# Mohd Maaz Khan

**Adm. No.** 24JE0560 +91 8864934918
kmaaz7256@gmail.com linkedin.com/in/maaz-khan github.com/maaz7409

## Education

**Indian Institute of Technology (Indian School of Mines) Dhanbad** Jul 2024 - May 2028 (expected)
*Bachelor of Technology, Chemical Engineering* *Dhanbad, Jharkhand*

## Experience

**Research Intern** Oct 2025 – Present
*Prof. Saurav Sengupta, Department of Chemical Engineering, IIT (ISM) Dhanbad* *Dhanbad, Jharkhand*

- Reconstructed the Lorenz dynamical system utilizing two distinct sparse regression techniques and conducted comparative performance analyses.
- Developed a surrogate model for 1D diffusion processes leveraging Deep Operator Networks (DeepONet).
- Modeled a spring-mass system using Hamiltonian Neural Networks (HNNs), evaluating predictive accuracy and physical consistency against baseline neural network architectures.

## Projects

**Enterprise RAG System with LLM-as-a-Judge** | *Python, FastAPI, React, TypeScript, LangChain, PostgreSQL, Qdrant, Docker*

- Architected an end-to-end enterprise document QA platform using **FastAPI** and **LangChain** to perform low-latency semantic search over unstructured PDFs.
- Implemented a hybrid retrieval pipeline combining **Qdrant** dense vector embeddings (text-embedding-3-small) and **BM25** keyword search to maximize context recall.
- Engineered a secondary **LLM-as-a-Judge** evaluation pipeline using Pydantic structured outputs to evaluate faithfulness, context relevance, and toxicity before user delivery.
- Designed an asynchronous **PostgreSQL** schema to log query trails, source attribution metadata, execution latency, and automated quality evaluation scores.
- Developed a responsive **React** + **TypeScript** frontend with real-time streaming, interactive source citations, and an analytics dashboard built with Recharts.
- Containerized the frontend and backend microservices using **Docker** Compose and configured a GitHub Actions **CI/CD** pipeline for cloud deployment.

**E-Commerce Customer Churn Prediction** | *Python, Scikit-Learn, Imbalanced-Learn, XGBoost, FastAPI, React, Docker*

- Orchestrated an end-to-end data preprocessing pipeline using **Scikit-Learn** and **Imbalanced-Learn**, mitigating noise in real-world data via **KNN imputation** and **Winsorization.**
- Standardized numerical distributions, applied **One-Hot Encoding** for categorical features, and utilized **SMOTE** to resolve extreme class imbalance within the training set to ensure robust model generalization.
- Engineered a high-performance stacking ensemble classifier, leveraging Logistic Regression as a meta-model to optimally combine predictions from **KNN**, **SVM**, **XGBoost**, and **Random Forest** base learners.
- Achieved an exceptional **96% accuracy** and a **0.93 macro F1-score** on a stratified test set, rigorously validating the pipeline to prevent data leakage during scaling and imputation.
- Containerized the machine learning model using **Docker** and developed a highly responsive REST API with **FastAPI**, seamlessly deploying the backend architecture to **Render**.
- Designed and integrated an interactive frontend interface utilizing **React**, deploying the client-side application to Vercel to enable real-time churn prediction for end-users.

**Automated Email Classification Engine** | *Python, Scikit-Learn, Support Vector Machines (SVM), NLP*

- Developed a custom **Scikit-Learn** pipeline to efficiently read, parse, and clean an unstructured, noisy dataset containing thousands of real-world emails.
- Implemented custom **Natural Language Processing (NLP)** operations to generate a comprehensive vocabulary dictionary, applying Count Vectorization to transform raw text into structural feature matrices.
- Trained and optimized a **Support Vector Machine (SVM)** classifier on a stratified training distribution to effectively distinguish and categorize complex email intents.

- Reached a **0.97 macro F1-score** and **97% accuracy** on the test set, ensuring zero data leakage between train and test splits during the custom vectorization phase.

**Inverse Heat Transfer Estimation using Physics-Informed Neural Networks** | *Python, Scientific Machine Learning, PyTorch, SciPy*

- Replicated a cutting-edge scientific machine learning paper, generating a synthetic benchmark dataset by computationally solving 1D heat transfer equations utilizing **Green's functions**.
- Architected a **Physics-Informed Neural Network (PINN)** featuring a custom learning rate scheduler to dynamically decouple and optimize both neural network weights and thermophysical parameters.
- Achieved **99% accuracy** in estimating thermophysical parameters under time-invariant heat flux, and **97% accuracy** for complex time-variant heat flux scenarios.
- Optimized the computational efficiency of the training loop, successfully reducing the parameter convergence time by a **factor of 10x** compared to the original published research.

## Technical Skills

**Languages**: C, C++, CUDA, Rust, Java, Python, TypeScript, JavaScript, SQL, HTML/CSS, R, Julia, MATLAB.
**Technologies & Frameworks**: PyTorch, Scikit-Learn, LangChain/LangGraph, FastAPI, React.js, PostgreSQL, Qdrant, Docker, NumPy, Pandas, SciPy.
**Concepts & Methodologies**: Data Structures & Algorithms, Machine Learning, Deep Learning, Retrieval-Augmented Generation (RAG), LLM Evaluation, MLOps, DevOps, Scientific Machine Learning (PINNs), Statistical Modelling.

## Achievements

- **American Express Strategy Track Challenge :** Placed in the top 10% among 17,000+ competitors from premier Indian institutes. Achieved an accuracy score of 0.882 by developing a quantitative model to rank customer profitability based on premier credit card data.
- **AMS DERIVE (PRIOR) :** Competed in advanced problem-solving challenges, successfully tackling complex quantitative questions encompassing financial modeling, Bayesian statistics, stochastic modeling, and probability theory .

## Social Engagements

**Sports-Engagements**: chess, cricket, table-tennis.