Design of translation software with multi-inputs to translate English to Hindi OR vice-versa

A Project Work Synopsis

Submitted in the partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE WITH SPECIALIZATION IN

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:

20BCS6819 – Aniket Mittal
20BCS6809 – Bipandeep Singh Pannu
20BCS6801 – Ankit Kumar
20BCS6728 – Akshit Jindal

Under the Supervision of:

Lata Gupta (E13365)

CHANDIGARH UNIVERSITY
Discover. Learn. Empower.

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

PUNJAB

April, 2023

Abstract

In an increasingly interconnected world, bridging language barriers is paramount for effective communication and access to information. The "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project addresses this challenge by developing a versatile and user-friendly translation tool capable of seamless translation between English and Hindi, two widely spoken languages in diverse contexts.

The core objective of this project is to provide an accessible and efficient solution for users who require real-time translations in both directions. By integrating the power of modern language processing techniques and web technologies, this software aims to offer an intuitive interface accommodating various input methods. Users can input text for translation through text fields, file uploads, or even voice recognition, making it adaptable to different user preferences and needs.

The software leverages the Django web framework, which ensures robustness and scalability, while also providing a secure platform for users to interact with the translation tool. The translation logic incorporates machine translation models, allowing for rapid and accurate conversion of text. Additionally, error handling mechanisms and user feedback features are implemented to ensure a reliable and user-friendly experience.

Extensive testing procedures are applied to validate translation accuracy and system performance, guaranteeing high-quality output for users. Future enhancements may include language detection capabilities and support for additional languages, further expanding the software's utility and reach.

Keywords

Translation, Django, NLP, Application Programming Interface, Deployment

Table of Contents

1. INTRODUCTION

In our increasingly globalized world, language diversity presents both a rich tapestry of cultures and a significant communication challenge. Effective communication across language barriers is vital for personal interactions, business collaborations, and access to information. The "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project is a response to this challenge, aiming to provide a versatile and accessible solution for bridging the gap between English and Hindi, two languages spoken by millions worldwide.

The primary goal of this project is to develop a user-friendly and efficient translation software that empowers individuals and organizations to effortlessly translate text between English and Hindi, regardless of their language proficiency. This software offers a multifaceted approach to input, accommodating various methods such as text input, file uploads, and voice recognition. By incorporating state-of-the-art machine translation models and the robust Django web framework, we intend to deliver a reliable and high-quality translation experience.

This project is not only about enabling direct communication between individuals who speak different languages but also about facilitating access to valuable information in both languages. It contributes to cultural exchange, education, and inclusivity, ultimately promoting global connectivity and understanding. In the following sections, we delve into the features, technologies, and methodologies employed to realize this vision of breaking down language barriers through technology.

In summary, the "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project endeavours to simplify cross-language communication, making it accessible to a wider audience. This project contributes to breaking down language barriers, fostering cultural exchange, and facilitating access to information, ultimately promoting greater global connectivity.

1.1 Problem Definition

The problem at hand revolves around the challenges posed by language diversity and the need for effective translation solutions in a multicultural and multilingual society. Communication barriers arising from language differences often hinder seamless interactions between individuals, organizations, and communities.

Specifically, the problem centres on the translation needs between English and Hindi, two languages with distinct linguistic and cultural contexts. While English is a global lingua franca, Hindi is one of the most widely spoken languages in India and neighbouring regions. Bridging the language gap between these two languages is essential for various purposes, including international business, academic research, content localization, and personal communication.

Existing translation tools may lack the versatility and user-friendliness required to cater to a broad user base with diverse needs. This project addresses these limitations by designing a translation software solution that provides users with multiple input methods and accurate translations between English and Hindi. By doing so, it aims to overcome the language barriers that often hinder effective communication and access to information in our interconnected world.

1.2 Problem Overview

The challenge at the heart of this project lies in the need for a user-friendly and efficient translation solution between English and Hindi. Language diversity is a significant barrier to effective communication and information access in today's interconnected world. Bridging this gap requires a versatile translation tool that can handle various input methods and deliver accurate translations, enabling individuals, businesses, and communities to transcend linguistic boundaries. The project aims to address these issues by designing a translation software capable of seamlessly translating between English and Hindi, catering to a wide range of users and contexts.

