Journal of Industrial Engineering and Applied Science
Journal Home: jieas.suaspress.org
Vol. 2, No. 4, 2024 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

VGCN: An Enhanced Graph Convolutional Network Model for Text Classification

REN, Zheng [1*]

[1] Georgia Institute of Technology, USA

**** REN, Zheng is the corresponding author, E-mail: superpy001@gmail.com***

Abstract: The rapid growth of textual data has made text classification a critical research area in the field of Natural Language Processing. Traditional Graph Convolutional Networks (GCN) face challenges in text classification due to the limitations of fixed edge weights and the inability to explore high-dimensional graph structural features, which hampers effective semantic information capture. To address these issues, we propose an improved graph convolutional network model, Variational Graph Convolutional Network (VGCN). In VGCN, we design a variational structure within the GCN to learn latent representations of textual data, forming a continuous, high-dimensional graph structure distribution. This feature representation captures complex graph structural features, thereby enhancing semantic information extraction and improving classification performance. Experimental results on datasets such as THUCNews and Toutiao demonstrate that VGCN significantly outperforms traditional models like CNN, RNN, and GCN in text classification tasks. These findings underscore the effectiveness and superiority of VGCN in capturing semantic information in texts.

Keywords: Graph Convolutional Neural Networks, Text Classification, Variational Structure.

DOI: https://doi.org/10.5281/zenodo.13118430 **ARK:** https://n2t.net/ark:/40704/JIEAS.v2n4a16

1 INTRODUCTION

Text classification, as a pivotal research task within the field of Natural Language Processing (NLP) [1], has found widespread applications [2] in sentiment analysis [3], news categorization [4], and intelligent question-answering systems [5]. Traditional text classification methods often require manual feature extraction and struggle to capture the complex relationships between words, which limits the performance of the models. However, the advent of deep learning has led to the development of neural network-based approaches that automatically learn text representations and extract features, achieving commendable results. Yet, these methods face challenges in preserving semantic features of text when transforming it into word vector matrices [6].

Graph Convolutional Networks (GCN) have demonstrated their capability to effectively capture the relationships between nodes within a graph structure [7], constructing semantic relationship graphs between words or documents, and learning node representations for end-to-end processing on graph data. This enables effective preservation of textual semantic features, leading to GCNs being widely applied in tasks like node classification and link prediction within graph data.

Despite the growing popularity of GCN-based text classification methods, they face significant challenges. Firstly, the use of fixed edge weights restricts the model's ability to adapt to changes in semantic relationships between different texts, thereby limiting its expressiveness in capturing textual semantic information. Secondly, traditional GCN models fail to fully exploit the high-dimensional graph structural features of textual data, hindering effective encoding of textual semantic information into the graph structure, and thus limiting further performance enhancements [8].

To address these challenges, this paper introduces an innovative text classification model named Variational Graph Convolutional Network (VGCN), which incorporates a variational structure and a dynamic edge weight learning mechanism. This model not only learns the latent representations of text data but also forms a continuous, high-dimensional graph structure distribution, significantly enhancing the semantic expression of textual data.

To validate the effectiveness of VGCN, we conducted evaluations on common text classification datasets such as THUCNews, Toutiao, and Sougou, comparing it against other models like TextCNN, TextRNN, and TextGCN. The experimental results demonstrate that VGCN achieves substantial improvements in text classification tasks, effectively confirming its superiority and effectiveness in capturing semantic information in texts.

The principal contributions of this paper are as follows:

Journal of Industrial Engineering and Applied Science
Journal Home: jieas.suaspress.org
Vol. 2, No. 4, 2024 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

We propose the Variational Graph Convolutional Network (VGCN) model, which incorporates a variational structure to learn latent representations of textual data. By forming a continuous, high-dimensional graph structure distribution, the VGCN model effectively learns complex graph structural features, thereby enhancing the performance of text classification.

We evaluate the VGCN model on text classification datasets including THUCNews and Toutiao. The experimental results demonstrate that the VGCN model achieves significant performance improvements in text classification tasks. Compared to other classical models, VGCN exhibits higher accuracy and generalization capabilities, confirming its superiority and effectiveness in capturing semantic information from texts.

