Numeral Tense Detection in Chinese Financial News

Yuxuan Liu
School of Computer Science and Technology, Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology
Wuhan, China
L823500180@163.com

Maofu Liu*
School of Computer Science and Technology, Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology
Wuhan, China
liumaofu@wust.edu.cn

Mengjie Wu
School of Computer Science and Technology, Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology
Wuhan, China
1749565249@qq.com

ABSTRACT

Time information is a very important dimension in information space, which can be shown as tense expressions in natural language. Meanwhile, numerals play an important role in financial texts, which is the embodiment of fine-grained information, and most of financial events contain numerals. We have observed that Chinese does not express the tense of texts intuitively at the lexical level of verbs, but through some adverb or auxiliary tense operators, and there has not further research on numeral tense in Chinese financial texts yet. However, the tense of numerals in the texts is very crucial for the financial fields which pays attention to time series. Therefore, to assist Chinese tense understanding in financial texts, in this paper, we propose a novel task for numeral tense detection in Chinese financial fields. We firstly annotate a numeral tense dataset based on Chinese finance news texts, named CFinNumTense, which defines the numeral tense categories into “past tense”, “future tense”, “static state” and “time”, and then conduct Chinese finance numeral tense detection task on CFinNumTense. We employ RoBERTa (Robustly Optimized BERT Pretraining Approach) pre-trained model as the embedding layer and use four baseline models, i.e., FNN (Feedforward Neural Network), TextCNN (Text Convolutional Neural Networks), RNN (Recurrent Neural Networks) and BiLSTM (Bidirectional Long Short-Term Memory), to detect numeral tenses, respectively. In the ablation experiments, we design NE (Numeral Encoding) to improve the information on target numeral in the texts, and design an auxiliary learning model based on BiLSTM. Experiments show that the multitask learning of target numeral tense detection and tense operator extraction can strengthen the understanding ability of target numeral tense in Chinese financial texts.

*corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
WWW '22 Companion, April 25–29, 2022, Virtual Event, Lyon, France
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9130-6/22/04…$15.00
https://doi.org/10.1145/3487553.3524639

CCS CONCEPTS

• **Computing methodologies** → *Information extraction*; *Phonology / morphology*.

KEYWORDS

Financial numeral, Tense detection, Multitask learning

ACM Reference Format:
Yuxuan Liu, Maofu Liu, and Mengjie Wu. 2022. Numeral Tense Detection in Chinese Financial News. In *Companion Proceedings of the Web Conference 2022 (WWW '22 Companion), April 25–29, 2022, Virtual Event, Lyon, France.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3487553.3524639

1 INTRODUCTION

Numerals play a very important role in text semantics, especially in some specific fields, such as financial field. Chen et al. [5] has pointed out that the lowest probability of numerals is in earnings call, reaching 55.3%, and the highest probability of numerals is in financial news, up to 99.8%, which implies that numerals play an extremely significant role in financial news texts. Numerals are also essential for stock market prediction. Typically, historical stock price or technical indicators are usually used to predict future stock prices [8, 16]. Nowadays, many studies predict stock prices through news titles or news texts encoding [1, 14, 18], since the financial news texts contains a large number of events in listed companies, which can have a certain impact on the company's stock price. In these financial events, we find that numerals are important elements in the most of financial events, and we list two financial events and event roles in Table 1. In the “Buy shares” event and “Repurchase” event, “Time” role, “Share” role and “Amount of money” role are all in the form of the numeral, which shows that numerals plays a very important role in financial events.

Table 1: Financial event and event roles

Finance event	Event roles			
Buy shares	Time	Subject	Shares	Amount of money
Repurchase	Time	Subject	Shares	Amount of money

Numerals provide more fine-grained information for financial events, and financial numerals in different tenses will also give investors different judgments and further affect stock prices. Tense is a central linguistic phenomenon and a very important dimension of information space. However, in the process of modeling, these studies [1, 8, 14, 16, 18] have not paid attention to the numeral tense information in financial news texts. Therefore, we propose a novel task, called numeral tense detection in Chinese financial news texts, which plays a very crucial role in financial prediction based on time series.

