FAIZEL KHAN

Minneapolis, MN | faizelkhan11@gmail.com | +1 (573) 200-4012 | www.linkedin.com/in/faizelkhan/ | Google Scholar Page

SUMMARY

Experienced Software Engineer with a specialization in Deep Learning. Strong expertise in Machine Learning Model Management and hands-on experience in deploying applications in production. Proficient in quantitative analysis and building fast and secure enterprise software solutions. Proven track record of utilizing complex data sets to provide valuable insights.

EDUCATION

M.S. Computer Science (***Focus:*** *Artificial Intelligence*) Sep 2021 - May 2023
University of Minnesota, Minneapolis, MN **GPA**: 3.87

B.S. Computer Science Jan 2015 - May 2018
Southeast Missouri State University, Cape Girardeau, MO **GPA**: 3.77

SKILLS

Languages: Python (Django), Javascript (Vue.js, Node.js)
Databases: Microsoft SQL, PostgreSQL, AWS
Dev Tools: Git, Docker, Airflow, Spark, ETLs, APIs
AI: Scikit-learn, PyTorch, Numpy, Pandas

PROFESSIONAL EXPERIENCE

Data Scientist *University of Minnesota (Minneapolis, MN)* Dec 2021 – May 2023

- Collaborated with cross-functional teams to develop data-driven solutions for various medical research studies.
- Pre-processed and analyzed electronic health records (EHR) data to identify patterns and trends, optimizing data pipelines for efficient use in data analysis.
- Co-authored two research papers, contributing to the analysis, interpretation, and visualization of research data.

Data Science Intern *Boston Scientific (Minneapolis, MN)* May – Aug 2022

- Developed a semantic search engine to index documents efficiently using NLP and software engineering methods.
- Implemented an ETL pipeline to refine the text corpus for the NLP model, utilizing NLP libraries such as NLTK and Spacy.
- Built and deployed a full-stack web application on AWS, ensuring scalability and reliability in a production environment.

Lead Software Engineer *BioKyowa Inc (Cape Girardeau, MO)* Jun 2018 – May 2021

- Build a full-stack web app using Django and PostgreSQL, managing data workflows for the company's manufacturing operation.
- Saved 3,000 work hours annually by developing custom system monitoring tools using Python, improving analytics capabilities.
- Designed an AI-powered chatbot as a proof-of-concept using NLP techniques to respond to employee queries effectively.
- Led a team of junior software engineers to deliver high-quality and scalable web applications using Agile methodology.
- Proposed and developed technical solutions based on customer requirements and product goals.
- Improved system performance by 28% by implementing an ETL pipeline using Django and Pandas.
- Saved $200k annually in overtime pay by taking initiatives to introduce Learning and Work Management systems.
- Negotiated up to 30% cost reduction in yearly contracts with vendors like Monday.com.

RESEARCH PROJECTS

Meta-Analysis of Differentially Private Fine-tuned Language Models (NLP Research) Feb 2023– May 2023

- When applying privacy to Large Language Models (LLMs), accuracy decreases, and computational cost increases.
- We demonstrated that the private finetuning of LLMs using parameter-efficient BitFit can achieve the same accuracy as non-private finetuned LLMs.

Measuring and improving supervision in the clinical learning environment (Medical Research) Jan 2022– May 2023

- No validated research exists on the impact of resident-doctor interactions in medical education using EHR data
- Developed and implemented a statistical learning model to extract meaningful insights from the data, which included interaction patterns, feedback quality, and supervision levels.

Choosing the right experts for BioMedical Question Answering (BioQA) tasks (NLP Research) Sep – Dec 2022

- Build a sparsely gated Mixture of Expert based QA methods aimed to simplify the diversity in biomedical questions.
- Utilized data augmentation and parameter tuning techniques to compete with the current state-of-the-art model.

ML-driven Short Video Streaming (AI Research) Sep – Dec 2022

- Build a Reinforcement Learning (RL) approach to decide which video to download, at what quality, and when to pause.
- Improved scores by 87% - 164% compared to the baselines, using an actor-critic RL algorithm, PPO, with action masking.

RESEARCH PUBLICATIONS

Academic and Wellness Outcomes Associated with use of Anki Spaced Repetition Software in Medical School Journal of Medical Education and Curricular Development https://doi.org/10.1177/23821205231173289

Associated With Health Care Professionals' Choice to Practice in Rural Minnesota. JAMA Netw Open. 2023 https://doi.org/10.1001/jamanetworkopen.2023.10332