2 RELATED WORK

The field of text classification can be chronologically segmented into three main phases in the context of its evolution: rule-based classification, machine learning models, and deep learning models — the latter being the current mainstream research direction.

In the rule-based classification era, the core approach was to utilize predefined rules. This method heavily relied on domain experts or manually defined sets of rules and features, which, while effective in specific contexts, often lacked the flexibility to adapt to varied textual data and could not capture complex word relationships or deep semantic features.

With the advent of machine learning, text classification methods such as Naive Bayes [9], Support Vector Machines (SVM) [10], and K-Nearest Neighbors (KNN) [11] became popular. These techniques learn features and patterns from labeled text data, but they typically require manual feature extraction and struggle to capture the intricate relationships between words, which limits their ability to mine semantic features within texts, thereby constraining their classification performance.

The advent of deep learning ushered in a significant shift, with neural network structures becoming capable of effectively extracting textual features. Among these, Convolutional Neural Networks (CNN) have been widely used. For instance, the TextCNN model proposed by Kim et al. [12] captures local features through convolution operations but faces limitations when dealing with longer texts due to its inability to capture long-range sequence information. To address these issues, Recurrent Neural Networks (RNN) and Long Short-Term Memory networks (LSTM) [13] gained attention. Liu et al.'s [14] TextRNN can capture long-distance dependencies in text, mitigating some of the drawbacks found in CNNs. However, it still faces challenges in capturing interactions between distant and non-sequential words. Subsequently, Johnson et al. [15] introduced the DPCNN in 2017, employing a pyramid-like convolution structure that stacks multiple convolution and pooling layers to gradually reduce sequence length while effectively capturing global semantic information, thus improving the handling of long texts. Teng Jinbao [16] and colleagues proposed utilizing a multi-channel attention mechanism combined with CNN and LSTM to highlight the importance of different words within a text, although these models have limited impact on improving text classification as they often prioritize local textual features over global relationships and semantic relevance.

In recent years, Graph Neural Networks have made significant strides in the domain of text classification. Yao et al.'s [8] TextGCN model constructs an entire corpus as a heterogeneous graph and utilizes graph neural networks to jointly learn embeddings for words and documents, transforming text classification into a node classification problem. This approach effectively learns text representations but struggles to encode textual semantic information into the graph structure and fully explore high-dimensional graph structural features. Following this, Liu et al. [17] in 2019 proposed the TensorGCN, which constructs a graph representation of textual data and applies graph convolutional operations to learn node representations. By using tensor representations to depict the complex interactions between a node and its neighbors, TensorGCN captures higher-order semantic relations between nodes. Ding et al. [18] enhanced text understanding by transforming texts into hypergraphs, yet still faced challenges in encoding textual semantic information into the graph structure and adequately considering contextual relationships.

3 ALGORITHM AND MODEL

3.1 OVERALL OVERVIEW OF VGCN

The input layer of VGCN processes pre-constructed graph structures from the text data, including feature matrices, adjacency matrices, and labels. The graph convolutional layer takes the preprocessed word-text heterogeneous graph as input. The weighted feature matrix is fed into the variational structure to learn low-dimensional latent representations of the text data, reconstruct the graph structure in the latent space, handle complexity, and learn a continuous high-dimensional graph structure distribution. This involves non-linear transformation and dimensionality reduction of features, resulting in more abstract and higher-level feature representations, enhancing the model's understanding and performance.

The processed feature matrix is then input into the graph convolutional network for graph convolution operations. During forward propagation, information is passed and features are updated using the adjacency matrix and node features. Each node's representation is updated by aggregating its neighbors' features, creating new feature representations that combine local node information with its neighbors', enabling end-to-end learning on the graph structure and better capturing the semantic information of the

Journal of Industrial Engineering and Applied Science
Journal Home: jieas.suaspress.org
Vol. 2, No. 4, 2024 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

text data.

