ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

TRABAJO DE FIN DE GRADO

OPINION MINING OF ONLINE PRODUCT REVIEWS USING A LEXICON-BASED ALGORITHM

Autor: Ignacio Martín-Borregón Musso

Director: Marina Bagic Babac

MADRID

Julio 2018

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. IGNACIO MARTÍN-BORREGÓN MUSSO

DECLARA ser el titular de los derechos de propiedad intelectual de la obra: "OPINION MINING OF ONLINE PRODUCT REVIEWS USING A LEXICON-BASED ALGORITHM",

que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar "marcas de agua" o cualquier otro sistema de seguridad o de protección.
b) Reproducirla en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
e) Asignar por defecto a estos trabajos una licencia Creative Commons.
f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente)*.

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

a) Que la Universidad identifique claramente su nombre como autor de la misma
b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
c) Solicitar la retirada de la obra del repositorio por causa justificada.
d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e

intereses a causa de la cesión.

d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusive del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- La Universidad adoptará las medidas necesarias para la preservación de la obra en un futuro.
- La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a 17 de JULIO de 2018

ACEPTA IGNACIO MARTÍN-BORREGÓN MUSSO

Fdo. Nacho M-B

Motivos para solicitar el acceso restringido, cerrado o embargado del trabajo en el Repositorio Institucional:

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título OPINION MINING OF ONLINE PRODUCT REVIEWS USING A LEXICON BASED ALGORITHM en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el curso académico 2017 - 2018 es de mi autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Fdo.: Ignacio Martín-Borregón Fecha: 11 / 07 / 2018

Nacho M-B

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Marina Bagić Babac Fecha: 12 / 07 / 2018

SUMMARY

Opinion mining of online product reviews using a lexicon-based algorithm

The worldwide social media is a rich resource of user-generated data, which can help organizations to formulate their business strategies, or affect the process of decision making in product or service design and implementation. This data is characterized by its massive size, its complexity and variability, and its growth speed. This makes it very complicated to manage, process and analyze the data with conventional tools. Therefore, new techniques are being developed to collect and use this data, in order to help companies improving their business.

The focus of this thesis is on extraction and analysis of unstructured product reviews for training predictive models, which recognize a specific range of human affective states. These affective states include emotions, moods, opinions, attitudes, as well as continuous dimensions for sentiment characterization, such as valence or intensity.

In this bachelor thesis, a methodological approach is used: first, a dataset with more than 250,000 customer comments and thousands of reactions is collected. Then, a domain-specific sentiment dictionary is built from the product posts, comments and reactions, in order to code and test a simple lexicon-based algorithm to predict the user opinions. Finally, the results will be analyzed and a new algorithm will be proposed, in order to improve even more the results obtained.

Keywords: social networks, operations management, sentiment analysis, lexicon dictionary.

RESUMEN

Análisis de opinión en críticas online de productos usando un diccionario basado en el léxico

Las redes sociales son un recurso muy rico en datos generados por los usuarios de las mismas. Esto puede ayudar a las empresas a formular sus estrategias comerciales o afectar al proceso de toma de decisiones en el diseño e implementación de productos o servicios. Los datos de los que hablamos se caracterizan por su tamaño masivo, su complejidad y variabilidad, y su velocidad de crecimiento. Esto hace que sea muy complicado administrar, procesar y analizar los datos con herramientas convencionales. Por lo tanto, se están desarrollando nuevas técnicas para recopilar y utilizar estos datos, a fin de ayudar a las empresas a mejorar sus negocios.

El objetivo de este proyecto es la extracción y el análisis de críticas no estructuradas de productos para su posterior uso en modelos predictivos, que reconozcan un rango específico de estados afectivos. Estos estados afectivos incluyen emociones, estados de ánimo, opiniones, actitudes, así como dimensiones continuas para la caracterización del sentimiento, como su valor o intensidad.

En este proyecto de fin de grado, se utiliza un enfoque metodológico: primero, se recopila un conjunto de datos con más de 250,000 comentarios de clientes y miles de reacciones. Luego, se construye un diccionario de sentimientos, específico para el mercado bajo estudio, a partir de las publicaciones, los comentarios y las reacciones de Facebook, con el fin de codificar y probar un algoritmo simple basado en léxico para predecir las opiniones de los usuarios. Finalmente, se analizarán los resultados y se propondrá un nuevo algoritmo para mejorar aún más los resultados obtenidos.

Palabras clave: redes sociales, gestión de operaciones, análisis de sentimientos, diccionario de sentimientos.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

The worldwide social media is a rich resource of user-generated data, which can help organizations to formulate their business strategies, or affect the process of decision making in product or service design and implementation. This data is characterized by its massive size, its complexity and variability, and its growth speed. This makes it very complicated to manage, process and analyze the data with conventional tools. Therefore, new techniques are being developed to collect and use this data, in order to help companies improving their business.

The focus of this thesis is on extraction and analysis of unstructured product reviews for training predictive models, which recognize a specific range of human affective states. These affective states include emotions, moods, opinions, attitudes, as well as continuous dimensions for sentiment characterization, such as valence or intensity.

In this thesis, a methodological approach is used: first, a dataset with more than 250,000 customer comments and thousands of reactions is collected. Then, a domain-specific sentiment dictionary is built from the product posts, comments and reactions, in order to code and test a simple lexicon-based algorithm to predict the user opinions. Finally, the results will be analyzed and a new algorithm will be proposed, in order to improve even more the results obtained.

Keywords: social networks, operations management, sentiment analysis, lexicon dictionary.

1.1 Problem Statement

People nowadays spend a lot of time on the internet, and the amount of time is constantly increasing. For example, recent studies show that teens spend up to nine hours a day on social platforms (more than one-third of the day). We tend to make more purchases online, and most of the companies have profiles on this social networks to share new information such as updates or product releases and marketing, as well as providing some customer service. Also, users of this social networks, who are millions of people from all around the world, now have the possibility of giving their opinion to every person willing to read it with just a few clicks. This has been a revolution for customer-to-customer communication, and this information is stored on the internet, as everything else.

At first, this was a huge amount of information that couldn't be managed, but new ways of collecting and analyzing this data have been developed in order to take advantage of it, as it is very useful for Operations Management.

The main purpose of this bachelor thesis is to explore a new way of analyzing data from Facebook and relate it to certain operations management applications. It is structured as follows: first, there is an overview of the studies related to social media data mining and its applications for operations management; then, data collection and some basic sentiment analyses are explained; finally, the development of the sentiment dictionary proposed in this thesis, its results and the conclusions will be provided.

1.2 Related Work

It is not surprising that we can find plenty of studies on social media data mining, as it is one of the revolutionary techniques for marketing purposes. However, companies still don't get the most out of the potential benefits of using social media data (Culnan, McHugh, & Zubillaga, 2010). (Pang & Lee, 2008) presented a general view of the existing work related opinion mining and sentiment analysis for blogs and social media. Most of the work on short text sentiment classification concentrates around Twitter and different machine learning techniques (Wang et al., 2011), (Kouloumpis et al., 2011), (Saif et al., 2012), (Sarlan et al., 2014). Not many sentiment analyses have been made using Facebook posts, because obtaining a labelled dataset for this purpose is much more complicated.

Social media platforms are rich sources for sentiment analysis, because there are millions of users expressing their opinions on different topics, using syntactic structures to describe emotions or state facts (Pak & Paroubek, 2010). Although web-blogs are slightly different than social media platforms -social media platforms such as Facebook or Twitter may be referred as microblog (Kwak et al., 2010)- some interesting research has been made in this field too. In (Yang et al., 2007), the authors based their sentiment analysis on data collected from web-blogs to conclude that a good strategy to determine the overall sentiment of the document is to consider the sentiment of the last sentence of the document as the sentiment of the whole document. A similar study was performed by (Wen & Wan, 2014). Also, (Wilson, Wiebe & Hoffman, 2005) presented “a new approach to phrase-level sentiment analysis that first determines whether an expression is neutral or polar and then disambiguates the polarity of the polar expressions”.

The way users express themselves in social media platforms is sometimes defined by emoticons too, not only words. Some sentiment analysis using emoticons from Twitter was performed in (Go et al., 2009), obtaining up to 81% of accuracy. Also, with the release of ‘Facebook reactions’, some studies included them on their tests. For example, (Tian et al., 2017) concluded that “there is a reliable correlation between Facebook reactions and emoji usages” and also demonstrated that “Facebook reactions and comments are a good data source for investigating indicators of user emotional attitudes”. Social media platforms are also full of product reviews and therefore, analysis of customer feedback is an area which gains interest for many companies over the years. Although there are some studies related to this topic, almost none of them use data from Facebook. (Yang and Fang, 2004), (Hu and Liu, 2004), (Cambria et al., 2013) analyzed customer reviews, but none of them are dealt with the specific nature of Facebook (or social media in general). (Krebs et al., 2017) combined sentiment analysis and reactions from Facebook posts using a customer feedback dataset from various supermarket's Facebook posts, although they focused on how to predict the Facebook reactions to some posts using neural network architectures. As (Yi, Nasukawa, Bunescu & Niblack, 2003) state on their work, there are two challenging aspects of sentiment analysis: the overall sentiment is useful but is only a part of the information of interest ("I am generally satisfied with the phone, although the battery life is short"), and it is difficult to associate the overall sentiment to a specific topic.

Nowadays, not only all this information is used by companies, but also by customers who base their purchases on the satisfaction expressed by other users in their reviews. People seem to like/dislike a specific product because of some feature associated with the product (Eirinaki, Pisal & Singh, 2012). The authors proposed a framework which not only classified a review as positive or negative, but also extracted the most representative features of each reviewed item. As it is explained in their work, (Dave, Lawrence & Pennock, 2003) claim that there exist some issues performing opinion mining about product reviews because of some reasons, such as ambivalence and comparison, because "Mixed reviews introduce significant noise to the problem of scoring words". Similar to the problem that (Farooq et al., 2016) try to handle in their study: "the inability to accurately determine the effect of negation on other words".

It is important to keep in mind that data from social networks has its advantages and disadvantages. For example, as social networks are based on the transmission of word-of- mouth information, the data being used for this thesis is completely subjective (Litvin, Goldsmith, & Pan, 2008; Jansen et al., 2009; Shih, Lai, & Cheng, 2013). This subjectivity may affect in a good way the decision of online consumption (Cheung, Lee, & Rabjohn 2008), but it can also be given less credibility or persuasiveness (Cheung et al. 2009; Zhang, Craciun, & Shin, 2010).

In addition, (Chan, Lacka, Yee, & Lim, 2017) focused on a very similar objective as this thesis, although the analysis of the data is performed in much different way.

