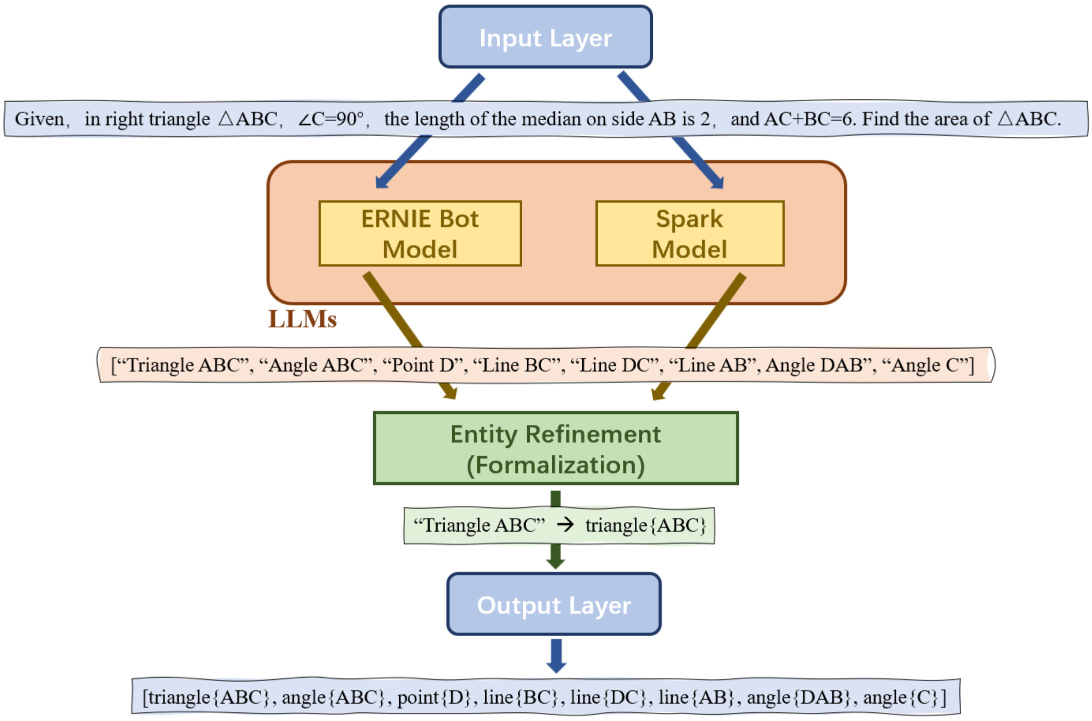

Figure 2: The overall architecture of geometric entity formalization based on large language models

4. Experimental Conclusions and Summary

4.1. Experimental Results

The experimental hardware environment: i5-11320H CPU, 16GB RAM, and a 1TB SSD. The experimental software environment: Windows 11, Python 3.9.13, and PyTorch 1.13.0. The key parameter settings for the BiLSTM-CRF model are shown in Table 3:

Table 3: Model Hyperparameter Settings.

Parameter	Setting Value	Detailed Description
EPOCH	50	Number of Iterations
LEARNING_RATE	1e-3	Learning Rate
EMBEDDING_DIM	100	Word Embedding Dimension
HIDDEN_DIM	256	LSTM unit Hidden Layer Dimension
TARGET_SIZE	41	Model Output Dimension
VOCAB_SIZE	183	Vocabulary Size
BATCH_SIZE	10	Batch Size
OPTIMIZER	Adam	Optimizer

The Basic Parameters of LLMs is shown in Table 4:

Table 4: Basic Parameters of LLMs.

