How Emotions of Customers Diffuse over Social Media Platforms after a Major Service Failure

Exploring Supervised Machine Learning to Recognize Basic Emotions in Customers' Social Media Data and using Outcomes for Competitive and Event Analysis in the Dutch Telecommunication Market

H.J.C. de With

Underlined

5 August 2016

Master Thesis Marketing Research

Marketing Department

Tilburg School of Economics and Management

Tilburg University

Master Thesis Supervisors:	dr. Hannes Datta	Tilburg University
	dr. George Knox	Co-reader Tilburg University
	Gerdien Ridderbos MSc	Underlined

Management Summary

In this thesis, we studied the effect of a major service failure on the emotions of customers. Emotions are an important factor with regards to customers' affective and behavioral states towards a company. However, insights in customer emotions are difficult to assess with traditional market research methodologies due to the essence of emotions and the experiential nature of services. We aim to overcome these problems by collecting historical social media data (user-generated content) related to the Dutch telecommunications market and developing an algorithm to automatically recognize emotions from text. We used a supervised machine-learning technique to develop the method. In this way, feedback from customers is gathered without using surveys or panel data. The average prediction accuracy of our method is 65.85%. With the output of our algorithm, we investigated a technical outage of telecommunications provider KPN that lasted 26 hours. During the time of the failure and after, we compared user-generated content related to KPN with that of its competition, for short term and long term effects.

We provided insights in the diffusion of six basic emotions (i.e. happiness, anger, sadness, fear, surprise and disgust) in different time periods after the failure. Effects are measured for four categories: social media platforms Twitter, Facebook, blogs and forums. We are able to see in which time periods during the service failure people were surprised, when they became sad, when they were angry and at which point they were relieved. The observed patterns show how some of our preconceptions are confirmed, while others might provide useful new insights into the effect of a major service failure on customers' emotions on social media, and how the actions of the company influence these emotions. Furthermore, our results indicate that emotions diffuse differently across social media platforms and that they have a strong relationship with activities of the company during and after the service failure. The service recovery effort by KPN is clearly visible in the results, the results indicate a negative evaluation of this effort. Next to studying events in the Dutch telecommunication market, the output of our emotion classification method can be used to for competitive analysis.

Table of Contents

Chapter 1. Introduction

With the rise of social media in different forms and shapes, customers have huge online reach and power to share their experiences (Kaplan & Haenlein, 2010). They can now express themselves on social networking sites, photo- and video sharing services, blogs, discussion forums and other types of social media. Marketing knowledge long held the belief that an unsatisfied customer tells ten people about a bad experience but in the new period of social media, customers have the tools to tell hundreds of people (Gillin & Moore, 2009). Consequently, their expressions can have huge influence in shaping the opinions of other customers and ultimately their loyalty to brands, their purchase decisions, and their own brand advocacy.

Commercial use of social media is encouraged by marketers and researchers, urging companies to take advantage of it in order to succeed in the modern marketplace (Kaplan & Haenlein, 2010; Kietzmann, Hermkens, McCarthy & Silvestre, 2011; Mangold & Faulds, 2009). The proclaimed benefits of social media resulted in a massive adoption amongst companies, service providers without any form of social media usage can hardly be found. Companies attempt to use social media and online brand communities to their advantage, which leads to various ways of social media implementation. Companies might engage with their customers by sharing brand information, reinforcing brand history and culture or by providing online support (Muniz & O'Guinn, 2001). Similar to other marketing channels, companies pursue to influence customer perceptions and behavior (Williams & Cothrell, 2000) and increase brand loyalty, possibly the most important incentive for companies to be present on social media (McAlexander, Schouten & Koenig, 2002).

