Scribble-The handwritten recognition project using NLP

Abstract:

The handwritten recognition project using NLP aims to develop a system that can accurately recognize and transcribe handwritten text into digital format. This system utilizes natural language processing (NLP) techniques to identify and classify the handwritten text, and then convert it into machine-readable text. The project involves training a deep learning model using large datasets of handwriting samples to improve the accuracy of recognition. The system can be applied to a wide range of use cases, including digitizing handwritten notes, converting handwritten documents into digital archives, and improving accessibility for individuals with visual impairments. The project has significant potential to streamline workflows and improve efficiency in various industries, and may also have applications in fields such as forensics and document analysis.

Requirements:

To implement the proposed handwritten recognition system, the following requirements are needed:

Handwritten dataset: A dataset of handwritten text is required to train the system. The dataset should be large and diverse to ensure that the system can recognize different styles of handwriting.

Image pre-processing: The system should be able to pre-process the input image to remove noise, normalize the image, and segment the text into individual characters.

Convolutional Neural Networks (CNNs): The system should include a CNN to extract features from the input image.

Recurrent Neural Networks (RNNs): The system should include an RNN to generate a sequence of characters from the extracted features.

Language models: The system should include a language model to convert the sequence of characters into words and sentences.

Beam search decoding: The system should use beam search decoding techniques to find the most probable sequence of words and sentences.

Training environment: The system should be trained on a high-performance computing environment with GPUs to reduce the training time.

Evaluation metrics: The system should be evaluated using standard metrics such as character error rate (CER) and word error rate (WER) to measure the accuracy of the recognition system.

Guided by

Raj Kumar Singh

Project Coordinator

K Subhashini Spurjeo