Alma Mater Studiorum
University of Bologna

SCHOOL OF SCIENCE
Computer science applied to Management

Measuring corporate reputation through online social media

Author:
Massimo-Maria Barbato

Supervisor:
Danilo Montesi

Academic year 2015-2016
December 2016

"*Instead of wanting to be like this or that, make yourself into a silent, immovable giant. That's what the mountain is. Don't waste your time trying to impress people. If you become the sort of man people can respect, they'll respect you, without your doing anything.*"

Eiji Yoshikawa, Musashi

Abstract

What is corporate reputation? How can it be measured? These two questions have been widely discussed by academics, without coming to a shared definition or evaluation methodology.
Each research defined corporate reputation according to its specific field: economic literature views it as the results of the firm's past strategic choices, while strategic researchers define it as a resource. In marketing, corporate reputation is improperly associated with the concept of brand image. Sociologists focus on the role of the interpretation of a company's actions done by intermediaries in a world of incomplete information. Organizational theories partially share the sociological studies definition, but it focusses on institutions support for the company's long-term success. Although there isn't a common definition, all the studies agree on one point: corporate reputation is the result of the relationship between a company and its stakeholders. On the stakeholder's opinion rely most of the corporate reputation measurement techniques, such as the Fortune's American Most Admired Companies or the Reputation Institute's indicators (Reputation quotient and its "child" RepTrack), that have been proposed during the past years, techniques that were criticized.
In this work, we have investigated if *corporate reputation can be evaluated from social media data*
The web 2.0 and the emergence of social media created new form of communication, they enabled people to share their opinion with others on a global scale, often by talking with strangers about subject that are barely known. Even if the communication is mostly "impersonal" still an underlying shared opinion can be evaluated.
We focused on the Volkswagen scandal and the buzz it created within the Twitter social network. VW's scandal was chosen because its widely covered evolution through time and its broad effects on VW's financial performance.
In order to fulfill the research goal, tweets about VW were collected. The dataset contains tweets from three weeks before the scandal breakout date until nine months after it came out.
This vast dataset was firstly analyzed and not VW's related elements were removed. The remaining part of it was then classified in two main groups: tweets about VW, but not related to the scandal, and the ones that specifically referred to VW's wrongdoing. Once the two sets were obtained, each of their elements were evaluated with a sentiment analysis software and after the opinion extraction was calculated the daily aggregated sentiment through a custom-built formula, defined

to adapt to the Twitter domain. This aggregation produced two different daily sentiment score: one for the general public opinion about VW and one for the users' judgement about the scandal. In the final part of the research we analyzed the correlation between VW's stock price and the scandal sentiment.
The work led to excellent results in the not-relevant elements removal phase and the classification one, but the opinion aggregation did not produce significant outcomes. This final results should not be considered as a research drawback, instead they represent a starting point for further analysis which will explore in more depth the on-line community opinion creation process.

Contents

List of Figures

List of Tables

List of Abbreviations

DB	**D**ata**b**ase
DBMS	**D**ata**b**ase **M**anagement **S**ystem
GUI	**G**raphical **U**ser **I**nterface
JSON	**J**ava**S**cript **O**bject **N**otation
NLP	**N**atural **L**anguage **P**rocessing
NLTK	Python's **N**atural **L**anguage **T**ool**k**it
SA	**S**entiment **A**nalysis
SQL	**S**tructured **Q**uery **L**anguage
SW	**S**oft**W**are
UM	**U**ser **M**ention
URL	**U**niform **R**esource **L**ocator
VW	**V**olks**w**agen

To all those people that didn't allow me to stop seeing the world through the eyes of a child

Chapter 1

Introduction

1.1 Background and motivations

The globalised economy raised the attention on the identification of the sustainable competitive advantage sources. The search for these advantage drivers broadened arriving to embrace not only the tangibles, but also those that are located in the field of intangibles [1]. This is not surprising, since most United States executives consider corporate reputation one of the most influential factor for the firm's success [2].
Corporate reputation is a reflection of how an organisation is regarded by its stakeholders [3] and it affects the way various stakeholders behave towards an organisation [4]. It can help the organisation obtain trust and credibility in society, which will assist in the achievement of its objectives and goals [5] [6]. It is an important source of goodwill when dealing with crises; it can be a competitive advantage, constitute a mobility barrier [7] and allows the organisation to attract the best employees and ensures their loyalty [8]. A favourable reputation encourages shareholders to invest in a company; it attracts good staff, retains customers and correlates with superior overall returns [9] [10].
However, many of these benefits have been criticised, because based on flawed measures of reputation or rely on conceptualisations of reputation that are unclear [4]. Wartick [11] argued that "one cannot talk about measuring something until one knows what that something is" [11].
This work presents a different approach to measure corporate reputation. The main idea is that general public opinion could be considered a good proxy of how the company is regarded by its stakeholders.
In order to obtain this result big data analysis and sentimen extraction were combined. Millions of users share their thoughts and considerations about different aspects and events through social media. This research focus on data from the

popular micro-blogging platform Twitter, because it could be considered as a rich source of information for decision making and sentiment analysis. Sentiment analysis refers to a classification problem where the main focus is automatically classify a text into positive and negative feelings with the aim of identifying attitude and opinions that are expressed in any form or language. Sentiment analysis over Twitter offers organisations a fast and effective way to monitor the publics' feelings towards their brand, business, directors, etc.
The data collected refer to the September 2015 Volkswagen diesel scandal, because of its duration, majour press coverage and effects it had on VW's stock prices. The proposed methodology divides the dataset in two majour category: tweets about VW's wrongdoing and the others, which do not refer to the scandal. Experimental results in the end link the general public opinion to VW's stockmarket performance.

1.2 Aims and objectives

This study has three aims. Firstly, divide the dataset in the two category: scandal related and not scandal related. Secondly, evaluate and aggregate the sentiment from the collected tweets. Thirdly, link the obtained sentiment to VW's financial performance.
In order to achieve these goals several research objectives are needed:

- Analysing the data
- Remove not relevant data from the dataset
- Defining a set of classification rules to discriminate the data
- Evaluate the single tweet sentiment
- Define a sentiment aggregation rule

1.3 Research questions

This study tries to answer the following questions:

1. Which is the best methodology to divide the dataset into the two groups?
2. How can the sentiment be aggregated?
3. Sentiment analysis could help to understand the financial market behaviour?

Chapter 2

Corporate reputation

2.1 Introduction

During the last few decades the consideration of intangible assets as an important factor in the pursuit of competitive advantage has gained an increasing interest in both the academic and industry sectors [12]. Regarding the industries, in his work Hall observes that CEOs consistently rank reputation as one of the most important intangible assets, and recommends that this issue should receive constant management attention [13].

Porter's [14] analysis of generic strategies raised the first attention about corporate reputation, but Fombrun and van Riel's [15] work started a new era of reputation studies based on the Corporate Reputation Review and the identification of several key research problems related to the theme.

Fombrun [16] provides the following definition of corporate reputation:

> A corporate reputation is a collective representation of a company's past actions and future prospects that describes how key resource providers interpret a company's initiatives and assess its ability to deliver valued outcomes.

From Fombrun's research comes to light that the lack of research was partially due to a problem of definition, as corporate reputation was defined in different ways by different schools of thought [15].

The concept of reputation has appeared in a large number of contexts, in both the academic literature and popular use. For example, often for the public, firms offering warranties are often said to be cultivating reputations for high quality, advertising campaigns are designed to create a reputation for trendiness, forecasters are said to have reputations for accuracy, or advisors for giving useful counsel [17].

In literature numerous theories have been used to examine the concept of corporate reputation, theories that often continued to use the following list of constructs as synonym to 'corporate reputation': image, identity, prestige, goodwill, esteem, and standing [11]. Depending on the field of study each of the previous terms has been offered as

- a broader term that encompasses reputation,
- an important component within reputation,
- the equivalent of reputation.

In other words, these terms are intended as bigger than, smaller than, or just the same as corporate reputation.
Despite the fact that reputation assumes different names in different context, one common feature should be considered: reputation and its literature synonyms all refer to a perceptual representation of a company's past actions and future prospects which describe the firm's overall appeal to all of its key constituents when compared with other leading rivals. The key points of this definition are that corporate reputation: i) has a perceptual nature, ii) is the result of the aggregation of the perceptions of all stakeholders and iii) is comparative.

2.2 Stakeholder

Stakeholders were initially defined as "those groups without whose support the organisation would cease to exist [18]." Over time, the definition of the word 'stakeholder' has broadened to comprehend "any group or individual who can affect or is affected by the achievement of the organisation's objectives." This includes not only the directors or trustees on its governing board (who are stakeholders in the traditional sense of the word), but also all entities which might be directly or indirectly affected by the company's actions, such as governments, media, local communities and the natural environment. The role of stakeholders in relation to the organisation is to:

1. Set expectations about the future organisation's behaviour and products, expectations which are rooted in the previous knowledge that the stakeholder has about the organisation;
2. Direcltly or indirectly experience the effects of the organisation's actions;
3. Evaluate outcomes, according with the personal judging criteria; and

4. Act on these evaluations, redefining their expectations by incorporating the newly acquired knowledge derived from the previous steps [19].