Social media offer opportunities for companies, but also bring challenges. Tremendous efforts are put in maintaining a good atmosphere on social media. Since the online environment has been becoming increasingly important, academics are interested in providing so-called "best practices" of how to act towards customers online (Malthouse et al., 2013; Risius & Beck, 2015; Sparks & Bradley, 2014). Nonetheless, the contribution of customers on social media channels is independent of objectives and strategies of the companies that use them (Cova & Dalli, 2009). Social media are dynamic and therefore hard to control in a business setting. The best thing companies can do is pay close attention to it and monitor it. Subsequently, they have to react to imminent issues since there are situations in which social media activity is more difficult to handle. A critical situation for service companies is when a major technical failure occurs. Not only do customers often react in an emotional way to service failures (Smith, Bolton

& Wagner, 1999), they respond to service failures by expressing themselves on social media (Kietzmann et al., 2011). A technical failure may be inevitable but the way it is handled by service providers can also provoke sincere emotional reactions. It is of great importance for companies to address their customers and handle a service failure situation the best way possible to prevent harmful consequences, as the main reason for switching service providers is a service failure (Keaveney, 1995). As Bitner (1990) argues, giving customers explanations for the experienced failures and compensating them in an appropriate way can lessen the possibility of negative consequences for behavioral intentions.

Especially in service failure situations, emotional behavior matters. Statements by customers are often negative and might have consequences for the company if not carefully managed. Stieglitz and Dang-Xuan (2013) find that emotionally charged Twitter messages tend to be retweeted more often and more quickly compared to neutral ones. Moreover, emotional states can transfer to other customers through emotional contagion on social media, leading them to experience similar emotions as in the messages they see from other people (Kramer, Guillory & Hancock, 2014). Companies should therefore strive to keep negative user-generated content that is related to them to a minimum. Nonetheless, they can infer insights from social media through monitoring and analysis and in this way respond by adjusting marketing messages and other activities on social media (Mangold & Faulds, 2009).

There are several reasons for studying emotions on social media. First of all, the importance of emotions in consumer behavior. Oliver, Rust and Varki (1997) state that emotions, next to cognitive judgments, produce satisfaction and that they are central to understanding experiences of customers. Besides, emotions are expressions that have a strong relationship with the stimuli that provoke them (Batson, Shaw & Oleson, 1992). Bagozzi, Gopinath and Nyer (1999) also suggest that emotions typically have a specific referent. As a result, social media analytics can be used by companies to systematically monitor and analyze user-generated content to keep as much control as possible in a field that is typically hard to manage.

One of the current social media analysis methods is sentiment analysis. It is the analysis of written text, for extracting the general opinion of the customer about specific entities or topics. For companies, sentiment analysis provides an insight into the overall polarity of emotions (positive or negative) and intensity (weak or strong) in messages about the company and its services (Pang & Lee, 2008). In general, sentiment analysis is about determining the primary attitude of customers towards a certain topic. There has been an extensive amount of research on this topic in the social media environment (Agarwal et al., 2011; Kiritchenko, Zhu

& Mohammad, 2014; Pak & Paroubek, 2010). Although its insights are beneficial to companies, sentiment analysis is limited to sentiment polarity and intensity, which may miss important nuances in emotional expression. A deeper insight into the emotions behind brand-related messages is therefore crucial for a better understanding of consumer behavior in an opinion-rich resource as social media.

In this thesis, we investigate the effects of a major service failure and its impact on different types of social media. The effects on six different emotions are analyzed. We aim to provide insights into these basic emotions, that are recognized in social media, prevail over the various social media platforms. Our results might provide a better understanding of customers that express emotionally loaded messages about the brand online. Going this step further in text analysis, identifying additional emotions from text, has recently been explored for measuring customer satisfaction from company data (Ren & Quan, 2012) and email conversations with customers (Gupta, Gilbert & Fabbrizio, 2013). However, a similar method has not yet been applied to user-generated content on social media, even though this is a source with constant feedback from customers that is relatively easy to obtain. Social media and especially a micro-blog like Twitter stand out from more formal ways of communication because of their short size and ease of use. This has resulted in new forms of expression and because of the short time to use, social media are enhancing users to express their thoughts in real-time, often resulting in far more emotional statements than might normally occur (Roberts et al., 2012).

