Chengyuan Deng

Address: Hill Center
Department of Computer Science
Rutgers University, NJ, USA

Email: malus.dns@gmail.com
Phone: (+1) 848-391-2312

Education

Rutgeres University, New Jersey, USA — 2021 Fall - present
Ph.D in Computer Science

- Advisor: Prof. Jie Gao

Rutgers University, New Jersey, USA — Sep 2018 - Dec 2020
Master of Science in Computer Science

- **Graduate with Honor:** Outstanding Publication Award
- *Advisor:* Prof. Dipankar Raychaudhuri

Tongji University, Shanghai, China — Sep 2014 - Sep 2018
Bachelor of Engineering in Electronics and Information Engineering
Minor in Applied Mathematics

Highlights

- ◇ School of Arts and Science Research Travel Award, 2025
- ◇ Nomination of IBM Fellowship, 2024
- ◇ Travel Award, *WADS/CCCG 2023*
- ◇ Scholar Award, *NeurIPS 2022*
- ◇ Graduate with Honor: Outstanding Publication Award, *Rutgers University*
- ◇ ACM Programming Contest, Shanghai Regional, **First Prize**, *2018 Shanghai.*
- ◇ Yamaha Asian Music Scholarship, **Piano Performance**, *2017 Shanghai.*

Research Experiences

Applied scientist intern, Amazon
Hosted by Dr.Chengwei Su, Kechen Qin, Emre Barut — May 2023 - Aug 2023

- Topic: Large language models

Research intern, NEC Labs
Hosted by Dr. Zhengzhang Chen — Jan 2023 - May 2023

- Topic: Time-series analysis. Two patents filed.
- Minor topic: Survey on LLM Domain Specification. Paper submitted.

Research Associate & Student Guest, Los Alamos National Lab
Hosted by Dr.Youzuo Lin — April 2021 - present

- Topic: Deep Learning and Computational Imaging. Three papers published.

Selected Publications

"On the Price of Differential Privacy for Hierarchical Clustering". **ICLR 2025**

"Low-sensitivity Hopsets". **ITCS 2025**

"Neuc-MDS: Non-Euclidean Multidimensional Scaling Through Bilinear Forms". **NeurIPS 2024**

"Domain Specialization as the Key to Make Large Language Models Disruptive: A Comprehensive Survey".

"Deconstructing The Ethics of Large Language Models from Long-standing Issues to New-emerging Dilemmas".

"**RIO-CPD**: Correlation-aware Online Change Point Detection". **ICML 2024W**

(In Submission) "LEMMA-RCA: A Large Multi-modal Multi-domain Dataset for Root Cause Analysis".

(In Submission) "Impossibility of Depth Reduction in Explainable Clustering".

"The Discrepancy of Shortest Paths". *International Colloquium on Automata, Languages and Programming*, **ICALP 2024**

"$\mathbb{E}^{FWI}$: Multiparameter Benchmark Datasets for Elastic Full Waveform Inversion of s Geophysical Properties", *Website: efwi-lanl.github.io*, **NeurIPS 2023**

"Evaluating Stability in Massive Social Networks: Efficient Streaming Algorithms for Structural Balance". **RANDOM 2023**

"Differentially Private Range Query on Shortest Paths". **WADS 2023**

"OpenFWI: Large-Scale Multi-Structural Benchmark Datasets for Seismic Full Waveform Inversion", *Website: openfwi-lanl.github.io*, **NeurIPS 2022 (Spotlight)**

"On the Global Self-attention Mechanism for Graph Convolutional Networks", **ICPR 2020 (Oral)**

"SAG-VAE: End-to-end Joint Inference of Data Representations and Feature Relations", **IJCNN 2020 (Oral)**

"Imbalance-XGBoost: Leveraging Weighted and Focal Loss for Imbalanced Binary Classification with XGBoost", **Pattern Recognition Letter**

Selected Projects

Distribution Testing in Multi-pass Streaming Model
Advised by Prof. Sepehr Assadi, course project for Sublinear Algorithms

- Literature Review of Distribution Testing in Single-pass Streaming Model and Distributed Communication Model.
- Proved a lower bound in Multi-pass Streaming Model for Uniformity Testing

Imbalance-XGBoost
Self-motivated

- **Open-source python library**, availabe on Github and PyPi, **star 240+**.
- The library leverages weighted and focal loss for imbalanced binary classification with XGBoost. State-of-the-art performances were achieved on a recently collected Parkinson disease dataset by Focal-XGBoost. Paper published.

(Kaggle) Intersection Congestions Prediction
Advised by Prof. Saed Sayad

- 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**.

Micro-Me
Collaboration with designers at University of College in London

- Implemented Faster-RCNN, YOLO3 and Facebook Object Detection API with Tensorflow on image datasets of foods and fruits, the models are able to detect 110 kinds of foods and fruits with accuracy of 90+%.
- Built a web app for demo using Bootstrap and Flask. The app is able to detect foods in the uploaded image and return fibre content.

Industrial Experiences

Machine Learning Intern, Newark — Jan 2020 - Mar 2020
Haystack.ai

- Engaged in developing deep learning models from cutting-edge academic papers for real-world applications
- Example Project: Selfie-to-anime. Collected anime images for training, implemented a cutting-edge paper published in ICLR 2020 "Unsupervised Generative Attentional Networks" and built the API with Flask.

Data Analysis Intern, Shanghai — Jun 2016 - Sep 2016
Haitong Securities, International

- Analyzed the daily stock quotation and cyclical data by setting up models then predicted trends
- Proposed financial models for cutting-edge companies and wrote reports, with **200**+ pageviews daily

Talks

◇ "Low Sensitivity Hopsets", NYC Theory Day

◇ "The Discrepancy of Shortest Paths", DIMACS Theory of Computation Seminar

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.*

Professional Services

◇ Reviewer, SIGIR 2023, KDD 2023, NeurIPS 2023

◇ Reviewer, IEEE Transaction on Neural Networks and Learning System

◇ Student Membership, ACM (Association for Computing Machinery)

◇ Student Membership, IEEE (Institute of Electrical and Electronics Engineers)

Teaching

◇ Introduction to Programming, Rutgers University Fall 2019

◇ Computer Architecture, Rutgers University Fall 2021