

Rindhuja Treesa Johnson

Baltimore, MD (Willing to Relocate) ♦ 914-746-5465 ♦ rindhuj1@umbc.edu ♦ [LinkedIn](#) ♦ [GitHub](#) ♦ [Portfolio](#)

SKILLS

- **Tools:** SQL, Python, C++, MS Power BI, Tableau, MS Excel, Streamlit
- **Cloud & Big Data:** Google Cloud Platform, Snowflake, AWS, Azure, Databricks, Spark, Hadoop
- **Database Management:** MySQL, MS SQL Server, Azure Data Studio, PostgreSQL
- **IDE & Project Management:** Docker, Jupyter NB, Visual Studio Code, GitHub, Jenkins, Google Colab, SharePoint
- **Python APIs/Lib:** Pandas, NumPy, Matplotlib, Seaborn, StatsModels, Sci-Kit Learn, PySpark, TensorFlow, XGBoost, PyTorch, NLTK, Keras, LangChain, Transformers, HuggingFace, Flask

EXPERIENCE

Data Scientist, Smart Ecosystems Inc. (Outlier) Jan 2024 - Present

- Developed **ETL pipelines** in Python and SQL to clean and transform **large** datasets.
- Conducted **regression analysis** and **Bayesian inference** to analyze customer trends and KPIs for strategic planning.
- Deployed **ML models** with TensorFlow and Sci-Kit Learn to improve operational efficiency.
- Designed **Tableau dashboards** visualizing KPIs, leading to data-driven decisions by stakeholders.
- Enhanced **AI model accuracy** by 25% through feature engineering and pattern analysis.
- Achieved 95% model accuracy via advanced **QA frameworks** and validation protocols.
- Automated **content validation workflows**, ensuring high-quality, human-like AI responses.
- Improved AI-generated content accuracy by 30% using **Supervised Fine-Tuning (SFT)** and **RLHF** techniques.
- Built **recommendation systems** using TensorFlow and PyTorch, increasing conversions by 18%.
- Collaborated with cross-functional teams to align data strategies with business objectives.

Junior Data Scientist, Leveragai Inc. Dec 2023 - Present

- Built **HIPAA-compliant Power BI** dashboards for patient recovery and treatment metrics.
- Reduced ER wait times by 25% using **predictive models** in Python and SQL.
- Analyzed **CMS datasets (Payroll-Based Journal)** in SQL, delivering actionable staffing insights for administrators.
- Deployed **LLM models** based on **RAG** with AWS SageMaker, improving clinical documentation by 35%.
- Led the development of an **AI Tutor for a Generative AI** using Azure OpenAI and tested for **quality assurance**.
- Designed and implemented a **Retrieval-Augmented Generation (RAG)** pipeline for real-time query resolution.
- Fine-tuned **LLM models** to generate LaTeX-formatted responses for math and technical content.
- Optimized **prompt engineering** strategies to improve AI accuracy and user engagement by 30%.
- Analyzed user interactions to refine model outputs, boosting content relevance and precision.

Data Analyst, Redwood Algorithms Jan 2021 - Dec 2022

- Increased **customer engagement** by 180% through predictive analytics and A/B testing.
- Automated **social media performance metrics** reporting with SQL and BigQuery.
- Designed **Power BI dashboards** for real-time user engagement and sentiment analysis.
- Conducted **CTR analysis**, boosting ad campaign ROI by 25% on digital platforms.
- Analyzed **customer behavior trends** with Python, deriving insights for marketing efficiency.
- Implemented SQL-based **data pipelines** for seamless integration of social media metrics.
- Presented insights to stakeholders, **driving strategic decisions** and improving communication.
- Optimized **ad targeting strategies**, enhancing customer reach and conversion rates.

TEACHING EXPERIENCE

Graduate Teaching Assistant, UMBC Jan 2024 – May 2024

- Tutored 50+ students in Apache Spark, Hadoop, Databricks, and AWS for big data analytics.
- Graded and provided feedback on data-intensive assignments, improving project quality by 20%.
- Guided students on scalable data engineering workflows and Spark MLlib implementations.
- Recommended AI-driven study aids, integrating LLM-generated explanations into class support.
- Enhanced learning outcomes through real-time debugging of Python and PySpark assignments.
- Maintained transparent communication between professors and students, reporting data integrity issues.

RTTP Math Coach, UMBC Sherman Scholars' Program Oct 2023 – May 2024

- Improved arithmetic proficiency of K-12 students by 70% using adaptive learning strategies.
- Implemented AI-enhanced practice tools, leveraging prompt engineering for custom problem generation.
- Collaborated with AI models to generate LaTeX-formatted content for advanced math exercises.
- Designed personalized learning frameworks informed by performance analytics and student data.
- Increased engagement by 50% through data-informed session planning and interactive activities.

PROJECTS

[CLV Prediction for Insurance Companies with Python, PowerBI, AWS, SQL, Streamlit, Gemini, and LangChain](#)
[Steam Review Analysis on Big Data using Apache Spark and Hadoop in Python](#)

[Time-series Analysis and Forecasting of Forex Rates with SARIMAX models](#)
[Nutrition Dashboard for US-based food produces in Tableau](#)
[Stock Portfolio Optimization and Dashboard Using Microsoft Excel, SQL, and Python](#)

EDUCATION

University of Maryland - Baltimore County (UMBC)

Baltimore, MD

Master of Professional Studies in Data Science

GPA: 4.0/4.0

Pondicherry University

Puducherry, India

Master of Science in Physics

GPA: 3.61/4.0 Valedictorian

HOBBIES

- Blogging: [Medium](#)
- Badminton
- Freelance Prompt Engineering – RLHF and SFT