The idea of the sentiment analysis is the following: whenever we read some text, we use our language capacities to understand the emotions that are being transmitted by the author. Now we are also able to understand the emotions expressed on a text programmatically, using some tools developed for text mining. One way of doing so is by calculating the sentiment content of each word of the text, and then considering that the sentiment content of the text is the sum of these calculated sentiments. In this thesis, a brief sentiment analysis using the tidytext R package (De Queiroz, Keyes, Robinson, & Silge, 2018) will be shown. This package contains a sentiment dataset based on different sentiment lexicons. Three of them were used: NRC (Mohammad, & Turney, 2010), Opinion Lexicon (Liu, 2004) and AFINN (Nielsen, 2011). They work with many English words as follows: NRC classifies words into one or more categories such as anger, fear, sadness, disgust, surprise or joy, categories which are also classified as positive or negative; Liu (2004) classifies them into positive or negative, and AFINN gives each word a value between -5 and 5, the positive values indicating a positive sentiment while the negative values indicate a negative sentiment.

Due to the fact that the analysis is based on calculating the sentiment content of each word instead of the whole text, qualifiers are not taken into account (e.g., “this phone is not good”). Also, some people use sarcasm to express their opinion, which confuses the algorithm (e.g., “I hate you :)”).

However, these sentiment lexicons don’t make a great job at tackling the harder task of emotion analysis, which is a natural evolution of sentiment analysis (Staiano & Guerini, 2014). For a better “buzz monitoring” model, classifying comments into positive or negative is not enough, and the lexicons available for classifying words into more emotions were built in a generic way, being poorly accurate.

Every type of text has its own unique characteristics. There is not a unique writing style in a scientific text, in the newspaper, in poetry or in social media. In the latter, for example, we can also find different writing styles, due to different users or different purpose of the social media. This is the main reason why “prepared” lexicons may not be efficient enough, and the reason for building our own lexicon dictionary (based exclusively in the the information that this thesis aims to analyze). This dictionary should have some benefits for the analysis. For instance, different words order affects the sentiment analysis: “He escaped but then he was caught :)” and “He was caught but then he escaped :(” express different emotions using the same words. (Wang & Manning, 2012) propose the use of word bigrams instead of isolated words, which helps tackling this problem. Using a specific domain dictionary doesn’t solve the problem directly but should help, considering that users from the same social media tend to express themselves in similar ways and, specially, because its based on Facebook reactions.

1.3 Objectives

This bachelor thesis will consist on a program based on R language that aims to reach this main objectives:

1) Collect a huge, representative and tidy dataset of consumer’s reviews to smartphone products.

2) Build a domain-specific sentiment dictionary using Facebook’s reactions and make some optimizations.

3) Build a lexicon-based algorithm to predict user’s opinions and test it.

In order to perform a sentiment analysis, a dataset is obviously needed. Collecting some specific representative data from Facebook is not an easy task, but it is possible thanks to some new tools and packages. This thesis uses Facebook data because it is full of unstructured product reviews, posted by users all around the world. The dataset is collected from some famous smartphone companies, due to the quantity and quality of reviews that they provide (it’s a very demanded market nowadays, and the product itself can have many different characteristics, like the quality of the camera, the screen, the design, the reliability, etc.). Based on this dataset, a domain-specific sentiment dictionary is built in order to improve the sentiment analyses performed using generic sentiment lexicons. Some adjustments are then made in order to improve even more the results. The program then calculates the emotions expressed by the user's reviews using a lexicon-based algorithm and, furthermore, it predicts user's opinions to new product releases posted on the smartphone company Facebook's posts.

1.4 Work Methodology

The program is completely based on R language. This language provides some tools and packages like Rfacebook (Barbera, Piccirilli, Geisler, A., & van Atteveldt, 2017), which is basically the cornerstone of this thesis. It allows downloading the information needed from Facebook to collect the dataset for this thesis. In order to use the package functions, creating an app in "Facebook for developers" is needed, as the functions require a special API key which is given by Facebook to its developers. Currently, Facebook is involved in some privacy issues which made them stop giving this keys (at least to non-regular developers), but the dataset of this thesis was collected before this issue. This problem will only result in the incapacity to improve the dataset, if needed.

Data from four smartphone companies was collected: Samsung Mobile, Huawei Mobile, Sony Mobile and LG Mobile. The dataset was collected in January 2018. All consumer's comments published during the 2017, and not any from other years, were downloaded for the analysis being 260.210 comments in total. It is important to remark that each company writes different number of posts or receives different amount of comments to one post. Therefore, when the analysis compares one company with another, it is needed to keep in mind that this analysis is based on different amount of data. Samsung retrieved 167.940 comments (more than half of the dataset); Sony did 37.556; LG retrieved 53.363 and Huawei only 1.351. Moreover, Samsung only posted 188 times, while Sony did it 500 times. LG and Huawei posted 366 and 301 times respectively.

Once the data is collected and cleaned, a brief sentiment analysis based on generic sentiment dictionaries is shown. These sentiment analysis is based on the previously mentioned tidytext R package.

After that, the focus is on the main task of this thesis: building the domain specific sentiment dictionary and lexicon-based algorithm. In order to build the dictionary, two matrices are created and multiplied, obtaining the dictionary. Before creating the matrices, it is needed to classify all the comments to the posts as documents. For the purpose of this thesis, each document represents each post (all the comments from users to a post from the company).

The first matrix is a word-by-document matrix. Basically, with the help of Rstudio and its packages it is obtained, for each word appearing in all documents, the normalized frequencies of that word appearing in each document. The second matrix is a document- by-emotion matrix (how much of each emotion is shown in each document, with normalized frequencies). The emotions for this matrix are obtained using the reactions to Facebook posts. Therefore, the sentiment dictionary for this thesis has the following emotions: *anger*, *fun*, *love*, *sadness* and *surprise*. However, it should be taken into account that Facebook reactions are expressed as emojis, and each emoji can express more than one emotion. For example, the “wow” emoji can express surprise or disbelief, which are actually very different emotions. Furthermore, the surprise could be for something positive or negative, etc.

Once both matrices are created, it is needed to multiply them (and normalize them again) to obtain the word-by-emotion matrix, the dictionary. This matrix has the “score” of expressed sentiment for each word.

Finally, the focus of this thesis is on developing an algorithm to predict the emoji distribution to a post. The basic method would be to sum the emotion weights of all the words in the post and normalize the output. However, the idea is to improve this algorithm to obtain more accurate results. The results of the predictions will be analyzed using a statistical approach.

CHAPTER 2: DATASET COLLECTION

The first step towards performing a good sentiment analysis and creating a reliable sentiment dictionary is collecting a huge, tidy and representative dataset. Due to the purpose of this thesis, the election was to collect posts from Facebook with some of their respective information: the user comments to the post and the number of each of the reactions to the post. Furthermore, the decision was to collect them from the Facebook page of four of the most well known smartphone companies: Samsung, Huawei, LG and Sony. Other companies such as Apple were not considered for the dataset because their Facebook page is not focused only on smartphone devices, but also on other products of different nature (laptops, music devices, etc.). The decision was to collect them from smartphone companies because of various reasons.

To begin with, smartphones are a world scale product. It is calculated by the GSMA that nowadays there's more than five thousand million smartphone users around the world, which represents almost three fourths of the world population. Moreover, it's a good thing for the thesis that the smartphone business, in spite of being massive and varied, its generally controlled by a few big companies. This, indirectly, entails that more data will be found for the same companies. Specifically, each of these Facebook pages count with millions of likes: 45M, 51M, 4'3M and 23M respectively. This helps collecting a huge dataset.

Lastly, smartphones have many different attributes. This is useful for the thesis, because it helps correlating different attributes of the same product to different sentiments expressed by the users. For example, if Huawei is well known for offering a good quality-price rate, words related to price will acquire a higher positive sentiment. The same will happen with LG's screen, etc. In general, a more accurate sentiment analysis can be performed to see the pros and cons of each company.

The dataset was collected in January 2018. It was decided that the dataset would consist only in the posts from 2017, in order to have a clear time frame and be able to use it for other applications too, such as the impact on the economic growth of these companies in the stock market.

The collection of the dataset was possible through an R package called Rfacebook (Barbera, Piccirilli, Geisler, A., & van Atteveldt, 2017), which has some functions that retrieve all the information regarding the posts from the public Facebook page that you give as an input. It cannot return information about private pages or public user profiles. This last part is remarked now but will be commented later. The output information is basically a data frame with the message of the post, the ID code of the post, the number of comments, likes and each of the reactions to the post, and the time when it was published. The ID code of the post will be used as an input for another function of this package, which retrieves all the information about a public Facebook post, including the list of comments and likes. This thesis is only interested in the post message, the number of its reactions and the messages commented to the post. The idea is to build different data frames with this information, for the different analyses that will be made. Each of the data frame's structure will be explained before each type of experiment. However, one thing in common to all the experiments is the necessary cleaning of the data.

Not all the data retrieved by this functions will be useful. Posts with no reactions (and their respective comments) are deleted from the dataset and *stop words* (in English) are also deleted from the messages (either posts or comments). Only English stop words are removed for two main reasons: non-English messages will not affect the different experiments (sentiment analysis and the construction of the sentiment dictionary) and there's no other stop words dataset built in R for other languages, specially not a dataset of stop words for "every other language". The reason why non-English messages don't affect the experiments will be exposed in their respective chapters. *Stop words*, in case there's an explanation needed, it's a dataset that contains all the words in English that are too common to take into account when processing natural language data. Words such as "the", "a", etc. that are not used to express anything. Adding or removing words to that list is possible if it is considered necessary for the analysis.

Let’s take an overview of the dataset collected:

- **Samsung Mobile:** 188 posts, 167.940 comments.
- **Huawei Mobile:** 301 posts, 1.351 comments.
- **LG Mobile:** 366 posts, 53.363 comments.
- **Sony Mobile:** 500 posts, 37.556 comments.

It is important to remark that each company writes different number of posts or receives different amount of comments to one post. Therefore, when the analyses compare one company with another, it is useful to keep in mind that these analyses are based on different amount of data, which influences the difference in the performance of the analyses for each company. The more information you have, the more accurate the results will be (if they are the same quality).

These numbers add up to a total of 1.355 posts and 260.210 comments, numbers that will be useful later on. Having more than a million for both would honestly be better for the analyses and, for this reason, a better collection of the data was attempted months later on this thesis, finding an unexpected issue that prevented me from achieving it.

The functions mentioned before use an API key as an input to give them access to the data that they retrieve. This key is given by Facebook to its developers, the users of the “Facebook for Developers” platform. In order to get the API key, it is just needed to create a free account in this platform and create an “app”. Nonetheless, Facebook was involved in some legal issues regarding the privacy of their users in April 2018, what caused the platform to close the access to it. Therefore, up to this day this task still cannot be achieved in order to improve the dataset collection.

CHAPTER 3: OPINION MINING OR SENTIMENT ANALYSIS

3.1 Calculating Sentiment from Data

Sentiment analysis, or opinion mining is defined as *"the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral".* The term was added to Oxford dictionaries in August 2014, and this "is evidence that sentiment analysis isn't going anywhere anytime soon, and is only going to become a more widespread and important tool for business and technology." (Mekkin Bjarnadottir, 2014).

