ML-GAT: MULTI LABEL NODE CLASSIFICATION USING ENHANCED GRAPH ATTENTION NETWORKS

A PROJECT REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF BACHELORS OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

Submitted by

Ashi Gupta (Roll number: 2k16/CO/065)

Prachi Garg (Roll number: 2k16/CO/220)

Performed under the guidance of

Prof. Rajni Jindal

Head of Department

Department of Computer Science and Engineering

MULTI-LABEL NODE CLASSIFICATION USING ENHANCED GRAPH ATTENTION NETWORKS

We propose a novel architecture, Multi Label Graph Attention Network (ML-GAT) that leverages the applicability of the attention based GAT network to efficient inductive semi-supervised multi-label classification by augmenting complex label-label and node-label dependencies implicit in the graph structure to the learning process. We compare our method with ML-GCN [1] and GAT [2] and GCN [1] baselines to examine the influence these losses have on the models and perform three empirical studies on the datasets to make a comparative analysis over the methods.

RESEARCH CONTRIBUTION

01

Firstly, we leverage graph attention networks for multi-label node classification and study the influence that the inclusion of label and node correlations has on the training process and learned models.

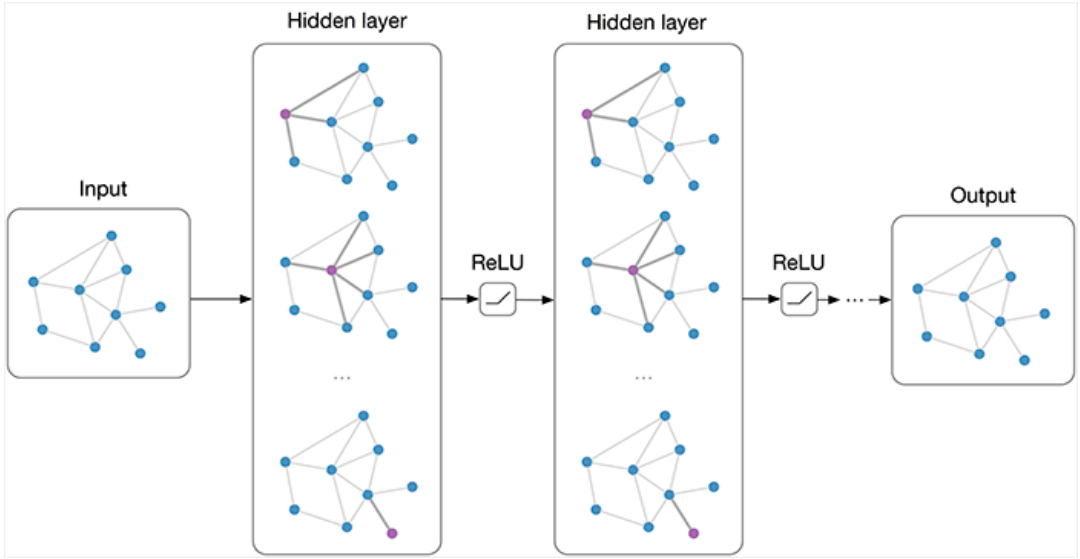
02

Secondly, we compare the relative importance of inter-label and node-label correlation with respect to ML-GCN and ML-GAT models.

03

Thirdly, we provide a comprehensive study that establishes state-of-the-art benchmarks for MLC using GNNs and standard datasets.

GRAPH NEURAL NETWORKS



Multi-layer Graph Convolutional Network (GCN)

