Naman Vipul Shah

+1 (930) 333-2658 ◇ shahnam@iu.edu ◇ linkedin.com/in/naman-shah0506/ ◇ github.com/Naman009 ◇ Atlanta, GA

EDUCATION

Master of Science in Data Science, Indiana University Bloomington   Aug 2023 – Expected May 2025
Coursework: Applied Machine Learning, Applied Algorithms, Elements of AI, Data Mining, Statistics, Advanced Database Concepts, Data Visualization.

Bachelor of Technology in Information Technology, Mumbai University   Aug 2019 – May 2023
Coursework: Deep Learning, Exploratory Data Analysis (EDA), Database Management Systems, Discrete Mathematics, Computer Vision.

TECHNICAL SKILLS

Languages/DB: Python, R, PHP, JavaScript, C, C++, HTML, CSS, Bootstrap, SQL, MySQL, PostgreSQL, MongoDB, Neo4j
ML/DL algorithms: CNN, RNN, LSTM, Random Forest, SVM, K-Means, XGBoost, SGD, Linear Regression, Logistic Regression, Naïve Bayes
Libs/Frameworks: Keras, Tensorflow, Pytorch, Scikit-learn, Pandas, NumPy, OpenCV, NLTK, Matplotlib, Plotly, Seaborn, Flask
Tools: Tableau, Microsoft Power BI, Microsoft Excel, Alteryx, Git, Github, Microsoft Office, Word and Powerpoint
Statistics: Hypothesis testing, Decision-making, P-Value Concept, Regression Analysis, Co-relation, Time Series Analysis

WORK EXPERIENCE

Graduate Teaching Assistant | Indiana University Bloomington | Bloomington, IN   *Aug 2023 – Present*
- Mentored 15 students in capstone project management, addressing **PHP, MySQL, UI/UX, and Git** challenges for a class of 200.
- Led interactive problem-solving sessions for **Applied Algorithms** course, empowering 150+ students to achieve substantial academic gains and mastery of data structures and complex concepts.

Backend Development Intern | Appectual IT Solutions | Mumbai, India   *Jan 2023 - May 2023*
- Reconstructed the backend architecture for a website using **PHP and MySQL**, resulting in a robust and scalable infrastructure.
- Adopted industry best practices in coding, database design, and security, significantly improving site reliability and maintainability.
- Developed **optimized SQL queries** to reduce database read-write times, achieving improved data retrieval and storage efficiency.
- Designed and developed a dynamic dashboard for **data visualization**, meeting client requirements for KPIs.

Machine Learning Intern | KJ Somaiya College of Engineering | Mumbai, India   *Mar 2021 - Dec 2021*
- Worked on **multiple Machine Learning and Deep Learning projects** involving the tech stack **Python, Tensorflow, Keras, Scikit-learn and NLTK**.
- Built a custom web scraper with Python and **BeautifulSoup** to curate a 60000-row database for "Website Type Classification" project.
- Performed **data analytics** like **data cleaning, feature engineering and text preprocessing** using **Natural Language Processing (NLP) Techniques** like lemmetization, stopword removal, text normalization and tokenization.
- Developed high-accuracy data classification models with **state-of-the-art ML algorithms**, achieving **97.89% accuracy** using **LinearSVC**.
- **Published a paper** on this project in the **IEEE Xplore Digital Library (DOI: 10.1109/ICAST55766.2022.10039588)**.
- Managed a **deep learning** and **image processing**-based project focused on failure detection using thermal images of solar panels.
- Analyzed solar panel images using hotspot detection, color space optimization, and data augmentation **reducing false positives by 20%.**
- Tested various deep learning model architectures and performed rigorous hyperparameter tuning resulting in **training and testing accuracy of 91.63% and 84.01%** respectively on Blue Channel images on ResNet50v2 model.

Software Development Intern | KJ Somaiya College of Engineering | Mumbai, India   *Apr 2021 - Jun 2021*
- Automated the process of online lecture scheduling by using **Optical Character Recognition** to convert PDF timetables to **pandas**' dataframe.
- Utilized Google Calendar's API to schedule recurring meetings based on extracted lecture details achieving the streamlined scheduling of daily lectures for a semester within 20 seconds.

PROJECTS

Online Food Delivery Industry | *Tableau, GCP, Apache Beam, Apache Airflow, Google Cloud Storage, BigQuery*
- Engineered and maintained a robust batch **ETL pipeline on Google Cloud Platform (GCP)**, managing daily transactions and achieving a **92% improvement in efficiency** using **Apache Beam, Apache Airflow and Dataflow.**
- Designed and implemented data flow strategies that ensured seamless data ingestion and processing, enabling rapid extraction of actionable insights for the online food delivery industry.
- Employed **Tableau** to create detailed **data visualizations and reports**, driving business intelligence and strategic decision-making.
- Optimized data storage and retrieval using **Google Cloud Storage & BigQuery**, accelerating data processing & cutting operational costs.

Intraday Algotrading Bot: trAIde | *Python, Pandas, NumPy, Matplotlib*
- Used **object-oriented programming** in **Python** to structure trading algorithms and leveraged **random forest and time series analysis** to shortlist 10 new companies every day for daily trades, simulating and calibrating strategies to optimize outcomes.
- Implemented **multi-threading** to significantly increase computational speed and enable efficient simultaneous trade execution, resulting in increased trade volume and achieving a **17.89% annual compounded profit**.
- Created engaging **data visualizations**, including profit/loss graphs, trade distributions, and real-time stock prices, to effectively communicate investment growth and trading performance to users.