Karim Nadim

Data Science | Machine Learning & Optimization

Montreal, Quebec
karim_ossama94@hotmail.com
+1 (438) 835-5250
LinkedIn: linkedin.com/in/karim-nadim
GitHub: github.com/karim-nadim
Website: karim-nadim.github.io

Motivation

I am passionate about solving **business problems** using Data Science & Machine Learning. With a Ph.D. in mathematics & industrial engineering, I creatively use my skillset to add **tangible value** to the team, the business, and the end-user. Strong background in statistical analysis, forecasting, and decision support using Python and SQL. I am constantly learning and always looking to improve.

SKILLS & Tools

- **Programming**: Python, SQL, Bash/Shell
- **Tools**: Tableau, Github, MLflow, Docker, DVC, AWS (S3, Lambda, IAM, EC2, SageMaker)
- **Machine Learning**: Scikit-learn, XGBoost, RandomForest, Association Rule Learning, Causal Impact Analysis
- **Deep Learning**: PyTorch, TensorFlow, Keras
- **Data Analysis & Visualization**: Pandas, NumPy, Matplotlib, Seaborn
- **Generative AI**: RAG, LangChain, Prompt Engineering

WORK EXPERIENCE

Data Scientist | CanmetENERGY - Natural Resources Canada | Varennes, QC | Aug 2024 – Present

- Designed and validated a recommendation system using deep learning for industrial process control, leading to a **15% reduction** in energy consumption.
- Built a fault diagnosis model based on Decision Trees and Random Forest in a large-scale industrial system, achieving **93% accuracy** in identifying fault categories and their root causes.
- Developed a forecasting model to predict degradation of industrial components using LSTM, achieving **R^2 up to 0.98**, supporting data-driven operating condition recommendations.
- Developed an indicator to track the health state of industrial equipment based on deep learning autoencoders **with a correlation coefficient of 0.91**, which led to enhanced equipment monitoring.
- Developed a schedule optimizer using reinforcement learning to improve production planning, achieving **10% reduction** in operational costs under realistic simulation scenarios.
- For enhancing information retrieval from technical industrial reports, I developed an agent using LangChain with RAG, reducing manual review time for engineering teams by **60%**.
- Prepared and engineered large-scale industrial datasets using SQL and Python (Pandas), enabling reproducible modeling and experimentation workflows.

Research Assistant | CanmetENERGY - Natural Resources Canada | Varennes, QC | Jan 2020 – Apr 2024

- Developed a predictive maintenance model for industrial equipment using LSTM networks (PyTorch), demonstrating **30% reduction** in downtime in simulation-based evaluations.
- Built ensemble regression model to forecast industrial system behavior with up **to 98% accuracy**, supporting scenario analysis and production planning decisions.
- Conducted exploratory data analysis, feature engineering, and statistical validation on large-scale time-series datasets, improving model interpretability and robustness.
- Optimized model hyperparameters using Optuna and tracked experiments with MLflow to ensure reproducibility and transparent comparison of modeling approaches.
- Applied dimensionality reduction (PCA, t-SNE) to relate process variables to economic and operational KPIs.
- Collaborated with engineers and non-technical stakeholders to translate modeling results into actionable insights and recommendations.

Projects

Measuring Campaign Effectiveness via Causal Inference

- Quantified the incremental sales uplift from a customer subscription program by applying causal impact analysis to transactional data, constructing a counterfactual using non-subscribing customers and demonstrating a 41.1% statistically significant increase in post-campaign spend (95% confidence).

Visual Similarity Search for Product Discovery

- Built an image similarity search pipeline using deep learning-based feature embeddings and cosine similarity to retrieve visually similar products, addressing customer search friction and enabling data-driven product discovery.

LLM-Powered SQL Agent for Self-Service Business Analytics

- Designed and implemented a natural-language SQL agent using LangChain that translates plain-English business questions into validated PostgreSQL queries, enabling self-serve analytics across customer and transaction data while enforcing safe query execution and SQL best practices.

EDUCATION

- **Doctorate in Industrial Engineering | Ecole Polytechnique de Montréal | Canada | Aug 2018 – Apr 2024**
- **Bachelor's in Mechatronics Engineering | German University in Cairo | Egypt | Jan 2012 – Jan 2017**