

VATSAL SHETH

(878) 834-9437 | vus@andrew.cmu.edu | <https://www.linkedin.com/in/sheth-vatsal/>

Education

CARNEGIE MELLON UNIVERSITY

Masters of ISM - Business Intelligence and Data Analytics, GPA: 3.7/4.0

Pittsburgh, USA

August 2023 – December 2024

- Coursework: Advanced Machine Learning, Unstructured Data Analytics, Inter-Statistics, Distributed Systems, Big Data

NMIMS UNIVERSITY

Bachelor of Technology (Honors), GPA: 3.47/4.0

Mumbai, India

July 2017 – August 2021

- Major: Electronics and Telecommunication Engineering | Minor: Artificial Intelligence and Machine Learning
- Coursework: Machine Learning Foundations, Fuzzy Logic, Neural Networks, Computer Vision, Object-oriented Java

Work Experience

Carnegie Mellon University

Researcher

Pittsburgh, USA

November 2023 – Current

- Collaborating with researchers at CMU on a data-driven hub for policymakers, employers & workers in emerging sectors
- Implementing dynamic dashboards with R Shiny, AWS and Docker to enhance data visualization and analytics capability
- Engineering machine learning models with a focus on optimizing data processing for faster dashboard load times

Accenture

Software Engineer

Mumbai, India

June 2021 – August 2023

- Led frontend implementation and architecture systems design of an IAM platform creating/implementing technical roadmaps
- Performed automation using Python and Splunk APIs for QA team members to help with production testing & test coverage
- Built an auto-pay feature employing AWS glue jobs in Python and PostgreSQL improving consumer experience
- Implemented full-stack development for an internal ticketing system leveraging Java 8, Servlets, Angular and Spring MVC
- Streamlined systems, saving the client approximately 67% in resources and finances by identifying and optimizing gaps

Projects

Financial Equity Research Analysis using Large Language Models (Research Project)

Technologies/Libraries – OpenAI API, LangChain, Python, BeautifulSoup, Requests, Pandas, React, Pinecone

- Utilized web scraping and cron job, converting text to embedding vectors for vector database, improving data accessibility
- Created a ReactJS chatbot interface with embedding vector conversion for seamless user interaction and query efficiency
- Leveraged LangChain and OpenAI API to extract relevant data chunks and generate answers, enhancing the analysis system

Location Analytics for Business Markets (Capstone Project for IBM Professional Certification)

Technologies/Libraries - Python, SAS, JSON, Geocoder, Foursquare API, Folium

- Leveraged geo-tagging, statistical packages and support vectors to cluster neighbourhoods for dataset creation and EDA
- Utilized predictive analysis to gain new insights from location information retrieved translating to business recommendations

Genre Classification and Recommendation System (Independent Research Study)

Technologies/Libraries - Python, TensorFlow, Librosa, SQL, Numpy, Seaborn

- Conducted Extract, Transform and Load (ETL) on sound waves using time series analysis techniques to prepare data sets
- Visualized data using Seaborn for exploratory data analysis, identifying crucial data points and refining the dataset
- Achieved 78% accuracy with ML for small datasets and 92% with a convolutional neural network for large datasets
- Orchestrated a recommender system for song recommendations using collaborative filtering and statistical modeling

Leadership & Activities

Software Engineer, Accenture:

- Headed a 4-person team, teaching conflict resolution, decision-making, abstract thinking and the importance of feedback
- Improved written and verbal communication skills through close collaboration with a global group of people
- Cultivated the ability to convey complex technical concepts to non-technical stakeholders

Tech-Tank Head, NMIMS University:

- Conceived and led 'Tech-Tank,' a competition linking startups with venture capitalists, resulting in funding for top winners
- Developed skills like adaptability, self-starter, strong work ethic and organizational skills, in a cross-functional team context
- Guided a team of 15 people cultivating issue-discovery skills, problem-solving skills and an eagerness to ask questions

Publications & Certifications

- "Genre Based Music Classification using Machine Learning and Convolutional Neural Networks" ICCCNT 2021 under IEEE
- IBM Data Science Professional Certificate (9-course certification)
- Machine Learning Specialization by University of Washington (4-course certification)

Technical Skills

Programming Languages: Python, R, SQL, Java, JavaScript, Swift, C++, HTML5/CSS3

Technologies: TensorFlow, Numpy, Pandas, Scikit-learn, PyTorch, Hadoop, Docker, Spring Boot, Angular, JSP, Keras, OpenCV

Source Control & CI/CD: GIT, Bamboo, Bitbucket

Databases & Analytics: PostgreSQL, Amazon RDS, SAS, Tableau, Excel Solver, Pinecone