Aayush Sharma

Phone: 469-785-1493
Email: Aayush.Sharma@utdallas.edu
LinkedIn: www.linkedin.com/in/aayushsharma97
GitHub: https://github.com/sharma-aa

SUMMARY:

Backend software engineer with professional experience of Java, Python, Applied Machine Learning and Big Data analytics to real world business problems. Built node build timeline prediction for Verizon and Intelligent smart scheduling system for NURSD.

EDUCATION

The University of Texas at Dallas May 2021
M.S., Business Analytics (Courses: Machine Learning, Big Data, Econometrics, Statistics, Data Visualization) *GPA 3.9/4.0*
Awards: Dean's Excellence Scholarship

Delhi Technological University, India May 2019
B. Tech, Software Engineering (Courses: Object oriented programming, Data structures and algorithms) GPA 4.0/4.0
Awards: First division with distinction

TECHNICAL SKILLS

- Programming Languages: Java, Python, C++, R, Git, C, JavaScript.
- Database Management: MySQL, PostgreSQL, MongoDB, Hadoop (Spark, MapReduce, Flume, Sqoop, Hive, Pig).
- Web Technologies: HTML, CSS, Bootstrap, Flask, Spring Boot, React.js.
- Machine Learning: Regression, SVM, Random Forest, RNN, Bagging, Boosting, PyTorch, TensorFlow, Scikit-Learn, Transformers.
- Tools and Framework: Pandas, Numpy , NLTK, PuLP, Selenium, Docker, Tableau, Matplotlib, Seaborn, Excel.

WORK EXPERIENCE

VERIZON, Texas, USA January 2021-May 2021
Software Development Engineer (Data Science Intern)

- Build a **predictive model** for VoIP node timelines using a **Neural Network architecture** with **word embedding** designed from scratch as input to increase build efficiency by 30%. Automate weekly report generation using **Python** for missing data.
- Write anomaly detection algorithm in Python to analyze misconfigured router data and predict failovers. Build **parser** to migrate data from excel workbooks to **JSON** files. Predict router misconfigurations to cease more than 1200 failovers.
- Build a **system design** for a web application using **Spring Boot** and **React.js** to analyze efficiency of engineers over **Jira** tickets by **clustering** semantically similar tickets. Build a recommender system for engineers using **skicit-learn** for tickets using scores.

NURSD HEALTHCARE, Illinois, USA June 2020- November 2020
Software Development Engineer (Machine Learning Engineer)

- Design automated workforce management system to schedule workers for healthcare facilities using multiple **Linear Programs**. System generates weekly schedule for 100 nurses for a large scale facility in 2.88 seconds.
- Optimize solution using objective function parameters like geolocation distance, worker schedule and work-day preferences, overtime, and requested days off. (**PATENT PENDING- Primary Contributor**)
- Build a recommender system using rating generated from service feedback based on **Google BERT** sentiment analysis algorithm. Create ER diagram for facilities and workers on **MySQL**. Populate dummy data by scraping webpages using **selenium**.

DELOITTE USI, Mumbai, India
Software Development (Technology Consultant Intern) June 2018 – July 2018

- Automate back-end data management using **Apex** triggers as a controller class, Structure object query language queries, and validation rules to increase productivity by approximately 30% in a team project.
- Prepare a refined travel management application U.I. by creating **VisualForce** pages using Salesforce Lightning Design System, **HTML**, and **CSS**. Improve the data representation using the dashboard to create visual reports ensuring efficient inferences.

ACADEMIC PROJECTS

Expense Tracker System April 2021

- Build back-end on **Mongo DB** using Java **Spring Boot** to track income, expenses, and balance with options to filter data.
- Font-end designed using **React** API. Create pie charts and histograms for data visualization.

Twitter Sentiment Analysis using Big Data March 2021

- **PySpark** used to create **ETL pipeline** and Data processing for tweets related to Demonetization in India from twitters' 2016 DB.
- Analysis performed on **Pig Grunt shell** to assign average rating to a tweet using sentiment values from multiple dictionaries.
- **Beeline shell** used for running **Hive** Queries. PySpark used for Data Visualization to find overall impact of Demonetization.

Real or Fake Disaster Tweet Detection, Kaggle December 2019 – January 2020

- Implement a hard-voting classifier with **SVM**, **LSTM**, **Random Forest**, **MLP**, and **boosted Ridge Regression** classifier to obtain a public score of 0.81 on the validation data set.
- Create word embeddings using **GloVe** for list of disaster related words scraped from webpages using **BeautifulSoup**.

ADDITIONAL INFORMATION

- Mathematics and Physics teacher for "Teach For India" initiative.
- Team captain 2nd position in Inter School Table Tennis Competition (Zonal) in 2012.