

Supriya Khadka

 Supriya090 |  supriyakhadka |  supriyakhadka.com.np |
 khadkasupriya76@gmail.com |  +977 9861446632

Summary

Postgraduate Researcher at Research Centre for Computational Science and Mathematical Modeling, Coventry University, working on lowering the cost of healthcare by clinical coding with LLMs.

Education

Master by Science in Research, AI

Sept 2024 - Sept 2025 (Expected)

Coventry University, United Kingdom

Supported by British Council Women in STEM Scholarship

Bachelors in Computer Engineering

Nov 2018 - Apr 2023

Pulchowk Campus, Tribhuvan University, Nepal

(Agg. 85.60%)

Professional Experience

Postgraduate Researcher, Coventry University

September 2024 - **now**

Project Title: Lowering the Cost of Healthcare by Clinical Coding with LLMs

Supervisors: Dr. Xiaorui Jiang, Dr. Pengpeng Hu, Prof. Vasile Palide

- Investigating generalization of biomedical pre-trained language models in clinical coding.
- Researching Large Language Models zero-shot and few-shot setups for explainability of clinical codes.
- Analyzing performance and cost-effectiveness of different few-shot learning strategies for adapting pre-trained and large language models to the clinical coding task.

Conversational AI Engineer, Diyo.AI

May 2023 - Aug 2024

- Building **multilingual chatbots** tailored for domains such as sexual and maternal reproductive health, safe abortion, e-governance, and sales.
- Researching and implementing **large language models** like *GPT3.5*, *Llama2*, *BLOOM* and *Okapi* to develop and improve conversational models and their language understanding capability for **low resource languages** like Nepali, under the supervision of Dr. Bishesh Khanal, in collaboration with NAAMII as a Research Assistant.
- Curating, refining, and enhancing NLP datasets for low-resource languages to develop language technologies for Nepali and other languages of Nepal, as a continuation of my [Internship at NAAMII](#)
- Successfully pitched Diyo.AI at **ICT Startup Awards 2023**, securing a spot in the **top 5** finalists.

Computer Vision Research Engineer, Rara Labs (Contract)

Jan 2023 - Mar 2023

- Worked on the discovery phase of implementing *AI for NWash*, a project initiated by The Ministry of Water Supply, Nepal in collaboration with the British Embassy, Frontier Tech and Rara Labs.
- Successfully implemented object detection and recognition algorithms on the dataset collected through the NWash Project, seamlessly integrating them with a web application.
- Conducted extensive research on computer vision algorithms like YOLO and R-CNN to identify the most suitable solutions based on the NWash dataset requirements.

Publications

Khadka, Supriya, and Bijayan Bhattarai. “Gender Bias in Nepali-English Machine Translation: A Comparison of LLMs and Existing MT Systems”. *OSF*, 2024, (Preprint). <https://doi.org/10.31219/osf.io/69und>.

Khadka, Supriya, Ranju G.C., Prabin Paudel, Rahul Shah, and Basanta Joshi. “Nepali Text-to-Speech Synthesis using Tacotron2 for Melspectrogram Generation”. *Proc. 2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL 2023)*. 2023, pp. 73–77, <https://doi.org/10.21437/SIGUL.2023-16>.

Paudel, Prabin, Ranju G.C., Rahul Shah, and **Supriya Khadka**. “Decentralized Book Publishing Platform using Blockchain and IPFS: An NFT-based Solution to Royalty Theft”. *International Conference on Engineering & Technology*. 2022.

Internships

NLP Research Intern, NAAMII

Oct 2022 - Apr 2023

- Researched the work done in Nepali NLP (Text Summarization, Nepali Transliteration and Nepali Text-to-Speech) under the supervision of **Dr. Bishesh Khanal**.
- Curated Nepali NLP datasets with the study of their features and limitations (Sentiment Analysis, Nepali Text-to-Speech and Machine Translation).

Research and Development Intern, CARD

Jul 2022 - May 2023

- Researched on **Nepali Text-to-Speech Synthesis** under the supervision of [Dr. Basanta Joshi](#) at the **Center for Applied Research and Development** at the Institute of Engineering, Pulchowk Campus.
- Prepared a comprehensive Text-to-Speech dataset for the Nepali language and made it publicly available on [OpenSLR](#).
- Conducted numerous experiments with multiple models and speech synthesis techniques, and integrated the best result into a [mobile application](#).
- Published our work at the 2nd Annual Meeting of the Special Interest Group on Under-resourced Languages(SIGUL2023), a Satellite Workshop of Interspeech 2023 ([Khadka et al.](#)).

