Aayushma Pant

aayushmapant50@gmail.com | 0494357414 | LinkedIn | Website | Medium | Sydney, NSW, Australia

Summary

AI and Data Engineer with 4+ years of experience building statistical and machine learning models, scalable Python and SQL pipelines, and production ML systems. Prepared and validated 70+ million records, tuned predictive models with Optuna, and delivered ML workflows with Azure, CI/CD and monitoring. **Published at WACV 2026 and ECML PKDD 2026.** Recipient of the **AI in Mining Award, Women in AI AUS + APAC 2026**.

Technical Skills

MLOps & Data Engineering: Databricks, MLflow, DVC, Apache Airflow, Docker, Docker Compose, Kubernetes, CI/CD (GitHub Actions)

Programming & Model Evaluation: Python (NumPy, Pandas, scikit-learn), SQL, FastAPI, REST APIs, C/C++, statistical modelling, time-series forecasting (ARIMA), controlled experiments, Optuna

Cloud Platforms: Azure AI, Azure OpenAI, Azure Cognitive Services, Azure AI Search, Azure Container Apps, Microsoft Entra ID, AWS (EC2, S3, Lambda, API Gateway, ECR), Google Cloud (Vertex AI, Cloud Build, Cloud Run)

Computer Vision & Deep Learning: PyTorch, OpenCV, TensorFlow, Keras, anomaly detection, image segmentation, spatial-spectral feature fusion, 3D reconstruction, Structure-from-Motion

Content Understanding: CNN classifiers, audio and video feature extraction (DCT, MFCC), OCR, image segmentation (SAM), OpenAI API, Claude API, Claude Code, Copilot Studio, Power Automate, LLM-based document classification and verification, Vision Transformers (ViTs), VLMs, LLMs, hyperspectral and multispectral model pre-training, efficient training

Geospatial & Data Formats: GeoPandas, GDAL, Rasterio, GIS data processing, spatial joins, remote sensing data, GeoTIFF, shapefiles

Tools & Methodologies: Git/GitHub, JIRA, Agile/Scrum & Kanban project management, Grafana, Notion, PostGIS, QGIS

Experience

AI & Data Engineer May 2026 – Present
Earth AI *Sydney, Australia*

- **Large-Scale Data Pipelines:** Built reusable Python, SQL and DuckDB pipelines to clean and validate 70+ million geological assay records for analysis and modelling.
- **Geospatial Predictive Modelling:** Developed and evaluated mineral-targeting models across 7+ million national grid cells using XGBoost and LightGBM, using Optuna for hyperparameter optimisation and MLflow to track experiments.
- **Spatial Feature Engineering:** Built feature-engineering and spatial-labelling workflows for 3D geochemical modelling across 30+ million drillcore records, using XGBoost.
- **Model Integration:** Integrated an anomaly detection model into an internal FastAPI application, adding interpretable maps and saved results for later review.
- **Workflow Automation:** Developed and tested AI workflows for task tracking and information retrieval, enabling hourly monitoring and automated updates.
- **Investment Series Demos:** Built AI demonstrations for Series C fundraising to explain how the technology worked.

AI Engineer Jun 2025 – May 2026
HyperGen *Geelong, Australia*

- **Document Intelligence and Food Compliance:** Designed and deployed a client-facing visual document verification application for Australian food compliance checks. Used SAM for segmentation, Mistral OCR to extract text from PDFs, and vision-language models (VLMs) for document classification and quality checks.
- **Production Deployment:** Deployed container artifacts to Azure Container Apps with SQL, full-text indexing and scalable workers. Built GitHub Actions CI/CD workflows and implemented Microsoft Entra ID access controls and audit logging.
- **Client Feedback and Testing:** Updated document verification rules based on client feedback and reran checks after each change. Used Databricks for testing and experimentation.
- **CRM and AI Agent Platform:** Designed and deployed a multi-user web app for timesheet and opportunity tracking, with AI agents for CRM workflows and secure access via Cloudflare Tunnel. Integrated Grafana dashboards for opportunity, product and profitability data and shared them with stakeholders.
- **Employee Workflow Automation:** Built Copilot Studio agents and Power Automate flows to automate employees' routine tasks.