	ERNIE Bot	Spark
Base Model Version	ERNIE-4.0-Turbo-8k	Spark4.0
Service Name	ERNIE-4.0-8k	Spark4.0 Ultra
Temperature	0.5	0.5
Response Format	JSON OBJECT	JSON OBJECT
Service Interface Type	HTTP	HTTP

The training loss function trend of the BiLSTM-CRF model, along with the performance curves of Accuracy and F1-score on the test set as training iterations increase, are shown in Figure 3:

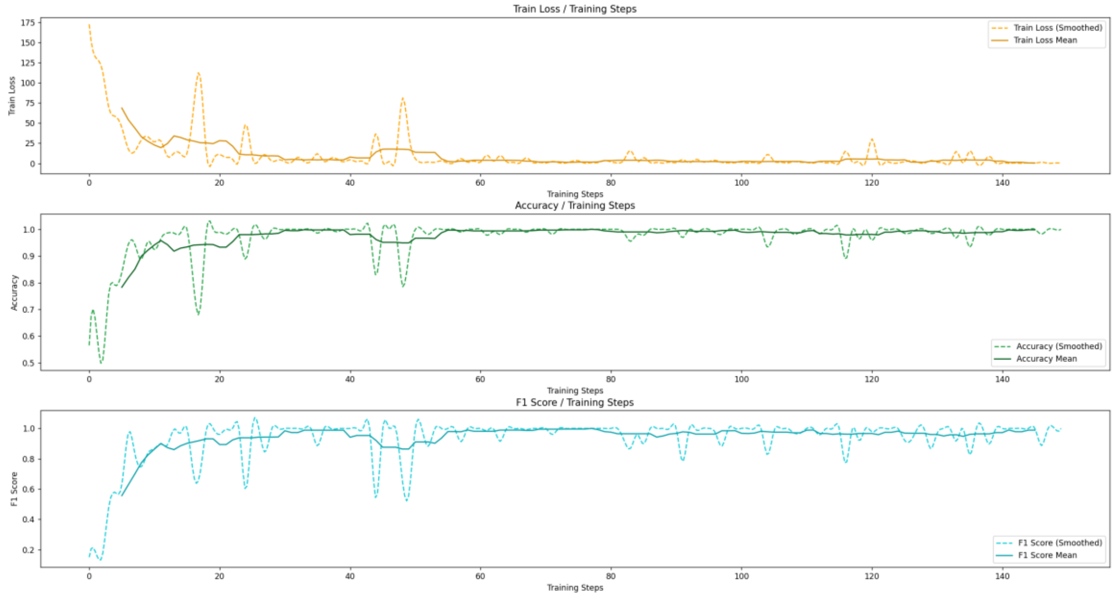

Figure 3: Training Performance of BiLSTM-CRF

Table 5 presents the performance of the BiLSTM-CRF model compared to the large language models ERNIE-Bot and Spark on the same test set, in terms of geometric entity recognition for Chinese plane geometry proposition texts:

Table 5: Experimental Results.

	Accuracy	F1-score
BiLSTM-CRF	0.98	0.97
ERNIE-4.0-Turbo-8k	0.99	/
Spark	0.96	/

4.2. Experimental Summary

This paper utilizes the BiLSTM-CRF model for the task of geometric entity recognition in Chinese plane geometry proposition texts. During the training process, the model's loss function gradually decreased and stabilized with the increasing number of training steps. Simultaneously, both Accuracy and F1-score showed a rapid upward trend and eventually stabilized, reflecting the model's continuous optimization and good convergence. On the test set, the model achieved high accuracy and an excellent F1-score, indicating its ability to efficiently capture the semantic features of geometric entities and perform accurate recognition.

When testing the same task using large language models ERNIE-Bot and Spark, ERNIE-Bot outperformed on the test set, achieving an accuracy of 99%. The errors were mainly in expressions like "Rt△ABC," where it failed to accurately recognize "right triangle ABC," instead identifying it as "triangle ABC." Additionally, "radius" in the phrase "radius of the circle" was also mistakenly recognized as a separate geometric entity. The Spark model performed slightly worse in this task, with errors such as repeated recognition of the same geometric entity, incorrect identification, and missing recognition of geometric entities. Compared to traditional neural network models, large language models can access larger and more complex training datasets, providing the model with stronger text processing and understanding capabilities. This advantage has led to a significant improvement in performance for geometric entity recognition tasks.

