SUCHITA SHARMA

sharma.such@northeastern.edu | 857-335-7669 |linkedin.com/in/suchitasharma99

EDUCATION

Northeastern University, Boston, MA — Sept 2023 - Aug 2026
Master of Science in Data Science - Khoury College of Computer Sciences
Related Courses: Natural Language Processing, Machine Learning, Human Computer Interaction

Savitribai Phule Pune University, Pune, India — Aug 2017 – May 2021
Bachelor of Engineering in Information Technology
Related Courses: Data Structures, Cloud Computing, Web Development, Distributed Systems, Data Science

TECHNICAL SKILLS

Languages and Databases: Python, R, SQL, Shell Script, Java, HTML, CSS, Figma
Technologies and Tools: Pandas, ScikitLearn, Numpy, Matplotlib, Seaborn, PyTorch, Docker, Git, Sklearn, Tableau
Cloud Computing: Amazon Web Services (AWS), Google Cloud (GCP)

WORK EXPERIENCE

Flentas Technologies Pvt Ltd, Pune, India — Jan. 2021 – Aug. 2023
Cloud Engineer

- Optimized existing cloud configurations leading to improved application deployment speeds by an impressive rate of 40%, enhancing overall project timelines across diverse client engagements while maintaining budget compliance at all levels.
- Implemented streamlined **CI/CD pipeline** that enhanced code integration times by reducing build errors to fewer than 5 incidents monthly, leading to more reliable deployments.
- Designed and optimized **SQL databases / Amazon RDS** to improve query performance, reducing response times and ensuring seamless data retrieval for business-critical applications.
- Coordinated with cross-functional teams to develop and deploy cloud solutions that directly supported three major business objectives, resulting in enhanced project efficiency and smooth workflows across departments.

ACADEMIC PROJECTS

Knowledge Distillation — Sept 2024 - Dec 2024

- Teacher Model: Framework from a research paper for the **multilingual case model based on BERT**. Trained in 104 languages, using cased tokens to maintain multilingual robustness.
- Fine-Tuning: Applied **transfer learning techniques** to fine-tune the BERT-based multilingual model on domain-specific datasets, improving accuracy and contextual understanding for targeted translation tasks.
- Student Model: Implemented a **sequence-to-sequence (seq2seq) RNN** for Spanish-to-English translation. Leveraged RNNs to effectively handle and retain longer sequences, ensuring accurate translations.

Hallmarks of Alzheimer's Disease — May 2024 - Aug 2024

- Conducted comprehensive analysis to pinpoint critical proteins exhibiting notable variations among Alzheimer's patients, categorized by age groups: under 75 years, between 75-85 years, and over 85 years.
- Applied **feature engineering techniques** to create new predictive variables, improving model accuracy and revealing deeper insights into demographic mortality trends.
- Modeling Techniques: **Support Vector Machine** was identified as the best performing model, offering high accuracy and low error rates.
- Generated precise graphs illustrating key trends related to Alzheimer's symptoms by age group and different identified proteins.

Road Traffic Mortality Analysis — Jan 2023 - Apr 2023

- Uncovered three primary drivers behind reducing mortality metrics: access improvements, preventive care initiatives, and lifestyle changes; this initiative simplified focus areas for future research.
- **Feature Engineering**: Created and selected meaningful features, such as age-adjusted mortality rates and socioeconomic factors, to enhance model performance and reveal deeper insights into demographic trends.
- Data Preparation and Modeling: Cleaned dataset for accuracy and used **regression analysis** to predict death rates based on sex and year.
- Developed detailed visual representations of data using ggplot2 and tidyverse, producing over 15 distinct graphs and charts that clearly illustrated complex findings related to demographic patterns in death rates.