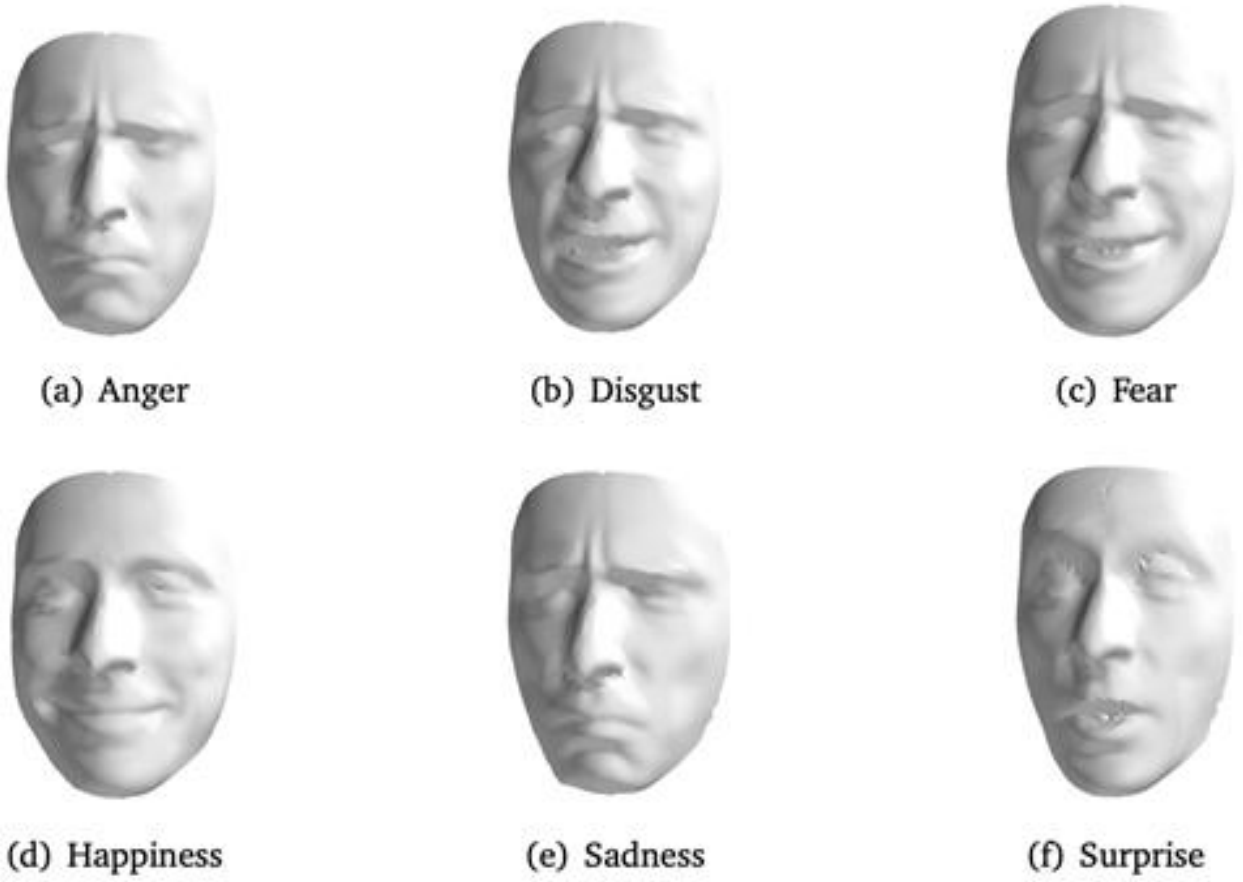
- We explore the field of geometric deep learning which is a generalization of deep learning to graphs and manifolds.
- Graph neural networks are a promising class of neighborhood aggregation based neural network models that learn rich representations on graph-structured data in its raw form capturing complex entity relationships from the graph topological structure.
- Traditional approaches like graph embeddings (Deepwalk [4], Node2vec [5], etc) extract structural information from graphs. They mostly rely on summary graph statistics, kernel functions, or hand designed features to measure local neighborhood structures and hence lack flexibility, scalability and are time consuming

MULTILABEL CLASSIFICATION

- In multi-label node classification, we can assign multiple labels to each instance. This is different from single label node classification where each instance can belong to a single class from the possible classes.
- Multi-label node classification on graphs is a relatively unexplored area.
- Since many real-world graph based problems require the assignment of more than one label to each node instance in the graph, we study the multi-label node classification problem.

