Harsh Atha

Bloomington, IN | (812) 325-4082| atha.harsh3@gmail.com | LinkedIn Profile: Harsh Atha

Education

Indiana University, Bloomington — CGPA: 3.9/4
Master of Science in Data Science — Graduation: May 2023
Course Topics: *Machine Learning, Computer Vision, Graph Analytics.*

K.J. Somaiya College of Engineering, Mumbai — CGPA: 7.45/10
Bachelor's of Engineering in Electronics and Telecommunications — Graduation: May 2017
Course Topics: *Image Processing, Data Structures, Applied Mathematics, Operating System.*

Technical Skills

- Programming Languages: Python, PySpark, SAP ABAP, R, and Embedded C.
- Database Technologies: SAP HANA, SQL, Google Cloud Platform, Hive, and Hadoop.
- Visualization Tools: Tableau, SAP BI, Bex, plotly, ggplot, matplotlib, seaborn, and PowerBI.
- Machine Learning Models/Libraries: Linear Regression, Perceptron, Logistic Regression, kNN, SVM, K-Means, Gaussian Mixture Models, PyTorch, NLTK, and TensorFlow.

Professional Experience

Research Assistant, Indiana University Bloomington, US. Dec 2021- Present.

- Creating a pipeline to visualize the Tableau dashboard - "US Utility Disconnections", as a part of the **Energy Justice Lab** at IU, by scraping utility data from the web using Selenium and modeling data in Python.
- Implemented automation to cut down manual workhours by 40% as of now.

Associate Instructor, Indiana University Bloomington, US. Aug 2021- Dec 2021.

- Teaching assistant for the course **INFO-I 535 MANAGEMENT, ACCESS, AND USE OF BIG AND COMPLEX DATA.**
- Graded assignments, led doubt solving sessions and aided the professor in setting up the course material.

Data Analyst, Jio Platforms Ltd., Reliance Corporate Park, Navi Mumbai, India. July 2017- July 2021.

- Implemented and enhanced over 117 automated data models and live dashboards by collaborating with the EXIM and Accounting team, which cut down week long manual tasks to a few hours.
- Aided in migration of SAP data to Hadoop and lead the team in shifting data modelling scripts from SAP HANA to PySpark. Used Apache Kafka, Sqoop, AirFlow, HIVE and Spark to achieve the ETL and automation process.

Academic Projects

Handwritten Digit Recognition Dec 2021 – Jan 2022.

- Trained a CNN model on MNIST dataset containing handwritten digits from 0 to 9, created a GUI using TKINTER to predict the hand-drawn digits.

CIFAR-10 Object classification Dec 2021 – Jan 2022.

- Implemented object classification from images by developing a CNN model from scratch for low resolution CIFAR-10 dataset. Improved the original model by using regularization and data augmentation.

POS Tagging and Text Classification Nov 2021-Dec 2021.

- Developed NLP based projects to perform POS Tagging on a book dataset using Hidden Markov Models and Bayes Networks.
- Sentiment Analysis for text datasets and multi class classification using Bayes Theory.

Voice for the Voiceless Dec 2016 - April 2017.

- Created a wearable device that facilitated communication with speech impaired people
- Designed the circuit and programmed the Arduino to convert American sign language to audio using ink-based flex sensors, speakers, and speech synthesizers.

Certifications

- **Python Programming Specialization** (4 courses + 1 Capstone) by **University of Michigan on Coursera**. (Mar '21)
- **Crash Course on Python and Interaction with OS using Python** by **Google on Coursera.** (Aug '20)