Sunday, March 12, 2023

DL4HC - Final Project Proposal - Team 61

Pre-training of Graph Augmented Transformers for Medication Recommendation

---

We choose for our final project, the GBERT graph learning model described in the 2019 paper https://www.ijcai.org/proceedings/2019/825 for the goal of drug recommendation.

The paper "Pre-training of Graph Augmented Transformers for Medication Recommendation" aims to solve the problem of medication recommendation, which is a critical task in healthcare. The paper presents a new approach to this problem using pre-training of Graph Augmented Transformers (GATs) on a large corpus of medical texts.

The authors argue that existing medication recommendation methods often ignore the complex relationships among medications, symptoms, and diseases, which can lead to suboptimal recommendations. To address this issue, the proposed model leverages medical knowledge encoded in large-scale medical corpora to pre-train the GAT model to learn the underlying structure of medical concepts and relationships.

The pre-trained GAT model is then fine-tuned on patient medical records to personalize the medication recommendation for each patient. The model can capture not only the direct relationships between medications and symptoms but also the indirect relationships between medications through shared diseases or symptoms.

The authors evaluate their proposed approach on several real-world datasets and show that it outperforms state-of-the-art medication recommendation methods. The MIMIC3 Dataset is used by the implementation of this paper and that is what the team will use for our coursework final project.

The specific enhancements discussed are :

- Pre-training to leverage more data - adapt the framework of BERT and pre-train the model on each visit of the EHR data to leverage the single-visit data that were not fit for training in other medication recommendation models.

- Medical ontology embedding with graph neural networks
- The framework of G-BERT consists of three main parts: ontology embedding, BERT and fine-tuned classifier.
  - Ontology embedding for medical code laid in leaf nodes by cooperating ancestors information based on graph attention networks.
  - Next, input set of diagnosis and medication ontology embedding separately to shared weight BERT which is pre- trained.
  - Finally, concatenate the mean of all previous visit embeddings and the last visit embedding as input and fine-tune the prediction layers for medication recommendation tasks.

The proposed approach is interesting and innovative because it addresses the key limitations of existing medication recommendation models, such as selection bias, the lack of consideration of complex relationships among diagnosis, medications, symptoms, and diseases. By using a pre-trained GAT model, the proposed approach can provide more personalized and effective medication recommendations for patients, leading to improved patient outcomes.

As part of the project, we will be reproducing the results of the paper using the base code https://github.com/jshang123/G-Bert on MIMIC3 synthetics data provided in the github repository. We plan to develop code and on need basis reuse some of the existing code used by the paper. As an extension we will apply the same on MIMIC IV dataset.

We will use our personal M1/M2 Macs and will transition to google cloud service if we hit performance issues. That being said, the paper provides some information on the hardware used for their experiments. They used an NVIDIA V100 GPU with 16GB memory for pre-training the transformer model on a large corpus of medical records. For fine-tuning the model on the medication recommendation task, they used a NVIDIA TITAN Xp GPU with 12GB memory.Based on this information, it is likely that a GPU with at least 12GB memory would be needed for fine-tuning the model on the medication recommendation task.

For ablation analysis, is a stretch goal for us for this project and we had the transformer architecture in mind for this. The proposed method uses a transformer model as the backbone for medication recommendation. To study the impact of the transformer architecture, we might compare the performance of the proposed method with other transformer-based models or non-transformer-based models.