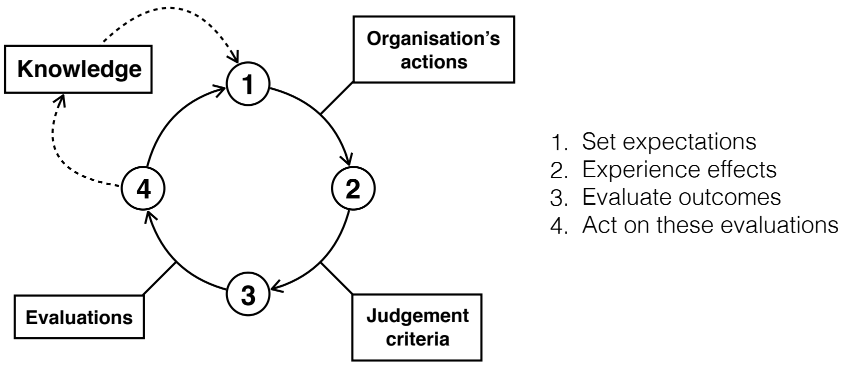


FIGURE 2.1: Stakeholder's expectation evolution cycle

From the previous stakeholder's expectation evolution cycle (see figure 2.1) is possible to deduce that organisations will be likely to have different reputations with different stakeholder groups. The evaluation criteria stakeholders use to judge an organisation's reputation will differ depending on the particular stakeholder's expectations of the organisation's role. Is important to notice that stakeholders' expectations are dynamic, and thus likely to change over time [20]. Additionally, as an organisation's reputation increases, so do stakeholders' expectations [21]. Thus, the organisation will have a different reputation with each of the stakeholder group, and may have different reputations with single group members, as expectations vary individually, but, even if there might not be an unanimity, an overall corporate reputation still results. Corporate reputation in Fombrun's definition is an unidimensional 'global perception', which crystallises the collective judgments into reputational orderings of firms [22] and that can be seen as "the degree to which the overall appeal is shared throughout the stakeholders [11]."

2.3 Corporate reputation: the combination of Identity and Image

In various literatures different terms have been used as synonyms to corporate reputation, but a more careful look at those words suggests that each construct should be looking at perceptions originating from different types of groups, while corporate reputation is the overall perception.
Bromley [11] divides the different perceptions into two main groups:

1. Corporate Image: how the organisation presents itself,

2. Corporate Identity: the concept of the organisation.

The 'corporate reputation', which results from the combination of these two groups of perceptions, is how the key stakeholders and other interested parties actually conceptualise that organisation, which implicitily specifies that image and identy are less important than reputation (see figure 2.2).

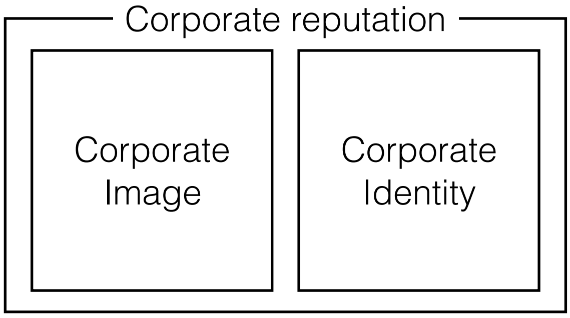


FIGURE 2.2: The relationship between corporate reputation, corporate image and corporate identity

2.3.1 Corporate image: how the company is seen

It is important to distinguish corporate image from corporate reputation as usually people consider these terms interchangeably. Corporate image is the internal and external observers' general impression of a corporation's distinct collection of symbols, such as the company's name and logo, its advertising campaigns and press releases, the required dress-code and members' standing.
Usually, a common definition of image in the context of reputation is a "summary of the impressions or perceptions held by external stakeholders" [23]. As can be observed, the definition mainly focuses on the external stakeholders, so the company's image is defined not as the company's belief, but as what external stakeholders believe or feel about the company from their experiences and observation. This corporate reputation definition is close to the definitions of 'corporate image' used in marketing such as "attitudes and feelings consumers have about the nature of the company" [4] or image is "what comes to mind when one hears the name or sees the logo" [24] of a particular firm. This similarity explains why in the marketing literature the terms image/brand and reputation are often used interchangeably without making any clear relationship between what can be usefully seen as two distinct concepts [4].
The creation of the corporate image is mainly a communication process. Process which involves public relations, marketing and other organisational activities that

attempt to shape people's impression of the firm.
Corporate reputation distinguishes from corporate image also because it has an accumulated historical meaning. Reputation evolves over time as a result of consistent performance whereas image is fashioned more quickly through well-conceived communication programmes [24]. Reputation here differs from image in that: image concerns the public's current belief about an organisation, while reputation presents an opinion built over time and focused on the company's behaviour. This distinction between image and reputation comprehends the deeper consideration that stakeholders can have in mind the image of an organisation without any real experience of it, while reputation implies something grounded in experience.
The corporate image can be shaped but not controlled by an organisation because factors such as media coverage, governmental regulations and surveillance, industry dynamics and other external forces also geatly influence the overall impression of the firm.

2.3.2 Corporate identity: what the company is

Identity is not just the image of the firm held by an inside stakeholder (e.g. employees), but rather, the underlying 'core' or basic character of the firm [25]. Corporate identity doesn't indicate only the process of internal stakeholder identification with the company, but rather refers to what the organisation actually is [26].
In literature there are two main expressions, often used as synonyms: organisational identity and corporate identity. The definition of the concepts given in literature are similar, reflecting the existence of an interdependence between the two, but should be noted that the expressions are on two different abstraction levels. Organisational identity is the answer to the questions "who are we?" or "how do we see ourselves?", in other words, the employees' perception of the firm. It indicates what company members perceive, feel and think about their company [27] and concerns those organisational characteristics that are most central, enduring, and distinctive. Organisational identity is a similar concept to company culture: organisational identity is "how we see ourselves" and culture is "how we do things around here". As defined by Albert and Whetten corporate culture is an "enduring and distinctive characteristic" of a firm and Barney [28] states that company culture can be a source of competitive advantage when it involves a unique personality, history and experiences of those who work within the organisation. Due to its pervasive nature within the company, culture is not something that can be easily changed by top management strategy [29] nor something that is readily manipulatable [30]. On the other hand, Hatch defines the link between

organisational identity and corporate culture saying that "how we define and experience ourselves is influenced by our beliefs which are grounded in and justified by cultural assumptions and values" [31]. Organisational identity can be seen as how people understand themselves in relation to the company culture and shared values. Thus, culture is a consequence of organisational identity and then culture can be changed only when organisational identity changes. While organisational identity refers to organisation members' perceptions of the firm influenced by the company culture, corporate identity refers to visual cues, such as name, logo and symbols and the strategic signals of organisational identity, such as vision, mission and philosophy, which are conceptualised as part of the strategic process linking corporate strategy to company image and reputation. Corporate identity reflects the unique characteristics or 'firm personality' rooted in the behaviour of members of the organisation and should reflect how employees identify themselves with the company [32]. Follows that corporate identity is a broader concept than organisational identity. Corporate identity is the composition of both the members' perceptions of the company (organisational identity) and the part of the external corporate image involving any public relations effect [33], sometimes referred as "corporate desired identity" [4].

2.3.3 The corporate reputation definition problem

Corporate reputation has been viewed from complementary perspectives by economists, strategists, sociologists, marketers and organisation theorists [15].

The economic perspective considers corporate reputation a dynamic construct based on a series of sequential strategic moves, in which firms' strategic choices work as signals sent to observers, who use these signals to form impressions about the firms [34].

Strategy theories define corporate reputation as strategic resources that are able to assure competitive advantage by creating mobility barriers or providing differentiation opportunities [34]. Reputation can constitute mobility barriers, as it can become part of the industry structure and is difficult to imitate and modify [12], [16]. Porter's [14], in his analysis of generic strategies, points out that good reputation is a key element of differentiation, as an organisation that pursues this kind of strategy needs the public to know what makes its offer better than others in the market.

Sociologists call attention to the process of social construction implicit in reputations [16]. In a world of incomplete information, the public has to interpret signals sent by firms and often relies on intermediaries to do so [22] [35]. According to

sociologists' logic, this is all part of a socio-cognitive process that has to be taken into consideration when studying reputation [34].
Marketers tend to see reputation as the result of companies' efforts to induce purchases and create loyalty [16]. Reputation is often treated by marketers as synonymous of brand image or brand equity, and the focus is on the process of building this image through the use of various marketing tools [34].
Organisational-study theories, such as institutional theory and stakeholder theory view reputation as being focused on the process of gaining legitimacy with actors in institutional environments. Rao [35] observes that firms have to gain legitimacy and cultural support in their institutional contexts in order to build reputations. Stakeholder theory shows that a good reputation with key stakeholders is necessary to guarantee their support, which is essential for long-term success [36].
The lack of one shared perspective led to a variety of definition of corporate reputation, which are a direct consequence of the fact that, as stated by Barnett, "[researchers] often were not aware of the diverse perspectives" [26].

