Journal of Computers Vol. 32 No. 3, 2021, pp. 150-157
doi:10.3966/199115992021063203011

A Novel Strategy Based on Temporal Date Mining for Improving the Integration of Urban and Rural Public Services and Management

Lun Li*

School of Marxism, Hohai University, Nanjing Jiangsu 211100, China
No. 1 Xikang Road, Nanjing City, Jiangsu Province, people's Republic of China
lilun@hhu.edu.cn

Received 5 December 2020; Revised 15 December 2020; Accepted 16 December 2020

Abstract. Urban and rural public services and management is the key problem for the development of modern society against the urban and rural resource imbalance and counter urbanization. To discover its actual contradiction, the questionnaire is the main and conventional way to survey the current situation. In this paper, we collect the three parts of Jiangsu province's urban and rural public service data, and use the Chinese microblog comment dataset NLPCC2013 to augment the dataset. We give a two-stage framework for the text sentiment analysis of subjective answers and the global answer vector classification. The former is based on a bidirectional gated recurrent unit network for a hierarchical analysis of text sentiment from character embedding to contextualize embedding. It can output a soft label for each sub-sentence, and then locate the potential answers. According to the full scores of the subjective items, the text sentimental labels with higher probabilities are tailed with the answers of the objective items to conduct the isometric inputs of deep forest. Based on the reweighted forests' contributions, the output layer can give a binary classification label for each input questionnaire. Experiment results demonstrate the proposed framework can predict the text sentiment accurately. Meanwhile, we analyze the source of the hot topics, and offer the policy suggestions for the future of urban-rural integration

Keywords: text sentiment, BiGRU, deep forest, Urban and rural public services and management

1 Introduction

Integrating the resources of the urban and rural public service and management is a new development strategy for the modern society. Both for the developing and developed countries, the strategy can furthest rebalance the resources facing to the growing of reverse urbanization and aging. In practice, the questionnaire survey is the common approach for collecting the reaction of the resident. To meet the voting environments and marking machine, most of questions of the surveys are always setting the objective items. The setting is easy to ensure the recovery rate of the questionnaires but not avoid the wild answering. So the combination of the objective and subjective items is the mainstream counterpart for a reasonable survey. However, it bring a series of problems. Different from the objective items, the subjective items are time-consuming for the reading and understanding. For the existing conventional methods, such as BOW [1] and context inferring [2-4], their performance cannot support the analysis of the complex input of text information due to the weaker generalizations of their architectures.

To overcome the difficulties of the generalization, the data-driven methods are successively proposed. For diversified feature extraction, the multi-layer processing is introduced as an automatic feature extractor. For an example of CNN (Convolutional Neural Networks)-like DL (Deep Learning) architecture [5], each multi-layer block could contain the sampling layers, convolutional layers (CL), and

* Corresponding Author

activation layers. The entire architecture cascades a sequence of the multi-layer blocks, and their outputs are connected and flatted with at least one fully-connected layers (FC) ending with a Sigmoid or Softmax function to output a predicted label for each input word vector. Many derived CNN-based end2end solutions are proposed to perfect the performance of word understanding but still have a big challenge of text analysis due to the subtle sentimental variation of the ordered word sequences. Hence, RNNs (Recurrent Neural Networks) is used to organize the outputs of the multi-layer blocks for a sequence analysis by integrating the temporal factor. Based on this methodology, the great strides are occurred in the fields of machine translation, image/video caption, and have almost broken all the bottlenecks in the existing feature engineering based methods. As a successful variant of RNN units, LSTM (Long and Short Term Memory) /GRU cells can relieve the vanishing gradient leading by the overlong chain effect of RNN units. Therefore it is the current dominant method in complex text, image, and scene analysis, despite containing lots of redundant computation.

