**Project Name**



Semantic Web Group Project

# **TDT4215**

**Web-intelligence**

April 2016

[Storvold, John-Olav](https://ntnu.itslearning.com/Person/show_person.aspx?PersonId=95926&Customer=105)

[Sellæg, Jørgen Lien](https://ntnu.itslearning.com/Person/show_person.aspx?PersonId=405940&Customer=105)

[de Miranda de Matos Lourenco, Ines](https://ntnu.itslearning.com/Person/show_person.aspx?PersonId=408705&Customer=105)

[Tidemann, Nicholas](https://ntnu.itslearning.com/Person/show_person.aspx?PersonId=109247&Customer=105)

**Abstract**

The aim of this report is to take a better look at sentiment analysis. The proposed project consists on the creation of an application that works as a movie reviews database. It is possible to search for a specific review, create a new one, and know that scores associated with them.

[

Include this in your report:

● Motivation for your project topic

● Explain the structure of your system and how relevant parts of the curriculum have been applied

● Implementation and usage of application

● Conclusion and remarks Make sure to include all your decisions in the report. Why you chose to go for that exact method instead of others, etc. The report should also include relevant information (ontology, specific algorithms you’ve created, etc.)

]

Contents

**Introduction1**

**Theory4**

Sentiment analysis5

Methods6

**Tools4**

Escrever título do capítulo (nível 2)5

**Product description4**

Goals5

Methods used5

**Results and discussion4**

**Future work4**

**References4**

# Introduction

The internet became a very important platform where people can participate, express their opinions and views, share emotions and comments about a lot of everyday life aspects. It became a huge platform of opinions and sentiments’ change. And due to this the concept of Sentiment analysis was introduced. The applications are countless, from obtaining feedback from a certain product to important information regarding public opinion.

This analysis ends up being a win-win situation for both consumers and producers about products and services. If we know that a lot of people liked a certain product, then it must have some features worth paying for. We can also know which are its strengths and weaknesses, and the best and worst characteristics. For the producers it is beneficial in the sense that they can find out what should be improved in the future.

When it comes to the reviews, the idea of scoring them came up so that instead of reading all thousands of reviews of a product any user can just concern about the scores.

Given the possibility of choosing any theme inside in the semantic web topic, these are the reasons why we decided to mainly focus our attention on sentiment analysis.

We found a way to apply them to movies reviews because it seemed to be the most difficult of several domains for sentiment classification, reporting an accuracy of 65.83% on a 120- document set.

# Theory

**2.1 Sentiment analysis**

It is important to distinguish between what is a rational evaluation (facts) and what is an emotional one (subjective). The base to do sentiment analysis is the interpretation of feelings. These can not only emotions but also attitudes and opinions.

We can want to classify users, texts, sentences/paragraphs, words, etc.

It is increasingly more important to have techniques that allow us to capture complex contextual phenomena. There are a lot of techniques that have been implemented in this field. The focus is now moving towards unsupervised approaches because of the huge amount of opinionated data.

A substantial number of sentiment analysis approaches rely greatly on an underlying sentiment (or opinion) lexicon. A *sentiment lexicon* is a list of lexical features (e.g., words) which are generally labeled according to their semantic orientation as either positive or negative. They can be Semantic Orientation (Polarity-based) Lexicons or Sentiment Intensity (Valence-based) Lexicons.

**2.1.1 Term-counting method**

Lexicon based

¡ Unsupervised ¡

Use of sentiment lexicon

The term counting method consists on

While a few tools out there rely on scoring sentiment words based on their “polarity” (i.e. their position on a scale between positive and negative), accurate sentiment analysis doesn’t look at only keywords or individual words alone, because of **semantics**.

* (+2) emotional positive
* (+1) rational positive
* (0) neutral
* (-1) rational negative
* (-2) emotional negative
* The first approach (term-counting method) is to count positive and negative terms in a review, where the review is considered positive if it contains more positive than negative terms, and negative if there are more negative terms
* The semantic orientation of terms and phrases can be used to determine the sentiment of complete sentences and reviews.
* We augment this method by taking contextual valence shifters into account

There seems to be some relation between positive words and postive reviews ● Can we come up with a set of keywords by hand to identify polarity?

The lexicon-based approach involves calculating orientation for a document from the semantic orientation of words or phrases in the document

we use dictionaries of words annotated with the word’s semantic orientation, or polarity

First, a list of adjectives and corresponding SO values is compiled into a dictionary. Then, for any given text, all adjectives are extracted and annotated with their SO value, using the dictionary scores. The SO scores are in turn aggregated into a single score for the text.

aspects of the local context of a word need to be taken into account in SO assessment, such as negation (e.g., not good) and intensification (e.g., very good), aspects that Polanyi and Zaenen (2006) named contextual valence shifters. Research by Kennedy and Inkpen (2006) concentrated on implementing those insights. They dealt with negation and intensification by creating separate features, namely, the appearance of good might be either good (no modification) not good (negated good), int good (intensified good), or dim good (diminished good).

t this lexicon-based method performs well, and that it is robust across domains and texts

t individual words have what is referred to as prior polarity, that is, a semantic orientation that is independent of context; and that said semantic orientation can be expressed as a numerical value.

