JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

CONVOLUTIONAL NEURAL NETWORK MODELS FOR IMPROVED SOCIAL EMOTION CLASSIFICATION

Dr Selvapriya M
Associate Professor, Department of Computer Science, Hindusthan College of Arts & Science (Autonomous), Coimbatore,
Tamilnadu, India.

ABSTRACT

Social Media has gained remarkable attention. This is attributed to the affordability of accessing social network sites such as Twitter, Google+, Facebook and other social network sites through the internet and the web 2.0 technologies and in social media services has enormously helpful for various users to depict their feelings and opinions through news articles, blogs and tweets. Twitter is one of the important and popular social media in which anybody can leave tweets about anything that occurs. However, these emotions also include noisy instances and it is a enormous challenge to get the textual inference of brief messages. It is a difficult task to identify the identical user documents from the complete social media text message. Also, online comments are classified by attributes using a sparse feature space, which increases the complexity of the corresponding emotion classification task. Therefore, feature selection is a desirable solution for finding a solution to this problem. Hybrid Neural Network (HNN) has been proposed for Social Emotion Classification. But, reducing sparse feature space from the emotion dataset tends to be an extremely cumbersome task. Three important contributions were made in this work to address these concerns. In the first contribution of the work, Quantum Behavior Particle Swarm Optimization based Sparse Encoding (QBPSO-SEn) approach is presented to choose the Optimum features from dataset which is helpful in increasing the robustness of the CNN model. The results of the experiments show that the proposed CNN model yields an increased classification accuracy in comparison with other techniques like Hybrid Neural Network (HNN) and also Neural Network (NN) schemes. However, this scheme does not support multi-label emotion classification.The second contribution of the work, Mutation Bat Optimization based Sparse Encoding (MBO-SC) is presented for translating the sparse low-level features into tight high-level features.However, the feature selection approach is not used in this mechanism and it may influence the classification accuracy.The third contribution of the work, Integrated Feature Selection (IFS)-EWCNN is presented to boost the performance of social emotion classification done on twitter reviews.The proposed IFS-EWCNN classifier yields superior performance in comparison with the available techniques in terms of precision, recall, sensitivity, specificity, F-measure and accuracy. The experiments are performed in SemEval 2016 Task 4A for sentiment and SemEval 2018 Task 1C emotion classification using the MATLAB tool. The proposed IFS-EWCNN classifier techniques help in the accurate classification of social emotions. The results of experiments reveal that the proposed IFS-EWCNN classifier achieves 89.31% of precision, 96.40% of recall, 92.45% of f-measure, 96.23% of specificity, and 96.20% of accuracy correspondingly.

Keywords: Social Media, Emotion Classification, Feature Selection, Sparse Encoding, Mutation Bat optimization, Twitter Sentiment Analysis.

1. INTRODUCTION

The universal rise of social media facilities has been a huge help for several users to convey their emotions and ideas via news articles, twitter, etc., [3]. Twitter is one of the significant and common social media in which anyone can post tweets regarding any occurrence. People may convey their opinions or emotions with no pressure through public platform. Detection and analysis of the emotions conveyed in social media content is advantageous for several applications such as industry, social welfare, etc. This is very useful in determining if particular item or service is good or bad. Sentiment Analysis (SA) refers to the automated process of analysis of text data and then sorting it into sentiments that are positive, negative or neutral [1]. Emotion Analysis extends sentiment analysis. It refers to the process of evaluating the text and categorizing the text into various emotion classes like anger, disgust, trust, sadness, joy, surprise, etc. For mining the emotions and polarity in tweets, different text mining techniques are utilized in the earlier works. But, it faces problems with multi label classification and accuracy rate. This technical work highlights on social emotion and sentiment classification on twitter data. The present research work focuses on social emotion classification using efficient deep learning techniques.