1.3 Hardware Specification

1. RAM: 6GB or more
2. PROCESSOR: 64bit
3. Laptop with GPU with more than or equal to 4cores.

1.4 Software Specification

1. Python
2. Anaconda software
3. Jupiter Notebook
4. Virtual Studio Code

Tools Required:

1. Django
2. Python
3. Voice Recognition API’s
4. APIs for Machine Translation
5. Deployment platforms
6. Database
7. Machine Translation Models

2. LITERATURE SURVEY

Translation software and natural language processing have made significant advancements in recent years, enabling the development of versatile tools for overcoming language barriers. This literature survey delves into existing research and technologies related to language translation, multi-input methods, and machine translation models, providing valuable insights into the project's context and potential innovations.

Machine Translation Models
State-of-the-art machine translation models, such as BERT and GPT-3, have shown remarkable improvements in translation accuracy and fluency. These models are essential to the project's success by enabling more precise translations between English and Hindi.

Translation APIs:
Widely used translation APIs like Google Translate and Microsoft Translator offer accessible and reliable machine translation services. Understanding their strengths and limitations can inform decisions regarding the integration of external translation services into the software.

Web Development Frameworks:
Frameworks like Django provide a solid foundation for building web applications. Studying best practices in web development, especially in the context of multilingual applications, is crucial for creating a user-friendly interface.

User Experience (UX):
UX design principles and studies on user behaviour can guide the development of an intuitive and efficient user interface, ensuring that users can easily interact with the translation software.

By synthesizing insights from these areas, the project aims to create a translation software that leverages the latest advancements while addressing specific challenges associated with translating between English and Hindi. This literature survey serves as a foundation for informed decision-making and innovative solutions in the project's development.

2.1 EXISTING SYSTEM

The existing systems for English-to-Hindi or Hindi-to-English translation often lack the versatility and user-friendliness required for a broad user base. Many rely on outdated technology or fail to offer multi-input options. Machine translation models may not always deliver the desired level of accuracy. Additionally, user interfaces can be cumbersome, hindering effective communication. This highlights the need for an improved solution that combines the latest advancements in natural language processing, machine translation, and web development to create a more efficient and accessible translation tool.

2.2 PROPOSED SYSTEM

The proposed system, "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa," is a cutting-edge web application that addresses the need for versatile language translation between English and Hindi. Leveraging the power of Django and modern machine translation models, it offers users a user-friendly interface with multiple input options, including text input, file uploads, and voice recognition. The system aims to provide highly accurate translations, overcoming language barriers for individuals and organizations. With its dynamic and responsive design, it seeks to enhance user experiences and accessibility, ultimately promoting seamless cross-language communication.

2.3 LITERATURE REVIEW SUMMARY

The literature review conducted for this project revealed valuable insights into the current state of translation software and related technologies. It underscores the significance of bridging language barriers, particularly between English and Hindi, two languages with global relevance. Key findings and trends in the reviewed literature include:

Machine Translation Models: Recent advancements in machine translation, notably the Transformer-based models like BERT and GPT-3, have led to substantial improvements in translation accuracy and fluency. These models serve as a cornerstone for the development of our translation software.

Translation APIs: Widely used translation APIs, such as Google Translate and Microsoft Translator, have been instrumental in enabling real-time translation services. These services offer practical solutions for quick and accessible translations.

Multi-Input Methods: Research in human-computer interaction (HCI) emphasizes the importance of offering diverse input methods. Voice recognition, file uploads, and text input have emerged as versatile approaches to cater to a wide range of user preferences.

Web Development Frameworks: The reviewed literature highlights the relevance of web development frameworks like Django for building robust and secure web applications. This aligns with our choice of Django as the foundation for our translation software.

Language-Specific Challenges: Studies on translation between English and Hindi emphasize the need to address language-specific challenges, including syntactic and grammatical differences. These insights inform our approach to translation logic.

User Experience (UX): User experience design principles have a pivotal role in ensuring the usability and accessibility of translation software. Incorporating UX best practices is integral to our software's success.

Error Handling: Effective error handling and user feedback mechanisms are crucial for user satisfaction. Existing translation software often provides valuable lessons in implementing robust error management.

In summary, the literature review illuminates the project's contextual landscape, showcasing the potential for innovation and improvement in the development of our translation software. It validates the project's objectives and emphasizes the importance of combining the latest technologies and user-centric design principles to create a versatile and effective solution for bridging language barriers between English and Hindi.

3. PROBLEM FORMULATION

The global landscape is marked by cultural diversity and multilingual interactions, emphasizing the necessity of effective language translation solutions. Specifically, the problem at hand pertains to the challenges associated with language barriers, particularly in the context of translating between English and Hindi. While English serves as a ubiquitous global language, Hindi boasts millions of speakers across diverse regions, making bridging the gap between these languages of paramount importance.