APPLICATIONS OF GNNS FOR MULTI LABEL CLASSIFICATION

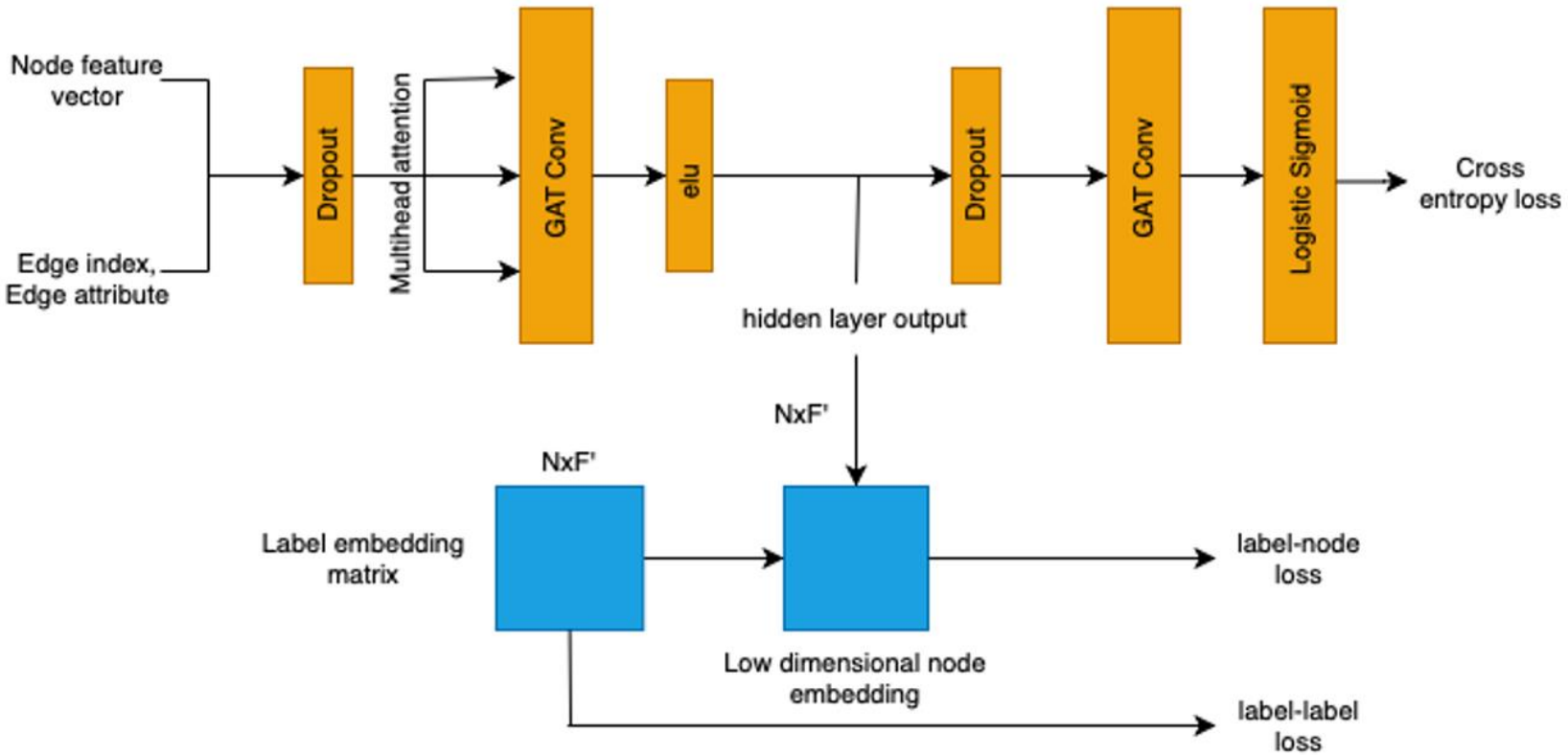
- Protein protein interaction networks are commonly used to study amino acid interactions and GNNs can solve several problem statements on this dataset.
- Graph based learning architecture could be used for problems like video annotation, a video clip can be annotated with multiple labels at the same time, such as ‘person’, ‘people march’, ‘pedestrians’.
- In facial expression classification, a person may express happiness and relaxation at the same time. Hence, it is a multi-label classification problem.



PROPOSED METHODOLOGY AND NOVELTY

- We propose a novel architecture, Multi Label Graph attention Network (ML-GAT) that leverages the applicability of the attention based Graph Attention Network (GAT) to efficient inductive semi-supervised multi-label classification by augmenting complex inter-label and node-label dependencies implicit in the graph structure to the learning process.
- This is done by incorporating node-label and label-label correlation losses in the objective function
- We implemented the GAT baseline in ML-GAT for weighted graphs.
- Main objective function optimized during training:

$$L_{sum} = \lambda_1 L_{n-l} + \lambda_2 L_{l-l} + l_{entropy}$$



Network architecture for Multi-Label Graph Attention Network (ML-GAT)

DATASET

For the purpose of this research, we use these multi-label graph-structured datasets:

FACEBOOK

Facebook is a **social network dataset** consisting of nodes representing the Facebook user and the edges between the nodes represents friendship relations between the nodes. The goal is to identify the **'social circles'** each user belongs to. It has 710 nodes, 56824 edges and 60 classes.

