Sahasra Iyer

413-512-3502 iyersahasra@gmail.com linkedin.com/in/sahasraiyer

Education

University of Massachusetts Amherst **Sep 2021 – May 2023 (Expected)**
College of Information and Computer Sciences, MS in Computer Science *GPA: 3.89/4.0*

Cummins College of Engineering for Women, Pune **Aug 2014 – May 2018**
Department of Information Technology, Bachelor's of Engineering in Information Technology *GPA: 3.82/4.0*

Coursework :

- Advanced Natural Language Processing, Neural Networks, Machine Learning, Information Retrieval, Methods of Applied Statistics, Advanced Databases, Algorithms for Data Science, Intelligent Visual Computing

Professional Experience and Internship

NLP Engineer Intern - Data Science For Common Good, UMass Amherst, MA **May 2022 – Aug 2022**

- Developed an application filter model for MassChallenge, a non-profit startup accelerator, to predict the startup's success based on their write-up attributes. A proprietary dataset with over 100 attributes was used to train the model.
- Developed a content-based recommendation model that provides a mentor-mentee matching for startups registered in a MassChallenge program using Python, ML libraries, and Apache PySpark.

Technology Analyst, Citicorp Services, Pune **Aug 2018 – Jul 2021**

- Utilized the Agile methodology to migrate a Commodities feed, initially running on PL/SQL, to a more efficient Spark/Scala framework. The upgraded system facilitated the generation of risk numbers for trades executed on the day.
- Improved reliability in Volcker Exchange-Traded transactions, reducing trade price mismatches from 70,000 to 100 per day, thus attaining a 97% error rate reduction. (earned Citi Gratitude Award for the same)
- Supported trade risk evaluation on a production-level Spark Cluster for the Simpliciti Risk Engine.
- Utilized Spark's in-memory computing capabilities to enable the Dynamic Distribution Engine (DDE) product processor to process payments more efficiently, resulting in notable improvements in efficiency and execution times.

Technical Skills

Languages: Python, Scala, XML, R, C, C++, Java
Technologies/Frameworks: Spark, PySpark, Scikit-Learn, PyTorch, TensorFlow, Pandas, Numpy, Matplotlib, Django, XGBoost, NLTK, spaCy, Tableau
DBMS packages: MySQL, Oracle DB, PostgreSQL, MongoDB, Cassandra, Domo
DevOps and Cloud: AWS, Docker, Git

Projects

Retrosynthesis - Molecule Route Prediction, Graduate Researcher - IBM **Jan 2022 - May 2022**

- Developed and optimized transformer-based multi-step models for molecule synthesis route generation using artificial intelligence, resulting in a 18% efficiency improvement compared to traditional methods.
- Implemented a self-supervised transformer-based atom mapper to analyze over 50,000 synthesis routes, resulting in the identification of key areas for improvement and an increase in overall valid route prediction accuracy by 30%.

Analysis of CGM time series data **Jan 2022 - May 2022**

- Worked on CGM time-series data analysis to extract features via various methods like statistical analysis, Fourier & power transforms. Implemented and fine-tuned multiple classification (KNN, SVM, SGD, Logistic Regression) & supervised-clustering algorithms (K-Means, DBScan) to achieve 70% accuracy improving over the baseline of 60%.

Political bias mitigation using attribute transfer **Sep 2021 - Dec 2021**

- Explored and implemented systems for bias mitigation in political news articles using text attribute transfer based on paraphrasing models implemented through a pretrained GPT2 medium model.
- Experimented with Reverse Attention Networks and XLNet Classifiers for attribute and content disentanglement.
- Achieved an accuracy of 58.39% on the test dataset, with fluency rate of 97.7% and similiarity score of 90.07%.

Text Detection from an Image and Translation **Apr 2017 - May 2018**

- Developed a machine translation tool that recognizes Hindi text from an image and produces the corresponding English translation using the Viterbi algorithm, an extension of the Hidden Markov Model, which incorporated syntaxes and semantics of the English language to produce grammatically correct translations using NLTK library.
- Authored and published papers in international journals Springer and IEEE for the same.