Besides, decision trees (DTs) and random forests (RFs) based framework is an alternative scheme against the learning-based methods. Due to the difficulty for extracting more discriminative representations, they always tail with the DL-based in their performance comparison. But on some small-scale datasets, GBDTs (Gradient-based Decision Trees)-based ensemble learning (EL) frameworks can gain an equal performance with the SOTA (State-Of-The-Art) DL-based methods in a faster speed.

Both of the above methods are ill-posed for an automatic marking system to understand and score the objective and subjective items of the questionnaires. However, we consider to use a mixture architecture with GBDTs and LSTM units to solve two categories of items, respectively. Thus, an EL framework with a sequence of build-in LSTM cells is proposed to tracking and analyzing the answers in this paper. It aims to eliminate the unqualified questionnaires, while give a text sentiment orientation class. The main contributions of the paper are as follows:

(1) We proposed an attentive BiGRU model to analyzing the Chinese text sentiment, while using the extracted word vectors, character vectors and contextualized vectors to predict the sentimental label for the answers of the subjective items;

(2) We formulate the problem of scoring the subjective answers as a machine reading comprehension process. The sentiments of all the potential answers are input into an inferring framework based on LSTM units to conduct a fixed-sized sentimental representation for cascade structure;

(3) We employ deep forest framework to classify the final answer vectors of the recovered questionnaires, and obtain a considerable progression for policy information mining.

2 Related Works

Corresponding the complexity and connotation of the subjective items, the scoring process seems as a machine reading comprehension (MRC) which consists of many complex tasks, such word representation, semantic understanding, and text sentiment analysis [6].

For word representation, it is aim to take the vectors transformed from the words as the learnable representation, which is considered as a fundamental work for NLP (Neural Language Processing). The conventional methods of word representation uses one-hot annotation for the key word in the given sentence to transform the entire sentence into a 0-1 vector. The methods gives a sparse but high-dimensional representation for the larger input vocabulary. Besides, lacking for words' relations is a critical defect to limit the further applications [7]. The shortcoming had been solved until the emergency of word embedding technologies. More elements are packaged into the vectors named CBOW [1], including numeral, verb, adj., and adv. Skip-gram is proposed to link these elements for a continuous low-dimensional vector. However, Dhingra *et al.* [8] revealed that the semantic differences by revising minor words can misguide the reader. Especially for the verbs, both the words and their characters need be used to pre-train a model for word embedding [2-3, 9-11]. Moreover, Peters *et al.* [12] found the change of contextual information cannot always follow with re-representing of the target words, and proposed ELMo a downstream-based fine tuning method for pre-train a hierarchical language model. Considering the bidirectional effect of the sub-sentence structure, Devlin *et al.* [13] design a transformer named BERT to encode the contextual information to high-level representations.

For the semantic understanding, the end2end Neural Networks (NNs) are the most popular methods for semantic segmentation. All the word vectors are encoded into a vocabulary. The target word vectors are orderly input to a NN-based architecture to classify their semantic labels, and the inferring model is working in its ending part to give a semantic label for a sentence. The attentive mechanism is always used to reweight the word vector to ensure the key word can be selected to organize a backbone of the input sentence. Wang et al. [9] designed a similar model to match multi-perspective context for selecting the significant information for scoring the answers. Their follow-up study [3] added an attentive GRU (Gated Recurrent Units) network to build a bijection between the key paragraphs and the answers of some particular questions. Yu et al. [14] designed multi-pass reasoners for checking the relations among the passages and their questions. Liu et al. [15] improved the inferring results by stochastic dropout. The abovementioned method focus on improving the inferring architecture by neuron units and attention mechanism. But there still have some researchers thinking the extractive model is insufficient, and bring a generative model for an elaborate answer [16].

For text sentiment analysis, the problem is formulated as an ensemble of the word selection and the sentence relation. The main approaches can be roughly divided into two typical methodologies, consists of the rule-based methods and the learning-based method. The front group mines a fixed-sized rule set as the grammar to package the words in a sequence with a given sentiment. Due to the large scale of the sentimental sub-sentence, the rule and the sentimental dictionaries are too difficult to cover all the words of the given corpus. Thus, almost all the SOTA methods are learning-based for preforming a better result.