**2.1.2 Infinitive form**

The method of counting positive and negative terms has a big boost if we start by transforming the terms from their current form into their base forms (lemmas). This increases the efficiency of the analysis due to the fact that without it we need to have all the word’s forms in our dictionary.

This is for example the case of words such as *good*, *better* and *best.* Instead of having the three forms in our dictionary, we can simply tell the program that *better* and *best* are only different forms of the word *good*. What happens then is that when any of those words appear on the text to be analyzed, they get the same score as the infinite form of the verb (the only form that needs to exist on the dictionary).

**2.1.3 Valence shifters**

Valence, or in this case sentiment shifters, are very important to take in consideration because they change the polarity of a term. In this case, the polarity of a term represents whether it is positive or negative.

*Not, no, none, never, nobody, nowhere,* neither are some examples of sentiment shifters. If any of them appear before the word, then it should change the polarity of that word’s score.

“This is *not* a good movie”. If to the adjective *good* it is given a score of +3, then because of the presence of a sentiment shifter before the score would change to -3.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sentiment shifter | + | positive word | = | negative word |
| negative word | positive word |

**2.1.4 Intensifiers and diminuishers**

|  |  |
| --- | --- |
| Intensifier + positive word = more positive word | Word’s score + % of intensifier |
| Intensifier + negative word = more negative word | Word’s score - % of intensifier |
| Diminisher + positive word = less positive word | Word’s score - % of intensifier |
| Diminuisher + negative word = less negative word | Word’s score + % of intensifier |

Very: Intensifier with 100%

Less: 50%

e.g: “I find this tool good and useful”

good = +3

useful = +2

So the score of the sentence is 3 + 2 = 5

e.g: “I find this tool good but **less** useful than yours”

less = 50%

So the score of the sentence is 3 + 50%\*2 = 4

**2.1.5 Connectives and modals!!!!!??????????????**

**2.1.6 Problems**

The analysis methods are pretty limited when it comes to some kind of situations.

First of all, it doesn’t care about the context.

“This movie is *excellent* if you want to waste your money.”

Another problem is that a word, such as an adjective, with a positive or negative meaning, doesn’t always have and associated sentiment. It corresponds to a sentiment ambiguity.

“Is this a *good* movie?”

The opposite can also happen. A sentence without any sentiment word can also express sentiments. IN this case it express a clearly negative sentiment.

“This movie shouldn’t have been seen that much”

Irony and sarcasm are not recognized by the algorithms.

“I’m really *happy* that n”

1. **Sarcasm**: a positive or negative sentiment word can switch sentiment if there is sarcasm in the sentence (e.g. “Sure, I’m *happy* for my browser to crash right in the middle of my coursework”).
2. **Language**: a word can change sentiment and meaning depending on the language used. This is often seen in slang, dialects, and language variations. An example is the word “*sick*“, which can change meaning based on context, tone and language, although clear to the target audience.

**~**

**2.2 Recommender systems??????**

A sentiment predictor system can be naturally considered to aid a recommender system. The recommender system will not recommend items that receive a lot of negative feedback.

# Tools

* Mongodb
* We used AFFIN lib,

A list of 2477 words that are classified between -5 and 5. The word ache is classified as -2, adore is 3, amused 3, amazing is 4, superb 5, breathtaking 5, outstanding 5, catastrophic -4, fraud -4

|  |  |  |
| --- | --- | --- |
| The movie is cool, I love it | 4 | Very positive sentiment |
| The movie is uncool, I hate it | -3 | Very negative sentiment |
| The movie is not cool, I do not love it | 4 | Very positive, same score as the first example. Sentiment is not aware of grammar or negation |

In order to correct this we tried to introduce sentiment shifters and intensifiers.

* Vader

Polarity

Intensity

* Stack

Orientation of words

* General Inquirer ?
* Linguistic Inquiry and Word Count ?
* Wordnet ?
* SentiWordNet

Orientation of sentences

* The PMI-IR method

# 4 Product description

**2.16. Goals**

**2.16. Methods used**

Expand the contraction, normalized the text, simplified

1. We begin by constructing a list inspired by examining existing well-established   
   sentiment word-banks (LIWC, ANEW, and GI).

we next incorporate numerous   
lexical features common to sentiment expressio

- a full list of Western-style emoticons, for example, :-) denotes a smiley face   
and generally indicates positive sentiment)   
- sentiment-related acronyms and initialisms (e.g., LOL and WTF are both examples of   
sentiment-laden initialisms)   
- commonly used slang with sentiment value (e.g., nah, meh and giggly).

