**Krithika Rajendran**

Software/Data Engineer

**rkrithika1993@gmail.com**| **(682)-306-2302**| <https://www.linkedin.com/in/krithika-rajendran>

**Summary**

Motivated Data Engineer with 4+ years of experience building and maintaining ETL pipelines, working with database technologies, and improving data quality and reliability. Skilled in Python, SQL, and big data technologies like Spark and Hadoop. Passionate about mentoring junior team members and continuously advancing data engineering practices.

**Skills**

* **LLM Expertise:** LLM Integration, Prompt Engineering, Fine-tuning, Model Deployment, Open-source LLM Optimization (Hugging Face), LLM-powered Chatbot Development
* **Machine Learning/AI:** PyTorch, TensorFlow, Scikit-learn, XGBoost, NLP, Deep Learning, Model Optimization, AutoML
* **Data Engineering:** Python, SQL, Spark, Hadoop, Kafka, Data Pipelines, AWS (Lambda, S3, Redshift), Google Cloud Platform
* **Software Engineering:** Java, Scala, Algorithms, Git, Agile Methodologies
* **Other:** Data Visualization (Tableau), Communication, Teamwork, Problem-solving, Creative Thinking

**Experience**

**RadicalX** | Artificial Intelligence Engineer | Aug 2023 – Present

* **Enhanced Developer Productivity with LLMs:** Spearheaded the development of AutoGen, a personalized pair programming tool utilizing Vertex AI Embeddings and LangChain, leading to a **10% improvement in developer efficiency**.
* **Showcased Advanced LLM Capabilities:** Implemented a Q&A bot with RAG and Chroma Vector database, achieving **80% faster query handling** and **95% retrieval accuracy**. Created an interactive chat interface with StreamLit and Gemini to explore and demonstrate the power of LLMs for educational applications.
* **Built Podcast Summarization Tool**: Developed a podcast summarization tool leveraging OpenAI's GPT-4 Function calling feature, enabling users to quickly grasp key points and insights from podcasts. Implemented data pipelines using AWS Lambda, S3, and Redshift for efficient data processing and storage.

**Ballotpedia** | Data Engineer Intern | Sep 2022 – Jan 2023

* **Streamlined Data Extraction:** Engineered Python scripts with Selenium, Beautiful Soup, and XPath, achieving a **67% improvement in data extraction efficiency** and contributing to faster project delivery.
* **Developed Innovative OCR Solution:** Led the development of an OCR tool using LayoutLMV3, enabling accurate data extraction from complex ballot sheet images and **automating a previously manual process**.
* **Explored Cloud-based Data Management:** Researched and proposed the use of AWS Textract, SageMaker, and other cloud services for efficient document processing and data management, demonstrating initiative and a commitment to exploring new technologies.
* **ETL and Data Analysis:**

• Created indexes, materialized views, and optimized complex SQL and Pandas queries to improve data pipeline performance by 30%  
• Continuously improved quality, efficiency, and scalability of data pipelines through workflow automation and modular design  
• Mentored junior team members by conducting code reviews and creating detailed documentation to aid future optimization efforts  
• Performed in-depth analysis on Kickstarter campaign data to uncover key success factors and provide actionable recommendations to stakeholders

**Tata Consultancy Services** | Big Data System Engineer | Dec 2015 – Dec 2018

* **Successfully Migrated Terabytes of Data:** Led a large-scale data migration project using Spark-SQL and Scala, ensuring data integrity and exceeding client expectations by delivering results ahead of schedule.
* **Optimized Real-time Data Pipelines:** Enhanced real-time data streaming pipelines for credit card approval processing, resulting in a **20% reduction in processing time** and improved decision-making capabilities.
* **Developed Efficient Data Processing Solutions:** Created and maintained custom MapReduce jobs for data processing and analysis, handling diverse data formats and contributing to improved data insights.

**Projects**

* **Developed AI-powered Personal Assistant for Digital Journaling:** Built a prototype using Bi-GRU and Bi-LSTM models, achieving 97% AUC accuracy. Further enhanced the system with **Hugging Face Transformers** to analyze user diary entries, extract insights, identify patterns, and provide personalized recommendations. Explored the application of iterative feature processing and time series analysis to deepen user behavior understanding and enhance the assistant's capabilities.
* **Big Data Graph Analysis:** Implemented graph substructure discovery on large datasets using Scala Spark, achieving a **50% improvement in application performance** through optimization techniques. This project highlights my ability to work with big data technologies and deliver impactful results.

**Education**

**The University of Texas at Arlington** | Master of Science in Computer Science | Aug 2023

* R for Data Science (Uplimit)
* Information Systems Security Professionals (NSTISSI No. 4011)

**Publications**

**Learning to Grade Short Answers using Machine Learning Techniques**

Women in Computing and Informatics, ACM Digital Library International Conference Proceeding

Krithika Rajendran, Jayasree Narayanan : <http://dl.acm.org/citation.cfm?doid=2791405.2791508>