Sid Mangalik

Richmond, VA 23238 | 508-762-8390
mangaliks@gmail.com | linkedin.com/in/sid-mangalik | smangalik.github.io

Education

Stony Brook University | Stony Brook, NY — May 2018
Bachelor of Science, Computer Science — GPA: 3.65/4.0

- One of ten Honors Computer Science program members
- Concentration in Human-Computer Interaction
- Two year-long theses in Natural Language Processing

Projects

Cryptocurrency Miner: A mining node for a BitCoin analog using blockchain in Java, by using concurrency the miner was able to compete and successfully mine blocks faster than competitor miners.
Language Modeler: NLP Pipeline with python for generating POS tagging, PCFG parsing, and n-grams which recreated the functionality present in most NLP tagging and mapping pipelines.
Predator Detection: NLP Research into whether chat participants were predators with Stanford CoreNLP which furthered research insight into how models of control and abuse can be represented using linguistics.
LUSId Paper: Research paper in the NLP domain into the effects of social demographics on scientific writing style with the Mallet API. The model generated was able to predict writer demographics with 95% accuracy.
EZRoomFix: Android app using QR Codes for easy logging of issues to Dell facilities management. This facilitated faster turn around on work orders and faster reporting of incidents in the office.
Commissioned Website: Commissioned website for concept artist Ray Chen to display their art collection. Implemented with minimalist and responsive design, making their art available on mobile and desktop.
Unix Shell: CLI that mimics most of the core functionality of the UNIX terminal in C, capable of signal handling and instruction routing/piping. Built in a well architectured design to increase response speed from queries.
Jeopardy! Data Visualization: Visualized data collected and synthesized with data science techniques from the hit show *Jeopardy!* since 1982. By using d3.js the data was presented in an accessible and usable format.
GANetflix: Analog for Netflix using SQL, JSPs, and JavaScript used to learn about strong database building. Built around a clearly defined visual language that translated to improved user experience.
Gerrymandering System: Political web tool in Java, with SQL, ReactJS, Spring Boot, and Leaflet. Allowed local political representatives to make more informed decisions about electoral boundaries and their fairness.
Instrument Recognition: Using unsupervised clustering to recognize musical instruments by their timbre as captured in their segmented fourier transforms. Achieved over 90% accuracy for predicting instruments.
DOTA2 Predictor: Coded multiple linear regression and perceptron comparable to an SVM in classification. Used highly efficient python libraries to achieve speeds equal with state of the art implementations.

Experience

Monitaur, MA | Research Data Scientist (Part-Time) — 2019 - Present

- Wrote the base code of a machine learning model auditing system central to our platform
- Added functionality for the tool using state of the art convolutional neural network architectures adding the ability to read in images and run predictions on them.
- Built the system around AWS with testing suites implemented over Docker to allow to quickly develop on a CI/CD platform on a cloud architecture.

Capital One, VA | Senior Data Engineer — 2018 - Present

- Working on applications using cutting-edge open source frameworks enabling the latest security policies and network protocols to minimize customer wait times.
- Pipelining data for streaming data and CICD to allow us to do real time banking/reporting.
- Bringing solid experience in emerging and traditional technologies
- Using Java, Python, Scala, Spark, AWS, Spring, Pandas, and databases

Stony Brook Dept. of Computer Science, NY | Research Assistant — 2016-2018

- Completed under Dr. Ritwik Banerjee at Stony Brook University over the course of two years.
- Research on predator detection that resulted in new scientific understanding about how online predators use manipulation and control to abuse their victims. Insights from this research can be used to create system to flag abusers before they act. The NLP pipeline, backend, and visualizations were made with CoreNLP, Java, d3.js respectively.
- Research on social demographics in scientific writing to demonstrate that even writers in the ACL community are writing with detectably different styles.The NLP pipeline, backend, and visualizations were made with Mallet, Java, and Python notebooks respectively.
- The latter research project was formally written in a paper submitted to CoNLL

Dell, MA | Data Science Intern — Summer 2017

- Analyzed customer's responses to Dell products using Natural Language Processing. This allowed tracking of customer sentiment of Dell's product over time.
- Performed usage and health analytics on Internet of Things platform maintained by Dell to monitor and preemptively arrange repairs for the server systems infrastructure.
- Implemented a Swagger specification that increased learning speed by 60% for a previously undocumented and underused Dell API for controlling provisioned IOT servers.

Advalent Corporation, MA | Android Development Intern — Summer 2015

- Wrote the base code of an Android application to keep employees updated on the database of all internal medical transactions for monitoring.
- Added functionality for the app to verify users and then connect to a SQL backend which pinged subscribers about critical table updates.
- Followed the Material Design guidelines specified for Android Development to create an engaging and unobtrusive user experience.

Skills

Programming Languages:	Object Oriented: Python, Java, C, JavaScript Markup: HTML5, CSS3, jQuery, SQL, LaTeX, JSON, XML, CSV Other: MIPS Assembly, SML, Prolog, MATLAB, R
Frameworks:	Amazon Web Services (AWS), Docker, Jenkins, React, Angular

Certifications

AWS Solutions Architect | Amazon Web Services
Machine Learning | Coursera.com (Andrew Ng)
Artificial Intelligence | Udacity.com (Peter Norvig / Sebastian Thrun)