YEAST

This is a biological dataset. In the Yeast dataset nodes represent genes in the form of micro-array expression data and edges indicate gene interactions, the graph represents protein interaction in the yeast organism. **The task is to predict the cellular localization sites of proteins based on genes and their interactions.** It has 1240 nodes, 4466 edges and 13 classes.

EXPERIMENTATION & RESULTS

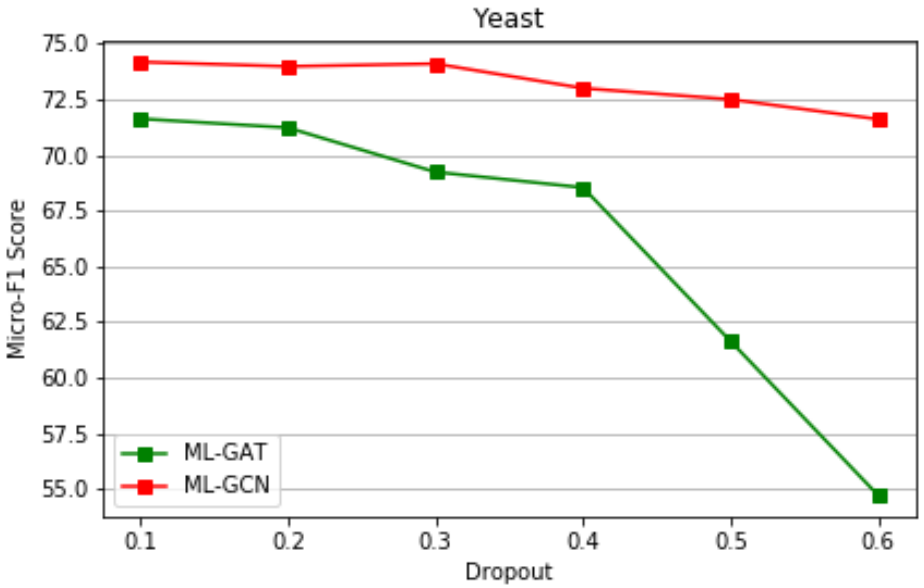
We conduct 4 types of experiments on the Facebook and Yeast datasets for multi-label node classification:

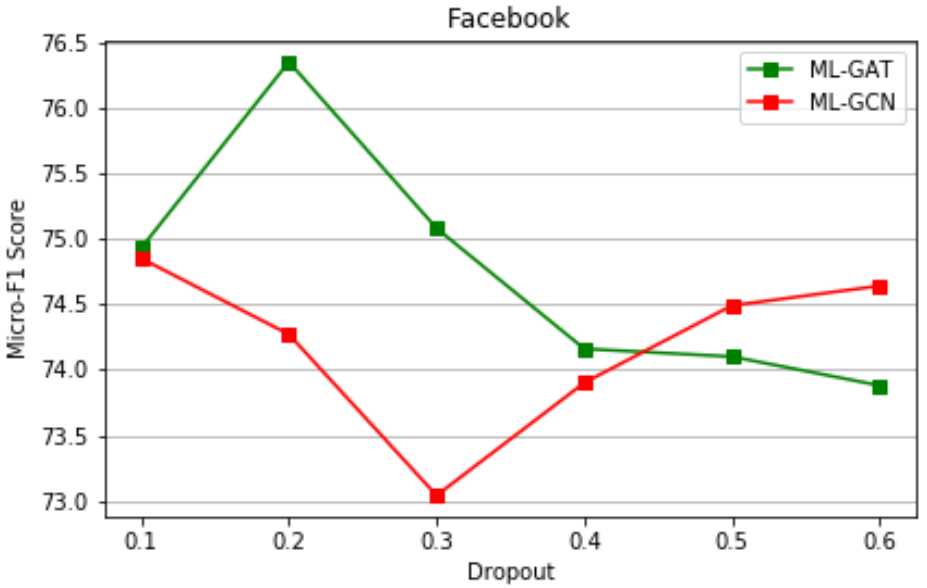
1. We first reproduced the Multi label Graph Convolution Network baseline as given in the original paper. ***We report test accuracy as 71.64 and 64.78 for Facebook and Yeast respectively which are both greater than the mentioned test accuracy in the MLGAT[1] paper.***
2. Next we add label embedding matrix to GAT backbone so as to include label-label and node-label loss in loss calculation.
3. We then analyze performance with different hyperparameter combinations.
4. Finally we perform experiments to find weightage of label-label and node-label loss with respect to accuracy.

