Chenhui Hu

chenhui.hu@gmail.com
LinkedIn: chenhui-hu-51066519
(617)-913-6436
Lexington, MA

Experienced machine learning and data science professional with excellent leadership, interpersonal, and product development skills. Seeking for principal/senior level machine learning engineering, data science, or software engineering roles.

Work Experience

Data Science Tech Lead **June 2020 — Present**
Azure Global, Microsoft Cloud & AI *Cambridge, MA*

- Leading the development of a platform for demand forecasting and supply chain optimization adopted by 40+ ISVs & customers
- Helped our pilot customers achieve 30% reduction of time to build forecasting models on average
- Responsible for architecture design, implementation, and testing of the product and launching it on Azure Marketplace
- Working across 4 teams to integrate internal tools developed by other teams into the product for building forecasting pipelines based on common data model, model interpretability, and forecasting millions of time series in parallel
- Contributed to the improvement of the tools by issuing 6 pull requests to fix bugs and providing 12 feedback items

Data Scientist II **Sept. 2017 — May 2020**
Azure Data Science, Microsoft Cloud & AI *Cambridge, MA*

- Led the design and development of Microsoft Forecasting repo containing 15 best practice examples and a Python utility library with 2.2K stars (most popular repo in time series forecasting)
- Made major contribution to the winning of a multi-hundred-million deal with one of the largest multinational retailers by building an end-to-end energy forecasting solution for over 11k stores using Spark and Azure IoT, reducing hourly forecast error from 8.5% to 5.5% through model optimization, and achieving 20x speed up of model training using multi-thread computation
- Contributed to the winning of 10-billion JEDI contract with U.S. Government by building a large-scale predictive maintenance solution using synthetic data to showcase Azure capability

Data Scientist **June 2016 — Aug. 2017**
Azure Data Science, Microsoft Cloud & AI *Cambridge, MA*

- Developed a customer churn prediction model for Portland Trail Blazers using AzureML Workbench and improved campaign conversion rate from 4% to 24.8% (covered by The TWIML AI Podcast)
- Built new product demand forecasting models for a food manufacturer to help them improve the forecasting accuracy by 15%

Software Engineer **Apr. 2010 — Aug. 2010**
Intel Corporation *Shanghai, China*

- OProfile tool development for ARM Cortex-A8 system using Shell scripts

Education

Harvard University, *Ph.D. in Electrical Engineering* May 2016
Shanghai Jiao Tong University (SJTU)
M.S. in Communication and Information Systems Mar. 2010
B.S. in Electrical Engineering (Honor Class) June 2007

Selected Projects

NLP for Recommendation System **Aug. 2021 — Nov. 2021**
Improved 13.6% F1-score of product recommendation by fine-tuning RoBERTa model on an online retail dataset and an Amazon dataset to understand the product similarities from descriptions

Predicting Poverty using Satellite Images **Sept. 2020 — Nov. 2020**
Achieved an R2 score of 72% of predicting poverty levels in three African countries using satellite images, transfer learning, and CNN model; explored feature importance of the CNN model using perturbation based algorithm

Ensemble Model for Malicious URL Detection **Mar. 2018 — Apr. 2018**
Obtained 97% accuracy and 92% F1-score of malicious URL detection by creating an ensemble model with TF-IDF features, feature selection, data resampling and model tuning

Skills

Tools and Languages	Python, R, Spark, C/C++, Git, DevOps, MLOps, SQL, Unix Shell
Machine Learning	Deep Learning, NLP, Computer Vision, Reinforcement Learning, AutoML, Responsible AI

Research

Research Assistant **June 2012 — Apr. 2016**
Massachusetts General Hospital & Harvard Medical School *Boston, MA*

- Inferred brain network structure by learning graphical models from terabytes of medical images
- Improved classification accuracy of Alzheimer's disease (AD) by 21% using graph-signal processing and deep neural networks
- Studied progression and origin of AD by proposing a network diffusion model with impulsive sources

Research Assistant **Sept. 2010 — May 2012**
School of Engineering and Applied Sciences, Harvard University *Cambridge, MA*

- Devised graph-based regularization algorithms for color image demosaicking
- Invented an adaptive binary sensing scheme to enhance the dynamic range of imaging by over 20 times

Research Assistant **Sept. 2007 — Mar. 2010**
Institute of Wireless Communication Technology, SJTU *Shanghai, China*

- Participated in 3 projects founded by Qualcomm Inc. and National Natural Science Foundation of China
- Designed efficient multicast protocols for mobile wireless networks and derived theoretical performance bounds

Team Leader **Mar. 2006 — Aug. 2007**
Lab of Cybernetic Technologies for Cars in Chinese Cities, SJTU *Shanghai, China*

- Designed pattern recognition and control algorithms for self-driving cars equipped with Freescale S12 MCU, CCD camera, and speed sensors to adapt the car to complex road conditions
- Designed technical strategies with supervisors for competitions and coordinated with 5 team members

Selected Awards

2014, **The 3rd IEEE ComSoc Asia-Pacific Outstanding Paper Award**

2011, Mou-Shiung Lin Fellowship, Harvard University

2010, Outstanding Graduate of Shanghai Government

2007, **Championship, National Intelligent Car Contest of College Students in China (242 teams)**

2005 - 2006, Excellent Academic Scholarship (5%, Twice) of SJTU, China

2005, The 1st Prize of Math Contest in Modeling of East China (1%)

Selected Publications

- **Chenhui Hu**, V. Paunic, "**Building Forecasting Solutions Using Open-Source and Azure Machine Learning**," *KDD*, 2020.
- R. Ju, **Chenhui Hu**, P. Zhou, Q. Li, "**Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning**," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2017. (cited by 86)
- **Chenhui Hu**, X. Hua, J. Ying, P. Thompson, G. Fakhri, Q. Li, "**Localizing sources of brain disease progression with network diffusion model**," *IEEE Journal of Selected Topics in Signal Processing*, 2016. (cited by 28)
- **Chenhui Hu**, R. Ju, Y. Shen, P. Zhou, Q. Li, "**Clinical decision support for Alzheimer's disease based on deep learning and brain network**", in *IEEE International Conference on Communications (ICC)*, 2016. (cited by 58)
- **Chenhui Hu**, J. Sepulcre, K. Johnson, G. Fakhri, Y. Lu, Q. Li, "**Matched signal detection on graphs: theory and application to brain imaging data classification**," *NeuroImage*, 2016. (cited by 35, conference version cited by 20)
- **Chenhui Hu**, L. Cheng, J. Sepulcre, K. Johnson, G. Fakhri, Y. Lu, Q. Li, "**A spectral graph regression model for learning brain connectivity of Alzheimer's disease**," *PLOS ONE*, 2015. (cited by 30, conference version cited by 38)
- **Chenhui Hu**, X. Wang, D. Nie, J. Zhao, "**Multicast scaling laws with hierarchical cooperation**," in *IEEE International Conference on Computer Communications (INFOCOM)*, 2010. (cited by 32, journal version cited by 79)
- **Chenhui Hu**, X. Wang, F. Wu, "**MotionCast: on the capacity and delay tradeoffs**," in *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2009. (cited by 64, journal version cited by 215)

Activities

Presentations: KDD 2020, AI Conference 2019, Strata 2019, MLADS 2016-2019, MICCAI 2014, ISBI 2013
Reviewer: PLOS ONE, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions of Biomedical Engineering, IEEE Transactions on Wireless Communications, IEEE Transactions on Circuits and Systems for Video Technology, etc.
Volunteer: 100 Black Men of America, Winchester School of Chinese Culture, Chinese American Association of Lexington, etc.