**SENTIMENT VALUE AND EXPRESSIONS**

**Paper 1:**
Exploring fine-grained sentiment values in online product reviews

Reference:

Description of this paper:
We hypothesise that it is possible to determine a fine-grained set of sentiment values over and above the simple three-way positive/neutral/negative or binary Like/Dislike distinctions by examining textual formatting features. We show that this is possible for online comments about ten different categories of products. In the context of online shopping and reviews, one of the ways to analyse consumers’ feedback is by analysing comments. The rating of the “like” button on a product or a comment is not sufficient to understand the level of expression. The expression of opinion is not only identified by the meaning of the words used in the comments, nor by simply counting the number of “thumbs up”, but it also includes the usage of capital letters, the use of repeated words, and the usage of emoticons. In this paper, we investigate whether it is possible to expand up to seven levels of sentiment by extracting such features. Five hundred questionnaires were collected and analysed to verify the level of “like” and “dislike” value. Our results show significant values on each of the hypotheses. For consumers, reading reviews helps them make better purchase decisions but we show there is also value to be gained in a finer-grained sentiment analysis for future commercial website platforms.

**Paper 2:**
Sentiment Analysis Tools Should Take Account of the Number of Exclamation Marks !!!

Reference:

Description of this paper:
There are various factors that affect the sentiment level expressed in textual comments. Capitalization of letters tends to mark something for attention and repeating of letters tends to strengthen the emotion. Emoticons are used to help visualize facial expressions which can affect understanding of text. In this paper, we show the effect of the number of exclamation marks used, via testing with twelve online sentiment tools. We present opinions gathered from 500 respondents towards “like” and “dislike” values, with a varying number of exclamation marks. Results show that only 20% of the online sentiment tools tested considered the number of exclamation marks in their returned scores. However, results from our human raters show that the more exclamation marks used for positive comments, the more they have higher “like” values than the same comments with fewer exclamations marks. Similarly, adding more exclamation marks for negative comments, results in a higher “dislike”.

**Paper 3:**
Reversing the Polarity with Emoticons

Reference:
Technology advancement in social media software allows users to include elements of visual communication in textual settings. Emoticons are widely used as visual representations of emotion and body expressions. However, the assignment of values to the “emoticons” in current sentiment analysis tools is still at a very early stage. This paper presents our experiments in which we study the impact of positive and negative emoticons on the classifications by fifteen different sentiment tools. The “smiley” :) and the “sad” emoticon :( and raw-text are compared to verify the degrees of sentiment polarity levels. Questionnaires were used to collect human ratings of the positive and negative values of a set of sample comments that end with these emoticons. Our results show that emoticons used in sentences are able to reverse the polarity of their true sentiment values.

**Paper 4:**

*Machine Learning Classifiers: Evaluation of the Performance in Online Reviews*

Reference:


**Description of this paper:**

Objectives: This paper aims to evaluate the performance of the machine learning classifiers and identify the most suitable classifier for classifying sentiment value. The term “sentiment value” in this study is referring to the polarity (positive, negative or neutral) of the text. Methods/Analysis: This work applies machine learning classifiers from WEKA (Waikato Environment for Knowledge Analysis) toolkit in order to perform their evaluation. WEKA toolkit is a great set of tools for data mining and classification. The performance of the machine learning classifiers was evaluated by looking at overall accuracy, recall, precision, kappa statistic and few visualization techniques. Finally, the analysis is applied to find the most suitable classifier for classifying sentiment value. Findings: Results show that two classifiers from Rules and Trees categories of classifiers perform equally best comparing to the other classifiers from categories, such as Bayes, Functions, Lazy and Meta. Novelty/Improvement: This paper explores the performance of machine learning classifiers in sentiment value classification of the online reviews. Data used is never been used before to explore the performance of machine learning classifiers.

**Paper 5:**

*Text Segmentation for Analysing Different Languages*

Reference:


**Description of this paper:**

Over the past several years, researchers have applied different methods of text segmentation. Text segmentation is defined as a method of splitting a document into smaller segments, assuming with its own relevant meaning. Those segments can be classified into the tag, word, sentence, topic, phrase and any information unit. Firstly, this study reviews the different types of text segmentation methods used in different types of documentation, and later discusses the various reasons for utilising it in opinion mining. The main contribution of this study includes a summarisation of research papers from the past 10 years that applied text segmentation as their main approach in text analysing. Results show that word segmentation was successfully and widely used for processing different languages.

**Paper 6:**

*Text Segmentation Techniques: A Critical Review*

Reference:
Description of this paper:

Text segmentation is a method of splitting a document into smaller parts, which is usually called segments. It is widely used in text processing. Each segment has its relevant meaning. Those segments categorized as word, sentence, topic, phrase or any information unit depending on the task of the text analysis. This study presents various reasons of usage of text segmentation for different analyzing approaches. We categorized the types of documents and languages used. The main contribution of this study includes a summarization of 50 research papers and an illustration of past decade (January 2007−January 2017)'s of research that applied text segmentation as their main approach for analysing text. Results revealed the popularity of using text segmentation in analysing different languages. Besides that, the word segment seems to be the most practical and usable segment, as it is the smaller unit than the phrase, sentence or line.

**Paper 7:**

Value of Expression behind the Letter Capitalization in Product Review

Reference:


Description of this paper:

Product reviews from consumers are the source of opinions and expressions about purchased items or services. Thus, it is essential to understand the true meaning behind text reviews. One of the ways is to analyzes sentiments, expressions and emotions behind the text. However, there are different styles of writing used in the text. One of widely used in the text is letter capitalization. It is commonly used to strengthen an expression or louder tone within the text. This paper explores the value of expression behind letter capitalization in product reviews. We compared fully capitalized text, text with one capitalized word and text without capitalization through the readers' perspective by asking them to rate the text based on Likert scale. Furthermore, we tested two samples of text with and without capitalization on 27 available online sentiment tools. Testing was done in order to check how current sentiment tools treat letter capitalization in their sentiment score. Results show that of letter capitalization is able to enforce the different level of expression. If the nature of the review is positive, the capitalization makes it more positive. Similar for the negative reviews, the capitalization tends to increase negativity.

**Paper 8:**

Hidden Sentiment Behind Letter Repetition in Online Reviews

Reference:


Description of this paper:

Minimal research has been done on how letter repetition affects readers' perception of expressed sentiment within a text. To the best of the researchers' knowledge, no studies have tested samples of text with letter repetition using sentiment tools. The main aim of this paper is to investigate whether letter repetition in product reviews are perceived to have any sentiment value, based on ratings by individual participants and analyses using sentiment tools. This study collected and analysed 1,041 consumer reviews in the form of online comments using the UCREL Wmatrix system, and simulated emotional words within the comments to contain repeated letters. A group of 500 participants rated 15 positive comments and 15 negative comments and their respective simulated counterparts, while 32 sentiment tools are used to analyse a pair of positive comment and its...
simulated counterpart and a pair of negative comment and its simulated counterpart. Results indicate that readers perceive letter repetition to amplify a comment’s sentiment value, in which the effect was found more strongly in negative comments than positive comments. On the other hand, analyses using sentiment tools show that a majority of these tools are unable to detect letter repetition within a word and instead, treats the word as a spelling mistake. As consumers or online users, in general, have been found to use letter repetition to intensify and express their sentiments in their comments, this study’s findings suggest that letter repetition processing in any text-based mechanism needs to be enhanced. The outcome of this paper is useful for improving the measurement of sentiment analysis for the use of marketing applications.