Intelligent Framework for Sentiment Analysis of Movie Reviews - A Comprehensive Survey

Eshwari G. Kulkarni
CSE dept., Walchand Institute of Technology, Solapur, 413006, India
Email: Kulkarni.eshwari73@gmail.com

Dr. Anita M. Pujar
CSE dept., Walchand Institute of Technology, Solapur, 413006, India
Email: arkulkarni10@gmail.com

Abstract—Sentimental research, also known as "Opinion mining," is concerned with evaluating attitudes and classify opinions based on reviews or comments. SA has emerged as a leading research area of Natural Language Processing (NLP). The role of Sentiment Analysis (SA) is to categorize people's views as positive or negative based on a specific statement or evaluation. Review and opinions play a significant part in determining the degree of happiness of consumers with a specific organization. Since human perceptions help to improve product quality, and the popularity or loss of a movie is determined by its ratings, there is an increased demand and need for building a good sentiment analysis model that classifies movie reviews.

The research article offers a study of emerging methods for categorizing sentimental analysis in general, as well as a summary of the key research problems posed in recent domain like sarcasm detection. We discovered that machine learning-based methods, such as supervised learning, unsupervised learning techniques, and Lexicon-based techniques, are the most commonly utilized. Research on different data sources has been supported using various algorithms such as Naive Bayes, SVM, and Ensemble classifier.

Keywords: Sentiment analysis, Natural Language Processing, machine learning approaches; Lexicon-based approaches, sarcasm detection

I.　INTRODUCTION

Today, the proliferation of the internet and social networking platforms such as Facebook, Instagram, Twitter, and others has resulted in a massive amount of consumer feedback and comments on specific facts of services.[3] People want to express their views, observations, ideas, emotions, about the services based on their perception.

Movie reviews are a significant way to assess a film's success. Although assigning a numerical/stars ranking to a movie informs us about its quantitative performance or disappointment, a selection of movie reviews provides us with a deeper qualitative perspective into various factors of the film. [1, 2] A textual movie review teaches us about the movie's positive and weak points, and a closer study of a movie review will inform us whether the movie satisfies the reviewer's standards in general.

II.　OVERVIEW OF SENTIMENT ANALYSIS

Sentiment analysis is a customer evaluation process for a website or product review or comment. Sentiment research is divided into three categories: positive comment, negative comment, and neutral comment respectively [1, 4, 6, 7, and 16]. The figure below demonstrates the entire sentiment analysis procedure, including how the feedback is categorized at each stage.
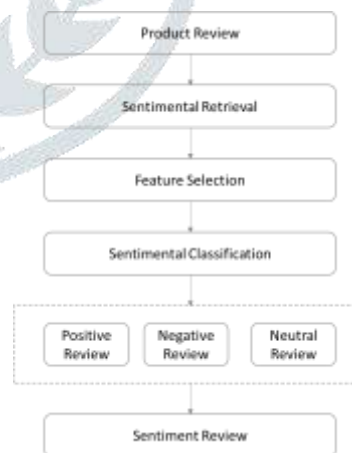

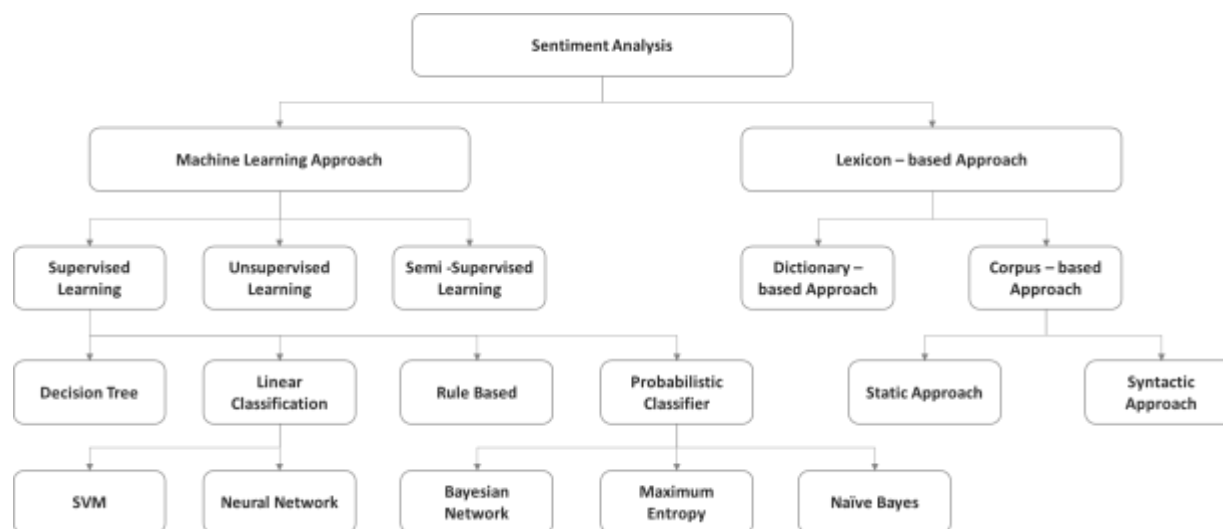**Figure: - Process of Sentiment Analysis on Reviews. [11]**