In linguistics, tenses generally include three kinds, i.e., past, present, and future. The languages of the world differ greatly in whether and how they express tense. For examplein, English, as an inflectional language, the tense is always grammaticalized as an inflectional morpheme attached to a verb. Chinese, on the other hand, as an isolated language, such morphological cues are rare or non-existent, and the underlying semantic tense has to be inferred from the context by adding some adverbs or auxiliary words. Briefly, Chinese cannot express the tense intuitively at the grammatical level like English and other inflectional languages. And it generally uses adverbs and auxiliary words, which we call tense auxiliary operators, to express the tense.

Example 1:

In Chinese: 我要去听讲座

In English: I am going to listen to a seminar.

From Example 1, it can be seen that English forces English speakers to habitually use language tenses (am going to) to distinguish among past, present and future time, while Chinese speakers do not. This does not mean that Chinese speakers cannot understand the differences among the past, present and the future time, but judge tenses through contextual semantics or some adverbs and auxiliary words called tense operators in linguistics. Based on this, we propose the task of numeral tense detection in Chinese financial news texts.

Example 2:

In Chinese: 齐鲁银行的资本将会剧烈下降**6.8**亿元，这将直接导致齐鲁银行资本充足率的下降，并给齐鲁银行的经营带来压力。

In English: The capital of Qilu bank will drop sharply by **6.8** hundred million yuan, which will directly lead to the decline of capital adequacy ratio and put pressure on the operation of the bank.

Example 3:

In Chinese: 公司充分利用**5**G的高速发展时期，推动安世半导体Alpha业务增加处于强劲增长阶段的Beta业务。

In English: The company would seize the high-speed development period of **5**G, and promote the Anson semiconductor Alpha bussiness to increase the Beta business which is in strong growth.

We notice that in Chinese financial news the “past tense” and “future tense” account for the majority. There are “time” and numerals representing “static state” in the news as well. The numerals representing time points and time periods are considered as “time”, while the numerals representing stock codes and proper nouns such as “600392” and “5” in “5G”, that is, in the real world, the financial numerals whose factual meaning will not change for a long time, are considered as “static state”. In Example 2, according to the context semantics, the number “6.8" is classified as “future tense”, and the tense operator is “将会”. In Example 3, the number “5” appears in a proper noun and is therefore classified as “static state”.

The main contributions of this paper are in threefold:

1. We have collected the financial news texts from the JRJ website, and constructed the CFinNumTense dataset. The numeral tense category and the tense operator of the target numeral have been manually annotated in CFinNumTense dataset. In the experiments, we divide the dataset into training set and test set according to the ratio of 7:3.

2. We adopt FNN (Feedforward Neural Network), TextCNN (Text Convolutional Neural Networks), RNN (Recurrent Neural Networks), BiLSTM (Bi-directional Long Short-Term Memory) to carry out experiments on CFinNumTense dataset to validate the performances of these baseline models on our proposed numeral tense detection task.

3. We add NE (Numeral Encoding) on the basis of baseline models to improve the performances. We also design an auxiliary learning task to extract tense operators i.e., tense clues, to promote the BiLSTM model understanding capability of tense semantics by sharing parameters.

