An approach for Aspect Based Sentiment Analysis using Deep Learning

CS 585, UMass Amherst, Fall 2016

Satya Narayan Shukla, Utkarsh Srivastava
satyanarayan@umass.edu, usrivastava@umass.edu

1 Outline

In this project, we present a deep learning approach for aspect based sentiment analysis (ABSA). Sentiment analysis is an important task in natural language processing and has a lot of applications in real world. The typical sentiment analysis is a process of classifying opinions expressed in a text as positive, negative or neutral. A more general task would be to predict the sentiments of each aspect mentioned in the text. In recent years with the growth of internet, social networking sites, discussion forums and blogs, e-commerce websites have gained immense importance. To enhance customer shopping experience, these websites generally provide platform for people to express their views about the product and its different aspects. There is a review and an overall score available for each product, but this doesn't provide the complete information. For example, two products might have same overall rating but with different unsatisfactory aspect opinions. This problem has inspired a new line of research on aspect level opinion mining since early 2000s [2]. Also, a major source of encouragement behind this project comes from the increasing popularity of ABSA task present in SemEval since 2014.

2 Literature Review

Paper [2] introduces us with the main motivation for finding a solution to ABSA by providing an insight into its application in AI. It proposes a trivial algorithm based on linguistic features to extract aspects present in a sentence. This proved to be a good starting point for further research in this area.

Paper [3] addresses two major problems with identifying semantic orientations of opinions expressed by reviewers on the features of the product. It presents a holistic approach that can accurately infer the semantic orientation of a review word based on the review context. It also addresses the problem of aggregating multiple aspects in a single sentence. Furthermore, it also considers implicit opinions building on the previous research works where only explicit opinions (expressed by nouns and adjectives) were considered. But given it uses a set of lexical rules, in the absence of a good coverage of such sets of rules, a lexicon based approach can only present limited accuracy.

Paper [5] gives a brief insight by evaluating the strengths and weaknesses of a variety of methods being applied for solving the problem at hand. By drawing a comparison between the 4 types of approaches (i.e. frequency-based, relation-based, supervised learning and topic analysis), it emphasizes on the requirement of either a large volume of data or a collection of large number of relation rules to mine appropriate content. Neither of the methods has generated extremely good results on real-world data. Further, the paper introduces us with the research challenges presented in the form of finding implicit aspects; mining multiple-aspects from a single sentence; cross-domain application of the same system; and the presence of highly unstructured text as input. This in turn motivates us to find a better alternate model to find improvements in both flexibility and accuracy.

Paper [4] provides an insight on how a convolutional neural network can be applied in dealing with sentence classification tasks in NLP. With the correct representation of the feature word-vector as input and slight tuning of model hyper-parameters, CNN provides better results than the conventional methods.

Poria et al. [1] presents the first deep learning approach to aspect based sentiment analysis. It uses a 7-layer deep convolutional neural network to tag each word in the review data as aspect word and non-aspect word. They have also comprehended a small set of grammar constructs to use in combination with neural nets to improve the accuracy. Here, features are word embedding along with their part of speech tags. This work shows that a deep CNN is more efficient than existing approaches for aspect extraction.

With the application of deep CNNs promising good results, we were encouraged to look deeper into the current research going on in this area and come up with a solution to validate the claim.

3 Approach

We plan to use a model architecture comprising of a deep neural network to solve this problem of aspect based sentiment analysis. We will try to tackle this task by following 2 steps with a deep learning model : Aspect Extraction; and finding the Sentiment associated with the extracted aspect. Additionally, we intend to improve the system accuracy by incorporating basic language/grammar patterns in correctly identifying features that could have been missed out by our main model. Given the train/test set has grammatically correct input text, we can leverage the use of a variety of linguistic constructs to improve accuracy. Finally, we plan to compare the accuracy of our method with prior metrics submitted for the corresponding SemEval task submissions.

The datasets (described below) have different train and test sets. As each set belongs to one specific product, we can directly divide approx. 10-15% of the training dataset into development set to tune the hyperparameters (e.g. number of hidden layers, number of neurons in each hidden layer) of the neural network using k-fold cross validation. Moreover, this a ready-made dataset i.e. our datasets already contain annotations.