Figure: - Taxonomy of Sentiment Analysis Methods [11].

Different Classes of Sentiment Analysis

Sentiments can be classified into three categories .i.e. positive, negative, and neutral sentiments [5, 9, 15]

a. Positive Sentiments: There are some positive remarks regarding the goal under consideration. It is referred to as successful if optimistic emotions rise. In the case of product ratings, if the product has a high number of positive reviews, there is a high probability to purchase the products by the customers.

b. Negative Sentiments: These are the negative remarks regarding the goal under consideration. If negative emotions arise, the item is removed from the priority list. In the case of product opinions, where there are more bad opinions than positive reviews, nobody wants to purchase it.

c. Neutral Sentiments: There are neither positive nor negative statements regarding the topic. As a result, it is neither chosen nor ignored.

Feature Selection- There are four forms of feature selection. Syntactic Feature, Semantic Feature, Link-based Feature, and Stylistic Feature. The first two elements are mostly applied. The syntactic function makes use of word tags, patterns, words, and punctuation. Semantic features, on the other hand, focus on the association between words, signs, and symbols. [4] Linguistic semantics should be used to correctly understand human sentiments through vocabulary.

Sentiment Classification- [11] Classification is often divided into 2 "Unsupervised learning" and "Supervised learning" methods. Linear Classifiers: Classification algorithms for logistical regression/naive bays, support vector machine, decision trees, random forest, and neural networks

Sentiment Analysis Using Different techniques

Sentiment Analysis may be performed using the techniques mentioned below.

1. Machine learning technique involves both supervised and unsupervised learning [11]

- **Unsupervised Learning** is dependent solely on inputs and has no sense of goals. It just makes use of clustering.

- **Supervised Learning** identifies pre-specified goals that must be met, as well as the inputs. When data collection is experienced during decision-making, it is conditioned to provide significant results.

2. Lexicon-Based Approaches:

Based on the sentiment of every Individual word, the Lexicon approach builds positive or negative polarity, and based on this a dictionary is then developed. We may use a combined feature, such as a sum or average, to figure out a document's general sentiment.

3. Using Natural language processing and Sentiment Dictionaries [8, 11]-

NLP is a platform for programming natural language sentences that is aided by ontologies. NLP is a subfield of artificial intelligence dealing with the automatic representation and generation of human language.

i.   POS Tagging wise-

The key role of the Part-of-Speech (POS) is to senses or determine the word token belongs to noun, verb, or adjective. [3] The term is delegated to its syntactic purposes. The key aspects of the speech in English are noun, pronoun, adjective, determinant, verb, preposition, adverb, conjunction, and interjection.

ii.   NER Detection Wise-

Named Entity Recognition (NER) marks word sequences in a text that are the names of objects, such as people and companies' names, or gene and protein names [12]. Classify designated individuals into a text predefined divisions, such as names of people, organizations, places, time expressions, numbers, numerical amounts, ratios, and so on.

4.   Emoticon Wise Sentiment Analysis

The existence of at least a single emoticon is used as a guideline to choose Twitter sentiments for classification in this method. Based on the message's emoticons, the text's sentiment can be inferred [2]

Emoticons	Feeling	Sentiments
:) :-)	Happy	Positive
:(:-(	Sad	Negative
:D :-D	Very Happy!	Positive
:| =| :-|	Straight face no expression	Neutral

Table: Sample of Emoticons Used.

