

Mohit Kaushik

DATA SCIENTIST - Python, Machine Learning & SQL

✉ mohit.coop2024@gmail.com

☎ (+1) 250-885-7280

📍 British Columbia

in [LinkedIn](#)

📁 [Portfolio](#)

SKILLS

- **Programming & Databases:** Python, SQL, ETL pipelines, event tracking, data modeling, cloud (AWS, GCP).
- **Machine Learning Models:** Regression, classification, clustering, neural nets, bandits, time series, model deploy.
- **Statistical Techniques:** A/B testing, causal inference, hypothesis testing, experiment design, statistical significance.
- **Analytics & Visualization:** Power BI, Tableau, Amplitude, Excel – real-time dashboards, KPIs, trend and cohort.

WORK EXPERIENCE

Data Analyst

December 2024 – Present

Aterio.io

British Columbia

- Developed robust ETL pipelines using SQL, Python, Airflow, and dbt in GCP BigQuery; directed pipeline latency by 28%, and boosted reporting refresh rates by 35% for 20+ real-time dashboards using Pub/Sub, Git, and Composer.
- Parsed unstructured data using BeautifulSoup, LangChain, OpenAI API, JSON transformation, and Pandas; uncovered 3 new customer segments, driving 25% improvement in campaign ROI using spaCy, Regex, Vertex AI, and YAML.
- Integrated REST APIs with FastAPI, Docker, and Airflow DAGs to build ML-ready datasets; decreased model deployment cycle by 10% and proposed data reliability for predictive scoring engines using CI/CD, Marshmallow, Redis.
- Designed KPI dashboards using Tableau, Looker Studio, and SQL views; enabled product teams to track usage trends and lifted weekly retention visibility by 40% via custom filters, data blending, BigLake, LookML, and Sheets connector.

Data Analyst Intern

September 2022 – December 2022

Maruti Suzuki India Ltd

India

- Analyzed over 1M rows using SQL, scikit-learn, K-means clustering, and linear regression; helped drive 12% growth in state-wise sales by optimizing pricing and inventory allocations with SMOTE, seaborn, joblib, T-SQL, and Numpy.
- Created Power BI dashboards using DAX, Power Query, RLS, bookmarks, and KPIs; improved decision speed by 22% and reduced manual reporting efforts by 18% using Datamarts, Paginated Reports, Drillthrough, KPI Cards, Azure.
- Built Python web crawlers using Requests, BeautifulSoup, Pandas, CronJobs, and SQLite; captured and cleaned 10K+ data points weekly, supporting 8% uplift in lead targeting effectiveness using XML parsing, Airbyte, Faker, Pandera.

Research Analyst Intern

May 2022 – August 2022

Vellore Institute of Technology (VIT)

India

- Engineered Alzheimer's classifier using logistic regression, scikit-learn, stratified K-fold, and ROC-AUC metrics; achieved 85% accuracy and improved recall by 13% over baseline using SHAP, LightGBM, LabelEncoder, and XGBoost.
- Preprocessed MRI scans using OpenCV, NumPy, SciPy, skimage, histogram equalization, grayscale thresholding, and array masking; enhanced image clarity by 85%, reducing training noise via dilation, CLAHE, TensorFlow, and imgaug.
- Reduced model complexity by 30% using ANOVA, PCA, variance thresholding, z-score scaling, and multicollinearity checks; accelerated training time by 40% without compromising F1 score using StatsModels, Yellowbrick, RobustScaler.

PROJECT EXPERIENCE

Netflix Data Analysis – Content Optimization

Data Analyst

- Liaised SQL to analyze engagement & segment users, boosting content visibility and retention by flagging weak genres.

Retail Store Sales Data Analysis – Revenue Improvement

Data Analyst

- Drove 7% revenue growth by linking sales KPIs to external trends and visualizing behavior shifts via Excel dashboards.

Bank Customer Churn Prediction

Data Scientist

- Reduced churn by 10% using a classification model and dashboards highlighting churn drivers and behavior trends.

Air Passenger Traffic Forecasting

Forecast Analyst

- Forecasted five-year air traffic with ARIMA (RMSE: 0.140), revealing growth trends via time series decomposition.

Medical Insurance Cost Prediction

Data Scientist

- Built a regression model (RMSE: 10,270) using MICE and polynomial features, identifying smoking as a key cost driver.

EDUCATION

Master of Engineering in Applied Data Science

September 2023 – October 2024

University of Victoria, British Columbia

Bachelor of Technology in Computer Science and Engineering

July 2019 - May 2023

Vellore Institute of Technology, India

CERTIFICATIONS

- **Databases and SQL for Data Science with Python**, IBM – Coursera