Cole McIntosh

AI and Data Engineer

Foreword

I am an innovative AI and Data Engineer specializing in developing cutting-edge AI systems and data solutions. With expertise in Large Language Models (LLMs) and AI Agents, I create sophisticated, scalable systems that push the boundaries of what's possible in artificial intelligence and data processing. My work focuses on leveraging the power of LLMs to build intelligent, autonomous systems capable of complex decision-making and problem-solving. By combining technical prowess with a strategic mindset, I deliver AI and data solutions that not only solve current challenges but also pave the way for future innovations.

Experience

Pay Ready - Data Engineer *(2023 - Present)*

- Developed agentic AI systems using Langchain for end-to-end automation of complex business processes.
- Engineered AI-driven knowledge systems, enhancing information access and assistant evaluation company-wide.
- Architected end-to-end LLM solutions for automating business workflows and analytics.
- Implemented multimodal AI systems to process and analyze diverse data types, improving decision-making capabilities.
- Created a RAG (Retrieval-Augmented Generation) Slack bot operating over internal documentation, significantly improving information accessibility and team productivity.

Otter - Scaled Tooling Analyst *(2022 - 2023)*

- Built scalable Python data pipelines for cross-team analytics and reporting, incorporating machine learning models for predictive insights.
- Implemented AI-driven internal tools using natural language processing techniques to enhance operational efficiency and data-driven insights.
- Developed a RAG chatbot using Langchain, integrating Jira team spaces to enhance cross-team collaboration and provide better visibility and understanding of various business aspects.

RealPage - Business Analyst *(2021 - 2022)*

- Designed and maintained comprehensive business intelligence reporting systems.
- Developed SQL data models for advanced analytics and automated client reporting.
- Created automation pipelines to streamline business processes, significantly reducing manual workload and improving operational efficiency.

Projects

Llama 3.2 1B Mango
Developed a fine-tuned version of the Llama 3.2 1B model, optimized for chain-of-thought reasoning. This model was trained 2x faster using Unsloth and Hugging Face's TRL library, incorporating techniques such as LoRA, QLoRA, and RoPE scaling. The model was trained on the SkunkworksAI/reasoning-0.01 dataset, containing 29.9k examples to improve step-by-step problem-solving abilities.

https://huggingface.co/colesmcintosh/Llama-3.2-1B-Instruct-Mango

Structured Output with Multimodal Agents

Developed a system for generating structured output using multimodal agents, showcasing advanced AI integration and task routing capabilities.
https://github.com/colesmcintosh/structured-output-with-multimodal-agents

Smol Vision

Built a local LLM vision pipeline for efficient image analysis without GPU requirements, demonstrating expertise in optimizing AI models for resource-constrained environments.
https://github.com/colesmcintosh/smol-vision

Reranked RAG

Developed a Streamlit app showcasing a Retrieval-Augmented Generation (RAG) system with reranking, designed to answer questions about me and my work based on my personal website content. Utilizes web scraping, LangChain, Cohere Rerank, Groq's LLM, and FAISS for efficient information retrieval and question answering.

https://cole-mcintosh-rag.streamlit.app

Finance Fine Tuned Mistral 7B

Developed a fine-tuned version of Mistral 7B, specifically tailored for finance-related tasks. This model was trained on the alpaca finance dataset, enhancing its capabilities in financial analysis, prediction, and domain-specific language understanding.

https://huggingface.co/colesmcintosh/mistral_7b_finance_finetuned

Skills

- Programming Languages: Python, SQL
- Cloud Platforms: AWS, Google Cloud
- Tools & Frameworks: Langchain, AG Grid, Salesforce, Slack Automations
- Machine Learning: PyCaret (AutoML), Huggingface, LLM Evaluation, Fine-Tuning
- AI Techniques: Natural Language Processing, Retrieval-Augmented Generative AI
- Data Management: Data Pipelines, ETL Solutions

Education

- AWS Cloud Technical Essentials - Amazon
- Google Cloud Fundamentals: Core Infrastructure - Google
- Python Specialization - University of Michigan
- SQL for Data Science - University of California, Davis

Afterword

Thank you for taking the time to review my resume. I'm always eager to engage with like-minded individuals and those curious about the possibilities of technology. Please feel free to reach out!