The fully connected layer processes the output features of the graph convolutional layer through linear and non-linear transformations to further extract and learn high-level feature representations. The output layer uses the Softmax function to normalize the output from the fully connected layer, producing a prediction probability distribution for each class. The loss function is calculated based on the prediction probabilities, and the model parameters are updated using the backpropagation algorithm to minimize the loss function, completing the model training process.

3.2 Variational structure

The objective of the variational structure in our method is to learn the latent representations of data and generate new samples that are similar to the original data. The core concept involves mapping the input data to mean and variance in the latent space and then generating new samples from this space through sampling.

In our approach, the graph structure with dynamically updated edge weights serves as the input (denoted as A) to the variational structure. This graph structure is then mapped into a low-dimensional latent space. The encoder, typically consisting of multiple fully connected layers, is responsible for mapping the nodes of the input text feature graph to the mean and variance vectors in the latent space. The outputs of the encoder represent the distribution parameters of the data in the latent space.

Sampling a latent sample from the output mean and variance utilizes the reparameterization trick to ensure the sampling process is differentiable. This involves sampling a random variable from a standard normal distribution, which is then linearly transformed to obtain the latent variable.

The latent sample is then remapped back into the original data space to reconstruct a high-dimensional graph structure of the sample. The reconstruction process focuses on minimizing the reconstruction error, which is pivotal for learning the features of the data.

By minimizing the reconstruction error, the model effectively learns to capture and recreate the underlying data characteristics, enhancing the overall performance of the text classification task.

4 EXPERIMENTS

4.1 Datasets

The THUCNews dataset, utilized in our study, is derived from a collection of approximately 740,000 news documents sourced from Sina News, from which 100,000 news items were extracted. These items span 10 categories, ensuring a diverse representation of content. The categories include economics, real estate, stocks, education, science, society, politics, sports, gaming, and entertainment, with each category comprising 20,000 documents. The dataset is structured into training, validation, and testing subsets, divided according to an 8:1:1 ratio, facilitating a comprehensive evaluation of our text classification models.

The Toutiao dataset, utilized for our experiments, comprises approximately 380,000 news articles collected from the popular Chinese news platform Toutiao. From this pool, 150,000 articles were selected, representing a wide array of 15 categories. These categories include public welfare, culture, entertainment, sports, finance, real estate, automobiles, education, technology, military, tourism, international news, securities, agriculture, and eSports. Each category features short text content. The dataset is systematically divided into training, validation, and testing sets with a distribution ratio of 8:1:1, ensuring a balanced approach to model training and evaluation.

4.2 Evaluation metrics

In our experiments, we evaluate the performance of text classification models using Accuracy (Acc), Precision (Pre), and F1 Score as the primary metrics. Here, TP (True Positives) and TN (True Negatives) represent correctly predicted positive and negative samples, respectively, while FN (False Negatives) and FP (False Positives) represent incorrectly predicted negative and positive samples, respectively.

Accuracy measures the proportion of correct predictions among the total number of cases evaluated and is calculated using the formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+FN} \tag{1}$$

Precision represents the ratio of correctly predicted positive observations to the total predicted positives and is calculated as:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

Recall measures the proportion of actual positives that are correctly identified and is computed as:

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F1 Score is the harmonic mean of Precision and Recall, providing a balance between the two by considering both the Precision and the Recall. It is particularly useful when the class distribution is uneven. The F1 Score is calculated as follows:

$$F1 = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{4}$$

These metrics collectively provide a comprehensive assessment of model performance, considering both the accuracy of positive predictions and the rate at which positives are correctly identified.

Journal of Industrial Engineering and Applied Science
Journal Home: jieas.suaspress.org
Vol. 2, No. 4, 2024 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

4.3 Results

To validate the effectiveness of the proposed Variational Graph Convolutional Network (VGCN), we compared it against a range of widely used and representative models in the field of text classification. These models include TextCNN, TextRNN, TextRNN_Att, TextRCNN, DPCNN, FastText, TextGCN, and TensorGCN.