This sentiment analysis can be performed in multiple ways. Firstly, it is needed to classify each word appearing in the text that is being analyzed into one or more sentiment categories. These categories can be simple (positive, negative) or more complex (joy, sadness, surprise, etc.). Inside this categories, the words can also have a certain value of belonging to the category, due to the nature of languages (some words can express sentiments in a stronger way than others). Furthermore, the classification of these words into the categories can be done manually (for instance, rating how positive or negative the words are, or classifying them into the different categories), or programmatically (through the use of an algorithm using different types of information as an input). Once the words are classified, various algorithms can be used to try to pull out the best of the analysis, depending on the data that this thesis wants to analyze. One of the simplest methods is to consider that the sentiment content of the text is the sum of the sentiment values of the words. However, many improvements of this algorithm have been developed.

The purpose of performing a sentiment analysis in this thesis is to show the outcome of a simple sentiment analysis performed on its data and to give a general idea of the scope of this technique. As it will be shown, the analysis has its own limitations that could be improved, which gives this thesis its direction.

In this thesis, the tidytext R package is used (De Queiroz, Keyes, Robinson, & Silge, 2018), which contains a sentiment dataset based on different sentiment lexicons. Three of them were used: NRC (Mohammad, & Turney, 2010), Opinion Lexicon (Liu, 2004) and AFINN (Nielsen, 2011). They work with many English words as follows: NRC classifies words into one or more categories such as anger, fear, sadness, disgust, surprise or joy, categories which are also classified as positive or negative; Liu (2004) classifies them into positive or negative, and AFINN gives each word a value between -5 and 5, the positive values indicating a positive sentiment while the negative values indicate a negative sentiment.

The algorithm for the analysis is based on calculating the sentiment content of each word instead of the text as a whole (simple algorithm) and consider the last one as the sum of the sentiment content of all words. Due to this fact, qualifiers are not taken into account (e.g., “this phone is not good”). Besides, some people use sarcasm to express their opinion, which confuses the algorithm (e.g., “I hate you :)”).

Comments published in other languages than English don’t affect the sentiment analysis. The reason for this is that neither of the three sentiment datasets have scores for non-English words, and therefore they don’t affect the sum of the sentiment scores.

The dataset used for this analysis is in the tidytext format. First of all, a data frame with the structure shown in Figure 1 is obtained. This data frame contains three columns: “text”, which contains the text of every comment collected in the dataset (of every company), “PostNumber” which contains the number of the post that the comments belong to (every company has its own numeration), and “Company” which has the name of the company that the comments belong to. The last two are for separation purposes when needed. The data frame is then converted to a tidytext format (only one word per column), obtaining a new data frame with the structure shown in Figure 2.

	text	PostNumber	Company
1	When can a girl get a pen back? My note exploded an...	1	Samsung
2	God... everytime I read these manufacturer posts I se...	1	Samsung
3	Is samsung a once in a lifetime purchase for anyone e...	1	Samsung
4	I'm a new Samsung user and I just tried to block the a...	1	Samsung

Figure 1. Data frame structure for sentiment analysis

	PostNumber	Company	word
1	1	Samsung	girl
2	1	Samsung	pen
3	1	Samsung	note
4	1	Samsung	exploded

Figure 2. Data frame structure in tidytext format

Before starting the analysis, this last data frame is cleaned by deleting the English stop words. Finally, it is first calculated the most common words that appear in the comments of each company. It is very common for text mining to look at word frequencies, and it is necessary for the sentiment analysis. It is also very helpful, as it helps understanding the main topics of discussion.

3.2 Results and Interpretation

The most common words from each company's posts, based on the word frequencies analysis for each company, are shown below:

```
      word     n
     <chr> <int>
 1 samsung 20813
 2   phone 10920
 3      s8  7739
 4    note  7487
 5  galaxy  4653
 6       8  4455
 7      s7  3237
 8  mobile  3007
 9 service  2844
10    edge  2605
```

Figure 3. The most common words for the comments to Samsung's posts

```
     word     n
    <chr> <int>
 1  phone   330
 2 huawei   321
 3   mate   237
 4      9   167
 5   love   144
 6 phones    98
 7  watch    75
 8    win    70
 9     10    66
10 device    65
```

Figure 4. The most common words for the comments to Huawei's posts

```
      word     n
     <chr> <int>
1     sony 10424
2   xperia  4854
3    phone  3595
4       xz  2320
5     love  1641
6  premium  1629
7   mobile  1335
8    price  1108
9       z5  1083
10  camera  1060
```

Figure 5. The most common words for the comments to Sony's posts

```
     word     n
    <chr> <int>
1      lg  9130
2   price  2414
3    nice  2336
4   phone  2244
5      g6  2171
6     v30  1195
7      g4  1022
8    love   904
9      g5   885
10 mobile   803
```

Figure 6. The most common words for the comments to LG's posts

Not plenty of information is given by this common words, because mostly the common words for each company are related to the company's name or main products. However, some basic ideas can still be extracted. For example, a quick look on the Internet can confirm that Samsung's most selling smartphones are the Samsung Galaxy Note8, the Samsung Galaxy S8, and then the old model, the S7. And that seems to be exactly what the people are talking about. Also, the word "service" appears very often on the comments, which may be caused by complains or compliments about the service (this will be checked later on). For Huawei, people are mostly commenting about the Huawei Mate 9 and 10, and people seem happy with their purchase (e.g., the 5^{th} most common word). Regarding Sony, it can be intuited that customers may like their Sony Xperia XZ, maybe because of the good quality of the camera, maybe or the price (or maybe they complain about it). Finally, it can be extracted from the comments to LG that the most selling phones would be the LG g6 (also g5 and g4) and the LG v30, which probably are "nice phones at a good price". Nevertheless, there is a need for further analysis on these comments to get a better understanding of what people think about these products. A sentiment analysis is needed.

As mentioned before, there are three sentiment lexicons for general purposes. To begin with the analysis, it is important to check which one is the most appropriate for the thesis. For this purpose, the three of them are used to analyse how the sentiments change across the different posts published, this time only by Samsung.

From the dashboard in Figure 7, it can be noticed that the three lexicons retrieve results that have similar forms, with the dips and peaks located on the same posts, although the absolute value of the sentiments is much different. The graph may be confusing or inaccurate, because taking a brief look at the posts that generate these peaks, it can be seen that they are the most commented posts, with over 20,000 comments each. One of them was a video introducing the new Samsung Galaxy Note8, and another one was a live video (which usually generates thousands of comments). Therefore, in order to take a better look at the general idea, instead of comparing the lexicons using the posts from one company, the sentiments for the four company's comments are compared in Figure 8.

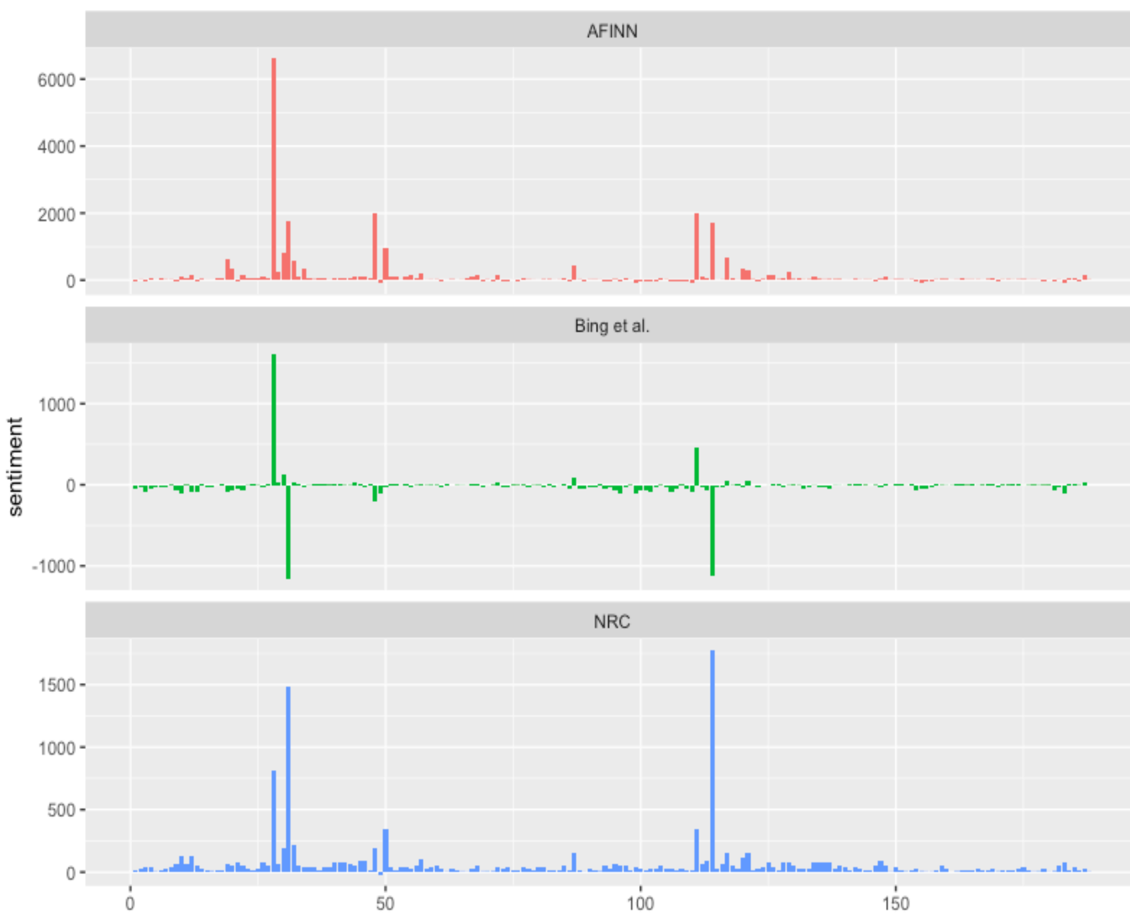

Figure 7. Comparison of three sentiment lexicons using Samsung's comments

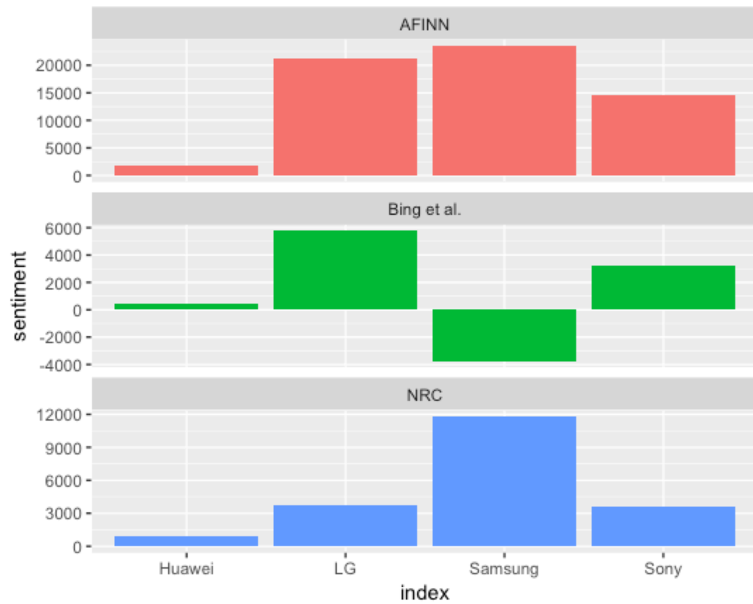

Figure 8. Comparison of three sentiment lexicons using the comments of each company

While AFINN and NRC show high positive sentiments for all the companies, Liu's shows that sentiments for Samsung are negative. This happens because Liu's lexicon has 6,783 words (much more than the other lexicons), where 4,783 are negative and 2,006 are positive (the ratio of negative/positive words is much higher than the other lexicons). This lexicon has some useful properties, as it includes misspellings, slang, and social-media mark-up. Therefore, from now on this thesis will use Liu's lexicon for the sentiment analysis. Before moving on, let's emphasize that Samsung seems to be getting some negative feedback on their posts, and remember that the most common words that were seen previously led us to think that Samsung may had some trouble with the service being provided to the customers.