2 RELATED WORK

Analyzing the numeral information in documents is an emerging topic, and attracts more and more attention. Chen et al. [6] attempt to explore the question of whether neural network models can learn numeracy and numeracy refers the ability to predict the magnitude of a numeral at some specific position in a text description. They also present an important application scenario, i.e., detecting exaggerated information. Spithourakis et al. [17] explore different strategies for modelling numerals with language models, such as memorisation and digit-by-digit composition, and propose a novel neural architecture that uses a continuous probability density function to model numerals from an open vocabulary. Chen et al. [4] design a novel task for argument mining in the financial domain, and provide an expert-annotated dataset, NumClaim. For the proposed task and their experimental results show that encoding numeral and co-training with the auxiliary task of the numeral understanding, i.e., the category classification task, can improve the performance of the proposed task under different neural network architectures. Liu et al. [12] present CFinNumAttr, a financial numeral attribute dataset in Chinese and their experimental results on the CFinNumAttr dataset show that the numeral attributes in social reviews or comments contain rich semantic information.

Based on our observation, there are some natural language processing related tense studies as well. Pustejovsky et al. [15] present a corpus, called TIMEBANK, which is richly annotated to indicate events, times, and temporal relations, making the utility of machine learning techniques can be tested. Asgari et al. [2] investigate the typology of Tense in 1000 languages with computation. Xue et al. [20] describe a method of annotating the tense of a Chinese sentence by annotating the tense of its English translation and then projecting this annotation back onto the Chinese sentence. Chambers et al. [3] propose a fully automatic two-stage machine learning architecture that learns temporal relations between pairs of events and define tense as one of the attributes of events. Vashishtha et al. [19] research Fine-grained temporal relation extraction aims to recognize the durations and timeline of event mentions in text. However, these works related to tense research does still stay in a general field and not go deep into the tense detection task in a specific field. Since there is no research on the tense of Chinese financial fields, we propose the numeral tense detection task based on Chinese finance news for the first time. In this paper, we propose a numeral encoding method suitable for the experiments, and achieved remarkable results. And we design an auxiliary task called tense operations extraction to assist the main task learning. The experiment results are shown in Section 4.

3 DATASET

3.1 Data Collection

We crawl Chinese finance news texts from the JRJ, which is a leading financial information provider in China and one of the largest Chinese financial websites in the world. The JRJ website involves a lot of financial information, such as finance, stocks, funds, futures, bonds, foreign exchange, banking, insurance, precious metals, real estate and so on. We have collected financial news texts which are released by February 2022 in JRJ website. After data cleaning, we select news texts with moderate text length and as many numeral as possible to form our CFinNumTense dataset, and then ask annotators to annotate the tense category of the target numeral. The CFinNumTense dataset has 10,930 instances of data. We refer to the tense related knowledge in linguistics and define four categories, i.e., "past tense", "future tense", "static state" and "time".

3.2 Data Annotation

In the annotation process, we find the tense information contained in the target numeral according to the semantics of financial news texts, and then judge whether it contains the tense operator that can indicate the certain tense category of the target numeral. In this paper, the BIO labeling system is used for the sequence labeling task, that is, each token is labeled as "B-CLU", "I-CLU" or "O", where "B-CLU", "I-CLU" and "O" represent beginning of the tense operator, rest of the tense operator and out of the tense operator, seperately.

The thorough annotation process with specific samples is as follows.

Step 1: For a given number, annotators first determine if the target number represents a time point or time period in the text, and if the answer is "Yes", it will be annotated as "time". In Example 4, "1" and "25" represent a time point and "10" denotes a time period.

Example 4:

In Chinses: 公司股票与可转债将于**1**月**25**日开市起停牌，停牌时间不超过**10**个交易日。

In English: The trading of the company's shares and convertible bonds will be suspended from the opening of the market which is on **1**/**25** for no more than **10** trading days.

Step 2: Next, if the target number in the text represents a proper noun or will not change for a long time, and it will be annotated as "static state". In Example 5, the target numbers "225" and "100" are the financial terminology.

