

Harshita Pasumarthu

469-825-9341 | harshitapasumarthu@gmail.com | [LinkedIn](#) | [GitHub](#) | [Website](#)

Education

University of Texas at Arlington, Arlington, TX Jan 2023 - Dec 2024
Masters in Computer Science (Specialization: Intelligent Systems and Big Data) - GPA: 4.0/4.0

- **Courses:** Design Analysis and Algorithms, Database Systems, Machine Learning, Data Mining, Data Analysis and Modelling Techniques, Computer Vision, Pattern Recognition, Cloud Computing and Big Data

Vellore Institute of Technology Vellore, TN, India Jul 2015 - Jun 2020
MTech Integrated Software Engineering - GPA : 8.99/10

- **Courses:** Web Technologies, Software Engineering, Agile Software Development, .Net Programming

Skills

Programming Languages & Frameworks: Python, JavaScript, Java, C++, SQL, React.js, Node.js, Spring Boot
Data Analysis & Visualization: Sci-kit Learn, Pandas, NumPy, TensorFlow, dtw (Dynamic Time Warping), Time Series Analysis, Bayesian analysis, Plotly, Matplotlib, NLTK, Tableau, Power BI, Monte Carlo Simulation,
Tools & Technologies: Docker, Postman, Git, Splunk, MapReduce, Airflow, Kafka, REST API, ETL processes, Data Warehouse
Certifications: C3.ai Advanced Application Development, Hacker Rank Intermediate Problem Solver, Data Science Orientation, Machine Learning with Python, Academy Accreditation - Generative AI Fundamentals

Experience

Office of Provost-Academic Affairs, UT Arlington Arlington, TX, United States
Graduate Student Data Analyst Aug. 2023 – Current

- Automated survey data extraction, reducing report generation time by 85%. This streamlined process allowed for faster decision-making regarding course offerings and resource allocation.
- Enhanced the UTA Administration website UI using Sitecore CMS, resulting in a improvement in user satisfaction.
- **Technologies Used:** Python pandas, numpy, SiteCore, matplotlib

Royal Dutch Shell Bangalore, India
Software Developer - Data Jul 2020 - Dec 2022

- Designed and developed robust ETL pipelines for seamless data integration into the C3.AI platform, enabling over 10,000 equipment pieces from multiple Shell assets to go live efficiently.
- Developed real-time data retrieval jobs from Azure Blob storage, seamlessly integrating with the C3.AI platform. Implemented alert scripts via EBS and ensured pipeline reliability through Splunk monitoring.
- Demonstrated proficiency in data integration and management, enhancing the reliability and accessibility of data within the C3.AI platform ecosystem through cross-functional collaboration.
- Boosted user engagement by 30% with a React-based redesign, enabling TAO specialists to train ML models for equipment failure detection.
- Created a React.JS front-end for real-time monitoring of 10,000+ equipment pieces using C3.ai and Plotly.js.
- Designed RESTful APIs and GraphQL endpoints for seamless communication in the backend. Increased code coverage to 80% by developing test strategies and writing unit tests with Jest.
- **Technologies Used:** JS, Python, Docker, C3.AI, Azure Blob, React.JS, Node.JS, Jest, Plotly.js, Postman, Azure DevOps, Github, Splunk

Software Developer -Data Intern(Coop) Aug 2019 - Jun 2020

- Developed a novel oil exploration POC using Dynamic Time Warping (DTW) for seismic data analysis.
- Validated equipment modeling results in Python during code migration for reliable performance forecasting
- Optimized Python scripts for equipment modeling, improving code readability and maintainability for potential integration with machine learning pipelines, resulting in a 20% reduction in processing time.
- **Technologies Used:** Python, Pandas, Numpy, Scikit, Pickle, matplotlib, dtw

Projects

Philia | [link](#) Oct 2023
• Developed an innovative cross-platform application utilizing Generative AI, empowering users to craft engaging social media content, fostering meaningful connections with loved ones.

Image Segmentation using KMeans | [link](#) Jan 2024- May 2024
• Improved image segmentation using KMeans clustering and Histogram of Oriented Gradients (HOG) features, achieving higher accuracy with IoU metric.

Automated Receipt Classification using CV, NLP, and ML Techniques | [link](#) Jan 2024-May 2024
• Developed a system utilizing computer vision (CV), natural language processing (NLP), and machine learning (ML) to automate receipt categorization for efficient expense tracking and financial analysis.

Phonebook API Input Validation | [link](#) Aug 2023-Dec 2023
• Implemented Dockerized phonebook application in IntelliJ using Maven, validated CRUD operations with Postman, and ensured seamless deployment and efficient project management through containerization and version control