3 Materials and Methodology

3.1 Data

In the paper, we collect data with intercept interviews and questionnaire surveys. First, we visit leaders, experts, and grassroots cadres inside and outside the province to consult on the integration of urban and rural public services and management. In terms of the interview implementation, according to the research objectives, the research team divided the respondents into two groups according to the nature of their work and the different living areas. Among them, the functional objects are divided into two groups: citizens and farmers. It is expected that the horizontal and vertical cross-comparison can obtain the understanding and evaluation of grassroots on the integration of urban and rural public service and management from multiple angles. The staff is mainly from government agencies or institutions engaged in medical and health and social security-related work in Jiangsu province. Second, the survey is divided into three regions, namely northern Jiangsu, central Jiangsu, and southern Jiangsu, to understand the status quo and development trend of integrating public service and management in urban and rural areas and various regions. Huai'an was taken as the research site in northern Jiangsu, And Taizhou and Nantong were taken as the research sites in central Jiangsu. Changzhou and Nanjing were selected as the research sites in the southern part of Jiangsu province to comprehensively understand the achievements and existing problems of the urban-rural integration of Jiangsu Province.

In the aspect of questionnaire design, the questionnaire on the Integration of Urban and rural Public Service and Management in Jiangsu is divided into three parts, namely, the basic information of the respondents, their understanding of urban and rural medical construction, and their understanding of urban and rural social security construction. The paper starts with the current urban and rural public service and management integration situation to analyze the existing problems from multiple levels and dimensions. The problems' crux and the related constructive suggestions are put forward to provide intellectual support for the urban and rural integration process's in-depth exploration.

A total of 360 questionnaires were issued, and 350 were effectively recovered, with an effective rate of 97.22%. Among them, 155 questionnaires were sent to rural residents, and 152 were collected effectively, with an effective rate of 98.06%. One hundred questionnaires were issued to urban residents, and 96 valid questionnaires were collected, with an effective rate of 96%.In the survey, 105 questionnaires were issued by government staff, and 102 valid questionnaires were collected, with an effective rate of 97.14%. We also survey the demography for urban and rural residents, containing gender, age, profession, household income per capita, and education. The distributions are illustrated in Fig. 1.

Data annotation: At the beginning of data annotation, we use the results of the demographic survey to cluster the questionnaire dataset corresponding to the subjects with multi-condition. For the self-collected corpus, we segment 3576 sub-sentences from the subjective answers for sentiment analysis. Three experts give a sentimental label for each input sub-sentence.

