ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

Different Approaches for Humor Recognition: A Survey

Anisha T S[1], Rafeeque P C[2]

PG Student, Dept of Computer Science & Engineering, Government Engineering College, Palakkad, Kerala, India[1]

Head of the Department, Department of Computer Science and Engineering, Government Engineering College, Palakkad, Kerala, India[2]

Abstract: Humor is the most gripping factor in human to human communication which provokes laughter and or amusement. It also have a great impact on human beings mental and physical health and hence, it is an essential element in communication. The purpose of this paper is to survey different approaches for humor recognition. Various methodologies are proposed for recognition of humor and they are Machine Learning, Discourse Parsing and Deep Learning. The paper also discuss the advantages and limitations of methodologies used. The various datasets used for humor recognition is also discussed in this paper.

Keywords: Convolutional Neural Network (CNN), Humor Recognition, Long Short-Term Memory (LSTM), Rhetorical Structure Theory (RST)

I. INTRODUCTION

Humor could be recognized as one of the most intelligent activity in personal communication. Thoughtful use of humor can help in eliminating embarrassment, in eliminating social barriers, and to create positive affection in social interactions between people and so, the role humor plays in life can be seen as a social and functional phenomenon. If computers can understand humor to some extent, it would make it easier to predict the intention of a person in a conversation. Humor can be seen as a cognitive process that causes laughter and entertainment. It promotes not only the success of human interaction, but also has a positive effect on mental and physical health.

The task of humor recognition is to determine whether a sentence expresses a certain degree of humor in a given context. The recognition of humor is a classification task in which we differentiate between humorous and non-humorous instances. Recently, humor recognition has drawn more attention as described by R. Mihalcea and C. Strapparava [1]. The main trend is to design interpretable and computable features, which can be easily explained and implemented in practice by humor theories. There are four semantic structure behind humor from four perspectives: incongruity, ambiguity, interpersonal effect and phonetic style. Design a set of features for each latent structure to capture possible humor indicators. Humor recognition is viewed as a classification problem as described in [1]. The main objective is to determine if a given text contains humorous expressions. For example, consider a sample in each category (True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN)) to get an idea of what kinds of sentences are predicted correctly and incorrectly.

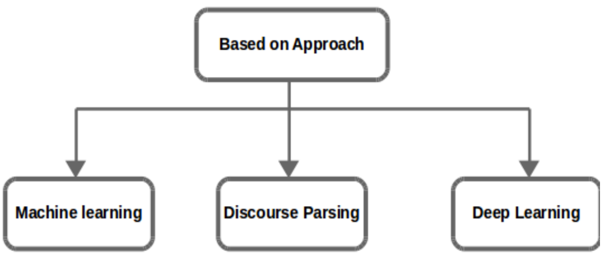

Fig. 1. Classification of Methodologies based on Approach Used in Humor

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

The TP sentence when he gave his wife a necklace he got a chain reaction shows that the model seems to be able to catch not only the literal meaning between the necklace and got a chain reaction. The TN sentence the barking of a dog does not disturb the man on a camel means that if you are lucky enough to own your own camel, a little thing like a barking dog won't bother you. It's a proverb, but not a joke, and the model recognizes it correctly as a non-humor.

Humor recognition is a challenging problem in natural language understanding. Numerous factors makes the task of automated humor recognition difficult. Firstly, it is difficult to achieve a universal definition of humor, because different people have different understandings of the same phrase. It requires lots of external knowledge since humor is situated in a broader context.

There are various types of humor such as wordplay, irony and sarcasm, but there are few formal humor taxonomies which is well explained by Y. Raz [2]. In general, humor is loosely defined. It is therefore impossible to create rules to identify humor. However, it is difficult for computers to build a mechanism to understand humor like human beings from both theoretical and computational perspectives.

There are various applications of humor recognition such as user intention systems. That is, if computers can understand humor in some way, it would make it easier to predict the intention of people in human conversation and thereby increase the proficiency of many machine-human interaction systems. Automatic humor recognition is also important in joke-generation systems, riddle-generation systems, dialogue-generation systems. It also helps to detect irony and for making the presentation effective. So, it is relevant in public speaking which reduce the social phobia, tension and helps to attract audience.