2.3.4 Corporate reputation in the era of Social media

The advent of Web 2.0 and the emergence of social media has empowered consumers and has led to a change of behaviour, since they have become the key contributors on the Web's contents by sharing information and experiences, thus, influencing peers in their decision-making process of products and services [37] and make buying decisions dependent on this information without companies being able to influence such processes [38]. Gurau [39] summarises this impact of social media on the behaviour of online audiences as follows:

- Social media constitute a network which allows a bidirectional communication between companies and their audiences;
- Users are connected to one another, enabling discussions and debates about companies and their product;
- Users are now able to filter the information that they receive
- Users can easily access to various sources of information through the spread of the Web and discussions with other stakeholders of the companies

The previous considerations imply that there has been a partial shift of dominance over information flows from the companies to stakeholders [40], especially consumers and other on-line opinion leaders. Also Aula [41] argues that companies lose their control over stakeholder relations and communications between their various stakeholder groups. This loss of control over information flows, can affect corporate reputation significantly. Aula [41] summarises three scenarios in which social media can increase the reputation risk that companies are exposed to:

- The content in social media is user-generated, usually that produced content did not undergo a revision process, therefore it might be false or might differ from what the companies are willing to share;
- Social media can enhance expectations (e.g. in the ethical behaviour and transparency of companies), which companies might not be able to fulfil;
- The dialogues and behaviour of companies in social media can endorse reputational risk, including reactions to conversations held, or the manipulation of information and activist influencers in social media.

Besides these reputation risk dynamics, companies can actively influence online users, and, thus, protect their reputation by engaging in social media [40].
A concept simultaneously appearing with social media is "online reputation". Online reputation represents the reputation of companies established in the Internet [42] and is predominantly built by community participation and through the so-called 'reputation aggregators', such as search engines like Google, which enable people to find content online [40].
Community participation is achieved by a consistent communication towards the companies' stakeholders, but reputation can even be formed through indirect experiences with the firm triggered through word-of-mouth, the media, or other platforms. It is important that companies' communicators know what is being said about their company on the Internet and that they participate in communicating their positions on key issues in forums, blogs, podcasts and other ways that reach their intended stakeholders. Employee blogs and e-mail are increasingly a means of enhancing corporate reputation or they can generate problems for the company.
Clearly, companies are just beginning to learn how to deal with these new communication channels and how to manage the influence on stakeholders' opinion that anyone with a computer linked to the Internet can generate.

2.4 Measuring corporate reputation

2.4.1 Existing methodologies

Corporate reputation is not an "easy variable to accurately measure" [34]. Multiple disciplines studied the key aspects of the definition of corporate reputation [43] and various heterogeneous approaches have been used to measure the corporate reputation construct. Both, definitions as well as measuring tools have been lively discussed with regard to their usability and validity[44]. Basically every measurement approach is linked to a definition of what actually is measured [11].
Geller observes that "although some rankings are verifiable and replicable, they tend to give disproportionate importance to a few stakeholders, resulting in a biased perception as other key stakeholders are excluded from the analysis" [34].
In this scenario fall the methodologies mostly used to evaluate corporate reputation: Fortune's indicators (AMAC and GMAC) and the Reputation Institute's indicators (Reputation Quotient and RepTrack).
The following part of this section briefly analyses those methodologies, together with other quantitative techniques, and presents the main issues raised by academics.

Fortune's indicators

Fortune's american ratings AMAC and its global counterpart, the GMAC, are published by the North American magazine Fortune. Until 1997 the Fortune's AMAC was the only reputation ranking available on a global level, but it was restricted to US firms.
The AMAC indicator rates 1000 U.S. manufacturing and service firms [45], while the GMAC evaluates 1,000 leading North American companies, complemented with 500 international companies. Both indexes rank the organisations in terms of financial results, best performance and turnover. The evaluation group includes executives, managers and financial analysts of each sector who are familiar with the companies that are being evaluated.
The overall ratings are obtained through the aggregation of the evaluations by attributes made by the respondents.

Reputation Institute's indicators

In 1999, Fombrun *et al.* highlighted some limitations of the indexes used to measure corporate reputation, especially: *(i)* the lack of content validity and *(ii)* the restricted perceptions elicited by the surveys, both points due to the the fact that the group of respondents is constituted by corporate leaders and financial analysts [45].
Recognising a growing need by both practitioners and academics for a better conceptual and empirical tool for assessing and managing reputation and to overcome the available methodologies' inadequate representation of all stakeholders, the Reputation Institute launched a global project in 1998 to understand and measure the diverse factors associated with corporate reputation [46]. The first measurement instrument that resulted from the exploration was the Reputation Quotient, which evolved into RepTrak in 2005.

The Reputation Quotient is an index based on surveying general population and which aims to find out which companies are liked and respected by individuals, and for what reasons.
The RepTrak methodology emerged in 2006 as a replacement for its predecessor and "was created to provide executives with an analytical instrument that could be used, not only to track and assess stakeholder perceptions of companies, but that would also enable a more comprehensive understanding of the underlying informational drivers of reputation that elicit emotional attachment" [46].
RepTrak contains new dimensions and new attributes from the Reputation Quotient.
The evaluation is held in the form of a poll in different countries, with respondents looking at one, two or three companies that they are familiar with, continuously throughout the year, grading it on a scale from 1 to 7.

Quantitative methodologies

Beside the previously presented methodologies exist quantitative approaches that tend to overcome the weaknesses of the qualitative methodologies.
The most recognised quantitative methodologies are:

- Intellectual Capital approach
- Accounting approach
- Marketing approach

The first two approaches rely on the consideration that there is a gap between the market price of a listed company and its book value. The difference between the two values relates, in part, to the value of the company's intangible assets, among which reputation [47]. The *Intellectual Capital* approach is evaluates the difference between market price and book value through the appropriate estimation of 5 dimensions: trademark, service marks, copyrights, authorisations and exclusive rights [48]. The *Accounting approach* is more focused on the assessment of corporate intangible assets. This approach requires an analysis of reputation associated to assets and liabilities and their valuation at fair value. The difference among those generates the reputation value [47].
Finally the *Marketing approach* suggests to measure Corporate Reputation through the concept of brand equity. Among the possible approaches to evaluate the company's brand, the most visible is through the amount of royalties that the market would pay to gain the grant of a trademark [47].

Considerations on the presented methodologies

All the presented methodologies have been criticised in the academic literature, because some lacks in their formulation.
Criticisms of the two methodologies developed by Fortune magazine are similar and point to the fact that: *(i)* the indexes consider only some stakeholders, *(ii)* measure assesses little beyond financial performance, *(iii)* valuations may not correspond to the reality and *(iv)* do not incorporate a multistakeholder vision [1], [4], [22], [44], [45], [49], [50].
The general public interview approach of the two models developed by the Reputation Institute has been argued because *(i)* it embeds excessive subjectivity in the methodology [48], *(ii)* often respondents lack of sufficient information to have an valuable opinion and *(iii)* for many companies and sectors, consumers can not be considered the most important stakeholders [50]. Geller also argues that, even if the Reputation Institute's methods are a clear step forward in measuring corporate reputation from various stakeholders' point of view, they "still fail to respond to the issue of how to use them with different stakeholder groups, and how to aggregate these perceptions" [34]. According to Cerchiello [48] another important criticism is the static measurement nature and the absence of adjustable weights in the evaluation of the opinions of the different stakeholder groups within the Reputation Quotient methodology.
The quantitative methodologies have been criticised too.

The intellectual capital approach's evident limit lays on the heterogeneity of the differences in the balance sheets, which do not allow a comparison among several companies. Moreover, it doesn't cover sudden events that can seriously affect the reputation of an organisation [48].
The accounting approach is based on the evaluation and analysis of intangible assets, however there is not a clear a criteria for fair value assessment [48].
The marketing approach relies on the evaluation of the corporate brand, which represents only one dimension and thus can not explain all the aspects related to the reputation concept [48].

2.4.2 Measuring corporate reputation through social media

The Reputation of an entity has been described as "the result of a public judgment that increases (or decreases) over the time and it is socially shared by different stakeholders" [51]. In a social network, members influence each other by sharing their experiences and opinions, constantly modifying the overall network's judgement on the discussed entities [52].
This constant evolution of the stakeholder's perception is stronger within web-based social networks, because they are characterised by "easy searching, open participation, a minimal publishing threshold, dialogue, community networking and the rapid and broad spread of information and other content via a wide range of feedback and linking systems" [41].
Social media and web 2.0 have strengthened the role of reputation in a wide variety of economic decision making processes: consumers consult social media based reputation in their choice of brands, talented employees are sensitive to social media based reputation in deciding to exit or remain with a company and investors increasingly use analysis of social media sentiment as part of their investment decisions [51].
Attitude towards companies is increasingly shaped by the opinions and feelings that circulate within digital networks. Traditionally news media were the main channel for stakeholders to gain knowledge about corporate reputation, however, were difficult to directly experience or observe [53]. Today, more and more people gain knowledge about a company by searching and interpreting online signals.
These signals are no longer only based on the range of comparisons between companies with similar offerings, but also on how a social network perceives the performance and quality of a company. Once people have built a picture, they share

their opinions and feelings with others and "the subjective truth turns into a collective truth about what an organisation is and what it should be" [41].
The evaluation of users' opinions and sentiments in online social media can be considered a good proxy of company reputation, therefore from the data produced in social media sites it is possible to extract, monitor and even predict corporate reputation trends by aggregating subjective opinions using data mining techniques [51]. Among the various techniques to analyse big data, sentiment analysis represents the best tool to interpret the large amount of data available on social media.

