Shuyang Li

shuyangli94@gmail.com ▪ shuyangli.me ▪ (314)-660-6660
Office #4146, 9500 Gilman Dr, San Diego, CA 92122

EDUCATION

University of California, San Diego, San Diego, CA (PhD)   *Exp.* 2023
    Area: Computer Science and Engineering, Machine Learning
Princeton University, Princeton, NJ (BSE)   June 2016
    Major: Operations Research and Financial Engineering   GPA: 3.71
    Certificates: Applications of Computing; Statistics and Machine Learning
    Thesis: *Exploring Rich Features for Sentiment Analysis with Various Machine Learning Models*

PUBLICATIONS

Majumder, B.*, **Li, S.***, Ni, J., McAuley, J. (2019). Generating Personalized Recipes from Historical User Preferences. *Empirical Methods in Natural Language Processing.*

Huang, D., **Li, S**., Dhaka, A., Story, G.M. and Cao, Y.Q., 2012. Expression of the transient receptor potential channels TRPV1, TRPA1 and TRPM8 in mouse trigeminal primary afferent neurons innervating the dura. *Molecular pain*, 8 (1), 66-85.

RESEARCH AND WORKING EXPERIENCE

University of California San Diego, AI Lab   September 2018-Present
Graduate Student Researcher
- Researching and developing models for structured, data-to-text, and dialogue text generation
- Researcher, architect, and lead developer for the UCSD Amazon Alexa Prize Socialbot Grand Challenge team
- Researching models for dialogue modeling and recommender systems under Professor Julian McAuley

Google, Kaggle Datasets (Google Cloud AI)   June 2019-September 2019
Software Engineering Intern (PhD)
- Built a framework for automatically generating semantic tags for datasets based on free-text metadata
- Implemented metrics for dataset discoverability and search success
- Doubled size of tag ontology and tripled tag coverage across all public datasets on Kaggle.com

Bloomberg, Structured Products Waterfall   June 2017-September 2018
Senior Software Engineer
- Designed and implemented a Spark-based infrastructure for high bandwidth data processing jobs
- Modeled time series data for 2.5 million asset-backed securities (ABS)
- Designed and implemented anomaly detection for ABS data passing through ETL and analytics pipelines

Goldman Sachs, Operations Automation and Analytics Technology   July 2016-June 2017
Technology Analyst
- Helped build an application to consume, parse, & standardize market messages for fixed income instruments
- Designed & built a machine learning platform to create metrics & predictive models for Operations division

Princeton University, Senior Thesis Research   September 2015-May 2016
Senior Thesis Research
- Investigated SVMs, Naïve Bayes, and ensemble methods for binary sentiment analysis on movie reviews
- Created a manually labeled corpus from 2004 Cornell IMDB data for subjective and summary sentences

Goldman Sachs, Operations Analytics Strategies   June 2015-August 2015
Summer Analyst
- Led team working on automated invoice recognition using Tesseract and Python to process and automate template matching with noisy and tilted images containing structured and semi-structured text

Princeton Laboratory for Energy Systems Analysis / CASTLE Lab   June 2014-August 2014
Summer Research Intern
- Created simulator in Java for the unit allocation problem of introduction of wind energy to the power grid
- Analyzed the role and performance of different classes of learning rates in reinforcement learning