EXPERIMENTATION & RESULTS

- Our MLGCN results exceed MLGCN* by **12.44%** and **7.94%** for Facebook and Yeast respectively.
- Our proposed model MLGAT exceeds MLGCN* by **15.01%** and **6%** for Facebook and Yeast respectively.
- Our proposed model MLGAT exceeds GAT by **4.68%** and **0.5%** for Facebook and Yeast respectively.

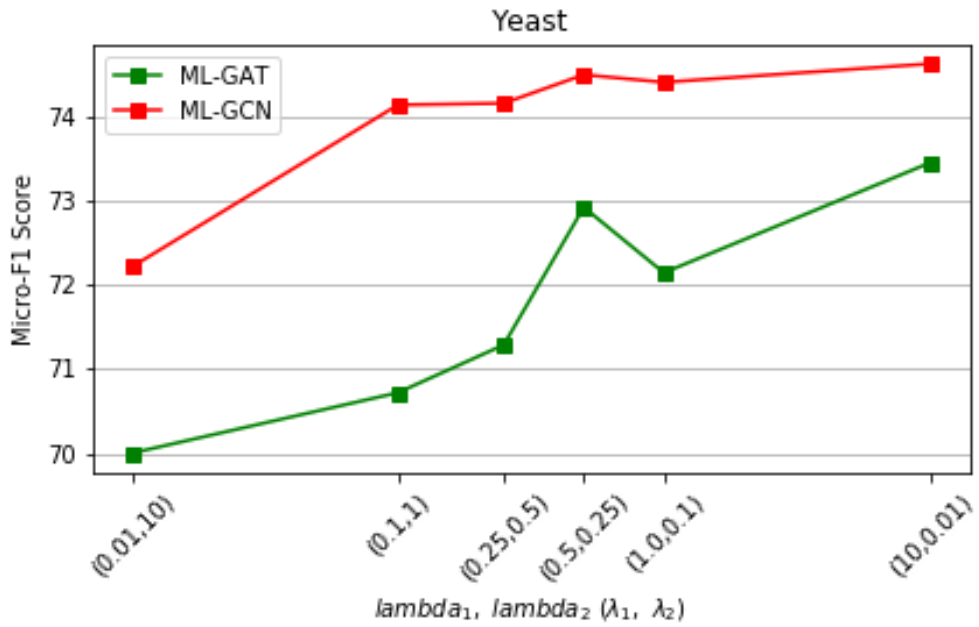
MODELS	FACEBOOK	YEAST
GCN*[2]	58.13	63.16
GCN	70.76	73.25
GAT	70.18	71.56
ML-GCN*[2]	59.85	66.06
MLGCN	72.29	74.0
MLGAT	74.86	72.06