2. LITERATURE SURVEY

Li et al (2017) developed a Hybrid Neural Networks (HNN) for classification of social emotions. This system proposed a new transfer learning scheme based on a KullbackLeibler (KL) divergence based stochastic gradient descent technique for transferring high-level semantic features extracted from unsupervised learning models to artificial neural networks to improve the efficiency of their emotion classification [4]. Katta and Hegde (2019) presented a Hybrid Adaptive Neuro-Fuzzy Inference System (ANFIS) and Support Vector Machine (SVM) based sentiment analysis on Political Twitter Data. The result revealed that the ANFIS- Non Linear SVM based Sentiment analysis of social network data has lesser computationalcomplexityand yields betteraccuracy[2]. Li et al (2019) presented a social emotion classification that depends on noise-conscious training. The results of experiments proves that the developed social emotion classification performance can be improved significantly[5]. Pandey et al (2017) presented a hybrid cuckoo search technique for Twitter sentiment analysis. The effectiveness of proposed technique has been tested on various Twitter datasets [6].

3. PROBLEM DEFINITIONS

Identifying emotions and sentiments from text is a classification job which falls under text mining. Emotion analysis refers to the function of deciding the attitude shown towards a target or topic. The attitude could be the polarity (positive or negative) or an emotional state like happiness, anger, or despair. Several statistical and machine learning classifier approaches have been designed for effective computing, which consists of Support Vector Machine (SVM), Neural Network (NN), k-Nearest Neighbors (KNN), Random Forest (RF) etc. But, the Nearest Neighbor classifier is very slowand complex in the high-dimensional data. Thus, training and working process of effective neural networks istough. Random Forest (RF) classification process is very complexand consumes large amount of time for construction compared to decision trees. Many of the earlier works on sentiment and emotion analysis have only concentratedon single-label classification. The earlier works did not considerthe co-presence of several emotion labels in one instance. Thus,newmachine learning based system model is essentialto handle several emotion classification problems found in Twitter.

4.OBJECTIVES OF THE RESEARCH

The important goals of this research work are:

- To design and deploy efficient Swarm Intelligence (SI)algorithms for reducing sparse space for achieving improved classifier accuracy.
- To implement efficient deep learning algorithms for better emotion classification with enhanced emotion accuracy
- To design an efficient sentiment classification model using fuzzy clustering algorithm for achieving better sentiment accuracy
- To achieve higher emotion classification accuracy with reduced space dimensional and increased detection rate.

5. CONTRIBUTION OF THE RESEARCH

5.1 Quantum Behavior Based Sparse Encoding For Feature Reduction and Social Emotion Classification with Semantically Rich Convolutional Neural Network

Convolutional Neural Network (CNN) based classification method is used in this work for categorizing emotions in social site with the help of online comments.SemEval-2007 dataset is considered as input. It includes 1,250 news headlines obtained from several sources such as Google news articles, CNN news articles, etc. In the pre-processing stage, Biterm Topic Model (BTM) and Word2vec are utilized. The Quantum Behavior Particle Swarm Optimization based Sparse Encoding (QBPSO-SEn) approach is used to reduce sparse feature space. As per QBPSO, the particle takes the sparse data and fitness evaluation yields the best possible λ value. In QBPSO, each particle makes its position up to date for the evolution period. The important objective of this approach is to sparsely reducethe number features provided by unsupervised teaching models for boosting the stability of the classifier model. CNN with Latent Semantic Machine (LSM) is developed for emotion classification. The proposedCNN modelincludes an input layer, convolutional layer, pooling layer, one or more LSMs, feature layer and an output layer. But, this approach does not provide support to multi-label emotion classification.

5.2. Enhanced Weight based Convolutional Neural Network (EWCNN) and fuzzy clustering for semantically rich multi-label social emotion classification

To improve the performance of multi-label emotion classification, Enhanced Weight based Convolutional Neural Network (EWCNN) and Fuzzy Clustering algorithm is applied. The word embedding vectors indicate the semantic associations among words that can be exploited for improving the text classification tasks.To reduce the dimensionality of the feature space the Mutation Bat Optimization based Sparse Encoding (MBO-SC) is used. Bat algorithm operates on the basis of the echolocation skills of micro bats and it has been formulated by their foraging behaviour. For emotion classification, EWCNN algorithm is used as a classifier.The proposed system influences the semantically EWCNN classifier to include semantic domain knowledge into the neural network to bootstrap its inference power and interpretability. But, Feature Selection technique is not used in this approach that may affect the classification accuracy.