Collecting status and feedback for optimistic/Positive ':)' and negative ':(' emoticons creates the Emoticon data collection. Messages are categorized depending on Positive and negative sentiment. Use Unicode's to analyze sentiment.

Sarcasm detection:-

The term Sarcasm detection is one of the important challenge in many areas like sentiment analysis, NLP etc. In sentiment analysis, sarcasm affects the polarity detection process. Sarcasm detection plays a crucial role. It is the unconventional way of conveying a message, which conflicts the context. The research is going on sarcasm detection by different approaches like Feature Based approach, Rule-Based approach, learning approaches like machine learning and Deep learning, Dictionary-based approach, Hybrid approach, etc.

Challenges Faced In Sarcasm:-

Some researchers are proved that correctly detecting sarcasm in a sentence can increase the sentiment analysis of that particular sentence. The main challenges in sarcasm detection are followed:

- It is an easier task to detect sarcasm from speech when it is compared to the sarcasm detection from text. Because, the use of a certain tone of speech, body language, and facial expression can be useful while identifying sarcasm from speech.
- The quality of the data set also a crucial factor in sarcasm detection. The general nature of sarcastic sentences are ambiguous and doubtful. The presence of hashtags which indicates the sarcasm solves this ambiguity. But without hashtags, sarcastic sentences are complicated to understand.
- Feature selection is another important task in sarcasm detection. So, introducing new features and using them with already existing features can increase the accuracy of sarcasm detection. Selecting an appropriate new feature should involve deeper study abut semantic, punctuation-based and hyperbole features, etc.
- Generally, a sarcastic sentence delivers a negative sentiment by using only positive words. So sarcastic detection needs more additional features like semantic features, features related to text author, etc.

III. LITERATURE REVIEW

Sentiment Analysis:-

Eissa M.Alshari et.al (IEEE-2018) suggested a tool for extending the scale of opinions through learning the polarity of certain non-opinion terms in SentiWordNet-based vocabulary. System efficacy is tested using the Internet Movie Review Dataset. The outcome is positive, demonstrating that the proposed Senti2Vec approach may be more accurate than the lexical resource SentiWordNet.

Dataset: - Internet Movie Review Dataset (IMDB) - The dataset contains 100,000 movie reviews, 50,000 of which are labeled.

Method: - Senti2Vec method

Accuracy: - The accuracy of the Senti2Vec method is 85.4%

Key Observation: - From our observation we can conclude that the proposed method's success is promising, indicating that it could be more successful than the traditional SentiWordNet.

Future Scope: - In future work the more research is focused on the effect of the various distance measures employed in the Senti2Vec on the success of sentiment analysis

Cahyanti et al. (IEEE-2020) discuss the need for Sentiment Analysis (SA) to categorize movie reviews or comments as favorable i.e. positive or negative. Machine learning models, such as Support Vector Machine, are applied for sentiment analysis classification since they provide the strongest results (SVM). The study also used key feature selection of Term Frequency – Inverse Document Frequency (TF-IDF) as a method of weighting terms, which is then further coupled with the feature extraction using Latent Dirichlet Allocation (LDA) as a method of modeling topics that can solve the limitations of SVM.

Dataset: - Internet Movie Review Dataset (IMDB)

Method: - It will use Term Frequency - Inverse Document Frequency (TF-IDF) and Latent Dirichlet Allocation (LDA) to extract features before classifying with Support Vector Machine (SVM)

Accuracy: - The accuracy of the lexicon-based method is 82.16%

Key Observation: - The fusion of TF-IDF and LDA extraction with 240 topics and 82.16 percent effective results was the outcome of the best performance solution based on the whole test scenario. It is therefore simpler to increase system performance by integrating TF-IDF and LDA feature extraction.

Future Scope: - Future analyses can be expanded by introducing new feature extraction variants or by combining TF-IDF and LDA features to reduce processing time.

The key focus of Hameed Zabit (IEEE-2019) in this paper was to deal with publicly accessible datasets on movie analysis, namely Movie Review and Stanford Sentiment Tree (SST2), in a computationally effective way to classify the sentiment into positive and negative respectively. We used only a single approach of bidirectional long-term memory layer (BiLSTM) with a global max-pooling layer.