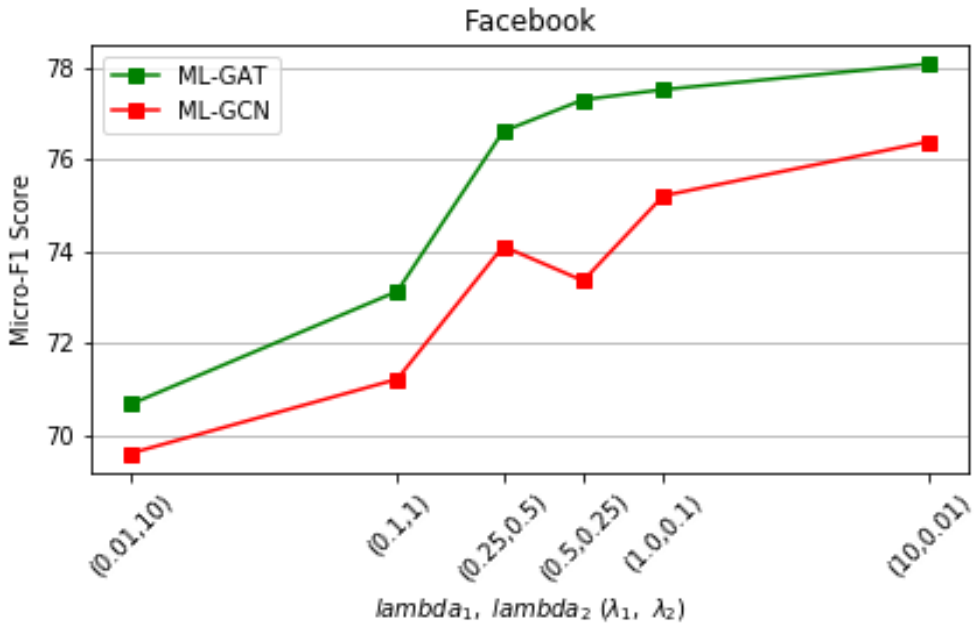




EXPERIMENTATION & RESULTS

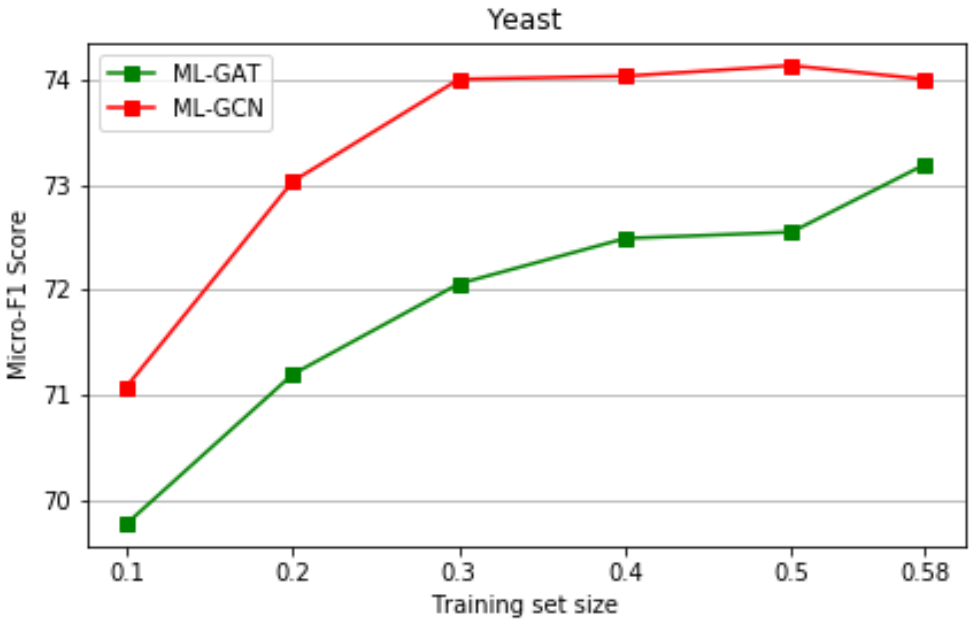
- Graph plots showing influence of layer dropout variation on classification performance. We show the results of our experiments on both the standard ML-GCN model and proposed ML-GAT model.
- Hyperparameter – Dropout

