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) Aug 2024 - Present

- Experimented with **5+ prompt engineering** techniques to optimize the LLM's response latency and quality.
- 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) techniques.
- Led **RLHF initiatives**, mitigating loss categories and refining AI model performance.
- Conducted **qualitative analysis** on RLHF outputs, enhancing model feedback loops.
- Collaborated with cross-functional teams and QMs to be up-to-date on the client's requirements.

Junior Data Scientist, Leveragai Inc. Aug 2024 - Present

- Led the development of an **AI Tutor for a Generative AI** course using Azure OpenAI.
- 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.
- Developed lesson plans and curated **AI-generated content** to enhance the learning experience.
- Integrated **Bing Search APIs** to enrich AI Tutor responses with up-to-date contextual knowledge.
- Optimized **prompt engineering** strategies to improve AI accuracy and user engagement by 30%.
- Conducted **quality assurance testing**, ensuring AI Tutor responses met educational standards.
- Collaborated with cross-functional teams to align AI content with curriculum objectives.
- Analyzed user interactions to refine model outputs, boosting content relevance and precision.

Data Analyst, Redwood Algorithms Jan 2022 - 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.

Graduate Coordinator/ Orientation Advisor, UMBC Jul 2023 – Feb 2024

- Mentored 70+ first-year incoming undergrad students, leveraging data-driven insights to enhance onboarding.
- Organized workshops promoting STEM skills, including data analytics and introductory Python coding.
- Facilitated peer learning communities, using AI tools for personalized learning paths.
- Integrated LLM-assisted knowledge bases to answer student queries with precision.

- Achieved a 4.75/5 satisfaction score by customizing guidance through structured interviews.

PROJECTS

CLV Prediction for Insurance Companies with Python, PowerBI, AWS, SQL, Streamlit, Gemini, and LangChain
AI Chatbot tailored to K-12 students with Local RAG using LangChain, Ollama, and Chroma DB
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