

Karush Pradhan

Data Scientist

Data Scientist with hands-on experience in Python and SQL, specializing in time-series analysis, ML modeling (XGBoost), and dashboard development with Streamlit. Skilled in building end-to-end data pipelines, extracting and transforming raw datasets, and deploying predictive models using Google Cloud Functions. Passionate about applying data science to real-world systems like HVAC control, energy analytics, and automation. Background in accounting supports sharp business insight and KPI-focused analysis.

Work Experience

| | |
|--|---|
| Data Scientist | Menteru Inc - Tokyo, Japan Oct 2024 - Present |
| Architected and maintained modular data pipelines in Python to process large-scale time-series sensor data. Utilized Google Cloud services including Dataflow, Pub/Sub, and Cloud Storage for scalable stream and batch workflows. | |
| Performed data extraction and transformation from Cloud SQL (MySQL/PostgreSQL) and BigQuery, creating model-ready datasets through complex joins, aggregations, and feature engineering. | |
| Developed and optimized XGBoost models in Vertex AI for HVAC performance simulation, focusing on energy efficiency, thermal comfort, and productivity. Conducted feature selection, outlier handling, and evaluation using metrics such as RMSE and residual analysis. | |
| Built interactive dashboards in Looker Studio and Streamlit to visualize KPIs, system behavior, and environmental metrics, enabling clear communication of data insights to stakeholders. | |
| Automated real-time device actuation through REST API integrations with building systems (e.g., SwitchBot) using Cloud Functions for dynamic, serverless control. | |
| Collaborated with cross-functional teams across cloud infrastructure, data engineering, and business units to ensure technical solutions aligned with operational objectives. | |

Selected Data Science Projects

| | |
|--|------------------------------------|
| Resume Chatbot | AI & Full Stack Application |
| github.com/karushp/resume-chatbot https://karushp.github.io/resume-chatbot/ | |
| An AI-powered chatbot that answers questions about my resume using RAG (Retrieval-Augmented Generation) | |
| Tools: Python, Flask, FAISS, Cohere API, Groq API, HTML/CSS/JavaScript, Github Pages | |
| EMS Energy Consumption Dashboard | Energy Management System Dashboard |
| github.com/karushp/ems-dashboard | |
| An interactive Streamlit dashboard for analyzing energy consumption data across Kansai and Kanto regions in Japan, featuring dynamic filtering and comprehensive visualizations. | |
| Tools: Python, Streamlit, Pandas, Plotly, Parquet | |
| YEN-USD Analysis and Prediction Model | Data analysis and Prediction model |
| github.com/karushp/yen-usd-analysis | |
| Exploratory analysis of YEN to USD and prediction using Random Forest Classifier | |
| Tools: Python, Scikit-learn, Pandas, Seaborn, Jupyter | |
| Heart Disease Prediction | Data analysis and Prediction model |
| github.com/karushp/heart-disease-predict | |
| This project predicts heart disease, optimizing K-Nearest Neighbors through scaling, normalization, and grid search. | |
| Tools: Python, Scikit-learn, Jupyter Notebook, Pandas, Seaborn | |
| Safety Helmet Detection | Computer Vision Model |
| Real-time detection ML Model identifying instances of safety helmet usage. | |
| Tools: YOLO, OpeanCV, Python, Streamlit | |

Previous Work Experience – (2014–2024)

Accounting & Finance | Sydney, Australia & Kathmandu, Nepal
Conducted financial analysis, budgeting, and reporting across healthcare, aesthetics, and public sector organizations. Reconciled and extracted large volumes of transactional data to identify trends, anomalies, and process improvements. Managed collections, credit evaluation, aged debt, and bad debt forecasting across multi-client environments. Led process audits, mentored junior staff, and contributed to internal tool and system development for AR/AP efficiency.

| | | | |
|------------------------------|---------------------------------|-------------------|---------------------|
| Organizations: | | | |
| LifeHealthcare | Accounts Receivable Coordinator | Sydney, Australia | Jul 2023 – May 2024 |
| Allergan Aesthetics | Accounts Receivable Coordinator | Sydney, Australia | Jul 2022 – Jul 2023 |
| Medlab Pathology | Accounts Receivable | Sydney, Australia | Jun 2021 – Jul 2022 |
| Mastercare Australasia | National Credit Manager | Sydney, Australia | Mar 2019 – Mar 2021 |
| Bajra Technologies Pvt. Ltd. | Account Manager | Kathmandu, Nepal | Jul 2014 – Oct 2015 |

Contact

karushpradhan@gmail.com
Matsudo, Chiba 271-0046

Socials

linkedin.com/in/karushpradhan
github.com/karushp

Education

Le Wagon Tokyo
Data Science and AI Bootcamp
2024

Macquarie University, Syndey (Australia)
Master in Accounting (CPA Extension)
2018

Christ University, Bangalore (India)
Bachelor of Business Management
specialized in Marketing
2014

Skills

Programming & Tools: Python, SQL, Pandas, NumPy, Jupyter, Git, GCP (Cloud Functions)

Machine Learning: Scikit-learn, XGBoost, KNN, Random Forest, Time-series forecasting, Hyperparameter tuning, Cross-validation, Model evaluation (RMSE, R², residuals), Wellness-based control logic

Visualization: Streamlit, Matplotlib, Seaborn, Looker Studio, Plotly, Custom KPI dashboards, Multi-tab UI design

Cloud & Automation: Google Cloud Functions, Pub/Sub, Cloud Dataflow, BigQuery, SwitchBot API, REST APIs, GCP IAM & roles

Data Engineering & ETL: Data preprocessing, Feature engineering, Outlier detection, Interpolation, Resampling, Time alignment, Multi-sensor data handling, Class-based pipeline design

Languages

English – Fluent
Nepali – Native
Hindi – Fluent

Additional

Japanese Driver's License
(普通自動車免許) valid until 2028