NLP Translation

For Educational Content

Abstract

Traditionally, lot of challenges are faced by vernacular-medium students as compared to English-medium students, in terms of learning, opportunities and overall personality development. Hence, inline with the new National Education Policy 2020 promulgated by GoI, it is need of the hour to bridge the above gaps faced by Indian students using disruptive technology solutions — NLP, AI, IoT, etc. These solutions may cover audio/visual/textual/other modes of education content processing and delivery to the students.

INTRODUCTION

Here we are going to develop an web application for Natural Language Processing (NLP) technology solutions for mass-effective conceptual learning and education in vernacular languages (in mother tongues), which consists of some modules such as,

- Text to Text ,
- Audio to Audio,
- Video to Video ,
- Image to Text translation in vernacular languages.

System Flow

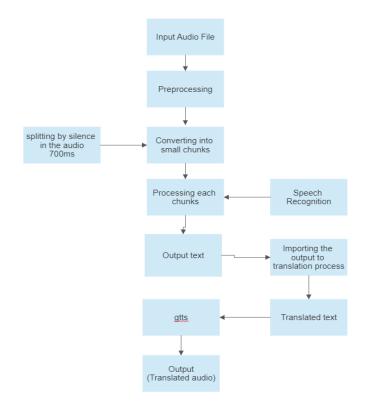


MODULES

Text to Text translation:

Input Text Preprocessing the input text Cleaning (spaces b/w words and Tokenization puntuation) Padding Word Embedding (same length) Generate the enocoded vectors Decode Output text

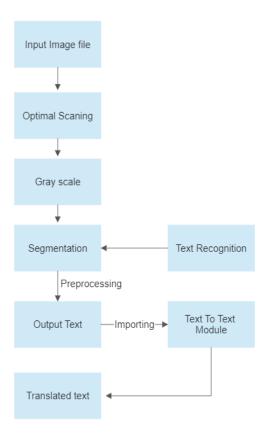
Audio to Audio translation:



Video to Video Translation:

Extract audio Audio file Input Video file Preprocessing splitting by silence Converting into in the audio small chunks 700ms Processing each chunks Speech Recognition Importing the Output text output to translation process Translated text gtts merge the audio → Output video file Translated audio

Image to Text(OCR):



Software Specification

We are developing an Web application using NLP, deep learning.

Frontend - HTML,CSS

Framework - Flask

Language - Python