Computer Vision Research Engineer Aug 2021 – Dec 2024
Angelswing Inc. *Seoul, South Korea*

- **Automated Inspection System:** Designed and deployed an AI-powered inspection system for aerial survey data, improving defect detection accuracy by 60%. Operationalised with Docker and CI/CD pipelines, enabling automated retraining and monitoring.
- **Geospatial ML Data Pipeline:** Designed and deployed an end-to-end ETL pipeline to automatically detect and process aerial Ground Control Points (GCPs) from raw geospatial survey data, reducing manual data processing by 65%. Applied MLOps best practices (MLflow, DVC, CI/CD) for smooth deployment, validation, and reproducibility.
- **Aerial Image Stitching and Data Processing:** Built a real-time aerial image stitching and transformation pipeline, achieving 50% faster processing and 40% reduced memory usage by using computer vision algorithms, GIS data, and mathematical optimisation, producing clean, standardised geo-referenced outputs from heterogeneous raw imagery

- **3D Mapping & Reconstruction:** Developed an MVP data pipeline for 3D mesh generation by integrating computer graphics with Structure-from-Motion techniques, enabling richer geospatial data products from raw aerial datasets
- **3D Visualisation Tool:** Created a prototype UAV imagery visualisation system using Three.js and 3D vision techniques, resulting in a 25% increase in user engagement compared to previous 2D tools.
- **Team Leadership:** Led a research initiative coordinating a distributed team of engineers, improving project development efficiency by 20% through agile collaboration (90% of quarterly OKR targets met).

Education

Master of Information Technology (Research) Thesis submitted
Deakin University, Science, Engineering & Built Environment *Melbourne, VIC, Australia*
- **Research Focus:** Hyperspectral Anomaly Detection using AI/Data Science methods (Computer Vision & ML) and Research Assistantship on CEC prediction via machine learning.
- **Scholarship:** Awarded a 100% full scholarship for the duration of the program in recognition of academic merit.
- **Graduate Teaching Assistant (Mar 2025 – May 2026):** Taught tutorial sessions on database concepts and SQL programming, graded assignments, provided one-on-one mentoring, and hosted coding workshops to reinforce key concepts.

Bachelor of Electronics and Communication Engineering Nov 2017 – Apr 2022
Tribhuvan University, Institute of Engineering *Thapathali, Nepal*
- Achieved First Class Honours. College Topper of the batch and three-time semester topper (Academic Excellence Awards).
- Committee member, Robotics and Automation Centre. Organised workshops on AVR/Arduino, KiCad, and AI.

Publications and Research

Noise-Robust Band Selection via Explainability-Constrained Learning for Hyperspectral Anomaly Detection
European Conference on Machine Learning (ECML PKDD 2026), Springer LNCS vol. 16950, pp. 53–70
- Proposed End-to-End Explainability-Constrained Learning (E2E-ECL), which jointly optimises band selection and anomaly detection using an explainability consistency loss. Maintained 89–93% noise rejection and $\geq$0.96 AUC under four noise types, and reduced 200+ spectral bands to 20 while retaining $\geq$98% of full-band accuracy across nine benchmarks.

DMS2F-HAD: A Dual-branch Mamba-based Spatial-Spectral Fusion Network for Hyperspectral Anomaly Detection
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2026)
- Developed and evaluated a dual-branch Mamba architecture that fuses spatial and spectral features for hyperspectral anomaly detection.

Deepfake Audio and Video Detection System
Research project, Tribhuvan University
- Co-developed CNN classifiers for deepfake detection. Combined video frames with discrete cosine transform (DCT) features for video classification, and used filter-bank features and mel-frequency cepstral coefficients (MFCCs) to classify authentic and synthetic audio.

Time-Series Forecasting: Stock Market Prediction
Personal project *Python, statsmodels*
- Built an ARIMA stock-price forecaster with rolling one-step-ahead evaluation and measured forecast error using mean squared error.

Nepali Voice-Controlled Home Automation System
Electronics & IoT
- Built a voice-controlled home automation system integrating speech recognition with microcontroller hardware.

Awards

AI in Mining Award, Women in AI AUS + APAC 2026.

Women in STEM Mentor Recognition: Awarded for volunteering and promoting tech education among young women.

Academic Excellence Awards Semester topper

Certifications

Deep Learning Specialization (Deeplearning.ai); Machine Learning (Coursera, Stanford University); Generative AI with LLMs (Deeplearning.ai); Data Analytics Internship (KPMG)

Volunteering

STEM Education for Girls (Smart Cheli), Lounge International Student Centre