

Hangrui Cao

☎ (+1)734-604-2766 | ✉ hangrui@umich.edu | 🏠 DiegoCao.github.io | 📷 DiegoCao

Education

University of Michigan

B.S.E. COMPUTER SCIENCE, MINOR IN MATH

Ann Arbor, Michigan

Sep 2020 - May 2022(expected)

- **Cumulative GPA: 4.0/4.0, Major GPA: 4.0/4.0, Minor GPA: 4.0/4.0**
- **Selected Courses:** Machine Learning(A), Computer Vision(A), Data Structure & Algorithms (A), Database & Management System(A+), Computer Organization (A), Computer Foundations (A), Matrix Algebra (A), Numerical Analysis (A), Computer Networks (ongoing), Conversational AI (ongoing), Human Computer Interaction (ongoing)

University of Michigan - Shanghai Jiao Tong University Joint Institute

B.E. ELECTRICAL AND COMPUTER ENGINEERING

Shanghai, China

Sep 2018 - Aug 2022(expected)

- **Cumulative GPA: 3.874/4.0, Ranked 8/248; Major GPA 3.892, Ranked 8/248**
- **Selected Courses:** Programming and Intro to Data Structures (A+), Intro to Computer and Programming (A+), Intro to Engineering (A+), Honors Mathematics III(A+), Honors Physics I & II (A), Discrete Math (A)

Research Interest

- Machine Learning; Large-scale data analysis and data mining; Federated and Distributed Learning; Optimization in ML

Publication

- **Hangrui Cao***, Qiying Pan*, Yifei Zhu. "Birds of a Feather Help: Context-aware Client Selection for Efficient Federated Learning". 2021. (In Submission to FL-AAAI-22) [PDF]
- **Hangrui Cao**, Xinyu Lu, Zhima Su, Jing Wu, Chengnian Long. "Simulation of DoS Attack in Networked Control System Using TrueTime". 2020. (2020 CAC)

Research Experience

Research Assistant, Foreseer Group of University of Michigan

ADVISOR: PROF. QIAOZHU MEI

Ann Arbor, Michigan

April 2021 - Present

Large data analysis and prediction with Github Archive Dataset | Data Mining, Large-scale data Analysis, Text embedding, Social network embedding

Description: We hope to present a large-scale empirical study to how sentimental emoji and text usage relates to Github workers' behaviour with different metrics

- Constructed social network based on Github 2018 year dataset, applied GNN on social graph and cluster embedding vector to build user working groups
- Implemented large-scale Github dataset processing with Spark and Hadoop and extract data of emoji usage and work activities, calculated the correlation between metrics of working groups' productivity and their emoji usage data
- Classified users into working groups by cluster GraphSage Embedding and extracted correlation between working stats and emoji usage.
- Applied cluster, regression and PCA on the extract work/emoji usage conversation vectors and further do prediction for users' work

Research Assistant, Intelligent Networked System Lab of Shanghai Jiao Tong University

ADVISOR: PROF. YIFEI ZHU

Shanghai, China

Sep 2020 - June 2021

Dependence-aware client selection for federated learning | Federated Learning, Contextual Bandit, NLP, Distributed ML

Description: The research project aims at resolving non-IID issues of clients local data in federated system, improving federated learning performance by using a context-aware selection algorithm to select clients with high data utility

- Invented and implemented a combinatorial context-aware client selection algorithm combined with SimHash to select clients of with high utility to reduce non-IID data in each federated round, and increased global model's final accuracy by 17%, speed to reach target accuracy 51% compared with benchmark FedAvg, 4.6% final accuracy and 7.1% speed improvement compared with state-of-art method Oort
- Conducted literature review on different federated learning papers on client selection and communication efficiency, and implemented relevant codes for the project, including the task of train and test, standard benchmark(FedAvg), state-of-art method comparison (Oort) and other related benchmarks like clusterFL.
- Processed social media NLP dataset (Stackoverflow, Yelp, and Superuser) for task of next word prediction and implemented stacked-LSTM for local FL model
- Finished writing the first author paper submitted to FL-AAAI-22

DEVIATE Group, Transportation Research Institute, University of Michigan

ADVISOR: PROF. CAROL FLANNAGAN, DR. KATHLEEN KLINICH

Ann Arbor, Michigan

Jan 2021 - Present

CNN of Shaky Ground Truth | Data Analytics, CNN for driver behaviors detection, Uncertainty in ML model, Shaky Ground Truth

Description: The project focused on developing methods for automating video data extraction to record vehicle occupant behavior and reduce bias, and used context increase network training accuracy and efficiency, with further information on Deviate website.

