Yanyi Chu

PhD Candidate

Full-time Intern

December 1996

15941896746

a96123155@sjtu.edu.cn

Shanghai, China

Skills

Languages： Python, R, Matlab, SQL, Latex, Markdown, C++, etc.

Development： Linux, Git, Shell, etc.

Frameworks： Scikit-learn, Tensorflow, Keras, Pytorch, Scikit-multilearn, XGBoost, Mlxtend, Networkx, Deep Graph Library, Requests, etc.

Interests

- Heterogeneous graph network
- Recommendation system
- Deep ensemble learning
- Graph neural network
- Multi-label learning
- Data mining

On-campus

- Vice-chairman of the college postgraduate students' union
- Head of the College Employment Social Practice Group

Strengths

- Multi-field Knowledge Reserve (AI medicine, Recommendation System, NLP, etc.)
- Good Understanding for AI papers
- Rich Experience in model coding
- Fast Learner
- Innovator

Education

2017-2022 Shanghai Jiao Tong University, China — Honorary PhD candidates
Major in Bioinformatics, GPA 3.83/4

2021(EXP) University of Calgary, Canada — Visiting PhD
Major in Cheminformatics

2013-2017 Shenyang Pharmaceutical University, China — Bachelor
Major in Pharmaceutics, GPA 3.5/4, TOP 5%

Project Experience

2020.7-2020.9 Lightspeed & Quantum Studios Group Intern Tencent

First, I build graph, extract node and edge features based on user's information. Then, construct a variety of graph neural network models (GCN, GraphSAGE, R-GCN, GraphSAINT, GCMC, PinSAGE, etc), and optimize the processing methods of unbalanced data (such as oversampling, downsampling, etc.). Finally, the model performance achieves the ideal result (precision and recall are 0.9 and 0.6, respectively), and the performance is comparable to that of LSTM using complex timing-sequence features.

2019.1-2019.12 Machine learning technology Intern Intel Asia-Pacific Co.

We work on *drug retrosynthesis which based on MCTS and deep learning.* First, the drug data set is expanded by deep generative model, drug features are extracted through graph neural network. Next, MCTS with deep learning are applied to return the feasible route for drug retrosynthesis.

2018.10-2018.12 Datacenter Intern Databerry

- Independently completed *the Douban website movie crawler project,* and made user software for colleagues.
- Use the recommendation algorithm based on association rules to independently complete *the project of film and television placement advertising prediction.*

PhD candidate project: Application of artificial intelligence in drug discovery and drug relocation

First, I use *data mining* techniques such as web crawlers to construct new data sets. Then use *graph neural network, autoencoder*, and *Word2Vec* to extract features. Finally, building classification models, such as *ensemble learning, deep learning, multi-label learning, outlier detection, recommended algorithms, etc*.

Honors and Awards

- 2020 National scholarship for Postgraduates
- 2020 Jiangsu Simcere Pharmaceutical Co., Ltd. Scholarship
- 2013-2017 Shenyang Pharmaceutical University：
  Outstanding graduates, Top ten outstanding college students, Advanced individual, Outstanding league member pacesetter
- 2016 Excellent college student in Shenyang
- 2015 3st Prize, National College Students Mathematical Modeling Competition Liaoning Division
- 2015 Shanghai Xinyi Bailuda Pharmaceutical Co., Ltd. Scholarship

Publications

1. **Yanyi Chu**, et al. Predicting drug-target interactions using multi-label learning with community detection method (DTI-MLCD). BioarXiv preprint BioarXiv: 2020.05.11.087734. (2020)
2. **Yanyi Chu**, et al. DTI-CDF: A Cascade Deep Forest Model Towards the Prediction of Drug-Target Interactions Based on Hybrid Features *Brief. Bioinform.* doi:10.1093/bib/bbz152 (2019, **IF: 9.101**)