AAYUSH MALLIK

774-622-0835 • aayushmallik.contact@gmail.com • http://www.linkedin.com/in/aayush-mallik • Portfolio

SUMMARY

Data analyst with expertise in machine learning, predictive modeling, and data-driven decision-making. Proficient in Python, SQL, and cloud platforms, with hands-on experience in big data analytics, A/B testing, dashboards, and ETL automation.

EDUCATION

M.S. in Data Analytics | GPA: 3.98/4.00 | *Clark University, USA* **May 2025**

Honors: Member, Alpha Epsilon Lambda Honor Society (Leadership, Character, Academic Excellence)

B. Tech in Computer Science and Engineering | GPA: 8.45/10.00 | *Vellore Institute of Technology, India* **Jul 2020**

PROFESSIONAL EXPERIENCE

Quantitative Research Analyst Intern | *Cutler Capital Management LLC, USA* **Mar 2025 – Present**

- Developed and automated data processing pipelines using the SISM neural network model for text extraction and summarization, reducing manual report generation time by 60% and enhancing financial insights.
- Constructed and optimized predictive models (XGBoost, Random Forest) for trend analysis, leveraging distributed data processing (Dask, Spark) and advanced feature engineering to enhance forecasting accuracy and model interpretability.
- Designed interactive dashboards and automated reporting systems using SQL, Python, and Power BI, streamlining real-time analytics for stakeholders by developing custom visualizations to track key trends and enhance reporting efficiency.

Research Assistant — Strategic Analytics & Institutional Research | *Clark University, USA* **Sep 2024 – Mar 2025**

- Engineered large-scale survey data processing, refactoring Python code with Dask and Parquet, reducing execution time from 20+ minutes to 3 minutes, enabling real-time analytics and data-driven decision-making.
- Optimized survey data using statistical techniques, A/B testing, and Tableau dashboards, generating data-driven insights to enhance product analytics and decision-making, improving KPI-driven strategies by 23% at the leadership level.

Machine Learning (ML) Engineer | *Cedar Gate Technologies, Nepal* **Aug 2020 – Mar 2022**

- Built and optimized scalable AWS-based data pipelines, automating ETL workflows with Apache Airflow, reducing processing time by 30%, and integrating real-time data streaming with Kafka for low-latency decision-making.
- Deployed predictive ML models for patient length-of-stay forecasting, leveraging feature engineering, hyperparameter tuning (Optuna, GridSearchCV), and ensembling, achieving 70% accuracy while integrating SHAP for model transparency.
- Streamlined ML deployment on AWS (SageMaker, Lambda, EC2, S3) and Azure (Azure ML, Databricks) with CI/CD pipelines (Docker, Kubernetes), integrating TFX for automated model retraining and deployment, improving scalability.

TECHNICAL SKILLS

Programming & Frameworks: Python, R, Excel, Pandas, TensorFlow, PyTorch, Keras, Dask, NLP.
Cloud & Big Data Platforms: AWS, Azure Databricks, Apache Spark, Hadoop, Google Cloud Platform (GCP), Snowflake.
Database Management: Microsoft SQL Server, MySQL, PostgreSQL, Oracle, MongoDB, NoSQL.
Machine Learning & AI: Predictive Analytics, Advanced Statistical Modeling (ANOVA, Hypothesis Testing), ETL Processes.

PROJECTS

Automated AI-Powered Recruitment System | USA

- Curated and implemented an end-to-end recruitment system leveraging spaCy and TfidfVectorizer to parse and evaluate resumes, enhancing hiring efficiency by 25%, and cutting manual evaluation time by 40% for HR teams.
- Enhanced transformer-based NLP for job-candidate matching by fine-tuning SentenceTransformer with domain-specific embeddings, improving match precision by 18% and refining recommendations through contextual similarity analysis.
- Integrated RLHF into a coding practice system for technical interviews, enabling adaptive learning and real-time feedback, improving problem-solving efficiency by 16% and enhancing learning adaptability by 34%.

ML-Powered Risk Assessment for Terrorism Trends | USA

- Implemented multi-class classification using LGBMClassifier and Neural Networks to predict attack types, achieving 87.89% average accuracy with F1 scores of 0.96 (training) and 0.87 (testing) with LGBMClassifier.
- Optimized model stability reducing class imbalance impact and improving training distribution, while increasing test accuracy by 12% through hyperparameter tuning with GridSearchCV and Repeated Stratified K-Fold Cross-Validation.