Preliminary Experiment (for Progress Report):

We plan to build a "naive" deep net (which may not be optimally tuned), without incorporating any improvements with the use of linguistic patterns or data-preprocessing, and test it on a subset of the full data-set to validate the working of the proposed model from end-to-end.

This preliminary experiment of ours, that is capable of running the whole pipeline, can then be treated as a baseline algorithm to test and compare against any improvements made in terms of speed or accuracy on the full data-set.

4 Possible Datasets

1. In improving the bag-of-words representation, to use each word as a feature, we are planning to use "Google News Word Embedding" as a word-to-vector representation in a 300-dimensional

space. We chose this dataset as the number of dimensions are considerable and the vocabulary size is relatively large (approx. 3M words).

2. For the current problem statement at hand, we came across the following data sets that seemed relevant for our problem statement. Each data set is already annotated and contains separate files for training and testing.

   (a) SemEval 2014 - ABSA (Task 4): Each review is tagged with (i) aspect and (ii) polarity
   http://alt.qcri.org/semeval2014/task4/index.php?id=data-and-tools

   (b) SemEval 2015 - ABSA (Task 12): Each review is tagged with (i) aspect entity type (ii) aspect attribute type (iii) polarity
   http://alt.qcri.org/semeval2015/task12/index.php?id=data-and-tools

   (c) SemEval 2016 - ABSA (Task 5): Each review is tagged with (i) aspect entity type (ii) aspect attribute type (iii) phrase corresponding to aspect (iv) polarity
   http://alt.qcri.org/semeval2016/task5/index.php?id=data-and-tools

Currently we haven't finalized on the data-set representation that we will work on, out of the 3 sets defined above. But we are thinking of moving forward with SemEval 2015 ABSA dataset as neither it is as general as (a) nor does build up on more complex details like (c).

5 Scope

The dataset to be used in the project, as described above, pertains to annotated information for Laptop and Restaurant domains in the form of reviews. The scope of this project is to build and tune a deep learning model that can extract features from the provided review text and tag the corresponding sentiment associated. Additionally, it will try to identify multiple aspects and their associated sentiment present in a single sentence (for e.g.: I liked the ambience but the staff was not cordial -> "ambience"::positive and "staff"::negative). We also plan to provide an overall sentiment for the product under consideration, but there is no benchmark provided in the data-set to test it against.

The scope in turn is constrained by the data set we are using. The test accuracy is heavily dependent on the kind of review-entity being used for training (for e.g.: testing for accuracy on Restaurant reviews with a trained model on Laptop reviews will perform badly, given the features found in the 2 domains are quite different). The approach constraints itself in finding the category of the aspect only from the aspect tags provided in the dataset. Moreover, each sentiment or opinion about the feature is restricted to 3 class labels: positive, negative and neutral. The reviews are assumed to be grammatically correct.

6 Pre-existing Software Systems which can be used

(a) Implementation of the code in : Python

(b) Useful information on word relations from : WordNet/NLTK, Stanford CoreNLP (for parser and identifying dependencies) etc.

(c) Machine Learning functions and utilities in Python from : sklearn, theano (for CNN/Deep Learning etc.), numpy, matplotlib etc.

Other software systems will be added/removed as per requirements once we start the project.

References

[1] Poria S., Cambria E., & Gelbukh A. (2016), "Aspect extraction for opinion mining with a deep convolutional neural network," Knowledge-Based Systems, 108, 42-49.

[2] Hu M. & Liu B. (2006), "Opinion extraction and summarization on the web," Proceedings Of The 21St National Conference On Artificial Intelligence - Volume 2, 1621-1624.

[3] Ding, X., Liu, B., & Yu, P. (2008), "A holistic lexicon-based approach to opinion mining," Proceedings Of The International Conference On Web Search And Web Data Mining - WSDM '08.

[4] Kim, Y. (2014), "Convolutional Neural Networks for Sentence Classification," arXiv.org. 1408.5882

[5] Pratima More & Archana Ghotkar, "A Study of Different Approaches to Aspect-based Opinion Mining," International Journal of Computer Applications 145(6):11-15, July 2016.