5.3 Integrated Feature Selection (IFS) algorithm and Enhanced Weight based Convolutional Neural Network (EWCNN) for Social Emotion Classification

For social emotion classification, the proposed approach formulated is an Integrated Feature Selection (IFS) algorithm and EWCNN. This approach aims to provide an emotion and sentiment classification model that consists of four steps. In the pre-processing stage, Biterm Topic Model (BTM) and Word2vec are utilized. Later**,** IFS is performed using filter and wrapper techniques. Filtering algorithm is performed by using the information gain and Pearson’s Correlation techniques. Wrapper selection algorithm is performed by using Expectation Maximization with Forward Search (EMFS), and Exhaustive feature selection. Then, Animal Migration Optimization based Sparse Encoding (AMO-SC) scheme is utilized for dimensionality reduction. The quality of the fitness (accuracy) is replaced by the probability of some new animals using the selected feature vectors approach. Finally, emotion classification is performed by using EWCNN. In this stage, for efficient classification of the sentiment samples fuzzy clustering is utilized. The proposed EWCNN approach is evaluated in terms of accuracy, precision, recall, specificity and f-measure and is compared with other existing approaches.

6. RESULTS AND DISCUSSION

In the MATLAB simulation environment the experimental assessment is carried out for both available and discussed research approaches. For English, a popular Twitter dataset obtained from SemEval 2016 Task 4A is used as source sentiment classification task. Task A includes 3 sentiment classes, which include positive, neutral and negative. For target emotion classification task, the Twitter dataset, currently published by SemEval 2018 Task 1C containing 11 emotions is used. The word embedding size d is fixed to be 300 for E1. In this, metrics such as accuracy, precision, recall, specificity and f-measure are considered. The results of the proposedIFS-EWCNN classifier are evaluated and compared with various methods such as Hybrid Neural Networks (HNN), Semantic Emotion Topic Model (SETM), 1-NN, CNN, EWCNN. Table 1 illustrates the overall performance comparison results of the proposed and existing techniques.

Table 1. Metrics Comparison vs. Classification Methods

Metrics	Methods					
	1-NN	SETM	HNN	CNN	EWCNN	IFS- EWCNN
Precision (%)	84.21	84.24	84.56	85.23	86.28	89.31
Recall (%)	77.29	79.28	79.68	89.24	95.37	96.40
F-measure (%)	87.09	87.71	87.27	89.33	90.18	92.45
Specificity (%)	81.35	86.56	86.82	90.43	94.96	96.23
Accuracy (%)	80.90	82.40	82.50	84.70	94.90	96.20

The results of experiments reveal that the proposed IFS- EWCNN approach yields 96.20% accuracy, 89.31% precision, 96.40% recall, 92.45% f-measure and 96.23% specificity that are observed to be significant when compared to 1-NN, SETM, HNN, CNN and EWCNN approaches.

7. CONCLUSION AND FUTURE WORK

The major aim of this work is to effectively classify the social emotion and sentiment classification on twitter data. In this proposed research work, Integrated Feature Selection (IFS) algorithm and Enhanced Weight based Convolutional Neural Network (EWCNN) is used for improving the classification accuracy of social emotion. In order to select optimal features, IFS algorithm is utilized. Then the sentiment classification is performed by fuzzy clustering and emotion classification is performed with the help of EWCNN classifier. The results of experiments reveal that the designed system yields 96.20% of accuracy, 89.31% of precision, 96.40% of recall, 92.45% of f-measure and 96.23% of specificity which are higher than 1-NN, SETM, HNN, CNN and EWCNN. The future scope of this research work is to boost the accuracy of emotional classification by presenting a generalized index of document significance. The proposed technique is intended to expand to emotionally conscious recommendation of events in social media improved systems and multimedia retrieval systems by combining with low-level features.

