Amit Prajapati

aprajapati@wpi.edu | +1 (508)410-9274 | LinkedIn | GitHub | United States

EDUCATION

Master of Science in Data Science, Worcester Polytechnic Institute, MA, USA **Aug 2023 - May 2025**

- Relevant courses: Data Structures and Algorithms, Big Data Management, Generative AI, Business Intelligence, Machine Learning, Deep Learning, Artificial Intelligence

Bachelor of Technology in Data Science, NMIMS, India **Aug 2019 - May 2023**

SKILLS

Languages: Python, R, JavaScript, TypeScript, PL/SQL, C, C++, Swift, MySQL, MongoDB, Apache Spark

Machine Learning: TensorFlow, PyTorch, Scikit-Learn, Langchain, OpenCV, NLTK, NumPy, Pandas, Dask, SciPy

Cloud Technologies: MLflow, Docker, Git, AWS EC2, Azure, Snowflake, CI/CD, Kubernetes

EXPERIENCE

Software Development Engineer –True Light Energy, Boston, MA **Jan 2025 – Present**

- Migrated workflows from Excel to a PostgreSQL database on AWS, improving scalability and reducing data errors by 40%, while enabling faster queries that cut reporting time from minutes to seconds.
- Engineered automated ETL pipelines in Python to fetch, clean, and store over 500K+ energy data records weekly, eliminating manual processing and ensuring 99% data accuracy.
- Built a web-based UI with interactive charts and dashboards, giving stakeholders access to real-time insights and reducing manual report preparation by 15+ hours per week.

Software Engineer Intern – OCA, Boston, MA **May 2024 – Aug 2024**

- Engineered multiple end-to-end data pipelines to ingest and process real-time wave sensor data from buoy APIs and internal document repositories, storing outputs in SQL databases and integrating with visualization tools like Power BI.
- Developed automated alerting and monitoring workflows with API-driven services, enabling early weather warnings that reduced wasted ship hours and operational delays by 30%.
- Implemented a RAG-based customer service chatbot using FAISS and LLaMA 3, cutting query response times from hours to less than 2 seconds and scaling customer support to 24/7 global availability.

Data Analyst – Munich RE, India **Dec 2022 – May 2023**

- Built a fraud detection pipeline on Azure Data Lake to ingest and process large-scale claims data, applying association rule on historical records to flag suspicious patterns.
- Integrated processed outputs into Power BI dashboards, enabling stakeholders to quickly visualize anomalies and make informed decisions.
- Improved claims validation efficiency by 60% and reduced fraudulent claims by 40%, significantly cutting down manual hours.

PROJECTS

Event Search Web and iOS App | Swift, Angular, Node.js, REST, HTML, CSS, GCP **|Link**

- Constructed a scalable full-stack app for event searches, integrating Ticketmaster & Spotify APIs for real-time data retrieval.
- Utilized Angular & Bootstrap to ensure seamless cross-platform functionality; deployed on GCP App Engine for high scalability
- Programmed robust server-side operations using Flask and Node.js, enabling efficient data processing and user interaction.

Object Detection with Satellite Images | Python, Computer Vision, GitHub, Yolo | **Link**

- Optimized the trade-off between image resolution and object detection accuracy in satellite imagery by testing models at varying resolutions, reducing data requirements by 70% while maintaining 80% detection accuracy.
- Implemented a YOLO-based training pipeline on the xView dataset, logging class-wise performance at different resolutions and generating resolution–performance curves to detect “knee points” for optimal thresholds.

WPI-Bot | AWS, LLMs, Python | **Link**

- Developed a campus chatbot using Stream-lit and FAISS, indexing 500+ WPI webpages so students could get instant, accurate answers instead of manually browsing.
- Engineered the backend pipeline with web scraping, text cleaning, and vector embeddings of data, powering a retrieval-augmented AI model (LLaMA-3) that improved response relevance and reduced search effort by 70%.