MOHAMMAD SAAD KHAN

Toronto, Ontario, Canada | (709) 728-3370 | Saadkhan4k@gmail.com | mohammadsaadkhan.cc

PROFILE SUMMARY

- Data Scientist with 5 years of experience and a Master's in Computer Engineering, building machine learning and AI solutions for real-world challenges. Strong background in predictive modeling, statistical analysis, forecasting, and developing production-ready AI systems using Python, SQL, and cloud technologies.

- Experienced in taking projects from initial data exploration to model development, deployment, and performance improvement. I focus on understanding why a problem exists, choosing the right approach, and building solutions that are reliable, efficient, and useful beyond the development stage.

- Actively use AI tools such as ChatGPT, Claude, and Codex in my daily work to accelerate development, test ideas, and explore better approaches. I am particularly interested in Generative AI and continuously experiment with new technologies to understand what they can do and where they can create real value.

TECHNICAL SKILLS

Programming Languages: Python, SQL, R

Machine Learning & AI: Regression, Classification, Deep Learning, NLP, Forecasting, Generative AI, LLMs, RAG, AI Agents, Prompt Engineering, Ensemble Methods

Data Engineering & Analytics: Data Cleaning, Feature Engineering, Hyperparameter Tuning, MLOps, Cross-Validation, Real-Time Data Analytics, API Integration, Time Series Analysis, Descriptive and Inferential Statistics, ETL Pipeline

Data Tools & Libraries: Jupyter Notebook, TensorFlow, Scikit-Learn, PyTorch, Pandas, NumPy, FastAPI, LangChain, LangGraph, Google Colab, Excel

Cloud Platforms & Containerization: AWS, Azure, GCP, Docker, Kubernetes

Big Data Technologies: Hadoop, Apache Hive, Apache Kafka, Apache Spark

Database Technologies: MySQL, PostgreSQL, MongoDB

Data Visualization: Power BI, Tableau, Matplotlib, Seaborn, Plotly

PROFESSIONAL EXPERIENCE

Data Scientist **Navsim Technology | St John's, NL, Canada** **June 2021 – June 2026**

- Owned end-to-end production machine learning systems for maritime navigation, improving pilot accuracy by 20% across low-connectivity environments using scalable Python architectures and optimized numerical computing libraries.
- Deployed production-ready ML models through containerized environments and FastAPI-based inference services, enabling secure prediction serving and autonomous heading estimation that reduced satellite reliance by 35%.
- Implemented LLM and RAG pipelines with prompt engineering for maritime data interpretation and contextual document retrieval, improving response quality and reducing manual analysis effort.
- Developed scalable MLOps infrastructure for monitoring, drift detection, automated retraining, model versioning, validation pipelines, and CI/CD deployment, increasing model adaptability by 20% and supporting long-term production stability.
- Benchmarked Random Forest, Extra Trees, LightGBM, XGBoost, and CatBoost, achieving a 15% performance uplift through ensemble modeling plus 3% through hyperparameter optimization and automated experiment tracking.
- Engineered predictive features using statistical modeling, feature selection, and dimensionality reduction to improve generalization and robustness.
- Optimized training workflows through performance profiling, parallel processing, and algorithmic optimization, delivering a 30% performance gain and improving runtime efficiency by 15%.
- Developed time-series forecasting and probabilistic models to identify operational trends and improve navigation predictions.
- Built distributed data pipelines and ETL workflows that transformed raw maritime telemetry into structured training datasets, reducing processing latency by 40% and enabling faster experimentation and reproducibility.
- Implemented anomaly detection, validation rules, and data-quality monitoring to improve dataset integrity and prediction reliability.
- Applied hypothesis testing, confidence intervals, and controlled A/B testing, improving decision accuracy by 12%.
- Enhanced legacy navigation platforms with AI-driven intelligence layers, increasing system reliability and operational efficiency by 17% while maintaining backward compatibility.
- Built executive-ready dashboards and analytical visualizations using Plotly and Tableau to communicate insights to cross-functional stakeholders.
- Analyzed model failures across different navigation conditions and used the findings to refine evaluation methods, feature design, and deployment decisions.
- Worked with engineering and operations stakeholders to prioritize model improvements based on user impact, technical feasibility, and production needs.

EDUCATION

Master of Science (M.Sc.) in Computer Engineering

Memorial University of Newfoundland | St. John's, NL, Canada | 2018 – 2021

Bachelor of Technology (B.Tech.) in Information Technology

Jamia Hamdard University | New Delhi, India | 2013 – 2017