Dataset: - Pang and Lee [14] introduced the movie review (MR) dataset in 2005. It's a well-balanced dataset of 10662 binary movie ratings.

Stanford Sentiment Treebank (SST2) is an expansion of the MR Dataset that was implemented by R. Socher et al. [15] in 2013. It is an unbalanced dataset that contains 9613 binary movie ratings.

Method: - Bidirectional Long Short-Term Memory (BiLSTM)

Accuracy: - The accuracy of this method is 80.21%

Key Observation: - We concluded that our experiments, with complex structures, are comparable with recently published complex frameworks. Also, our solution needs minimal computing costs and may aid in general opinion categorization in real-time applications.

Future Scope: - This study's potential directions involve multilingual and multi-label view labeling correlated with real-time application of our enhanced framework.

Sarcasm Detection:-

Many researchers have studied and performed different methods for sarcasm detection. Most of the sarcasm detection works done on a textual data set. Relatively lesser works are done on the sarcasm detection on voice, images, and videos, etc., compared to the textual sarcasm detection.

Tepperman et al. was the first paper that touches on the issue of sarcasm. According to them, sarcastic speech violates at least one of Grice's maxims for cooperative dialogue ("Do not say what you believe to be false"). However, because of technological limitations at the time, they only used Natural

Language Processing (NLP) techniques to focus on online sentences that include the phrase "Yeah right" as the indicator of sarcasm.

Carvalho et al. stresses that sarcasm detection needs certain oral or gestural clues. In the case of text, these will be substituted by emoticons, punctuation marks, quotation marks and interjections. All these features are signs that textual comments or posts on social media are trying to find ways to imitate physical conversations.

In [23], sarcasm detection is based on context incongruity. The research unsurprisingly produced a good result. The popular sitcom "Friends" is used as the dataset to imitate real life conversation. Using an "all-words" approach and "incongruous words-only" approach, this work detected sarcasm by looking at sentence completion. All these existing research imitated what happens in a physical conversation with goal to understand sarcasm better.

Filatova discuss the terms irony and sarcasm. Sarcasm is deemed to be a subset of irony. It has positive literal meanings, negative intended meanings, and clear victims. This defines the identification of one type of sarcasm that is common in tweets: contrast between a positive sentiment and negative situation. A rather conflicting view is provided where they state that irony, satire, parody and sarcasm are similar, where the differences are just a matter of usage, tone, and obviousness.
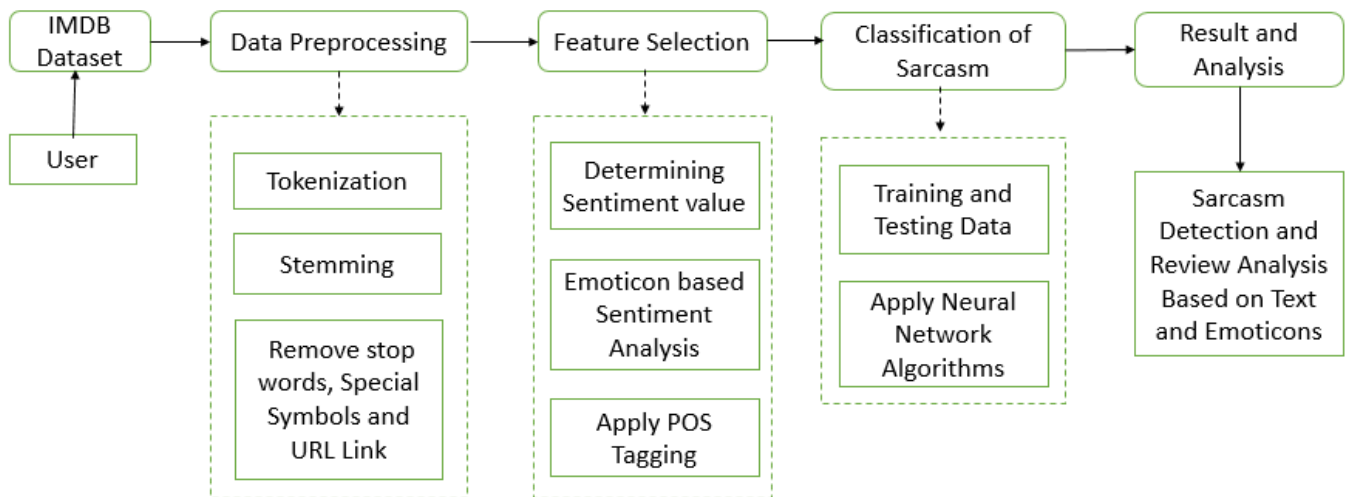

