

Pragyan Pahadi

+977 9764022836 | pahadipragyan@gmail.com | [GitHub](#)

PROFESSIONAL SUMMARY

Software Engineering student specializing in Data Science, Machine Learning, and empirical data pipelines. Strong foundation in Python, deep learning frameworks (PyTorch), and exploratory data analysis. Passionate about leveraging machine learning models, statistical processing, and analytical tools to extract actionable insights and engineer scalable AI-driven solutions.

EDUCATION

Nepal College of Information Technology (NCIT) <i>Bachelor of Engineering in Software Engineering (8th Semester) SGPA: 3.3/4</i>	Lalitpur, Nepal 2022 - Ongoing
Caspian Valley HSS <i>High School (11th & 12th Standard)</i>	Lalitpur, Nepal Graduated
Budhanilkantha School (BNKS) & Bhagawati HSS <i>Primary and Lower Secondary Education</i>	Kathmandu & Sindhuli, Nepal Graduated

TECHNICAL SKILLS

Languages: Python, SQL, C/C++, TypeScript, JavaScript
Machine Learning & Data Science: PyTorch, Pandas, NumPy, Matplotlib, Grad-CAM, RAG Pipelines
Tools & Infrastructure: Jupyter Notebook, Git/GitHub, Docker, Dialogflow (NLP)
Frameworks: Node.js, React, Next.js

PROJECTS

Agentic Framework for Alzheimer's Detection <i>Python, PyTorch, OpenCV, Grad-CAM</i>	Ongoing
- Co-developed a multimodal clinical decision support system combining Convolutional Neural Networks (CNNs) for visual feature extraction and a Retrieval-Augmented Generation (RAG) pipeline. - Spearheaded Vision Model Engineering by preprocessing raw Optical Coherence Tomography (OCT) images and training a baseline CNN to classify Alzheimer's biomarkers. - Implemented Grad-CAM algorithms to generate anatomically grounded saliency maps, prioritizing diagnostic transparency and explainability. - Evaluated model performance using rigorous machine learning metrics, including F1-scores, precision, recall, and AUC-ROC.	
Data Analytics & Physical Simulations <i>Python, Pandas, Jupyter Notebook, Matplotlib</i>	Jan. 2025 – May 2025
NYC Public Schools SAT Analysis: Developed multivariate data aggregations and engineered borough-level metrics to highlight socio-educational trend anomalies and score deviations. Netflix Content Mining: Designed advanced boolean filtering queries to analyze massive catalog matrices, identifying historical shifts and underlying patterns in production data.	
NLP Integration & Server Pipelines (Furever) <i>TypeScript, Dialogflow, Node-cron</i>	July 2025
- Integrated a custom Dialogflow NLP Agent to establish a context-aware interface that handles interactive, real-time user inquiries and intent classification. - Engineered automated server-side data pipelines using <code>node-cron</code> for asynchronous database updates and automated task scheduling.	
Cloud-Native & Network Data Architectures <i>Docker, C, Socket API</i>	May 2025 – June 2025
- Deployed real-time data ecosystems within a cloud-native environment using Docker containerization (Cricksy). - Engineered low-level networked client-server systems optimizing custom UDP data transmission and explicit IP addressing protocols.	

RESEARCH EXPERIENCE

Multimodal Agentic Framework for Alzheimer's Detection <i>PyTorch, LangGraph, RAG</i>	Ongoing
- Co-authoring an ongoing research study on an interpretable clinical decision support system utilizing OCT images and patient metadata. - Trained a baseline CNN (ResNet) for visual feature extraction and implemented Grad-CAM to generate anatomically grounded saliency maps of retinal layers. - Built a LangGraph-driven Retrieval-Augmented Generation (RAG) pipeline to fuse visual vectors with structured metadata, synthesizing context-aware diagnostic reports.	

EXPERIENCE

Data Science Fellow

Nepalese Society of Student Researchers (NSSR) × Sunway Student Representative Council (SSRC)

Jan. 2025 – Present

Kathmandu, Nepal

- Selected for a competitive data fellowship program centered on applied Python-based Data Science, exploratory analysis, and statistical processing.
- Built and benchmarked end-to-end analytical data pipelines involving deep data manipulation, data cleaning, and programmatic feature engineering.

CERTIFICATIONS & CONTINUOUS LEARNING

Technical Certifications & Training

DataCamp & Kaggle

Ongoing

Online

- **DataCamp Certifications:** Successfully completed certified coursework in **Introduction to Python**, **Intermediate Python**, and **Data Manipulation with pandas**.
- **Kaggle Certifications:** Certified in **Intro to Programming** and **Python**.
- **Continuous Learning:** Maintained active daily learning streaks (140+ days on Kaggle, 2+ months on DataCamp), demonstrating sustained technical discipline in data analytics.