Example 5:

In Chinese: 近日，在深交所与日本交易所集团合作的ETF互通项目中，两国基金管理人各自取得相关产品批准，将分别设立ETF产品，投资于对方市场的单只目标ETF，实现跟踪投资日经**225**指数、粤港澳大湾区创新**100**指数。

In English: Recently, in the ETF interworking project cooperated by Shenzhen Stock Exchange and Japan stock exchange group, the fund managers of the two countries have obtained the approval of relevant products and will set up ETF products respectively to invest in a single target ETF in the other market, so as to track the investment in the Nikkei **225** index and the innovation **100** index of Guangdong, Hong Kong and Macao Greater Bay area.

Step 3: And then, annotators further determine whether the target number is an element in the financial-related events contained in the text.

Step 4: If the event of which the given number is an element has occurred in the past, the annotator will annotate as "past tense", and if there is a tense operator representing the past tense, the clue text span will be annotated. In Example 6, the extreme decline event of Suning Tesco's share price is mentioned. The two elements of "7" and "6.46" are the roles of "decline range" and "post decline price", respectively. "一度" has the meaning of "once" and we can know that the event of extreme decline of Suning's share price occurred in the past.

Example 6:

In Chinese: 苏宁易购一度跌近**7**%触及**6.46**元，股价创2013年8月以来新低。

In English: Suning Tesco once fell nearly **7**% to **6.46** yuan, which made its share price hit a new low since August 2013.

Step 5: If the semantics of the event in which the given number is represents the occurrence or prediction of the occurrence in the future, annotators will annotate it as "future tense", and if the context contains the tense operator representing the future tense, the clue text span will be annotated. In Example 7, the decline in the growth rate of social financing is mentioned. The two elements of "11" and "12" are both the role of "growth rate after decline" in this event. "预计" means that the event has not happened but is predicted to happen.

Example 7:

In Chinese: 宏观流动性温和回归，预计2021年社融增速缓慢回落至**11**%-**12**%。

In English: Macro liquidity returns moderately and it is expected that the growth rate of social finance will slow down to **11**% - **12**% in 2021.

Step 6: If the given number does not belong to the above four categories, it is annotated as "other". In Example 8, since the given number "1.5" is neither a role in financial events nor a proper noun or time, it is annotated as "other".

Example 8:

In Chinese: 公司主要从事风塔及风塔零部件的生产和销售,主要产品是用于**1.5**MW及以上功率风机的风塔。

In English: the company is mainly engaged in the production and sales of wind towers and wind tower parts. Its main products are wind towers for fans with power of **1.5**MW and above.

Step 7: In the end, our constructed dataset only collects the data instances with the labels of financial numeral tense defined in this paper.

Through the above data annotation process, we summarize an annotation flow chart to show the annotation logic clearly and the common tense operators in Chinese financial news texts, as shown in Figure 1 and Table 2, respectively.