Figure 9 is shows a closer look at the feedback given to the four companies. It can be noticed that most of the Samsung's posts receive comments with negative sentiment content, while the other companies keep a low positive feedback. The questions that should be answered are, e.g. what are the problems that each company may have, or which are their strengths.

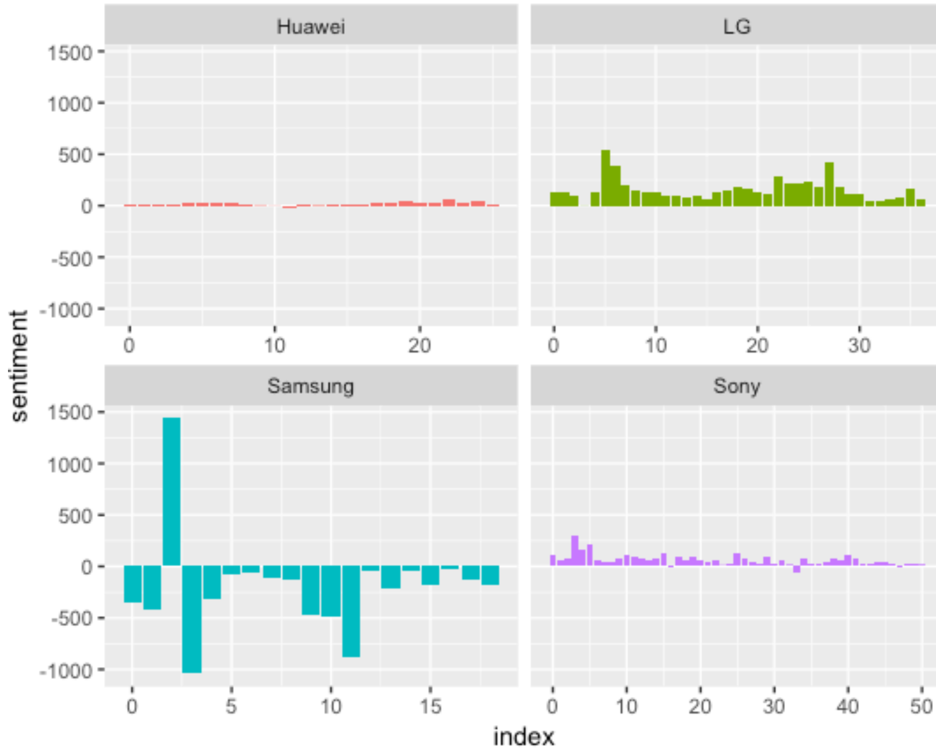

Figure 9. Comparison of the companies using Liu's lexicon

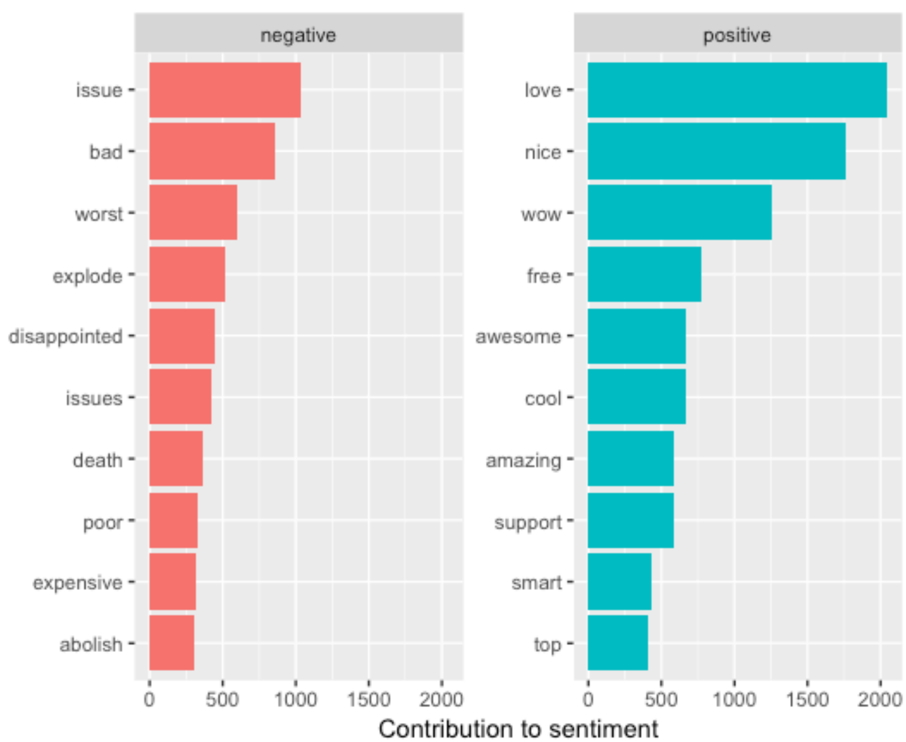

Figure 10. Words that contribute the most to positive and negative sentiment in Samsung

From Figure 10, the problem that Samsung has can be observed: Explosions! In 2016, Samsung introduced to the market the Samsung Galaxy Note 7, which was withdrawn two months later because of some issues that made the battery explode. It can also be seen that some customers find that their smartphones are too expensive. It is interesting to pay attention to the rate of positive to negative words, because it looks like around a 40% of the customers are not happy with Samsung's products. However, Samsung is still the leading company for the smartphone business because of their variety of smartphones and their prestige built over the years.

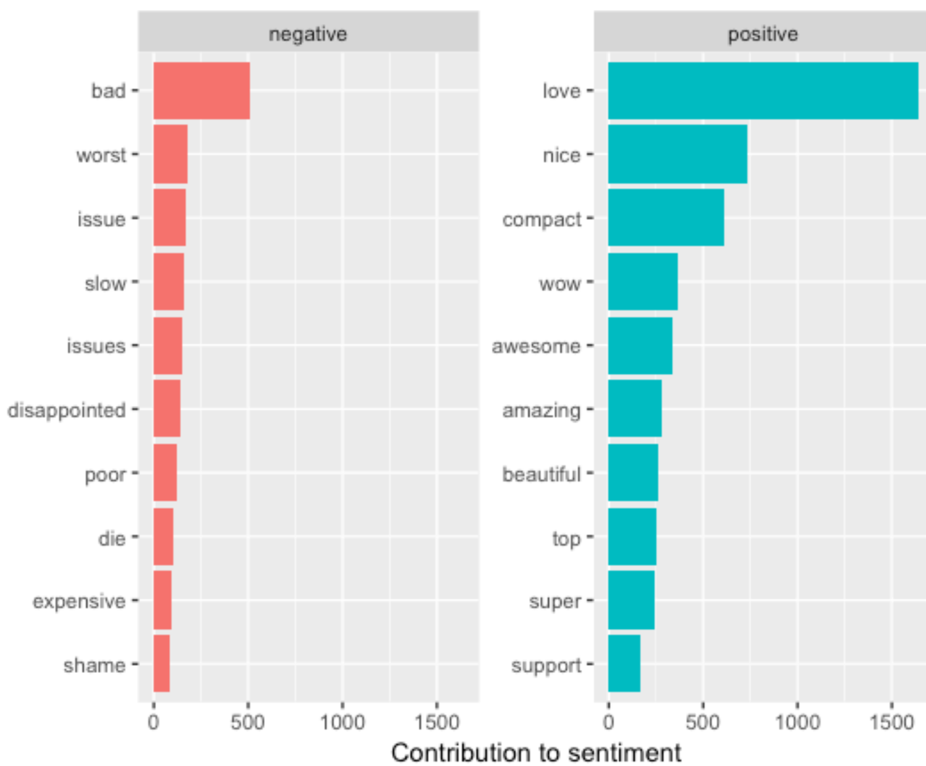

Figure 11. Words that contribute the most to positive and negative sentiment in Sony

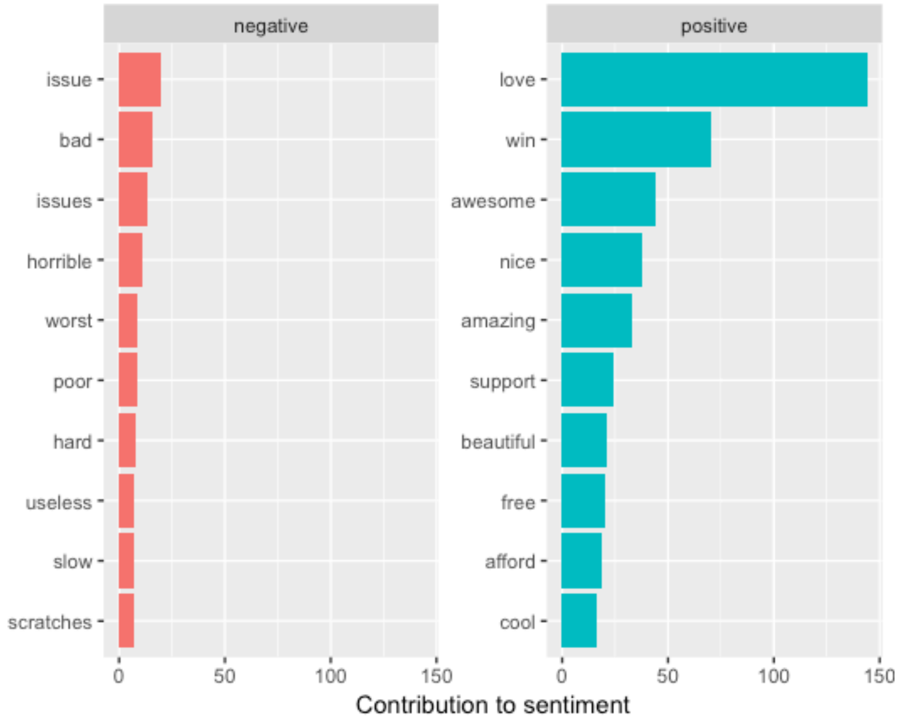

Figure 12. Words that contribute the most to positive and negative sentiment in Huawei

From Figure 11, it can be extracted that Sony's strength is that their smartphones are good-looking, although they may be a little bit expensive and slow. However, words with a positive sentiment content have higher values because of their frequencies in the comments. This shows that a high percentage of the customers are satisfied with their smartphones.