As depicted in Table 1, VGCN outperforms these established models on the THUCNews dataset across three crucial metrics: accuracy, precision, and F1 score. Specifically, VGCN achieved scores of 92.79%, 93.00%, and 92.81%, respectively, marking the highest values among the models tested. These results highlight the benefits of integrating a variational structure, which not only enables the model to learn latent representations of text but also facilitates the formation of a continuous, high-dimensional graph structure distribution. This advanced distribution significantly enhances the model's capability to capture deep textual features, thereby improving classification performance effectively.

Table 1: The experimental results are compared in the THUCNews datasets

Model	THUCNews		
	Acc	Pre	F1
TextCNN	90.92	90.97	90.93
TextRNN	90.95	91.01	90.92
TextRNN_Att	90.58	90.68	90.60
TextRCNN	90.73	90.93	90.77
DPCNN	90.90	91.03	90.94
FastText	91.92	91.88	91.85
TextGCN	86.91	86.94	86.94
TensorGCN	91.10	91.65	91.10
VSD-GCN	92.79	93.00	92.81

Further, to assess the robustness of VGCN, we employed the Toutiao dataset, focusing on short text comparisons with TextCNN, TextRNN, and TextGCN. As illustrated in Table 2, VGCN recorded an accuracy of 91.70%, a precision of 92.39%, and an F1 score of 91.11%. These results show substantial improvements over the TextGCN model, with accuracy and precision enhancements of 4.43% and 4.81%, respectively. VGCN's superior performance across different text lengths demonstrates its high accuracy and generalization capabilities, underscoring its superiority in text classification tasks. These outcomes further validate the effectiveness and performance advantages of the VGCN model in capturing textual semantic information.

Table 2: The experimental results are compared in the Toutiao datasets

Model	Toutiao		
	Acc	Pre	F1
TextCNN	85.38	84.51	83.75
TextRNN	82.41	81.03	80.67
TextRNN_Att	79.01	72.57	75.42
TextRCNN	82.73	82.96	89.76
DPCNN	78.81	78.85	75.13
FastText	88.93	89.88	88.08
TextGCN	87.27	87.58	87.32
TensorGCN	89.95	89.94	89.98
VSD-GCN	91.70	92.39	91.11

To assess the performance of the VGCN, we monitored the changes in loss rates on both the training and validation sets across two datasets. We observed a consistent decrease in the loss function over the course of iterations, indicating that VGCN effectively captured the features and patterns of the data, demonstrating stable performance on the validation set.

Specifically, in the initial phase, there was a significant drop in the loss rates within the first 10 epochs, suggesting that the model was rapidly adapting to the training data and generalizing well to the validation set. Subsequently, while the loss continued to decrease, the rate of reduction slowed, indicating that the model had entered a more stable phase of learning. In later iterations, the loss rate gradually declined and eventually stabilized, signifying that the model had thoroughly learned and adapted to the primary characteristics and patterns of the data.

To prevent overfitting and further enhance the performance of the model, we implemented an early stopping strategy. This approach utilizes the performance on the validation set to determine when to halt training, thereby ensuring the model ' s generalizability. This careful monitoring and adjustment of the training process have been crucial in optimizing VGCN's effectiveness across diverse textual datasets.

5 CONCLUSION

In this paper, we propose an improved Graph Convolutional Network (GCN) algorithm, which we call Variational Graph Convolutional Network (VGCN), aimed at addressing challenges related to feature extraction instability, high dimensionality, and the insufficient mining of latent semantic information in text classification tasks. By employing a variational structure to learn the latent representations of textual data, we effectively tackle the issue of extracting features from high-dimensional graph structures. Our experiments on the THUCNews and Toutiao datasets demonstrate that VGCN outperforms existing models in classifying both long and short texts. The comparative results not only confirm the superior performance of our model but also underscore its potential.

Looking forward, we plan to further explore the generalization capabilities of this model by testing it across a more diverse and extensive array of datasets, including multilingual and cross-domain datasets, to assess its adaptability and robustness in various textual environments. This work aims to broaden the applicability of VGCN, ensuring it delivers high performance across different linguistic and contextual challenges.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

REN, Zheng

College of Computing, Georgia Institute of Technology, North Avenue, Atlanta, GA 30332.

