

HANG GAO

http://hangg.me

Mobile : 646-460-5552

Email : hang.gao@columbia.edu

Address : 209 W 108th Street, NY. NY, 10025

EDUCATION

- **Columbia University** New York, NY
M.S. in Computer Engineering *Aug 2017 – Dec 2018 (Expected)*
- **Shanghai Jiao Tong University (SJTU)** Shanghai, CN
B.E. in Computer Science; Major GPA: 3.82/4.0, Overall GPA: 3.62/4.0 *Sep 2013 – Jul 2017*
 - **Core Courses:** Introduction to Programming in C&C++ (100), Information Theory (100), Digital Signal Processing (100), Artificial Intelligence (99), Data Mining (95), Probability Theory and Statistics (94), Machine Learning (93), Algorithm Design and Analysis (92)
- **Carnegie Mellon University** Pittsburgh, PA
Full-time visiting research intern in Prof. Anind K. Dey's group *Jul 2016 – Sep 2016*

PUBLICATIONS

- **DeepCoach: An Online Correction System for Aggressive Drivers Leveraging Contextual Information**
N. Banovic, H. Gao, V. Rao, J. Mankoff, A. Dey *Submitted to ACM International Joint Conference on Pervasive and Ubiquitous Computing (UBICOMP), 2017.*
- **ER: Early Recognition of Inattentive Driving Events Leveraging Audio Devices on Smartphones**
H. Gao, X. Xu, J. Yu, Y. Chen, Y. Zhu, G. Xue, and M. Li. *IEEE International Conference on Computer Communications (INFOCOM) 2017.*
- **Distribution Modeling for Smart Citys Taxi Resources Allocation**
H. Gao, F. Qiu, N. Yang. *International Journal of Engineering Mathematics (Best Undergrad Paper), 2016.*

EXPERIENCE & PROJECTS

- **4 Paradigm Co., Group of Machine Learning and Algorithm** Beijing, China
Research Intern in Machine Learning | Supervised by: Yuqiang Chen and Weiwei Tu *May 2017 – Present*
 - **Diabetes Prediction:** Inferred on a long-term-retrieved healthcare dataset consisting of 1m records for diabetes prediction, using Gradient Boosting Decision Tree (GBDT) on large clusters.
 - **Knowledge Distilling:** Distilled complex model into human-readable rules to guild clinical prescriptions. Beated long-standing rule of thumbs in related area with simpler sketch but better AUC.
 - **Multi-task Learning:** Proposed a MTL-based GBDT algorithm which improved performance in all 21 subtasks using novel multi-task boosting techniques with sparse feature encoding.
 - **Large Scale Tuning:** Experimented with multiple settings of more than 500 cross validations on Hadoop cluster to prove superiority of MTL over STL using T-test.
- **Alibaba AI Lab, Group of Machine Learning** Hangzhou, China
Research Intern in Data Mining | Supervised by: Jichao Zhen *Feb 2017 – May 2017*
 - **Distributed Crawler System enabled of By-passing Firewalls:** Developed and deployed a highly distributed crawler system which would by-pass most of firewall blocking. Built up a dataset of personal information with more than 100 million records in less than two weeks.
 - **Lifelong Learning for Abnomaly Detection:** Improved detection of abnomal user behaviors along with the data collection of crawlers. The model was able to discover new category during its life span.
- **Google, Group of Ad Serving** Shanghai, China
Engineer Intern in Machine Learning | Supervised by: Yu Shen *Oct 2016 – Jan 2017*
 - **Click-Through Rate (CTR) Prediction:** Proposed multi-stage strategy to maximize online performance by first using offline model ensembling, then building LR/FM on top of soft probabilities of base learners.
- **Carnegie Mellon University, Lab of Ubiquitous Computing** Pittsburgh, PA, USA
Research Assistant in Deep Learning | Supervised by: Prof. Anind K. Dey *Jul 2016 – Sep 2016*
 - **Online Correction for Aggressive Drivers:** Built a smartphone-based online correction application for aggressive driver by predicting driving states sequence based on individual history and street information. According work has been submitted to UBICOMP'17.

- **End-to-end LSTM Framework:** Developed and abstracted an end-to-end LSTM framework - EasyLSTM, for tasks on temporal signals. The work was largely inspired by workflow of online correction model training.
- **Dashboard Application with RESTful Services:** Set up RESTful services to track remediation process, along with an admin application to maintain sensors on the fly. Note that both Proj 2 and 3 has been incorporated into Giotto Infrastructure at CMU.

• **Shanghai Jiao Tong University, Lab of Intelligent Perception**

Shanghai, China

Research Team Leader in Machine Learning | Supervised by: Prof. Minglu Li

Aug 2015 – Jul 2016

- **Inattentive Driving Detection:** Put forward a fine-grained and low-cost inattentive driving detection system using grouped SVM. The work has been published on INFOCOM'17.
- **Exclusive Audio-based System:** Eliminate restriction on smartphone posing as such in traditional Computer Vision methods by transferring raw sound signal into doppler representations.
- **Early Recognition:** Designed an early recognition algorithm by measuring distance between action points on feature space. Achieved over 90% accuracy using data of about half the gesturing duration.

• **Shanghai Jiao Tong University, Lab of Artificial Intelligence and Smart City**

Shanghai, China

Research Team Leader in Mathematical Modeling | Supervised by: Prof. Xiaotie Deng

Aug 2015 – Apr 2016

- **Optimization of Taxi Resource Allocation:** Proposed a dynamical optimization algorithm for allocating taxi though out urban area based on current and historical records of supply and demand distribution. The task has been modeled as a non-trivial Reinforcement Learning problem.
- **On-policy Control with Trace Approximation:** Simulated taxi population by Monte Carlo Method given centre set (state) and allocation (action) set and update policy iteratively.
- **Tradeoff Balancing:** Explored the optimal policy of subsidy strategy for vendor platform (like Uber) by simulating bilateral dynamics between passengers and drivers.

HONORS & TEACHING

- **Course Assistant, COMS 203 (Probability Theory and Statistics)** *2017*
- **Rong Chang Scholarship (5/80) for academic excellence of SJTU** *2016*
- **Finalist Winner Prize (~1%) of Mathematical Contest in Modeling** *2016*
- **Gold Medal (7/400) of China Undergraduate Mathematical Contest in Modeling** *2015*
- **National Scholarship (~3%)** *2015, 2016, 2017*
- **Silver Medal (10/370) of Intel International Science and Engineering Fair (ISEF)** *2012*

PROGRAMMING SKILLS

- **Programming Languages:** Python, C++, C, Java, Bash, L^AT_EX, SQL, HTML/CSS, JavaScript, C#, Verilog
- **Libraries:** Pytorch, Keras, Caffe, OpenCV, XGBoost, Scrapy, Flask, Android
- **Tools & Platforms:** Vim, Visual Studio, IntelliJ, TMUX, CMake, Redis, MongoDB, Hadoop, Docker, FPGA, git, ssh

EXTRACURRICULUM

- **Shanghai International Marathon** Shanghai, China
Top 200 over more than 3000 participants consisting of fans, semi-pros, and professionals *2016*
- **weMet - Welcoming Gala for 2018er, SJTU** Shanghai, China
President director of the gala, taking charge of more than 80 people *2014*
- **Volunteer Society Nepal** Katmandu, Nepal
Volunteer teacher in Chinese and Mathematics *2013*