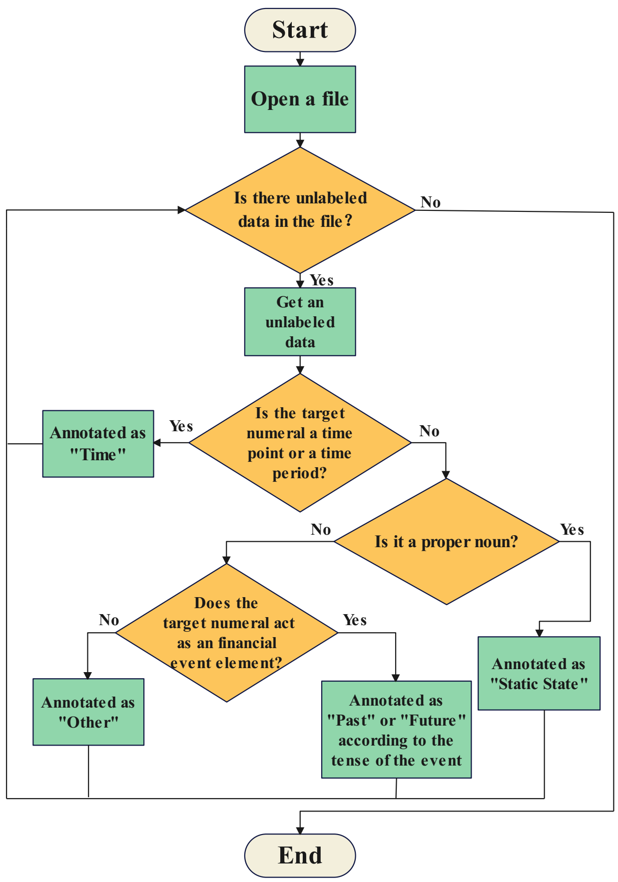

Figure 1: Annotation Flow Chart

Table 2: Common tense operators in Chinese Financial news

Tense	Tense operators
Past	了，已经，昨日，曾，一度
Future	拟，将，未来，预计

3.3 Tense Category Distribution in Dataset

Table 3 and Table 4 show the statistics and proportion of training set and test set of the CFinNumTense dataset. In the 10,930 instances in the CFinNumTense dataset, we separate about 70% of the data as the training set, the remaining 30% as the test set, and take 10% of the test data as the development set. In training set, 20.65% and 31.45% of instances with target numeral are annotated as "past tense" and "future tense", 11.12% and 36.78% of instances with target numeral are annotated as "static state" and "time". In test set, 15.57% and 32.74% of instances with target numeral are annotated as "past tense" and "future tense", 12.48% and 39.21% of instances with target numeral are annotated as "static state" and "time". In the training set and test set, the proportion of various tense category is relatively balanced, and the distribution of each tense categories in the training set is basically similar to that in the test set.

Table 3: Distribution of training data

	Past	Future	Static state	Time	Total
Quantity	1549	2359	834	2758	7500
Proportion	20.65%	31.45%	11.12%	36.78%	100%

Table 4: Distribution of test data

	Past	Future	Static state	Time	Total
Quantity	534	1123	428	1345	3430
Proportion	15.57%	32.74%	12.48%	39.21%	100%

4 MODELS

4.1 Robustly Optimized BERT Pretraining Approach

RoBERTa (Robustly Optimized BERT Pretraining Approach) [13] is based on the improvement of BERT (Bidirectional Encoder Representations from Transformers) [7]. BERT is a bi-directional encoding representation model derived from the transformers model, and RoBERTa is improved on the basis of BERT with three training aspects, which make RoBERTa better representation for the downstream tasks than BERT.

In this paper, we use the RoBERTa model as embedding module without fine-tuning. In the tokenizing phase, we need to add some tokens to the vocabulary of RoBERTa to ensure that the target numeral can be segmented into single characters to promise the correct offset information of the target numeral and adapt to the sequence labeling task.

In the FNN model, we choose the "CLS" token vector as the embedding. In TextCNN, RNN and BiLSTM models, we use text vector as the embedding.

4.2 Neural Network Architecture

We conduct experiments employing Feedforward Neural Network (FNN), Text Convolutional Neural Networks (TextCNN), Recurrent Neural Networks (RNN), Bi-directional Long Short-Term Memory (BiLSTM) as baseline models. We call batch size as bz. In the FNN model, the size of input matrix is bz × 768 without NE and bz × 772 with NE. In the TextCNN model [10], we set the input channel as 768 without NE and 772 with NE. The output channel is 512. We use a maximum pooling layer and a multi-layer perceptron to classify the "past tense", "future tense", "static state" and "time" of target numerals. The RNN model consists of a RNN layer and a multi-layer perceptron. The hidden size of the RNN layer is set to 500. The BiLSTM model [21] consists of a BiLSTM layer with hidden size of 500 and a multi-layer perceptron.

