International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

# GRAPH CONVOLUTIONAL NETWORK BASED ON MULTI HEAD POOLING FOR SHORT TEXT CLASSIFICATION

**Mr. M. Uday Kumar, Nandhini Mamidi**

Associate professor Department of Computer science and Engineering.
M.Tech, Department of Computer science and Engineering (21JJ1D5806), nandhumamidi08@gmail.com

**ABSTRACT:** Short texts, few features, and not having enough training data, among other things, are still the main problems that make it hard for standard text classification methods to work well. To solve these issues, we created the MP-GCN for classifying short texts with some help from a human. We also offer its three forms, which teach node representation in 1-order heterogeneous networks, 1-order isomorphic graphs, and 1&2 order isomorphic graphs. It checks the structure of the text network without using training word embedding as the first node trait. The multi-head technique provides the pooling representation subspaces without trainable parameters, whereas a self-attention based graph pooling methodology identifies and assesses the most significant nodes. Without using pre-training embedding, the results of the experiments showed that MP-GCN did better than the best models on five standard datasets.

***Keywords*** *– Multi-head-Pooling-based Graph Convolutional Network (MP-GCN), Short text classification, Semi-supervised learning, Isomorphic graphs, Heterogeneous graphs, Node representation learning, Graph convolutional networks, Structural information, Graph pooling, Self-attention.*

## INTRODUCTION

One of the earliest NLP challenges is text categorization. The purpose of NLP is to tag words, queries, paragraphs, and documents. Researchers have developed DL models that categorize text better than ML approaches in recent years. RNN, CNN, transformer, and capsule net models are examples. Few academics have investigated semisupervised graph convolutional networks (GCNs) for text recognition in recent years [1, 2]. The major reason is that it works in a lot of real-life situations. To begin, it works better with short texts by making more connections between word nodes. It can also be used with scenes that have few or unclear meanings and no background information [2]. Second, it works well in situations where there isn't a lot of named training data, which is something that makes traditional neural networks not work very well [3]. Thus, we must immediately explore semi-supervised GCNs for text categorization. Also, semi-supervised GCNs are difficult to utilize. Due to varied situations, the pre-training word vector may not enhance text categorization. Instead, it may hinder graph creation. Second, it generates a network for the collection to add additional node connections (often using algorithms). Graph creation and feature extraction must consider memory and processing use.

## LITERATURE REVIEW

Graph Convolutional Networks (GCNs) have gotten a lot of attention lately because they work well in many situations, especially when it comes to classifying text. Kipf and Welling (2016) were the first to use GCNs in semi-supervised classification. Their work laid the groundwork for more study to come in this area [1]. Taking this idea a step further, Yao, Mao, and Luo (2019) looked into how GCNs can be used for text classification and showed that they are good at working with graph-structured data that comes from text [2].

Linmei et al. (2019) created Heterogeneous Graph Attention Networks (HANs) to help with the difficult task of semi-supervised short text classification. These networks show that GCNs can be used with a variety of graph shapes [3]. Bruna et al. (2013) and Defferrard et al. (2016) helped us understand spectral networks and locally linked networks on graphs. This set the stage for future progress in graph-based neural designs [4, 5].

In 2017, Hamilton, Ying, and Leskovec did more research on GCNs using inductive representation learning on big graphs. This made the models more scalable and useful in a wider range of situations [6]. Velickovic et al. (2017) created Graph Attention Networks (GATs), which stress how important attention processes are for understanding how nodes are connected [7].

There are also in-depth reviews of graph neural networks (GNNs) and how they can be used in the literature. Zhou et al. (2018) gave a thorough summary of GNN methods and the many ways they can be used. They showed how flexible these models are when it comes to finding complex relationships in data [8].

Text Level Graph Neural Networks (TL-GNNs), which were suggested by Huang et al. (2019), show how useful GNNs can be for text classification tasks, especially when the goal is to find connections between words and sentences [9]. Zhang, Li, and Song (2019) created Aspect-Specific Graph Convolutional Networks and showed how well they work for classifying sentiments based on aspects [10].

Tensor Graph Convolutional Networks were created by Liu et al. (2020). They made GCNs work with tensor data, which opens up more options for text classification jobs [11]. Peng et al. (2018) studied large-scale hierarchical text classification using recursively regularized deep graph-CNNs. They showed that graph-based models can handle very big datasets [12].

In conclusion, the research on using graph convolutional networks to classify text shows a wide range of models and methods. These methods are highly flexible, scalable, and useful in a wide range of situations and data types. All of these improvements help the area of graph-based neural models for natural language processing grow.

