

Article

Multi-Feature Fusion Method for Chinese Shipping Companies Credit Named Entity Recognition

Lin He, Shengnan Wang * and Xinran Cao

School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China

* Correspondence: 13009467815@163.com; Tel.: +86-130-0946-7815

Abstract: Shipping Enterprise Credit Named Entity Recognition (NER) aims to recognize shipping enterprise credit entities from unstructured shipping enterprise credit texts. Aiming at the problem of low entity recognition rate caused by complex and diverse entities and nesting phenomenon in the field of shipping enterprise credit, a deep learning method based on multi-feature fusion is proposed to improve the recognition effect of shipping enterprise credit entities. In this study, the shipping enterprise credit dataset is manually labeled using the BIO labeling model, combining the pre-trained model Bidirectional Encoder Representations from Transformers (BERT) and bidirectional gated recurrent unit (BiGRU) with conditional random field (CRF) to form the BERT-BiGRU-CRF model, and changing the input of the model from a single feature vector to a multi-feature vector (MF) after stitching character vector features, word vector features, word length features, and part-of-speech (pos) features; BiGRU is introduced to extract the contextual features of shipping enterprise credit texts. Finally, CRF completes the sequence annotation task. According to the experimental results, using the BERT-MF-BiGRU-CRF model for NER of shipping enterprise credit text data, the F1 Score (F1) reaches 91.7%, which is 8.37% higher than the traditional BERT-BiGRU-CRF model. The experimental results show that the BERT-MF-BiGRU-CRF model can effectively perform NER for shipping enterprise credit text data, which is helpful to construct a credit knowledge graph for shipping enterprises, while the research results can provide references for complex entities and nested entities recognition in other fields.



Citation: He, L.; Wang, S.; Cao, X. Multi-Feature Fusion Method for Chinese Shipping Companies Credit Named Entity Recognition. *Appl. Sci.* **2023**, *13*, 5787. <https://doi.org/10.3390/app13095787>

Academic Editors: Yu-Dong Zhang and Douglas O'Shaughnessy

Received: 6 March 2023

Revised: 1 May 2023

Accepted: 3 May 2023

Published: 8 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: shipping enterprise credit; named entity recognition; multi-features; bidirectional gated recurrent unit network; conditional random field

1. Introduction

The shipping industry is an important pillar of China's economic development and an important guarantee for the development of international trade [1]. Since the 18th Party Congress, General Secretary Xi Jinping has repeatedly given important instructions on the development of the shipping industry, emphasizing that "a strong economic country must be a strong maritime country and a strong shipping country" when linking to Shanghai Yangshan Port on 6 November 2018, and on 17 January 2019, when inspecting Tianjin Port, that "for the economy to develop and the country to be strong, transportation, especially shipping, must first be strong." Since the 19th Party Congress, General Secretary Xi Jinping has repeatedly given important instructions on the construction of a social credit system, requiring the institutionalization of integrity building, making the construction of a credit system an important cornerstone of healthy economic development and emphasizing the need to build a new credit-based regulatory mechanism. To carry out credit assessment and management in the field of shipping, it is of great significance to further promote the construction of a transportation credit system, maintain the market order of fair competition in the field of shipping, and promote the progress of China from a large shipping country to a strong shipping country [2].

Accurate and effective credit evaluation for shipping enterprises can promote the healthy development of shipping enterprises. The existing shipping enterprise credit assessment model overly emphasizes the financial credit status of shipping market subjects, and the mining and analysis of unstructured data such as government regulatory information and judicial judgment information are relatively few, which cannot objectively reflect the credit status. The inability to efficiently and accurately mine the unstructured information in the credit field of shipping enterprises makes the current credit assessment in the shipping field generally suffer from incomplete credit information and inaccurate assessment results [3]. A knowledge graph is essentially a semantic network that reveals the relationship between entities, which is one of the main forms of knowledge representation in the era of big data and also an important basic resource for artificial intelligence applications [4]. Facing the disorderly shipping enterprise credit information existing on the Internet, by constructing the shipping enterprise credit information knowledge graph, the key shipping credit information can be extracted from the massive shipping big data, and a comprehensive and accurate shipping credit semantic relationship can be formed by processing the data. Constructing a shipping enterprise credit knowledge graph and conducting shipping enterprise-related credit analysis based on the graph is beneficial for financial institutions to conduct comprehensive and accurate credit assessments of shipping enterprises. One of the most important parts of the knowledge graph construction process is NER tasks, and the accuracy of NER determines the quality of relationship extraction and knowledge graph, which will affect the effect of downstream tasks. Therefore, in view of the characteristics of shipping enterprise credit entities, this paper proposes NER for shipping enterprise credit domain texts, which lays the foundation for the subsequent knowledge graph construction and provides more comprehensive data support for shipping enterprise credit-related risk analysis and shipping enterprise credit evaluation.

NER research in the field of shipping corporate credit is still in its infancy due to the lack of publicly available datasets. The data in this field have the following characteristics:

1. Entity names are long and complex in form. The name of a shipping enterprise credit entity consists of a word, phrase, alphabet, or symbol. For example, “CFLP-AAA Logistics Enterprise” is composed of “C”, “F”, “L”, “P”, “-”, “AAA”, “Logistics”, “Enterprise”, “CFLP-AAA”, and “CFLP-AAA”. “P”, “-”, “AAA”, “logistics”, and “enterprise” are combination of letters, words, and symbols. In the process of NER, it is difficult to accurately identify the entity if only a single feature vector is used as input for the model.
2. Entity nesting exists. For example, the company entity “Anhui Eddyang County Shipping Company” contains two entities: “Anhui Province” and “Eddyang County”, so it is very easy to make a mistake in the process of entity recognition.

Therefore, according to the textual characteristics of the shipping enterprise credit domain, this paper proposes the BERT-MF-BiGRU-CRF model for the NER task. The multi-features are integrated with the contextual features obtained from BiGRU and introduced into the CRF layer for classification. Compared with the model with a single feature, the model proposed in this paper can make full use of shipping enterprise credit text information for NER.

The main contributions of the paper are as follows: first, the crawled shipping enterprise credit data are pre-processed, and the data are annotated by BIO. Next, multi-feature extraction is performed, and character-level features are extracted by the BERT model; the shipping enterprise credit word list is obtained by jieba word separation and removal of repeated words, and no additional user dictionaries are introduced in the word separation process. The words that all characters in the sentence may form are identified according to the generated word list, and then the characters are added to the four word sets B, M, E, and S according to their positions in the words, and when there is no matching word in the set, it is filled with the marker NULL. The four set representations corresponding to each character in the text sequence are counted, and the vectors of the four sets are stitched as the word vector representation corresponding to the character; the pos and

word length of the words in the sets are counted, and the pos vector representation and the word length vector representation corresponding to each set are calculated by weighted average, and the word vector feature, pos feature, and word length feature are obtained, respectively. The extracted four types of features are stitched as the input of BiGRU, and the contextual features are extracted in BiGRU to obtain the text-rich semantic features. Finally, the sequence annotation is completed by CRF to fully explore the hidden information in the text labels and further improve the accuracy of shipping enterprise credit NER. The shipping enterprise credit entities recognition in this paper can also be used for relationship extraction to further lay the foundation for building a knowledge graph of shipping enterprise credit information.

This paper is structured as follows: in Section 1, the background of credit entity recognition of shipping companies, data characteristics, and the contribution of this paper and the structure of the paper are briefly described. In Section 2, the current status of research on named entity recognition is described. In Section 3, the research content, the relevant theories, and methods used in this study are described. In Section 4, the data are pre-processed and annotated, and three types of experiments are conducted, which are the experimental comparison with existing research solutions, the experimental comparison including different feature models, and the experimental fine-grained entity experiment for shipping enterprise credit, and the following conclusion is drawn: the deep learning model based on multi-feature fusion proposed in this paper has the best NER effect on shipping enterprise credit. In Section 5, the main content of this paper is summarized, and the outlook of future research is outlined.