2.4.3 Sentiment analysis

Sentiment analysis is the area of research that attempts to make automatic systems to determine human opinion from written natural language. Is part of the affective computing paradigm and refers to the process of categorisation of unstructured human-authored documents "based on their affective orientation, meaning the emotional attitude of the person expressing the opinion" [51]. It is a discipline at the crossroads of information retrieval and computational linguistics. Unlike text mining, sentiment analysis is concerned with the opinion expresses in a text instead of its topic [52].
There are three common basic approaches:

- *Machine learning*: the sentiment evaluation in the machine learning method use linguistic and/or syntactic features. This approach can be further divided into the two categories of supervised and unsupervised methods. Labelled training data is used in supervised learning methods and unlabelled data is used in unsupervised learning methods. The result is an algorithm able to detect and classify new objects according to their sentiment [51]. The main difference between the two approaches is that in the former an program for automatic detection of the sentiment is trained with the previous classified data, while in the latter this training phase is omitted.

- *Linguistic analysis*: is about inferring the sentiment valence of a text based on its grammatical structure. Linguistic analysis attempts to identify superlatives, negations, context and idioms as part of the prediction process [54].

- *Lexicon-based methods*: the most common approach for text classification, the lexicon-based approach, involves the use of a lexicon. This approach requires "the creation of a knowledge base-lexicon of affective words, with additional data characterising emotional states and relations" [51].

In this case, the input of the approach are lists of words that are pre-coded for polarity and sometimes also for strength. The algorithm uses the occurrence of the lexicons' word within texts to estimate its overall sentiment [51], [54].

Chapter 3

Methodology

3.1 Introduction

In this section will be firstly presented the phases that led to the result and after the main tools used during the analysis.

3.2 Research phases

The research could be divided into the following phases:

1. Data collection;
2. Data analysis;
3. Data cleaning;
4. Data pre-processing;
5. Classification;
6. Sentiment extraction;
7. Sentiment aggregation.

3.2.1 Data collection

The data used was a set of Twitter posts from August 28th, 2015 to June 6th, 2016 bought from the social media aggregation company GNIP.
The data consisted of 933,037 English-language tweets from 227,353 different accounts.
The collected tweets had to:

```
{
  "firstName": "John",
  "lastName": "Smith",
  "age": 25,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021"
  }
}
```

Figure 3.1: Example of a JSON object

- Contain the words vw or volkwagen or at least one of the hashtags volkswagenscandal, vwemission, vwscandal, vwgate, vwdieselgate
- Be written in English
- Be generated from a US account

The first rule tried to narrow the number of tweets to only those relative to VW company.The restriction to English was chosen to remove the complication of multiple languages and the account country location was chosen because VW scandal originated in the US.

3.2.2 Data analysis

The collected data were returned as a series of files in the JSON format, an open-standard format that uses human-readable text to transmit data objects consisting of attribute–value pairs. Figure 3.1 shows an example of a JSON object.
Each of the JSON objects collected contains:

- **id**: a unique IRI for the tweet;
- **actor**: an object representing the twitter user who tweeted, together with all metadata relevant to that user;
- **verb**: the type of action being taken by the user, which could be “post”, when a user posts a new tweet, and “share”, when a user retweets another user’s tweet using the retweet functionality;

- **generator**: an object representing the utility used to post the tweet (e.g. the indication that the user posted the tweet using his phone)
- **provider**: an object representing the provider of the activity, such as Twitter;
- **inReplyTo**: an object with the link referring to the tweet being replied to;
- **location**: an object representing the Twitter "Place" where the tweet was created;
- **geo**: point location where the Tweet was created;
- **twitterEntities**: lists of urls, mentions and hashtags contained in the tweet;
- **twitterExtendedEntities**: an object from Twitter's native data format containing "media", such as multi-photos;
- **link**: a permalink for the tweet;
- **body**: the tweet's text;
- **object**: an object representing tweet being posted or shared. When the json object refers to a retweet in this section is presented the original tweet retweetted;
- **postedTime**: the time the action occurred, e.g. the time the tweet was posted.

Not all of the contained information are relevant for the analysis, therefore in the DB were stored only a subset consisting in: id, actor, body, verb, postedTime, inReplyTo, location. From the object section were also extracted the urls, user-mentions and hashtags included in the tweet. See appendix A for the structure of the used DB.

3.2.3 Data cleaning

Data cleaning, data cleansing, or data scrubbing is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data. Data cleansing may be performed interactively with data wrangling tools, or as batch processing through scripting.
During the analysis phase emerged that some of the collected data do not refer to

VW, so before advancing to the next phase (data pre-processing) those tweets had to be removed.
GNIP data collection rule is structured in a way that:

- is case *insensitive*;
- matches all the texts that contain the specified keywords (*character sequence matching*).

Case insensitivity makes words in upper case (e.g. VW) and in lower case (e.g. vw) the same, while character sequence matching treats isolated words (e.g. "... vw messed up...") and sequences that contain the keyword as part of a larger sequence (e.g. URLs, hashtags and user mentions) equally.
The main issue within the dataset arises from the second rule, expecially when some character sequence are part of an URL. Figure 3.2 presents an example of a text that is not relevant to the analysis. In order to remove these data a dataset cleaning was need.
The cleaning was performed in two steps: firstly, the URLs were removed from

Release news: Kim's Tech Picks: Wireless chargers, Digital cameras, New movie releases and more https://t.co/cEJDfvwAIY

FIGURE 3.2: Example of a not compliant text

the text, then each text was analysed to check if it still contained the collection keywords. To perform this first cleaning step, the list of texts and their URLs were extracted from the DB, each URL in the text was replaced with a blank space and then was checked if the text still contained at least one of the collection keywords. Secondly, because some of the texts could contain the keywords not as part of a URL, a second compliance check was performed. In this second step the tweet's text was analysed using a matching rule for each keyword and, when there was a positive match, the text was evaluated to check if the match was a false positive. Figure 3.3 shows an example of a text that is a false positive match.
The cleaning phase highligthed that 252.569 data were not about VW, thus the

@vw_ward: In the race to save species, is it ethical to use GMOs? http://t.co/IXx595MhKI via @ensiamedia

FIGURE 3.3: Example of a text that produces a false positive match

dataset was reduced to 680.468 relevant tweets. Figure 3.4 shows the percentages of compliant and not compliant tweets within the dataset.

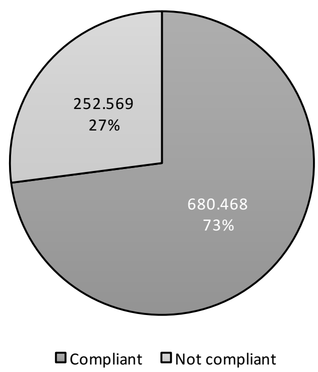


FIGURE 3.4: Data cleaning results

3.2.4 Data pre-processing

Beside of the structure of the collected data, the raw elements in the dataset need to be preprocessed in order to obtain the information necessary for the analysis. In text mining this phase is considered a crucial one, because from it depends the accuracy of the subsequent analysis.
For each raw tweet, the entity was analysed to extract the useful information and after the extraction, the information were stored in the DB, to make them available for the next analysis phases.
The tweet text was elaborated in order to align it to the required sentiment analysis tool format. The operation that were performed on the text are:

- Lowercase conversion;
- URLs removal;
- Retweet indication removal;
- User mention removal;
- Hashtag removal;
- Emoticon removal;
- Stop-words removal;
- Tokenisation;

Lowercase conversion, also known as case folding, is the process which converts the text into lowercase letters in order to group a word that has been written in uppercase and lowercase letters as the same token. It could be argued that in unformal English language, such as the one under analysis, a word with uppercase

letters might have a different meaning than if the same word is in lowercase letters, for example when the word happy is written in uppercase (HAPPY) usually is because the user wants to add more enphasis on that specific word. In this stage although the aim is not to judge the emphasis of each word, but to reduce the possible set of different words presented to the classifier algorithm, thus enhancing its efficiency.
URLs are part of the text which carry no information. Because URLs are frequently included in tweets but are useless in text mining, they can be removed from the text. The same argument applies to the indication that a tweet is a retweet automatically added by Twitter. In this latter case the need of indication removal is made more evidend by the fact that the identification is stored within a field, the verb field, in the DB.
Tokenization is the process of breaking a sentence into words, symbols or phrases; these are then referred to as tokens.
The previous text manipulation were performed using specific scripts written in Python language.
The first one simply used the string conversion function provided with the programming language. The following two instead relied on a combination of DB interrogations and regular expressions application. By querying the DB the list of URLs associated with each tweet was created and then the specific portion of the text which contained the URL was removed, but some of the tweets contained other incomplete URLs which were removed by using a specific regular expression that substituted the URL with a blank space.
The tokenization was performed by using a method contained in an additional Python's library specific for natural language processing, NLTK (please refer to subsection 3.3.2 for a description of the NLTK).