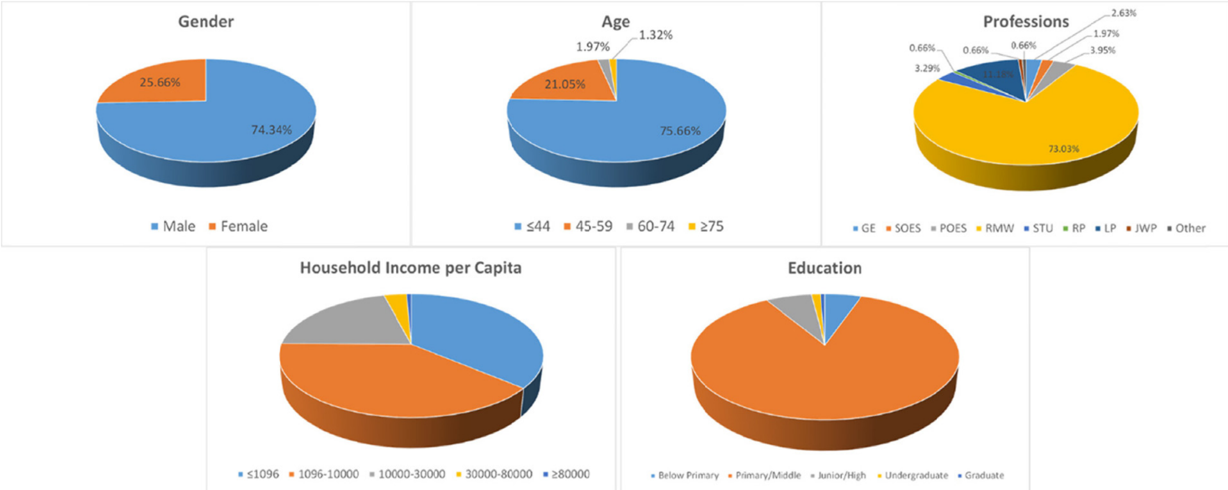

Fig. 1. The distribution of the demographic survey for gender, age, profession, household income per capita, and education

Data augmentation: To rebalance the ratio and improve the model's generalization, we merged a Chinese microblog corpus in a sentiment classification with the self-collected corpus. The Chinese microblog corpus was collected for the sentiment task in NLPCC2013, which the 13049 sub-sentences were extracted from 4000 posts as the samples to be classified.

Data cleaning: A data cleaning mechanism is executed to eliminate the counterfactual and arbitrary answers. It is a DT-based method to classify the cascade answer vectors of the objective and subjective items.

3.2 The Proposed Framework

Contextualized embedding encoder: Following the ELMo-based embedding approach, we first used a skip-gram (shown in Fig. 2) to compute a vector representation for each word. Each word vector is independent with its context, and pre-trained in a BiLSTM (Bidirectional Long Short Term Memory) network with residual connecting layer. For a given sentence of N words, the outputs of the head sub-network in Fig. 3 can be encoded with the forward and backward LSTM layers, which are denoted as $\vec{h}_n^i$ and $\overleftarrow{h}_n^i$, $i = 1, ..., l, \ldots, L$. Thus, the context-dependent representation of the n-th word can be written as

$$R_n = \{x_n, h_n\} = \{x_n, \vec{h}_n^i, \overleftarrow{h}_n^i\}. \tag{1}$$

where x_n is the n-th word context-independent representation. For the L layers of BiLSTM, a linear combination is used to package the outputs of all the layers into a vector. A parameter set of r_n is weighting the contextualized information by

$$x_n^{\varepsilon} = r_n \sum_{i=0}^{L} s_i(w_n) \cdot h_n \tag{2}$$

where s_i is a softmax function to output a soft label for the input word. According to its context information, contextualized embedding is complete.

Word and character embedding encoder: A convolutional neural network is running for each word to transform the characters of each word into a 64-dimension vector. The pre-trained BiLSTM network is used to assign the words into a 350-dimension vector, because most of the sub-sentences has less than 250 word, occupying 98.4% of the corpus. Thus, we design a length of the 350 LSTM units with 512 hidden neurons ending with a ReLU activator, to finish word embedding. Integrating the output of contextualized embedding, a newborn linear vector $\{w_n^c, w_n^w, w_n^{\varepsilon}\}$ is employed to update the representation of the sub-sentence. Based on a bidirectional GRU, the inferring results can be expressed as $\vec{h}_t^i$ and $\overleftarrow{h}_t^i$ for the t-th sub-sentence.

