Waiz Khan

(443) 378-2030 | wkhan12@jh.edu | LinkedIn | GitHub | Portfolio

Education

Johns Hopkins University — Baltimore, MD
M.S.E. in Data Science — *Aug. 2024 – May 2026*
- Coursework: Investment Science, Financial Derivatives, Deep Learning, Artificial Intelligence, Statistics

GD Goenka University — Gurugram, India
B.Tech. in Computer Science & Engineering — *Aug. 2019 – Nov. 2023*
- Coursework: Optimization Techniques (Operations Research), Numerical Methods, Probability & Statistics, Linear Algebra, Data Mining & Warehousing, Distributed Systems

Experience

Research Assistant — Dec. 2025 – Present
Johns Hopkins University, Center for Language and Speech Processing — *Baltimore, MD*
Advisor: Dr. Philipp Koehn
- Developing multimodal neural machine translation system for extreme low-resource languages; designing encoder architectures that fuse visual and symbolic representations
- Building preprocessing pipelines for domains with limited parallel corpora and sparse annotation availability

Research Assistant — Aug. 2022 – Nov. 2022
GD Goenka University — *Gurugram, India*
Advisor: Dr. Yogesh Kumar
- Built end-to-end ML, causal inference, and optimization pipeline for educational risk prediction on 47,027 records, achieving 84.1% ROC-AUC with DoWhy causal validation
- Developed ILP allocator that doubled expected impact per budget dollar; stress-tested with 10,000 Monte Carlo simulations and VaR reporting

Data Scientist Intern — May 2022 – Sep. 2022
EBMS Solutions Pvt Ltd — *Chennai, India*
- Engineered churn prediction pipeline using XGBoost with time-series CV; deployed to production CRM with automated retraining; achieved +15% retention lift and AUC 0.82
- Built self-serve analytics layer with curated cohort tables and data quality monitoring; reduced reporting cycle time by 30%

Projects

Black-Scholes Options Pricer | *Python, NumPy, SciPy, Streamlit* — Jan. 2025
- Built production-grade European options pricer with Greeks calculation (Delta, Gamma, Vega, Theta, Rho) and implied volatility solver using Newton-Raphson with Brent fallback
- Implemented arbitrage diagnostics (put-call parity, butterfly spreads, strike monotonicity); 91+ tests, 85% coverage

Monte Carlo Risk Engine | *Python, pandas, Plotly, Streamlit* — Jan. 2025
- Developed portfolio risk simulator computing VaR, CVaR, Sharpe, Sortino, and Max Drawdown using multivariate normal and GBM models
- Applied Ledoit-Wolf covariance shrinkage and walk-forward validation to prevent look-ahead bias; 90%+ test coverage

Technical Skills

Languages: Python, R, SQL, Rust, Swift, Java, C++, JavaScript, Bash
ML/AI: PyTorch, TensorFlow, scikit-learn, XGBoost; CNNs, transformers, NLP; causal inference (DoWhy); time-series CV
Quant/Finance: Black-Scholes, Greeks, VaR/CVaR/ES; Monte Carlo simulation; portfolio optimization; backtesting
Data/Cloud: pandas, NumPy, SciPy; PostgreSQL, Spark; AWS (EC2, S3, SageMaker); Git, Docker; Tableau, Power BI
Security/iOS: Digital forensics, memory/log analysis, incident response; SwiftUI, Core ML, HealthKit