

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 innovative because it addresses some of the key limitations of existing medication recommendation methods, such as the lack of consideration of complex relationships among 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. As an extension we will apply the same on MIMIC IV dataset. TODO: Why is this interesting ? Computational feasibility ?