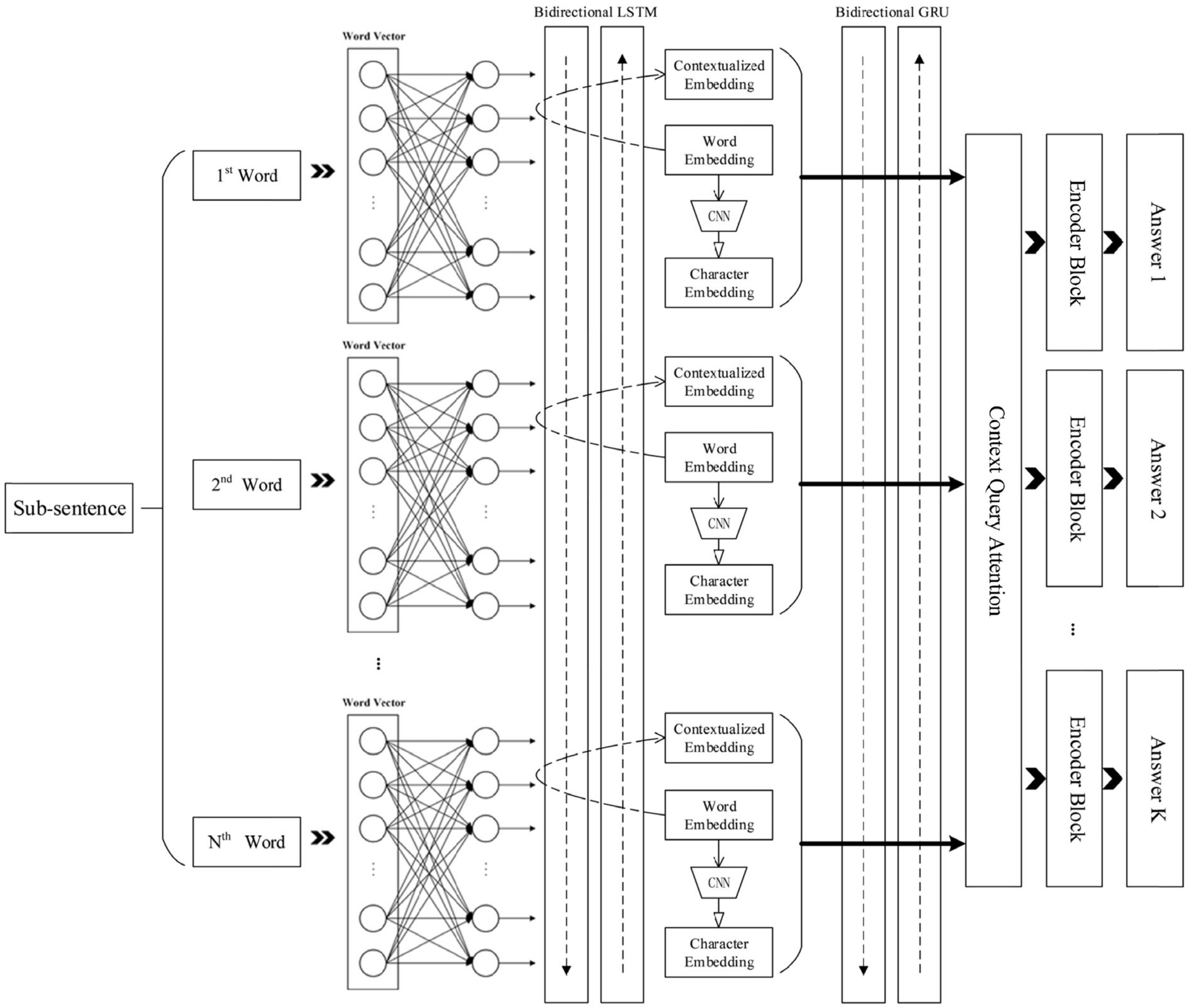

Fig. 2. The framework of text sentiment analysis for the answers of the subjective items