Huawei's customers seem to be satisfied too, as is shown in Figure 12. What really catches the attention is that Huawei is the only company that doesn't have the word "expensive" among the words that contribute the most to the negative sentiment. Instead, there are two words which contribute to the positive sentiment that have the completely opposite meaning: "free" and "afford". Searching for news about Huawei, it will be seen that this Chinese company was the second leading company in the smartphone business, and their business policy is based on high quality at the best price.

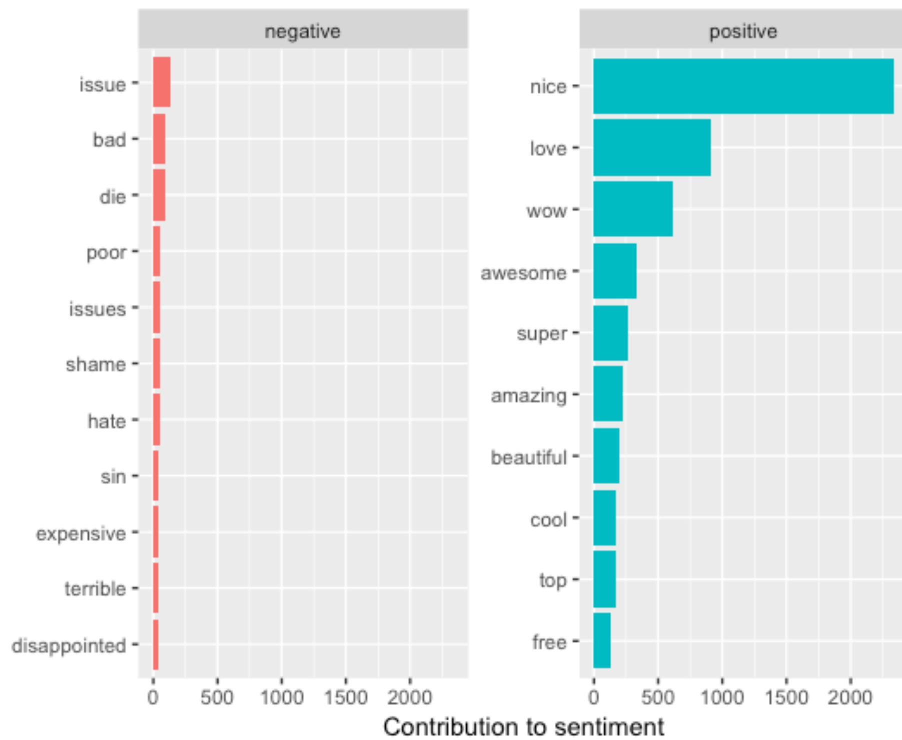

Figure 13. Words that contribute the most to positive and negative sentiment in LG

LG basically doesn't have very negative feedback, which is why from Figure 8, it can be observed that it is the company with the higher positive sentiment. From Figure 13, it can be concluded that customers are generally satisfied with their purchase.

From the outcomes of a simple sentiment analysis, it can be noticed how powerful this technique can be, although the results are still not accurate. In order to use it for operations management purposes, some improvements should be made. Some of the possibilities are: improving the dataset, improving the sentiment dictionary that is used to perform the analysis or improve the method for calculating the sentiment content of the whole text. This thesis considers the second one because the words appearing in the comments have an important relation with the names of the product, and that is not considered in this generic-domain sentiment dictionaries. Moreover, because Facebook reactions could play a very important role in the analysis of the user's reviews. It is also a novel approach that can be performed in a fully automated way. Finally, because it is foreseen that this not only could allow the thesis to analyse reviews but to predict them.

CHAPTER 4: DOMAIN-SPECIFIC SENTIMENT DICTIONARY

The domain-specific sentiment dictionary for the task is built by using Facebook's reactions to the posts related to the product. These reactions determine the sentiment content of the words appearing in the dictionary, and they are chosen by the users, not the managers of the Facebook page that is publishing the posts. Therefore, it could be useful to build **two different dictionaries** and compare the accuracy of the results for both of them. One will determine the sentiment content of the words appearing in the posts, while the other will take into account only the words appearing in the comments published by the users. Therefore, **two different datasets** are prepared for this task.

The final outcome of these dictionaries is going to be a matrix. The rows are the words of the dictionary, while the columns are the normalized score for each sentiment. As Facebook distinguishes five different reactions, those are the sentiments taken into account for the sentiment dictionary.

4.1 Posts Sentiment Dictionary

The first sentiment dictionary that is created is the posts dictionary. In order to do so, the dataset is prepared as follows:

a) Create a Posts-Reactions data frame:
   A data frame containing every post from every company and its reactions:

from_name	message	angry_count	haha_count	love_count	sad_count	wow_count
Samsung Mobile	Experience the essence of the Galaxy ...	8	12	314	3	121
Samsung Mobile	The #GearFit2, #GearS2 and #GearS3 ...	5	5	145	1	45
Samsung Mobile	More music to your ears. Spotify is no...	8	5	148	2	39
Samsung Mobile	Introducing Samsung LEVEL Box Slim,...	11	18	651	4	370
Samsung Mobile	Capture better selfies with #GalaxyA2...	8	10	301	3	196
Samsung Mobile	Play with water. #GalaxyA2017 #IP68	8	31	282	3	109
Samsung Mobile	Take your group selfies to another lev...	4	9	188	1	48
Samsung Mobile	Fitness, motivation & style now come ...	3	6	143	1	36

Figure 18. Posts-Reactions data frame structure

b) Create a Term Document Matrix (TDM) from the message column (character vector) of the Posts-Reactions data frame:

A TDM is "*a mathematical matrix that describes the frequency of terms that occur in a collection of documents.".* In a TDM, columns correspond to documents in the collection and rows correspond to terms. In this case, each document corresponds to each post (and each term is a word). As is shown in Figure 19, it has 3.844 rows (different words appearing in the posts) and 1.355 columns (total number of documents). This matrix has three vectors: "i", "j" and "v". "I" is row index, and its maximum value is the number of words appearing in the document. It assigns a number to every different word, matching same words with the same number. "j" is a column index, and its maximum is the number of documents. It assigns every word to its document. "v" is vector with the values of absolute frequency of the words in its document. Before creating the TDM, converting the non-readable characters (hashtags, emojis, etc.) to a readable format is needed.

Name	Type	Value
dtm_Posts	list [3844 x 1355] (S3: TermD	List of length 11
i	integer [20142]	1 2 3 4 5 6 ...
j	integer [20142]	1 1 1 1 1 1 ...
v	double [20142]	1 1 1 1 1 1 ...
nrow	integer [1]	3844
ncol	integer [1]	1355

Figure 19. Term Document Matrix for Posts

c) Create a matrix only with reaction's values:

The normalized values of the reaction's votes need to be added to the TDM. Therefore, the matrix is created and normalized. In this matrix, the rows are still the posts, not the words. The structure is shown below:

	angry_count	haha_count	love_count	sad_count	wow_count
1:1	0.017467249	0.026200873	0.6855895	0.006550218	0.26419214
1:2	0.024875622	0.024875622	0.7213930	0.004975124	0.22388060
1:3	0.039603960	0.024752475	0.7326733	0.009900990	0.19306931
1:4	0.010436433	0.017077799	0.6176471	0.003795066	0.35104364

Figure 20. Normalized reactions votes matrix (Document-Emotion matrix)

d) Convert TDM to a matrix:

Using the R function data.matrix(), the TDM is converted to a matrix format. This matrix will have for each row (different terms) the absolute frequency of the term appearing in each document (columns). It looks like this:

galaxya2017	1	0	0	0	1	1	1	0
introducing	1	0	0	1	0	0	0	0
totally	1	0	0	0	0	0	0	0
apps	0	1	0	0	0	0	0	1
armour	0	1	0	0	0	0	0	1
connected	0	1	0	0	0	0	0	1
fitness	0	1	0	0	0	0	0	2
gearfit2	0	1	1	0	0	0	0	1

Figure 21. Word-Document matrix of absolute frequencies

e) Multiplication of both matrices:

Multiplying the Word-Document emotion by the Document-Emotion matrix will return the Word-Emotion matrix that is the sentiment dictionary. Every word is given a value for each sentiment. Before using this dictionary, column-wise and row-wise normalization is applied. Column-wise normalization is used for over representation of happiness. This is because of the fact that people tend to express more positive moods on social networks (Quercia et al., 2011; Vittengl and Holt, 1998; De Choudhury et al., 2012). The first specific-domain sentiment dictionary is now ready, and it looks like this:

	angry_count	haha_count	love_count	sad_count	wow_count
attitude	0.20239208	0.07130073	0.19836038	0.30832251	0.21962429
different	0.11155626	0.62597006	0.14371761	0.02333254	0.09542354
dna	0.26596280	0.08278233	0.20054512	0.14929417	0.30141557
essence	0.21205999	0.08913706	0.20183449	0.23940550	0.25756297
experience	0.13304017	0.13108148	0.15434888	0.38470089	0.19682858
galaxy	0.19882707	0.31230322	0.12405114	0.21175877	0.15305980
galaxya2017	0.28164490	0.11398942	0.18736337	0.14399470	0.27300762
introducing	0.25887873	0.05871096	0.13256110	0.36582643	0.18402277
apps	0.25575550	0.30237129	0.16461877	0.08356583	0.19368860
armour	0.37628196	0.09509606	0.22258876	0.07939177	0.22664146
connected	0.21333904	0.36841723	0.19752076	0.04501242	0.17571055
fitness	0.22767497	0.24140970	0.22769649	0.15411695	0.14910188
gearfit2	0.41732997	0.07744318	0.18817034	0.11566491	0.20139159

Figure 22. Posts Sentiment Dictionary structure

This dictionary has a total of 2.029 terms annotated with their sentiment score.

