HAMZA ABUBAKAR KHERUWALA

Phone: +1 (716) 235-6752 | *Email:* habubakar89@gmail.com
GitHub: github.com/habubakar89 | *LinkedIn:* *linkedin.com/in/whyser* | *Medium:* *medium.com/@habubakar89*

WORK EXPERIENCE

Morgan Stanley - *SDE II* *Nov 2023 - Present*

- Invented a holistic, real-time behavior analysis engine ***(patent pending)*** analyzing **300K+** daily events to identify + contextualize fraud.
- Built a Kafka + Apache Flink pipeline with **~200ms** latency, enabling real-time, scalable risk evaluation and fraud signal processing.
- Automated stream and batch ETL pipeline provisioning to accelerate data activation, contributing to **$25M+** in annual cost savings.
- Achieved **~70%** fraud reduction in ATO & synthetic identities by leveraging RAGs to surface behavioral context & historical risk signals. Reduced triage time by **55%** and false positives by **50%**, enabling cross-platform behavior correlation across applications.
- Architected a scalable savings platform for high-value client profiling using targeted data flows, built to support **99.99%** uptime.
- Accelerated project launch by **40%** through Terraform-based infrastructure and scalable data orchestration, unlocking **$500M+** AUM potential and **$25M+** in projected annual revenue.
- Automated privacy request workflows, reducing processing time from **20 weeks to a few hours** & improving SLA adherence by **60%.**

Morgan Stanley - *SDE I* *Feb 2023 - Nov 2023*

- Automated modernization of **50M+** lines of legacy code using RAG-enhanced LLMs, achieving **~95%** accuracy and eliminating manual translation across **450+** programs.
- Engineered automated validation pipelines powered by AWS Lambda for orchestration and Aurora for state tracking, cutting code validation time by **87%,** saving **129** person-years, and speeding up modernization by **88%.**
- Delivered real-time data reporting via AWS Kinesis, cutting data latency by **95%,** and improving leadership decision turnaround by **3x,** unlocking **$10M+** in data-driven operational value.
- Streamlined high-volume data access on service APIs, reducing response latency by **25%** and increasing stakeholder adoption by **30%.**

Hult Prize Foundation - *Regional Associate (I) | Accelerator Intern (II)* *July 2021- Sep 2021 (I) | Summer 2021, 2022 (II)*

- Engineered a distributed, fault-tolerant voting system with **99.7%** success rate under high concurrency to handle real-time user surges.
- Optimized concurrent processing for **10000+** active users, reducing response latency by **40%** during peak load periods.
- Streamlined backend workflows and interfaces, cutting setup time by **30%** and accelerating live events with **100K+** attendees.

IoTIoT.in - *Artificial Intelligence Intern* *Jan 2021- May 2021*

- Built a real-time, device-agnostic gesture recognition framework. Improved input reliability by **45%** through advanced motion tracking and signal processing, saving **~$100K** in operational costs annually.
- Engineered parallelized model training, cutting latency by **35%** & achieving **90%+** accuracy, leading to **15%** boost in product usage.

MediaPro Innovations Ltd. - *SDE Intern* *June 2020 - Dec 2020*

- Applied user behavior analysis and ML-driven content filtering, driving a **20%** increase in user retention and accelerating development velocity by **30%** through efficient workflows.
- Improved backend efficiency through caching and indexing, achieving **25%** higher learner engagement and supporting **50,000+** active users with **20%** fewer defects.

SKILLS

Languages & Backend: Python, Java, C++, TypeScript, SQL, React, GraphQL
GenAI & Cloud: RAG (Hybrid. + rerank), agent/workflow orchestration (LangGraph), LLM evaluation & reliability, guardrails/prompt injection defense, vector search, NLP, AWS (IAM, Lambda, Step Functions, Bedrock), Observability, Cost Optimization
Data & Platform: PostgreSQL, MongoDB, PySpark, Terraform, Docker, Observability (Tracing/Telemetry)
Relevent Certifications: AWS Certified Machine Learning Specialty

PUBLICATIONS

Citations: 272 | Source: Google Scholar, January 2026

Comparative Study of Sentiment Analysis and Text Summarization for Commercial Social Networks	*3 cites, July 2020*
BioUAV: Blockchain - envisioned framework for digital identification to secure access in next-gen UAVs	*26 cites, Sep 2020*
Comprehensive review of text-mining applications in finance (Impact Factor: 2.964 - Q1)	*229 cites, Nov 2020*
Interplay of Machine Learning and Software Engineering for Quality Estimations **(Top 10%)**	*16 cites, Nov 2020*
Context-Enriched Machine Learning-Based Approach for Sentiment Analysis	*2 cites, April 2022*

EDUCATION

Master of Science (M.S.) *in Computer Science, University at Buffalo* *Buffalo, New York*
Bachelor of Technology (B. Tech) *in Computer Engineering, Nirma University* *Ahmedabad, India*

PROJECTS

Citation-Grounded Knowledge & Learning Platform (Governed GenAI)

- Built a citation-first Q&A system over authenticated domain texts using multi-stage retrieval and reranking with strict answer gating (“no source → no answer”), with **>95%** citation coverage & reducing unsupported claims by **~70%** in offline evaluations.
- Implemented governed tool access for personalized learning plans, enforcing least-privilege IAM, write-gated actions, and end-to-end audit trails to ensure deterministic, reviewable recommendations.
- Deployed on AWS with control/data plane separation (API Gateway, Lambda, Step Functions, EventBridge, Bedrock), meeting **<2s P95 latency**, enabling replayable traces for compliance, and cutting **~40%** inference costvia model tiering and caching.