In the context of this paper, the domestic large language model ERNIE-Bot demonstrated exceptional geometric entity recognition abilities. Its performance further validates the potential and advantages of large language models in specialized domain tasks, offering new technological directions and breakthroughs for geometric entity recognition and, more broadly, for natural language processing tasks.

5. Conclusions

Based on the experimental results, it can be observed that while traditional neural network models have lower training and inference costs compared to large language models, providing certain resource advantages, they are limited by the size and quality of the training datasets. In comparison, large language models exhibit greater flexibility and generalization capabilities. In this experiment, the geometric entity recognition task for planar geometry proposition texts was performed with a training dataset of approximately 420 examples and a test dataset of about 200 examples. Moreover, given the high diversity and complexity of natural language descriptions of planar geometry propositions, especially in the Chinese context, where the expression forms of propositions are highly variable, the issue of insufficient dataset size becomes more prominent. For complex propositions or expressions that have not appeared in the training set, traditional neural network models still struggle to process them effectively, which is a significant cause of accuracy loss.

For large language models, their superior text comprehension and processing abilities offer a marked advantage. However, this high performance comes with extremely high hardware and software requirements. As a result, the use of large language models primarily depends on commercial computational services provided by enterprises, and thus incurs certain operational costs. Nonetheless, given the powerful computational capabilities and outstanding performance of large language models, they are invaluable for specific tasks, especially those involving highly complex or diverse text processing needs. In practical applications, a balance can be struck between traditional neural network models and large language models based on the specific task requirements and available resources, thereby optimizing the trade-off between cost and performance.

In the future, improving the performance of the geometric entity recognition task can be approached from several angles. First, expanding the training dataset size and using techniques such as data augmentation to improve the diversity and quality of the dataset will enhance the model's ability to adapt to unseen expression forms. Second, exploring the integration of small models with large models could leverage the pre-training capabilities of large language models and the cost-efficiency of traditional models, achieving a balance between performance and efficiency. Third, by studying the specific characteristics of Chinese planar geometry proposition texts, more tailored feature extraction and model optimization strategies can be designed to enhance the overall performance of the task.

References

[1] ZHANG Jinbao, JI Lingyan. (2018). Artificial Intelligence Enhanced Education or Education for Development of Intelligence——An Analysis on Connotation and Target of Intelligent Education in the Age of Artificial Intelligence. Modern Distance Education Research (02), 14-23.

[2] Chen, W., Yin, M., Ku, M., Lu, P., Wan, Y., Ma, X., ... & Xia, T. (2023, December). Theoremqa: A theorem-driven question answering dataset. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 7889-7901).

[3] Liu, W., Hu, H., Zhou, J., Ding, Y., Li, J., Zeng, J., ... & He, L. (2023). Mathematical language models: A survey. arXiv preprint arXiv:2312.07622.