4.2 Comments Sentiment Dictionary

The second sentiment dictionary that this thesis wants to create is the comments dictionary. In order to do so, it is needed to prepare the dataset as it was done for the other dictionary. However, due to organization purposes, the method is a bit different: this time, the documents will be all the comments to one post.

a) Create a large vector of comments:
   Storing every comment to each post in a character vector position.
b) Create a TDM with the vector:
   This TDM has the same structure as the other one, but has 150.165 rows (number of different words appearing in the comments) and 1.355 columns (number of documents).

c) Convert the TDM to a matrix:

This matrix will have for each row (different terms) the absolute frequency of the term appearing in each document (columns).

d) Multiplication of matrices:

The new Word-Document matrix is multiplied by the Document-Emotion matrix (the same that was used for the other dictionary) and the output matrix is normalized, obtaining the new dictionary, which looks like this:

accept	0.2090252	0.10858982	0.23027907	0.22678389	0.22532198
accepting	0.2583688	0.04979184	0.26434023	0.17985198	0.24764715
accessible	0.2295355	0.07191300	0.25915969	0.23290611	0.20648571
achet	0.2483793	0.08810098	0.22485244	0.22857485	0.21009241
acheter	0.2232250	0.21377287	0.19207432	0.20042469	0.17050313
acquisition	0.2064618	0.06778591	0.23395793	0.24161491	0.25017942
acreditam	0.3377137	0.08023435	0.21047092	0.18481970	0.18676132
across	0.3337597	0.07792422	0.19452878	0.20170634	0.19208091
act	0.1672572	0.11270320	0.16150473	0.37548200	0.18305292
activate	0.2197644	0.13106406	0.21827785	0.21426041	0.21663327
active	0.2292606	0.09082357	0.21422823	0.24206240	0.22362519

Figure 23. Comments Sentiment Dictionary structure

This second sentiment dictionary has 35.575 terms with their annotated sentiment score. It is 17 times bigger than the first one. However, some of the words appearing in this dictionary belong to other languages or could even be names of Facebook users. The latter is due to comments in which users tag another user. These words won't affect the outcome of the subsequent prediction because they wont be used. The percentage of non-used words for this dictionary is unknown.

Both dictionaries will be used for the same purpose in the next chapter. Nevertheless, they come from a different data and have different properties. For example, the comments dictionary might have spelling mistakes or slang, while the posts dictionary is probably more formal. Also, the comments dictionary probably contains more words that are commonly associated with negative sentiments than the post dictionary. Finally, as it was mentioned before, the comments sentiment dictionary is multilingual. These properties make each dictionary appropriate for different purposes and the thesis will focus on one. The rest are out of the scope of this thesis.

CHAPTER 5: PREDICTION OF REACTIONS

The thesis has the objective of coding an algorithm that, when given a new post as an input, calculates the normalized distribution of user's reactions to the post. This could allow having an idea of the impact of the new products before even releasing them (in this case, before even announcing their release). This has many useful applications in the field of operations management, and could improve significantly the process of decision making. The steps to build this algorithm are the following:

a) Extract the new post and convert it to a tidy format (one word per row).
b) For each word, find the exact same word in the dictionaries and copy the columns related to its sentiment score.
c) Sum the total score of each sentiment and normalize the output
d) The result will be a vector containing the normalize scores for each sentiment.

The coding for this algorithm will be shown in the cd attached at the end of this thesis.

5.1 10-fold Cross Validation: Explanation and Results

In order to study the results obtained with this algorithm, a technique known as 10-fold Cross Validation is performed. This technique evaluates predictive models by dividing the original dataset into two different sets: the train set, which is used to train the model; and the test set, to evaluate it. Specifically, 10-fold Cross Validation will randomly partition the dataset into ten equal size folds. While one of them will be used as the test set, the other nine will be used as the train set.

Additionally, it is needed to build the dictionaries again using only the train set. This way, the test set is used just for testing and won't take part in the dictionaries.

At this point, the thesis has the train set (1/10 of the posts extracted) and two different dictionaries (one built from 9/10 of the posts and the other one built from the comments corresponding to those 9/10 of the posts). The purpose now is to calculate the normalized distribution of reactions predicted for each of the posts in the train set and compare them to the normalized distribution of the real reactions of those posts. In order to do so, the function cor() is called, which calculates the Pearson correlation coefficient between both matrices. Due to the fact that the data sample is randomly mixed to make the folds, every time the algorithm is run different results are found. It will be shown the first four Pearson Coefficient matrices for both cases, to get a better idea of the results.

Testing the accuracy of the predictions is usually calculated using the MAPE (Mean Absolute Percentage Error). The MAPE is calculated using this formula:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Real - Prediction}{Real}\right| \cdot 100$$

If the MAPE is less than 10% it means that the predictive model is good enough. However, it has a drawback that can be observed from the formula: It cannot be calculated when the real value is zero. This scenario is possible in our dataset (posts with zero reactions in one or more sentiments) and would give infinite values. Only valid results will be shown.

The Pearson correlation coefficient results obtained using the posts dictionary are shown in Figure 24, while the results obtained from the comments dictionary are shown in Figure 25.

The Mean Average Percentage Errors will be shown only for the improved algorithm, which will be explained further in this chapter.

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.07104048	0.06931641	-0.007043227	0.05997303	-0.108381065
haha_count	0.33551953	0.45999838	-0.460423044	-0.09473409	-0.044267374
love_count	-0.23487933	-0.36541404	0.526362392	-0.11651214	-0.197357044
sad_count	-0.14455697	-0.08691709	0.095850264	0.12756695	0.003449232
wow_count	-0.23102062	-0.34719990	0.117276049	0.07976661	0.371519067

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.05203364	-0.03277477	0.28690683	-0.10875990	-0.35729251
haha_count	0.09310912	0.40870384	-0.31249876	-0.00941762	-0.03987510
love_count	-0.26112680	-0.30207797	0.30332483	-0.12731274	-0.04522269
sad_count	0.18731946	-0.16069494	0.09123417	0.13755422	0.03130415
wow_count	-0.10722851	-0.15258602	-0.20486702	0.13298332	0.45781238

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.0960999811	-0.07547202	0.07076597	0.112793783	-0.03252610
haha_count	-0.0758171970	0.39207206	-0.30131487	-0.047033844	-0.03053077
love_count	-0.1100938828	-0.32589755	0.40758447	-0.076610867	-0.15312485
sad_count	0.1217959593	0.03583908	-0.13837123	0.001476254	0.12753894
wow_count	0.0002251086	-0.17299226	0.06338990	0.018099095	0.11443416

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.27665170	-0.01283143	0.02509230	0.11156295	-0.1770128
haha_count	-0.04338782	0.47053932	-0.23830107	0.10450534	-0.1837482
love_count	-0.12443932	-0.27955309	0.43210086	-0.17815938	-0.1891275
sad_count	0.01532888	-0.04998490	-0.17275810	0.05103047	0.2721932
wow_count	-0.14017098	-0.24098298	-0.02334844	-0.12119737	0.3825516

Figure 24. Pearson Correlation matrices using Posts dictionary

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.18714179	-0.03378558	-0.05651772	0.18922503	0.07256986
haha_count	-0.12655896	0.39454500	-0.26539753	-0.16081493	-0.06605348
love_count	-0.04264925	-0.42513863	0.41053698	0.05977775	-0.06189135
sad_count	0.09541214	-0.18466173	0.13013408	0.06963290	0.01727846
wow_count	0.07625261	-0.33524380	0.14524740	0.06623744	0.17635811

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.17992607	-0.06808064	-0.01784701	0.16481169	0.06271060
haha_count	-0.11955299	0.41309959	-0.27131692	-0.15009442	-0.08160622
love_count	-0.05147964	-0.43912220	0.42582945	0.05649917	-0.06477896
sad_count	0.12514237	-0.22896471	0.09546016	0.10084012	0.10954639
wow_count	0.07687556	-0.36159150	0.18153131	0.06387031	0.15757158

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.18455693	-0.01244858	-0.08632873	0.22833240	0.08600863
haha_count	-0.12545318	0.38207607	-0.25625978	-0.15890806	-0.06428857
love_count	-0.03170152	-0.41261015	0.39780889	0.06955611	-0.06163292
sad_count	0.08079516	-0.19050680	0.14099756	0.03374501	0.01424500
wow_count	0.06551959	-0.30401450	0.14493005	0.02497254	0.14496874

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.17700354	-0.04666834	-0.04310222	0.18095423	0.07154209
haha_count	-0.12635737	0.37406926	-0.25769272	-0.14940284	-0.05348762
love_count	-0.03259632	-0.41087608	0.42061651	0.03349447	-0.09215157
sad_count	0.06771477	-0.18271554	0.11315517	0.07344953	0.04253309
wow_count	0.09524137	-0.27171013	0.11788414	0.06060440	0.13672192

Figure 25. Pearson Correlation matrices using Comments dictionary

From the matrices shown in figures 24 and 25, this thesis is interested only in the diagonal, which represents the Pearson correlation coefficient between the sentiments predicted and the real ones. It can be observed from the results obtained that both dictionaries have a similar accuracy in predictions. However, the Posts Sentiment Dictionary gives significantly more varied results. This is probably due to the amount of data that makes up the dictionary. The more information the dataset has, the more accurate the results are. This is reflected by the difference between these results and the ones from the Comments Sentiment Dictionary, which has 17 times more information and, therefore, the results vary very little. Before analyzing the quality of the results obtained, the average Pearson Correlation Coefficient is calculated for every emotion from the four cases (Figures 26 and 27).

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,0710405	0,4599984	0,5263624	0,1275670	0,3715191
0,0520336	0,4087038	0,3033248	0,1375542	0,4578124
0,0961000	0,3920721	0,4075845	0,0014763	0,1144342
0,2766517	0,4705393	0,4321009	0,0510305	0,3825516
Average:				
0,1239565	**0,4328284**	**0,4173431**	0,0794070	**0,3315793**

Figure 26. Average Pearson Correlations for Posts Sentiment Dictionary

It can be noticed that for the Posts Dictionary, the results are much better for the positive emotions (“Haha”, “Love” and “Wow”). The results for this emotions show an average of about 0.4 in Pearson correlation coefficient., which conform a moderate relationship between the prediction and the real distribution of reactions.

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,18714179	0,394545	0,41053698	0,0696329	0,17635811
0,17992607	0,41309959	0,42582945	0,10084012	0,15757158
0,18455693	0,38207607	0,39780889	0,03374501	0,14496874
0,17700354	0,37406926	0,42061651	0,07344953	0,13672192
Average:				
0,182157083	**0,39094748**	**0,413697958**	0,06941689	0,153905088

Figure 27. Average Pearson Correlations for Comments Sentiment Dictionary

For the Comments Dictionary, the results are also better for positive emotions (“Haha”, “Love”) but this time, “Wow” gets worse results. In general, it could be affirmed that the Posts Dictionary is better for prediction of reactions as it gives better results for “Haha” and “Love” and “Wow” emotions, at least using this simple algorithm.

5.2 Other algorithms for the reactions prediction

In order to improve the results obtained, some new algorithms for the prediction of the reactions are considered.

The first consideration is to take into account possible spelling mistakes. This is done by calculating the sentiment score of each word as the sum of the sentiment scores of all the words similar to the word from the post, and normalizing the result. However, this consideration doesn’t improve the results and is rejected.

The first try was to make the calculations considering only the first 50% of the words of the post, because people might be reacting without reading all the post. However, this didn’t improve the results even when considering higher percentages like 80%.

