Shoaib Ahmed Shaikh

Toronto, ON | Willing to relocate | 437 566 4584

shoowaib95@gmail.com | https://www.linkedin.com/in/shoaib084/ | https://www.cruzai.tech |
https://www.github.com/shoaib646

SUMMARY

- Experienced Machine Learning professional with 5+ years in Financial Services; healthcare; sports and Retail domain. Proficient with Python, ML deployment, NLP, and Agile methodology while indulging as Research Scientist, Python Developer & Machine Learning Engineer & Data Scientist and Testing existing AI Application and developing features.
- Strong quantitative and qualitative skills with ability to communicate technical ideas and make easy-to-understand and actionable solutions.

TECHNICAL SKILLS

- Database   : Relational - My SQL, PostgreSQL, Oracle, MongoDB, NoSQL
- Programming   : Python (Pandas, NumPy, Scikit-Learn, Keras, Pytorch, TensorFlow etc.) VBA, HTML, CSS, Django, Streamlit
- Data visualization : Tableau, Microsoft Power BI, Built-in Libraries (Matplotlib, Seaborn, Plotly)
- Cloud Services   : Databricks, Amazon ec2, Data Factory, Amazon SageMaker, GCP (Virtualization), Docker
- Code Editor   : Jupyter, Google Colab (GPU), PyCharm, Visual Studio Code.
- Additional   : ETL, Tensorboard, Generative AI, CI/CD, Airflow, Dialog flow, DVC pipeline, GIT, Spark, MS excel, MS word

WORK EXPERIENCE

Research Assistant | Durham College - AI HUB (Sports Media - Domain) | Jan 2024 - June 2024
- Gathered and performed cleaning activities on extensive raw datasets containing player statistics, game logs, and other relevant NBA data to develop end-to-end machine learning lifecycle to ensure accuracy and consistency with outlier detection, and feature engineering.
- Identifying key performance indicators (KPIs) and features that significantly impact player performance using domain knowledge and statistical analysis.
- Researched and built fine-tuned LLM application for Developing and implementing machine learning models, leveraging LLM (Large Language Models) to analyze NBA player performance, game strategies, and team compositions.

Machine Learning Engineer | eClerx Services (Domain - Retail) | Nov 2022 - August 2023
- Develop and implement Marketing Mix Modeling (MMM) using Python libraries such as Stats models and Scikit-learn to analyze the impact of different marketing strategies on hotel bookings and revenue, optimizing marketing spend.
- Implement machine learning algorithms such as collaborative filtering and matrix factorization to personalize hotel recommendations based on user preferences and booking history.
- Utilize natural language processing techniques like sentiment analysis and topic modeling to extract insights from customer reviews and feedback, enhancing the recommendation system's understanding of guest preferences and improving service quality.
- Explore deep learning architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to analyze unstructured data like images and text, improving the accuracy of hotel image recognition and guest preference identification.
- Conduct research on transfer learning approaches to leverage pre-trained language models like GPT-3/4 and BERT for fine-tuning on domain-specific datasets, enabling the recommendation system to adapt to evolving guest trends and preferences.
- Employ techniques such as word embeddings and sequence-to-sequence models to generate personalized hotel descriptions or responses, enhancing the guest experience and engagement with the booking system.
- Integrate conversational AI algorithms like Rasa into the system architecture to develop a chatbot capable of engaging in natural language conversations with guests, providing real-time assistance, booking support, and tailored hotel recommendations.

- Utilizing Multi-Touch Attribution (MTA) models to evaluate the contribution of various marketing touchpoints across the guest booking journey using Python libraries. Analyze data to improve marketing strategies and enhance customer acquisition and retention.

Research Scientist | Tata Consultancy Services (Domain – Financial Services) | June 2018 - Nov 2022

- Applied deep learning and machine learning architectures, including transformers, large and small dedicated language models to solve business problems and speed up innovation and PyMC3 and Stan for Bayesian modeling and Hypothesis testing including A/B testing and other statistical methods.
- Utilizing Statistcal techniques to solve complex problems and improve efficiency.
- Implemented Computer Vision solutions for Image processing, masking and Object Detection using Faster RCNN, Yolo and likewise pre-trained models, Image Classification/Segmentation.
- For a Sports Media based client where I was part of the Max-Auto-Optimization team. My primary responsibilities included constructing precise, scalable, and resilient prediction models using extensive datasets. These models were instrumental in real-time bidding and recommendation systems.
- Participated and contributed to sprint planning, estimation, and code reviews, advocating innovative solutions and long-term technical investments that will drive innovation.
- Developed and maintained Directed Acyclic Graphs (DAGs) in Apache Airflow to orchestrate complex workflows for data collection, preprocessing, model training, evaluation, and deployment.
- Collaborated with studying comparative performance between classical machine learning model and quantum computing models such as VQR and VQLS

PROJECTS

Optimal Path Finder @ https://www.cruzai.tech
- Developed a real-time navigation app for Scarborough, Ontario using A Star Algorithm & LSTM model for traffic prediction. Seamlessly deployed on cloud for global accessibility, available on request to demonstrate.

Cartridge Masking @ https://github.com/shoaib646/Cartridge-Masking
- Automated masking of fired ammunition images using a custom algorithm with image segmentation, improving efficiency and accuracy in forensic analysis.

Recommendation System at https://github.com/shoaib646/Recommendation-System
- Deployed personalized course recommendation system using ML & NLP techniques. Integrated with Stream lit for intuitive UI. Continuously refined for enhanced user experience.

LLM Application (On hold) at https://github.com/shoaib646/LLMTesting
- Demonstration to build a LLM Application from scratch an Initiated utilizing Bi-gram concept; paused due to GPU constraints.

EDUCATION

April 2025	Artificial Intelligence & Cloud Computing- PGD	Durham College, Oshawa, Canada
Jan 2021	Information Tech & Systems - MBA	ICFAI School, Tripura, India
April 2017	Information Tech (computer services), bachelor's degree	University of Mumbai - VPM