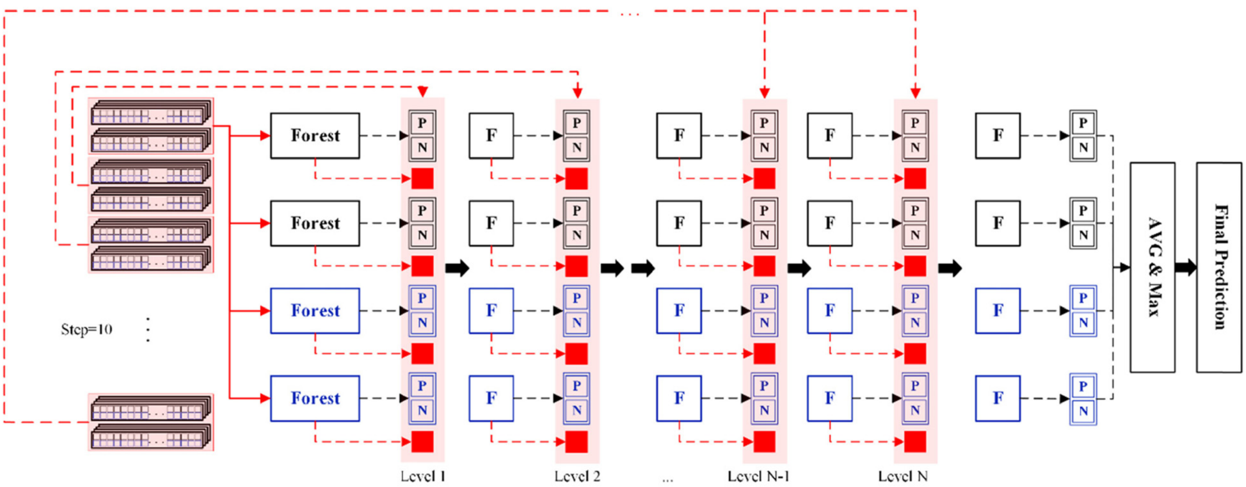

Fig. 3. Schematic diagram of the level-wise cascade structure: the left vectors cover the normalized answers of the objective and subjective items; the red squares are the variance of the top-5 features of the RFs to estimate the contributions of the RFs in the voting step

We analyze the relations of the context and the query of the sub-sentence with a similarity matrix of $C \times Q$. Thus, the attention can be computed as a cosine distance of $s(c_i, q_j)$ for the i-th contextualized vector and the *j*-th query vector to find a reasonable answer. After reweighting the encoders, we can predict the start and end points of the subjective answers, which are not unique for a given question. Furthermore, they are fed into a series of 50- dimension LSTM units to output the final sentimental labels of the *k*-th answer by minimizing the sum of the negative log probability:

$$L = \frac{1}{N}\sum_{n=1}^{N}[\sum_{k}\log(p_k^b) + \sum_{k}\log(p_k^e)] \tag{3}$$

where p_k^b and p_k^e is the predicted positions of the starting and ending words in the *k*-th answer. Meanwhile, we also cascade the outputs of *T* answers for predicting a successive sentimental outputs.

Multi-grained scanning: After the inferring process, we gain a series output of the answers' sentiments. We concatenate the outputs with the answers of the objective items. Due to the fixed-sized LSTM units, the linked answer vectors are isometric. Two random forests (RFs) and complete random forests (CRFs) are paralleled in a cascade level of deep forest framework to classify the sentiments of the questionnaires [7]. A sliding windows with a step of 10 is used to sample the inputs for all the forests.

Level-wise cascade structure: We follow the level-wise growing strategy for the cascade structure. The training data is poured into the sampling layer to conduct the multi-grained features for the inputs of all the forests. We set the reciprocals of standard deviations of the top-5 features give a normalized weight to the different forests. The cascade is growing until the predicted result on the training data is less than that of the validation data [7, 9].

4 Experiments

4.1 Evaluation Metrics

In our experiments, we employ CPM and F1 score to evaluate the performance of the proposed framework, in which F1 score is a good evaluation metrics for the detection tasks. Meanwhile, Competition Performance Metric (CPM) [17] is an average value of the seven operating points at Free Receiver Operating Characteristic (FROC) curve, which can further estimate the robustness for the detection methods.