The next try was to consider only the last 80% of the words of the post, in order to give more importance to the end of the post, which may include keywords such as the name of the smartphone. But this didn’t manage to improve the results neither.

The first consideration that achieved an improvement of the results using both matrices was to remove the stop words from the post that was being analyzed. This way words that are to common to actually express a sentiment would not be taken into account.

Also, for some reason, when avoiding the column-wise normalization that was done to the sentiment dictionary for over representation of happiness (see Chapter 4), results are also a bit better, but only using the posts dictionary. For the comments dictionary, the results get worse instead of better.

Last but not least, it is also taken into account possible spelling mistakes. This is done by calculating the sentiment score of each word as the sum of the sentiment scores of all the words similar to the word from the post, and normalizing the result. This improves the results, but only when using the comments dictionary and only for the “haha” reaction. It also lowers the time for computing the algorithm.

The results are shown in the next figures. First of all, it is shown the results using the Posts Dictionary, when the stop words are removed and the column-wise normalization is avoided (Figure 28). Figure 29 and 30 show the results when using the Comments dictionary, first deleting the stop words and then considering also spelling mistakes.

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.40441350	0.1986435	-0.2805603	-0.03206179	0.003574582
haha_count	0.01015852	0.6951782	-0.4532970	0.02612757	-0.303662118
love_count	-0.03319241	-0.5208992	0.6403536	0.01478720	-0.174897944
sad_count	0.05924272	0.1097500	-0.2334232	-0.05744683	0.174931545
wow_count	-0.06996737	-0.1951845	-0.2010187	-0.03709989	0.550884062

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.25894843	0.2459830	-0.1496126	0.18092132	-0.25394133
haha_count	0.04896172	0.5792694	-0.3155416	0.02955737	-0.34463570
love_count	-0.07094726	-0.3368759	0.4374670	-0.10241693	-0.13591414
sad_count	0.05953240	-0.1105958	0.0413574	0.12221471	0.04508529
wow_count	-0.02250731	-0.2598032	-0.1886264	0.04742393	0.62367485

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.25470061	0.16435024	-0.1631257	0.16021360	-0.09621865
haha_count	0.05772481	0.56864473	-0.2536408	0.02497853	-0.22293351
love_count	-0.02610352	-0.45731692	0.4870144	-0.15618990	-0.16747724
sad_count	0.01411857	-0.01419164	0.0811557	-0.05816315	-0.09002864
wow_count	-0.07345862	-0.03453890	-0.3352834	0.15244188	0.46876307

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.51720341	0.1667675	-0.09847582	-0.04196871	-0.23793249
haha_count	0.02258426	0.5860372	-0.52386876	0.13561086	-0.05946524
love_count	-0.04581570	-0.3813405	0.54979392	-0.14044181	-0.22392546
sad_count	0.03356781	-0.0374674	0.05384410	-0.04456251	-0.03068190
wow_count	-0.08017963	-0.1031745	-0.19241431	0.06538170	0.40405022

Figure 28. Improved Pearson Correlation Matrices using Posts Dictionary

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.25312567	0.00396662	-0.0743865	-0.01781289	0.08635232
haha_count	-0.13246201	0.41655773	-0.3984313	-0.22372678	-0.03033714
love_count	0.01536517	-0.36144104	0.4551311	0.01116601	-0.11943743
sad_count	-0.09847732	-0.19202213	0.1685742	0.23649193	0.04172313
wow_count	0.10920964	-0.30392451	0.2295832	0.30819293	0.10441286

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.2311005856	0.0154876	-0.06785468	-0.010779136	0.05956804
haha_count	-0.1190733802	0.3994862	-0.37864157	-0.219461513	-0.03470182
love_count	-0.0002870577	-0.3526394	0.44001601	-0.008803522	-0.10783664
sad_count	-0.0593093798	-0.1726209	0.12815891	0.282575836	0.06597893
wow_count	0.0778465158	-0.2911187	0.22087105	0.269093224	0.10251568

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.22636066	-0.09732376	-0.02198438	0.13336426	0.09337615
haha_count	-0.09064775	0.51926277	-0.32087063	-0.15365845	-0.09809732
love_count	-0.06650197	-0.47681876	0.46106649	-0.06688738	-0.09642452
sad_count	0.03599821	-0.23782774	0.07076232	0.21989131	0.13625025
wow_count	-0.01613570	-0.37762342	0.20017254	0.10276485	0.12558747

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.265945952	0.0268913	-0.09245858	-0.0005350996	0.07520723
haha_count	-0.124618303	0.3951475	-0.36244064	-0.2368153399	-0.05081189
love_count	-0.009808315	-0.3555153	0.44971884	0.0018025614	-0.11788483
sad_count	-0.062210910	-0.1837646	0.12655428	0.2965825463	0.08542788
wow_count	0.065599935	-0.2929467	0.21842347	0.2632190889	0.11054829

Figure 29. Improved Pearson Correlation Matrices using Comments Dictionary (1)

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.21994730	0.0008426037	-0.05874870	0.09971471	0.04247752
haha_count	-0.06425080	0.5071044625	-0.35111176	-0.13397630	-0.05037365
love_count	-0.06812827	-0.4578990823	0.45125354	-0.08509273	-0.10189523
sad_count	-0.05099108	0.0429487728	-0.08004059	0.11776087	0.05939050
wow_count	-0.00678219	-0.4902329709	0.27534813	0.15822876	0.13947017

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.22592674	-0.08131754	-0.01611322	0.16629377	0.06710591
haha_count	-0.09398944	0.51321168	-0.33492376	-0.15478965	-0.07292380
love_count	-0.04411526	-0.45259875	0.45145085	-0.12548709	-0.10817996
sad_count	-0.06047799	-0.11834144	0.01173885	0.19522512	0.10307056
wow_count	0.02401903	-0.33225140	0.17935759	0.08125452	0.10204738

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.247019997	-0.11055651	0.02072115	0.11561095	0.04969138
haha_count	-0.102641839	0.56619226	-0.32941281	-0.16425028	-0.13348211
love_count	-0.034027576	-0.49837659	0.42131068	-0.05795817	-0.02658189
sad_count	-0.020159572	-0.07332663	-0.04240802	0.13927811	0.12564512
wow_count	0.001695284	-0.51445641	0.25933031	0.20288660	0.18179918

	angry_count	haha_count	love_count	sad_count	wow_count
angry_count	0.22636066	-0.09732376	-0.02198438	0.13336426	0.09337615
haha_count	-0.09064775	0.51926277	-0.32087063	-0.15365845	-0.09809732
love_count	-0.06650197	-0.47681876	0.46106649	-0.06688738	-0.09642452
sad_count	0.03599821	-0.23782774	0.07076232	0.21989131	0.13625025
wow_count	-0.01613570	-0.37762342	0.20017254	0.10276485	0.12558747

Figure 30. Improved Pearson Correlation Matrices using Comments Dictionary (2)

The average Pearson correlation coefficients are calculated and shown below:

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,4044135	0,6951782	0,6403536	-0,0574468	0,5508841
0,2589484	0,5792694	0,4374670	0,1222147	0,6236749
0,2547006	0,5686447	0,4870144	-0,0581632	0,4687631
0,5172034	0,5860372	0,5497939	-0,0445625	0,4040502
Average:				
0,3588165	**0,6072824**	**0,5286572**	-0,0094894	**0,5118431**
Previous Average:				
0,1239565	0,4328284	0,4173431	0,0794070	0,3315793

Figure 31. New Average Pearson Correlation Coefficients for Posts Dictionary

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,2531257	0,4165577	0,4551311	0,2365919	0,1044129
0,2311006	0,3994862	0,4400160	0,2825758	0,1025157
0,2263607	0,5192628	0,4610665	0,2198913	0,1255875
0,2659460	0,3951475	0,4497188	0,2965825	0,1105483
Average:				
0,2441332	**0,4326136**	**0,4514831**	**0,2589104**	0,1107661

Figure 32. New Average Pearson Correlation Coefficients for Comments Dictionary (1)

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,2199473	0,5071045	0,4512535	0,1177609	0,1394702
0,2259267	0,5132117	0,4514509	0,1952251	0,1020474
0,2470200	0,5661923	0,4213107	0,1392781	0,1881799
0,2263607	0,5192628	0,4610665	0,2198913	0,1255875
Average:				
0,2298137	**0,5264428**	**0,4462704**	**0,1680389**	0,1388212
Previous Average:				
0,182157083	0,39094748	0,413697958	0,06941689	0,153905088

Figure 33. New Average Pearson Correlation Coefficients for Comments Dictionary (2)

From Figure 31, it can be concluded that the new method for calculating the reactions has improved significantly the results of four reactions out of five. The results show a weak-moderate correlation between the prediction and the real distribution for the “angry” reaction. For “haha” they indicate a moderate-strong relation. For “love” and “wow” a moderate relation is found, and for “sad” there is a very weak relation.

In Figures 32 and 33, it can be seen that both considerations for the algorithm have a positive impact on the results obtained. When stop words are removed, the results already indicate a slightly better performance for four of the five reactions. The same happens when considering spelling mistakes, too. The effects of the latter are a higher performance for the “haha” reaction, which indicates a moderate relation between the prediction and the real distribution, at the cost of a lower performance for the “sad” reaction. The rest of the reactions can be considered to be equal.

In general, these results indicate a better performance of the Posts Dictionary. This means that the words from the posts have more impact on the distribution of the reactions than the comments expressed by the users.

Furthermore, the Mean Average Percentage Errors, which are calculated for the predictions using the Posts Dictionary, are shown in the Figure 34 below:

Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
0,863707	0,737561	0,114600	0,628038	0,378825
0,928744	0,835266	0,148423	1,477296	0,438151
1,668469	0,767449	0,112345	0,767701	0,368845
0,804394	0,905522	0,129270	0,657623	0,367590
Average:				
1,066329	0,811450	0,126159	0,882665	0,388353

Figure 34. Mean Average Percentage Errors using Posts Dictionary

These results show a good performance of the predictive model proposed in this thesis. The predictions vary from the real values between a 0,13% ("love" reaction) and a 1,07% ("angry" reaction). These results go according to the ones observed before.

