SEP AMINIAN

Portfolio: Sephml
Linkedin: amirsepehr-aminian/ Location: London, UK
Contacts: +447878807771 amirsepehr.aminian@gmail.com

SUMMARY

Software and Machine Learning Engineer with expertise in Data Science, SQL, and NLP, developing scalable solutions for enterprise applications. Saved £2.8m annually by developing scalable NLP solutions. Skilled in end-to-end Machine learning development, NLP, and cloud-based applications, looking for opportunities to build scalable NLP applications.

EXPERIENCE

Software consultant - Assistant Manager - S&W Jun. 2024 - Present
London, United Kingdom

- Performing requirements engineering and scoping for POC stage and providing a timeline of delivery
- Engineering of a performant end-to-end machine-learning pipeline leveraging GPT-4 with LangChain in a public cloud micro-service environment to help with core business reporting tasks to reduce time spent for delivery by more than 20 hours per engagement (40% of total time spent), reducing the costs by £2.8m yearly.
- Leveraged Azure SQL for structured data extraction and transformation to support business intelligence reporting.
- Designed and implemented a robust process for monitoring and automating model performance updates, ensuring continuous improvement and efficiency.
- Leveraging my expertise in Python, JavaScript and Cloud technologies to offer strategic guidance and implement efficient software solutions that align with industry best practices for various use cases in professional services.

Software consultant - RSM UK Mar. 2023 - May 2024
London, United Kingdom

- Facilitated the incorporation of Microsoft Copilot into the business as a part of a wider team by designing detailed use case scenarios and integrating considerations on privacy policies and data sharing as well as accuracy. This improved the overall productivity of a team of 50, increasing it by 18% based on project completion time.
- Handling client-facing duties for advisory work by holding discussions with senior technical stakeholders and drafting relevant technical documentation about the projects.

Machine Learning Engineer - Definely Apr. 2022 - Mar. 2023
London, United Kingdom

- Developed NLP classifiers at scale with fine-tuned transformer models (BERT, XGBoost, and Spacy), achieving over 0.9 F1 Score on legal contract analysis.
- Led a team of 2 to engineer and implement a comprehensive text extraction pipeline using Python, and Azure SQL for data preparation pre-training data preparation.
- Engineered a robust data preprocessing pipeline using Azure SQL, optimising query performance and storage efficiency.
- Spearheaded the creation of an efficient pipeline for labelling data, improving and optimising engineer productivity through labelled data management, repository oversight, and task automation.
- Optimised inference pipeline, reducing latency from 800ms to 250ms using ONNX and quantisation.

EDUCATION

MSc. Artificial Intelligence with Distinction- Queen Mary University of London Sept. 2021 - Sept. 2022

- **Relevant coursework**: Introduction to NLP, Advanced NLP, Statistical Machine learning

BSc. Geomatics Engineering with Distinction- University of Tehran Sept. 2016 - Feb. 2021

- **Details**: Minor in Computer Engineering.
- **Relevant coursework**: Advanced Programming, Databases, AI, Data Structures and Algorithms

SKILLS

- **Programming Languages:** Python, JavaScript, C++
- **Machine Learning:** PyTorch, TensorFlow, Keras, LangChain, HuggingFace, LLMs, Transformers, Scikit Learn, NumPy, Pandas, Matplotlib, NLTK, Spacy, Weights and Biases, fine-tuning
- **Frameworks and technologies:** Node.JS, Express.Js, React.Js, Django, Azure cloud, SQL (PostgreSQL, Azure SQL, MySQL), MongoDB, Git, Docker, CI/CD with GitHub Actions, ML Flow
- **Certifications:** Intro to Software Engineering(IBM), Machine learning Summer School 2023(OxML), Neural Networks and Deep Learning(Coursera), Python for Data Science(IBM)

SELECTED PROJECTS

Planning AI Assistant Apr. 2024 - Present

- Led a team of 2 to deliver a scalable ML pipeline handling almost a million documents. Engineered an optimised document retrieval system using vector databases (Pinecone) and distributed processing.
- Implemented a data collection pipeline for gathering planning applications in PDF format with high reliability and stored in a data warehouse.
- Leveraged Flask and MongoDB Atlas to create a micro-service application to extract over 140 useful features from the PDF data by leveraging LangChain and GPT4 and prompt engineering.
- Engineered an AI assistant to answer planning questions and provide references accurately.
- Developed a performant facet search engine to search through for relevant documents from scraped data using Express.Js and Mongo DB.

Machine Learning & Data Science projects Aug. 2022 - Mar. 2024

- Created a pipeline to recover LLM prompts using Gemma7b with a score of 0.68.
- Implemented robust neural network solutions for a variety of NLP tasks, including Natural Language Modeling, Machine Translation, Named Entity Recognition, and Co-reference Resolution.
- created a Seq2Seq Dialogue model with precision, closely adhering to the original paper's specifications. Demonstrated expertise in model development using TensorFlow Keras and PyTorch, incorporating ELMO and BERT for embedding layers. Explored attention-based methods and traditional LSTMs to enhance overall model performance across diverse projects.
- I Implemented a vanilla version of GoogLeNet, VGG16 and ResNet DNNs based on the original papers for MNIST and CIFAR10 datasets.
- I implemented the SRCNN network for a convolutional super-resolution task. Generative adversarial networks were also implemented for an image generation task.
- I implemented a fine tuned YOLO object detection model to efficiently perform real-time image analysis and classification within a real-time detection using camera feed application.

Data Science & SQL Analytics (Personal Project) Aug. 2021

- Engineered a scalable data pipeline using SQL (PostgreSQL) to process and clean 2M+ transactional records, leveraging advanced queries (JOINs, CTEs, window functions) to extract key customer insights efficiently.
- Developed customer segmentation models using K-Means clustering, applying feature engineering techniques (purchase frequency, recency, and average order value) to classify high-value, occasional, and churn-risk customers.
- Optimised SQL queries and indexing strategies, reducing data retrieval time from 5s to 1.2s.