CSE/LING567 Project
Sarcasm Detection

Niharika Raut
Department of Computer Science
State University of New York at Buffalo
nraut@buffalo.edu

Fall 2018

Abstract

The purpose of the project is to understand the linguistic nature of sarcasm and do a comparative analysis of difference machine learning models that are used to detect sarcasm and propose a new model that is at par or better than the existing models. Sarcasm manifests with a contrasting theme, between positive and negative sentiments or literal and figurative meanings. The proposed model captures both the incongruity and the contrast of sarcasm text.

1 Introduction

Wikipedia defines Sarcasm as "a sharp, bitter, or cutting expression or remark; a bitter gibe or taunt". It is usually an ironic statement that is used to express contempt. The reason why most text classification or sentiment analysis algorithms fail in detecting sarcasm, along with a lot of human beings is the contrasting nature of it.Also due to the disparity between the written text and utterance of such a sentence.Sarcasm manifests with a contrast between positive and negative sentiments, or literal and figurative scenarios all in on sentence. The use of sarcasm is prevalent is social media and heavily skews the sentiment analysis and opinion mining results.

Consider a few examples of sarcastic statements:

1. I absolutely loved to be ignored!
2. Yeah. It's snowing sooner here in the Midwest. I think the next Ice Age is coming.
3. Yeah, right!

Joshi et al.,2015 explores the incongruity, i.e the juxtaposition of positive and negative terms and phrases. For example, in sentence one *love* and *ignored* richly captures the nature of sarcasm. A comparative analysis of different deep learning models is done on the benchmark dataset from the Internet Argument Corpus.

2 Previous Work

There is a lot of previous work using rule based and satistical approaches using unigrams and pragmatic features(such as emoticons, etc.) (Gonzalez-Ibanez et al., 2011; Carvalho et al., 2009; Barbieri et al.,2014), (b) extraction of common patterns, such as hashtag-based sentiment (Maynard and Greenwood,2014; Liebrecht et al., 2013), a positive verb being followed by a negative situation (Riloff et al., 2013), or discriminative n-grams (Tsur et al.,2010a; Davidov et al., 2010). There are sarcasm detection systems that mainly rely on deep and sequential neural networks (Ghosh and Veale, 2016; Zhang et al.,2016). In these works, compositional encoders such as gated recurrent units (GRU) (Cho et al., 2014) or long short-term memory (LSTM)(Hochreiter and Schmidhuber, 1997) are often employed, with the input document being parsed one word at a time. Since there is no word pair interaction, contrast or juxtaposition of contrasting situations isn't captured in the such models. Since simple sequential models don't capture long range dependencies, working with recurrent neural networks help capture the intra-sentence relationships. Models have been implemented that captures both the contrast and incongruity in sarcastic text using attention models. Such an method can be thought of as a self targeted co-attention model which has captured both long-range dependencies and word-word relationship.

2.1 Attention Models

The idea of neural attention is to soft select a sequence of words based on relative importance to the task at hand. Early innovations in attentional paradigms mainly involve neural machine translation (Luong et al., 2015; Bahdanau et al., 2014) for aligning sequence pairs. Attention is also commonplace in many NLP applications such as sentiment classification (Chen et al., 2016; Yang et al., 2016), aspect-level sentiment analysis (Tay et al., 2018s,2017b; Chen et al.,2017) and entailment classification (Rocktaschel et al.,2015).

3 Proposed Method

In this project, the new proposed method is to use a bi-directional lstm model with attention stacked with a convolution layer with max pooling to classify if a text is sarcastic or not. This model is then compared with other machine learning models (existing) for the same data set.

3.1 LSTM

Recurrent neural networks are networks with loops in them , allowing information to persist, making them appealing when the application needs to connect previous information to the present task. For example, a language model trying to predict the next word based on previous words. But when the application needs one to looks at information which is at a certain distance from the current information, RNNs are unable to connect this gap. LSTMs, which are capable of learning long-term dependencies, help in such situations. In sarcasm, the contrasting situations don't necessarily occur right next to each other, LSTMs help capture the long term dependencies.