2. Related Work

NER is an important step in building a knowledge graph and is the basic work of information extraction, which lays the foundation for entity alignment, relationship extraction, etc. [5]. NER methods are divided into three main categories: rule dictionary [6]-based methods, machine learning-based [7] methods, and deep learning-based methods.

Most of the early NER used rule-based approaches, mainly by manually constructing limited rules to find strings in the text that match the rules. The rule-based approach has high accuracy in specific domains and is suitable for domains with small data sets that do not require frequent updates, but the approach requires a more comprehensive set of rules to be constructed manually, which has the problems of time-consuming manual operation and poor portability. With the rise of machine learning, the NER methods started to rely on various statistical models, marking a small number of samples by manually setting features and then training the model. Therefore, this method has good portability. It mainly includes Hidden Markov Model (HMM) [8], CRF [9], and Support Vector Machine (SVM) [10] for entity recognition. HMM is an indirect observation of the state of Markov chains through a sequence of observation vectors, which generates the corresponding observation vectors from each probability density distribution of the state sequence and contains two basic assumptions. Since the model uses the Viterbi algorithm, the training speed is relatively fast. However, its output relies on the independence assumption, which leads to the inability to consider long-range dependence and cannot truly describe the information of the office letter in the sequence, which limits the selection of features to a certain extent and can only achieve local optimality and is more suitable for short textual entity recognition. The principle of SVM is to map the input vector x into a high-dimensional feature space by means of a pre-selected nonlinear mapping and construct the optimal classification hyperplane to separate the two classes of samples without errors and to maximize the classification gap between the two classes. Although SVM can handle high-dimensional features, the training time is relatively long. CRF is a conditional probability distribution model for another set of output sequences given a set of input sequence conditions, with no strict independence assumption conditions; it can accommodate arbitrary contextual information and has a flexible feature design. Among the traditional machine learning methods, CRF is seen as the dominant model for NER, with the advantages that in the

process of labeling each location, CRF can output the global optimal sequence using internal and contextual feature information, and its high accuracy in predicting entity types makes it a commonly used machine learning-based method for NER.

Dan Wang et al. [11] proposed to use HMM for initial entity recognition using lexical properties as observations for the semantic incompleteness feature of short texts and to construct a phonetic cognate relationship database to identify potential entities based on the recognition results. For example, Pan Huashan et al. [12] used the SVM approach for Vietnamese news text and selected words and lexicality as features for entity recognition at syntactic and semantic levels, respectively, to build a Vietnamese news text classification model. Mena et al. [13] proposed a method based on a combination of HMM and SVM for place name recognition. Firstly, the labeled place name training set is used to train HMM for place name recognition, and then the recognized place names are matched with GeoNames' place names list, and the informative and relevance features of the place names are calculated and input to SVM to get the final place name recognition results. Kai Liu et al. [14] implemented NER in TCM electronic medical records based on the CRF model, which introduced contextual indicator words, lexicon, and other features and achieved 80% F1 value for symptoms. Shi-kun, Wang et al. [15] used a CRF approach for NER using TCM medical cases, and the results showed that the CRF approach was significantly better than the maximum entropy and SVM approaches. Hong-ping Hu [16] et al. used CRF as the Chinese NER model, comparing two different levels of models based on character level and word level.

Traditional machine learning methods convert NER tasks into sequence annotation problems, requiring a large amount of feature engineering. Due to the limitations of manually designed features, when the dataset is large, it is easy to generate over-fitting problems. The NER tasks in the field of shipping enterprise credit belong to domain-specific NER, and many long entities and complex entities exist in the shipping enterprise credit dataset, which require a large amount of domain knowledge as well as the need to manually design and extract a large number of features using the traditional machine learning approach, increasing the labor cost and time cost.

In recent years, with the rise of deep learning, scholars have started to apply deep learning to NER tasks. The NER method based on deep learning can automatically extract data features and complete classification and train the model in an end-to-end manner. This approach does not require a lot of time and labor costs to design and extract features and can automatically extract features for each word or phrase. In addition, since deep learning is a data-driven approach, the larger the amount of data, the more accurately the model is trained and the better the entity recognition. Therefore, the use of a deep learning-based approach can effectively solve the problems of credit entity recognition of shipping companies with traditional machine learning.

Huang et al. [17] systematically compared the performance of multiple neural network models and proposed the Bidirectional Long Short-Term Memory (BiLSTM) and CRF model for three tasks of sequence annotation, pos annotation, chunking, and NER experiments on three different datasets, respectively, and the BiLSTM-CRF model had the best or near-best accuracy compared to the baseline model. Ma and Hovy [18] combined BiLSTM, CNN, and CRF to propose a novel neural network architecture that utilizes both word- and character-level representations without feature engineering or data preprocessing and is therefore suitable for a wide range of sequence labeling tasks, and the model achieved the best performance in experiments on both datasets. Xiang Jiang [19] et al. proposed a BiLSTM-IDCNN-CRF model of word embedding for an ecological governance technology domain dataset, which is obtained by a BiLSTM network and IDCN network to different granularity of features, and the experimental results verified the effectiveness of the word embedding method. According to the characteristics of data in the field of finance and taxation, Yu Qiu [20] et al. proposed a BiLSTM-CRF model combining character features and word vector features to identify entity boundaries, using an integrated learning approach for entity type labeling and entity class labeling as a multi-label multi-category classification

task, and the experimental results show that the combined word vector approach had higher recognition performance for entities of longer length and rare entities. De-Ping Chu [21] et al. proposed a BiLSTM-CRF model based on pre-trained word vectors to compensate for the lack of word vector specificity by adding word dynamic features and word character level features to improve the recognition level of complex multi-word meanings in geological texts and the extraction ability of local features of geological entities, and the experimental results proved to be better in the recognition of geological entities in small-scale corpus. Le-jun Gong [22] et al. proposed a two-level annotation model based on the combination of domain dictionary and CRF to recognize four types of entities: disease, symptom, drug, and operation. However, the word vector generated by Word2Vec is static and cannot solve the problem of polysemy. To solve this problem, BERT [23] has been widely used in NER tasks. Lin [24] investigated the NER tasks in the Chinese medicine domain based on various deep learning models and selected BERT models and traditional Bidirectional BiLSTM models to recognize symptomatic entities, and the results show that character-based BERT models were better for Chinese medicine entity recognition. Peng-fei Zhao [25] et al. proposed a NER model of BiLSTM-CRF fusing BERT word-level features with external lexicon features for the problem of low recognition rate of rare entities in the agricultural entity recognition process and conducted comparison tests with BiLSTM-CRF, CNN-BiLSTM-CRF, and BERT-BiLSTM-CRF models, and the results show that the method outperformed the other three mainstream models.

Compared with the general text data, the shipping enterprise credit text contains a large number of professional words related to intellectual property rights, license certificates, adjudication documents, administrative penalties, and so on, and there are a large number of long entities and complex entities, with the phenomena of remote dependency, nested entities, and multiple meanings of words. Although some existing models have better recognition effect, the models are complex and prone to errors on the boundary of corporate credit entity recognition. Although some models use feature combination for NER, the granularity of extracted entities is relatively coarse for long and complex entities in shipping enterprise credit text, and the semantics of the text are not rich enough. In order to obtain a richer semantic representation of this text, more fine-grained features of entities are extracted in the shipping enterprise credit NER task.