Figure: - Proposed System Flow

IV. PROPOSED IDEA

Step 1: - Data Collection-- Fetch and Upload Movie Review Data

The user collects Movie Review data from the IMDB dataset. The retrieved dataset includes the movie reviews or comments, as well as their hashtag comments. The gathered data includes information such as the title of the film, ratings

for the film, and movie reviews in the form of text and emoticons.

Step 2: - Preprocessing—Noisy data cleaning

It is one of the important steps before analyzing any dataset. Data preprocessing includes cleaning of the data such as tokenization, stemming, lemmatization and removal of stop words as well as eliminate special symbols and URL link.

These are the three basic data preprocessing techniques in the textual data. Apart from this, there are additional data preprocessing techniques; lowercasing all the text in the text data; it is helpful in case when a same word is written in the different case at different places. Noise removal and null value removal are also a step used of data preprocessing.

Step 3:- Feature Selection

This step includes analyzing different feature of the text, which will help in classification of the tweets, whether the tweets are sarcastic or not. The different feature extracted from the text are:

Feature 1: The initial feature that was determined from the text data was the sentiment value of the text. The process of sentiment analysis was done on the text and stored as separate feature. Sentiment value of the text retrieved are categorized on the basis of the value.

-1 Negative sentiment

0 Neutral sentiment

1 Positive sentiment

Feature 2: Once the sentiment value has been identified, the next feature that is extracted in whether there are emoticons present in the text or not. People always have a tendency to use emoticons while expressing any sarcastic comments.

Feature 3: The next extracted feature is POS tagging. POS tagging help in determining the use of various part of speech in the text such as verbs, nouns and adjective. With the help of POS tagging, four features are extracted which are nouns, verbs, adjectives and miscellaneous.

Step 4: - Classification of Sarcasm

Our Approach of the Sarcasm detection is based on Sarcasm headline dataset and/or Amazon review dataset that has the majority of the Sarcasm in it.
Once all the features are extracted, the next step is to apply NN Algorithms on the data to detect whether the reviews or comments, are sarcastic or not.

The dataset is initially split into three different sets, first is for the training of the NN model, the second set is for the testing the NN model and the third set is for the prediction and analysis with the help of the trained and tested model.

Step 5: - Final Outcomes:

Finally, this section include the Prediction of Sentiment Analysis outcomes using the procedures described above. Positive, negative, and neutral reviews are classified. This section also deals with the evaluation of different algorithms that were used in classification of the reviews into sarcasm and non-sarcasm respectively.

V. CONCLUSION

Sentiment analysis is one of today's most popular research topics among academics. The information gathered from online data sources such as blogs, microblogs, forums, and review sites, among others, has played an important role in expressing people's feelings, thoughts, emotions, and opinions about a specific topic, event, or issue.

Sentiment analysis aims to extract the perspective concealed behind the user's comment and identify the user's interests, priorities, and thoughts about the specific thing. This article identifies and analyzes sentiment analysis methodologies through a review and analysis.

The suggested approach includes critical phases for determining if a text's tone is positive or negative. This article would be useful to academics engaged in the fields of sentiment analysis and opinion mining. With the help of better sarcasm detection approaches, the decision making systems will have a better performance which will benefit the organization to understand the views, thoughts and opinions of their users.

REFERENCE

[1] Eissa M.Alshari, Azreen Azman∗, Shyamala Doraisamy, Norway Mustapha and Mostafa Alkeshr "Effective Method for Sentiment Lexical Dictionary Enrichment based on Word2Vec for Sentiment Analysis" 2018 IEEE-Fourth International Conference on Information Retrieval and Knowledge Management

[2] Rachana Bandana "Sentiment Analysis of Movie Reviews Using Heterogeneous Features" IEEE 2018

[3] Azilawati Azizan, Nurul Najwa SK Abdul Jamal, Mohammad Nasir Abdullah, Masurah Mohamad, Nurkhairizan Khairudin "Lexicon-Based Sentiment Analysis for Movie Review Tweets" IEEE 2019

[4] Fitri Eka Cahyanti, Adiwijaya, Said Al Faraby "On the Feature Extraction for Sentiment Analysis of Movie Reviews Based on SVM" IEEE 2020- 8TH International Conference on Information and Communication Technology (ICoICT)

[5] Zabit Hameed, Begonya Garcia-Zapirain, Ibon Oleagordia Ruiz "A computationally efficient BiLSTM based approach for the binary sentiment classification" IEEE- 2019.

