Akanksha Wagh

USA | awagh2410@gmail.com | LinkedIn | Github

SUMMARY

Data Scientist with 3+ years of experience building Machine Learning, NLP, and predictive analytics solutions using Python, SQL, Scikit-Learn, PyTorch, TensorFlow, and cloud platforms (Azure, AWS, GCP). Skilled in end-to-end model development—data collection, feature engineering, model training, tuning, deployment, and monitoring.

EDUCATION

Master of Science in Data Science August 2023 - May 2025
Stevens Institute of Technology
Bachelor of Technology in Information Technology August 2019 - June 2023
Pune University – JSPM College of Engineering

WORK EXPERIENCE

Business Analyst, Compass Group LLC February 2024 - May 2025
Tech: *Python, SQL, Power BI, Azure Synapse, Scikit-Learn, NLP, Cloud Analytics*

- Integrated multi-source CRM, campaign, survey, and feedback data using **Python + SQL,** reducing fragmentation by **60%** and accelerating **reporting** by **40%**—creating unified datasets for **predictive modeling**.
- Conducted **sentiment analysis** on **1,000+** feedback entries; generated topic clusters improving customer experience **decision-making**.
- Visualized **KPIs** and **operational trends** through **analytics dashboards**, boosting cross-team decision-making speed by 40%.
- Performed **statistical analysis** on engagement and demand trends to support predictive **insights** across **marketing and operations**.

Data Science Intern, Stevens Institute of Technology May 2024 - August 2024
Tech: *Python, PySpark, LangChain, GPT-4, LLMs, NLP, MLflow*

- Designed, trained, and optimized **ML / NLP models** on large-scale conversational data and built automated monitoring pipelines—including drift detection, performance tracking, and scheduled retraining workflows—improving resolution accuracy by **30%**.
- Launched real-time dashboards for **conversation analytics**, tracking metrics like intent accuracy, escalation rate, and response quality.
- Conducted **QA** evaluation of chatbot outputs and implemented feedback-based improvement loops to enhance production performance.

IT Data Analyst, Tata Motors Limited January 2023 - June 2023
Tech: *PySpark, Python, PostgreSQL, Django REST, Power BI, ML Analytics*

- Engineered scalable ETL pipelines for **unstructured** IT support data, reducing manual effort by **40%**, and enabling real-time analytics.
- Performed **statistical modeling** and **trend forecasting** to detect recurring support issues, to improve predictive maintenance planning.
- Built analytics dashboards highlighting anomalies, issue patterns, and ticket trends, enhancing operational planning efficiency by 25%.

Data Engineer & Analytics | Freelance June 2021 - December 2022
Tech: *AWS S3, GCP, Azure ML, Python, Qlik, Power BI*

- Integrated **cloud-based data warehouses** and automated **ETL pipelines** consolidating 12M+ transactions for a finance client.
- Developed **predictive readmission models**, achieving 92% accuracy and reducing readmissions via automated **retraining workflows.**
- **Analyzed 8TB+** clickstream to uncover behavioral insights that boosted campaign ROI, **conversions rate** and **decision-making** speed.

PROJECTS & PUBLICATIONS

Time Series Forecasting for NJ Transit On-Time Performance (**Tech:** *R, ARIMA, SARIMA, ACF/PACF, AIC/BIC)*
Implemented forecasting models improving prediction accuracy by **18%**, uncovering seasonal + post-COVID transit behavior patterns.

Deep Learning for Soccer Highlight Generation (**Tech:** *GANs, Data Augmentation, PyTorch, OpenCV, CNNs, Computer Vision)*
Built an event detection model with **95% accuracy**, improving highlight-generation precision by **25%**.[Published in IEEE]

Business Intelligence for Vendor Evaluation and Selection (**Tech:** *Power BI, Python)*
Analyzed student feedback using sentiment & trend analysis; generated dashboards for decision-making in vendor evaluation and selection.

Intelligent Vehicle Tracking System for Campus Transportation (**Tech:** *Python, OpenCV, OCR, Computer Vision, Pandas*)
Created an Intelligent Vehicle Tracking System tracking system to automate bus logging and optimize campus transportation operations.

YouTube Watch History Analysis (**Tech:** *Python, Streamlit)*
Designed and created an interactive dashboard analyzing personal YouTube watch patterns, content preferences, and time-of-day usage.

TECHNICAL SKILLS

Programming & Data Systems: Python, SQL, ETL Pipelines, Airflow, BigQuery, AWS S3, Azure Synapse, GCP, REST APIs
Machine Learning & Analytics: Predictive Modeling, LLM, Feature Engineering, Model Evaluation & Optimization, A/B Testing
Frameworks, MLOps & Visualization: Scikit-Learn, TensorFlow, PyTorch, PySpark, MLflow, Docker, Power BI, Streamlit, Qlik, Tableau