**Algorithms.**

In this we used algorithms like BERT + MP-GCN – GRU – LSTM – CNN - Bi-LSTM

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

BERT + MP-GCN:

The BERT system is free and open source for ML that works with NLP. BERT is meant to help computers figure out what unclear language in text means by using the text around it to set the scene.

A Graph Convolutional Network, or GCN, is a way to learn on graph-structured data with some help from a teacher. It is built on a type of convolutional neural networks that work well and directly on graphs.

GRU:

Gated recurrent units (GRUs) are a way to control how recurrent neural networks work. They were first described by Kyunghyun Cho et al. in 2014. An LSTM with a forget gate is like a GRU. However, the GRU has fewer factors than an LSTM because it doesn't have an output gate. In some situations, the GRU is better than LSTM. It is a type of RNN. If you use datasets with longer patterns, LSTM is more accurate than GRU. GRU is faster and uses less memory.

LSTM:

LSTM is a DL architecture based on an artificial recurrent neural network. Time series and pattern issues can be solved using LSTMs. Recurrent neural networks like Long Short-Term Memory (LSTM) can estimate sequences by order. This is essential in problem-filled fields like voice recognition, machine translation, and more. LSTMs are hard in DL

CNN:

CNNs are DL network designs. It processes pixel data to identify pictures and other tasks. CNNs are the greatest DL neural networks for discovering and identifying items. CNN learned spatial feature structures organically and adaptably via backpropagation. Fully linked layers, convolution layers, and pooling layers are used to achieve this.

Bi-LSTM:

Bidirectional LSTM layers understand how time steps in a time series or chain are connected across time in both directions. These associations may help the network learn from the complete time series at each time step. We can make any neural network recall sequences from the past or future. Bidirectional long-short-term memory.

**ARCHITECTURE**

The design of the system includes Traditional text classification methods have some flaws. The Multi-head-Pooling-based Graph Convolutional Network (MP-GCN) fixes these problems by adding three designs for semi-supervised short text classification. The main goal of these designs is to learn how to describe nodes for 1-order isomorphic graphs, 1-order mixed graphs, and 1-order isomorphic graphs with two orders. MP-GCN only uses the structure

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

knowledge of the text graph, so it doesn't need to learn word embeddings ahead of time. It has a self-attention-based graph sharing method for figuring out which nodes are important and choosing them. The multi-head method improves representation variation without adding more factors that can be trained. The results of the experiments show that the model does better on five standard datasets than the best models that don't use pre-training embeddings.

