Exploring the Effectiveness of Convolutional Neural Network with Ensemble Technique

CURRENT STATUS: UNDER REVISION

Journal of Big Data © Springer

Monalisha Ghosh
NIT Durgapur

monalisa_05mca@yahoo.com Corresponding Author
ORCiD: https://orcid.org/0000-0001-5827-8222

Goutam Sanyal
National Institute of Technology Durgapur

DOI:
10.21203/rs.2.21664/v1

SUBJECT AREAS
Theoretical Computer Science

KEYWORDS
Deep Learning framework, Sentiment Analysis, Binary PSO method, Convolutional Neural Network (CNN), Support Vector Machine
Abstract
Sentiment Analysis has recently been considered as the most active research field in the natural language processing (NLP) domain. Deep Learning is a subset of the large family of Machine Learning and becoming a growing trend due to its automatic learning capability with impressive consequences across different NLP tasks. Hence, a fusion-based Machine Learning framework has been attempted by merging the Traditional Machine Learning method with Deep Learning techniques to tackle the challenge of sentiment prediction for a massive amount of unstructured review dataset. The proposed architecture aims to utilize the Convolutional Neural Network (CNN) with a backpropagation algorithm to extract embedded feature vectors from the top hidden layer. Thereafter, these vectors augmented to an optimized feature set generated from binary particle swarm optimization (BPSO) method. Finally, a traditional SVM classifier is trained with these extended features set to determine the optimal hyper-plane for separating two classes of review datasets. The evaluation of this research work has been carried out on two benchmark movie review datasets IMDB, SST2. Experimental results with comparative studies based on performance accuracy and F-score value are reported to highlight the benefits of the developed frameworks.

Full Text
Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures
Figure 1

The Architecture of a proposed framework for the sentiment classification task.

Figure 2

CBOW: Continuous bag of words model
Figure 3
The proposed CNN Architecture for Feature Extraction.

Figure 4
Accuracy and Loss of CNN during training and validation.
Experimental Results (% Accuracy) of the proposed model for movie review data set

Cross-dataset (IMDB[SST2]) comparison through Classification Accuracy

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
Figures.pdf
Tables.pdf