

Hari Prasannaa Thangavel Ravi

LinkedIn | hariprasannaa2001@gmail.com | 202-743-8633 | Portfolio | GitHub

WORK EXPERIENCE

MyEdMaster

Ashburn, VA

Data Science Intern

May 2025 – Aug 2025

- Designed and deployed a recommendation engine using embeddings and engagement signals, improving candidate–startup match relevance by 12%.
- Refined ranking models with iterative feedback integration, increasing connection success rate by 3% while ensuring accessibility compliance.

Data Science for Sustainable Development

Washington, DC

Data Scientist

Apr 2025 – Present

- Built time-series forecasting models achieving 91% accuracy in campus energy demand prediction, reducing energy waste by 16%.
- Designed anomaly detection workflows on IoT telemetry across 50+ buildings, optimizing resource allocation and lowering operational costs.
- Integrated telemetry dashboards with Tableau and Python APIs to monitor model drift and performance within SLA thresholds.

Facilio Technology Solutions

Chennai, India

Member of Technical Staff

May 2022 – Jul 2024

- Engineered a custom analytics platform with JavaScript and Java, enabling real-time dashboards and driving an 8% customer base increase.
- Designed and deployed REST APIs to support scalable data access and reporting, improving integration with client systems and enhancing reliability.
- Optimized SQL queries and ETL workflows for time-series data, reducing reporting latency by 15% and decreasing manual workload.
- Applied statistical and ML models on telemetry KPIs to forecast consumption trends and detect inefficiencies in building systems, enabling proactive maintenance interventions and improving system reliability.
- Collaborated with cross-functional engineering and product teams to resolve architectural-level issues, improving platform stability and increasing retention by 8%.

PROJECTS

Intelligent Document Q&A Assistant with RAG

Mar 2025 – May 2025

- Built an LLM-powered retrieval system using embeddings and FAISS vector DB, enabling context-aware Q&A over enterprise knowledge bases.
- Designed and deployed a RAG pipeline, improving answer relevance by 8% compared to keyword search baselines.
- Implemented monitoring for latency, token usage, and hallucination rates, reducing inaccuracies by 12% while maintaining sub-500ms response times.

Churn Prediction Pipeline

Jan 2025 – Mar 2025

- Developed an automated retraining pipeline with Apache Airflow for 5M+ HR records, orchestrating ingestion from BigQuery through model scoring with version control.
- Trained and evaluated XGBoost models with categorical encoding and SHAP explainability, achieving accuracy within 2% of benchmark models.
- Implemented continuous monitoring and weekly refreshes to mitigate data drift, ensuring reliability of live production models.
- Integrated outputs into Looker dashboards with ROI simulations, enabling leadership to launch targeted retention campaigns.

EDUCATION

The George Washington University, Columbian College of Arts & Sciences

Washington, DC

Master's in Data Science

Aug 2024 – May 2026

- Cumulative GPA: 3.88/4.0, Global Leaders Award
- Relevant Coursework: Machine Learning, Reinforcement Learning, Cloud Computing, NLP

TECHNICAL SKILLS & CERTIFICATIONS

- Languages:** Python, SQL, Java, R, C++
- AI/ML:** Neural Networks, Scikit-learn, TensorFlow, NLP, Time Series Forecasting, RAG, Fine-tuning
- Data Analysis:** Feature Engineering, Statistical Modeling, A/B Testing, Hypothesis Testing
- Visualization:** Tableau, Power BI, Looker
- Databases:** MySQL, PostgreSQL, MongoDB, Neo4j
- Cloud & Tools:** AWS, GCP (BigQuery, GCS), Airflow, Git, Jupyter Notebook, Postman
- Deployment & Reliability:** REST API design, Unit/Integration Testing, CI/CD, ETL, Model Drift Monitoring, SLA Compliance