Compared with traditional deep learning models, the BERT pre-training model is based on the bidirectional transformer encoder, which has more powerful language representation and feature extraction capabilities. In the shipping enterprise credit NER task, the BERT model is used to extract character-level features to enrich the semantic representation of the text and alleviate the problem of multiple meanings of words.

Therefore, in this paper, based on the BERT-BiGRU-CRF deep learning model, we innovatively propose a method to incorporate multiple text features in the model for the NER task based on the text characteristics of the shipping enterprise credit domain.

3. Method

For the textual characteristics of shipping enterprise credit domain, this paper proposes the NER model BERT-MF-BiGRU-CRF that fuses multiple feature vectors. Firstly, the text data in the field of shipping enterprise credit is manually annotated; the multi-features of the text are extracted; the pre-trained character features are spliced with word vector features, pos features, and word length features and input to the BiGRU layer to extract the contextual features of the text; and finally, the output is performed by the CRF to obtain the optimal prediction label sequence of shipping enterprise credit. The NER process is shown in Figure 1. The BERT-MF-BiGRU-CRF model adopts a hierarchical structure with bottom-up layers: input layer, feature fusion layer, BiGRU layer, CRF layer, and output layer. For the input sequence $S = \{c_1, c_2, c_3, \dots, c_n\}$, c_i generates the word vector x^b with rich information through the BERT layer; x^b is spliced with the word vector feature x^v , word length feature x^l , and pos feature x^p to get the vector x_i , and then x_i is input to the

BiGRU layer for contextual information learning; and finally, the global output is output through the CRF layer. The optimal sequence is finally output through the CRF layer.

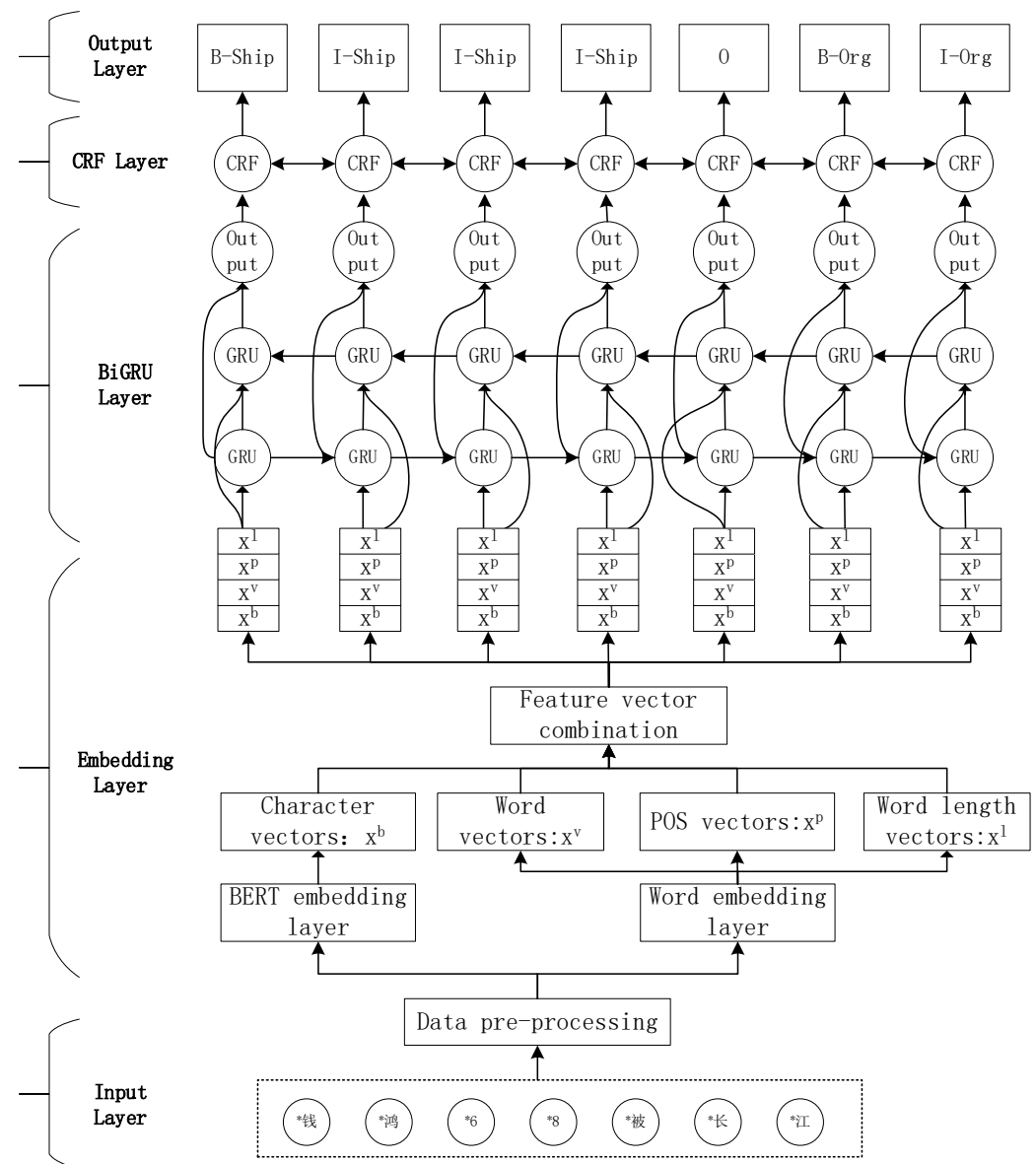


Figure 1. BERT-BiGRU-CRF model structure diagram. * Notes: The English translation of the input layer is “Qian Hong 68 by Yangtze River”.

3.1. Embedding Layer

The model performs word-level feature extraction at the embedding layer through the BERT model, converts the features into a dense vector of fixed dimensions, and uses stitching to fuse word-level features as the input vector for the encoding layer.

3.1.1. Character Level Features

In order to obtain rich word-level semantic representations, this paper uses a BERT pre-training model to obtain word-level feature vector representations of the corpus. The BERT pre-training model is based on the bidirectional transformer encoder, which is used to obtain word-level feature representation by masking the language model and next sentence prediction model to learn the sentence-level semantic relationship of text sequences for better extraction of text feature information [25].

During BERT pre-training, the sequence $S = \{c_1, c_2, c_3, \dots, c_n\}$ corresponds to the input $E = \{E_1, E_2, E_3, \dots, E_n\}$, as shown in Figure 2, which is superimposed by three embedding features, namely character embedding e_i^t , sentence embedding e_i^s , and position embedding e_i^p , where $E_i = e_i^t + e_i^s + e_i^p$ and e_i denotes the embedding feature of each character. Two special markers [CLS] and [SEP] are added to the initial position of sequence X and between sentences, [CLS] identifies the start position of the sequence and [SEP] identifies the split between sentences. After multiple bidirectional transformer encoders, the vector E obtains the vector $X^b = \{x_1, x_2, x_3, \dots, x_n\}$, which contains rich semantics, where $X \in \mathbb{R}^{n \times h_{bert}}$, h_{bert} denotes the hidden state dimension of BERT, and x_i^b denotes the BERT vector of the i -th character in the sequence.

