



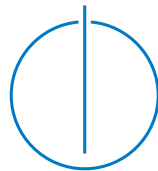
DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

# Deep Learning for Sentiment Analysis

Ankit Bahuguna





DEPARTMENT OF INFORMATICS

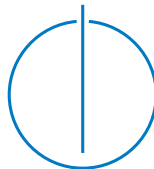
TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

## **Deep Learning for Sentiment Analysis**

### **Deep Learning Methoden für Sentiment Analysis Probleme**

Author:	Ankit Bahuguna
Supervisor:	Prof. Georg Groh
Advisor:	Prof. Georg Groh
Submission Date:	September 15, 2015



I confirm that this master's thesis in informatics is my own work and I have documented all sources and material used.

Munich, Germany,  
September 15, 2015

Ankit Bahuguna

## Acknowledgments

I would like to wholeheartedly thank my supervisor and adviser **Mr. Georg Groh**, who advised my Master's Thesis at TUM and encouraged me the whole time. He played a pivotal role in me continuing the study in the area of my interest and help me chase the state of the art systems with utmost frivolity. I would also thank, **Prof. Hinrich Schütze**, with whom, I started this study during an interdisciplinary project at Ludwig Maximilian's University (München) in April 2014 . His supervision and guidance helped me to get started with the new area of Deep Learning and its application to problems in Natural Language Processing. Additionally, I would thank my previous guide, **Prof. Pushpak Bhattacharyya**, (Indian Institute of Technology, Bombay) who introduced the field of Natural Language Processing and helped kick-start my interest in the area and encouraged me continuously to aim higher and helped me open my mind to look at problems from various perspectives.

In the end, I would also like to thank my family, who gave me this opportunity to study at Technical University of Munich, Germany. Without them, nothing of this would be possible.

# Abstract

Word vector representations or word embeddings, learned using deep neural network models prove extremely useful at capturing semantically similar words and phrases. Recent work in the area, utilize semantic word vector spaces and sentiment compositionality to classify sentiment better. One of the biggest constraints with these methods, is need of manually labeled sentiment for phrases in parse trees of sentences making its use challenging for languages with constrained resources.

Moreover, traditional methods do not efficiently leverage the use of semantic word vector spaces and word context in overall sentiment classification. In this report, we first present a broad overview of the problem, discuss traditional methods and then introduce Deep Learning based methods for Natural Language Processing which help in understanding the problem of Sentiment Analysis, better than traditional methods. Also presented is the analysis of concatenating various word vector representations which do not necessarily require phrase level labeling, comparisons of their results in various experimental configurations and showing improvement over traditional unsupervised baseline methods and state of the art results. In the same experiment, we also explore in-domain, out-of-domain and mixed-domain data towards training these word vectors and show how each one affects sentiment classification.

We also present promising future work, which may lead to better deep neural network based word representations for various NLP tasks.

# Contents

<b>Acknowledgments</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Information Gathering, Opinions and Sentiment . . . . .	1
1.2 Sentiment Analysis: An Introduction . . . . .	2
1.3 Applications . . . . .	3
1.3.1 Review-oriented Search Engine . . . . .	3
1.3.2 Websites which host Reviews . . . . .	3
1.3.3 Sub-Component Technology . . . . .	4
1.3.4 Business and Government Intelligence . . . . .	4
1.3.5 Applications across domains . . . . .	5
1.4 General Challenges . . . . .	5
1.4.1 Difficulty in Sentiment Classification? . . . . .	6
<b>2 Sentiment Analysis</b>	<b>8</b>
2.1 Supervised Machine Learning Techniques . . . . .	8
2.2 Unsupervised Machine Learning Techniques . . . . .	8
2.3 Domain Adaptation . . . . .	8
<b>3 Deep Learning</b>	<b>9</b>
3.1 Foundations: Neural Networks . . . . .	9
3.2 From Neural Networks to Deep Neural Networks . . . . .	9
3.3 Applications . . . . .	9
<b>4 Deep Learning and NLP: Word Representations and Language Modelling</b>	<b>10</b>
4.1 Language Modelling . . . . .	10
4.2 Word2Vec . . . . .	10
4.3 Glove . . . . .	10
4.4 Polyglot . . . . .	10
4.5 Paragraph Vectors . . . . .	10

