

Chengyuan Deng

312 Hampshire Court, Piscataway, NJ 08854

🌐 hibara-ai.github.io

✉ charles.deng@rutgers.edu

☎ (+1) 848 3912312

EDUCATION

Rutgers University, New Jersey, USA

Master of Science in Computer Science, GPA: 3.58

Sep 2018-Oct 2020

Tongji University, Shanghai, China

Bachelor of Engineering in Electrical Information Engineering

Minor: Applied Mathematics, GPA: 3.62

Sep 2014-Jun 2018

Hong Kong Polytechnic University, Hong Kong

Exchange student in Computer Science

2017

RESEARCH INTERESTS

Machine Learning Methodology

Graph Neural Network, Variational Autoencoders, etc.

Natural Language Processing

Representation Learning, Machine Comprehension, etc.

Music Information Retrieval

Human Interaction with Music, Modelling Music as Language

(Tentatively in Future)

Intersection of Mathematics and Machine Learning

PUBLICATIONS

(Under Review) C.Wang, **C.Deng**, "GSA-GCN: Global Self-attention Graph Convolution Networks", available on my webpage <https://hibara-ai.github.io/blogs/GSA-GCN.pdf>.

C.Wang, **C.Deng**, V.Ivanov, "SAG-VAE: End-to-end Joint Inference of Data Representations and Feature Relations", arXiv preprint at <https://arxiv.org/abs/1911.11984>.

C.Wang, **C.Deng**, S.Wang, "Imbalance-XGBoost: Leveraging Weighted and Focal Loss for Imbalanced Binary Classification with XGBoost", arXiv preprint at <https://arxiv.org/abs/1908.01672>.

RESEARCH EXPERIENCES

Rutgers University, New Brunswick

Research Assistant at NLP Group, advised by Prof. Karl Stratos

Jan 2020 - Present

- Working on Our-of-vocabulary word representation learning and inference.
- Working on Ordered-neuron LSTM.

Rutgers University, New Brunswick

Self-motivated, three papers published

Aug 2019 - Present

- Imbalance-XGBoost now serves as an open-source python library, guidance and codes are available on github.
- Proposed Self-attention Graph Variational Autoencoders, which can simultaneously learn feature relations and data representations in an end-to-end manner.
- Proposed Global Self-attention Graph Convolutional Networks, which outperformed on benchmark graph kernel datasets in node classification and graph classification tasks. We also provided a theoretical proof that GSA-GCN is able to alleviate over-smoothing.

Recurrent.ai, Remote

Research Intern(NLP), advised by Zhilin Yang

Sep 2019 - Nov 2019

- Pre-processed and trained a new-collected Chinese text-to-speech dataset with Tacotron and Tacotron2
- Introduced a novel approach of leveraging Pinyin in the training process
- Outperformed former results and achieved long-text generation, reduced EOS into 2 seconds.

Tongji University, Shanghai

Undergraduate Research Assistant

Feb 2018 - May 2018

- Implemented Bi-CoPam algorithm, which synthesized three clustering algorithms, namely K-Means, hierarchical clustering and SOM into consensus partition matrices to optimize results, corresponding to specific brain zones.
- Located commonly responded brain zones precisely by training fMRI images of people listening to music pieces, which can be classified into liked and happy, liked and sad, disliked and happy, disliked and sad.

INDUSTRIAL EXPERIENCES

Haystack.ai, Newark

Deep Learning Engineer

Dec 2019 - Present

- Engaged in developing deep learning models from cutting-edge academic papers for real-world applications

Rutgers University, New Brunswick, New Jersey

Teaching Assistant, Intro to Computer Science

Sep 2019–Present

- Graded assignments, exams for computer science undergraduate classes, assisted with course logistics

Tongji Fintech and Big Data Research Institute, Shanghai

Software Engineering Intern

June 2017–Aug 2017

Haitong International Securities, Ltd, Shanghai

Data Analysis Intern

July 2016–Sep 2016

SELECTED PROJECTS

Intersection Congestions Prediction

Kaggle contest

Oct 2019

- Implemented multiple regression models, neural networks, CatBoost, LightGBM, XGBoost to predict waiting time and distance at intersections in four cities: Atlanta, Boston, Chicago and Philadelphia.
- XGBoost outperformed other approaches, leaderboard 25/432.

Hybrid Neural Network Based Movie Recommendation System

Data mining course project, advised by Prof.Yongfeng Zhang

May 2019

- Proposed a novel recommendation system with sliding-window convolution and various neural networks training on different movie features, also integrated matrix factorization methods for comparison.
- The system included following functions: predicting ratings, top n recommendation list, top n similar movies, top n other favorite movies.
- **ESSAY**

Semi-supervised Sound Separation from Single-channel Mixtures

Machine learning, spectral analysis

Mar 2018

- Semi-Supervised Separation of Sounds by Probabilistic Latent Component Analysis, which is essentially a EM-style non-negative matrix factorization algorithm.
- Given sample of one sound, this algorithm learns its frequency features from spectrogram and can factor out and extract it from mixture of other sounds. Experimented on piano and drum sounds.

HONORS AND AWARDS

- Yamaha Asian Music Scholarship of Honorable Mention, **Piano Performance**, 2017 Shanghai.
- ACM Programming Contest, Shanghai Regional, **First Prize**, 2018 Shanghai.
- Mathematical Modeling Invitation of U.S.A. **Second Prize**, 2018 Shanghai.
- National Undergraduate Contest in Mathematical Modeling, **First Prize**, 2017 Shanghai.
- National Undergraduate Contest in Electrical Design, **Third Prize**, 2017 Shanghai.

LEADERSHIP

- **President**, Student Pianist Association of Tongji University, 2017-2018.
Held two anniversary concerts in classical music, presented performances in multiple campus concerts.
- **Program Manager**, Junior Achievement, 2016-2017.
- **Volunteer**, mathematics teacher in elementary school, 2015.