There are various methodologies for recognizing humor automatically and they are Machine Learning, Discourse Parsing, Deep Learning. Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) and Highway Networks are the major framework used in deep learning based approach for recognizing humor. Fig. 1. shows the classification of methodologies based on the approaches they use for humor recognition. This survey aims to discuss different ways of recognizing humor and thus highlight the advantages and disadvantages of each. Such a comparative study is very important as there are wide range of applications using humor recognition systems. This survey will provide some insights for choosing the right methodology to develop a humor recognition system.

This paper is organized as follows: Section II gives a formal definition of the humor recognition system and an example. Section III discusses about various approaches used for recognizing humor and their advantages and disadvantages. It provides a comparison between the methodologies. Section IV discusses the future scope of humor recognition. Section V gives a brief conclusion of humor recognition.

II. HUMOR RECOGNITION

Humor recognition is a traditional text classification problem which distinguishes a sentence as humor or not and a created computational models to discover the latent semantic structure behind humor from four perspectives: incongruity, ambiguity, interpersonal effect and phonetic style. Automatic humor recognition is a system which gives the output as whether the sentence contains humorous instances or not. Research in humor has defined many different theories of humor and it has proven that these theories are very important in recognizing humor. The highly recognized and dominant theories which has been used in the existing works are superiority theory, relief theory and incongruity theory. Fig.2. shows the basic idea of humor recognition systems.

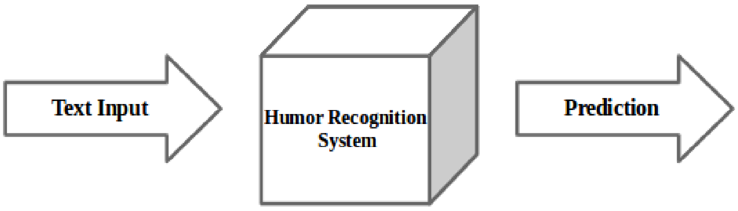

Fig.2. Basic idea of humor recognition system

The problem of humor recognition system can be formally defined as: Given an input which contains humorous and non-humorous instances, the goal of the humor recognition system is to classify whether a sentence in a given context expresses a certain degree of humor or not.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

Let us consider an example for humor recognition. Let the sentences be: 1. When he gave his wife a necklace he got a chain reaction and 2. The barking of a dog does not disturb the man on a camel. The system identifies sentence 1 has humorous sentence and sentence 2 has a proverb.

III. DIFFERENT METHODOLOGIES OF HUMOR RECOGNITION

A. Machine Learning Approaches

Classification algorithms in machine learning are widely used in natural language processing application. Classifiers like Random Forest, K-nearest neighbours are used for recognizing humor. Diyi Yang et. al. [3] proposes an approach to identify several semantic structures and design sets of features for each structure and use a computational approach to recognize humor. A simple and effective way of extracting anchors that allow humor in a phrase is also developed in this methodology. The data sets used in this methodology are Pun of the Day 1 and the 16000 One-Liner dataset which is explained in detail in [1].

Formulated humor recognition as a traditional problem of text classification and use Random Forest to perform 10 fold cross validation on two datasets. Random Forest is an ensemble of decision trees for classification (regression) that builds decision trees during training and outputs the class mode of individual trees.

Random forests use averaging to find a natural balance between the two extremes. K Nearest Neighbor (KNN) features that uses the humor classes of the K sentences (K = 5) that are the closest to the sentence are also designed as a feature. For comparison with the classifier such as RF and KNN, several baselines are used. They are Bag of Words, Language Model, Word2Vec, SaC Ensemble.

To understand humor recognition thoroughly, Anchors are extracted. Anchors are words that prompt humor in a sentence. Scoped humor anchor candidates for the syntactic categories of words or phrases of Noun, Verb, Noun Phrase, Verb Phrase, ADVP or ADJP. It is achieved via a parse tree. In order to generate anchor candidates, we have analyzed each sentence and selected words or phrases that meet one or more of the latent structure criteria by first extracting the minimal sub trees of NP, VP, ADVP. And finally adding remaining nouns and verbs into candidate sets.

