

• **E-mail** : hamaljyotsan@gmail.com

• **Contact no** : +9779824595613

Experience	<div><div>Intern at Subiz Innovations as Python Automation Engineer</div><div><ul style="list-style-type: none">Jun 2023 - Nov 2023 · 6 months<ul style="list-style-type: none">Data Extraction : Made a python bot which extracts the Nepal stock exchange stocks data into a csv file.OCR : Used a PaddlePaddle OCR which converts Bank statement images, pdf into excel sheet.Data Cleaning : Extracted data from different sites which i was able to clean for ML model training</div><div><div>ML Engineer at Palm Mind Technology</div><div><ul style="list-style-type: none">Dec 2024 - Present<ul style="list-style-type: none">Production-level Chatbot : Built and deployed chatbots for real-time contextual responses, with features like booking, lead post, recommendation based on real-time location.Real-time Web Scraping Pipeline : Engineered a live data extraction system using socket connections.OpenAI & Twilio Integration : Integrated OpenAI APIs with Twilio to enable real-time user interaction and automationCG Motors Chatbot Revamp & Hyundai Deployment : Migrated Hyundai chatbot from LangChain to LangGraph, integrated SQLite for memory, and optimized the RAG pipeline. Deployed and maintaining with improved performance and scalability.</div></div></div>
Projects	<div><div><div>Phone Call Agent: Developed a real-time phone call agent using Twilio and OpenAI APIs. The agent can book, reschedule, and cancel appointments through natural conversation, leveraging WebSocket for real-time communication. Github : Link</div><div>Libraries used:<ul style="list-style-type: none">WebSocket • FastAPI • Twilio</div></div><div><div>Build Llama.cpp for Custom GPU Architecture: Built and configured llama.cpp for custom GPU architectures.Compiled llama.cpp for GTX 1650 Ti Mobile and RTX 3060 (via Salad), enabling support for efficient inference.Successfully served the gemma-3 4B GGUF multi-model setup.Github : Link</div><div>Libraries used :<ul style="list-style-type: none">Llama.cpp • Cmake • CUDA Toolkit • Multer • REST API</div></div><div><div>PPOCRv5 Fine-tuning Backend: Built a FastAPI backend to receive images and PDFs from a frontend dashboard, enabling users to easily fine-tune a PPOCRv5 model. The backend handles data preprocessing, configuration, and training, streamlining the fine-tuning process for document OCR tasks. Github : Link</div><div>Libraries used:<ul style="list-style-type: none">Pytorch • FastAPI • PaddlePaddle • Datasets</div></div><div><div>TempleVision : Developed TempleVision as part of my Final Year Project to enhance image colorization of black-and-white temple photos.Fine-tuned the DeOldify generative model using a dataset of 2,500 colorful temple images from Kathmandu, Nepal.Improved the model’s ability to produce vibrant, culturally rich colorizations that preserve the essence of traditional temples. Github : Link</div><div>Libraries used:<ul style="list-style-type: none">base64 • Pytorch • FastAPI</div></div></div>
Training	<div><div>AI Fellowship</div><div><ul style="list-style-type: none">Fusemachines Apr - Oct 2024<ul style="list-style-type: none">Engaged in a comprehensive AI fellowship focused on advanced machinelearning, deep learning, and computer visionApplying technologies like face detection, gaze tracking, and head pose detection in Fellowship projects.</div></div>
Education	<div><div><div>Lincoln University, Phoenix College of IT & Management Maitidevi, Kathmandu 2022-2025</div><div>Bachelor of Information Technology • BIT</div></div><div><div>Hope International College Karkando , NPJ 2019-2021</div><div>High School</div></div></div>