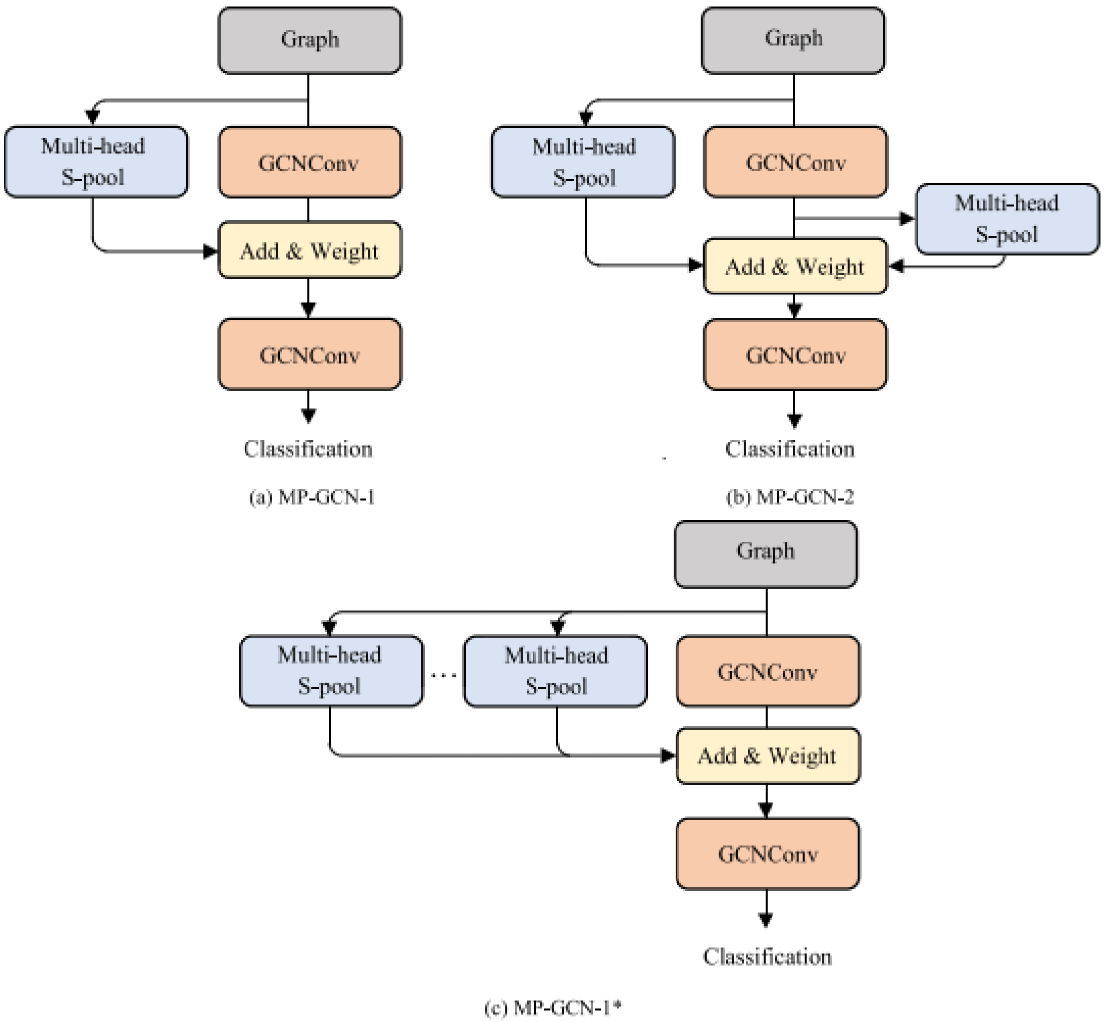


Fig System Architecture

**COMPARISON TABLE**

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in
IJARST

Table.1: Graph Convolutional Network Based On Multi Head Pooling For Short Text Classification

| S. No | Title | Author/Reference | Method/Algorithm implemented | Advantage | Disadvantage |
|---|---|---|---|---|---|
| 1 | Semi-Supervised Classification with Graph | Thomas N. Kipf, Max Welling [1] | Efficient variant of convolutional neural networks for semi-supervised learning on | Linear scalability with graph edges, encodes local structure and | Not explicitly mentioned in the provided information. |

| | | | | | |
|---|---|---|---|---|---|
| | Convolutional Networks | | graph-structured data, based on spectral graph convolutions. | node features, outperforms related methods significantly in experiments. | |
| 2 | Graph Convolutional Networks for Text Classification | Liang Yao, Chengsheng Mao, Yuan Luo [2] | Text GCN applies graph convolutional networks to text classification, using word co-occurrence and document word relations for learning embeddings. | Text GCN outperforms state-of-the-art text classification methods, learns predictive embeddings for words/documents, and exhibits robustness with less training data. | Dependence on word co-occurrence may limit performance; potential complexity and resource requirements in constructing and processing large, dynamic graphs. |
| 3 | Heterogeneous Graph Attention Networks for Semi-supervised Short Text Classification | Hu Linmei, Tianchi Yang, Chuan Shi, Houye Ji, Xiaoli Li[3] | Heterogeneous Graph ATtention networks (HGAT) for semi-supervised short text classification using a dual-level attention mechanism in Heterogeneous Information Networks (HIN). | HGAT effectively leverages few labeled and large unlabeled data through dual-level attention, outperforming state-of-the-art methods across six benchmark datasets significantly. | Potential complexity in implementing and tuning dual-level attention may require expertise; dependence on data quality and network structure for optimal performance. |
| 4 | Spectral Networks and Locally | Joan Bruna, Wojciech Zaremba, Arthur Szlam, Yann LeCun | Two proposed constructions for generalizing CNNs to non-translationally invariant | Efficient learning on low-dimensional | Limited to low-dimensional |

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

| | | | | | |
|---|---|---|---|---|---|
| | Connected Networks on Graphs | [4] | domains—hierarchical clustering and graph Laplacian spectrum-based. | graphs, enabling convolutional layers with parameters independent of input size for deep architectures. | graphs; may not perform optimally on high-dimensional or complex structures, limiting generalizability to diverse domains. |
| 5 | Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering | Michaël Defferrard, Xavier Bresson, Pierre Vandergheynst [5] | Formulating CNNs using spectral graph theory for efficient localized convolutional filters on irregular domains represented by graphs. | Achieves linear computational and constant learning complexity, applicable to diverse graph structures, demonstrated on MNIST and 20NEWS datasets. | Specific limitations not mentioned, potential challenges in handling certain graph structures or scalability issues not addressed explicitly in the summary. |

