conferenceseries.com

International Congress on AI AND MACHINE LEARNING

August 02, 2021 | Webinar

Forecasting probabilities of occurring the next possible word in a sentence in the online handwriting recognition systems

Harjeet Singh

Chitkara University, Punjab, India

In general, the prediction models are increasingly used for reasoning and decision making in various applications. Since, the demand of real-time based applications is increasing gradually due to the huge advancements in IT based devices such as Tablet-PC, touch-screen based smart phones, digital-pen/stylus based devices, digitizers etc. The present study describes about the forecasting probabilities of occurring the next possible Gurmukhi word in a sentence, which depends only on the immediately preceding word, written in the real-time environment. The online handwritten captured word information is first segmented into its individual strokes, which are recognized using Support Vector Machine (SVM) classifier. Thereafter, the bigram language model is utilized at stroke or character level in order to enhance the word recognition accuracy. The recognized word(s) is then used further to determine the occurrence of next possible word depending on their historical ability to forecast. Forecasting is a challenging task and totally dependent on the given data. In this study, the corpus, "Punjabi Monolingual Text Corpus-AnglaMT" (available at https://tdil-dc.in), containing approximately 83,000 sentences has been used for training the model. To overcome the data sparseness problem, the linear interpolation (Jelinek and Mercer, 1980) method is used for smoothing the n-gram estimates. The experiments show that the proposed online handwritten word forecasting framework significantly outperforms and produce consistent forecasts for the most likely word on the basis of given word.

harjeet.research@gmail.com