The main advantage of methodology [3] is both quantitative and qualitative experimental results are taken when anchors are extracted. It is a classification task which distinguishes between humorous and non-humorous instances. There are several issue in this methodology [3] such as difficult to understand the meaning of phrase, needs lots of external knowledge, categorization of humor is difficult since there are different types of humor.

The performance of different semantic structures are validated and how the combination of the structures contributes to classification. The data sets such as Pun of the Day 1 and the 16000 One-Liner dataset are balanced in terms of positive and negative instances giving a classification accuracy of 50%. For humor anchor extraction, both qualitative and quantitative results are obtained.

B. Discourse Parsing Approaches

Lizhen Liu et.al. [4] proposes an idea to exploit sentiment analysis for humor recognition by considering humor theories[5]. Here main idea is to model sentiment association between discourse units. In the methodology [4], the main indicators for recognizing humor are discourse relation, sentiment conflict and sentiment transition. RST(Rhetorical Structure Theory) style discourse parsing described in [6] are used to get discourse units. The dataset used in this methodology [4] contains 10,200 humorous short texts and 10000 non-humorous short texts from Reuters titles and Proverbs and British National Corpus (RPBN).

Features are designed by highly recognized and dominant humor theories. For comparison with the classifiers, several baselines are used and they are incongruity structure, ambiguity, phonetic style, ambiguity, KNN feature. These five feature sets are known as Humor Centric Features(HCF). Besides, these five features, there is Word2Vec features which uses average word embeddings as sentence representations for classification.

To automatically recognize humor, discourse parser is implemented. RST structure builds over the entire text a hierarchical structure. A coherent text is represented as a discourse tree, the leaf nodes of which are individual text units known as the elementary discourse unites (EDUs). It can be connected through relations. The parser can automatically separate a sentence into EDUs. And then it gives discourse relations between EDUs.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

Th main advantage of this methodology [4] is sentiment association in discourse is more useful than counting the number of emotional words. Discourse parsing and sentiment analysis is done to get sentiment association patterns. The main limitation is that non-humorous texts contain same-unit and attribution in the non-humorous texts are more. The numbers of emotional words are used as features but emotional word count (EWC) is short. In this methodology [4], 79% of the humorous instances contain more than one EDU and 38% of non-humorous messages contain more than one EDU. The humorous texts may conatin more complex sentence structures. According to sentiment analysis tool, 57% of humorous instances have non-neutral polarity and 47% of non-humorous instances have non-neutral polarity. Sentiment transition is the most useful indicators for humor recognition.

C. Deep Learning Approach

Deep learning are the hot areas in machine learning and artificial intelligence for past few years. It is proving that it is the best method for various problems in day to day life. With the advancement of deep learning, humor recognition is becoming promising. There are various frameworks of deep learning such as CNN, LSTM for recognizing humor.

Peng-Yu Chen et. al. [7] proposes a deep learning approach for humor recognition. Four datasets such as Pun of the Day, 16000 One-Liners, Short Jokes dataset and PTT jokes are used with distinct joke types in both English and Chinese. Convolutional Neural Network and Highway Networks are the methods used in this methodology [7]. Highway network increases the proficiency of the architecture.

Convolutional Neural Network (CNN) is used to extract features for image and speech signals. When it comes to natural language processing (NLP), it is used for various applications such as text categorization. In this methodology [7], converted tokenized input sentence to a 2D matrix using GloVe embedding vectors trained on 6B tokens and 400K vocabulary words of Wikipedia 2014 + Gigaword 5 as the embedding layer. Filter sizes in the range of 3 to 20 are used.

For each filter size, 100-200 filters are applied to the model. After convolutional layer, max pooling is performed and then flatten the output. To increase the performance of the architecture, highway networks is used. When training of deeper networks becomes more difficult with increasing depth, highway networks is used which increases the performance. It consists of gate units that regulate the flow of information
through the network.

The main advantage of [7] are humor recognition is done to different types of humor and to different languages. It uses concatenation of CNN with highway networks if there are more deeper netwoks. Performance is increased when usingthe concatenation of CNN with highway networks. But there is a difficulty in selecting the linguistic features. The dataset consists of four parts: Pun of the Day, 16000 One-Liners, Short Jokes dataset and PTT jokes. The results of the model helps increase F1-Score from 0.859 to 0.903 on 16000 One-Liners and from 0.705, 0.864 to 0.901 on Pun of the Day compared to previous work.