Table 1: Data Statistics

Avg number of words in an instance	31 words
Avg length of sarcastic sentences	231 chars
Number of positive examples	1630
Number of negative examples	1630

3.2 Bidirectional LSTM

A bidirectional neural network connects two hidden layers of opposite directions to the same output. This form of generative deep learning enables the output layer to get information from backwards as well forward states simultaneously.The BiLSTM structure preserves global information and helps capture the long term dependencies faster as compared to LSTM as they increase the amount of input information available to the network. The use of providing the sequence bidirectionally was initially justified in the domain of speech recognition because the context of the whole utterance is used to interpret what is being said rather than sequential interpretation. This is applicable in detecting sarcasm, as the whole context of the text is required to determine the sarcastic nature of text.

3.3 Implementation

This project implements two bidirectional LSTM with attention models, with CNN with max pooling and without CNN. A bidirectional LSTM layer is initiated with outputs being averaged together. There are 128 hidden units for this layer and the weights for words are picked from the preloaded embeddings. Two attention layers with activations tanh and softmax were added.The hidden states were multiplied with the attention coefficients to find the weighted average. This weighted vector was passed to single layer dense network with relu as the activation for classification.

For another model, a convolution layer was added before the lstm layer. The filter width is 3 and number of filters f = 100.

4 Experiments and Results

4.1 Dataset

The Sarcasm Corpus V2 is a subset of the Internet Argument Corpus (IAC), including response text from quote-response pairs annotated for sarcasm. It contains data representing three categories of sarcasm: general sarcasm, hyperbole, and rhetorical questions.The data is in form of 'quote-response' pairs, where quote is a dialogic parent to the response. The sarcasm annotations are related only to response, but quote acts as a context. The subset of the data used for sarcasm detection has 3260 quote-response pairs with 1630 sarcastic examples.

4.2 Embedding

The models we ran with two different embedding, Glove for common crawl and Fasttext. The dimensions for both the embedding is d=300. The Fasttext embedding have embedding for approximately 1Million words, whereas the embedding used from glove have 2 millions words. It was observed that the models gave better results with golve embeddings.

4.3 Compared Methods

The proposed model was compared with the following algorithms.

1. NBOW is a simple neural bag-of-words baseline that sums all the word embeddings and passes the summed vector into a simple logistic regression layer.

2. LSTM is a vanilla Long Short-Term Memory Network. The size of the LSTM cell is set to d = 100

3. ATT-LSTM (Attention-based LSTM) is a LSTM model with a neural attention mechanism applied to all the LSTM hidden outputs.

4. CNN-LSTM-DNN (Convolutional LSTM + Deep Neural Network), proposed by (Ghosh and Veale, 2016), is the state-of-the art model for sarcasm detection. This model is a combination of a CNN, LSTM and Deep Neural Network via stacking. It stacks two layers of 1D convolution with 2 LSTM layers.The output passes through a deep neural network (DNN) for prediction

4.4 Results

Table 2: Experimental Results on IAC v2 Dataset

Model	Accuracy
NBOW	66.09
Vanilla LSTM	62.22
Attention LSTM	69.96
CNN-LSTM-DNN (Ghosh and Veale)	64.38
Bidirectional LSTM	70.18%
Bidirectional LSTM with Attention	**70.34%**
CNN Bidirectional LSTM with Attention	**71.17%**

5 Conclusion and Future Perspectives (if any)

Based on the intuition of feeding model sequences from both directions so that it captures the long range dependencies of contrast, a new neural network architecture for sarcasm detection is proposed. This network incorporated bidirectionally, enabling it to detect contrasting sentiment, situations and incongruity. The proposed model outperforms strong state-of-the-art baselines such as GRNN and CNN-LSTM-DNN.

References

[Ghosh and Veale] Ghosh, A. and Veale, T. Magnets for Sarcasm: Making Sarcasm Detection Timely, Contextual and Very Personal. Technical report.

[Tay et al.] Tay, Y., Anh Tuan ψ, L., Cheung Hui ϕ, S., and Su δ, J. Reasoning with Sarcasm by Reading In-between. Technical report.

[Yang et al.] Yang, Z., Xu, G., and Shi, B. Bidirectional LSTM-RNN with Bi-attention for reading comprehension. Technical report.
(author?) (Yang et al.) **(author?)** (Tay et al.); (Ghosh and Veale)