<b>5</b>	<b>Deep Learning: Modern Apporaches</b>	<b>11</b>
5.1	RNTN : Recursive Neural Tensor Networks . . . . .	11
5.2	MSDA : Marginalized Stacked Denoising Auto Encoders . . . . .	11
5.3	CNN : Convoluted Neural Networks . . . . .	11
5.4	LSTM : Long and Short Term Memory Networks . . . . .	11
5.5	Comparative Evaluation . . . . .	11
<b>6</b>	<b>Concatenated Word Representations for Sentiment Classification</b>	<b>12</b>
6.1	Problem . . . . .	12
6.2	Experiments . . . . .	12
6.3	Observations . . . . .	12
6.4	Result . . . . .	12
<b>7</b>	<b>Moving from Text to Speech: Emotion Recognition</b>	<b>13</b>
7.1	Conventional Methods . . . . .	13
7.1.1	GMM: Gaussian Mixture Models . . . . .	13
7.1.2	I-Vectors: Total Variability Matrix . . . . .	13
7.2	Deep Neural Networks for Emotion Recognition . . . . .	13
<b>8</b>	<b>Conclusion</b>	<b>14</b>
8.1	Section . . . . .	14
8.1.1	Subsection . . . . .	14
8.2	Section . . . . .	14
	<b>Glossary</b>	<b>16</b>
	<b>Acronyms</b>	<b>17</b>
	<b>List of Figures</b>	<b>18</b>
	<b>List of Tables</b>	<b>19</b>
	<b>Bibliography</b>	<b>20</b>

# 1 Introduction

*Romance should never begin with sentiment. It should begin with science and end with a settlement. — Oscar Wilde, An Ideal Husband*

## 1.1 Information Gathering, Opinions and Sentiment

"What other people think", has been a very important element in the overall decision making process. This is more relevant in the modern era with the explosion of web and internet based services where the people have a new platform where they can voice their opinions and discuss about day to day things. With the advent of this platform, the subjective information too has grown extremely rapidly over the years.

It is interesting to note that social opinion of users impact a large corpus of audience. A simple case example is, that if a product is negatively perceived in the social circles of the internet, it will not be taken positively by a new user who is interested in purchasing the product. Hence, /this also is a tool for organizations to understand how the product is performing in the public. This, has given rise to a new area of study known as social media monitoring and analysis.

In the modern day, organizations have made significant investments to process this vast amount of data and make sense of the important information contained in it, so that they can leverage their products and services and improve their overall user experience or increase chances of making more profit, by identifying key areas which are causing an indirect impact on the overall profit margin. In the modern world, the organizations which makes the best use of data is the clear winner and it also is an opportunity for the researchers working in the area to aggressively pursue new techniques and algorithms which beat the current state of the art systems, giving the organizations slight edge over their competitors, which can in turn bring millions of dollars in revenue.



## 1.2 Sentiment Analysis: An Introduction

It's well said that the fundamental difference, between a human and a computer is that, the computer doesn't perceive or express emotions. If someday, this barrier can be overcome, then it will be very hard to distinguish a human conversation from a machine conversation. Thus, the broad goal for the study of Sentiment Analysis, is making computers recognize and express emotions.

First, we present some of the common terms, which are regularly used in context of Sentiment Analysis, viz., **opinion**, **sentiment** and **subjectivity** in text. These can be defined in the following manner:

**opinion** : *A view or judgment formed about something, not necessarily based on fact or knowledge.*

**sentiment** : *A view or opinion that is held or expressed.*

**subjectivity** : *Refers to how someone's judgment is shaped by personal opinions and feelings instead of outside influences.*

The term opinion mining appears in a paper by Dave et al. [69] that was published in the proceedings of the 2003 WWW conference. Interestingly, an ideal opinion-mining tool would "process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinions about each of them (poor, mixed, good)."

The history of the phrase sentiment analysis parallels that of "opinion mining" in certain respects. The term "sentiment" used in reference to the automatic analysis of evaluative text and tracking of the predictive judgments therein appears in 2001 papers by Das and Chen [66] and Tong [296], due to these authors' interest in analyzing market sentiment. It subsequently occurred within 2002 papers by Turney [298] and Pang et al. [235], which were published in the proceedings of the annual meeting of the Association for Computational Linguistics (ACL) and the annual conference on Empirical Methods in Natural Language Processing (EMNLP). These papers really increased the popularity of the term "Sentiment Analysis" in the Natural Language Processing research community. A number of papers mentioning "sentiment analysis" focus on the specific application of classifying reviews as to their polarity (either positive or negative), a fact that appears to have caused some authors to suggest that the phrase refers specifically to this narrowly defined task. However, nowadays many construe the term more broadly to mean the computational treatment of opinion, sentiment, and

subjectivity in text.

