

BSc Project Screening Form: Guidelines

Part 1 – Project Proposal

Student Name	Charitra Shrestha
Student Number	2214705
Degree Pathway (course)	BSc (Hons) CS &SE
Supervisor Name	Krishna Aryal
Title of Project	Real-Time English - Nepali Bidirectional Speech Translation
Abstract of the project	<p>This research focuses on the development of a real-time bidirectional speech translation system between English and Nepali, utilizing artificial intelligence (AI) and natural language processing (NLP). The project aims to bridge communication gaps by providing an efficient, speech-to-speech translation tool. With the increasing global interactions, a real-time translator that seamlessly converts spoken English to Nepali and vice versa is essential for travelers, businesses, and local communities.</p> <p>This project will leverage automatic speech recognition (ASR) to convert spoken language into text, machine translation (MT) to process and translate the text, and text-to-speech (TTS) to generate speech output. Spectral Subtraction, Wiener Filtering and Deep Learning-based Denoising can be integrated into the ASR module to enhance speech clarity and improve recognition accuracy. The model will integrate deep learning techniques, including transformer-based NLP models.</p> <p>This research also explores the challenges in handling complex sentence structures, linguistic nuances, dialect variations, and pronunciation differences in Nepali and English. The goal is to develop a translation system with high accuracy and low latency, making it viable for practical applications in day-to-day</p>

	<p>communication.</p> <p>Keywords: Speech-to-Speech Translation, ASR, TTS, Machine Translation, Deep Learning, Real-Time Speech Translation</p>
Project deliverables	<ul style="list-style-type: none"> - AI-Powered Speech Translation System - Contextual Report - Final Report - Academic poster
Description of your artefact	<p>Project Background</p> <p>Nepal's growing globalization and tourism highlight the need for real-time translation systems, as many visitors struggle with language barriers. While mobile translation apps exist, they often require an internet connection and lack accuracy for underrepresented languages like Nepali.</p> <p>AI-driven speech translation has improved cross-lingual communication, and integrating AI with IoT has led to portable, efficient translation systems suitable for travelers, businesses, and education. Systems combining Automatic Speech Recognition (ASR), Neural Machine Translation (NMT), and Text-to-Speech (TTS) have enhanced translation accuracy and fluency.</p> <p>With advances in microcontroller-based processing, a real-time English - Nepali bidirectional speech translation system can be developed to operate offline using Raspberry Pi. By integrating ASR, NMT, and TTS with noise cancellation, this project aims to bridge language gaps and enable seamless communication in various real-world conditions.</p> <p>Problem Statement</p> <p>Language barriers pose significant challenges for tourists, expatriates, and local communities in Nepal, affecting communication in sectors such as tourism, healthcare, and business (Anon., n.d.). Existing translation solutions, like mobile</p>

applications and web-based tools, often fail to provide real-time, accurate translations, especially for low-resource languages like Nepali (Zhong & Yang, n.d.).

Moreover, speech translation systems typically require high computational power, making it challenging to deploy them on low-cost, embedded hardware such as Raspberry Pi or Arduino-based microcontrollers (Sarkar, et al., 2024). Noise interference and regional dialects further reduce the effectiveness of automated speech recognition, leading to inaccurate translation (Milinkovic & Milinkovic, June 2015).

Aim:

To research about the appropriate methods and approaches using IoT and AI, to make real time bidirectional English - Nepali speech translation system.

Objectives:

- Research and implement one of the approach or methods for speech translation using AI and IoT.
- Develop a voice translation system using machine learning and NLP models for accurate speech-to-text and text-to-speech conversion.
- Implement real-time speech processing on a microcontroller.
- Ensure seamless translation with minimal latency, providing a near-instantaneous response.
- Make the solution portable and accessible.

List of Features

- Real-time speech-to-speech translation.
- Low-latency processing for immediate response
- High-accuracy translation using a customized dataset.