3.2.5 Classification

To distinguish betweet the tweets associated with the VW's scandal and the one that Naive Bayes classifier is a supervised learning algorithm based on applying Bayes' theorem with the "naive" assumption of independence between every pair of features. Bayes' Theorem is a formula that calculates a probability by counting the frequency of values and combinations of values in the historical data.
The theorem finds the probability of an event occurring given the probability of another event that has already occurred. If B represents the dependent event and

A represents the prior event, Bayes' theorem can be stated as follows:

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)} \tag{3.1}$$

To calculate the probability of B given A, the algorithm counts the number of cases where A and B occur together and divides it by the number of cases where A occurs alone.
In spite of its apparently over-simplified assumptions, this classifier have worked quite well in many real-world situations, famously document classification and spam filtering. It requires a small amount of training data to estimate the necessary parameters and can be extremely fast compared to more sophisticated methods.
Three differen version of the text were provided to the algorithm: the full text without URLs and retweet indication, the text epurated of the UM and the text without hashtags. Usually only one pre-processed tokenised text is used, but in this case, because of the sentence length limitation to 140 characters imposed by Twitter, there was the risk that some important information were lost.
To train the classifier a set of 1000 tweets was randomly extracted from the dataset and manually labelled. The labelling was done on a binary base, where:

- **Y**: if the tweet is related to the VW's scandal
- **N**: otherwise

Instead to test the accuracy of the classifier 100 randomly extracted tweets were used. This test set was manually labelled and the result of the classification were compared with the given label.
Together with the training set, a set of classification rules was developed. These rules are a combination of:

- a various length list of specific keyword
- lenght of the token
- presence or absence of emoticons
- presence or absence of stop-words

The list of specific keywords was created from the training phase of the classifier. NLTK has a method, show_most_informative_features, that allows the user to check the most informative features derived from training phase of the classifier. To obtain this list of features the classifier was firstly trained by using all the words in the training set as features and then the tokens with an information rate higher than 1 to 1 were inserted into a file. This approach led to a list of over 600 words,

which was reduced by evaluating the accuracy of the classifier while the number of element was gradually removing elements from the list.
The length of the token was set into a range where: (i) all the words were kept into the text before its tokenization, (ii) only the words with at least two charactes were keps and (iii) only the words longer than two charactes were kept.
The emoticons and stop-words removal was a delicate decision, because:

- *emoticons* became a big part of online communication, so it is very likely that most of the Twitter's texts contain at least one emoji, but on the other hand no emoticon was included in the list of features utilised as features;
- *stop-words* are functional words which carry no information. Because they are frequently used in written language but are useless in text mining, usually they can be removed from the text.

By evaluating the accuracy of the classification using different combinations of the four main rules presented above, the classification was done by:

- using a list of the 400 most informative features
- removing words shorter than three letters
- keeping both emoticon and stop-words

Text version	No URLs	No UMs	No hashtags
Accuracy	81%	82%	77%

Table 3.1: Classifier accuracy on the test set

Table 3.1 reports the accuracy of the classifier on the three versions of text.
To further check the accuracy of the classifier another subset of the dataset was used. All the tweets posted before the VW's scandal breakout date are not related to the scandal, so the classification algorithm was applied to this set in order to check its performances. The main difference in this case from the previous one is that the classification was checked only for negative labels. Table 3.2 reports the accuracy of the classifier on the pre-scandal set.

Text version	No URLs	No UMs	No hashtags
Accuracy	92,97%	93,04%	93,26%

Table 3.2: Classifier accuracy on the pre-scandal set

3.2.6 Sentiment extraction and aggregation

For the sentiment extraction was used the SW Sentistrength (see subsection 3.3.3 for the description of Sentistrength SW). From each version of the text stored in the DB the SW returned a couple: the overall positive sentiment and negative sentiment associated with the words contained in the text. The global sentiment of the text was measured by adding these two values.
After that each global sentiment was produced the classification done in the previous classification phase was used to derive the daily sentiment trend associated with the scandal and not scandal related tweets.
To do so a modified plurality voting methodology was used. The classic plurality methodology is "the method with which many of us are most familiar" or in other words the aggregated sentiment is the sum of the sentiment scores within the given time frame.
The classic plurality methodology performs well when dealing with one time event (e.g. an election), but it can lead to distorted results when the "voting" is reiterated over time.
The research described in this work falls into this second case, so in order to produce consistent results the daily sentiment was normalised by dividing the score by the number of elements involved.
Another problem that arised from the domain is that different users have different posting behaviours, for example there are users (e.g. commercial Twitter users) that post tweets almost every hour while there are others that post tweets every once in a while. If the former users post tweets with positive sentiment then influence of the possible negative sentiment of the latter ones could be obfuscated, thus once again leading to distorted results. To tackle this issue an alysis on the number of tweets posted by one user in the specific time frame was perfomed.
Whenever one actor produced more than one tweet in one day then was the average sentiment of all his tweets was used instead of the sum of the sentiment of all his tweets. Also whenever one user posted more than once the same tweet then only one of those was used in the aggregation.

3.3 Tools for the analysis

3.3.1 MySQL database

MySQL is an open-source relational database management system (RDBMS) based on Structured Query Language (SQL). This DBMS was chosen because it offers:

- Cross-platform support;
- Updatable views;
- Easy scalability;
- Various graphical user interface (GUI) to define and manage the DB;
- Backup support.

For this research were used the Community edition of MySQL, because its wide support and coverage from the online community, together with MySQL Workbench as GUI.

3.3.2 Python's Natural Language Toolkit

NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet and various corpus, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning and an active discussion forum.
NLTK is intended to support research and teaching in NLP or closely related areas, including empirical linguistics, cognitive science, artificial intelligence, information retrieval, and machine learning. It has been used successfully as a teaching tool, as an individual study tool, and as a platform for prototyping and building research systems.

3.3.3 Sentistrength

SentiStrength is a program designed to identify positive and negative sentiment strength in short informal social web text and has been applied to comments in the SNS MySpace.
SentiStrength algorithm classifies for positive and negative sentiment on a scale of 1 (no sentiment) to 5 (very strong positive/negative sentiment). Each classified

text is given both a positive and negative score, and texts may be simultaneously positive and negative. For instance, "Luv u miss u", would be rated as moderately positive (3) and slightly negative (-2). SentiStrength combines a lexicon - a lookup table of sentiment-bearing words with associated strengths on a scale of 2 to 5 - with a set of additional linguistic rules for spelling correction, negations, booster words (e.g., very), emoticons and other factors. The positive sentiment score for each sentence is the highest positive sentiment score of any constituent sentence. The positive sentiment score of each sentence is essentially the highest positive sentiment score of any constituent word, after any linguistic rule modifications. The same process applies to negative sentiment strength. The special informal text procedures used by SentiStrength include a lookup table of emoticons with associated sentiment polarities and strengths, and a rule that sentiment strength is increased by 1 for words with at least two additional letters (e.g., haaaappy scores one higher than happy). The algorithm has a higher accuracy rate than standard machine learning approaches for positive sentiment and a similar accuracy rate for negative sentiment strength [54].

Chapter 4

Results and discussion

4.1 Challenges

The novelty, for the author, of the argument analysed and the techniques involved were one of the major challenge of this study. Although the large literature written on text analysis and classification, to the best knowledge of the author, none applied to the specific argument discussed in this dissertation. Moreover, most of the literature about sentiment aggregation does not cover user sentiment aggregation, forcing the author to develop his own approach.
Another challenge worth mentioning was that the stage of applying pre-processing techniques to the dataset was one of the most time consuming parts in this research.

4.2 Results

The cleaning phase led to a drastical reduction of the dataset's dimension (see figure 3.4), although if the dataset was almost halved there were still enough data to process to obtain the results.
The classification phase tried to explore different set of rules to evaluate the accuracy of the classification. This process even if was conducted by a trial-and-error approach produced great results, with an accuracy rate above 80% during normal testing and up to 90% when applied to the larger pre-scandal set.
Three kind of text were kept during this phase: (i) a text most similar to the one collected (only URLs and retweet indication removed), (ii) a text without UMs and (iii) the most minimal text without any URLs, retweet indication, user mentions (UM) and hashtags.
The classifier performed very well with all the three versions of the text and when

