Off-Topic Text Detection System Using a Hybrid Model

Authors : Usama Shahid

Abstract : Be it written documents, news columns, or students' essays, verifying the content can be a time-consuming task. Apart from the spelling and grammar mistakes, the proofreader is also supposed to verify whether the content included in the essay or document is relevant or not. The irrelevant content in any document or essay is referred to as off-topic text and in this paper, we will address the problem of off-topic text detection from a document using machine learning techniques. Our study aims to identify the off-topic content from a document using Echo state network model and we will also compare data with other models. The previous study uses Convolutional Neural Networks and TFIDF to detect off-topic text. We will rearrange the existing datasets and take new classifiers along with new word embeddings and implement them on existing and new datasets in order to compare the results with the previously existing CNN model.

Keywords : off topic, text detection, eco state network, machine learning
Conference Title : ICCEIT 2023 : International Conference on Computer Engineering and Information Technology
Conference Location : Karachi, Pakistan
Conference Dates : December 25-26, 2023