Atharva Wandile

Boston, MA 02130 | Email: atharvaw31@gmail.com

GitHub| LinkedIn | Website

Technical Skills

- **Languages**: Python, Java, Go, Scala, JavaScript, Typescript
- **Tools and Technologies**: AWS (EMR, Athena, S3), MongoDB, Hadoop, Apache Spark, Redis, Docker, Kubernetes
- **Frameworks**: Tensorflow, PettingZoo, Flask, Django, JSwing, JUnit, Log4j
- **Libraries**: Pandas, Keras, PyTorch, Numpy, Matplotlib, OpenAI gym

Experience

Software Developer, Oak AI, *Boston, MA* **Jan. 2024 – Present**

- Strengthened data access controls and amplified system transparency through an orchestrated audit, documentation, and clarification of Supabase RLS policies and externally visible entry points.
- Engineered an efficient algorithm resulting in the identification of duplicate documents based on text content.

Research Intern, Lab for Learning and Planning in Robotics, *Boston, MA* **May. 2023 – Jan 2024**

- Multi-Agent Reinforcement Learning: Developed continuous macro-action environments for Multi-Agent Reinforcement Learning. Analyzed and tested continuous macro-action PPO algorithm for robotics and autonomous systems.
- Hierarchical Reinforcement Learning: Achieved state-of-the-art performance by training agents using option-critic and actor-critic algorithms for centralized multi-agent reinforcement learning on four rooms and petting zoo MPE environments.

Data Engineer, Redbus, *Bangalore, India* **Jan. 2020 - Jul. 2021**

- Big Data Pipelines: Improved customer retention by 10% through the design and implementation of big data pipelines, ETL and parallel processing models for customer life cycle management with cross-functional communication with marketing teams for analysis; leveraging AWS cloud and database technologies (MongoDB, Cassandra, Postgres).
- Distributed Parallel Processing in Spark: Slashed processing times by 200% for weekly and monthly statistical analyses implementing Spark programs on AWS platform, enhancing efficiency.
- Flask and Golang Web APIs: Developed low latency Flask and Golang API's handling millions of hits per second. Optimized legacy api's to cut down response time by 50%.

Machine Learning Intern, Allgo Embedded Systems , *Bangalore, India* **Jun. 2019 - Aug. 2019**

- Tracking Driver Actions: Led a research based project to identify activities of vehicle drivers given video input from dashboard cam and classify activities into distinct classes in realtime using computer vision and deep learning.
- Deep Learning and Computer Vision: Analysed and compared multiple models including 2 stream CNNs, LSTM and traditional CV techniques. Demonstrated final model with ability to correctly identify a smaller subset of actions with high accuracy up to 86%.

Software Engineer Intern, Edumerge Pvt. Ltd. , *Bangalore, India* **Jun. 2018 - Aug. 2018**

- Performance Prediction: Crafted a sophisticated regression model with 2% error margin tailored to anticipate the academic performance of a specific class. This model utilized historical performance data a from constrained dataset of 1000 students from the preceding batch and was particularly designed for proactive intervention. This precision contributed to data-informed decision-making and enhanced academic outcomes.

Education

Northeastern University, *Boston, MA* **May 2023**

- Master of Science in Computer Science *GPA: 3.8/4.00*

Related courses: Algorithms, Parallel Data Processing, Reinforcement Learning, Data Mining

JSS Science and Technology University (SJCE), *Mysore, India* **Jan 2020**

- Bachelor of Engineering in Information Science *GPA: 9.14/10.00*

Related courses: Data Structures, Object-oriented programming, Machine Learning, Cloud Computing

Projects

- **Distributed Real Time Collaborative Editor**: Developed responsive app for a distributed text editor (like google docs) that can support multiple users updating document in real-time and is fault tolerant by leveraging an architecture with data replication management and distributed transactions.
- **Distributed Hierarchical Agglomerative Clustering for Music Recommendation**: Evaluated distributed k-means clustering with average linkage clustering on Apache Spark using AWS EMR to group similar songs together from million song dataset for music recommendations.