If we look in a broad view, the terms sentiment analysis and opinion mining roughly imply the same field of study, which itself can be attributed as the sub field of subjectivity analysis. For the sake of simplicity, we will stick to "sentiment analysis" throughout this work.

## 1.3 Applications

*Sentiment without action is the ruin of the soul.* — Edward Abbey

There are a number of application areas of this study, although, in the introduction we have mentioned reviews related to websites and products but in general there are many other possibilities. It is because of all the possible applications, there are a good number of small startups to large organizations which have dedicated teams who look into this subject matter and have it as part of their mission statement.

### 1.3.1 Review-oriented Search Engine

A classic example can be a review oriented search engine. Consider a product search engine which also utilizes the reviews on the products made. And ranks the result based on the positive polarity of the products. Thus, the products which has maximum positive product reviews are ranked higher than the ones which receives lower reviews. This will help a client who is user of the search engine to have high confidence about the kind of products returned. Such a search engine will in a way give an assurance to the customer that the best reviewed product which are socially favorable and appreciated are shown first to the user, hence making it easy to make the final choice of eventually purchasing the product.

### 1.3.2 Websites which host Reviews

There is a growing number of opinion hosting websites, for example, epinions.com; along with the booming e-commerce industry which showcase products and offers, there are millions of reviews generated every week. Thus, they are a primary playground for sentiment analysis. These can be a good benchmark for the companies to evaluate how the product is performing in the market. By performing a comprehensive analysis of the reviews which are hosted on these websites, the companies get a fair idea about the pros and cons about their product and also about the competitor, and a general sentiment about the public perception can help organizations make important

decisions about how to go or not go about the future iterations of the projects. Some organizations, perform a real time analysis, so that if a feature is causing a universal cry, resources can be quickly put together into immediately fixing the issue and thus leveraging overall user experience, leading to a longer user loyalty retention.

### 1.3.3 Sub-Component Technology

Sentiment analysis can also be used a sub component leveraging the working of another intelligent computational system. One possibility is the augmentation to a *recommendation system*. Thus, a recommendation system will utilize a sentiment analysis system, which can then not recommend products which receive a lot of negative feedback.

*Detection of overheated or agnostic language* in emails or any other form of communication is another use of subjectivity detection and classification.

In *online systems and display advertisements*, where its required to display advertisements relevant to the content of the web-page and some may contain sensitive material, it always better is such sensitive websites can be recognized in advance. A more sophisticated system can display an advertisement when more positive statements in content are discovered and nix the same when negative statements begin to appear.

Also in *information extraction* which is heavily based on objectivity in text, can leverage subjectivity detection and discard any subjectivity content, thus improving overall quality of information extracted.

*Question Answering systems* can also be improved, as the opinion oriented questions can be handled differently. These questions differ in the subjective sense and a pre-analysis of the same can be useful to generate better answers. Similarly, *summarization* can also be made useful in the context of the multiple views points.

In the area of citation analysis, where potentially, one can detect whether the citation done by the researcher in his scientific work is a support or a rebuttal of the cited work.

### 1.3.4 Business and Government Intelligence

Sentiment Analysis is used effectively as a business intelligence tools. The reviews about the product can be indicative directly what is right or wrong about the product or it may be useful to mine the reviews from all the general domains like opinion websites,blogs, public forms and product pages and summarize each f the review entities corresponding to each product thereby giving a consensus of all the key points relevant in the eyes of the users with respect to the product. The data can also determine what are the current trends in the given business category and can be used to leverage the key insights which in-turn can benefit sales.

With a view point of government Intelligence, such a tool can be used to monitor hostile behavior in a communication. This form of government intelligence can help in also beefing up security if a known threat is determined in advance.

### 1.3.5 Applications across domains

Sentiment Analysis has not only been cherished by computer and data scientists but also in the political circles of the government. It has actually turned out to be a golden tool in political elections. Knowing what the people think about the candidates in a general election can allow a party to choose the right candidate and also work on its shortcomings in due time, to win the ever significant final elections.

