Samir Wagle

Satdobato, Lalitpur | waglesameer5@gmail.com | +977-9840032620 | samirwagle.com.np
samir-wagle-349591142 | SamirWagle

Introduction

AI and NLP Engineer focused on deep learning, multilingual language technologies, and data-driven problem solving. My work spans research, model development, dataset creation, and applied machine learning, particularly in low-resource language settings. In addition to my research experience, I have contributed to a broad range of engineering projects including system-level software, cross-platform development, and device integration—which strengthens my ability to design reliable, real-world AI solutions. I am passionate about building ethical, impactful, and scalable technologies that bridge the gap between research and practical deployment.

Education

Kathmandu University School of Engineering *Dec 2020 – June 2025*
BE in Computer Engineering

- GPA: 3.72/4.0

Capital Secondary School (CCRC) *Aug 2018 – Nov 2020*
+2/ Science

- CGPA: 3.54/4.0
- **Coursework:** Physics, Chemistry, Mathematics, English, Nepali, Computer Science

Adarsha Vidya Mandir Madhyamik Vidyalaya *2008 – 2018*
Secondary Education Examinition

- CGPA: 3.85/4.0

Experience

Research and Development (Software Engineering) *Kamaladi, Kathmandu*
CityTech Private Limited *March 2025 – Present*

- Worked on embedded C programming for the POS payment terminal powered by the Asino Q161 Pro microcontroller.
- Designed and implemented user interface features including image rendering, audio playback, and interactive screens for the POS terminal.
- Developed and tested TCP socket communication modules for real-time data exchange, including sending/receiving JSON data to/from backend servers.
- Worked on Android Based POS Device, Migrated the entire code base from Android 10 to Android 14 resolving compatibility issues, updating libraries,SDK's and system level components and ensuring full compatibility with new OS environment
- Integrated and stabilized MQTT-based communication on Android 14 using a modified client library due to lack of OS support for the existing implementation.
- Diagnosed and fixed major hardware integration issues including card reader, NFC, and PIN pad failures by collaborating with the manufacturer, updating SDKs, and modifying device-level code.
- Performed extensive bug fixing, optimization, and system verification to ensure a stable, production-ready POS workflow.
- Redesigned the UI for a 240×320 px POS display, improving usability on a small form-factor device.

Undergraduate Research Intern *Dhulikhel, Kavre*
Information and Language Processing Lab, Kathmandu University *Aug 2023 – Present*

- Developed deep learning models for sentiment classification and created an English Text dataset by scraping YouTube comments for Sentiment Analysis

- Published a Research Paper titled " Profanity and Offensiveness Detection in Nepali Social Media Using Bi-directional LSTM Models " at 21st International Conference of Natural Language Processing(ICON 2024), MIT Campus of Anna University
- Contributing to the development of Trilingual Machine Translation System (English-Nepali-Tamang) in Society-Centered AI Research Program by Google

Co Founder / Technology and Innovation Co Ordinator *Dhulikhel, Kavre*
Impact Nexus Nepal (Non Governmental Organization) *October 2024- Current*

- Administrate and manage the Microsoft Intra Admin Center, ensuring secure and efficient access management, identity governance, and compliance.
- Oversee user access controls, authentication policies, and security protocols while providing technical support and streamlining IT operations to enhance infrastructure and improve organizational efficiency.
- Conduct extensive online research to identify and apply for nonprofit services, grants, and support programs by analyzing eligibility criteria, preparing the necessary documentation, and streamlining application processes to maximize opportunities for nonprofit growth and sustainability.
- Conduct online research to identify nonprofit services, grants, and funding opportunities, while analyzing eligibility criteria, preparing documentation, and writing compelling proposals to secure financial support and streamline application processes for nonprofit initiatives.

Research and Publications

Profanity and Offensiveness Detection in Nepali Language Using Bi-directional LSTM Models Dec 2024

Proceedings of the 21st International Conference on Natural Language Processing (ICON)

Click Here for Full Paper

Evaluating Sentence Embedding Models for Nepali Sentiment Analysis: A Comparative Study Aug 2025

Accepted on National Conference on Computer Innovation 2025. Conference on Aguust 24 2025.

Retrieval-Augmented Generation Framework for the Nepali Legal Domain Question Answering Aug 2025

Submitted in SIGIR-AP 2025 Conference.