Teaching Experience

Technical Lead and Facilitator, Women Leaders in Technology

Sept 2020 - Nov 2023

- Served as a Technical Facilitator for three consecutive batches of a fellowship program aimed at empowering women leaders in technology.
- Instructed **over 50 fellows** in essential technical subjects, including Git, Basic JavaScript concepts, Asynchronous JavaScript concepts, Express, and MongoDB, during the fellowship bootcamp.
- Promoted to the Technical Lead in the latest batch, overseeing all technical sessions and projects, and teaching approximately half of the classes throughout the duration of the fellowship.

Awards and Honors

British Council Women in STEM Scholarship	2024
Ncell Excellence Scholarship - Yearwise Top Student (4 times winner)	2019 - 2023
GritFeat AI Health Hackathon - 1st Runner Up	January 2023
Hack-A-Week Winner - Category Fintech	February 2022
3rd Annual Nepal AI School - Full Scholarship	December 2021
Grace Hopper Celebration Student Scholar	July 2021

Selected Projects

Shruti - A Nepali Book Reader

[Link to Repo](#)

- Mobile application featuring integrated Nepali Text-to-Speech for audiobook functionality.
- Built using Tacotron-2 architecture (for spectrogram generation) and Hifi-GAN and WaveGlow (for vocoder), all of which implemented in PyTorch.

NepaliXlit

[Link to Repo](#)

- Adaptation of IndicXlit for Nepali systems. Prepared a parallel corpus of 3500 romanized-nepali words, incorporating the nuances of Nepali conversational styles
- Finetuned using fairseq with an incremental approach, finally obtaining Top@5 accuracy of 86% in the third iteration and deployed using FastAPI, reducing the inference time to less than a second.

NepaliSense DistillBERT

[Link to Repo](#)

- Implemented a Nepali variant of DistillBERT for sentence embeddings, drawing inspiration from the paper *Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation*.
- Utilized a student-teacher model requiring a parallel Nepali-English corpus for training. The necessary data was sourced from the TED2020 corpus.

Arthur - The Gladiator

[Link to Repo](#)

- A game featuring a knight rescuing princesses through three distinct levels, incorporating maze-solving with Dijkstra's algorithm, puzzle-solving with BFS and DFS, and puzzle sorting using Quick Sort and Merge Sort.
- Built the game in C++ with SFML as Graphics Library.

Technical Skills

- **Python** including libraries useful for machine learning and deep learning like **Numpy**, **Pandas**, **Scikit-Learn**, **Pytorch**
- Front end programming using **JavaScript** and its frameworks like **React**
- Server side coding using **Node.js** and **Express** framework
- SQL databases like **PostgreSQL** and **MySQL** and NoSQL database like **MongoDB**
- Version control systems like **Git**

Volunteer Experiences

Planning Coordinator, GitHub Field Day Nepal

Apr 2022 - Aug 2023

- Worked as the planning coordinator for the first and second GitHub Field Day in Nepal.
- Responsible for the planning and coordination of pre-events and main event day, and management of selected and waitlisted applicants.

Managing Editor, The Zerone

Jan 2019 - Feb 2023

- Served as the **editor**, **associate editor**, and **managing editor** for the department magazine, The Zerone, in the years 2019, 2020, and 2023, respectively.
- Responsible for managing junior editors and editing, proofreading, shortlisting incoming articles.

Event Manager, Hult Prize at IOE, Pulchowk Campus

Sept 2021 - Jul 2022

- Acted as the event manager of one of the top on-campus programs in the Central Asia Region.
- Organized 4 physical orientations and three virtual orientations for 12 colleges, 4 entrepreneurship talk series and two workshops for 600+ participants in total and 2 mentorship sessions for 47 semifinalists and 18 finalists respectively.

Technical Manager, Pulchowk Girls

Sept 2020 - Jun 2022

- Served as the technical manager for the all-girls campus club.
- Responsible for handling technical aspects of webinars, designing content, preparing survey reports and managing three technical coordinators, who were part of the technical team.

Language Proficiency

TOEFL - 116 (30R/29L/29S/28W)