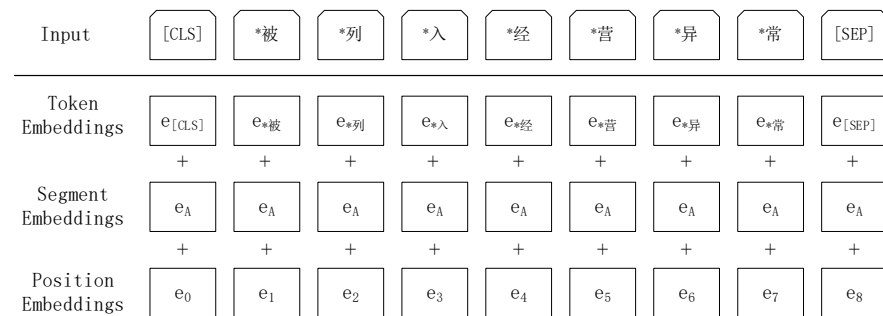


Figure 2. BERT model structure. * Notes: The English translation of the input layer is “Listed as business exception”.

3.1.2. Word Level Features

The further incorporation of word-level features in the character-level feature-based model helps to improve the recognition effect of the model. In this paper, three word-level features are selected, which are word vector features, word length features, and pos features. Because word-level feature extraction requires the use of word lists, the shipping enterprise credit word list in this paper is obtained by jieba word separation and removal of duplicate words, and no additional user dictionaries are introduced in the word separation process.

In order to avoid the influence of word separation errors on extracting word-level features, this paper introduces four sets B, M, E, and S to save the different words corresponding to the characters. Firstly, all the words that the characters in the sentence may form are identified according to the generated word list, and then the characters are added to the four sets of B, M, E, and S according to their positions in the words, with B representing characters appearing at the beginning of the words, M representing characters appearing in the middle of the words, E representing characters appearing at the end of the words, and S representing characters that can be seen as words themselves. When there is no matching word in the set, it is filled with the token NULL. The set of four words corresponding to each character can be expressed by the following Equations (1)–(4):

$$B(c_i) = \{w_{i,k}, i < k \leq n\} \quad (1)$$

$$M(c_i) = \{w_{j,k}, 1 \leq j < i < k \leq n\} \quad (2)$$

$$E(c_i) = \{w_{j,i}, 1 \leq j < i\} \quad (3)$$

$$S(c_i) = \{c_i\} \quad (4)$$

(1) Word Vectors Features

Word features are the features of words themselves, which are the most basic features in the field of natural language processing and are represented in the form of word vectors. The word vectors in this study are the vectors after word reorganization based on characters

after stuttering and splitting. The word vectors in this study are embedded directly through the embedding layer.

The four set representations corresponding to each character in the text sequence are counted, and the set vector representation corresponding to each set is calculated by weighted averaging, and the vectors of the four sets are stitched together as the word vector representation corresponding to that character, as shown in Equation (5).

$$x_i^v = v_{B(c_i)} \oplus v_{M(c_i)} \oplus v_{E(c_i)} \oplus v_{S(c_i)} \quad (5)$$

where x_i^v denotes the word vector corresponding to character i and $v_{B(c_i)}$ denotes the vector corresponding to the set of words B .

(2) Pos Features

Chinese pos characteristics include 12 categories such as nouns, verbs, prepositions, and quantifiers. In the texts related to shipping company entities and risk penalty entities or honor entities, verbs such as “obtain” and “subject to” often appear, and the pos characteristics can effectively assist in the recognition of company entities and honor entities as well as risk penalty entities in shipping enterprise credit texts. The recognition of corporate entities and honorary entities and risk penalties in shipping enterprise credit texts can be effectively assisted according to pos features.

Count the four set representations corresponding to each character in the text sequence, label all words using the pos annotation system after stuttering and splitting, calculate the pos vector representation corresponding to each set by weighted averaging, and stitch the lexical vectors of the four sets as the pos vector representation of the word corresponding to that character.

$$x_i^p = P_{B(c_i)} \oplus P_{M(c_i)} \oplus P_{E(c_i)} \oplus P_{S(c_i)} \quad (6)$$

where x_i^p denotes the pos vector of the word corresponding to character i and $P_{B(c_i)}$ denotes the pos vector corresponding to the set of words B .

(3) Word Length Features

The names of the entities in the credit field of shipping enterprises are relatively long, such as “Business Exception List of Bad Credit”, “Executtee of Bad Credit”, and “CFLP-AAA Logistics Enterprise Title”. Therefore, this paper uses the word length as the feature to identify the entities.

Similar to word vector feature and pos vector feature extraction, first count the four sets of representations corresponding to each character in the text sequence, count the word lengths of the words in the sets, calculate the word length vector representation corresponding to each set by weighted averaging, and stitch the word length vectors of the four sets as the word length vector representation of the word corresponding to the character.

$$x_i^l = L_{B(c_i)} \oplus L_{M(c_i)} \oplus L_{E(c_i)} \oplus L_{S(c_i)} \quad (7)$$

where x_i^l denotes the word length vector of the word corresponding to character i and $L_{B(c_i)}$ denotes the word length vector corresponding to the set of words B .

$$x_w = x^v \oplus x^p \oplus x^l \quad (8)$$

3.1.3. Embedding Layer Vector Splicing

The word-level features, including word vector features, word length features, and pos features, are obtained by BMES word reorganization based on stuttering subscripts, and these three word-level features are stitched with the character-level features obtained by BERT to obtain a fixed-length vector, which is used as the input to the coding layer BiGRU. The spliced vector can be represented by Equation (9):

$$x = x^v \oplus x^p \oplus x^l \oplus x^b \quad (9)$$

3.2. BiGRU Layer

Gated Recurrent Unit (GRU) is a special recurrent neural network developed on the basis of Long Short Term Memory (LSTM), which has a simpler structure and solves the problem of gradient disappearance and gradient explosion in LSTM and can better capture semantic dependencies and achieve long-term memory. Structurally, GRU combines the forgetting gate and the input gate into an update gate, so there are only update and reset gates and a mixture of cell states and hidden states, so GRU has fewer training parameters, which can speed up the training and improve the model performance [26].

The structure of the GRU is shown in Figure 3, where r_t is a reset gate to control whether the current content is remembered, and z_t is an update gate to control the proportion of preorder memory and candidate memory in the current state [27].

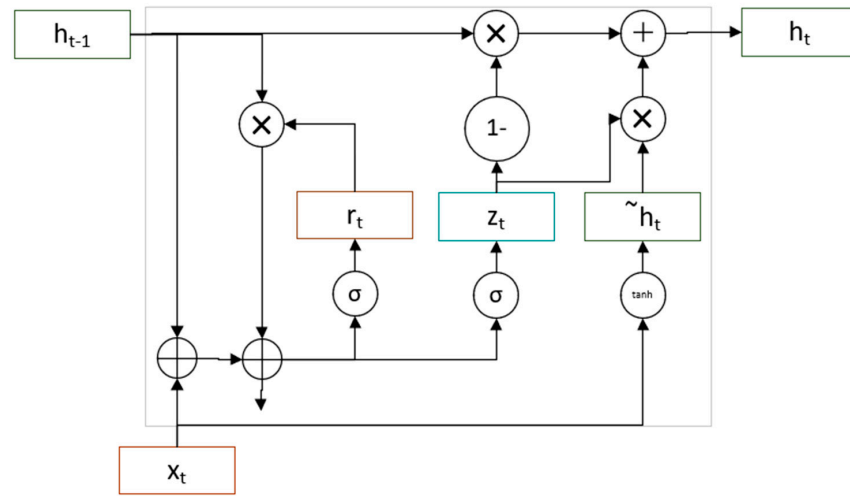


Figure 3. Structure of GRU.