REFERENCES:

1. Akhtar, M. S, Gupta, D, Ekbal, A and Bhattacharyya, P, (2017). Feature selection and ensemble construction: A two-step method for aspect based sentiment analysis. Knowledge-Based Systems, vol.125, no.6, pp.116-135.
2. Akkouche, N, Balistrou, M, Loubar, K, Kadi, M. E. A and Tazerout, M, (2020). Pyrolysis polybutadiene model including self-heating and self-cooling effects: kinetic study via particle swarm optimization. Waste and Biomass Valorization, vol.11, no.2, pp.653-667.
3. Aldayel, H. K and Azmi, A. M, (2016). Arabic tweets sentiment analysis–a hybrid scheme. Journal of Information Science, vol.42, no.6, pp.782-797.
4. Al-Radaideh, Q. A and Al-Qudah, G. Y, (2017). Application of rough set-based feature selection for Arabic sentiment analysis. Cognitive Computation, vol.9, no.4, pp.436-445.

5. AL-Rubaiee, H. S, Qiu, R, Alomar, K and Li, D, (2016). Sentiment analysis of Arabic tweets in e-learning. Journal of Computer Science, vol.12, no.11, pp.553-563.
6. Al-Thubaity, A, Alqahtani, Q and Aljandal, A, (2018). Sentiment lexicon for sentiment analysis of Saudi dialect tweets. Procedia computer science, vol.142, no.11, pp.301-307.
7. Ambar, S, Jan, S and Khan, F, (2019). Human Emotion Detection And Classification In Text Mining. Pakistan Journal of Science, vol.71, no.4, pp.235.
8. Aston, N, Liddle, J and Hu, W, (2014). Twitter sentiment in data streams with perceptron. Journal of Computer and Communications. Journal of Computer and Communications, vol.02, no.03, pp.11-16.
9. Baali, M and Ghneim, N, (2019). Emotion analysis of Arabic tweets using deep learning approach. Journal of Big Data, vol.6, no.1, pp.1-12.
10. Beigi, G, Hu, X, Maciejewski, R and Liu, H, (2016). An overview of sentiment analysis in social media and its applications in disaster relief. Sentiment analysis and ontology engineering, vol.639, no.1, pp.313-340.
11. Billal, B, Fonseca, A, Sadat, F and Lounis, H, (2017). Semi-supervised learning and social media text analysis towards multi-labeling categorization. IEEE International Conference on Big Data (Big Data), vol.430, no.1, pp. 1907-1916.
12. Bollegala, D, Weir, D and Carroll, J, (2012). Cross-domain sentiment classification using a sentiment sensitive thesaurus. IEEE transactions on knowledge and data engineering, vol.25, no.8, pp.1719-
1731.
13. Bouazizi, M and Ohtsuki, T, (2017). A pattern-based approach for multi-class sentiment analysis in Twitter. IEEE Access, vol.5, no.8, pp.20617-20639.
14. Boudad, N, Faizi, R, Thami, R. O. H and Chiheb, R, (2017). Sentiment Classification of Arabic Tweets: A Supervised Approach. J. Mobile Multimedia, vol.13, no. 3-4, pp.233-243.
15. Brahimi, B, Touahria, M and Tari, A, (2016). Data and Text Mining Techniques for Classifying Arabic Tweet Polarity. Journal of Digital Information Management, vol.14, no.1, pp. 15-25.
16. Bravo-Marquez, F, Frank, E, Mohammad, S. M and Pfahringer, B, (2016). Determining word-emotion associations from tweets by multi-label classification. In 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Omaha, NE, USA, no.10, pp. 536-539.
17. Brynielsson, J, Johansson, F, Jonsson, C and Westling, A, (2014). Emotion classification of social media posts for estimating people's reactions to communicated alert messages during crises. Security Informatics, vol.3, no.1, pp.1-11.

18. Cai, J, Luo, J, Wang, S and Yang, S, (2018). Feature selection in machine learning: A new perspective. Neurocomputing, vol.300, no.7, pp.70-79.
19. Cai, X, Wang, H, Cui, Z, Cai, J, Xue, Y and Wang, L, (2018). Bat algorithm with triangle-flipping strategy for numerical optimization. International Journal of Machine Learning and Cybernetics, vol.9, no.2, pp.199-215.
20. Cambria, E, Das, D, Bandyopadhyay, S and Feraco, A, (2017). Affective computing and sentiment analysis. In A practical guide to sentiment analysis, Springer, Cham, pp. 1-10.