Rohaan Khan

+92 312 2887469 | rohaank67@gmail.com | Karachi, Sindh

Summary

As a proactive and initiative-driven AI and Machine Learning Engineer, I have a demonstrated history of leading and executing innovative projects from conception to completion. Currently, I'm working as an Individual Contributor at Turing, where I evaluate responses of Large Language Models (LLMs) to ensure they meet quality and accuracy standards. Before this, I worked at Alexa Translations, where I specialized in designing state-of-the-art language models, significantly enhancing translation accuracy and efficiency. My expertise in leveraging advanced transformer models and optimizing large language models contributed to groundbreaking advancements in translation technologies. In my role at Centric, I pioneered the development of an AI-driven Applicant Tracking System and versatile AI solutions, including custom chatbots and scalable intra-company assistants. My hands-on approach and ability to innovate in AI and machine learning have consistently driven substantial improvements in operational processes and client satisfaction. My experiences have honed my skills in both technical innovation and project leadership, making me an ideal fit for roles that require a blend of AI expertise and proactive project management.

Experience

Turing | Remote
Data Scientist | 06/2024 - Present

As an Individual Contributor at Turing, I write effective Python code to solve complex problems. I also derive valuable insights from public datasets. I work closely with researchers to understand their needs, communicate results, and help the organization achieve its goals. My role involves evaluating responses from Large Language Models (LLMs), conducting comparative analysis, and providing feedback to improve model performance. Additionally, I utilize my data analysis skills to address important business queries, balancing technical expertise with effective communication.

Alexa Translations | Remote
Machine Learning Engineer | 07/2023 - 06/2024

As a Machine Learning Engineer at Alexa Translations, I was at the forefront of developing state-of-the-art (SOTA) custom language models, specializing in advanced machine learning techniques to enhance translation accuracy and efficiency. My role is pivotal in driving innovations in language translation technologies.

Key Responsiblities

- **Designing Translation Architectures**: Crafting robust architecture frameworks for translation models, ensuring they meet high standards of accuracy and efficiency.
- **Language Translation Development**: Actively working on language translation projects, applying cutting-edge machine learning methodologies to improve translation quality.
- **Research and Innovation**: Conducting comprehensive research in the field of machine translation to stay ahead of the curve in applying the latest advancements.
- **Fine-Tuning Large Language Models (LLMs)**: Expertly fine-tuning LLMs for translation purposes, tailoring them to specific linguistic nuances and requirements.
- **Developing Seq2Seq Models**: Utilizing the latest Performer/Reformer-based transformer models to design and implement complete sequence-to-sequence models for effective translation processes.
- **Data Preparation and Management**: Responsible for cleaning and preparing datasets to optimize model learning and performance.
- **Model Optimization for Performance**: Loading models on multiple GPUs to enhance computational efficiency and reduce inference time.
- **Model Pruning and Quantization**: Implementing advanced techniques such as pruning and quantization to streamline models without compromising their effectiveness or accuracy.

Achievements:

- Successfully developed and deployed custom language models with SOTA accuracy, significantly enhancing translation quality.
- Pioneered the use of Performer/Reformer-based transformer models in our translation workflows, setting new standards in efficiency and accuracy.
- Played a key role in optimizing model performance, achieving notable reductions in inference times while maintaining high levels of accuracy.

Centric | Remote
AI Developer | 01/2023 - 08/2023

As an AI Developer at Centric, my role involves spearheading the development of innovative AI solutions, including a cutting-edge AI-driven Applicant Tracking System (ATS). My expertise lies in applying advanced machine learning techniques to optimize recruitment processes and develop custom AI applications for diverse client needs.

Key Responsiblities

- **Development of AI ATS System**: Led the creation of an AI-powered ATS, leveraging a Named Entity Recognition (NER) based approach to fine-tune BERT models. This system efficiently identifies preferable candidates by analyzing resumesin-depth.
- **Resume Ranking with Siamese Networks**: Implemented Siamese Networks to rank resumes post-selection, enhancing the accuracy and relevance of candidate shortlisting.
- **Custom Chatbot Development**: Fine-tuned LLaMa 2 models to develop custom chatbots tailored to the requirements of multiple clients, enhancing customer interaction and support.
- **Developing Scalable RAG Architecture**: Innovated a custom Retriever-Augmented Generation (RAG) architecture to create scalable intra-company assistants. This involved:
  - Utilizing Vector Databases for storing embeddings.
  - Employing Hierarchical Navigable Small World (HNSW) techniques for efficient chunk selection.
  - Integrating Large Language Models (LLMs) for content generation.
  - Applying cross-encoder/bi-encoder models as rankers to evaluate and enhance the efficiency of the LLM's outputs.

Red Bull
Student Brand Manager | 02/2021 - 03/2023

- Developed and implemented tailor-made sales and marketing plans, leveraging local knowledge to engage with the target audience effectively.
- Acted as a direct point of contact for consumers and customers, promoting product trials and establishing Red Bull consumption on various occasions.
- Served as a brand ambassador, representing Red Bull in a premium manner, identifying market trends, and enhancing product visibility and distribution.
- Collaborated with opinion leaders and contributed to the global community by connecting and sharing innovative ideas.
- Achieved significant personal and professional growth, building key sales and marketing skills, and establishing valuable relationships for future career opportunities.

Skills

Natural Language Processing, Artificial Intelligence, Large Language Models

Education

FAST NUCES
Computer Science | 05/2023