**SUMMARY**

This study shows that The Multi-head-Pooling-based Graph Convolutional Network (MP-GCN) solves problems that traditional text classification methods have, like short text, few features, and not enough training data. It shows three ways to classify short texts with some help from a human, with a focus on learning how to describe nodes in isomorphic and mixed graphs. MP-GCN uses the structure information in the text graph instead of word embeddings that need to be trained first. Adding a self-attention mechanism to a graph pooling mechanism picks out important nodes, and the multi-head method improves representation subspaces. The MP-GCN model does better than the best models in five different datasets, and this is true even without pre-training embeddings.

**CONCLUSION**

In this work, we suggest MP-GCN as a way to sort short texts. Multi-head sharing is used in this network to improve the learning of representations for important nodes. We show three versions of MP-GCN designs that focus on learning how to model nodes in a 1-order graph, a 1&2-order isomorphic graph, and a 1-order mixed graph. The results of the experiments show that MP-GCN can do better than the best models on five different standard datasets without using pre-training embedding.

**FUTURE SCOPE**

MP-GCN's future lies in its ability to be used in a wide range of fields that need to classify short texts. More study can look into how well it can handle changing data sets and processing in real time, which would make it easier to use on a larger scale. Adding advanced pre-training anchoring methods may help you get even better results. Additionally, looking into how well MP-GCN works with busy or uneven datasets can help us understand how it can be used in real life. Working together to make the design bigger so it can handle written material in multiple languages and formats could lead to uses in more than one field. As the field changes, it will be very important to make MP-GCN work better for specific industries and deal with new problems.

**REFERENCES**

[1] T. N. Kipf and M. Welling, ''Semi-supervised classification with graph convolutional networks,'' 2016, arXiv:1609.02907.

[2] L. Yao, C. Mao, and Y. Luo, ''Graph convolutional networks for text classification,'' in Proc. Conf. Assoc. Adv. Artif. Intell. (AAAI), Sep. 2019, pp. 7370–7377.

[3] H. Linmei, T. Yang, C. Shi, H. Ji, and X. Li, ''Heterogeneous graph attention networks for semi-supervised short text classification,'' in Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP), 2019, pp. 4821–4830.

[4] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, ''Spectral networks and locally connected networks on graphs,'' 2013, arXiv:1312.6203.

[5] M. Defferrard, X. Bresson, and P. Vandergheynst, ''Convolutional neural networks on graphs with fast localized spectral filtering,'' in Proc. Conf. Neural Inf. Process. Syst. (NeurIPS), Jun. 2016, pp. 3844–3852.

[6] W. Hamilton, Z. T. Ying, and J. Leskovec, ''Inductive representation learning on large graphs,'' in Proc. Conf. Neural Inf. Process. Syst. (NeurIPS), Dec. 2017, pp. 1024–1034.

[7] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, ''Graph attention networks,'' 2017, arXiv:1710.10903.

[8] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, ''Graph neural networks: A review of methods and applications,'' 2018, arXiv:1812.08434.

[9] L. Huang, D. Ma, S. Li, X. Zhang, and H. Wang, ''Text level graph neural network for text classification,'' in Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLPIJCNLP), 2019, pp. 3444–3450.

[10] C. Zhang, Q. Li, and D. Song, ''Aspect-based sentiment classification with aspect-specific graph convolutional networks,'' in Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP), Nov. 2019, pp. 4568–4578.

[11] X. Liu, X. You, X. Zhang, J. Wu, and P. Lv, ''Tensor graph convolutional networks for text classification,'' in Proc. Conf. Artif. Intell. (AAAI), Apr. 2020, pp. 8409–8416.

[12] H. Peng, J. Li, Y. He, Y. Liu, and M. Bao, ''Large-scale hierarchical text classification with recursively regularized deep graph-CNN,'' in Proc. Conf. World Wide Web (WWW), Apr. 2018, pp. 1063–1072.

[13] Y. Kim, ''Convolutional neural networks for sentence classification,'' in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), Oct. 2014, pp. 1746–1751.

[14] T. Liu, X. Zhang, W. Zhou, and W. Jia, ''Neural relation extraction via inner-sentence noise reduction and transfer learning,'' in Proc. Conf. Empirical Methods Natural Lang. Process., Oct. 2018, pp. 2195–2204.

[15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ''BERT: Pre-training of deep bidirectional transformers for language understanding,'' 2018, arXiv:1810.04805.