Moreover, There are now e-portals which are facilitating *eRuleMaking*, where a new policy is first open to the public and then based on the various opinions of the general public, the decision to move forward and cement the proposed rules into a law or updating its various sections is taken into consideration. This has proved to be an effective tools to iron out any problems which may exist in a proposed law even before its presented to the democratic authority. Thus, saving both time and resources.

## 1.4 General Challenges

Often a simple question comes to mind, when discussing Sentiment Analysis: *How is it different from classic text mining and fact based analysis?*

Traditionally, text categorization involves categorizing text into categories. These categories can be topics which are either predefined or looked up as more data is processed. These categories can also be user or application specific. And, for a given task, we might be dealing with wither a two class or binary classification or a multi-class classification (say, among thousands of classes). In contrast, with sentiment analysis, we are usually focused with 3 classes: positive, negative or neutral (coarse grained) or, 5 classes: positive, slightly positive, negative, slightly negative and neutral (fine grained). Thus, the number of classification classes are fairly limited and it generalizes well to many information domains and users. moreover the topics or categories can sometimes be highly non-correlated whereas the sentiments classes are often opposite in nature (in case of binary: positive- negative classification.)

Also, there is fundamentally a difference when it comes to answering opinion oriented questions vs fact based questions. The traditional information extraction methods which work well for a number of facts based templates. An opinion oriented information extraction method too will be a generalized version of the traditional system, as they are focused on similar fields of an opinion expression (holder, type and strength) irrespective of the topic. Although, it may seem the propositions presented

above may make the task of sentiment analysis relatively easy, but its far from the actual truth. We will learn next about the difficulties in developing sentiment analysis systems, as compared to traditional fact based text analysis systems.

### 1.4.1 Difficulty in Sentiment Classification?

First, we start with a simple scenario and a connected question: Given, we have an opinionated sentence in some given language, our task is to classify it into one of positive, negative or neutral classes. How hard is this overall task? Let's start with some examples:

1. It was a great restaurant.
2. It should have been a great restaurant.
3. The restaurant was great in that it will make all future meals seem more delicious.
4. Despite a pleasant experience I can't support the many reviews that it was a great restaurant.

First sentence is a positive sentence, implied by "great restaurant", The rest of the sentences are all negative. But there exists a varying degree of negativeness among them. The second sentence with the phrase "should have been" implies the desired result. Which suggests that the restaurant was expected to be better than it actually was. The third sentence is a classic case of sarcasm. This is hardest to spot even for humans; On the first glance we might be mistaken to consider this as positive. Sentence 4, started with a positive vibe "pleasant experience", but the reviewer then retracts his support to all the positive reviews made about the restaurants by others, thus making his overall review negative.

These four sentences just provide a glimpse to how hard is to analyze sentiment. Moreover, apart from detecting polarity in subjective sentences, there is something, which is known as beyond polarity analysis, where the task is to detect more expressive features like anger, happiness, sadness and relaxation in text, these too are very hard to predict and are highly context driven. The context may be based on the expression in the sentence or say, in a conversation among two individuals, their relationships between them. For example, the sentence "Your're a Liar!", may mean either positive or negative given the two individuals and their relationships. Like, a candid conversation between a couple, one may take this in positive sense (not always!), but among politicians it certainly means seems negative.

Then there is also a problem of **entity level sentiment** and **domain specificity**. The sentiment of a reviewer of a product may be dependent on individual entities composing a product. Say, for example: A user is writing a review about a **digital camera**:

*"I consider this is as a decent looking camera. The lens is of high quality and performs well in low light condition. Although, the flash is timely, it just seems to fade away the colors. There is*

*also the issue with the battery which only takes about 100-150 photos on full charge, making periodic recharging necessary. Last, the price is the reason, I won't recommend this camera to anyone, because there are other better cameras for the same specs and lower price which are better buy in the category. "*

In the above review, the user is highly positive about the looks and lens of the camera. Thus, the review begins on a positive note and then the user comments about the flash, battery and the price, eventually making the review a negative one. This example is essentially useful in understanding the complexity involved in extracting the overall sentiment of a long review. Such long reviews are common in the real world and thus, this task is harder than it seems. Often, in such cases, it's best to understand what entity or sentence contributes the most to overall sentiment. Although, this is easier said than done!