The Dutch telecommunications market is chosen to study because of its high amount of data and wide selection of providers that trigger different emotions. We include the social media platforms Twitter and Facebook and categorize blogs and forums in this research. These four categories are included in order to investigate whether emotions differ according to the platform that is used. All brand-related user-generated content (UGC), from January 1st 2014 until February 29th 2016, is retrieved from those social media platforms. The biggest source of online messages was *KPN*, the market leader in the industry. Other UGC suppliers were *Vodafone*, *T-Mobile*, *Telfort*, *Tele2*, *Simyo* and *Hollandsnieuwe*. Our research is divided into two main parts. First, we use all social media data to develop a method to automatically recognize basic emotions. This method is created in such a way that it can recognize emotions in UGC related to telecommunications, but is generalizable so it can be extended to other service categories or products in the future. Combining big data analysis and text analysis, a method is developed to assess which emotion occurs in social media messages. A large set of messages is used as a training set in order to develop a supervised machine-learning method for information extraction. In our research, information extraction consists of the detraction of emotions from social media text. Secondly, a subset of data around a major service failure of telecommunication provider KPN is analyzed to assess its emotional consequences on different social media platforms. In order to do so, the output of the emotion classification method is used.

This thesis is organized as follows. First, we present an overview of existing literature related to emotions in marketing, service failures, user-generated content and text analysis, in order to provide a theoretical framework. Based on this, our research method and expectations for results are presented. Then, an overview of the data collection and the automated emotion classification procedure is given. Subsequently, the analysis of the major service failure of KPN is presented. Finally, we discuss our results and conclusions.

Chapter 2. Literature review

In this chapter, an overview of existing literature is given that relates to subjects we consider in our study. First, previous research on emotions and the implications of emotional behavior for companies is outlined in paragraph 2.1. We provide support for the use of six basic emotions in our classification and define each of these emotions. In paragraph 2.2, we look at the implications of service failures and how companies might compensate this by service recovery efforts. In paragraph 2.3 we introduce user-generated content as a promising source of customer opinions. The last paragraph gives an overview of current text analysis methods and the different utilizations for them.

2.1 The Role of Emotions in Consumer Behavior

Emotions have been of importance in marketing literature for a long time and there has been an extensive amount of research devoted to the role of affect and emotions in consumer behavior. Previous research implies that emotions influence consumers' decision making processes (Leventhal, 1982; Westbrook, 1987). Morris, Woo, Geason and Kim (2002) even conclude that emotions are almost twice as important as "facts" in explaining purchase intention amongst various product categories. Related to this thesis, a study on switching behavior of customers in the Swedish telecommunications market shows that emotions play a significant part in switching operators (Roos & Friman, 2008). Emotions were present in the majority of the switching processes of customers. The role of emotions in customer satisfaction is also studied. As Oliver, Rust and Varki (1997) state, emotions coexist with several cognitive judgments in producing customer satisfaction and they are central to understanding the customer experience. Wirtz and Bateson (1999) also suggest that customer experience is partly based on emotional assessment, and separating cognitive judgments from emotional judgments is needed for understanding consumer behavior. As Lench, Flores and Bench (2011) conclude; emotions change the way people feel and think, which drives evaluation and decision making.

Consequently, researchers' attention has also focused on studying the drivers of emotional behavior. Izard and Buechler (1980) state that emotional reactions are often long-term, are hard to predict and are difficult to control. However, as Batson, Shaw and Oleson (1992) state, emotions have a strong relationship with the stimuli that provoke them. Bagozzi, Gopinath and Nyer (1999) also suggest that emotions typically have a specific referent in a customer relationship. These emotions are reactive and are the result of an experience that either exceeds or fails the anticipations of the customer.