4.2 Results

At first, we report a detecting results to classify the passages with the apparent positive/negative sentiment in Table I. In the comparison, we select the four mainstream methods for sentiment analysis. The attentive LSTM is the most popular schemes for text analysis and image understanding. The attention mechanism is working on the word embedding vectors. A sparse counterpart of the original sentence is conducted by the high-weighted words. Different from attention-based methods, XGboost is an open source project of Microsoft, which offers a light EL (Ensemble Learning)-based solution for the small-scale data classification. It pre-sorts all the extracted feature from the input passages, and orderly sockets them into some decision trees. Obviously, both of them employ LSTM units to drive their inferring, and drop the RNN-based scheme to avoid the performance degeneration due to the longer chains sparse. XGboost performs worse than attention in the comparison, because the outputs of the LSTM units only gives the few information for the feature extraction of XGboost based on the sliding windows, leading to reduce the impact of EL-based method for the final classification task. GRU is a variant of LSTM unit with two gated controllers, which is easier to build a larger-scale network. Under the default setting, it shows a better F1 score than LSTM units.

Table 1. F1 scores of text sentiment analysis among the SOTA methods and ours

Schemes	All	Positive	Negative
LSTM+Attention	0.951±0.002	0.967±0.001	0.836±0.004
LSTM+XGboost	0.943±0.006	0.961±0.006	0.792±0.017
GRU+Attention	0.962±0.007	0.974±0.004	0.837±0.015
BiGRU+ Attention +DFs	0.977±0.004	0.982±0.003	0.921±0.001

The results of the proposed scheme are listed in the bottom row of Table I. It gives a satisfying performance both for the positive and negative samples. Besides, we find our scheme has lower variance, which benefits from the context-query attention mechanism and multi-grained scanning. The former give the larger weight for the key context and query in the given passage and the subjective answer. The similarity matrix effectively filters the answer candidates. The later offers multi-scale perceptive fields for inferring the text sentiment change.

Subsequently, we give a quantitative comparison based on FROC and CPM. The seven operating points on FROC curves are 1/8, 1/4, 1/2, 1, 2, 4, 8 FPp (False positives per Passage). To simplify the experiment, we select two methods, which show better performance in Table. I. Seen from Table. II, we can found that our scheme occupies in a leading position at the seven operating points. Especially after 1/2 FP/p, the sensitivities show a rocket trend. It demonstrate the proposed method cannot only achieve a satisfactory results for scoring subjective items, but also can use deep forest to classification the answers with the different text sentiment.

Table 2. CPM scores of text sentiment analysis among the SOTA methods and ours

Schemes	GRU+Attention	BiGRU+DF (Ours)
1/8 FP/p	0.814	0.827
1/4 FP/p	0.869	0.875
1/2 FP/p	0.915	0.926
1 FP/p	0.926	0.964
2 FP/p	0.942	0.972
4 FP/p	0.951	0.980
8 FP/p	0.957	0.987
CPM	0.911	0.933

4.3 Solutions

In the ending of the paper, we use text sentiment analysis to mine the hot topics of this questionnaire survey. Urban-rural integration is the hottest sub-sentence of the self-collected corpus, which come out from 92.1% of rural migrant workers. To solve the problem, we suggest clarifying the functions of urban and rural areas. With the metropolis of Nanjing, SuXiChang and Xuzhou, the coordinated development of large, medium and small cities and key central towns should be promoted. Endowment insurance is another hot topic, which be mainly collected from the private-owned enterprise staffs, rural migrant workers, and others. We note these social group lack for a stable income while facing to the self-endowment pressure. Aiming to the problem, we should construct the four security lines of the new rural cooperative medical care system and the rural endowment insurance system, and create conditions for the integration of urban and rural social security system.

