Praveen Ganapathy Ravi

+1 774-955-1379 · pravi1@umassd.edu · linkedin.com/in/pravi01 · github.com/PraveenGanapathy

About

Data Engineer and Software Developer with over 4 years of experience in backend development, ETL workflows, and cloud-native applications. Holds a Master's in Data Science from UMass Dartmouth with a focus on data pipelines, statistical modeling, and interactive visualization. Proficient in Python, SQL, AWS, Apache Airflow, Power BI, and D3.js. Known for building scalable, production-ready systems that support data-driven decision-making across diverse business domains.

Education

University of Massachusetts Dartmouth — Jan 2024 – May 2025
Master of Science in Data Science (GPA: 3.97) — *Dartmouth, MA*

- Coursework: Advanced Data Mining, Artificial Intelligence, Big Data Analytics, High Performance Scientific Computing, Database Design, Advanced Machine Learning, Advanced Mathematical Statistics, Visual Analytics, Statistical Analysis.
- Served as a Career Ambassador – Data and Reporting Specialist, and Grader for DSC 520 and MTH 280.
- Clubs: Data Science Club, Outdoor Club, GenAI Explorers, Cyber Secure Computing Club.

Anna University Chennai — Jun 2016 – Nov 2020
Bachelor of Engineering, ECE (GPA: 7.39) — *Chennai, India*

- Completed coursework in electronics, signal processing, embedded systems, and communication theory.
- Published paper on "Design of Microstrip Fractal Antenna for Vehicle-to-Vehicle Communication" in Journal of Xidian University.
- Activities: Eco Club, Tech Symposiums, Embedded Workshops.

Technical Skills

- **Programming:** Python, SQL, R, JavaScript, PHP, HTML/CSS
- **Data Engineering:** ETL Pipelines, Apache Airflow, Data Warehousing, Migration
- **Data Analysis:** EDA, Hypothesis Testing, Predictive Modeling, Statistics
- **Cloud/DevOps:** AWS (EC2, S3), Docker, CI/CD, Scalability
- **Databases:** PostgreSQL, MySQL, SQLite, DB Design, Optimization
- **Visualization:** Power BI, Tableau, Streamlit, Matplotlib, Seaborn, D3.js
- **Libraries/Frameworks:** Pandas, NumPy, Django, Streamlit, Next.js, Symfony
- **Tools:** Git, Jupyter, VS Code, Excel, Shell Scripting

Work Experience

University of Massachusetts Dartmouth — Sep 2024 – May 2025
Career Ambassador – Data and Reporting Specialist — *Dartmouth, MA*

- Developed a custom D3.js dashboard to visualize key Career Center performance metrics such as student appointment trends, resume reviews, and employer engagement, improving data accessibility for internal stakeholders.
- Created interactive Power BI dashboards to monitor student participation in events and appointments, streamlining reporting processes and supporting data-informed decision-making.
- Automated data extraction and preprocessing from Handshake using Python and Pandas, reducing manual work and enhancing data accuracy.
- Compiled and analyzed event data to drive strategic decision-making for job fairs and outreach initiatives.

University of Massachusetts Dartmouth — Jan 2025 – May 2025
Grader – MTH 280: Scientific Computation — *Remote*

- Evaluated student assignments and projects involving computational modeling, simulation, and visualization techniques using mathematical programming tools and scripting languages.

- Ensured accurate and timely feedback on coursework requiring a strong foundation in calculus and applied mathematics, supporting student learning in preparation for advanced computational math courses.

University of Massachusetts Dartmouth   Sep 2024 – Dec 2024

Grader – DSC 520: High Performance Computing   *Remote*
- Assessed graduate-level assignments on topics such as parallel processing, memory hierarchies, multi-core systems, and message passing in HPC environments.
- Provided constructive, technical feedback on student submissions related to the design, analysis, and implementation of high-performance computational applications.

Infosys Limited   Jun 2021 – Jan 2024

Data Engineer, Backend Developer   *Chennai, India*
- Designed and executed robust ETL pipelines to migrate 4TB of production data from SQL Azure Virtual Machines to Azure Managed Instances, ensuring zero data loss and improved query performance across client-facing analytics systems. (NTT Data)
- Led backend data engineering efforts for telecom applications at TalkTalk Business (UK), including legacy modernization from Zend/Symfony PHP stacks to newer frameworks.
- Built and maintained CI/CD pipelines using Bamboo, GitLab, and Azure DevOps, reducing deployment time by 40% and enhancing release stability.
- Managed SEO domain migration and infrastructure consolidation projects, maintaining over 99% data integrity and minimizing service disruption during large-scale rollouts.
- Awarded Infosys INSTA Awards (Sep 2023, Dec 2022) for outstanding contributions to project leadership, enabling $50K+ in client operational savings.

Wipro Limited   Feb 2020 – May 2021

Data Engineer   *Chennai, India*
- Completed intensive technical training in backend development and data center operations, with strong performance in Python-based application development modules.
- Developed and optimized sample web application modules in Python during training, achieving improved response time benchmarks in simulated environments.

PROJECT EXPERIENCE

Brain Tumor Detection Using Deep Learning - Python, TensorFlow, Streamlit *(Spring 25)*   [Project Demo](#)
- Developed an MRI-based CNN classification system using EfficientNet and other pretrained models with 94.87% accuracy.
- Deployed Streamlit interface for live inference; enhanced clinical decision-making utility.

DocViz – Interactive Document Clustering Platform - Python, Node.js, D3.js *(Spring 25)*   [Project Demo](#)
- Built a document analysis platform with TF-IDF, KMeans, and MDS visualized using D3.js.
- Enabled multi-document interaction, clustering, and visual insights in a browser-based UI.

Career Centre Analytics Dashboard - Flask, D3.js, Power BI, Azure *(Spring 25)*   [Project Demo](#)
- Created a dashboard with Flask and D3.js to track UMass career center metrics using Azure Blob Storage.
- Improved internal reporting by automating metric updates from Excel sheets.

Flight Delay Analysis - Apache Airflow, Python, Streamlit, AWS S3 *(Fall 24)*   [Project Demo](#)
- Built a data pipeline with Apache Airflow, Papermill, and Streamlit to analyze FAA flight delays.
- Used AWS S3 for data storage and Docker for deployment orchestration.

SimuChat – AI Roleplay Game (Hackathon Winner) - Next.js, Shadcn, AI/LLM *(Fall 24)*   [Project Demo](#)
- Developed a real-time roleplay game with 36 AI characters and 12 scenarios using Next.js and Shadcn.
- Won first place for innovation in Jia.Seed hackathon with 15K+ interactions and 95% engagement accuracy.

Local Library Management System - Django, SQLite, Ajax *(Summer 24)*   [Project Demo](#)
- Designed a Django-based system with user auth, book browsing, and permission-based lending.
- Hosted on PythonAnywhere; supported DB ops with SQLite and Ajax validation.

Bank Marketing Dataset Analysis - Python, Streamlit, XGBoost, Sklearn *(Spring 24)*   [Project Demo](#)
- Built predictive models (Random Forest, XGBoost) to classify term deposit subscriptions.
- Created a Streamlit dashboard for interactive client behavior prediction.