



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 3, March 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijrset@gmail.com

www.ijrset.com

Precision Farming Insights: Innovating Agricultural Text Classification through Dynamic Feature Fusion

V.S. Gayathri¹, S. Rohith Kumar², J. Naga Lakshmi³, V. Bhuvaneshwari⁴, S. Abhishek⁵

¹Assistant Professor, Dept. of AI, Annamacharya Institute of Technology and Sciences(AITS), Karakambadi, Tirupati, Andhra Pradesh, India

^{2,3,4,5}Student, Dept. of AI, Annamacharya Institute of Technology and Sciences (AITS),Karakambadi, Tirupati, Andhra Pradesh, India

ABSTRACT: Traditional text classification methods treat agricultural text values as characters, losing the semantic expression of numerical features. To fully mine the deep latent semantics in agricultural text, we propose a novel text classification method based on dynamic fusion of multivariate features. We use a bidirectional LSTM (Bi-LSTM) model with attention to extract global key text semantics. Multiple window convolutional neural networks (CNNs) extract local semantics at different levels. Numerical values expressing essential semantics are manually extracted into vectors. An attention mechanism then dynamically fuses all these features, enriching agricultural text semantic expression and improving phenotypic numeric text classification performance. In summary, this agriculture-focused approach combines Bi-LSTM, CNNs, attention, and numeric feature engineering to create a robust text classifier attuned to the unique characteristics and semantics of the agricultural domain.

KEYWORDS:Agriculture, Text Classification, Multivariate Convolutional Neural Networks,Semantics, Gradient Boosting Classifier.

1. INTRODUCTION

A vast number of agricultural technology websites, resource databases, and information platforms have surfaced in recent years due to the field's rapid expansion. agricultural technology websites, resource databases and information platforms have emerged. Text as the most common form of data carrying, contains a huge amount of information. The second aim of the text classification in the structural information is to conclude awarely and with a high level of confidence on the commercial data resources. Consequently, text classification is considered a huge subject including agricultural text classification in particular[1]. For example, text extraction can assist tasks like the arable farming production and research, taking out of conscise and loud agricultural text data the useful information. The heat tolerance, the vocable that represents as a theme in agricultural texts, is used to group together the inherent agronomic characters of wheat as described in agricultural texts. In view of the lack of agricultural corpus, this project uses Python to crawl the national audit wheat feature text, constructs the wheat text corpus, and classifies the text according to whether the wheat has cold resistance Recent Works. The present research suggests a multi-feature dynamic fusion technique for text classification related to agriculture. [2] one of the classic tasks of Bi-LSTM and MUL-CNN have achieved good results in the task of text classification in different fields. Whereas MUL-CNNconcentrates on the local feature information of the text in the convolution window, Bi-LSTM records the global feature representation of the text in the sequence. The combination of the two can obtain the comprehensive feature representation information of the text globally and locally, thus ensuring the effect of classification.

II. RECENT WORKS

[3] Direct marketing is necessary for many businesses in the credit card, insurance, bank, and retail industries. Institutions can develop marketing objectives with the aid of data mining. Target audiences are a natural fit for data mining tools, which also increase response rates. The Naive Bayes and C4.5 decision tree algorithms are two data mining strategies that have looked into in this paper. This effort aims to forecast a client's likelihood of subscribing to a term deposit. And also conducted a performance comparison between the two algorithms. The algorithms are trained and tested using publicly available UCI data. Additionally, extract useful information from decision trees that concentrates on making significant and intriguing company decisions.The fusion of CNN and Bi-LSTM allows for accurate representation of key points discussed. Specifically, the Bi-LSTM network captures sequential dependencies in the agricultural text, considering both past and future context. Meanwhile, the mul-CNN extracts local features, identifying relevant patterns within the text. The combination of these two architectures enriches the semantic expression of the agricultural data, enabling more precise classification.

[4] Neural network-based methods have demonstrated amazing promise in phrase modeling in recent years. Recurrent neural networks (RNNs) and convolution neural networks (CNNs) are the two primary types of neural network architectures. Long-term dependencies can be captured by RNN, and it can store the semantics of earlier data in a fixed-sized vector. But the fixed-sized vector limits the RNN's capacity to extract global semantics, making it a biased model. As an alternative, CNN can use convolutional filters to extract text's n-gram properties. However, convolutional filters' width limits how well they work. This research offers an multi Scale Convolutional Neural Network (Mul-CNN) for text categorization, which aims to ameliorate the weaknesses of the two types of networks while combining their capabilities.

[5] In order to solve issues with long text classification, particularly the existence of redundant and diverse information, the research presents a unique text classification model called attention-based BiLSTM fused CNN with gating mechanism (ABLG-CNN). The model uses word2vec for word vector representation, attention mechanism for keyword extraction, and BiLSTM for contextual feature capture. Convolutional neural networks (CNNs) are also used for topic salient feature identification. A gating mechanism is developed to weigh the output characteristics of BiLSTM and CNN, aiding the fusion of text aspects conducive to classification, in order to address the possibility of off-topic phrases being included in large texts. This all-inclusive method seeks to improve text categorization performance, especially when handling long and densely informational textual input.