- punctuation (!)

-caps HAPPY

This process provided us with over 9,000 lexical feature candidates

Then we looked for candidates to sentiment expressions using a  wisdom-of-the-crowd13 (WotC) approach to acquire a valid   
point estimate for the sentiment valence (intensity) of each context-free candidate   
feature.

We collected intensity ratings on each of our candidate lexical features   
from ten independent human raters (for a total of 90,000+ ratings). Features were   
rated on a scale from "[–4] Extremely Negative" to "[4] Extremely Positive", with   
allowance for "[0] Neutral (or Neither, N/A)". ~

We kept every lexical feature that had a non-zero mean rating, and whose standard   
deviation was less than 2.5 as determined by the aggregate of ten independent raters.   
This left us with just over 7,500 lexical features with validated valence scores that   
indicated both the sentiment polarity (positive/negative), and the sentiment intensity   
on a scale from –4 to +4. For example, the word "okay" has a positive valence of 0.9,   
"good" is 1.9, and "great" is 3.1, whereas "horrible" is –2.5, the frowning emoticon :(   
is –2.2, and "sucks" and it's slang derivative "sux" are both –1.5.

1. degree modifiers (also called   
   intensifiers, booster words, or degree adverbs) impact sentiment intensity by either   
   increasing or decreasing the intensity

contextual sparseness resulting from shortness of the text and a tendency to use abbreviated language conventions to express sentiments.