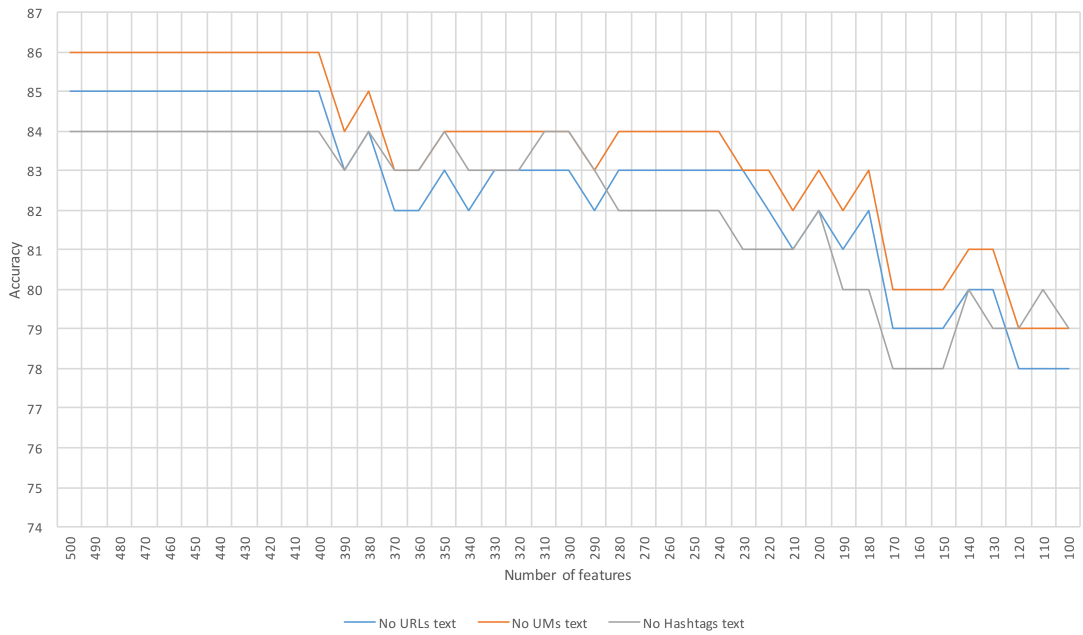


FIGURE 4.1: Classification accuracy versus Number of features used (Test set)

it was applied to the larger set of pre-VW's scandal tweets it outperformed the results obtained from the smaller test set.
The sentiment evaluation and aggregation phase, led to contradictory results:

- the general sentiment about VW remained stable during the entire evalutaion time frame. It shows some fluctuations around the neutral value, which were expected, but mostly stayed within the range -0.5 - 0.5.
- Twitter users' opinion about VW's scandal strongly changed over time, rapidly switching from very negative to positive in short time. Even if a negative trend could be observed, there are no links to any major events that could have caused the sudden changes.

When the aggregated sentiment was compared to VW's stock prices no correlation was observed (see table 4.1 for the correlation values). To evaluate the correlation the Microsoft Excel correlation built-in function was used. Next section presents some hypothesis about the causes of the observed results.

Text version	No URLs	No UMs	No hashtags
Correlation	0,0219	0,0078	0,0177

TABLE 4.1: Correlation between VW's stock price and daily sentiment

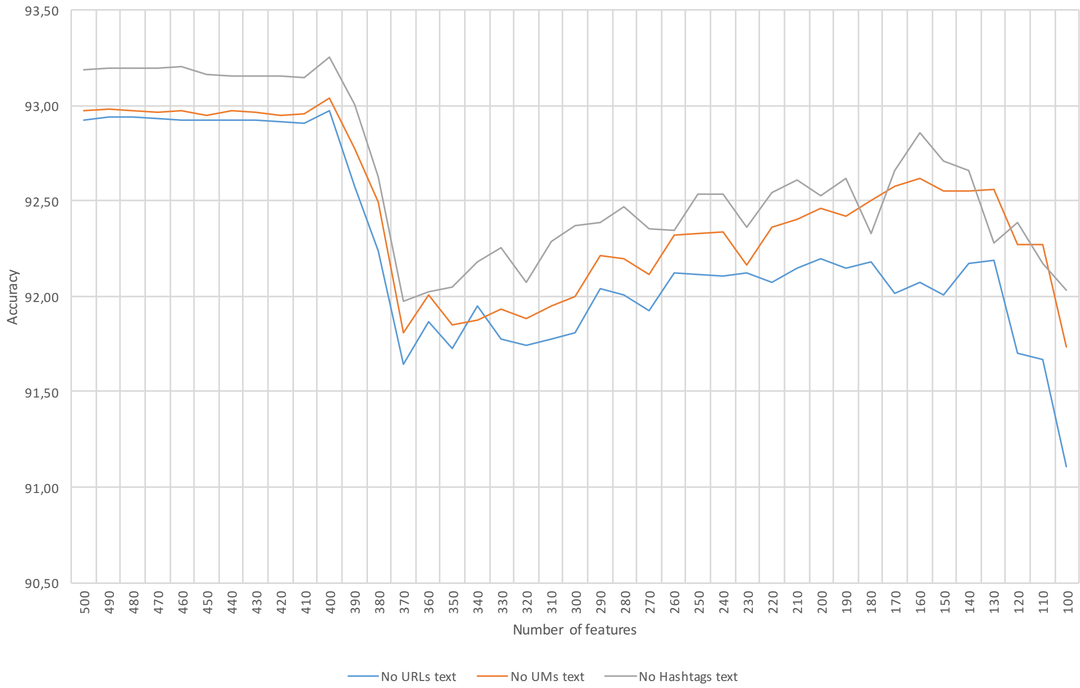


FIGURE 4.2: Classification accuracy versus Number of features used (Pre-scandal set)

4.3 Discussion

To the best knowledge of the author, the presented approach has not been applied in any other domain, thus it is a novel technique for evaluating corporate reputation.

When thinking about how people would react to a scandal is natural to think that their opinion would shift toward the negative side. Although this line of thought is generally accepted, there are things that should be considered during the analysis:

- people's opinion relies mostly on news reports, hence their opinion could be biased by the opinion expressed by the reporters;
- when a company is involved in a scandal tries to minimise the impact of the scandal on its reputation with actions (e.g. promotions and press releases) that can have a strong positive effect on people's opinion;
- word-of-mouth plays a crucial role in customers' opinion development;
- people reaction to a specific event is not only linked to the event under analysis, but it also involes their previous experience.

On the contrary financial market relies on more accurate and updated sources, such as news wires. Financial analysts have access to faster communication channels,

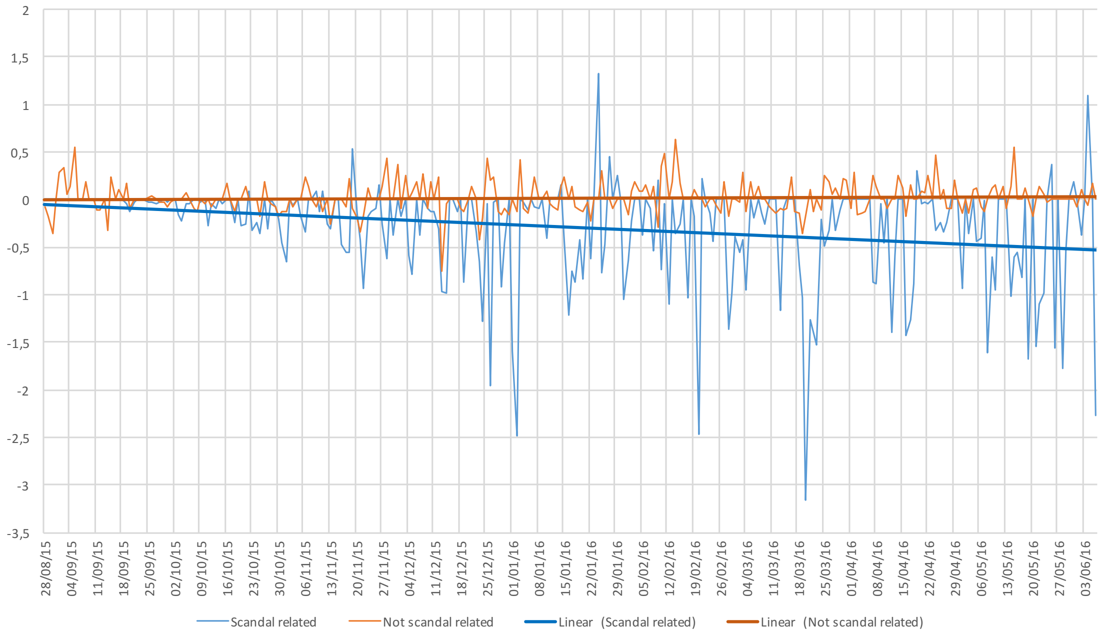


FIGURE 4.3: Daily aggregated sentiment

which are well reflected in their investment choices.
Could be concluded that a company financial performance is the result of expert analysis of a specific event (in this case VW's diesel scandal), while customers opinion is a blend of their knowledge of the event and their feeling about the entity involved in the event, feeling which not always are shared through social media.

Chapter 5

Conclusion and future work

5.1 Conclusion

To the best knowledge of the author, this is the first study done to measure corporate reputation using sentiment analysis applied to the Twitter domain.
This study blended various techniques to achieve its final goal: measuring Twitter users reaction to VW's scandal.
The most relevant information in the provided dataset were stored in a relational database (DB). The graphical user interface (GUI) provided by MySQL helped to design and refine the DB, allowing the author to store more information without reloading the entire DB and to easily check the results of the various elaborations performed. The GUI also helped in the non-compliance identification of many tweets, leading to a higher refinement of the dataset.
Python's language and its natural language processing toolkit (NLTK) turned out to be excellent for the performed tasks. Python programming language, due to its emphasis on the code readibility, led to easily debuggable scripts, limiting the need of a large usage of comments within the code. NLTK is expressly designed to work with human language data. Thanks to a hands-on guides and books introducing programming fundamentals alongside topics in computational linguistics, plus comprehensive API documentation, NLTK is suitable for linguists, engineers, students, educators, researchers, and industry users alike. The toolkit has been applied to many research topics and widely discussed within the online community.
The wide coverage of the processing techniques discussed on-line by researchers helped to identify the best strategies to pre-process the data, leading to very positive results, especially during the data subdivision phase. It might be argued that the classificaton was done naively, but, although the simple approach, the results were excellent.