III. PROPOSED WORK

Figure 1 shows that to develop an effective computational model for identifying disaster-related information by synergistically integrating features from visual and textual modalities.

The primary contributions of our work are:

1. Bi-LSTM-ATT model captures global text features with attention to semantic keywords, enhancing semantic representation.
2. Mul-CNN extracts local features from agricultural text sequences using multi-scale convolutional kernels.
3. Numerical values in agricultural text are separately extracted and encoded to construct phenotypic feature vectors.
4. An attention mechanism dynamically combines local and global features along with numerical vectors to enhance accuracy in agricultural text classification.

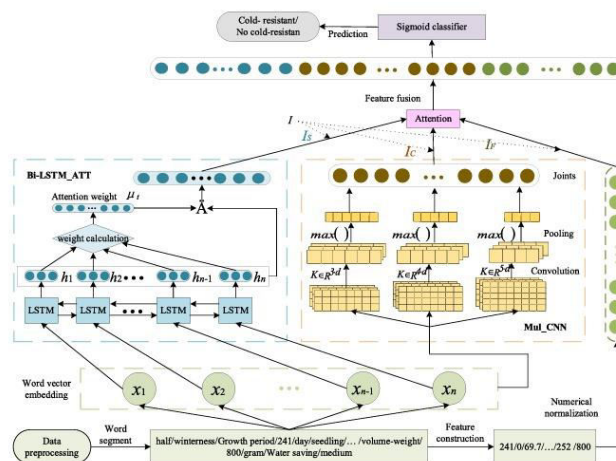


Fig 1: Agricultural text classification model based on multiple feature fusion.

IV. RESULTS AND DISCUSSION

The training set trains the model, the test set assesses its performance, and the validation set fine-tunes network parameters. Accuracy and Kappa coefficient are higher when CNN and Bi-LSTM are combined than when they are used alone. Mul_CNN + Bi-LSTM, for example, performs 88.34% better than Mul_CNN and 87.70% better than Bi-LSTM.

Bi-LSTM-ATT, which adds an attention mechanism, improves accuracy by 3.28% and the Kappa coefficient by 0.0746. By combining semantic and phenotypic variables, the Mul_CNN + Bi-LSTM_ATT demonstrates enhanced agricultural text categorization with an accuracy of 90.15%.

V.CONCLUSION

Texts through highlighting the dynamic mixture of numerous variables to appropriately constitute the complex combination of phenotypic numerical features discovered in agricultural texts. We can extract richer neighborhood semantic capabilities from texts with varying tiers of wheat cold tolerance through using convolution kernels of various sizes, which gives us a more nuanced view of agricultural facts.

Furthermore, MUL-CNN mixed with an interest-fusion mechanism improves our model's potential to recognize crucial semantic phrases in worldwide contexts, improving our understanding of agricultural texts and permitting more unique type.

Moreover, our method can be used to create separate numerical features from wheat trait texts which have numerical expressions, enlarging the function space and facilitating thorough exam of agricultural statistics. The multi-characteristic dynamic fusion method is similarly stronger by the addition of an attention mechanism, which dynamically modifies weights to appropriately and adaptively seize important records pertinent to textual content categorization throughout more than a few feature tiers.

Furthermore, our advised approach suggests promise in figuring out ability semantic elements hidden in agricultural texts, mainly those concerning phenotypic values, which might be crucial to agricultural choice-making and studies. In order to improve the accuracy of wheat bloodless resistance textual content class and reputation, we intend to extend the dataset and practice pattern equalization procedures inside the destiny. Our method has the ability to substantially advance agricultural textual content category strategies with in addition development and alertness, improving the understanding and alertness of agricultural facts in numerous fields inclusive of crop control, genetic analysis, and agricultural policy-making.

REFERENCES

- [1] F.-H. Tseng, H.-H. Cho, and H.-T. Wu, "Applying big data for intelligent agriculture-based crop selection analysis," *IEEE Access*, vol. 7, pp. 116965–116974, 2019
- [2] M. Zhao, C. C. Dong, Q. X. Dong, and Y. Chen, "Question classification of tomato pests and diseases question answering system based on BIGRU," *Trans. Chin. Soc. Agric. Machinery*, vol. 49, no. 5, pp. 271–276, May 2018.
- [3] Y. C. Zhou, T. Y. Xu, and H. B. Deng, "Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils," *Jiangsu Agric. Sci.*, vol. 46, no. 17, pp. 219–223, Sep. 2018.
- [4] Z. Liu, H. Huang, C. Lu, and S. Lyu, "Multichannel CNN with attention for text classification," 2020, arXiv:2006.16174.
- [5] J. Deng, L. Cheng, and Z. Wang, "Attention-based BiLSTM fused CNN with gating mechanism model for Chinese long text classification," *Comput. Speech Lang.*, vol. 68, Jul. 2021, Art. no. 101182.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details