Exploring the Ethical Implications of Clickbait Practices in Nepal June 2024

Analyzed the ethical impact of clickbait in Nepal's digital media through surveys and interviews. Findings emphasize the need for media literacy, ethical guidelines, and responsible content creation.

Click Here for Full Paper

Certification and Achievements

Machine Learning Fundamentals

- Basic Fundamentals of Machine Learning, Data Preprocessing, Machine Learning Algorithms, Tools and Libraries, Evaluation Metrics and Deep Learning

Postman API Fundamentals

- Use of Postman API,How API Works, API Testing and Automation, Creating and Managing API Requests, API Documentations, and Postman Monitors

Exploratory Data Analysis (Machine Learning Fundamental Boot camp)

- Data Cleaning and Preprocessing, Descriptive Statistics, Data Visualization, Feature Selection, Trend Analysis and Data Summary.

Event Leadership and Co Ordination

- KU HackFest 2022. Digital First Hybrid Hackathon (Discord Server management, Bot Integration,Server Management, Event Reporting to MLH)
- KU Hackfest 2023. (Pre event and Post Event Management,Social Lead, Content Creation, Social Media Management, SEO Analyst)
- Voluntered at TedX Kathmandu University 2024.

Projects

NepSAUL *NepSAUL, GitHub*

- Developed the first Retrieval-Augmented Generation (RAG) system for Nepali legal question answering using over 10,000+ case laws from the Nepal Kanun Patrika archive.
- Implemented and compared sparse (BM25) and dense (multilingual-e5-large) retrieval using FAISS indexing, optimizing for low-resource legal data environments.
- Achieved 91percent Precision@1 with BM25 on chunked data and 75percent with e5-large embeddings on dense retrieval.
- Assessed generation quality via LLM-as-Judge for truthfulness (85) and groundedness (74) percent, with a manual evaluation accuracy of 84percent by a legal expert.

  Demonstrated the viability of AI-powered legal QA in low-resource languages, setting a foundation for future Nepali legal LLM applications.
- Tools Used: Python, Tensorflow, Keras, Kaggle, Natural Language Processing, BM25 Retriever, PineCone

Nepsense *NepSense Github*

- A Nepali language NLP tool for detecting offensive content and sentiment analysis, designed with a Streamlit interface for accessibility.
- Tools Used: Python, Tensorflow, Keras, Google Colab, Natural Language Processing

Comment Sense *CommentSense Github*

- An NLP-based tool that classifies YouTube comment sentiments using machine learning and provides a React-based interface for moderation.
- Tools Used: Python, Tensorflow, Google Colab, Kaggle.

Mapping the UnMapped *Github*

- Developed a geospatial hashing algorithm tailored to Nepal's geography, integrating district codes, city types, and nearest landmarks. Implemented BASE64-encoded Vigenère cipher encryption for secure and compact location encoding. It is our hackathon Project.
- Tools Used: Python, Collab

Water fountain Simulation using PyOpenGL and PyGame *Github*

- Developed a real-time water fountain simulation using PyOpenGL and PyGame, implementing particle physics for realistic water movement and rendering. Optimized performance for smooth animation and interactive visualization.
- Tools Used: Python, PyOpenGL, PyGame

Contribution to Profanity and Offensive Words Dataset Collection *Github*

- Conducted data collection and analysis of Nepali profanity and offensive language to support NLP applications, content moderation, and sentiment analysis. Assisted in curating and preprocessing linguistic datasets for AI and machine learning models.
- Tools Used: Collab, Python, Flask, Selenium

Heart Beat Simulation *Github*

- Developed a heartbeat simulation system to monitor heart rate using embedded systems. The system simulates real-time heart rhythms and provides visual feedback for medical applications, demonstrating proficiency in real-time monitoring and embedded system integration.
- Tools Used: Circuit Simulation, Proteus, Embedded System Technologie

Job Hunt – Job Search Platform (C++ and Qt) *Github*

- Developed a desktop application connecting employers with job seekers, enabling job postings and applications. Implemented backend functionality to demonstrate job search platform operations using C++ and Qt.
- Tools Used: QT, C++

Skills

Tech Skills . C++, C, JavaScript, Python, Node, React, Machine Learning, Deep Learning, Figma, Data Science, Flask

Interests: Deep Learning, Content Creation, Gaming, Photography, Event Meetup

Soft Skills: Adaptability, Time Management, Problem Solving, Communication, Creativity.