

EDUCATION

- Syracuse University**, Syracuse, NY May 2025
Master of Science | Applied Data Science
Relevant Coursework: Database Management and Administration, Introduction to Data Science, Quantitative Data Science, Data Warehousing, Applied Machine Learning, Business Analytics, Lean Six Sigma, Financial Analytics, Big Data Analytics
- Thadomal Shahani Engineering College, University of Mumbai**, Mumbai, India May 2023
Bachelor of Engineering | Computer Engineering
Relevant Coursework: Database Management System, Machine Learning, Data Structures, Natural Language Processing

EXPERIENCE

- Graduate Research Assistant**, Syracuse University, Syracuse August 2024 – Present
- Exploring clustering of 3,000 combinations over 200 ns of 7-sequence peptoid molecules using VMD and Plotly for visualization and exploratory data analysis
 - Implementing deep learning models including GCNN, CNN, and LSTM to analyze 50 bootstrapped molecular datasets and identify complex patterns
- Graduate Technology Analyst Intern**, Upstate Medical University, Syracuse August 2024 – Present
- Leveraging PageRank on EPIC data to rank 5 clinics by no-show rates, patient volume, and revenue, optimizing clinic performance and reducing inefficiencies across 21 medical personnel
 - Analyzing patient trends using forecasting models like SARIMA and ETS, securing 90% accuracy, and creating dashboards to visualize insights, enhancing decision-making and amplifying patient satisfaction
 - Streamlining workflows using Git for a CI/CD pipeline and developing project management artifacts such as stakeholder analyses and process flow diagrams
- Artificial Intelligence Research Team Lead**, NEXIS Student Technology Lab, Syracuse February 2024 – Present
- Leading a team of 3 people to conduct sentiment analysis on 50,000 tweets adopting NLP techniques like stemming, lemmatization, and BERT transformers, achieving 91% accuracy in predictions
 - Furnishing data wrangling, feature engineering, and visualizations for 923 companies while engineering machine learning models for startup acquisition likelihood, selecting Gradient Boosting as the top model, resulting in an F1 score of 0.987
- Graduate Data Science Intern**, J. Galt Finance Suite, Indianapolis May 2024 – August 2024
- Incorporated an Auto dialer through Node.js and a REST API connection with Twilio, applying XGBoost for statistical programming of call audio and duration estimation, increasing conversion rates by 15%
 - Formulated an ETL pipeline with Python to enrich Zoho CRM records in Big Query leveraging data from Octoparse, adding 1000 data points, and upgrading data quality
 - Developed a performance dashboard for 40 sales professionals harnessing Python Dash, facilitating strategic collaboration with stakeholders through actionable financial insights
 - Automated lead generation using the RAG technique, Botpress, Google Cloud Function, and JavaScript, boosting qualified leads

PROJECTS

- Multi-Domain Enterprise Data Warehouse Using Snowflake and DBT** January 2024 - May 2024
- Defined data pipelines with SQL, Snowflake, and dbt to integrate data from 4 business units, deployed Power BI across domains to monitor KPIs, and improved problem-solving and customer targeting by 20%
- Counteracting Misinformation by Gen-AI using Neural Networks** January 2024 - May 2024
- Built a deepfake detection model with CNNs, AlexNet, and OpenCV for face detection, preprocessing 10,000 real and 6,900 fake images, gaining 98% training and 84.6% test accuracy for real-time content moderation

TECHNICAL SKILLS

Programming Languages: Python, R, Java, C, C++, HTML, CSS, JavaScript
Database: MS SQL Server, BigQuery, Snowflake
Tools: GCP, R Studio, Alteryx, Docker, Figma, Git, Tableau, PowerBI, MS Excel, Trello, Twilio, dbt, Lucid, Postman, Azure Data Studio
Frameworks: Django, Flask, Node.js
Libraries: TensorFlow, PyTorch, AutoML, AutoTS, Keras, Pandas, NumPy, Matplotlib, Plotly, NLTK, scikit-learn