[4] Chen, J., Tang, J., Qin, J., Liang, X., Liu, L., Xing, E. P., & Lin, L. (2021). GeoQA: A geometric question answering benchmark towards multimodal numerical reasoning. arXiv preprint arXiv:2105.14517.
[5] GAN Wenbin. (2018). Research on Automatically Solving Plane Geometry Problems (DOCTORAL DISSERTATION, Central China Normal University). Doctor https://kns.cnki.net/kcms2/article/abstract?v=x0sJsLzXXbifstQ4abWx-ScKx680f7nGI4SknzZ1b77oYizZSu6djQB2unmRDyZ1Blilue9fFj172EN0ftOI8z-Jr3z4yEyB3xVgdT0hXwgt8lgm0IQvxvcciqxI_szmpqGpcoNZ4aryBp8Pnz9hlnrARuCbiQTdi5AL9IHyMsSxO6d10vvwiyViFAcWrHij0tD0Lm2r-9M=&uniplatform=NZKPT&language=CHS
[6] Lample, G., & Charton, F. (2019). Deep learning for symbolic mathematics. arXiv preprint arXiv:1912.01412.
[7] Ning, M., Wang, Q. F., Huang, K., & Huang, X. (2023, October). A symbolic characters aware model for solving geometry problems. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 7767-7775).
[8] Jingzhong Zhang, & Yongbin Li. (2009). Thirty years of machine proof of geometric theorems. Journal of Systems Science and Mathematical Sciences, 29(9), 1155.
[9] Xia, R., Li, M., Ye, H., Wu, W., Zhou, H., Yuan, J., ... & Zhang, B. (2024). GeoX: Geometric Problem Solving Through Unified Formalized Vision-Language Pre-training. arXiv preprint arXiv:2412.11863.
[10] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences, 103, 102274.
[11] Lafferty, J., McCallum, A., & Pereira, F. (2001, June). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In Icml (Vol. 1, No. 2, p. 3).
[12] Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.
[13] Kenton, J. D. M. W. C., & Toutanova, L. K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of naacL-HLT (Vol. 1, No. 2).
[14] Joshi, M., Chen, D., Liu, Y., Weld, D. S., Zettlemoyer, L., & Levy, O. (2020). Spanbert: Improving pre-training by representing and predicting spans. Transactions of the association for computational linguistics, 8, 64-77.
[15] He, P., Liu, X., Gao, J., & Chen, W. (2020). Deberta: Decoding-enhanced bert with disentangled attention. arXiv preprint arXiv:2006.03654.
[16] Koroteev, M. V. (2021). BERT: a review of applications in natural language processing and understanding. arXiv preprint arXiv:2103.11943.
[17] Zhu, W., & Cheung, D. (2021). Lex-bert: Enhancing bert based ner with lexicons. arXiv preprint arXiv:2101.00396.
[18] Welsby, P., & Cheung, B. M. (2023). ChatGPT. Postgraduate Medical Journal, 99(1176), 1047-1048.
[19] Cosentino, R., & Shekkizhar, S. (2024). Reasoning in large language models: A geometric perspective. arXiv preprint arXiv:2407.02678.
[20] ZHANG Yu. (2018). Design and creation of a formalized geometric knowledge base. Information Technology (05), 139-144. doi:10.13274/j.cnki.hdzj.2018.05.031.
[21] Fu Fan. (2019). The Formal Proof of Geometric Theorem in Coq Case Study: The Formal Proof of Yanglu Theorem (MASTER DISSERTATION, Beijing University of Posts and Telecommunications). Master https://kns.cnki.net/kcms2/article/abstract?v=x0sJsLzXXbhUTm9QAaVaWw0sxdbtua_YJ9gcAi8jIlaLNl3yqiyusHcIUMmCTdFwI5NQAHM2rdOcrTLIfx4jl8ZNaxl_58h6iqFYAlzXRh16GVwHzA49UUasLg_QJccHbT7y2V4rB_goSEOpCRVEoQ3ZtEz5ibSWBGy5Ze3udB5eeFaC6K__cF4znQ2l2FqCuPX-usCsPwU=&uniplatform=NZKPT&language=CHS
[22] Lignos, C., Kruse, M., & Rueda, A. (2023, December). Improving NER Research Workflows with SeqScore. In Proceedings of the 3rd Workshop for Natural Language Processing Open Source Software (NLP-OSS 2023) (pp. 147-152).
[23] Qiu, Q., Xie, Z., Wu, L., Tao, L., & Li, W. (2019). BiLSTM-CRF for geological named entity recognition from the geoscience literature. Earth Science Informatics, 12, 565-579.
[24] SHAO Jingze*, LI Linna &NIE Yaoxin. (2022). A Chinese entity extraction method based on BILSTM_CRF. China CIO News (10), 124-127.*
[25] Ji, Y., & Gao, S. (2023). Evaluating the effectiveness of large language models in representing textual descriptions of geometry and spatial relations. arXiv preprint arXiv:2307.03678.
[26] Gao, J., Pi, R., Zhang, J., Ye, J., Zhong, W., Wang, Y., ... & Kong, L. (2023). G-llava: Solving geometric problem with multi-modal large language model. arXiv preprint arXiv:2312.11370.