# ESHA WANG

615 Pavonia Ave, Apt 3506
Jersey City, NJ 07306
esha.wang@gmail.com
(309) 530-9780

## EDUCATION

Sept 2018 – Apr 2020 **Stanford University** **Stanford, CA**
M.S. in Statistics, GPA: 3.9/4.0

Sept 2011 – Jun 2015 **California Institute of Technology** **Pasadena, CA**
B.S. in Applied and Computational Mathematics, GPA: 3.5/4.0

## WORK EXPERIENCE

Jul 2019 – Sept 2019 **Nokia Bell Labs** **Sunnyvale, CA**
***FxIn Machine Learning Summer Research Intern***
- Designed environment-invariant CNN models using PyTorch and CUDA to classify indoor human activity using WiFi Channel State Information data, with accuracy of 78% on test data with arbitrary furniture orientation.
- Full responsibility over data collection, cleaning, and processing using NumPy, Pandas, and sklearn.
- Presented live demo with real-time labeling of unscripted human activity to various teams across Bell Labs. Model correctly labeled all activities over three minutes of live testing.
- Conducted literature review of state-of-the-art research in unsupervised learning, and tested capabilities of k-means clustering and Gaussian mixture models for human activity recognition.

Aug 2016 – Sept 2018 **Oracle Corporation** **Redwood City, CA**
***Software Engineer II***
- Improved Oracle Linux reboot functionality by making executable file addresses position-independent.
- Fixed security vulnerabilities in Oracle Linux and Solaris 12 using RPM SPEC files and bash tools.
- Upgraded Free and Open Source Software components to latest versions and integrated them into Solaris 12.

Aug 2015 – Jun 2016 **SBB Research Group LLC** **Chicago, IL**
***Applications Developer***
- Used Python and R to implement mathematical models to estimate implied volatilities of structured notes.
- Analyzed SPX/RUT data using MySQL to predict future behavior of indices with inter/extrapolation techniques.

## RELEVANT PROJECTS

Winter 2020 **The Best Sarcasm Detector Ever Created (CS224n, Stanford)**
- Used PyTorch to developed content embedding CNN for generalized sarcasm detection over multiple corpus datasets, including Twitter/Reddit posts and The Onion/Huffington Post news headlines.
- Project proposal linked here (Google Drive) or here (GitHub).

Winter 2020 **Natural Language Processing Mini Projects (CS224n, Stanford)**
- Implemented skip-gram word2vec embedding model, neural transition-based dependency parser, and neural machine translation RNN with bidirectional LSTM and character-based encoder.

Fall 2019 **Last.fm Package: Analysis of Music Listening Habits in R (STATS290, Stanford)**
- Wrote R package from scratch using Last.fm API to query, visualize, and recommend music based on a user's listening history, including full documentation, use cases, and comprehensive tests.
- Functionalities include pulling listening history and track/album/artist information, and generating world map of artists' origins, word cloud of most listened to music genres, and new music recommendations.
- Report linked here (Github).

Fall 2018 **Facial Expression Recognition (CS229, Stanford)**
- Developed predictive models including a baseline 3-layer CNN, ResNet50, and fine-tuned models using pre-trained weights and image augmentation, powered by Keras and a single Nvidia GTX 960 GPU.
- Performance of 66.1% accuracy on the FER2013 dataset reached #5 on the 2013 Kaggle leaderboard.
- Report linked here (Google Drive) or here (GitHub).

Fall 2018 **Machine Learning Mini Projects (CS229, Stanford)**
- Implemented prediction of daily website traffic (Poisson regression), spam classification (naive Bayes), image compression (k-means), handwritten digit classification (CNN), and mixed audio component extraction (ICA).

Fall 2018 **Predicting the Success Rate of Speed Dating (MS&E226, Stanford)**
- Developed and compared several predictive models, such as linear regression, logistic regression, naive Bayes, and decision trees/random forests, with best model being linear regression with interaction terms.
- Statistical inferences conducted include normal and clustered bootstrapping, multiple hypothesis testing such as Bonferroni Correction and Benjamini-Hochberg, and statistical significance of coefficients.
- Report linked here (Google Drive) or here (GitHub).

Spring 2014 **Netflix Prize (CS156, Caltech)**
- Leveraged Singular Value Decomposition (SVD), SVD++, time SVD++, Restricted Boltzmann Machine, and the Weekday Model to achieve successful algorithm blend, surpassing Netflix's original algorithm by 7.75%.

## SKILLS & COURSES

Machine Learning, Computer Vision, NLP, Stochastic Processes, ANOVA, Optimization, Bayesian Statistics, Statistical Consulting

## PROGRAMMING LANGUAGES/TOOLS

Proficient – Python, R/R Markdown, PyTorch, CUDA, conda, git  Experienced – SQL, C/C++, Tableau