The most controversial phase was the sentiment aggregation one. On one side, as expected, the general opinion about VW as company is fairly stable, with some non distinctive fluctuations around the neutral value. On the other side, the overall daily sentiment about VW's scanadal did not present an easily explainable trend, so further research is needed in order to enlighten the reason of such behaviour.

5.2 Suggestion for future work

5.2.1 Dataset cleaning

The dataset includes tweets of different nature: thoughts, opinions and commercials. Commercials tipically carry positive sentiment, but do not reflect any true opinion about an entity, so another possible study would be one in which this kind of texts are identified and removed from the dataset. In order to do so two possible approaches could be used: (i) evaluating each text content trying to identify word patters used in commercials or (ii) gathering more information about the user that posted the tweet and identifying the ones that use their account for commercial communications.

5.2.2 Text classification

During the text classification phase only single words were used as classification features. A more refined approach to the subdivision task could involve a set of couples of words, *bigrams*, that tend to appear together when a text refers to a specific domain, and other features selection rules.
Moreover, online communication is characterised by a high ratio of typing mistakes and usage of contracted words, slangs and acronyms. The addition of other pre-processing pahses, such as spell checking, slang translation and acronyms expansion, would definitely led to a better classification accuracy.

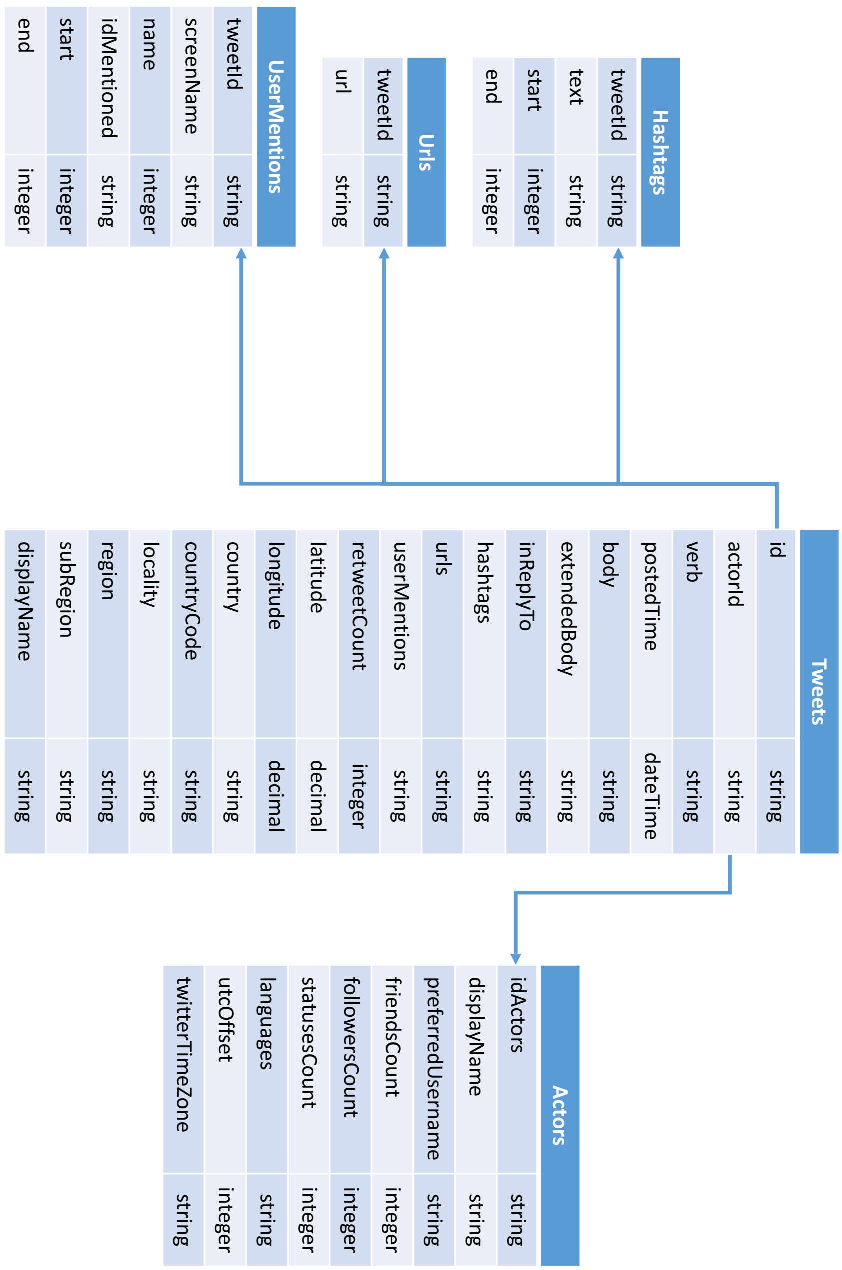
5.2.3 Sentiment aggregation

The sentiment aggregation formula marginally considered few of the information provided with the dataset: the originality of the tweet (posted or shared tweets), the influence that users have within the Twitter community (their followers count), the behaviour of the user (e.g. the amount of tweets posted). All these information

could be integrated in the aggregation formula.
Another dimension that could be worth analysing could be the daily media coverage of the scandal, the associated opinion and the link between media coverage and tweets volume.
Also would be interesting a comparison between different methodologies, evaluating their performances.

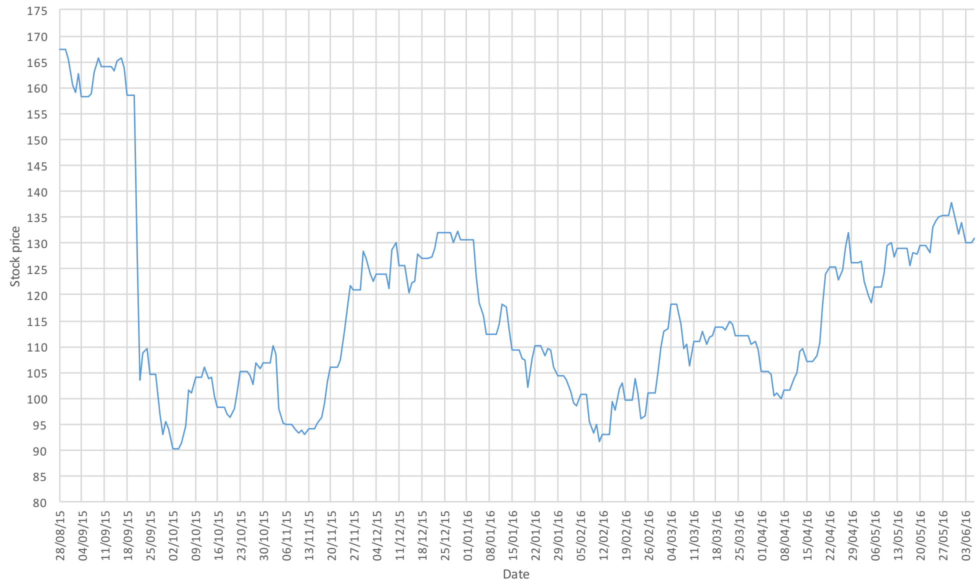
Appendix A

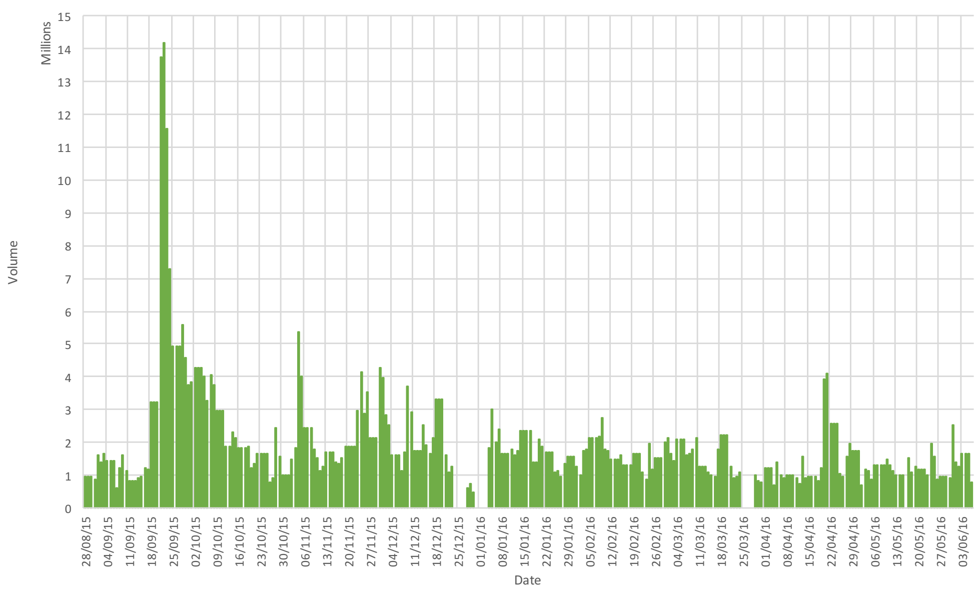
Database structure



Appendix B

VW's financial performance





Bibliography

[1] M. Schwaiger, "Components and parameters of corporate reputation - an empirical study", *Schmalenbach Business Review*, vol. 56, pp. 46–71, 2004.