One may, also find that the specific qualities of the entities, which seem good for a certain product may not be good for another. This is where domain specificity is required. For example,

- **"unpredictable"** may be negative in a car review, but positive in a movie review [Turney, ACL2002]
- **"cheap"** may be positive in a travel/lodging review, but negative in a toys review

## **2 Sentiment Analysis**

### **2.1 Supervised Machine Learning Techniques**

### **2.2 Unsupervised Machine Learning Techniques**

### **2.3 Domain Adaptation**

## **3 Deep Learning**

### **3.1 Foundations: Neural Networks**

### **3.2 From Neural Networks to Deep Neural Networks**

### **3.3 Applications**



## **4 Deep Learning and NLP: Word Representations and Language Modelling**

### **4.1 Language Modelling**

### **4.2 Word2Vec**

### **4.3 Glove**

### **4.4 Polyglot**

### **4.5 Paragraph Vectors**

## **5 Deep Learning: Modern Approaches**

**5.1 RNTN : Recursive Neural Tensor Networks**

**5.2 MSDA : Marginalized Stacked Denoising Auto Encoders**

**5.3 CNN : Convoluted Neural Networks**

**5.4 LSTM : Long and Short Term Memory Networks**

**5.5 Comparative Evaluation**

## **6 Concatenated Word Representations for Sentiment Classification**

### **6.1 Problem**

### **6.2 Experiments**

### **6.3 Observations**

### **6.4 Result**

## **7 Moving from Text to Speech: Emotion Recognition**

### **7.1 Conventional Methods**

#### **7.1.1 GMM: Gaussian Mixture Models**

#### **7.1.2 I-Vectors: Total Variability Matrix**

### **7.2 Deep Neural Networks for Emotion Recognition**

# 8 Conclusion

## 8.1 Section

Citation test [Lam94].

### 8.1.1 Subsection

See Figure 8.1.



Figure 8.1: An example for a figure.

## 8.2 Section

See Table 8.1, Figure 8.2, Figure 8.3, Figure 8.4.

Table 8.1: An example for a simple table.

A	B	C	D
1	2	1	2
2	3	2	3

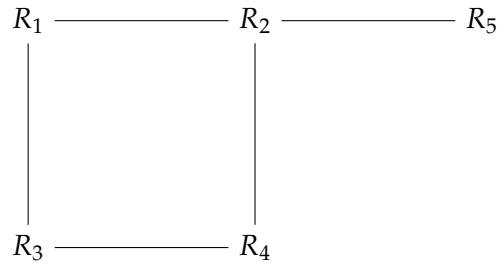


Figure 8.2: An example for a simple drawing.

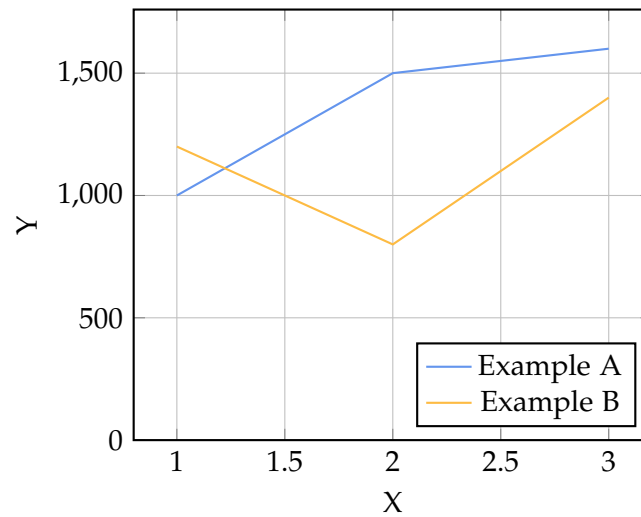


Figure 8.3: An example for a simple plot.

```
SELECT * FROM tbl WHERE tbl.str = "str"
```

Figure 8.4: An example for a source code listing.

# Glossary

**computer** is a machine that. . .

# Acronyms

**TUM** Technische Universität München.



## List of Figures

8.1	Example figure . . . . .	14
8.2	Example drawing . . . . .	15
8.3	Example plot . . . . .	15
8.4	Example listing . . . . .	15

# List of Tables

8.1	Example table . . . . .	14
-----	-------------------------	----

# Bibliography

- [Lam94] L. Lamport. *LaTeX : A Documentation Preparation System User's Guide and Reference Manual*. Addison-Wesley Professional, 1994.