The GRU network forward propagation weight parameter update formula [28] is:

$$z_t = \sigma(w_{xz}x_t + w_{hz}h_{t-1} + b_z) \quad (10)$$

$$r_t = \sigma(w_{xr}x_t + w_{hr}h_{t-1} + b_r) \quad (11)$$

$$\tilde{h}_t = \tanh(w_{xh}x_t + r_t \otimes w_{hh}h_{t-1}) \quad (12)$$

$$h_t = z_t \otimes h_{t-1} + (1 - z_t) \otimes \tilde{h}_t \quad (13)$$

where x_t denotes the input at moment t ; h_{t-1} denotes the input of the hidden state at moment $t - 1$; \tilde{h}_t denotes the candidate hidden state; h_t denotes the hidden state, which is the output of the GRU unit; \oplus denotes vector splicing; \otimes denotes Hadamard product, i.e., multiplying by the corresponding elements; σ is the sigmoid function; w is the weighting factor; b is the offset term.

The hidden state h_t of GRU can only obtain information from the past, but in the process of NER, the context of the text is important for predicting the labels, so this paper adopts BiGRU as part of the NER model. The forward GRU captures the above information of the text, the reverse GRU extracts the below information of the text, and the output of the network is obtained by stitching the output of the forward and reverse GRUs, which fully extracts the contextual information of the text and can effectively improve the effect of NER. The output of the decoding layer BiGRU can be expressed in Equation (14).

$$h_i = \left[\vec{h}_i \oplus \overleftarrow{h}_i \right] \quad (14)$$

3.3. CRF Layer

Two common methods of sequence labeling are direct classification using softmax function and CRF model. Although the former can output the labels corresponding to the maximum probability of words, the labels are independent of each other and thus prone to unreasonable sequences. The CRF adds some constraints to the prediction labels to ensure their accuracy, so the CRF is chosen for sequence labeling in this paper.

For a piece of text, the input text sequence is denoted by $S = \{c_1, c_2, c_3, \dots, c_n\}$, and the output annotation sequence is denoted by $y = \{y_1, y_2, \dots, y_n\}$. For the prediction sequence y , its score function is given by:

$$\text{Score}(S, y) = \sum_{i=0}^n A_{y_i, y_{i+1}} + \sum_{i=1}^n P_{i, y_i} \quad (15)$$

where P_{i, y_i} denotes the score of the y_i th tag of the i th character, A is the transfer score matrix, $A_{y_i, y_{i+1}}$ denotes the score transferred from tag y_i to tag y_{i+1} , and y_0 and y_n are the tags added at the beginning and end of the sentence. The output sequence with the highest score is predicted by Equation (9) at final decoding.

$$y^* = \operatorname{argmaxScore}(S, y) \quad (16)$$

The names of credit entities in the credit text of shipping enterprises are long and complex in composition, and there is the phenomenon of nested entities. For the characteristics of shipping enterprise credit text, a deep learning method based on multi-feature fusion is used to solve the above problems. BERT is introduced to enhance the character-level semantic representation; word-level features are obtained by reorganizing words through BMES after introducing stuttering subscripts, including word vector features, pos features, and word length features to fully extract text features and improve the entity recognition effect; and BiGRU is introduced to obtain the contextual features of the text to solve the recognition error problem caused by long entity names. Two one-way GRUs are combined together to form BiGRU to obtain the forward and reverse information, respectively, to fully extract the contextual features of the shipping enterprise credit text, and finally, by CRF into a sequence annotation task.

4. Experiments Results and Analysis

4.1. Experimental Data, Multiple Features, and Evaluation Metrics

4.1.1. Data Collection and Labeling

This paper selects 200 shipping enterprises in the Yangtze River Delta region, whose business scope involves domestic coastal and middle and lower reaches of Yangtze River general cargo shipping, foreign trade container feeder liner shipping, cargo transportation, domestic ship agency, domestic waterway freight forwarding, etc. as the research object. The 2612 textual data in the field of shipping enterprise credit crawled from the national and local maritime bureaus (<https://www.msa.gov.cn/>, accessed on 1 March 2021), magistrates' websites (<https://wenshu.court.gov.cn/>, accessed on 15 April 2021), official websites of shipping companies, and major news websites are used as the dataset for this experiment.

The shipping company credit text contains some expressions, spaces, and bulleted items that are not useful to NER and should be deleted. In this paper, we first use regular expressions to complete the work of format conversion and data cleaning, removing meaningless contents such as spaces, emoticons, and bulleted items from the text and converting data from different sources into a uniform format. The commonly used meta characters in regular expressions and their descriptions are shown in Table 1.

Before formally training the model, the textual data in the shipping enterprise credit domain are first annotated with entity types, and the commonly used annotation methods for the dataset are Beginning, Inside, Outside (BIO); Beginning, Inside, End, Outside, Single (BIOES); and Beginning, Middle, End, Single (BMES) [29]. With BIO, i.e., Beginning, Inside, Outside, the first token of an entity is labeled B-[type], tokens at other locations of the entity are labeled I-[type], and tokens not belonging to any entity are labeled O; With BIOES,

i.e., Beginning, Inside, End, Outside, Single, End is used to identify the end of an entity, and Single is used to identify an entity that contains only one token. With BMES, i.e., Beginning, Middle, End, Single, S denotes a single word. This paper uses the BIO [30] format for annotation with the help of the YEDDA [31] tool. The entity types include eight types of entities: company, honor, ship, risk penalty, organization, person, certificate, and intellectual property, and the partial annotation results obtained are shown in Table 2, where B-X denotes the beginning of the entity word, I-X denotes the middle or end of the entity word, and O denotes other words than the entity, i.e., non-entity. The BIO annotation samples are shown in Table 3 below. In this experiment, the shipping enterprise credit dataset is divided into training set, validation set, and test set according to 6:2:2.

Table 1. Metacharacter and description in regular expression.

Metacharacter	Description
^	Used to match the first line of an input word
\	Used to place the next character token, or an octal escape, or a backward reference
+	Match the previous subexpression one or more times
?	Match the preceding subexpressions zero or once
*	Match the preceding subexpressions any number of times
\$	Used to match the end of a line of input

Table 2. Category and samples of credit entities.

Label	NUM	Entity Samples	Entity Samples (English)
Company	3375	江苏永信船务有限公司	Jiangsu Yongxin Shipping Co.
Ship	1501	兴航化0058	Xinghanghua 0058
Honor	369	中国ESG 优秀企业500 强	China's Top 500 ESG Excellence Companies
Risk-Lisk	1129	重点跟踪航运公司	Focus on tracking shipping companies
Organization	156	芜湖海事局	Wuhu Maritime Bureau
Person	88	周明军	Mingjun Zhou
Certificate	187	水路运输许可证	Waterway transport license
Intel-pro	143	质量管理体系认证证书	Quality management system certification

Table 3. BIO labeled samples.

Dataset	Label	Dataset	Label
宁 *	B-Company	获 *	o
波 *	I-Company	得 *	o
银 *	I-Company	两 *	B-Certificate
星 *	I-Company	岸 *	I-Certificate
海 *	I-Company	直 *	I-Certificate
运 *	I-Company	航 *	I-Certificate
有 *	I-Company	许 *	I-Certificate
限 *	I-Company	可 *	I-Certificate
公 *	I-Company	证 *	I-Certificate
司 *	I-Company	书 *	I-Certificate

* Notes: The English translation of the sample data is “Ningbo Yinxing Shipping Co., Ltd. received the cross-strait direct shipping license certificate”.