- Performed data analysis of detected object detection result (Yolo model output) and human-label result (with heatmap and other plots) to explore shaky ground truth by labelers' disagreement
- Re-designed neural network model to train and predict based on proportional vector of multi-classes
- Implemented benchmark methods like Bayesian CNN to compare neural network approaches with our designed model
- Conducted eye gaze data analysis via GazeRecorder, PyGaze and GazeVisual and compared their performance

Research Assistant, Automation Department of Shanghai Jiao Tong University

Shanghai, China

ADVISOR: PROF. JING WU

Feb 2019 - Aug 2020

Simulation of Networked Control Systems Using Truetime | Networked Control Systems, Truetime Simulation, Neural Network**Description:** Simulate the network control system based on Truetime and apply different methods of design components in the system, and apply neural network system

- Developed a network control system to perform simulation based on MATLAB and Truetime Toolbox, and Simulated the denial-of-service attack on the system
- Applied neural network to analyze data from National Grid to fit the plant model.
- Implemented inverted pendulum system to test the effect of Dos attack on performance of the system and design effective controllers to enhance system's robustness and stability under attack.
- Wrote First-Author Paper published in China Automation Conference 2020.

Selected Projects**5G & SuperResolution Research Project**

Ann Arbor, Michigan

ROBUSTNET GROUP, UNIVERSITY OF MICHIGAN, ADVISED BY ADVISED BY PROF. MORLEY MAO

Oct 2021 - present

- Simulated different size of BasicVSR model on Redis video dataset,
- Implemented web-browsing function of App in Java to work on android 5G environment
- Working on deploying trained BasicVSR into tflite model, and inference models on android devices

Enhancing Text-Entry Experiences for Better Interaction Between Instructor and Students in Synchronous Hybrid Class

Ann Arbor, Michigan

RESEARCH COURSE GROUP PROJECT

Sep 2021 - Present

- Designed surveys and gathered data from remote and in-person student groups respectively
- Created prototype interface to support the user requirements of the remote students
- Performed simplistic user evaluation and heuristic evaluation on remote students and instructors to evaluate prototype design

Email Voice Assistant

Ann Arbor, Michigan

CAPSTONE PROJECT

Sep 2021 - Present

- Processed Email Request on the backend using Flask
- Designed User Interface and processed user input with Tkinter, React & Pyaudio (user voice input, text to speech)
- Utilized Dialogflow and other conversational AI tools to generate smart reply to users

Automatic Image Colorization with CNN and GAN

Ann Arbor, Michigan

COURSE GROUP PROJECT

Jan 2021 - May 2021

- Designed and implement GAN(Generative Adversial Networks) model structure, training and testing via Python codes.
- Compared GAN method with classification colorization by metrics of PSNR and SSIM
- Conducted literature review into different cutting-edge methods of image colorization via deep learning and wrote project paper.

Work Experience**Deep Learning Research Intern, Intel Corporation**

Intel Asia, Shanghai, China

MANAGER JASON DAI, MENTOR QIYUAN GONG

May 2021 - Aug 2021

Project: PPML(privacy preserving machine learning) and BigDL(Big Deep Learning) of Analytics-zoo

- Handled Intel-SGX product update support for new graphene 1.2 RC to support PPML
- Implemented classical and deep learning model to support PPML in Analytics including XGBoost Regressor, XGBoost Classifier, Yolo and ResNet in BigDL
- Implemented multiple unit tests for testing new machine learning training tasks in PPML.
- Applied Docker and container for support new version product bundles.

Selected Awards

2021	Dean's List , EECS Department, University of Michigan	Michigan, U.S
2021	University Honor , University of Michigan	Michigan, U.S
2020	Muriel & Jackson Lum Scholarship , University of Michigan	Michigan, U.S
2020	Undergraduate Excellence Scholarship , Shanghai Jiao Tong University	Shanghai, China
2019	John Wu & Jane Sun Scholarship , Shanghai Jiao Tong University	Shanghai, China

Skills**Programming** Python, C++, C, SQL, JAVA, Scala, JavaScript, MATLAB, LaTeX**FrameWork** Pytorch, Tensorflow, OpenCV, Networkx, Spark, Sqlite, Flask**Other**

2020	Teaching Assistant , Intro to Computer & Programming (VG101), Academic Writing I & II	Shanghai, China
2020	Advisor in UM-SJTU Advising Center , 3-hour advising for undergraduates, organize free lectures	Shanghai, China
2020	Volunteer Teaching , volunteer in Eryuan, Yunnan, China as primary school teachers	Dali, Yunnan, China