



Natural Language Processing for Sentiment Analyses: A Survey and Experiments

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im Studiengang
Wirtschaftswissenschaften Fakultät der Leibniz Universität Hannover

vorgelegt von:

Name: David
[Redacted]

Vorname: Jan
[Redacted]

Prüfer: Prof. Dr. Michael H. Breitner
Betreuer: Rouven Wiegard

Hannover, den 02.01.2018

List of Contents

List of Tables	V
List of Figures	V
List of Abbreviations	VI
1 Introduction	1
1.1 Relevance and Motivation	1
1.2 Research Question	4
2 Theoretical Background of Sentiment Analysis	5
2.1 The Problem of Sentiment Analysis – Definitions	5
2.1.1 Definition of Necessary Terms	6
2.1.2 Objective and Tasks of Sentiment Analysis	9
2.1.3 Opinion Summary	11
2.1.4 Evaluation Measures – Precision, Recall, F-Measure and Accuracy	12
2.2 Classification Level	14
2.3 Sentiment Lexicon and its Issues	15
3 Literature Review	17
3.1 Identifying the Relevant Literature and Structuring the Review	17
3.2 Methods and Approaches in the Field of Sentiment Analysis	18
3.3 Existing Literature Reviews and Surveys	19
3.4 Aspect Identification & Extraction	21
3.4.1 Unsupervised Learning Methods	21
3.4.2 Supervised Learning Methods	25
3.4.3 Hybrid Methods	26
3.5 Aspect-Based Sentiment Evaluation	29
3.5.1 Dictionary-Based Approach.....	30
3.5.2 Supervised Machine Learning Approaches	32
3.5.3 Unsupervised Machine Learning Approaches	36
3.6 Results of the Literature Reviews	40

4	Aspect Extraction Use Case	42
4.1	Convolutional Neural Network	43
4.1.1	Step 1: Convolution	43
4.1.2	Step 1b: Rectifier Linear Unit Layer	45
4.1.3	Step 2: Pooling.....	46
4.1.4	Step 3: Flattening	47
4.1.5	Step 4: Full Connection	47
4.2	Comprehensive Overview of a CNN	50
4.2.1	Training Process – Back-Propagation.....	51
4.3	Aspect extraction - Convolutional Neural Network by Poria et al. (2016)	54
4.4	Training CNN for Sequential Data	54
4.5	Network architecture	55
4.6	Dataset	56
4.6.1	Word Embeddings	56
4.6.2	Amazon Embedding.....	57
4.6.3	Evaluation Corpora	57
4.6.4	Features and Rules	57
4.6.5	Linguistic Patterns	58
4.7	Experimental Results	59
5	Aspect Evaluation Use Case	60
5.1	Long Short-Term Memory	61
5.2	LSTM with Aspect Embedding (AE-LSTM)	62
5.3	Attention-Based LSTM	62
5.4	Attention-Based LSTM with Aspect Embedding (ATAE-LSTM)	64
5.4.1	Model Training	65
5.4.2	Experiment.....	66
5.4.3	Comparison with Baseline Methods	67
5.5	Qualitative Analysis	68
5.6	Further Benefits	69
	Discussion, Limitations and Recommendations	70
	Conclusion and Future Research	77
	References	79

1 Introduction

1.1 Relevance and Motivation

Natural Language Processing (NLP) is an interface between computer science, artificial intelligence and computational linguistics. It deals with the interaction between human (natural) language and computers. More precisely it addresses programming computers to handle and operate with large natural language corpora. Even though there has been work before, it is generally agreed, that the history of NLP started with the paper “Computing Machinery and Intelligence” published by Alan Turing in 1950. In 1954 started an experiment of fully automatic translation of more than sixty sentences from Russian into English. At that time, it was thought, automatic translation would be solved in the near future (Koehn, 2009). As machine translation is still far from being perfect, it was very optimistic. Until the 1980s most NLP systems were based on a complex set of hand-written rules. In the subsequent years machine learning came into focus within NLP, due to the fact, inter alia, of the increasing computational power. In the beginning machine learning algorithms were similar to the already existing hand-written rules, because predominating hard if-then rules were used. In the following years, statistical models came to the fore. These methods are more reliable with unknown inputs. Most recently the research is focusing on unsupervised or semi-supervised learning algorithms. These algorithms learn from data, which has not been annotated with the desired answers before.