Dario Bertero et. al. [8] proposes another deep learning framework for recognizing humor. Long Short-Term Memory (LSTM) based framework is used to predict humor in dialogues. It is used first time in this methodology [8]. Data is taken from a popular TV-sitcom: ”The Big Bang Theory”. The classifier is made of a concatenation of a CNN followed by a LSTM. A vector of higher level syntactic, structural and sentiment features before the output softmax layer.

The input feature is fed into the first embedding layer to obtain a low dimensional dense vector. A sliding window is applied. To extract the salient features of all the tokens into a single vector for the whole utterance, max pooling is done. In this methodology [8], the three input features are word tokens, character bi-grams, Word2Vec.

This methodology [8] reduces the number of false positive and increase the F1 score. Predicting humor response from the canned laughter remains a challenging task. The audience must also be kept constantly amused. Built a corpus from thepopular TV-sitcom :”The Big Bang Theory” from seasons 1 to 6. Downloaded the subtitle files which is used to segment all the episodes into scenes and get the speaker identity of each utterance. The audio track of each episode is extracted and manual annotation is performed.

D. Discussion on Datasets for Humor Recognition

Pun of the Day dataset used in [7] and [3] is constructed from the Pun of the Day website. The pun also known as paronomasia is a form of wordplay that takes advantage of multiple meanings of similar words for a humorous effect. The negative samples of this dataset are taken from news website.Short Jokes dataset used in methodology [7] are collected from most of the jokes among four datasets of an open database on a kaggle project. It contains 231,657 short

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

jokes and length ranging from 10 to 200 characters. The negative samples of this dataset are taken from WMT16 english news.

PTT Jokes is a dataset written in chinese used in methodology [7]. PTT Bulletin Board System is the largest terminal-based bulletin board system (BBS) in Taiwan. It has more than 1.5 million registered users. The negative samples of this dataset are taken from Yahoo News in politics.

16000 One-Liners dataset used in [7] and [3] contains humorous samples from daily joke websites. A single-liner is a joke that usually has very few words with comic effects and an interesting linguistic structure in a single sentence. Longer jokes can have a relatively complex linguistic structure, a single-liner with very few words has to produce a humorous effect.These are the main four datsets used for humor recognition. There are other datasets also. Reuters titles and Proverbs and British National Corpus (RPBN) is a dataset that contains 10,200 humorous short texts and 10000 non-humorous short texts used in [4]. For recognizing humor from dialogues, a corpus from the popular TV-sitcom is used.Various other manually created datasets are used for humor recognition. Since the datasets and evaluation criteria used in the above are same for some methodologies. Table 1 shows the contribution of methodologies discussed and dataset used.

Table I: Overview of Compared Approaches and Datasets Used

Type of Approach	**Proposed Approaches**		
	Model	**Highlights**	**Dataset**
Machine Learning	Diyi Yang et.al. [3]	Random Forest and KNN	Pun of the Day 1 and 16000 One-Liner
Discourse Parsing	Lizhen Liu et.al. [4]	Rhetorical Structure Theory	Reuters titles and Proverbs British National Corpus (RPBN)
Deep Learning	Peng-Yu Chen et.al.[7]	Convolutional Neural network	Pun of the Day 1, PTT Jokes and 16000 One-Liner
	Dario Bertero et.al.[8]	Long Short-Term Memory	Data from TV siticom

IV. CONCLUSION AND FUTURE WORK

Proposed various approaches such as Machine Learning, Discourse Parsing, LSTM and CNN architecture with Highway Networks that can learn to distinguish between humorous and non-humorous texts. Also extended the techniques of automatic humor recognition to different types of humor as well as different languages in both English and Chinese. The proposed idea can be explained with major humor theories. Different types of datasets are used for various methodologies are also discussed in this paper. The datasets used are Pun of the Day, 16000 One-Liners, Short Jokes dataset and PTT jokes, Reuters titles and Proverbs and British National Corpus (RPBN), data from television channels for finding humor in dialogues. Humor recognition is the challenging task in natural language understanding. This survey focuses on the various kinds of methodologies for humor recognition.