[6] Rosy Indah Permatasari, M. Ali Fauzi, Putra Pandu Adikara, Eka Dewi Lukmana Sari "Twitter Sentiment Analysis of Movie Reviews using Ensemble Features Based Naïve Bayes" IEEE 2018

[7] Mais Yasen, Sara Tedmori, Amman, Jordan "Movies Reviews Sentiment Analysis and Classification" IEEE 2019 Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)

[8] Charu Nanda, Mohit Dua, and Garima Nanda "Sentiment Analysis of Movie Reviews in Hindi Language using Machine Learning" 2018-IEEE-International Conference on Communication and Signal Processing, April 3-5, India

[9] Malini R, Dr.Sunitha M.R, "OPINION MINING ON MOVIE REVIEWS" IEEE- 2019 1st International Conference on Advances in Information Technology.

[10] Suhariyanto, Ari Firmanto, Riyanarto Sarno "Prediction of Movie Sentiment based on Reviews

and Score on Rotten Tomatoes using SentiWordnet" IEEE-2018 International Seminar on Application for Technology of Information and Communication (semantic)

[11] Esha Tyagi AND Dr. Arvind Kumar Sharma "An Intelligent Framework for Sentiment Analysis of Text and Emotions – A Review" IEEE 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017)

[12] Tejaswini M. Untawale1, Prof. G. Choudhari2 "Implementation of Sentiment Classification of Movie Reviews by Supervised Machine Learning Approaches" IEEE-2019 Third International Conference on Computing Methodologies and Communication (ICCMC 2019)

[13] Yin Fulian, Wang Yanyan, Pan Xingyi, Su Pei "A Word Vector-based Review Vector method for Sentiment Analysis of Movie Reviews Exploring the applicability of the Movie Reviews" IEEE-2018 3rd International Conference on Computational Intelligence and Applications

[14] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," in Association for Computational Linguistics (ACL), 2005.

[15] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts, "Recursive deep models for semantic compositionality over a sentiment treebank," in Empirical Methods in Natural Language Processing (EMNLP), Seattle, Washington, USA, 2013.

[16] Wint Nyein Chan1, a, Thandar Thein2,b "A Comparative Study of Machine Learning Techniques for Real-time Multi-tier Sentiment Analysis" 1st IEEE 2018 International Conference on Knowledge Innovation and Invention

[17] Mitali Desai ; Mayuri A. Mehta "Techniques for sentiment analysis of Twitter data: A comprehensive survey" IEEE 2017.

[18] Parul Sharma and Teng-Sheng Moh "Prediction of Indian Election Using Sentiment Analysis on Hindi Twitter" 2016 IEEE International Conference on Big Data.

[19] Brandon Joyce ; Jing Deng "Sentiment analysis of tweets for the 2016 US presidential election" IEEE 2017.

[20] Jyoti Ramteke ; Samarth Shah ; Darshan Godhia ; Aadil Shaikh "Election Result Prediction Using Twitter sentiment Analysis" IEEE 2016.

[21] J. Tepperman, D. Traum, and S. Narayanan, "'Yeah right': sarcasm recognition for spoken dialogue systems.", INTERSPEECH, p. 4, 2006.

[22] E. Filatova, "Irony and Sarcasm: Corpus Generation and Analysis UsingCrowdsourcing," Lrec, pp. 392–398, 2012.

[23] A. Joshi, V. Sharma, and P. Bhattacharyya, Harnessing ContextIncongruity for Sarcasm Detection, vol. 2. 2015.

[24] P. Carvalho, L. Sarmento, M. J. Silva, and E. de Oliveira, "Clues fordetecting irony in user-generated contents," Proceeding 1st Int. CIKM Work. Top. Anal. mass Opin. - TSA '09, p. 53, 2009.