Large companies have strong NLP research labs, including IBM, Microsoft, Sun, Google, Facebook and many other. There are various classes of machine learning algorithms applied to NLP tasks. The NLP field can be subdivided into many subfields based on the many aspects of human languages, like automatic machine translation, question answering or sentiment analysis. One useful application within this area is Google Translate from Google. It translates one text from one human language into another. Thus, it becomes possible, that non-English-speakers get access to an extensive amount of scientific information on the Web in English. Not only Google is using NLP, but also IBM. They developed a question answering system called Watson. IBM Watson can understand and answer questions “more effectively than any other question-answering system currently available” (Lee et al., 2016: 1). However, this paper is focusing on the field of sentiment analysis. Sentiment analysis identifies, extracts, quantifies, and studies subjective opinion within a written text, such as reviews and social media.

Opinions have an essential impact on nearly all human activities, and therefore they are key influencers to our behavior. Our beliefs and everyday choices are, for a great part, dependent on the view and evaluation of others. Therefore, we are asking friends and family members for their opinions about nearly everything before taking decisions. Since the internet has such a big impact on our daily life, we are not constricted anymore to just asking friends or family. People all around the world can share their experience, knowledge, and opinion via the internet and

future customers have access to detailed information about a product with its different aspects. Not only future customers but also companies are very interested in reviews. “These reviews are like gold mines to the companies” (Mehto et al., 2016: 1). Customer-generated feedback can help companies to develop their product, and expose what is being liked and disliked. It helps for the further implementation of recommendations by identifying customers’ preferences. This research area is called sentiment analysis (SA) or opinion mining (OM).

According to Liu (2015), sentiment analysis “is the field of study that analyzes people’s opinions, sentiments, appraisals, attitudes, and emotions towards entities and their attributes expressed in written text” (Liu, 2015: 1). Following Medhat, Hassan, and Korashy (1993), it is the “computational study of people’s opinions (...) toward an entity” (Medhat et al., 2014: 1093). The entity can either be an individual, event or a topic. Most likely these topics are captured by reviews, which is the reason why this paper will mainly focus on customer generated comments or reviews that imply positive or negative sentiments. According to Medhat the terms sentiment analysis and opinion mining are interchangeable and have a mutual meaning (Medhat et al., 2014), even though Tsytarau and Palpanas mentioned that they have slightly different notions (Tsytarau et al., 2012). Following them the aim of OM is “extracting and further processing users’ opinion about products, movies, or other entities”, while SA was “initially formulated as the NLP task of retrieval of sentiments expressed in texts” (Tsytarau et al., 2012: 481). Nevertheless, these two expressions will be used as synonyms in this paper.

Sentiment analysis is a new research area, which can be justified as prior to 2000 there was not much user-generated and opinionated text available in a digital form. Without these data research in sentiment analysis simply is not possible. Therefore, there had been barely any research before this year (Liu, 2012). Since the World Wide Web appeared, it has changed our way of communication and interaction between each other dramatically. Nowadays we have nearly limitless connectivity. As a result, it is possible to get any information from different sources immediately and read it straight away on our smartphones or computer screens. Furthermore, not only the possibility of getting information changed, even more importantly, the way of sharing information also changed entirely. Internet users are not only passively consuming available content, but also annotates this content and consequently actively generate new information. Most likely every internet user shares his ideas, experience, and knowledge with a large community. There are many possibilities to do so. Among them are Facebook, Twitter, Amazon, Instagram, Tumblr but also blogs, wikis, forums and social media in general where users share their information, feelings, and opinions with the rest of the world. As a result, the volume of user-generated and opinionated data is growing rapidly. Moreover, it seems like it will increase even more soon (Schouten, 2016). Consequently, these days costumers are not dependent solely on the product description of the seller, who obviously has incentives to con-

concentrate on the good parts of their products, moreover, there is such a large amount of opinionated data that it is very challenging, laborious and time-consuming, if not impossible, to read all of them. Nowadays, due to the rapidly growing size of data on the internet, it is already hard to access required information without the help of search engines. It gets even more complicated if available information are from different sources in a different format. Additionally, different issues, like spam and irony, are very challenging if a sentiment analysis is required. Therefore, it is very beneficial to have a well-operating sentiment analysis system.

To achieve a well-operating sentiment analysis system at aspect-level, this papers look at different approaches and methods in both, aspect extraction as well as aspect evaluation, to find the state-of-the-art methods.

However, first of all, chapter 2 involves the most important terms and definition within the area of sentiment analysis. Furthermore, the differences between sentiment analysis at different levels will be given. Additionally, the most used measures of the correctness of methods are defined.