4.3 Numeral Encoding (NE)

We represent the target numeral with discrete representation and add a distributed representation to present the target numeral position information. In Figure 2, they are represented by NR (Numeral Representation) and PR (Position Representation) modules, respectively. That is, we use bag-of-words model and set the bag-of-words size to 12. For each target numeral, we get a 1×12 tensor to represent the digit (0-9) and two other symbols related to the target numeral (the decimal point and minus), and concatenate the 1×12 tensor with a 1×7 tensor of the target numeral position information. In the experiments, we use the FNN architecture which output dimension is 4 to encode the numeral information. In Figure 2, it is represented by NE (Numeral Encoding) module, and concatenate the encoded numeral information with the context information, then we get the embedding whose last dimension is 772.

4.4 Multitask Learning (ML)

Considering the tense operators in Chinese for understanding tense semantics in Chinese texts, we propose a new learning task called tense operators extraction to assist the main task. In this paper, we employ multitask learning architecture allowing tasks to share knowledge in the learning process, and improving the performance and generalization ability of the model by using the correlation of multiple tasks. The multitask learning architecture in this paper is shown in Figure 2. We use BiLSTM-CRF (Bi-directional Long Short-Term Memory-Conditional Random Field) [9] model to conduct sequence annotation task, and extract tense operators as an auxiliary task. BiLSTM-CRF has achieved the performance of state-of-the-art on POS, chunking and NER tasks. In the process of training, BiLSTM effectively captures the input contextual features. The two tasks share the parameters of the BiLSTM layer, which can improve the ability of the model to understand the tense information of the main tasks, so as to improve the performance of the model.

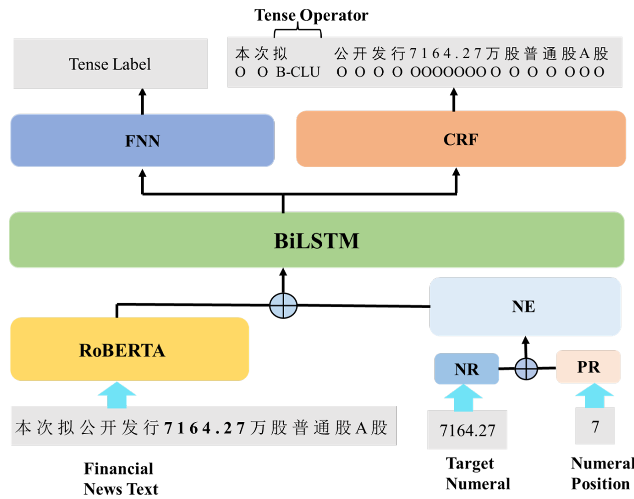

Figure 2: Multitask Learning Model Architecture

5 EXPERIMENTS

5.1 Setting

There are 10930 instances in the CFinNumTense data set, and the ratio of training set to test set is about 7:3. We use Adam [11] as the optimizer, and the learning rate is set to 0.0001. The batch size (bz) is 64. We adopt Macro-F1 and Micro-F1 as the evaluation metrics of the results.

5.2 Results

The experimental results measured by Micro-F1 and Macro-F1 metrics has listed in Table 5 and Table 6. We can see that Micro-F1 is better than Macro-F1 for the same model. Table 5 shows the performances of FNN, TextCNN, RNN and BiLSTM baseline models and Table 6 illustrates the performances of the baseline models with numeral encoding (NE), respectively. From the performance of Micro-F1 and Macro-F1, the model with numeral encoding is significantly better than the baseline models. These results show that encoding the symbol information of the target numeral provides the semantic understanding task of the target numeral, which is useful in the proposed task. This also shows that symbolic information and is not originally learned in RoBERTa representation. Table 6 also shows the results with numeral encoding and auxiliary task. On the model of auxiliary learning with sequence annotation task (BiLSTM + NE + ML), Macro-F1 and Micro-F1 both reach the best scores, which proves that the auxiliary learning in this paper can improve the finance numeral tenses understanding of the model.