Finally, the prediction of reactions is performed using the Posts Dictionary with the improved algorithm, although this time the dataset is divided for each company, to remark the differences of the results when focusing on only one company and watch how different amount of dataset can affect the performance of the predictive model. Figures 35-38 show the results of the Pearson Correlation Coefficients for eight different tests using only each of the company's posts:

	Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
1	0.30975279	0.69555656	0.04899546	0.52899514	0.4199946
2	0.14282657	0.68602183	0.29894576	0.51145337	0.6520689
3	0.34949504	0.85720930	0.55939807	0.32453591	0.6235327
4	0.45494122	0.91048951	0.74741745	0.02455504	0.5589405
5	0.40939149	0.51973375	0.39891247	0.59518100	0.1748563
6	0.07073285	0.35008826	0.22106820	0.01444831	-0.2095180
7	0.08382367	0.70309361	0.40537400	0.36644390	0.5264322
8	0.70581959	0.06355535	0.42941399	0.03585450	0.3265683

Figure 35. Pearson Correlation Coefficients for Samsung Company

	Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
1	-0.19823897	0.7285741	-0.075941858	0.31905894	0.08331622
2	0.35658401	0.3688692	0.245406188	0.22877791	0.24925078
3	-0.11965787	0.4914031	0.388965050	-0.09547843	0.32310516
4	0.29779033	0.2700268	0.516626816	0.12610096	0.01175472
5	-0.09025264	0.7780540	-0.006702772	-0.18866066	0.07495420
6	-0.04401358	0.3520343	-0.050919675	-0.05367217	-0.03190997
7	0.06390253	0.1992409	0.342191052	-0.09842267	-0.17051084
8	-0.14609136	0.4199364	0.423802178	-0.16299033	0.26878051

Figure 36. Pearson Correlation Coefficients for Huawei Company

	Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
1	0.40199450	0.220878283	0.3416250	0.270825352	0.31315479
2	0.09918976	0.133471972	0.5479636	-0.009816812	-0.05293389
3	0.29782935	0.063414155	0.4624052	0.212655687	0.31773119
4	0.47368752	0.173274031	0.4350896	0.118042545	0.23506911
5	0.31955932	-0.005506247	0.4816716	0.558312102	0.52581172
6	0.14360723	0.310733168	0.5390787	0.232606718	0.08763580
7	-0.12259683	0.465158611	0.3126706	0.301597264	0.57384961
8	-0.03874627	0.446305313	0.3574611	0.256360114	0.17607654

Figure 37. Pearson Correlation Coefficients for LG Company

	Angry_Count	Haha_Count	Love_Count	Sad_Count	Wow_Count
1	0.631311716	0.30965895	0.3667590	0.179178828	0.3433706
2	-0.004785958	0.72498786	0.7314798	-0.009635273	0.4476049
3	0.217873969	0.03489153	0.3047733	0.378106071	0.3340617
4	0.064317250	-0.13533440	0.3848649	0.192815364	0.3086639
5	0.451957530	-0.07631393	0.2958279	0.325433876	0.3101305
6	0.616692337	0.62799605	0.2252274	-0.112459065	0.1928594
7	0.672927129	-0.02929801	0.2360523	0.303866268	0.3366752
8	0.205729097	0.04800833	0.7029144	0.153604615	0.6476167

Figure 38. Pearson Correlation Coefficients for Sony Company

When analyzing the results, it is important to take a look at the amount of data that makes up the dataset. After the cleaning, Samsung presents 130 available posts (117 used as the train set and 13 used as the test set), Huawei has 234 posts (210 for train and 24 for test), LG has 290 posts (261 as train set and 29 as test set) and Sony 325 posts (295 for train set and 33 for test). In general, these numbers are very low compared to the 979 posts that would conform the dataset for all companies together.

As it can be seen from the figures, this low amount of data produces variations in the results that make the task of calculating the average results to be worthless, because it wouldn't be representative. This makes clear that for further improvements in the work presented in this thesis, the volume of the dataset is one of the keys to achieving better and more accurate results. This should be an easy task because the code is simple, although it is not possible at the moment due to the Facebook issues that were presented in Chapter 2.

It can be concluded from the experiments performed in this chapter that this model has a high potential for the prediction of Facebook's reactions. The collection of a bigger dataset is considered as a key point to obtain better and more accurate results. Nonetheless, the use of this naïve algorithms already show a high coverage and the improvements shown in the last part of the chapter show the huge potential of this predictive model. When studying the results, we see that negative reactions ("Angry" and "Sad") present worse performance results than the positive reactions. The reason for this could be the over representation of happiness (commented in Multiplication of both matrices:), as people tend to express more positive emotions, even in reactions. One way to demonstrate this is to sum all the number of reactions to every post, which are 7.365, 60.920, 466.538, 5.999 and 168.426. It is noticed that "love" is an over used reaction, and also that a higher number of reactions implies a better prediction. Additionally, it could be related to the "positive language" expressed by the posts from the Facebook companies.

In order to improve even more the results obtained in the predictions, the use of normalized frequencies should be considered as they provide better results (Staiano & Guerini, 2014). Also, new improvements in the algorithm and a better collection of the dataset are the new goals to expand the work presented in this thesis.

CONCLUSION

In this thesis, an extraction of posts, comments and reactions from Facebook is made in order to perform different analyses on it for operations management purposes. The data extracted is related to the smartphone business, collecting every post with its reactions and comments published in 2017 by four different smartphone companies: Samsung, Huawei, LG and Sony.

The first step after the collection of the dataset, is to perform a basic sentiment analysis of the data. It indicates how powerful this technique can be if developed correctly. As it was carried out using generic dictionaries, the thesis then proposes the creation of some specific-domain dictionaries, using Facebook's reactions to measure the sentiment value of each word. Two dictionaries are built, one from the words of the posts and another one from the words appearing in the user's comments.

Finally, the thesis focuses on the prediction of the reactions to the posts being analyzed. An algorithm is developed based on the two sentiment dictionaries built in the previous task. The results of the prediction are analyzed using the 10-fold Cross Validation technique and calculating the Pearson Correlation Coefficients between the predictions and the real values. Although the results are not very statistically significant, they show a high-coverage of the Sentiment Dictionaries. After implementing some improvements on the naïve algorithm, new results are presented, showing the high potential of this model. It is concluded that the Posts Dictionary is better for the prediction and the Mean Average Percentage Errors of a few predictions are calculated.

Nevertheless, there is even much more future work that could be done to expand this thesis, which could be a key for business management and decision making.

LITERATURE

Barbera, P., Piccirilli, M. Geisler, A., & van Atteveldt, W. (2017). RFacebook, R package, Available at: https://cran.r-project.org/web/packages/Rfacebook/index.html

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. IEEE Intelligent Systems, 28(2):15–21.

Chan, H. K., Lacka, E., Yee, R. W.Y., & and Lim, M. K. (2017). The role of social media data in operations and production management. *International Journal of Production Research*, 55 (17). 5027-5036.

Cheung, C. M., M. K. Lee, and N. Rabjohn. (2008). The Impact of Electronic Word-of-Mouth. *Internet Research*, 18 (3): 229–247.

Cheung, M. Y., Luo, C., Sia, C. L., & Chen. H. (2009). Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of on-Line Consumer Recommendations. *International Journal of Electronic Commerce*, 13 (4): 9–38.

Culnan, M. J., McHugh, P. J., & Zubillaga, J. I. (2010). How Large US Companies Can Use Twitter and Other Social Media to Gain Business Value. *MIS Quarterly Executive*, 9 (4): 243–259.

Dave, K., Lawrence, S., & M. Pennock, D. (2003). Mining the peanut gallery: opinion extraction and semantic classification of product reviews.

De Choudhury, M., Counts, S. & Gamon, M. (2012). Not all moods are created equal! exploring human emotional states in social media. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

De Queiroz, G., Keyes, O., Robinson, D., & Silge, J. (2018). tidytext R package, Available at: https://cran.r-project.org/web/packages/tidytext/index.html

Eirinaki, M., Pisal, S., & Singh, J. (2012). Feature-based opinion mining and ranking. *Journal of Computer and System Sciences*, 78(4), 1175–1184. doi:10.1016/j.jcss.2011.10.007

Farooq, U., Nongaillard, A., Ouzrout, Y., and Qadir, M. A. (2016). Negation Handling in Sentiment Analysis at Sentence Level. In Internation Conference on Information Management, Londres, United Kingdom.

Go, A., Huang, L., & Bhayani, R. (2009). Twitter sentiment analysis. Final Projects from CS224N for Spring 2008/2009 at The Stanford Natural Language Processing Group.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 168–177. ACM.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter Power: Tweets as Electronic Word of Mouth. *Journal of the American Society for Information Science and Technology*, 60 (11): 2169–2188.

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The Good the Bad and the OMG!. In Proceedings of the 5th International AAAI Conference on Weblogs and Social Media

Krebs, F., Lubascher, B., Moers, T., Schaap, P., & Spanakis, G. (2017). Social Emotion Mining Techniques for Facebook Posts Reaction Prediction.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). “What is Twitter, a Social Network or a News Media?” In Proceedings of the 19th International Conference on World Wide Web, 591–600. New York: ACM.

Litvin, S. W., Goldsmith, R. E. & Pan, B. (2008). Electronic Word-of-Mouth in Hospitality and Tourism Management. *Tourism Management* 29 (3): 458–468.

Liu, B. (2004). Opinion Lexicon, Available at: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html

Mekkin Bjarnadottir: *‘Sentiment Analysis’ Added to Oxford Dictionaries* (2014). https://www.lexalytics.com/lexablog/sentiment-analysis-added-to-oxford-dictionaries

Mohammad, S., & Turney, P. (2010). NRC Emotion Lexicon, Available at: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

Nielsen, F. A. (2011). AFINN, Available at: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010

Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In LREc, volume 10.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval.

Quercia, D., Ellis, J., Capra, L., & J. Crowcroft. (2011). In the mood for being influential on twitter. *Pro- ceedings of IEEE SocialCom'11*.

Saif, H., He, Y., Alani, H. (2012) Semantic sentiment analysis of twitter.

Sarlan, Aliza & Nadam, Chayanit & Basri, Shuib. (2014). Twitter sentiment analysis. 212-216. 10.1109/ICIMU.2014.7066632.

Shih, H.-P., Lai, K.-H., & Cheng. T. C. E. (2013). Informational and Relational Influences on Electronic Word of Mouth: An Empirical Study of an Online Consumer Discussion Forum. *International Journal of Electronic Commerce*, 17 (4): 137–166.

Staiano, J., & Guerini, M. (2014). DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News.

Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., & Sun, C. (2017). Facebook Sentiments: Reactions and Emojis.

Vittengl, J.R. & Holt, C.S. (1998). A time-series diary study of mood and social interaction. *Motivation and Emotion*, 22(3):255–275.

Wang, S. & Manning, C. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification.

Wang, X., Wei, F., Liu, X., Zhou, M., & Zhang, M. (2011). Topic sentiment analysis in Twitter: A graph-based hashtag sentiment classification approach. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management (CIKM'11), pages 1031–1040.

Wen, S. and Wan, X. (2014). Emotion classification in microblog texts using class sequential rules. In AAAI, pages 187–193.

Wilson, T., Wiebe, J., & Hoffman, P. (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis.

Yang, C., Hsin-Yih Lin, K., & Chen, H. (2007). Emotion classification using web blog corpora. In WI ’07: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, pages 275–278, Washington, DC, USA. IEEE Computer Society.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. International Journal of Service Industry Management, 15(3):302–326.

Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques.

Zhang, J. Q., G. Craciun, & D. Shin. (2010). “When Does Electronic Word-of-Mouth Matter? A Study of Consumer Product Reviews.” Journal of Business Research 63 (12): 1336–1341.