

Zade Alfalah

U.S. Citizen ❖ (715) 797-5940 ❖ Eau Claire, WI ❖ zadealfalah@gmail.com

EDUCATION

University of Colorado, Boulder, CO

MS, Data Science - 3.96 GPA

Dec. 2022

BA, Physics, Mathematics, Astronomy

Jan. 2020

WORK EXPERIENCE

Legrand, Eau Claire, WI

Data Analyst

Jun. 2023 - Dec. 2023

- Developed predictive models and forecasts for stocking predictions
- Led weekly strategy discussions based off of created models and analysis
- Built document ingestion pipelines automating previously manual tasks, saving over \$30,000/yr
- Designed and deployed sales dashboards using Power BI
- Introduced MySQL architecture reducing complexity and cost of analysis

Laboratory for Atmospheric and Space Physics, Boulder, CO

Junior Data Scientist

Sep. 2017 - Dec. 2019

- Architected ETL pipelines for processing large quantities of imagery data
- Optimized imagery analysis scripts, reducing compute times by 22%
- Built visualizations for conference presentations utilizing seaborn and matplotlib
- Onboarded junior scientists utilizing self-created documentation

PERSONAL PROJECTS

jobReqs

Automated Web Scraping and Job Posting Analysis

- Developed and maintained scalable backend for the project pipeline, built on AWS and Docker, running on ECS clusters
- Integrated open-source LLMs (BERT, Llama) for text classification and entity recognition
- Employed OpenAI API for zero-shot results, reducing costs
- Used Ray/Anyscale for distributed training with MLFlow
- Created API with FastAPI for speedy and efficient querying of model

localLibrary

Custom Retrieval-Augmented Generation (RAG) Chatbot

- Built a chatbot to help answer questions about books in my library
- Utilized Langchain with Chroma for vector storage and Ollama as a platform
- Integrated open-source LLMs (Mistral, Llama 2) incorporating prompt engineering for best results

TOOLS

Languages: Python, SQL, R, Bash, Git

Software: AWS, GitHub/GitLab, MySQL, Jupyter, MLFlow, Docker, ArcGIS, Tableau, Power BI, Excel

Libraries/Frameworks: Pandas/NumPy, NLTK/Spacy, XGBoost, PyTorch, Tensorflow/Keras, HuggingFace, scikit-learn, Selenium, Airflow, Scrapy, OpenCV, Snowflake, FastAPI, Langchain, Terraform