In this thesis, the interest is in emotional expressions on social media. First of all to recognize emotions from this type of text, secondly to provide insights into the effects of a major service failure. However, distinguishing emotions from text offers challenges. As Ortony, Clore and Foss (1987) state, emotions are in essence not linguistic things but: “the most convenient nonphenomenological access we have to them is through language” (p. 342). As Wierzbicka (1999) argues, language is a direct way of comprehending emotions as it expresses someone’s emotions anytime and anywhere. Furthermore, language that is emotionally loaded expose the intentions of a text (Das, Martinez-Jerez & Tufano, 2005). As a result, there is a clear motive for the analysis of emotions from text in UGC.

For distinction between emotions, the standard by Ekman is used in this study. Ekman (1992) identified six “basic emotions”: happiness, sadness, anger, fear, disgust and surprise. These emotional types are often referred to as the “Big Six” (Alm, 2009; Calix, Mallepudi, Chen & Knapp, 2010), since they are based on facial patterns that are recognized in cultures worldwide. However, the number of basic emotions and their interpretation have been frequently discussed in existing literature. Certainly, the discussion around the concept of “basic emotions” is rather existential and therefore still persistent. Ortony and Turner (1990) argue that basic emotions are neither psychologically or biologically primitive and that they are not the building blocks for producing a variety of other emotional experiences. Nonetheless, they state that: “a reluctance to accept the notion of basic emotions does not mean that it is unreasonable as a research strategy to classify emotions in certain ways” (p. 329). Expressed emotions are often in between fundamental emotions and varying in intensity. Basic emotions are evolutionary natural and triggered by unconditioned stimuli, while higher cognitive emotions are triggered through more sophisticated cognitive stimuli (Clark, 2013). A recent meta-study from the field of cognitive neuroscience on this topic finds strong evidence for the existence of basic emotions and especially for the categories anger, happiness, fear, sadness and disgust (Vytal & Haman, 2010). To recognize emotions from text, the standard by Ekman is chosen to be able to properly distinguish between six main dimensions of emotions. Moreover, more than half of existing emotion classification techniques make use of Ekman’s basic emotions (Krcadinac, Pasquier, Jovanovic & Devedzic, 2013). Therefore, this research will be based on the emotion categorization by Ekman and build on the definitions in Table 1 (p. 10).

Table 1: Definitions of basic emotions

Emotion	Antecedents	Characteristics
Anger	Frustration, restraint, pain or unrelieved distress (Izard, 2013) A constellation of specific uncomfortable subjective experiences (Kassinove & Sukhodolsky, 1995) Caused by external forces to the self (Chipperfield, Perry, Weiner & Newall, 2009)	Accompanied by a high degree of tension and impulsiveness (Izard, 2013) Closely associated with aggression (Averill, 2012)
Fear	A threat or the absence of something that normally provides comfort and security (Izard, 2013)	A person feeling uncertainty, insecurity, and imminent danger. Can interact with positive and negative emotions at moderate intensity levels (Izard, 2013)
Sadness	Typical antecedents are death, concern for others and loneliness or loss (Chipperfield et al., 2009)	Group of feelings like suffering, disappointment, shame, neglect, (Shaver, Schwartz, Kirson & O'Connor, 1987) depression, misery, helplessness (Laros & Steenkamp, 2005)
Disgust	Failure of cleanliness but can also be directed toward a person or an idea (Izard, 2013)	People want to get rid of the object or leave the situation that disgusts them. Disgust can help to detect harmful situations (Izard, 2013) Defensive response, feeling of rejection (Scherer & Ekman, 1984)
Surprise	Any sudden or unexpected event (Izard, 2013)	'Strictly speaking, it is neither positive nor negative' (p. 283 Izard, 2013)
Happiness	A positive event caused by the effort of others or by own effort (Kumar, Olshavsky & King, 2001) Sudden relief from sources of negative feelings (Scherer & Ekman, 1984)	Only emotion with clear positive valence Associated with 'move toward it' (Lee & Lang, 2009)