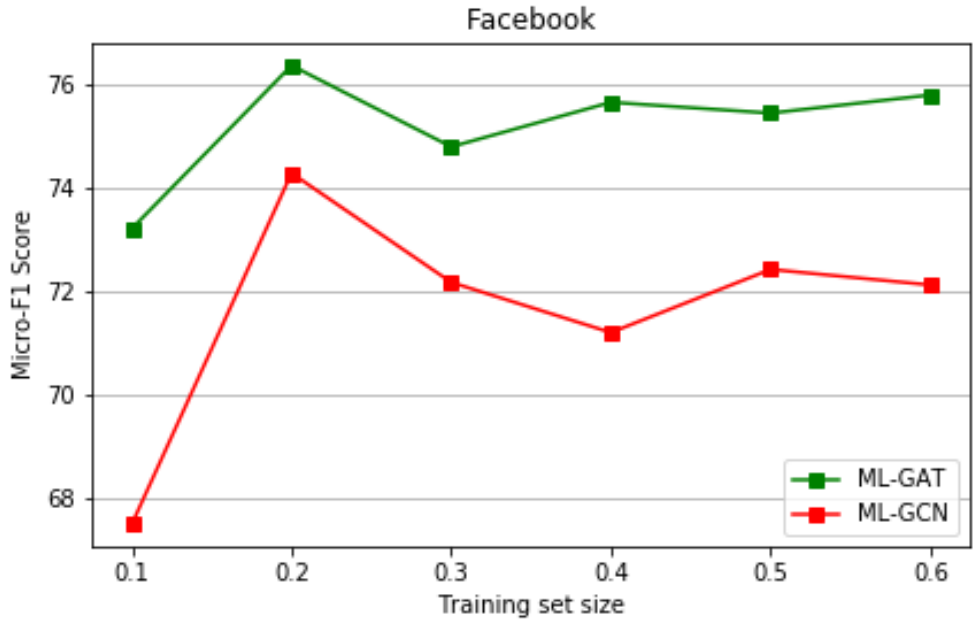




EXPERIMENTATION & RESULTS

- This experiment investigates the influence of varying the regularization hyperparameters lambda1 and lambda2 on model performance. This helps us draw inferences on the importance of the node-label and label-label losses.
- Hyperparameters – lambda1 and lambda2





EXPERIMENTATION & RESULTS

- Graph plots showing influence of training set sizes on test set performance for both ML-GCN and ML-GAT in our experiments.
- Hyperparameter- Training Size

FUTURE WORK

- A possible direction for future work could be performing these experiments over large-scale inductive learning tasks like MLC on the Protein-Protein Interaction dataset (PPI), which contains multiple graphs. GAT has a few inherent properties that make it suitable for large scale learning.
- We can also check the performance of the proposed framework on unweighted graph datasets.

TOOLS USED

1

Python, Pytorch; Tensorboard, Matplotlib to analyze training and performance evaluation

2

We majorly used the **Pytorch Geometric** Pytorch library to implement the graph neural network models.

3

All models were trained using Google Collaboratory K80 GPUs and NVIDIA GeForce GTX 1650

REFERENCES

- [1] Gao, Kaisheng, Jing Zhang, and Cangqi Zhou. "Semi-Supervised Graph Embedding for Multi-Label Graph Node Classification." International Conference on Web Information Systems Engineering. Springer, Cham, 2019.
- [2] Veličković, Petar, et al. "Graph attention networks." arXiv preprint arXiv:1710.10903 (2017).
- [3] Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." arXiv preprint arXiv:1609.02907 (2016).
- [4] Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014.
- [5] Grover, Aditya, and Jure Leskovec. "node2vec: Scalable feature learning for networks." *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 2016.

ICTIS 2020 ACCEPTANCE EMAIL

ICTIS -- 2020 : Notification for Paper : 672 : ML-GAT: Multi-Label Classi↑cation Using Enhanced Graph Attention Network Inbox ×

Springer ICTIS 2020 <info@ictis.in> Wed, 22 Apr, 12:59 (4 days ago)
to rajnijindal, me, Prachi

Dear Dr. Rajni Jindal, Ashi Gupta, Prachi Garg,

Ref: 672

Title: ML-GAT: Multi-Label Classi↑cation Using Enhanced Graph Attention Network

Congratulations! On behalf of the Program Committee of ICTIS 2020 - Ahmedabad, I am happy to inform you that your above mentioned paper has been ACCEPTED for oral presentation in ICTIS 2020 and publication in Springer SIST series subject to fulfillment of Guidelines by Springer.

Important Notification COVID 19 - We are aware of severe travel restrictions, travel advisories, lockdowns in India and across the globe due to pandemic COVID 19, and we are al monitoring Indian Government measures daily. We will continue to monitor the situation and will post regular updates. In response to travel concerns as of now, the requirement present accepted papers in person will be relaxed and will allow for virtual presentations.

For all updates and details please visit - https://ictis.in/COVID-19.php

---

An accepted paper will be published in the Springer proceedings (SIST) only if the final version is accompanied by the payment information (i.e transaction reference number) subject to qual check as per Springer Guidelines.

Kindly follow the below mentioned guidelines (strictly), related to preparation of final manuscript, copyright transfer form, payment and final submission. The procedure has been detailed as a step process (I)-(III):

Our paper was accepted at the Fourth International Conference on ICT for Intelligent Systems (ICTIS – 2020). The proceedings of ICTIS get published in springer SIST (ISSN 2190-3018) ; Scopus cite score 0.65; H-index 15 according to scimagojr.