The performance of the CNN model outperforms the previous work. The accuracy on PTT, political jokes in Chinese, and the short jokes dataset with various types of jokes in English are both as high as above 90%. The novel deep learning model relieves the required human intervention of selection linguistic features for humor recognition. There are few suggestion for building better automatic humor recognition systems. Discovering the characteristics of humor and then generation of humor. Use a virtual agent system to predict humor by using human robot humorous interactions. Various other methods or frameworks of deep learning can be used to recognize humor. Also find how the humorous texts can be generated using deep learning models as well. Comparative evaluation of deep learning models for generation of humor is not done yet.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

REFERENCES

[1]. R. Mihalcea and C. Strapparava, "Making computers laugh: Investigations in automatic humor recognition," in Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 531–538, Association for Computational Linguistics, 2005.
[2]. Y. Raz, "Automatic humor classification on twitter," in Proceedings of the 2012 Conference of the North American Chapter of the Associationfor Computational Linguistics: Human Language Technologies: Student Research Workshop, pp. 66–70, Association for Computational Linguistics, 2012.
[3]. D. Yang, A. Lavie, C. Dyer, and E. Hovy, "Humor recognition and humor anchor extraction," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 2367–2376, 2015.
[4]. L. Liu, D. Zhang, and W. Song, "Modeling sentiment association in discourse for humor recognition," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), vol. 2, pp. 586–591, 2018.
[5]. C. Gruner, "R.(1997). the game of humor: A comprehensive theory of why we laugh. new brunswick."
[6]. W. C. Mann and S. A. Thompson, "Rhetorical structure theory: Toward a functional theory of text organization," Text-Interdisciplinary Journal for the Study of Discourse, vol. 8, no. 3, pp. 243–281, 1988.
[7]. P.-Y. Chen and V.-W. Soo, "Humor recognition using deep learning," in Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), vol. 2, pp. 113–117, 2018.
[8]. D. Bertero and P. Fung, "A long short-term memory framework for predicting humor in dialogues," in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 130–135, 2016.
[9]. L. Chen and C. M. Lee, "Predicting audience's laughter using convolutional neural network," arXiv preprint arXiv:1702.02584, 2017.
[10]. T. Miller and I. Gurevych, "Automatic disambiguation of english puns," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), vol. 1, pp. 719– 729, 2015.
[11]. T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in Proceedings of the conference on human language technology and empirical methods in natural language processing, pp. 347–354, Association for Computational Linguistics, 2005.
[12]. C. Dyer, M. Ballesteros, W. Ling, A. Matthews, and N. A. Smith, "Transition-based dependency parsing with stack long short-term memory," arXiv preprint arXiv:1505.08075, 2015.
[13]. L. Devillers, S. Rosset, G. D. Duplessis, M. A. Sehili, L. Béchade, A. Delaborde, C. Gossart, V. Letard, F. Yang, Y. Yemez, et al., "Multimodal data collection of human-robot humorous interactions in the joker project," in Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, pp. 348–354, IEEE, 2015.
[14]. D. Bertero, P. Fung, X. LI, L. WU, Z. LIU, B. HUSSAIN, W. CHONG, K. LAU, P. YUE, W. ZHANG, et al., "Deep learning of audio and language features for humor prediction.," in LREC, 2016.
[15]. A. Joshi, V. Sharma, and P. Bhattacharyya, "Harnessing context incongruity for sarcasm detection," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume2: Short Papers), vol. 2, pp. 757–762, 2015.
[16]. F. Barbieri and H. Saggion, "Modelling irony in twitter," in Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics, pp. 56–64, 2014. A. Reyes, P. Rosso, and T. Veale, "A multidimensional approach for detecting irony in twitter," Language resources and evaluation, vol. 47,no. 1, pp. 239–268, 2013.
[17]. E. Miltsakaki, L. Robaldo, A. Lee, and A. Joshi, "Sense annotation in the penn discourse treebank," in International Conference on Intelligent Text Processing and Computational Linguistics, pp. 275–286, Springer, 2008.