Mahmud Hasan



Data Engineer | DE and ML Pipelines · Geospatial Analytics · Experimentation

Dhaka, Bangladesh linkedin.com/in/mahmud-buet15 github.com/Mahmud-Buet15

Data engineer, 5 years, specialising in the pipelines and production models behind real-time mobility. At Pathao — Bangladesh's largest ride-hailing and delivery platform — I own the Airflow and dbt pipelines that turn 0.6M trip events and GPS pings a day into the features driving ETA prediction across multiple cities and multiple countries, and the MLflow tracking, monitoring and alerting that keep those models honest in production. Raised on-time ETA prediction by up to 38pp.

EXPERIENCE

Pathao — Dhaka, Bangladesh

Bangladesh's largest on-demand digital platform (ride-hailing, food, parcel, courier)

Data Engineer III — Jan 2025 – Present

Data Engineer I — Oct 2023 – Dec 2024

ETA Prediction (production ML, millions of rides/month)

- Owned ETA model development across multiple cities and multiple countries, raising the share of rides predicted within ±30% of actual arrival — biggest gains +38pp on Chittagong bike and +24pp on Dhaka bike. Models serve 0.6M rides a month.
- Built automated data-generation and training pipelines in Airflow with MLflow experiment tracking, removing manual retraining effort and keeping model quality consistent across releases.
- Launched a centralized ETA monitoring dashboard plus automated degradation alerting, shifting the team from reactive ad-hoc investigation to proactive detection.
- Enhanced the speed-feeding mechanism in Open Source Routing Machine (OSRM) to improve routing accuracy feeding downstream ETA predictions.
- Led the initial end-to-end audit of the ETA process, mapping workflows, isolating the largest error sources, and converting findings into a prioritized roadmap adopted across teams.

Rastamapo — Trip Distance Measurement

- Rebuilt trip-distance reconstruction from raw driver GPS pings, raising route reliability by up to 34pp and automated fare-dispute resolution by up to 35pp — reducing manual customer-support intervention on distance disputes by making reconstructed routes trustworthy enough to arbitrate fares automatically.
- Mentored one engineer through the project end to end — set the technical approach, reviewed all code, and remained accountable for delivery. He now owns the service.

Parcel Batching

- Designed and built a parcel batching simulator to model multi-parcel dispatch before any production rollout.
- Identified optimal batching configurations across order-volume scenarios, giving product and ops calibrated expectations for supply efficiency and driver earnings.

Popular Pickup Points

- Generated popular pickup recommendations for Bangladesh and Nepal from historical trip data to cut driver navigation and wait time.
- Measured user impact through A/B testing and automated the generation and refresh pipeline.

Fraud Detection

- Analyzed ride and food transaction patterns to surface fraudulent users and drivers, feeding confirmed cases to stakeholders for enforcement.
- Developed detection rules targeting rides exploiting large incentive payouts.
- Developed an analytical dashboard giving an overview of fraudulent activity across verticals.

Business Analytics

- Produced cross-vertical performance analyses presented to the CEO and executive team on business health and improvement opportunities.
- Built targeted high-potential lead lists from multi-source data and partnered with Marketing to streamline onboarding.

Progoti Systems Ltd. Dhaka, Bangladesh

Fintech; TallyKhata and TallyPay digital financial products

Jr. Data Scientist — Dec 2021 – Oct 2023

Trainee Jr. Data Scientist — Sep 2021 – Nov 2021

Sugar Project (national subsidy disbursement program)

- Built Python ETL ingesting from Google Sheets, production databases, and CSVs into the data warehouse with cleaning and validation layers.
- Wrote PL/pgSQL analysis scripts powering stakeholder reports and dashboard datasets.
- Automated disbursement data sharing with partner banks via Python and shell, replacing a manual process.
- Ran reconciliation across sugar mills and maintained the program dashboard in Google Data Studio.

Analytics & BI

- Built and maintained six dashboards: Customer Journey Management (TallyPay), Investor, HR, Super QR, Survey Data, and Sugar.
- Ran cohort, funnel, and campaign-effectiveness analyses; performed root-cause investigation on dashboard anomalies.
- Produced user segmentation models and investor audit reporting.

SKILLS

Languages & Query	Python · SQL · PL/pgSQL · Shell scripting
ML & Data Science	Statistical analysis · A/B testing · ML model development · ML model training automation using MLflow
Data & ML Engineering Platform	Airflow · MLflow · Docker · Kubernetes · Terraform · Helm · dbt · Spark · Kafka · ETL/ELT pipeline design · Data modeling
Cloud	Google Cloud Platform (BigQuery and supporting GCP services)
Databases	PostgreSQL · MySQL · Google BigQuery
Visualization	Google Data Studio · Redash · Apache Superset · Power BI
Geospatial	OSRM · GPS trace processing · Routing systems
Version Control and Workflow automation	GitHub · GitLab · CI/CD

EDUCATION

B.Sc. in Naval Architecture and Marine Engineering (NAME) 2016 – 2021

Bangladesh University of Engineering and Technology (BUET)

CERTIFICATIONS

- Applied Data Science II: Machine Learning & Statistical Analysis *(with honors)*
- Applied Data Science I: Scientific Computing & Python *(with honors)*
- Databases and SQL for Data Science with Python *(with honors)*