After the basic knowledge is given, two systematic literature reviews, according to Webster & Watson, are performed in chapter 3, one of them for aspect extraction and the other for aspect evaluation. The implied papers are classified regarding their methods, which are supervised, unsupervised, and hybrid. The results of the literature reviews are two “state-of-the-art” methods for aspect extraction, the other for aspect evaluation. One of them uses a convolutional neural network, which will be discussed in chapter 4. The CNN is split into four steps, which will be explained in this chapter. On the basis of that, this paper will have a look at the approach of Poria et al. (2016) in detail.

Within the area of aspect evaluation, a method based on Long Short-Term Memory (LSTM) achieved state-of-the-art performance. More precisely, it is an attention-based LSTM with word embedding, proposed by Wang et al., (2016). Chapter 5 will focus on this paper in detail. In the end, a discussion will be given and the limitation of this paper. The paper will be completed by a conclusion and future research areas.

1.2 Research Question

An efficient aspect-based sentiment analysis system is a great benefit for both, the manufacturers as well as the consumers. Both have a high demand to know what is liked, respectively disliked about a specific product. The manufacturers may profit from this information because they can improve their product very precisely in the next production cycle, whereas the consumer gets a comprehensive overview of the experience of others and how they see this product. In this way, the consumer has the opportunity to inform in detail before purchasing. Recently a wide range of different approaches for aspect-based sentiment analysis have been proposed. All of them have different advantages and possibilities. Since these algorithms are the most important part of an accurate aspect-based sentiment analysis system, it has to be known which algorithm works best and is most suitable. Without this knowledge, it is not possible to develop an accurate aspect-based sentiment analysis system.

Therefore, the objective of this paper is to analyze algorithms for aspect extraction and aspect evaluation to find the most suitable algorithms. To find them, a literature review will be conducted, according to Webster & Watson. To compare the correctness of each approach, different measures are used.

Based on the literature review the state-of-the-art methods will be derived for both aspect extraction and aspect evaluation. Arising from the stated objectives this paper attempts to answer the following research question (RQ):

RQ: “Which are the state-of-the-art algorithms for aspect-based sentiment analysis? How good are they for aspect extraction and evaluation?”

Based on that, the state-of-the-art methods will be discussed in detail.

Conclusion and Future Research

This paper is about sentiment analysis, also called opinion mining. It presented some basic knowledge and gives in-depth information about specific techniques for aspect extraction and aspect evaluation. There is a real need for a well operating aspect-based sentiment analysis system in the industry, because every business wants to know what the consumers perceive about their products and services, and those of their competitors. This knowledge would gain a competitive advantage. The same also applies for consumers, because everyone wants to know the opinion and experience of existing users.

Due to the importance and the resulting benefit, there is a wide range of different approaches and methods. Objective of this paper was to survey the state-of-the-art methods in aspect extraction and aspect evaluation. In order to answer this objective, this paper gave an introduction and underlined the relevance of this topic, first. The second part was about the theoretical background of sentiment analysis, including definitions of important terms and different subtasks. Additionally, this section defined measures to compare the different approaches, which was done in the next part of this paper. The third part consists of two literature reviews, one for aspect extraction and the other for aspect evaluation. The aim was to find the state-of-the-art methods in both, with the result of a Convolutional Neural Network Method for aspect extraction proposed by Poria et al. (2016) and a Long Short-Term Memory Method for aspect evaluation proposed by Wang et al. (2016). The fourth part look at convolutional neural networks and gave in-depth information about the surveyed approach, whereas the fifth chapter dealt with the Long Short-Term methods of Wang et al. (2016). In the end of this paper a comprehensive discussion was given. Additionally, an alternative to a self-developed aspect-based sentiment analysis was given as well.

Beyond, a sentiment analysis system cannot just be used to analyze consumers' opinions and experience about products and services, it can be further used to analyze people's state of the mind using their post on social media platforms, like Facebook or Twitter, which could have a great benefit for the whole society. Furthermore, it could be used to prevent crimes and help the law enforcement authorities to detect potential terrorists with antisocial behavior, who might be able to cause damage to the society. In some cases, shooters announced their crimes in their social networks.

After all the research, which has been done so far, there are two promising research directions. First, there is an opportunity to develop a novel machine learning algorithm, which learns for a large amount of text data to mine domain-specifics on the one hand, but also to extract opinion independent of the domain.

Second, most of the current research is focusing on the English language. Of course, there are sentiment analysis systems in other languages. Nevertheless, these are mostly based on methods

developed for the English language, with some additions to handle language-specific issues. Thus, it is possible to develop a more specialized sentiment analysis system corresponding to the language. Besides, a more general sentiment analysis system might be developed, which works language independent and is, therefore, able to deal with different languages.