Siddhesh Jagtap

Tempe, AZ | (602) 570 6978 | LinkedIn | GitHub | sjagtap5@asu.edu | **Open to Relocation**

PROFESSIONAL EXPERIENCE:

Cognition and Intelligence Lab at Arizona State University — **Tempe, Arizona, United States**
Graduate Research Assistant — *August 2021 – Present*

- Designed and built a Language Model based on **RoBERTa** architecture, using a distributed GPU computing environment, for variable type (from source code) prediction task with an accuracy of 82.4%.
- Developed a scalable data processing solution in **Python** to process 120k textual files and run on a distributed system, as a requirement for various NLP related tasks.
- Collaborated with the Biodesign Center for Biocomputing, Security and Society, to implement a **unified Language Model** catering to over 7+ NLP tasks.

Decision Theatre Network at Arizona State University — **Tempe, Arizona, United States**
Data Analyst — *March 2021 – August 2021*

- Developed a **statistical model** to automatically cluster similar emails received by ASU's University Technology Office, which aided in reducing human efforts by 14%.
- Created a **regression model**, to efficiently predict overall sentiment and aspect-based sentiment for ASU's Bio Design Lab collaborative study, processing 1.3M tweets.
- Improved the processing speed of an in-built data processor system from 10% to 21% by using multiprocessing computing through **Python**.

Jio Platforms Limited — **Navi Mumbai, Maharashtra, India**
Software Engineer II (Machine Learning) — *August 2017 – December 2020*

- Spearheaded and developed a scalable, distributed, efficient free-text **search algorithm in Python**, for cross-disciplinary business teams, with salient features like real time search, faster query results irrespective of the volume of data and spell checker, thus improving manual search time by 20%.
- Built a scalable **statistical prediction model** to accurately estimate the demand quantity of construction materials for the current month based on the material's previous demand pattern and helped in reducing purchase costs by 60%.
- Collaborated with the Procurement team and built an innovative solution for auto-processing of 50k legal documents, in terms of data storage in a **distributed MongoDB system**, which provided a significant manual effort reduction by 20%.
- Designed and developed an **automated file processing system** for the Legal team, to extract fields of interest and populate extracted data into a tabular structure, which catered to 30+ vendor systems.

TECHNICAL SKILLS:

Programming Languages: Python | Java | C | C++ | Shell | Bash
Databases: RDBMS | MySQL | PostgreSQL | NoSQL | MongoDB | Cassandra | MS-SQL
Software/Cloud Tools: Excel | PowerPoint | Docker | Jenkins | Git | Kafka | Flask | Kubernetes | Microsoft Azure
Applications/Technologies: scikit-learn | pandas | NumPy | Pytorch | Tensorflow | Keras | OpenCV | NLTK
Methodologies: Agile | Scrum | Data Modeling

EDUCATION:

Arizona State University — **Tempe, Arizona**
MS in Computer Science (GPA: 3.94/4) — *Graduation Date: December 2022*

University of Mumbai — **Mumbai, India**
BEng in Electronics and Telecommunication (GPA: 8.03/10) — *Graduation Date: May 2017*

Relevant Coursework: Structured Programming Approach, Object Oriented Programming

ACADEMIC PROJECTS:

Model-based Semantic Evaluation (Vision Language Project)

- Goal: Utilize **NLI models (RoBERTa, BART and T5)** to evaluate the performance of a system on multiple tasks, which learns to compare the semantic meaning of two sentences.
- Designed an evaluation system in Python to compute BARTScore and NLIScore using NLI models and a distributed computing environment.

Tweet Stance Classification and Evaluation using Transfer Learning

- Goal: To predict the stance automatically while assessing whether the author is in favor of, against, or neither inference is likely based on language.
- Developed a system using **Transfer Learning** to predict the stance of a given tweet (text) and performed multiple evaluation tasks on the same.