Aditya Jain

jainadit@usc.edu | **+1 (323) 986-9260** | **Linkedin** | **Github** | **adityajain.me**

Education

University of Southern California, Los Angeles, CA — Jan 2021 - Present
Master of Science (Computer Science) — **GPA: 4.0/4.0 (Honors)**
Coursework : Analysis of Algorithms, Foundations of Artificial Intelligence, Data Mining, Database Systems, ML for Games

University of Pune, India — Aug 2014 - Jun 2018
Bachelor of Engineering (Computer Engineering) — **Aggregate: 74% (3.6/4.0)**
Coursework : Data Structures, Operating Systems, Data Warehousing, Cloud Computing, Software Engineering

Technical Skills

- **Languages:** Python, Java, C++, JavaScript, SQL, CSS, HTML, Object Oriented Programming
- **Databases:** MySQL, MongoDB, Redis, SQL Server
- **Machine Learning:** Scikit-learn, Pandas, Numpy, Keras, Tensorflow, Pytorch, SciPy, NLTK, matplotlib, seaborn, Statistics
- **Tools and Frameworks:** Flask, pySpark, Git, Docker, Kubernetes, gRPC, graphQL

Experience

Salesforce, Inc — May 2022 - Aug 2022
Software Engineering Intern | Python, Pandas, PySpark, Docker, Orchestration, Argus
- Built an automated monitoring system to track Einstein Search metrics over time and detect anomalies.
- Used pySpark to support parallelism over multiple metric definitions and org-ids to increase scalability by 150 times.
- Developed code for deploying the designed framework on hawking orchestration to support daily runs on production.

USC Institute of Creative Technologies — Feb 2021 - May 2022
Student Worker (RA - Machine Learning) | Python, pytest, JavaScript, mongoDB
- Researched techniques to improve performance of automatic answer grading system in Opentutor (a dialog based learning platform) and resolve the cold-start problem.
- Implemented a clustering algorithm to group and identify different type of correct answers for an expectation. Used cluster center of each group as features to improve grading system accuracy to 85%.
- Developed ML pipeline for grading system and contributed in developing python based web API interface for this system.
- Performed unit testing via pytest to compare the performance of models after different tweaks or improvements.

Cognizant Technology Solutions — Sep 2018 - Jan 2021
Associate Projects (Data Science)
- **Search-ad click prediction | Python, keras, pandas, tf-serving, gRPC, Docker, Redis**
  - Researched on various NLP techniques to improve ad click-through rate of an existing system by 10%.
  - Implemented NLP pipeline for a ML based advertisement search engine. Worked on text preprocessing functions and integrated them in pipeline to prepare input for a machine learning model during training as well as prediction.
  - Deployed CLSM model using tf-serving on Docker over gRPC interface for improving query response time to 5 ms.
  - Contributed in converting text preprocessing steps from Python to C++ to improve overall query response time.
- **Medicare Star Analytics (Healthcare) | Python, SQL, sklearn, pandas, numpy, matplotlib, seaborn**
  - Developed analytical and ML models to assist customers' marketing team in identifying the right and influential population for various outreach programs to improve the overall STAR rating of the Medicare plans.
  - Built ML models to get probability of patient's health decline (bladder issues, mental/physical health decline, etc) leveraging his/her medical records. Achieved .75 AUC score using gradient boosting on decision trees.
  - Worked on data visualization, report creation of analytical models to support our hypothesis and for client presentation.

Projects - *https://projects.adityajain.me*

AI Checkers Agent | Python | Link
- Implemented checkers AI Agent in Python using mini-max algorithm with α-β pruning to compete against other agents.
- Implemented agent ranked 6th among 250 students in checkers AI agent competition for "Foundation of AI" course at USC.

Kinship Detection using Faces | Python, Keras, numpy, OpenCV, Computer Vision | Link
- Used VGG-Facenet to generate face embedding and trained a siamese network architecture to get kinship probability.
- Performed image augmentation to increase dataset size and achieved 0.79 AUC using Kaggle Dataset.

NER/POS Tagging Web Application | Python, Keras, numpy, Flask, Docker, HTML/CSS, JavaScript | Link
- Trained an LSTM based seq2seq model using Keras to tag all words of paragraph to their Named Entity or Part of Speech.
- Developed an web interface using Flask API to interact with the model and finally deployed app as a docker container.

Biometric Authentication Android App | Java, Node.js, Express.js, mongoDB, XML, HTML, CSS | Link
- Developed an android app to perform fingerprint based authentication of candidates to prevent impersonation in exams.
- Built an API using Node.js to modify/fetch data on mongoDB server using API requests sent from Retrofit in android.

Leadership and Achievements

- Runner-Up in Smart India Hackathon '17 among 300 teams nationwide. **Apr 2017**
- Among top 8 finalist from 100 participants in Infosys Techzooka '16. **Oct 2016**
- Organizer and Lecturer in a national level event called "Linuxication" at MIT, Pune. **Mar 2016/2017/2018**