Esha Wang

esha.wang@gmail.com | 309-530-9780 | Jersey City, NJ
linkedin.com/esha-wang | github.com/eshawang | Personal Blog

Education

2018 – 2020 **M.Sc. Statistics**
Stanford University, GPA: 3.9

2011 – 2015 **B.S. Applied and Computational Mathematics**
California Institute of Technology, GPA: 3.5

Employment History

Senior Data Scientist, May 2022 – Present
Hartford Steam Boiler (Hartford, CT)
– Technical lead for Location Risk Score (LRS), a tool that blends XGBoost and CatBoost to evaluate riskiness of insuring a particular location. Oversaw project from inception to deployment, resulting in LRS being widely used by internal groups and external paying customers.
– Lead model developer for IoT sensor project, implemented Isolation Forest for anomaly detection, evidential kNN for time series classification, and Time2Normalcy following customer alert using causal analysis.
– Created HSB data science team's first official gold coding standards, including rules and guidelines for documentation, repository organization, code review, and coding style best practices.
– Mentored and created development plans for interns and junior data scientists of various backgrounds.

AI/ML Radio Systems Research Scientist, Jan 2021 – Apr 2022
Nokia Bell Labs (New Providence, NJ)
– Designed environment-invariant CNN models to classify indoor human activity using WiFi Channel State Information, with peak accuracy of 78% for unseen person, movement pattern, and furniture orientation. Full responsibility over data collection, cleaning, and processing using NumPy, scikit-learn, and pandas. Presented live demo with real-time labeling of unscripted human activity to Bell Labs head and various business units.
– Contributed to DeepRx, a CNN-based deep learning receiver that takes in a received signal as input and outputs the original transmitted 0/1 bits. Personal contribution resulted in up to 10% improvement in accuracy.
– Developed traffic sensing project, which uses millimeter-wave data to detect location, size, and speed of vehicles at a traffic intersection. Tools used include Facebook's Detectron2, Faster R-CNN, and semantic segmentation.

Software Engineer II, Aug 2016 – Sep 2018
Oracle Corporation (Redwood City, CA)

Applications Developer, Aug 2015 – Jun 2016
SBB Research Group LLC (Chicago, IL)

Independent Projects & Activities

New England Statistics Symposium (NESS) 2023 Statathon Judge
Created data science challenge for Statathon and represented HSB as competition judge. Link to challenge here.

Sarcasm Detection
Implemented content embedding CNN for generalized sarcasm detection over several corpus datasets, including Twitter/Reddit posts and The Onion/HuffPost news headlines. Link to project proposal here and final report here.

Facial Expression Recognition
Developed predictive models using baseline 3-layer CNN and ResNet50, and fine-tuned models using pretrained weights and image augmentation, powered by a single Nvidia GTX 960 GPU. Link to report here.

Effectiveness of Speed Dating
Compared linear/logistic regression, naive Bayes, and decision trees with analysis of normal/clustered bootstrapping, Bonferroni Correction and Benjamini-Hochberg, and statistical significance of coefficients. Link to report here.

Skills

Coding Python, R, SQL, Git, Azure DevOps.

ML & Stats Azure Databricks, Spark/PySpark, TensorFlow, PyTorch, Hyperopt, Optuna, CART, gradient boosting, neural networks, unsupervised learning, A/B testing, statistical inference, causal inference.

Misc. LaTeX, photography, videography, blogging.