

Akanksha Wagh

[LinkedIn](#) | awagh2410@gmail.com | Bellevue, WA

EDUCATION

Master of Science in Data Science

August 2023 - May 2025

Stevens Institute of Technology, NJ, USA | Peer Mentor

Coursework: Optimization for DS, Data Analytics, Applied Machine Learning, Time Series Analysis, Statistical Methods, Statistics, BIA for GenAI.

Bachelor of Technology in Information Technology

August 2019 - June 2023

Pune University, India | Outstanding Student Award 2023

WORK EXPERIENCE

Data Analyst - Marketing | Compass Group, USA

February 2024 – May 2025

- Designed and deployed production-grade machine learning models and **ETL pipelines** using **Python, SQL, and Natural Language Processing (NLP)** to monitor social media sentiment and campaign performance, improving content engagement by 20% through automated data transformation and data integration.
- Applied **statistical analysis, A/B testing, and Star Schema** modeling to optimize marketing datasets, boosting student engagement by **15%** and retention by **25%** via sentiment analysis of customer feedback.
- Built real-time dashboards and KPI trackers using **Azure Synapse Analytics, Power BI, and SQL**, driving a **25%** increase in lead conversion through scalable CRM insights and refined audience targeting.

Machine Learning Intern | Stevens Institute of Technology, USA

May 2024 – August 2024

- Built, deployed, and integrated** a context-aware **chatbot** for the **IT Department** using **GPT-4, LangChain, and Python**, connected via **ETL pipelines** to enterprise data sources, reducing support ticket volume by **20% monthly** and clearly articulating results.
- Designed and implemented automated **context-switching workflows, real-time monitoring dashboards, and data models** to track chatbot interactions, achieving a **30% improvement** in resolution accuracy and enabling **production-grade performance monitoring**.
- Conducted **QA testing and drift detection** on model responses, enhancing chatbot performance based on **real-time feedback loops**.

Data Engineer Intern | Tata Motors Limited, India

January 2023 - June 2023

- Engineered **data pipelines** (Python, PySpark) to centralize IT support desk structured/ unstructured data in **PostgreSQL**, conducted **root-cause analysis of ticket trends** to uncover recurring issues, and automated workflows to cut down manual intervention by **40%**.
- Designed and built an **employee portal dashboard (Figma, Django, Power BI)** with **REST API integration and advanced analytics** to track activity trends and IT ticket metrics, improving self-service access, **boosting operational efficiency by 28%**.

Freelance – Data Engineer & Analytics | India

June 2021 – December 2022

- Built a **BigQuery-based data warehouse** with **ETL pipelines** across **GCP and AWS S3**, consolidating **12M+ transactions** from six systems for a finance client, cutting reporting latency by **45%** and ensuring regulatory compliance.
- Developed **MLOps pipelines on Azure Machine Learning** for predictive patient readmission models, achieving **92% accuracy** and reducing readmission rates by **18%** through automated retraining and deployment workflows using high-dimensional data.
- Created **interactive dashboards** using **Qlik** and mined **8TB+ clickstream data** for an e-commerce client, uncovering key behavior patterns and boosting campaign ROI by 27%, lead conversion by 32%, and decision-making speed by 40%.

RELEVANT PROJECTS & PUBLICATIONS

Self-Generated Soccer Highlights [[GitHub](#)]: Researched, developed, and implemented a **deep learning model** to summarize soccer matches, achieving **95% event detection accuracy** and improving **highlight-generation precision by 25%**. [[Published in IEEE](#)]

Grader Bot [[GitHub](#)]: Developed a GenAI-powered grading assistant using **GPT-4o, LangChain, Python, OCR, and Streamlit**.

Forecasting NJ Transit Bus On-Time Performance [[GitHub](#)]: Built time series forecasting models using **ARIMA, SARIMA, R, ACF/PACF analysis, and AIC/BIC evaluation**, improving prediction accuracy by **18%** and uncovering seasonal and post-COVID transit trends.

AI-Powered Microsoft Stock Forecasting: Developed hybrid LSTM, ARIMA, and LLM forecasting models integrating sentiment analysis.

Business Intelligence Dashboard for Campus Vendor Selection [[GitHub](#)]: Evaluated student feedback from a tasting event hosted by Zaro's Bakery, a potential replacement for the existing vendor on campus.

My YouTube Watch History Analysis [[GitHub](#)]: Analyzed personal YouTube watch history data using **Python and Streamlit**.

TECHNICAL SKILLS

NLP, MLOps, Predictive & Time Series Modeling, Scikit-Learn, TensorFlow, PyTorch, Data Warehousing, Data Mining, BigQuery, Tableau.