[2] R. Hall, "The strategic analysis of intangible resources", *Strategic management journal*, vol. 13, no. 2, pp. 135–144, 1992.

[3] P. Feldman, R. Bahamonde, and I. Velasquez Bellido, "A new approach for measuring corporate reputation", *Revista de Administração de Empresas*, vol. 54, no. 1, pp. 53–66, 2014.

[4] R. Chun, "Corporate reputation: Meaning and measurement", *International Journal of Management Reviews*, vol. 7, no. 2, pp. 91–109, 2005.

[5] D. Baur and H. Schmitz, "Corporations and ngos: When accountability leads to co-optation", *Journal of Business Ethics*, vol. 106, no. 1, pp. 9–21, 2011.

[6] J. Mahon and S. Wartick, "Dealing with stakeholders: How reputation, credibility and framing influence the game", *Corporate Reputation Review*, vol. 6, no. 1, pp. 19–35, 2003.

[7] R. Caves and M. Porter, "From entry barriers to mobility barriers", *Quarterly Journal of Economics*, vol. 91, pp. 421–434, 1977.

[8] J. Forman and P. Argenti, "How corporate communication influences strategy implementation, reputation and the corporate brand: An exploratory qualitative study", *Corporate Reputation Review*, vol. 8, no. 3, pp. 245–264, 2005.

[9] P. Roberts and G. Dowling, "Corporate reputation and sustained superior financial performance", *Strategic Management Journal*, vol. 23, pp. 1077–1093, 2002.

[10] R. Vergin and M. Qoronfleh, "Corporate reputation and the stock market", *Business Horizons*, vol. 41, no. 1, pp. 19–26, 1998.

[11] S. Wartick, "Measuring corporate reputation: Definition and data", *Business and Society*, 2002.

[12] J. Barney, "Firm resources and sustained competitive advantage", *Journal of Management*, vol. 17, pp. 99–120, 1991.

[13] R. Hall, "A framework linking intangible resources and capabilities to sustainable competitive advantage", *Strategic Management Journal*, vol. 14, pp. 607–618, 1993.

[14] M. Porter, *Competitive strategy: Techniques for analyzing industries and competitors*. New York: The Free Press, 1980.

[15] C. Fombrun and C. van Riel, "The reputational landscape", *Corporate Reputation Review*, vol. 1, pp. 5–13, 1997.

[16] C. Fombrun, "Corporate reputations as economic assets", in *Blackwell handbook of strategic management*, Blackwell, 2005, 289–312.

[17] G. Mailath and L. Samuelson, *Repeated Games and Reputations*. Oxford University Press, 2006.

[18] R. Freeman, "Stockholders and stakeholders: A new perspective on corporate governance", *California Management Review*, vol. 25, no. 3, pp. 88–106, 1983.

[19] D. Wood and R. Jones, "Stakeholder mismatching: A theoretical problem in empirical research on corporate social performance", *International Journal of Organizational Analysis*, vol. 3, no. 3, pp. 229–267, 1995.

[20] D. Hanson and H. Stuart, "Failing the reputation management test: The case of bhp, the big australian", *Corporate Reputation Review*, vol. 4, no. 2, pp. 128–143, 2001.

[21] J. Mahon, "Corporate reputation: A research agenda using strategy and stakeholder literature", *Business and Society*, vol. 41, no. 4, pp. 415–445, 2002.

[22] C. Fombrun and M. Shanley, "What's in a name? reputation building and corporate strategy", *Academy of Management Journal*, vol. 33, no. 2, pp. 233–258, 1990.

[23] G. Davies and L. Miles, "Reputation management: Theory versus practice", *Corporate Reputation Review*, vol. 2, no. 1, pp. 16–27, 1998.

[24] E. Gray and J. Balmer, "Managing corporate image and corporate reputation", *Long Range Planning*, vol. 31, no. 5, 695–702, 1998.

[25] T. Melewar and E. Jenkins, "Defining the corporate identity construct", *Melewar and Jenkins, 2002*, vol. 5, no. 1, pp. 76–90, 2002.

[26] M. Barnett, J. Jermier, and B. Lafferty, "Corporate reputation: The definitional landscape", *Corporate Reputation Review*, vol. 9, no. 1, pp. 26–38, 2006.

[27] M. Hatch and M. Schultz, "Relations between organizational culture, identity and image", *European Journal of Marketing*, vol. 31, pp. 356–365, 1997.

[28] J. Barney, "Organizational culture: Can it be a source of sustained competitive advantage?", *Academy of Management Review*, vol. 11, no. 3, 656–665, 1986.

[29] A. Hochschild, *The Managed Heart: Commercialization of Human Feeling*. University of California Press, 1983.

[30] L. Smircich, "Concepts of culture and organizational analysis", *Administrative Science Quarterly*, vol. 28, 339–358, 1983.

[31] M. Hatch, "The dynamics of organizational culture", *Academy of Management Review*, vol. 18, no. 4, pp. 657–693, 1993.

[32] H. Stuart, "Employee identification with the corporate identity", *International Studies of Management and Organization*, vol. 32, no. 3, pp. 28–44, 2002.

[33] R. Abratt, "A new approach to the corporate image: Management process", *Journal of Marketing Management*, vol. 5, no. 1, pp. 63–76, 1989.

[34] G. Geller, "A review and critique on the relation between corporate reputation, value creation and firm performance", *Amazon, Organizations and Sustainability*, 2014.

[35] H. Rao, "The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the american automobile industry: 1895–1912", *Strategic Management Journal*, vol. 15, no. 1, 29–44, 1994.

[36] R. Freeman and J. McVea, "A stakeholder approach to strategic management", in *Blackwell handbook of strategic management*, Blackwell, 2005, pp. 189–207.

[37] E. Constantinides and S. Fountain, "Web 2.0: Conceptual foundations and marketing issues", *Journal of Direct, Data and Digital Marketing Practice*, vol. 9, no. 3, pp. 231–244, 2008.

[38] H. Schau and M. Gilly, "We are what we post? self-presentation in personal web space", *Journal of Consumer Research*, vol. 30, no. 3, pp. 385–404, 2003.

[39] C. Gurau, "Intergrated online marketing communication: Implementation and management", *Journal of Communication Management*, vol. 12, no. 2, pp. 169–184, 2008.

[40] A. Grutzmacher, "Reputation 2.0: The role of social media in corporate reputation - case nokia", Master's thesis, Aalto University: School of Economics, 2011.

[41] P. Aula, "Social media, reputation risk and ambient publicity management", *Strategy and Leadership*, vol. 38, no. 6, pp. 43–49, 2010.

[42] B. Jones, J. Temperley, and A. Lima, "Corporate reputation in the era of web 2.0: The case of primark", *Journal of Marketing Management*, vol. 25, no. 9, pp. 927–939, 2009.

[43] M. Rhee and M. Valdez, "Contextual factors surrounding reputation damage with potential implications for reputation repair", *Academy of Management Review*, vol. 34, no. 1, pp. 146–168, 2009.

[44] R. Horn, "Is the reputation quotient a valid and reliable measure for corporate reputation?", PhD thesis, Universität Kassel, 2014.

[45] C. Fombrun, N. Gardberg, and J. Sever, "The reputation quotient: A multi-stakeholder measure of corporate reputation", *Journal of Brand Management*, vol. 7, no. 4, pp. 241–255, 1999.

[46] C. Fombrun, L. Ponzi, and W. Newburry, "Stakeholder tracking and analysis: The reptrak® system for measuring corporate reputation", *Corporate Reputation Review*, vol. 18, no. 1, pp. 3–24, 2015.

[47] A. Trotta and G. Cavallaro, "Measuring corporate reputation: A framework for italian banks", *International Journal of Economics and Finance Studies*, vol. 4, no. 2, pp. 21–30, 2012.

[48] P. Cerchiello, "Statistical models to measure corporate reputation", *Journal of Applied Quantitative Methods*, vol. 6, no. 4, pp. 58–71, 2011.

[49] G. Berens and C. van Riel, "Corporate associations in the academic literature: Three main streams of thought in the reputation measurement literature", *Corporate Reputation Review*, vol. 7, no. 2, pp. 161–178, 2004.

[50] M. Carrió i Sala, "Creating a new multistakeholder methodology for measuring corporate reputation", Master's thesis, Pompeu Fabra University, Barcelona, 2011.

[51] E. Colleoni, A. Arvidsson, L. Hansen, and A. Marchesini, "Measuring corporate reputation using sentiment analysis", in *Proceedings of the 15th International Conference on Corporate Reputation: Navigating the Reputation Economy, New Orleans, USA*, 2011.

[52] T. Bhuiyan, A. Josang, and Y. Xu, *Trust and reputation management in web-based social network*. InTech, 2010.

[53] S. Einwiller, C. Carroll, and K. Korn, "Under what conditions do the news media influence corporate reputation? the roles of media dependency and need for orientation", *Corporate Reputation Review*, vol. 12, no. 4, pp. 299–315, 2010.

[54] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events", *Journal of the American Society for Information Science and Technology*, vol. 62, no. 2, pp. 406–418, 2011.