4.1.2. L Evaluation Metrics and Parameter Settings

1. Evaluation Metrics

The evaluation metrics of the model in this paper use the accuracy, recall, and F1 value of recognition results, and the calculation formula of each metric is [27]:

$$P = \frac{\text{Predict the number of correct entities}}{\text{Total number of predicted entities}} \quad (17)$$

$$R = \frac{\text{Predict the number of correct entities}}{\text{Total number of standard entities}} \quad (18)$$

$$F1 = \frac{2 * P * R}{P + R}$$

2. Parameter Setting

Although studies have been conducted using the BERT-BiGRU-CRF model for NER tasks, there is still a gap in NER research in the area of shipping enterprise credit. With the popularity of deep learning, deep learning can use more and more frameworks, the most popular today are the following three: TensorFlow, Keras, and PyTorch. PyTorch uses imperative programming with a simple design, so it is easy to implement, and the code is easy to understand; through the technique of reverse automatic derivation, it can achieve zero latency to arbitrarily change the behavior of the neural network to achieve dynamics, so it is highly flexible and easy to debug, and in many reviews, PyTorch's speed performance outperforms frameworks such as TensorFlow and Keras. TensorFlow is a combination of a lower-level symbolic computation library and a higher-level network specification library and is pure symbolic programming, so while it is more efficient, it is more complex to implement, and the graph definition is static, so it is not as flexible and convenient as PyTorch. Keras is a high-level application programming interface (API) that runs on top of other deep learning frameworks. It has grown rapidly due to its ease of use and syntactic simplicity but is lacking in performance due to its common usage and is not as flexible as TensorFlow and PyTorch. Therefore, the experiments are based on Windows platform, using python 3.7 as the runtime environment, and the NER model is built by pytorch framework. In the process of model training, the setting of parameters is very important. The optimal parameters of the model in this paper are obtained by training the parameters on the data in the training set. The specific parameter settings are shown in Table 4.

Table 4. Model parameters.

Parameters	Value
GRU dimension per layer	150
Dropout	0.5
Optimizer	Adam
Learning Rate	0.0005
Batch size	20

4.2. Analysis of Experimental Results

Using the trained parameters as model parameters for the validation and test sets, two experiments were done in order to verify the effectiveness of the BERT-MF-BiGRU-CRF model on the shipping enterprise credit dataset. This paper focuses on the BiGRU-CRF model, BiLSTM-CRF model, Word2Vec-BiGRU-CRF model, BERT-BiLSTM-CRF model, BERT-BiGRU-CRF model, and BERT-MF-BiGRU-CRF model. The same training framework and parameters are used for each model, and each model is defined as follows:

- (1) Model in this paper (BERT-MF-BiGRU-CRF): use BERT to obtain character features; sort that may be composed according to word matching into the sets B, M, E, and S; obtain word vector features, word nature features, and word length features; and stitch the obtained features as the input of the BiGRU-CRF model.
- (2) HMM model: borrowing from the cascading HMM model proposed by Yu Hongkui et al. [32], the model consists of two parts, the underlying HMM and the adapted HMM. The former is used to recognize non-nested entities, and the latter is used to recognize nested entities.
- (3) BiLSTM-CRF model: a classical model in the field of named entity recognition.

- (4) SVM-BiLSTM-CRF model: borrowing from the SVM-BiLSTM-CRF model proposed by Zhou et al. [33]. It has an additional step compared with the BiLSTM-CRF model, i.e., using SVM to filter out sentences containing named entities.
- (5) BiGRU-CRF model: similar to the BiLSTM-CRF model, BiGRU as a variant of BiLSTM with a simpler model structure.
- (6) BERT-BiLSTM-CRF model: Based on the BiLSTM-CRF model, BERT is added, and word vectors are obtained by BERT pre-training.
- (7) Word2Vec-BiGRU-CRF model: Word vectors are obtained by Word2Vec, and the word vectors are input to the BiGRU layer for contextual feature extraction, and finally, sequence annotation is performed by CRF.
- (8) BERT-BiGRU-CRF model: character encoding is obtained by BERT, the word vector is input to the BiGRU layer for contextual feature extraction, and finally, the sequence is annotated by CRF.

The results of NER experiments for different models are shown in Table 5 below.

Table 5. Experiment results of NER based on different models (%).

Model	P (%)	R (%)	F1 (%)
HMM	60.23	61.25	60.71
BiLSTM-CRF	69.84	68.16	69.01
SVM- BiLSTM-CRF	70.66	69.38	69.98
BiGRU-CRF	71.03	68.56	69.35
Word2Vec-BiGRU-CRF	73.35	72.74	72.99
BERT-BiLSTM-CRF	79.52	78.43	78.80
BERT-BiGRU-CRF	84.54	82.14	83.38
BERT-MF-BiGRU-CRF	92.84	90.46	91.75

The precision, recall, and F1 values of the BiLSTM-CRF model only reached 69.84%, 68.16%, and 69.01% on the shipping firm credit dataset. The BiGRU-CRF model has a simpler structure compared to the BiLSTM-CRF model, and the recognition on the shipping enterprise credit dataset is a little better, with an accuracy improvement of 1.19%; although the recognition effect is improved, there is great room for improvement.

In order to verify the effectiveness of Word2Vec and BERT for character-level feature extraction in shipping enterprise credit text, a comparison experiment between Word2Vec-BiGRU-CRF and BERT-BiGRU-CRF was conducted. The experimental results show that the precision, recall, and F1 values of the Word2Vec-BiGRU-CRF model for entity recognition are 73.35%, 72.74%, and 72.99%, respectively, and the precision, recall and F1 values of the BERT-BiGRU-CRF model for entity recognition are 84.54%, 82.14%, and 83.38%, respectively. It can be clearly seen that the BERT-BiGRU-CRF model entity recognition is better with significantly higher precision and recall, and the F1 value is improved by 10.39%. With the same base model, it can be concluded from this experiment that BERT is more suitable for shipping enterprise credit text character-level feature extraction.

By adding the BERT pre-training model to the BiLSTM-CRF model to form the BERT-BiLSTM-CRF model, compared to the BiGRU-CRF model and BiLSTM-CRF model, the accuracy, recall, and F1 values were improved by 14.66% and 9.68%, 13.58% and 10.27%, and 14.03% and 9.79%, respectively, which are all three significant improvements. The experimental results illustrate that the introduction of pre-trained models in the process of NER of shipping enterprise credit can obtain rich word-level semantic representations, alleviate the problem of multiple meanings of words, and improve the recognition effect of shipping enterprise credits. In addition, the experimental results of comparing the BERT-BiLSTM-CRF model with the BERT-BiGRU-CRF model show that the simpler structure of BiGRU can achieve better recognition results for the shipping enterprise credit dataset.

By adding the word vector, word length feature vector, and pos feature vector in the text to the BERT-BiGRU-CRF model, the F1 value of shipping enterprise credit NER reaches 92.84%; compared with the single BERT-BiGRU-CRF model recognition result, the F1 value

improves by 8.3%, accuracy improves by 8.32%, and recall improves by 8.37%. All three have been greatly improved, which can strongly demonstrate that for shipping enterprise credit data, the introduction of multiple features in the named entity recognition process can effectively improve the model entity recognition effect, especially for the entities with long entity names in shipping enterprise credit; the fusion of multiple features can better solve the complex nesting problem existing in shipping enterprise credit entities and improve the effect of entity recognition.

Compared with other models, this paper proposes using the BERT-MF-BiGRU-CRF model for shipping enterprise credit entity recognition, which has the highest accuracy, recall, and F1 value. This proves that the introduction of pre-trained models in the NER process can make full use of the contextual information of the utterance to obtain rich word-level semantics, and on this basis, the fusion of multiple features in the input of the model can obtain rich text semantic features, which effectively solves the problem of complex name structures of shipping enterprise credit entities and improves the effect of shipping enterprise NER.