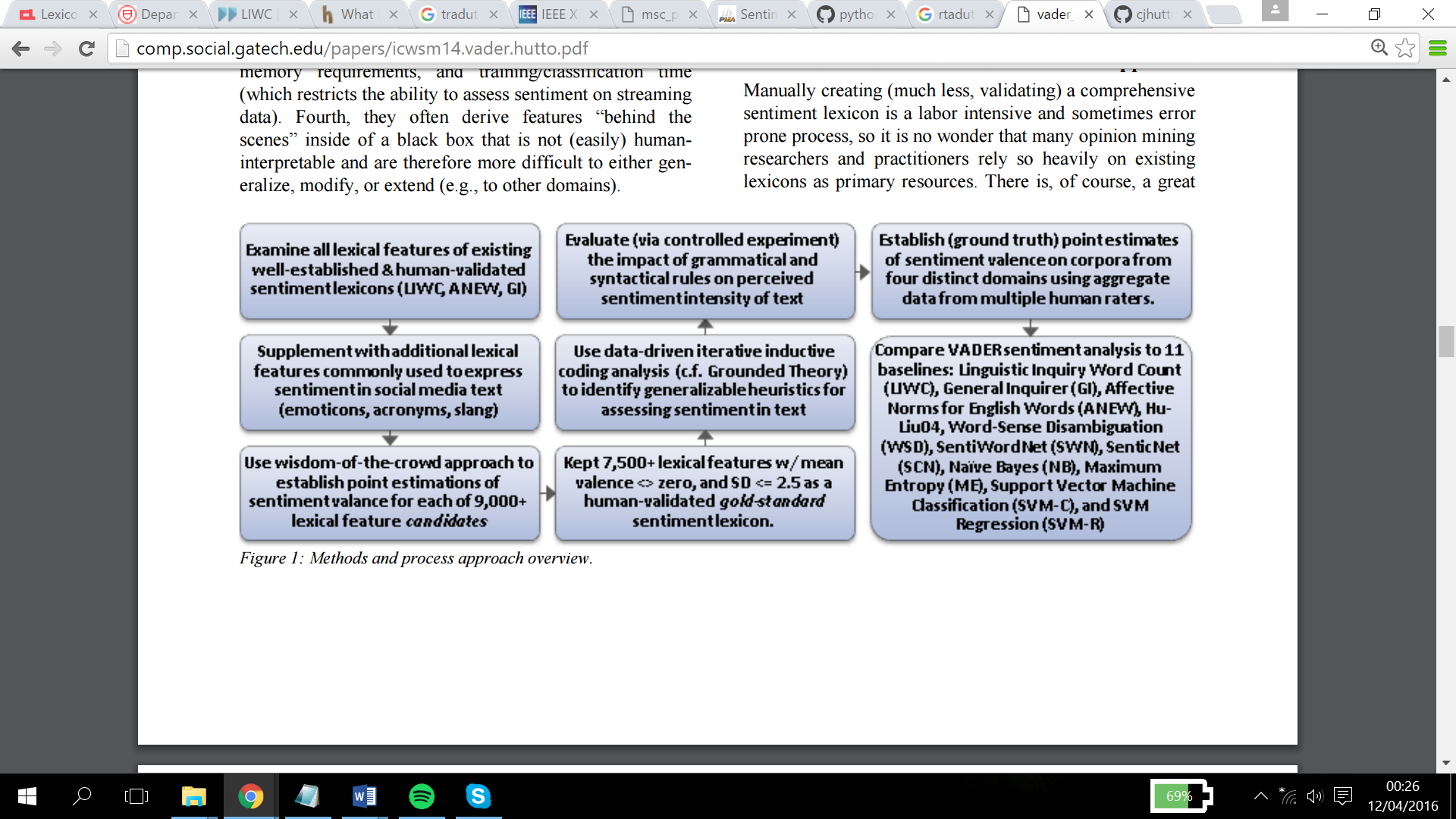
\_\_\_PAPER

Lexicons and Context-Awareness

Whether one is using binary polarity-based lexicons or more nuanced valence-based lexicons, it is possible to improve sentiment analysis performance by understanding deeper lexical properties (e.g., parts-of-speech) for more context awareness. For example, a lexicon may be further tuned according to a process of word-sense disambiguation (WSD)

Word-sense disambiguation refers to the process of identifying which sense of a word is used in a sentence when the word has multiple meanings (i.e. its contextual meaning). For example, using WSD, we can distinguish that the word catch has negative sentiment in “At first glance the contract looks good, but there’s a catch”, but is neutral in “The fisherman plans to sell his catch at the market”. We use a publicly available Python package9 t

Our approach: requires no training data, but is constructed from a generalizable, valence-based, human-curated gold standard sentiment léxicon that is sensitive both the polarity and the intensity of sentiments



1. Create a Valence-Aware Sentiment Lexicon

From existing well-established sentiment word-banks + smileys, initialitsms, slang

= 9,000 lexical feature candidates

2. We used a wisdom-of-the-crowd13 (WotC) approach to acquire a valid point estimate for the sentiment valence (intensity) of each context-free candidate feature

Features were rated on a scale from “[–4] Extremely Negative” to “[4] Extremely Positive”, with allowance for “[0] Neutral.

We kept every lexical feature that had a non-zero mean rating, and whose standard deviation was less than 2.5 as determined by the aggregate of ten independent raters. This left us with just over 7,500 lexical features with validated valence scores that indicated both the sentiment polarity (positive/negative), and the sentiment intensity on a scale from –4 to +4.

For example, the word “okay” has a positive valence of 0.9, “good” is 1.9, and “great” is 3.1, whereas “horrible” is –2.5, the frowning emoticon “:(” is –2.2, and “sucks” and “sux” are both –1.5

* 1. punctuation
  2. 2. Caps
  3. Degree modifiers (intensifiers)
  4. The constrative conjunction but

# 6 Results and discussion

5 stars: “The characters are so real and handled so carefully, that being trapped inside the Overlook is no longer just a freaky experience. You run along with them, filled with dread, from all the horrible personifications of evil inside the hotel's awful walls. There were several times where I actually dropped the boo

was too scared to pick it back up. Intellectually, you know it's not real. It's just a bunch of letters and words grouped together on pages. Still, whenever I go into the bathroom late at night, I have to pull back the shower curtain just to make sure.”

1 star: “The original Star Wars trilogy was a defining part of my childhood. Born as I was in 1971, I was just the right age to fall headlong into this amazing new world Lucas created. I was one of those kids that showed up early at toy stores [...] anxiously awaiting each subsequent installment of the series. I'm so glad that by my late 20s, the old thrill had faded, or else I would have been EXTREMELY upset over Episode I: The Phantom Menace... perhaps the biggest let-down in film history.”

This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can’t hold up” or “I hate the Spice Girls. ...[3 things the author hates about them]... Why I saw this movie is a really, really, really long story, but I did, and one would think I’d despise every minute of it. But... Okay, I’m really ashamed of it, but I enjoyed it. I mean, I admit it’s a really awful movie ...the ninth floor of hell...The plot is such a mess that it’s terrible. But I loved it.” 15 In these examples, a human would easily detect the true sentiment of the review, but bag-of-features classifiers would presumably find these instances dif- ficult, since there are many words indicative of the opposite sentiment to that of the entire review



* Only term- counting precision=X%

Good=2. Better =?

* With lemma form precision = X%
* With sentiment shifters precision = X%

Good=2. Not good = -2

overstatements and understatements

not very good doesnt work that well

Rhetorical devices/modes such as sarcasm, irony, implication, etc

REF

**Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.**