The existing translation tools often fall short in terms of versatility and user-friendliness. They may lack support for multiple input methods and fail to provide consistently accurate translations. Inadequate user interfaces can hinder seamless communication, thereby limiting the potential for cross-cultural interactions.

The problem formulation centres on the development of a sophisticated translation software that addresses these challenges comprehensively. This software will enable users to translate text seamlessly between English and Hindi, leveraging state-of-the-art machine translation models and a user-friendly interface with versatile input options. The overarching goal is to overcome linguistic barriers and facilitate efficient communication across diverse contexts.

By formulating this problem, we seek to harness the power of technology and innovation to create a translation solution that empowers individuals, businesses, and communities to transcend language limitations, thereby fostering cultural exchange, enhancing accessibility to information, and promoting greater global connectivity.

4. OBJECTIVES

The primary objectives of the "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project are as follows:

Translation Excellence: Develop a translation software that excels in translating text accurately and fluently between English and Hindi, addressing language-specific nuances and challenges.

Versatile Input Methods: Provide users with versatile input methods, including text input, file uploads, and voice recognition, ensuring accessibility for users with diverse needs and preferences.

User-Friendly Interface: Create an intuitive and user-friendly interface that promotes ease of use and fosters a positive user experience.

Error Handling: Implement robust error handling and user feedback mechanisms to ensure reliable and user-centric translation services.

Scalability: Build a scalable and adaptable software framework that can accommodate future enhancements and expansions, including support for additional languages and features.

5. METHODOLOGY

The development of the "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project will follow a systematic and iterative approach to ensure its successful implementation. The methodology can be broken down into the following key stages:

Requirements Analysis:
Gather detailed requirements from stakeholders to understand the specific needs and expectations for the translation software. Identify key features, user scenarios, and input methods.

Literature Review:

Conduct an in-depth review of existing translation tools, natural language processing techniques, machine translation models, and user interface design principles. This informs decision-making and best practices throughout the project.

Design and Planning:
Create a comprehensive project plan, including the system architecture, database design, and user interface mock-ups. Define the translation logic, error handling mechanisms, and input methods. Establish a timeline and allocate resources effectively.

Development:
Build the translation software using the Django web framework and Python programming language. Implement the chosen machine translation model or integrate with translation APIs. Develop the user interface with support for multiple input methods.

Testing and Quality Assurance:
Rigorously test the software to ensure translation accuracy, reliability, and usability. Perform unit testing, integration testing, and user acceptance testing. Collect user feedback to make improvements.

Error Handling and Optimization:
Fine-tune the error handling mechanisms and user feedback features to enhance user experience. Optimize the translation logic for improved performance and accuracy.

Deployment and Hosting:
Deploy the software on a suitable web hosting platform or server infrastructure. Configure security measures and monitor system performance.

Documentation:
Create comprehensive user documentation and help resources to guide users in effectively using the translation software.

User Training and Feedback:
Provide user training if necessary and gather ongoing feedback to identify areas for improvement.

6. EXPERIMENTAL SETUP

The experimental setup for the "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project is designed to rigorously test and evaluate the performance, accuracy, and usability of the translation software. It encompasses the following components:

Development Environment:

The project is developed and tested in a controlled development environment, ensuring consistency in codebase and dependencies. Virtual environments are employed to manage project-specific dependencies and isolate the development environment.

Translation Models and APIs:

Machine translation models, including pre-trained transformer-based models for English and Hindi, are integrated into the software. Additionally, external translation APIs, such as Google Translate and Microsoft Translator, are used as benchmarks for comparison.

Diverse Test Data:

A diverse dataset comprising English and Hindi text samples is collected to represent various linguistic styles, complexities, and contexts. This dataset includes text from different domains, such as literature, technical documents, and colloquial communication.

User Scenarios:

User scenarios are designed to simulate real-world usage, encompassing different input methods like text input, file uploads, and voice recognition. Users from various linguistic backgrounds are involved in the evaluation.

Testing Protocols:

Comprehensive testing protocols are established to assess translation accuracy, fluency, and error handling. Quantitative metrics, such as BLEU score and translation speed, are measured to gauge the system's performance.

User Feedback Collection:

User feedback is actively collected during testing to identify issues, usability challenges, and suggestions for improvement. Surveys, interviews, and usability testing sessions are conducted to gather qualitative insights.

