

KHAWAJA HUSSAIN AHMED

(702) 348-9820 | khawajaahmed_2025@depauw.edu | Greencastle, Indiana

[GitHub](#) | [LinkedIn](#) | [Portfolio](#)

EDUCATION

DEPAUW UNIVERSITY

Aug 2021 – May 2025

Bachelor of Arts in Computer Science, minor in Data Science

- **Cumulative GPA:** 3.50/4.00
- **Relevant Coursework:** Data Structures & Algorithms, Object Oriented Software Development, Discrete Mathematics, Data Mining, Data Visualization with Tableau, Artificial Intelligence, Statistics for Professionals, Foundations of Advanced Math's, Statistical Modeling.
- **Awards:** Kneale Scholarship, International Merit Grant, Margaret Noble Lee Debate Prize, Dean's List (6-Time Honoree).
- **Leadership:** Lead Resident Assistant, Founder & VP Table Tennis Club.

EXPERIENCE

Data Engineering Intern

Dec 2024 – Feb 2025

Stealth AI

- **Architected** an Azure based **healthcare revenue cycle management data pipeline** processing **10K+** daily records, reducing KPI computation time by **40%** and identifying **\$150K+** in recoverable **revenue**.
- **Established PySpark** validation frameworks that **flagged 5K+** monthly anomalies, **improving downstream analytics** reliability by **35%**.
- **Optimized** cloud infrastructure with **incremental ETL** and **dynamic scaling**, **cutting costs by 20%** while maintaining **99.9% SLA** for **50K+** daily transactions.
- **Resolved** cross-hospital **data inconsistencies** using **SCD2 dimensional modeling**, **increasing reporting accuracy** by **28%**.

AI/ML Engineering Intern

Feb 2024 – May 2024

Depauw University

- **Spearheaded** the development of a reinforcement learning system that **outperformed baseline models by ~11%**.
- **Designed** a memory-efficient training architecture **decreasing** computational requirements by **~20%** while **improving model convergence time by 40%**.
- **Implemented CI/CD pipeline** with automated testing that **reduced deployment time** from days to hours, enabling rapid iteration across 3 development cycles.
- **Delivered critical production fixes** that **resolved performance bottlenecks**, **accelerating** system response time by **15%** under peak loads.

Data Science Intern

Oct 2023 – Jan 2024

Depauw University

- Built an automated sports analytics platform using **Python** and **pandas** that integrated 5+ data sources, delivering insights that **improved team performance by 23%** across 12+ competitions.
- Designed custom **ETL & EDA** workflows **processing 50K+ player statistics** daily, **reducing analysis latency by 80%**.
- **Developed statistical research** for predictive models with **85% accuracy** that identified 7 high-potential recruits.

TECHNICAL PROJECTS

Agtech Agricultural Pipeline

- Engineered a **scalable ETL pipeline** that **ingested** and **normalized** IoT sensor data from **200+** field devices using **Python, SQL, and Apache Airflow**, enabling real-time crop yield predictions.
- Refined **data quality checks** and **anomaly detection** algorithms that **identified 12%** of sensor readings as anomalous, **improving model reliability by 25%**.
- **Optimized database queries** with indexing and partitioning strategies, **reducing average query latency from 5s to 300ms** and supporting near real-time dashboard visualizations.

SKILLS & EXPERTISE

Technical Skills

- **Data Engineering:** SQL, Azure Data Factory, Databricks, ADLS Gen2, Apache Airflow, ETL Pipelines, Data Modeling, Data Governance.
- **Data Science:** Python (pandas, NumPy, scikit-learn, matplotlib), R/RStudio (ggplot2, tidyverse), Statistical Modeling, Hypothesis Testing.
- **Machine Learning:** Reinforcement Learning, Regression Analysis, NLP.
- **Programming:** Python, Java, C++, Swift, C#, Assembly.
- **DevOps & Tools:** Git/GitHub, Linux, Azure Cloud, CI/CD, Excel, Tableau.