AMAAN VORA

New Brunswick, NJ | +1 732-640-(3024) | amaanvoraa@gmail.com
linkedin.com/in/amaan-vora | github.com/deadven7 | amaanvora.com

EDUCATION [portfolio.amaanvora.com]

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
New Brunswick, NJ

Masters of Science, Statistics (Data Science) | Cumulative GPA: 3.5/4.0
Expected May 2024

Relevant Coursework: Data Mining (Machine Learning), Statistical Modeling and Computing, Data Wrangling, Regression and Time Series Analysis, Statistical Software (NLP), Statistical Learning (Neural Networks)

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
Chennai, TN, India

Bachelor of Technology, Computer Science and Engineering | Cumulative GPA: 9.2 / 10.0
July 2018 - May 2022

Relevant Coursework: Data Structures and Algorithms, Calculus, Programming Languages, Neuro-Fuzzy and Genetic Programming, Database Management Systems, Artificial Intelligence, Computational Logic, Advanced Programming Practices

SKILLS [certifications.amaanvora.com]

Technical Skills: NumPy, Pandas, Matplotlib, Plotly, Seaborn, Scikit-learn, Linear Regression, Logistic Regression, Kernel SVM Classification, Decision Trees, Random Forests, K-Means Clustering, Neural Networks, Transformers, Natural Language Processing, TensorFlow, PyTorch, Keras, ggplot2, dplyr, tidyverse

Programming Languages: Python, R, SQL, C/C++, Java, HTML/CSS

Tools & Frameworks: AWS, Docker, Kubernetes, DynamoDB, MySQL, PostgreSQL, SQLite, Tableau, Hadoop, PowerBI, Spark, BigQuery, Git

Certifications & Training: Google Data Analytics Professional Certificate (Google Data Analytics), Introduction to Deep Learning (Nvidia DLI), Machine Learning Fundamentals (University of Washington)

EXPERIENCE [experience.amaanvora.com]

THE DAILY TARGUM
New Brunswick, NJ

Lead Front-End Developer
August 2023 – *present*

- Led the front-end development of a new cloud computing – backed website utilizing AWS, reduced cost of maintenance by 50%, enhanced user-friendliness, and increased website footfall by approximately 50%.
- Reduced monthly costs by 25% through streamlined AWS and DynamoDB deployment and optimized the efficient deployment of the website's frontend while decluttering the existing code and improved its readability

RUTGERS UNIVERSITY (Statistics Department Research)
New Brunswick, NJ

Graduate Research Assistant
January 2023 – July 2023

- Improved data efficiency and algorithm output by 15%, with a high accuracy of 0.93 for the Random Forest Algorithm, and a subsequent conversion of the algorithm to demonstrable Stata functions
- Enhanced Two-Stage Curvature Identification using a Hat Matrix as part of better tuning Random Forest Algorithms and aimed for a better understanding of Causal Inference Techniques as part of ongoing research

SIEMENS LIMITED (Global Division)
Pune, MH, India

Global Business Services Intern (Machine Learning Team)
June 2021 – August 2021

- Developed a highly efficient bash RC Automation System for driver files that supported over 30 file extensions and installed a vast library of 5000+ files, ensuring seamless integration with the system
- Facilitated the development of a SentenceBERT model for consumer comment data (100,000+ values) and generated comprehensive clusters based on sentiment, enabling insights for strategic decision-making
- Collaborated as part of a global team on improving the accuracy to 0.89 by embedding vectored sentences, encoding paragraphs, and using triple-network BERT structures for accessory hyper-parameter tuning

RESEARCH [International Research Conference on IoT, Cloud and Data Science]

COMPUTATIONAL LINGUISTICS APPROACH TO CLUSTERING SCIENTIFIC RESEARCH
May 2022

- Clustered scientific corpora via legacy and novel algorithms like BERT into broad subjects and niche sub-fields
- Resulted in an accuracy of 0.89 along with volume increments via feature engineering and hyper-parameter tuning

PROJECTS [amaanvora.com/projects]

ABSTRACT TEXT SUMMARIZATION USING GENERATIVE ADVERSARIAL NETWORKS
July 2023

- Extracted over 100,000 topic-based news coverage from reputed sources to perform feature extraction using TF-IDF documentation, lemmatization, and tokenization to prepare data for summarization
- Implemented hierarchical attention to increasing fluency of summaries, attaining a ROUGE-L score of 0.89 while proving the efficacy of such a system for query resolution and machine translation

GENERALIZED RECOMMENDER SYSTEM FOR ANY FORMATTED DATA
May 2023

- Conducted pre-processing, tokenization, and feature engineering on over 10,000 data points and converted data to a streamlined scalar value using TF-IDF documentation and One-Hot Encoding
- Incorporated legacy K-Means and novel Transformer algorithms to achieve an evaluation score of 0.81 for data of varied numerical and string values, ultimately deriving recommendations from streaming and college data