Table 5: Results of baseline experiments

Models	Micro-F1	Macro-F1
FNN	**59.59%**	54.88%
TextCNN	56.36%	51.38%
RNN	58.48%	54.87%
BiLSTM	59.30%	**55.21%**

Table 6: Results of ablation experiments

Models	Micro-F1	Macro-F1
FNN (NE)	78.02%	70.41%
TextCNN (NE)	77.14%	72.10%
RNN (NE)	**81.40%**	76.60%
BiLSTM (NE)	80.64%	**77.11%**
BiLSTM (NE+ML)	**83.27%**	**80.60%**

5.3 Discussions

As listed in Table 5 and Table 6, among the four baseline models, BiLSTM performs the best or close to the best in both Micro-F1 and Macro-F1 metrics, indicating that BiLSTM architecture can effectively capture the tense semantic information in Chinese financial news texts. Without NE, FNN model achieve the best performance on Micro-F1 scores, reaching 59.59%. With NE, TextCNN performs best on Micro-F1 scores, reaching 81.40%. We find the BiLSTM model performances best on Macro-F1, whether the target numeral information is provided or not, which shows that the BiLSTM model architecture is least affected by the uneven data distribution.

We further discuss the experimental results between the BiLSTM (NE) and BiLSTM (NE + ML) models. The Micro-F1 and Macro-F1 scores of BiLSTM (NE + ML) achieve 83.27% and 80.60%, respectively, which are better than the baseline models and the models only with NE. It reveals that the tense operator extraction can improve the performance of the model in target numeral tense detection to a certain extent. Numeral encoding and tense operator extraction tasks can provide different information for the model. In other words, the multitask learning with numeral encoding is suitable for the financial numeral tense detection and can achieve the best performance.

6 CONCLUSIONS

In this paper, we explore the numeral tense issues in Chinese financial news texts, and present our contructed CFinNumTense dataset. We also validate the performances of baseline models via experiments on CFinNumTense dataset. With the comparative and ablation experimental results, we find that learning with the numeral encoding can improve the performance on financial numeral tense detection task. Furthermore, the auxiliary task of tense operator extraction can make BiLSTM with numeral encoding module reach the best performance.

Our constructed CFinNumTense dataset and the numeral tense detection task in Chinese financial news texts will advance and facilitate the numeral tense study in financial documents, such as market analysis reports and financial tweets. It can also provide a new study direction for the stock market prediction. In the future, we plan to explore the event extraction from financial news texts integrating the tense attribute. Furthermore, we will also try to introduce tense information to encoding Chinese financial news texts to conduct the stock market prediction.

ACKNOWLEDGMENTS

The work presented in this paper is supported by the National Key Research and Development Program of China under Grant No.2020AAA0108501.

REFERENCES

[1] Ryo Akita, Akira Yoshihara, Takashi Matsubara, and Kuniaki Uehara. 2016. Deep Learning for Stock Prediction Using Numerical and Textual Information. In *Proceedings of the 15th IEEE/ACIS International Conference on Computer and Information Science (ICIS)*. 1–6. https://doi.org/10.1109/ICIS.2016.7550882

[2] Ehsaneddin Asgari and Hinrich Schütze. 2017. Past, Present, Future: A Computational Investigation of the Typology of Tense in 1000 Languages. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, 113–124. https://doi.org/10.18653/v1/D17-1011

[3] Nathanael Chambers, Shan Wang, and Dan Jurafsky. 2007. Classifying Temporal Relations between Events. In *Proceedings of the 45th annual meeting of the association for computational linguistics companion volume proceedings of the demo and poster sessions*. 173–176. https://aclanthology.org/P07-2044.pdf

[4] Chung-Chi Chen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2020. NumClaim: Investor's Fine-Grained Claim Detection. In *Proceedings of the 29th ACM International Conference on Information* (Virtual Event, Ireland) *(CIKM 20)*. Association for Computing Machinery, New York, NY, USA, 1973–1976. https://doi.org/10.1145/3340531.3412100