Hardware and Network Infrastructure:

The software is deployed on a suitable hosting platform or server infrastructure with sufficient computational resources and network capabilities to ensure reliable and responsive service.

Security Measures:

Security measures are implemented to protect user data and ensure the confidentiality and integrity of translations.

7. CONCLUSION

The "Design of Translation Software with Multi Inputs to Translate English to Hindi OR Vice-Versa" project represents a significant step toward overcoming language barriers and fostering effective communication between English and Hindi speakers. Through a systematic and comprehensive development process, we have created a versatile translation tool that excels in translation accuracy and user-friendliness.

This project successfully integrated state-of-the-art machine translation models, enabling precise translations, while offering a user-friendly interface that accommodates various input methods. The iterative development approach allowed us to fine-tune error handling mechanisms and optimize translation performance, ensuring a reliable and satisfying user experience.

The experimental setup, encompassing diverse test data and user scenarios, yielded valuable insights and confirmed the software's proficiency in real-world usage. User feedback played a pivotal role in identifying areas for improvement, reflecting our commitment to user-centric design. As we move forward, we recognize the ever-evolving nature of language and technology. This project serves as a foundation for ongoing enhancements and expansion, including support for additional languages and the integration of cutting-edge translation techniques.

In conclusion, the "Design of Translation Software with Multi Inputs" project contributes to global connectivity by breaking down language barriers, promoting cultural exchange, and facilitating access to information. It exemplifies our commitment to harnessing technology for the betterment of society, fostering inclusivity, and enabling effective communication across linguistic diversity.

8. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: INTRODUCTION

This chapter covers the basic overview of project “Short Term Crypto Currency Price Predictor using ML”, basic idea of how it works and use case, future scope.

CHAPTER 2: LITERATURE REVIEW

This chapter includes the literature available for the project work. The findings of the researchers are highlighted which is the basis of the current implementation.

CHAPTER 3: OBJECTIVE

This chapter provides introduction to the concepts which are necessary to understand the proposed system.

CHAPTER 4: METHODOLOGIES

This chapter covers the technical details of the proposed approach with various mathematical equations used in the project and some figures for the idea of how the code is working.

CHAPTER 5: EXPERIMENTAL SETUP

This chapter provides information about the subject system and tools used for evaluation of proposed methods.

CHAPTER 6: CONCLUSION AND FUTURE SCOPE

The major finding of the work is presented in this chapter. Also, directions for extending the current study are discussed.

REFFRENCES

[1] Aditi Kalyani, Priti S. Sajja, A Review of Machine Translation Systems in India and different Translation Evaluation Methodologies, Interna-tional Journal of Computer Applications (0975 8887) Volume 121 No23, July 2015.

[2] Suman K. Saksamudre, P.P. Shrishrimal, R.R. Deshmukh, A Review on Different Approaches for Speech Recognition System ,International Journal of Computer Applications(0975 888Volume115 No.22,April 2015

[3] Mouiad Fadiel Alawneh, Tengku Mohd Sembok Rule-Based and Example-Based Machine Translation from English to Arabic,2011 Sixth International Conference on Bio Inspired Computing: Theories and Applications

[4] F. Seide, G. Li, D. Yu, Conversational Speech Transcription Using Context-Dependent Deep Neural Networks In Inter speech, pp. 437440, 2011.

[5] y Keiichi Tokuda, Yoshihiko Nankaku, Tomoki Toda, Heiga Zen, Speech Synthesis Based on Hidden Markov Models.

[6] Dorfleitner G, Lung C (2018) Cryptocurrencies from the perspective of euro investors: a re-examination of diversification benefits and a new day-of-the-week effect. J Asset Manag 19(7):472–494

[7] Foley S, Karlsen JR, Putniņš TJ (2019) Sex, drugs, and bitcoin: how much illegal activity is financed through cryptocurrencies? Rev Financ Stud 32(5):1798–1853

[8] Hyun S, Lee J, Kim JM, Jun C (2019) What coins lead in the cryptocurrency market: using Copula and neural networks models. J Risk Financ Manag 12(3):132.

[9] Jiang Z, Liang J (2017) Cryptocurrency portfolio management with deep reinforcement learning. In: 2017 intelligent systems conference (intelliSys). IEEE, New York, pp 905–913

[10] Mallqui DC, Fernandes RA (2019) Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. Appl Soft Comput 75:596–606