Siddhartha Mishra

siddharthami@umass.edu • siddhartha1234.github.io

Education

MS in Computer Science
University of Massachusetts, Amherst, MA *Incoming Jan 2021*

B.Tech. in Computer Science and Engineering and Eng. Science
Indian Institute of Technology(IIT), Hyderabad, India, 8.78/10 GPA *2015 - 2019*

Central Board of Secondary Education, India
Higher Secondary, 95.4% *2013 - 2015*

Interests

Natural Language Processing, Statistical Machine Learning, Deep Learning

Experience

Goldman Sachs Private Ltd.
Analyst *May 2019 - Dec 2020*

- Working in Enterprise Machine Learning platform services team on services such as metric analysis, alert prediction and resolution using serverless frameworks like AWS lambda, AWS Stepfunctions.
- Worked on Grafana dashboards for monitoring stack using PromQL and model deployment to Kubernetes.
- Added features to big data log analysis and system health check platform for real-time alerting.

Goldman Sachs Private Ltd.
Summer Analyst *May-July 2018*

- Worked on a plugin for a graphical pipeline design tool of models used for prediction/auto-resolution of alerts using Angular.
- Built a compiler for validating the pipeline and added export as/import from payload features.

Coala, Python Software Foundation
Developer *May-July 2017*

- Contributed to the open source community by helping Coala improve the design of their caching framework and added other structural/performance optimizations in their codebase.

Research/Projects

*Additional details and reports available at my **website***

Multiclass Recurrent Gaussian Process for NLP problems [Code]
Advisor: Dr. Srijith P.K. *February 2019 - August 2019*

- Formulated a multi-class model for recurrent Gaussian process using Variational Inference and ELBO gradient optimisation.
- Applied the model on various NLP problems such as Parts of Speech Tagging, Sequence encoding etc. with lesser data to utilize a Bayesian model.

Query Segmentation using LSTMs [Code]
Advisor: Dr. Maunendra Deskar *October 2018*

- Designed a novel approach to Query Segmentation by mapping it to a sequence tagging problem.
- Used bi-directional LSTMs with/without CRF layers on webis-qsec dataset.

Deep reinforcement Learning model for self driving cars [Code]
Advisor: Dr. Vineeth N. Balasubramanian *March 2018*

- Simulated environment using Mario Kart game.
- Used Asynchronous Actor Critic Advantage (A3C) to increase exploration rate and hence reducing training time in comparison to methods such as DQN.
- Designed custom approximate reward function by a hybrid pipeline to increase granularity of reward to accelerate learning.

Story similarity Detection/Clustering [Code]
Advisor: Dr. Manish Singh *Feb 2018*

- Used NLP language models and LSH for efficient similarity hashing using TF-IDF.
- Used Community Detection for a non-euclidean model on Tweets and Reuters data for news to find similar stories.
- Used a hybrid model combining these techniques to obtain better performance.

Media Server implementing MPEG-DASH and HLS protocols [Code]
Advisor: Dr. Anthony Franklin *Jan 2018*

- Implemented widely used streaming protocols as part of a media server and compared their performance in different scenarios.
- Added a layer of END-to-END encryption to make the media server secure.

Parallel Johnson's Algorithm and Optimized Matrix Chain Multiplication [Code]
Advisor: Dr. Sathya Perri *Nov 2017*

- Deviced a distributed version of Johnson's Algorithm and the underlying data structures used in it.
- Implemented Concurrent Matrix Chain Multiplication to compute independent states of Dynamic Programming in parallel.

Academic Achievments

- Academic Excellence Award for highest SGPA in my major in academic year 2016-2017.
- Graduated 2nd (GPA wise) in my Department Class of 2019.
- Qualified for ACM ICPC Amritapuri regionals and Kharagpur regionals 2017. Secured 49th rank/265 teams in Amritapuri regionals 2017.
- Winner of "Honeywell Machine Learning Hackathon 2019" for the task of automated feature extraction of cockpit images in aircrafts.
- KVPY Fellowship by Indian Institute of Science, Bangalore; secured All India Rank 210.
- 97.66 percentile in JEE Advanced out of 150,000 candidates.
- Qualified INMO (Indian National Mathematics Olympiad) by clearing two stage regionals KVS-JMO and RMO; secured All India Rank 7.

Relevant Coursework

- **Machine Learning**: Applied Machine Learning, Deep Learning, Theoretical Deep Learning, Bayesian Data Analysis, Theory of Learning and Kernel Methods
- **Data Science**: Data Mining, Information Retrieval, Information Theory
- **Computer Science**: Distributed Computing, Parallel and Concurrent Programming, Computer Networks, Operating Systems, Database Management system, Algebra of Computer Science, Computer Network and Security, Compilers, Advanced Data Structures and Algorithms, Discrete Structures
- **Mathematics**: Calculus, Linear Algebra, Differential Equations, Probability, Statistics

Academic Responsibilities

MA2140: Statistics — **Prof. J. Balasubramaniam**
Teaching Assistant — *Spring 2019*

CS3530: Computer Networks-I — **Prof. Anthony Franklin**
Teaching Assistant — *Fall 2018*

MA2110: Probability — **Prof. J. Balasubramaniam**
Teaching Assistant — *Fall 2017*

- Worked closely with professors and fellow TAs to draft and evaluate assignments and exams. Assisted students with doubts and challenges faced throughout the course.

Technical skills

Languages: `Python, C++, C , Go, Java, Android, C#, Haskell, Prolog, MIPS,` LaTeX
ML/Data Science: `Tensorflow, PyTorch, Theano, NumPy, OpenCV, SciPy, SKLearn, Pandas, MATLAB`
Devops/Tools: `Kubernetes, Prometheus, Kafka, Kibana, Elastic Search`
Web: `HTML, CSS, JavaScript, AngularJS, NodeJS, Flask, MongoDB, SQLite`