4.3. Discussion

Facing the problems of long credit entities and nested entities in the shipping field, this paper proposes to use the BERT-MF-BiGRU-CRF model for shipping enterprise credit entity recognition, which fuses the feature vectors such as word vector features, word length features, and word nature features in the text with the character features obtained from the BERT model as the gated recurrent neural network input; the BERT pre-training model combined with multi-features and BiGRU has strong semantic representation capability, which can capture both contextual features and additional features, making full use of text information to capture all features of the text. The experiments prove that the multi-feature fusion has a large improvement on the recognition of shipping enterprise credit entities.

4.3.1. Effect of Different Features on the Model

In order to improve the recognition effect of the model, we introduce various features, including character-level features, word vector features, word length features, and pos features, to address the problem that the credit entities of shipping enterprises are long, complex in composition, and have the phenomenon of nested entities. In order to study the degree of influence of different features on the entity recognition effect in the credit domain of shipping enterprises, different features were added to the BiGRU-CRF model for experiments, respectively, and the entity recognition effect of the model is shown in Table 6.

Table 6. Experiment results of NER based on different features.

Model	P (%)	R (%)	F1 (%)
Word length Features-BiGRU-CRF	72.15	67.92	69.39
POS Features-BiGRU-CRF	73.35	71.09	72.21
Word vectors-BiGRU-CRF	74.42	74.50	74.33
MF-BiGRU-CRF	79.65	77.93	78.42
BERT-BiGRU-CRF	84.51	82.17	83.36
BERT-MF-BiGRU-CRF	92.85	90.64	91.73

According to the experimental results in Table 6, it can be seen that by adding the word length feature, pos feature, and word vector feature to the BiGRU-CRF model, respectively, although the effect of shipping enterprise credit entity recognition is similar, there are subtle differences, but the effect is not significant, so these three features have approximately the same degree of influence on the model. These three word-level features were spliced and input into the BiGRU-CRF model to form the MF-BiGRU-CRF model, and their precision, recall, and F1 values reached 79.65%, 77.93%, and 78.42%, respectively, all of which were higher than the recognition results achieved by incorporating any single feature in the

model, indicating that using the spliced features in the model is better than a single feature and more suitable for the NER task of shipping enterprise credit.

In order to compare the impact of character-level features with the spliced word-level features on the model, the experimental results of the BERT-BiGRU-CRF model are compared with the MF-BiGRU-CRF model in this paper, and the F1 value of the former is 4.94% higher than that of the latter; considering that the performance of the model incorporating multiple features is better than that of the model incorporating single features in the shipping enterprise credit NER task, the three features obtained at the word level and the character-level features obtained by BERT are spliced and input into the BiGRU-CRF model to form the BERT-MF-BiGRU-CRF model, the accuracy, recall and F1 values of which are higher than 90%, which is the best among all models for entity recognition. Therefore, for the characteristics of shipping enterprise credit text, we choose several suitable features to join the NER model and can obtain better recognition effect.

4.3.2. Different Types of Entity Recognition Results

The BERT-MF-BiGRU-CRF model proposed in this paper consists of an input layer, an embedding layer, a BiGRU layer, and a CRF layer. Firstly, the character-level features of shipping enterprise credit text are obtained through BERT; secondly, the unlabeled shipping enterprise credit data are separated by jieba, and word reorganization is performed using four sets of B, M, E, and S to extract word vector features, word length features, and pos features of shipping enterprise credit text; then, the obtained feature vectors are spliced and used as the input of BiGRU to obtain the contextual features of shipping enterprise credit text; finally, classification is performed through the CRF layer.

In order to further understand the credit NER task of shipping enterprises, the recognition effects of various types of entities are analyzed based on the BERT-MF-BiGRU-CRF model proposed in this paper. Table 7 shows the recognition of shipping enterprise credit entities by BERT-MF-BiGRU-CRF in detail.

Table 7. Experimental results of BERT-MF-BiGRU-CRF entities.

Model	P (%)	R (%)	F1 (%)
Company	80.11	82.39	81.54
Ship	85.52	83.46	84.91
Honor	73.34	73.79	73.48
Risk-List	87.49	87.32	87.40
Organization	90.94	92.88	91.89
Person	93.32	91.61	92.54
Certificate	74.75	75.58	75.24
Intel-pro	75.54	73.38	74.89

As can be seen from Table 5, the recognition of Honor, Intel-pro, and Certificate are not satisfactory, both in terms of accuracy and recall, which are only about 70%, and the F1 values of these three types of entities are 73.34%, 74.89%, and 75.24%, which are much lower than those of Person (92.54%) and Organization (91.89%), which are the two types of entities. Analysis of the reasons may be that the format of Honor, Intel-pro, and Certificate entities in the shipping enterprise credit text is not uniform and complex in composition, with both Chinese characters and English letters and numbers, for example, the Honor entity sample “中国ESG 优秀企业500强 (The English translation of the sample is “China ESG Top 500 Excellent Enterprises”)), which makes it impossible to identify the entity accurately in most cases. Secondly, for the recognition of Company, although the F1 value reached 81.54%, there is still a gap compared to the two types of entities, Person (92.54%) and Organization (91.89%). Company entities often have long names and nested entities, such as “马鞍山市江泰船务有限公司” (The English translation of the sample is “Maanshan Jiangtai Shipping Co”). In the NER process, it is likely to be recognized as “马鞍山市” (The English translation of the sample is “Maanshan City”) and “江泰船务有

限公司” (The English translation of the sample is “Jiangtai Shipping Co”). However, the recognition results are quite good for entities in other standard formats, which also shows the effectiveness of the model in this paper for the credit NER task of shipping companies.

5. Conclusions

- (1) Although studies have been conducted using the BERT-BiGRU-CRF model for NER tasks, there is still a gap in NER research in the area of shipping enterprise credit. Facing the problems of complex and diverse shipping enterprise credit entities and nested entities, this paper innovatively proposes to use the BERT-MF-BiGRU-CRF model with multi-feature fusion for shipping enterprise credit entity recognition, and the experimental results show that the method has better effect on entity recognition on the dataset of shipping enterprise credit domain.
- (2) Based on the BERT pre-training model to obtain character-level feature vectors, the semantic representation of shipping enterprise credit text is enriched to alleviate the problems of multiple meanings of words and low recognition rate of entities due to excessive length of entities, which greatly improves the recognition performance of the model. In future research, more models should be applied in this field, and the models should be optimized and improved according to the textual characteristics of the shipping enterprise credit field to promote the research progress of NER in the shipping enterprise credit field.
- (3) The NER model of the shipping enterprise credit domain proposed in this paper can provide reference for NER in other domains, and the entity information recognition by this model can lay the foundation for the later relationship extraction and the construction of a shipping enterprise credit knowledge graph. The research results have an important role in improving the credit analysis in the shipping field and are of great practical significance in promoting the healthy and rapid development of China's shipping industry. Of course, considering the small size of the dataset in this paper, in future research, in order to further improve the effect of shipping enterprise credit NER and enhance the quality of the shipping enterprise credit knowledge graph, the dataset in this paper should be continuously enriched and extended on the basis of this paper.
- (4) At present, credit assessment in the shipping field is in the initial stage, which is the intersection field of shipping big data and artificial intelligence. The current credit assessment models usually ignore the differences between regions and industries, and there are relatively few studies on credit assessment models for the characteristics of the shipping field, and the imperfect and opaque credit information prevailing in the shipping field is the most important challenge to build a high-quality shipping credit assessment model. However, the research methods for industry research ideas of credit assessment in related fields can be borrowed, and the credit assessment model suitable for the characteristics of the shipping field can be further researched on the basis of this paper in the future.