[5] Chung-Chi Chen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2021. *Numerals in Financial Narratives*. Springer Singapore, Singapore, 55–71. https://doi.org/10.1007/978-981-16-2881-8_5

[6] Chung-Chi Chen, Hen-Hsen Huang, Hiroya Takamura, and Hsin-Hsi Chen. 2019. Numeracy-600K: Learning Numeracy for Detecting Exaggerated Information in Market Comments. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 6307–6313. https://doi.org/10.18653/v1/P19-1635

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[8] Tingwei Gao, Yueting Chai, and Yi Liu. 2017. Applying Long Short Term Momory Neural Networks for Predicting Stock Closing Price. In *Proceedings of the 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*. 575–578. https://doi.org/10.1109/ICSESS.2017.8342981

[9] Zhiheng Huang, Wei Xu, and Kai Yu. 2015. Bidirectional LSTM-CRF Models for Sequence Tagging. *arXiv preprint arXiv:1508.01991*. https://arxiv.53yu.com/abs/1508.01991

[10] Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, 1746–1751. https://doi.org/10.3115/v1/D14-1181

[11] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*. arXiv:1412.6980 [cs.LG] https://arxiv.53yu.com/pdf/1412.6980.pdf?ref=https://githubhelp.com

[12] Maofu Liu, Xinxin Xia, and Wei Wang. 2021. A Chinese Dataset for Exploring Financial Numeral Attributes. In *Companion Proceedings of the Web Conference 2021*. Association for Computing Machinery, New York, NY, USA, 255–259. https://doi.org/10.1145/3442442.3451374

[13] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692* abs/1907.11692. arXiv:1907.11692 http://arxiv.org/abs/1907.11692

[14] Pisut Oncharoen and Peerapon Vateekul. 2018. Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators. In *Proceedings of the 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*. 19–24. https://doi.org/10.1109/ICAICTA.2018.8541310

[15] James Pustejovsky, Patrick Hanks, Roser Saurí, Andrew See, and Marcia Lazo. 2003. The TimeBank corpus. In *Proceedings of Corpus Linguistics*, Vol. 2003. Lancaster, UK., 40. https://www.researchgate.net/publication/228559081_The_TimeBank_corpus

[16] Kato Ryota and Nagao Tomoharu. 2012. Stock Market Prediction Based on Interrelated Time Series Data. In *Proceedings of the IEEE Symposium on Computers Informatics (ISCI)*. 17–21. https://doi.org/10.1109/ISCI.2012.6222660

[17] Georgios Spithourakis and Sebastian Riedel. 2018. Numeracy for Language Models: Evaluating and Improving their Ability to Predict Numbers. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Melbourne, Australia, 2104–2115. https://doi.org/10.18653/v1/P18-1196

[18] Manuel R. Vargas, Beatriz S. L. P. de Lima, and Alexandre G. Evsukoff. 2017. Deep Learning for Stock Market Prediction from Financial News Articles. In *Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. 60–65. https://doi.org/10.1109/CIVEMSA.2017.7995302

[19] Siddharth Vashishtha, Benjamin Van Durme, and Aaron Steven White. 2019. Fine-Grained Temporal Relation Extraction. *arXiv preprint arXiv:1902.01390* abs/1902.01390. arXiv:1902.01390 http://arxiv.org/abs/1902.01390

[20] Nianwen Xue, Yuchen Zhang, and Yaqin Yang. 2013. Distant Annotation of Chinese Tense and Modality. In *Proceedings of the IWCS 2013 Workshop on Annotation of Modal Meanings in Natural Language (WAMM)*. 47–55. https://doi.org/10.1.1.380.2093

[21] Shu Zhang, Dequan Zheng, Xinchen Hu, and Ming Yang. 2015. Bidirectional Long Short-Term Memory Networks for Relation Classification. In *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation*. Shanghai, China, 73–78. https://aclanthology.org/Y15-1009