5 Conclusions

This paper proposes a two-stage framework with BiGRU (Bidirectional Gated Recurrent Units) and DF (Deep Forest) for automatic scoring the questionnaires. Considering the absent of Chinese comment dataset, we collect 350 questionnaires, and segment words and sub-sentences from their passages to form a comment corpus like NLPCC 2013. By using a divide and conquer methodology, the BiGRU sub-network is used to analyze the input text sentiment, which can be formulated as machine reading comprehension. We embed character, word, contextualized information into a learnable representation, and word embedding model is pre-trained for a better performance. Based a simple CNN architecture the characters of all the words are embedded into the 64- dimensional vectors. Contextualized information are training based on ELMo to output a soft label for each answers. The outputs of the starting and ending positions are varying due to the different combination between the elements from the starting and ending point set, respectively. From global perspective, according to the score values of subjective items, we fulfill a fix-sized answer vector and feed them into a deep decision framework to predict the final label. Both in the comparisons of F1 scores and CPM scores, the proposed framework gains the desired results, respectively reaches 0.977 and 0.933. In our future work, we plan to introduce the leaf-wise cascade for the original gcforest to boost the computation speed and improve the detecting performance. Meanwhile, we trackback the hot topics and analyze their main source, and give a policy suggestion for the future of urban-rural integration.

References

[1] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Proc. Advances in Neural Information Processing Systems, 2013.

[2] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323(6088)(1986) 533-536.

[3] M. Seo, A. Kembhavi, A. Farhadi, H. Hajishirzi, Bidirectional attention flow for machine comprehension. <https://arxiv.org/abs/1611.01603>, 2016.

[4] W. Wang, N. Yang, F. Wei, B. Chang, M. Zhou, Gated self-matching networks for reading comprehension and question answering, in: Proc. 55th Annual Meeting of the Association for Computational Linguistics, vol. 1, 2017.

[5] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature 521(7553)(2015) 436-444.

[6] K.M. Hermann, T. Kocisky, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, P. Blunsom, Teaching machines to read and comprehend, in: Proc. Advances in Neural Information Processing Systems, 2015.

[7] R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval, Reading, MA, USA: Addison-Wesley, 2011.

[8] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323(6088)(1986) 533-536.

[9] Z. Wang, H. Mi, W. Hamza, R. Florian, Multi-perspective context matching for machine comprehension. <https://arxiv.org/abs/1612.04211>, 2016.

[10] D. Weissenborn, G. Wiese, L. Seiffe, Fastqa: A simple and efficient neural architecture for question answering. <https://arxiv.org/abs/1703.04816v1>, 2017.

[11] M. Hu, Y. Peng, Z. Huang, X. Qiu, F. Wei, M. Zhou, Reinforced mnemonic reader for machine reading comprehension. <https://arxiv.org/abs/1705.02798>, 2017.

[12] M. E. Peters, M. Neumann, M. Lyyer, M. Gardner, C. Clark, K. Lee, L. Zettlemoyer, Deep contextualized word representations. <https://arxiv.org/abs/1802.05365>, 2018.

[13] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding. <https://arxiv.org/abs/1810.04805v1>, 2018.

[14] A.W. Yu, D. Dohan, M.-T. Luong, R. Zhao, K. Chen, M. Norouzi, Q.V. Le, QANet: Combining local convolution with global self-attention for reading comprehension. <https://arxiv.org/abs/1804.09541>, 2018.

[15] X. Liu, Y. Shen, K. Duh, J. Gao, Stochastic answer networks for machine reading comprehension. <https://arxiv.org/abs/1712.03556v1>, 2017.

[16] C. Tan, F. Wei, N. Yang, B. Du, W. Lv, M. Zhou, S-NET: From answer extraction to answer generation for machine reading comprehension. <https://arxiv.org/abs/1706.04815v5>, 2017.

[17] M. Niemeijer, M. Loog, M.D. Abramoff, M.A. Viergever, M. Prokop, B. van Ginneken, On combining computer-aided detection systems, IEEE Transactions on Medical Imaging 30(2)(2011) 215-223.