	<ul style="list-style-type: none"> Portable and scalable hardware design. 																				
Risk analysis	<table> <thead> <tr> <th>Risk</th><th>Impact</th><th>Likelihood</th><th>Mitigation Strategy</th></tr> </thead> <tbody> <tr> <td>Inaccurate translation</td><td>High</td><td>High</td><td>Train models with a large English - Nepali corpus understanding.</td></tr> <tr> <td>Hardware failure</td><td>Medium</td><td>Medium</td><td>Use reliable microcontrollers.</td></tr> <tr> <td>Background noise affecting ASR</td><td>High</td><td>High</td><td>Implement noise-cancellation algorithms</td></tr> <tr> <td>Latency in translation</td><td>High</td><td>Medium</td><td>Optimize model and processing pipeline.</td></tr> </tbody> </table>	Risk	Impact	Likelihood	Mitigation Strategy	Inaccurate translation	High	High	Train models with a large English - Nepali corpus understanding.	Hardware failure	Medium	Medium	Use reliable microcontrollers.	Background noise affecting ASR	High	High	Implement noise-cancellation algorithms	Latency in translation	High	Medium	Optimize model and processing pipeline.
Risk	Impact	Likelihood	Mitigation Strategy																		
Inaccurate translation	High	High	Train models with a large English - Nepali corpus understanding.																		
Hardware failure	Medium	Medium	Use reliable microcontrollers.																		
Background noise affecting ASR	High	High	Implement noise-cancellation algorithms																		
Latency in translation	High	Medium	Optimize model and processing pipeline.																		
How does your project relate to your degree course and build upon the units/knowledge you have studied/acquired	<p>During my bachelor's in computer science and software engineering, I developed strong skills in programming, AI, and Machine Learning (ML), which are essential for building an English-to-Nepali Machine Translation system. This project integrates Natural Language Processing (NLP), AI-driven translation models, and software engineering principles to enhance real-time speech translation. My expertise in Python and Java has enabled me to implement Neural Machine Translation (NMT) models, process datasets, and optimize algorithms. Additionally, software engineering and computer architecture help structure efficient models, manage resources, and develop scalable APIs, while OOP principles ensure modular and maintainable code. Studying ML and NLP has deepened my understanding of deep learning, tokenization, and sequence-to-sequence models, crucial for training translation systems. Courses on research methodologies have strengthened my ability to evaluate datasets,</p>																				

	measure translation accuracy, and implement state-of-the-art techniques for improving translation quality.
Resources required in developing the artefact	<p>Hardware Requirements</p> <ul style="list-style-type: none"> - ESP32 Microcontroller - Microphone Module - Speaker Module

	<ul style="list-style-type: none"> - Remote server with 8GB ram, and Intel core I 7 for AI model <p>Software Requirements</p> <ul style="list-style-type: none"> - Python IDE (VS Code, Jupyter Notebook) - Python Interpreter 3.11+ - Git & GitHub - AI and NLP Libraries: TensorFlow, PyTorch, NLTK, SpaCy, Hugging Face Transformers, OpenNMT - Arduino IDE <p>Languages Used</p> <ul style="list-style-type: none"> - C/CPP for ESP32 programming - Python 3.11+
Have you completed & submitted your ethics form?	<input checked="" type="radio"/> YES <input type="radio"/> NO
If the project is a development of previous work by yourself or others, give details below. Failing to declare such previous work here may be treated as an academic offence	

Supervisor Signature: *[Signature]* *Feb 27, 2025* *[Signature]* *02/27/2025*

After the proposal has been signed off by both the supervisor and course coordinator scan the proposal and upload on BREO with signatures. Projects that follow proposals that have not been approved may be cancelled and there will be no compensation for any time lost

References

- Anon., n.d. Language and Communication: Overcoming Barriers in Nepal.
- Milinkovic, A. & Milinkovic, S., June 2015. *Continuous speech recognizer for low-end embedded devices*, s.l.: s.n.
- Sarkar, S., Babar, M. F. & Hasan, M., 2024. *Processing Natural Language on Embedded Devices: How Well Do Transformer Models Perform?*, s.l.: s.n.

Zhong, T. & Yang, Z., n.d. *Opportunities and Challenges of Large Language Models for Low-Resource Languages in Humanities Research*, s.l.: s.n.

Part 2 – List of relevant resources

1. Books

- a. Cabrera, X. G., 2008. *Real Time Speech*. s.l.:RESEARCH AND DEVELOPMENT CENTER.

2. Journal Papers

- a. Acharya, P. & Bal2, B. K., n.d. A Comparative Study of SMT and NMT: Case Study of English-Nepali Language Pair.
- b. Bangalore, S., Sridhar, V. K. R. & Kolan, P., n.d. Real-time Incremental Speech-to-Speech Translation of Dialogs.
- c. Divate, S., Biradar, G., Patole, A. & Attar, N., December-2023 . REAL TIME LANGUAGE TRANSLATOR. *International Research Journal of Modernization in Engineering Technology and Science* , 05(12).
- d. Greenstein, E. & Penner, D., n.d. Japanese-to-English Machine Translation Using Recurrent Neural Networks.
- e. Joshi, B. & Bhatta, B., 2023-02-17. End to End based Nepali Speech Recognition System. *Journal of the Institute of Engineering*, Issue 2023-04-03.
- f. Joshi, B. & Shrestha, R., December 2023. NEPALI SPEECH RECOGNITION USING SELF-ATTENTION NETWORKS. *International Journal of Innovative Computing, Information and Control*, Volume 19.
- g. Poudel, S. & Bal, B. K., 2018. Bidirectional English-Nepali Machine Translation System for the Legal Domain. pp. 90 -93.
- h. Prajapati, C., Nyupane, J., Shrestha, J. D. & Jha, S., 2008. Nepali Speech Recognition.

- i. Shakya, S. & Nemkul, K., 19-20 February 2021. English to Nepali Sentence Translation Using Recurrent Neural Network with Attention. Issue 12 April 2021.

3. Web Sites with relevant information

- a. Shrestha, A., 2021. *researchgate*. [Online]

Available at:

https://www.researchgate.net/publication/363213858_A_Reflection_on_Machine_Translation_Process_from_Nepali_to_English