REFERENCES

[1] Li Fangfang, Su Puzhen, Duan Junwen, Zhang Shichao, & Mao Xingliang.(2023). Multi-granularity information relationship enhancement for multi-label text classification. Journal of Software, 34(12), 5686-5703.

[2] Tang Qinting. (2020). Research and Implementation of Deep Learning Based Online News Text Classification System Master (Dissertation, Beijing University of Posts and Telecommunications). Master https://link.cnki.net/doi/10.26969/d.cnki.gbydu.2020.002057 doi:10.26969/d.cnki.gbydu.2020.002057.

[3] Wang, X., Li, F., Zhang, Z., Xu, G., Zhang, J., & Sun, X. (2021). A unified position-aware convolutional neural network for aspect based sentiment analysis. Neurocomputing, 450, 91-103.

[4] Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning--based text classification: a comprehensive review. ACM computing surveys (CSUR), 54(3), 1-40.

[5] Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., ... & He, L. (2022). A survey on text classification: From traditional to deep learning. ACM Transactions on Intelligent Systems and Technology (TIST), 13(2), 1-41.

[6] Sun Hong, Lu Xinrong, Xu Guanghui, Huang XueYang & Ren Libo. (2023). Fusing semantic and syntactic dependency analysis for graph convolutional news text classification. Journal of Chinese Information (07), 91-101.

[7] Wang, J., Liang, J., Cui, J., & Liang, J. (2021). Semi-supervised learning with mixed-order graph convolutional networks. Information Sciences, 573, 171-181.

[8] Yao, L., Mao, C., & Luo, Y. (2019, July). Graph convolutional networks for text classification. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 7370-7377).

[9] Zhang, L., Jiang, L., Li, C., & Kong, G. (2016). Two feature weighting approaches for naive Bayes text classifiers. Knowledge-Based Systems, 100, 137-144.

[10] Haddoud, M., Mokhtari, A., Lecroq, T., & Abdeddaïm, S. (2016). Combining supervised term-weighting metrics for SVM text classification with extended term representation. Knowledge and Information Systems, 49, 909-931.

Journal of Industrial Engineering and Applied Science
Journal Home: jieas.suaspress.org
Vol. 2, No. 4, 2024 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

[11] Abu Alfeilat, H. A., Hassanat, A. B., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., Eyal Salman, H. S., & Prasath, V. S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data, 7(4), 221-248.

[12] Kim, H., & Jeong, Y. S. (2019). Sentiment classification using convolutional neural networks. Applied Sciences, 9(11), 2347.

[13] Zhang, Y., Liu, Q., & Song, L. (2018). Sentence-state LSTM for text representation. arXiv preprint arXiv:1805.02474.

[14] Liu, P., Qiu, X., & Huang, X. (2016). Recurrent neural network for text classification with multi-task learning. arXiv preprint arXiv:1605.05101.

[15] Johnson, R., & Zhang, T. (2017, July). Deep pyramid convolutional neural networks for text categorization. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 562-570).

[16] Teng Jinbao, Kong Weiwei, Tian Qiaoxin, Wang Zhaoqian, & Li Long. (2021). A multi-channel attention mechanism text classification model based on CNN and LSTM. Journal of Computer Engineering & Applications, 57(23).

[17] Liu, X., You, X., Zhang, X., Wu, J., & Lv, P. (2020, April). Tensor graph convolutional networks for text classification. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 05, pp. 8409-8416).

[18] Ding, K., Wang, J., Li, J., Li, D., & Liu, H. (2020). Be more with less: Hypergraph attention networks for inductive text classification. arXiv preprint arXiv:2011.00387.

[19] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).

[20] Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759.

[21] Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., & Xu, B. (2016, August). Attention-based bidirectional long short-term memory networks for relation classification. In Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers) (pp. 207-212).

[22] Lyu, S., & Liu, J. (2021). Convolutional recurrent neural networks for text classification. Journal of Database Management (JDM), 32(4), 65-82.