Author Contributions: Investigation, S.W.; Data acquisition, S.W. and X.C.; Methodology, S.W.; Software, S.W.; Writing—original draft, S.W.; Writing—review & editing, L.H. and X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Kang, H.-W.; Wang, G.W.Y.; Bang, H.S.; Woo, S.-H. Economic performance and corporate financial management of shipping firms. *Marit. Econ. Logist.* **2016**, *18*, 317–330. [CrossRef]
- Woo, S.-H.; Kwon, M.-S.; Yuen, K.F. Financial determinants of credit risk in the logistics and shipping industries. *Marit. Econ. Logist.* **2021**, *23*, 268–290. [CrossRef]
- Clintworth, M.; Lyridis, D.; Boulougouris, E. Financial risk assessment in shipping: A holistic machine learning based methodology. *Marit. Econ. Logist.* **2021**, *25*, 90–121. [CrossRef]
- Yang, B.; Liao, Y.-M. Research on enterprise risk knowledge graph based on multi-source data fusion. *Neural Comput. Appl.* **2022**, *34*, 2569–2582. [CrossRef]
- Yang, Q.; Peng, Z.; Su, H.; Deng, W.; Zhang, H.; Li, M. Chinese electric power entity recognition based on Bi-LSTM-CR. *Inf. Technol.* **2021**, *09*, 45–50. [CrossRef]
- Cook, H.V.; Jensen, L.J. A Guide to Dictionary-Based Text Mining. In *Bioinformatics and Drug Discovery; Methods in Molecular Biology*; Larson, R., Oprea, T., Eds.; Humana Press: New York, NY, USA, 2019; Volume 1939, pp. 73–89. [CrossRef]
- Feng, X.; Zhang, H.; Hao, W. Named Entity Recognition for Military Text. *Comput. Sci.* **2015**, *42*, 15–18+47.
- Hu, W.; Tian, G.; Kang, Y.; Yuan, C.; Maybank, S. Dual Sticky Hierarchical Dirichlet Process Hidden Markov Model and Its Application to Natural Language Description of Motions. *IEEE Trans Pattern Anal. Mach. Intell.* **2018**, *40*, 2355–2373. [CrossRef]
- Lafferty, J.D.; McCallum, A.; Pereira, F. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In Proceedings of the 18th International Conference on Machine Learning, Williamstown, MA, USA, 28 June–1 July 2001; Available online: <http://www.aladdin.cs.cmu.edu/papers/pdfs/y2001/crf.pdf> (accessed on 10 March 2023).
- Haffner, P.; Tur, G.; Wright, J.H. Optimizing SVMs for complex call classification. In Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, Hong Kong, China, 6–10 April 2003. [CrossRef]
- Wang, D.; Fan, X. Named entity recognition for short text. *Comput. Appl.* **2009**, *29*, 143–145+171. [CrossRef]
- Pan, H.; Yan, X.; Yu, Z.; Guo, J. Vietnamese News Text Classification Method Based on Support Vector Machine. *Shanxi Univ. Nat. Sci. Ed.* **2013**, *36*, 505–509. [CrossRef]
- Habib, M.B.; van Keulen, M. A Hybrid Approach for Robust Multilingual Toponym Extraction and Disambiguation. In Proceedings of the Language Processing and Intelligent Information Systems, Warsaw, Poland, 17–18 June 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 1–15. [CrossRef]
- Liu, K.; Zhou, X.; Yu, J.; Zhang, R. Named Entity Extraction of Traditional Chinese Medicine Medical Records Based on Conditional Random Field. *Comput. Eng.* **2014**, *40*, 312–316. [CrossRef]
- Wang, S.; Li, S.; Chen, T. Recognition of Chinese Medicine Named Entity Based on Condition Random Field. *J. Xiamen Univ. Nat. Sci.* **2009**, *48*, 359–364.
- Hu, H.; Zhang, H. Chinese Named Entity Recognition with CRFs: Two Levels. In Proceedings of the 2008 International Conference on Computational Intelligence and Security IEEE Computer Society, Suzhou, China, 13–17 December 2008; Volume 2, pp. 1–6. [CrossRef]
- Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF Models for Sequence Tagging. *arXiv*, 2015; arXiv:1508.01991. [CrossRef]
- Ma, X.; Hovy, E.H. End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF. *arXiv*, 2016; arXiv:1603.01354.
- Jiang, X.; Ma, J.; Yuan, H. Named Entity Recognition in the field of ecological management technology based on BiLSTM-IDCNN-CRF model. *Comput. Appl. Softw.* **2021**, *38*, 134–141. [CrossRef]
- Yu, Q.; Li, C. Research on entity recognition and tagging in fiscal and taxation domain. *Comput. Eng.* **2020**, *46*, 312–320. [CrossRef]
- Chu, D.; Wan, B.; Li, H.; Fang, F.; Wang, R. Geological Entity Recognition Based on ELMO-CNN-BiLSTM-CRF Model. *Earth Sci.* **2021**, *46*, 3039–3048. [CrossRef]
- Gong, L.; Zhang, Z. Clinical named entity recognition from Chinese electronic medical records using a double-layer annotation model combining a domain dictionary with CRF. *Chin. J. Eng.* **2020**, *42*, 469–475.
- Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, 2–7 June 2019; Association for Computational Linguistics: Minneapolis, MN, USA, 2019; Volume 1, pp. 4171–4186, abs/1810.04805. [CrossRef]
- Lin, F.; Xie, D. Research on Named Entity Recognition of Traditional Chinese Medicine Electronic Medical Records. In *Health Information Science*; Springer: Cham, Switzerland, 2020; Volume 12435. [CrossRef]
- Zhao, P.; Zhao, C.; Wu, H.; Wang, W. Recognition of the agricultural named entities with multi-feature fusion based on BERT. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 112–118.
- Tang, X.; Liu, Z. Research on Text Sequence Tagging and Joint Extraction of Entity and Relation in Financial Field. *Inf. Sci.* **2021**, *39*, 3–11. [CrossRef]
- He, L.; Zhang, Y.; Ba, H. Named entity recognition of exotic marine organisms based on attention mechanism and deep learning network. *J. Dalian Ocean. Univ.* **2021**, *36*, 503–509. [CrossRef]
- Gan, C.; Li, M.; Wan, Y.; Zhang, Q. Named Entity Recognition Based on Character Embedding and Bidirectional Gated Recurrent Network. *Softw. Guide* **2023**, 1–5.
- Ghaddar, A.; Langlais, P.; Rashid, A.; Rezagholizadeh, M. Context-aware Adversarial Training for Name Regularity Bias in Named Entity Recognition. *Trans. Assoc. Comput. Linguist.* **2021**, *9*, 586–604. [CrossRef]

30. Sang, E.F.T.K.; Veenstra, J. *Representing Text Chunks*; ACL: Bergen, Norway, 1999; pp. 173–179. Available online: <https://aclanthology.org/E99-1023> (accessed on 10 March 2023).
31. Yang, J.; Zhang, Y.; Li, L.; Li, X. YEDDA: A Lightweight Collaborative Text Span Annotation Tool. In Proceedings of the ACL 2018, System Demonstrations, Melbourne, Australia, 15–20 July 2018; Association for Computational Linguistics: Melbourne, Australia, 2018; pp. 31–36. [[CrossRef](#)]
32. Yu, H.; Zhang, H.; Liu, Q.; Lv, X.; Shi, S. Chinese named entity identification using cascaded hidden Markov model. *J. Commun.* **2006**, *2*, 87–94.
33. Zhou, X.; Zhao, X.; Liu, T. Named Entity Recognition Method of Judgment Documents with SVM-BiLSTM-CRF. *Comput. Syst. Appl.* **2019**, *28*, 245–250. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.