MAGNET:

Multi-Label Text Classification using Attention-based Graph Neural Network

CS7.403.S22.Statistical Methods in AI

Team Number - 06
Persistent Perceptron
Annapoorani A(2021201017)
Debashish Roy(2021201034),
Malla Hemalata(2021201066),

DOCUMENTATION

- In this project, we try implementing the paper to obtain Multi-Label Classification using Attention Based Graph Neural Networks.
- The paper proposes a novel approach that doesn't rely on methods like Binary Relevance, Classifier Chains or Label Powers as they do not take into account the label correlation in a multi-label approach.
- It uses a GAN network with Graph Convolution network as basis.
- The correlation is understood by using Adjacency matrix and attention weights to represent the correlation matrix.
- Embeddings are taken as input, fed to our model and the output is our multi labels which classify our text.
- Link to the Github Repository: https://github.com/debashish05/SMAI_S22_06

PROJECT SCOPE

- We will be implementing bi-lstm and used BERT embedding for the articles. We will treat this as our baseline model and compare the result of MAGNET with this model.
- We will be implementing the MAGNET model with a single Graph Attention layer for multi label text classification and compare the results with the baseline.
- A graph attention based model is used to capture the correlation between the various labels in our problem.

PROJECT TIMELINE

April 1st - Project teams announcement.

April 4th, 5th - Paper reading.

April 6th - Paper and project discussion with our mentor.

April 12th - Embedding of the input text using Bert

April 15th - Bi-LSTM and preprocessing text

April 18th - Baseline model completed

April 30th (Expected Date for final Evaluation) - MAGNET Model

Task Completed

- We have used Reuters-21578. It is a collection of documents collected from Reuters News Wire in 1987. The Reuters-21578 test collection, together with its earlier variants, has been such a standard benchmark for the text categorization (TC) (Debole and Sebastiani, 2005). It contains 10,788 documents, which has 8,630 documents for training and 2,158 for testing with a total of 90 categories.
- We have implemented bi-lstm and used BERT embedding for the articles. We are treating this as our baseline model and compare the result of MAGNET with this model.
- Currently we are getting an accuracy of **57% on training data and 59% accuracy on the test data.**

FUTURE TASKS

- We will need to implement the Graph Attention Network.
- We will create an adjacency matrix and use attention weights with it, which would represent the correlation between the labels, and use GAT networks in order to learn these parameters.
- An activation function would be used on our network represented in our hidden-layers, and finally the output obtained can be used to mark our multiple layers.
- A dot product will be performed with our initial Bi-LSTM model output to predict labels and determine the loss.

DELIVERABLE FOR END EVAL

For our end eval, we will have our model of the bi-directional LSTM ready and the MAGNET model.

THANK YOU