

AARAV SUBEDI

📍 Kathmandu, Nepal ✉ aaravsubedi123@gmail.com ☎ +977 9860608631
🌐 aaravsubedi.com.np 🐙 github.com/subedi-aarav 🔗 linkedin/aarav-subedi

INTRODUCTION

I am an enthusiastic undergraduate pursuing B.Tech in Artificial Intelligence at Kathmandu University. I am deeply committed to harnessing AI for practical solutions and community empowerment. My academic journey covers Machine Learning, Natural Language Processing, Data Science, Computer Vision, Linear Algebra, Probability & Statistics, and Optimization, enriched by hands-on research in deep learning and multilingual NLP. My strengths lie in problem-solving, collaboration, and communication, driven by a passion for AI's transformative impact across industries.

EDUCATION

B.Tech in Artificial Intelligence Kathmandu University, Dhulikhel, Nepal	2022 – 2027 (Ongoing)
+2 Science Lord Buddha Secondary School, Biratnagar, Nepal	2019 – 2021 GPA: 3.43 / 4.00
Secondary Education (SEE) Elixir Academy, Kathmandu, Nepal	2019 GPA: 3.80 / 4.00

EXPERIENCE

KUAIC — AI Canvas Coordinator <i>Kathmandu University AI Club, Dhulikhel, Nepal</i>	2025 – Present
- Coordinated AI Canvas, a module under AI Convergence featuring Datacamp sessions, Kaggle Competitions, Foundational Models workshops, and IoT Bootcamp. - Managed planning, execution, and cross-team coordination across multi-stage technical events with 100+ participants.	

PROJECTS

SCARLET, Vision-Language Model for CS Research Figure Understanding <i>Supervised by Prof. Dr. Sudan Jha, Kathmandu University</i>	2026 – Ongoing
- Fine-tuning Qwen2.5-VL-3B-Instruct on 330,000+ CS research paper figures from arXiv using QLoRA (4-bit, rank 8) across a two-stage pipeline: domain SFT on diagram QA pairs followed by reasoning SFT on synthetic CoT traces. - Targets consumer-deployable inference (<8GB VRAM) with planned open release of dataset, model checkpoint, and prototype tools for figure summarisation, diagram QA, and visual similarity search.	
Multi-Language News Aggregation & Summarization <i>Supervised by Mr. Sandeep Gupta, Kathmandu University</i>	2025
- Built a multilingual news platform aggregating 450+ daily articles using mBART50 for Nepali–English abstractive summarization and translation, evaluated via a 50-participant user study. - Integrated XLM-RoBERTa for Nepali news classification (F1: 0.90) and sentiment analysis (F1: 0.89) with semantic retrieval and diversity reranking using Qdrant.	
REACO, Face Recognition App for Photographers <i>github.com/REACO</i>	2024
Built an end-to-end face recognition system to automate sorting and grouping of photographer image collections by individual.	

- Trained YOLOv8m on 32,000 images for face detection; applied Dlib landmark-based face alignment followed by a Siamese network with triplet loss (InceptionV3) for embedding generation; used KNN and SVM for recognition.

Parikar, Nepali Cuisine Classifier

2023

github.com/parikar

- Developed a mobile app to classify 12 types of Nepali cuisine using transfer learning with YOLOv8, InceptionV3, and VGG16; achieved over 90% accuracy across all three models.
- Converted models to TensorFlow Lite for on-device inference; deployed via Flutter with real-time camera-based classification.

RESEARCH

LipiNet, Multi-Script Handwritten Character Recognition

2025 – Ongoing

- Designed LipiNet, a unified architecture for handwritten character recognition across fifteen benchmarks spanning eight writing systems, introducing **Persistent Scaffold Injection (PSI)**, a script-specific stroke signal injected at every encoder stage as a structural prior.
- Achieved state-of-the-art **99.75% on DHCD** (Devanagari, 46 classes) with 3.9M parameters, surpassing ResNet-85 (99.72%, ~39M) and MallaNet (99.71%, 17M) at $10\times$ and $4.4\times$ greater parameter efficiency respectively.
- Demonstrated cross-script generalisation across Thai (99.83%), Persian (99.72%), Bengali (98.90%), and six other writing systems with a single shared architecture skeleton.

TigerMove, Deep RL Agents for Baghchal (*Accepted, ICCMSO Conference — In Press*)

2024 – 2025

- Conducted the first systematic deep RL benchmark on Baghchal, the traditional asymmetric Nepali board game, comparing DQN, REINFORCE, PPO, and MuZero across both player roles over 100,000 self-play episodes.
- MuZero achieved the strongest results (86% Tiger win rate, 62% Goat) via MCTS-based planning; PPO offered the best performance-to-compute trade-off at a fraction of MuZero's 24.5-hour training cost.
- Identified role-based learning asymmetry and emergent strategic behaviors including pincer movements, sacrifice plays, and tempo control across model-based and model-free paradigms.

CERTIFICATIONS

Introduction to Generative AI – IBM SkillsBuild

Issued July 2026

Credential URL: IBM SkillsBuild

Machine Learning for Data Science – IBM SkillsBuild

Issued Mar 2025

Credential URL: IBM SkillsBuild

Learn to Program: The Fundamentals – Python, University of Toronto

Issued May 2022

Credential URL: 2HQ856VYEP6Q

TECHNOLOGIES

AI / ML

Programming

Frameworks & Dev

DevOps & Cloud

Databases

Design & Productivity

Soft Skills

PyTorch, TensorFlow, Computer Vision, NLP, Transformers, API, Algorithms

Python (Advanced), C/C++ (Intermediate), SQL, PHP, HTML/CSS, Bash

Django, Django REST Framework, Flutter, Streamlit

Docker, Kubernetes, MinIO, Azure, Git/GitHub

MongoDB, MySQL, SQLite, OpenMediaVault, NextCloud

Figma, Microsoft Office (PowerPoint, Excel)

Leadership, Communication, Teamwork, Problem Solving, Presentation