2.2 Service Failures and Recovery Attempts

When the event of a service failure occurs, companies have to handle the situation in the best way possible to avoid negative consequences. As Keaveney (1995) states, the largest category of service switching is due to core service failures, where failures include: "all critical incidents that were due to mistakes or other technical problems with the service itself (p. 76)". Service failures influence perceived service quality, which determines behavioral intentions like word-of-mouth, service switching and service loyalty (Bitner, 1990). Furthermore, service failures could reduce customer satisfaction and cause customers to complain (Hart, Heskett, & Sasser, 1990). Prevention of service failures may be difficult or even impossible in service organizations, the aim should therefore be to lessen or eliminate harmful consequences. As Bitner (1990) argues, providing customers with explanations for failures and compensating them in an appropriate way can lessen negative consequences for behavioral intentions. The term "service recovery" refers to efforts made by the company in response to a service failure (Gronroos, 1988). These efforts are considered to be essential around service failures, as customers are usually more emotionally involved in and aware of recovery attempts than they normally are (Berry & Parasuraman, 1991). Evaluating, and consequently improving companies' service recovery efforts might result in an increase in customer satisfaction, a recovery of trust and a decrease in negative word of mouth after failures (De Witt & Brady, 2003). Concluding, when service failures occur, companies should handle the situation the best way possible to prevent harmful consequences. A possible way of gathering opinions and feedback of customers as a result of a service failure could be to collect them in user-generated content on social media.

2.3 User Generated Content

The emphasis of this thesis is to study emotions by analyzing customers' online messages. Of all online messages, a considerable amount is related to companies. An analysis of 150.000 tweets revealed that in 19% a brand name was mentioned (Jansen, Zhang, Sobel & Chowdury, 2009). User-generated content (UGC), also called electronic word-of-mouth (eWOM), is an important means through which consumers express themselves and communicate with others online (Boyd & Ellison, 2008). In relationship to companies, it is defined as: "a statement made by a potential, actual or former customer about a product or company, which is made available to a multitude of people via the Internet" (Hennig-Thurau, Gwinner, Walsh & Gremle, 2004, p. 39). Tirunillai & Tellis (2014, p. 3) state: "Relative to surveys of customers, UGC is spontaneous, widely available, low cost, easily accessed, temporally disaggregate, and live". Moreover, insights from UGC are based on hundreds of thousands of customer contributions from various online sources. UGC can therefore serve as a useful source of information or meaning for marketers about customers' experiences with quality (Tirunillai & Tellis, 2014) and therefore enables us to study the effects of service failures.

Furthermore, UGC has a certain degree of trustworthiness among social media users. For instance, a research of Nielson (2009) found that 70% of people trust customer reviews on social media. Because the sender of the information is not the company but another customer, like family, friends, acquaintances or even complete strangers (Feick & Price, 1987). Bruhn et al. (2012) found that UGC on social media even has a greater impact on consumers' brand image perceptions than information on traditional media, because of the source credibility. A negative review is found to have a greater negative impact compared to the positive impact that a positive review has (Lee, Rodgers & Kim, 2009). Moreover, emotional states can transfer to other customers through emotional contagion on social media, leading them to experience similar emotions as in the messages they see from other people (Kramer, Guillory & Hancock, 2014). It is therefore of vital importance for companies to study the online behavior of their customers and use their feedback as input for improvements, especially after the experience of a service failure.

2.4 Developments in Text Analysis of Affective and Emotional Content

In this section, existing literature on the intersection of text analysis and social media is outlined. First, the current progress of text analysis in marketing research is discussed. Secondly, the relatively new division of emotion classification is introduced and applications of this method are presented. The last part handles the specific characteristics of social media data and its challenges.

In recent years, the interest in social media text analysis has been growing for companies that pursue a good online relationship with their customers and want to compare their performance with competitors through social media (He, Zha & Li, 2013). The direct line between customer opinions and businesses that is provided by social media has valuable implications for marketing as a competitive analysis source (Jansen et al., 2009). Moreover, it is beneficial for companies to pay attention to its customers and use their feedback as an advantage. As Mohammad and Turney (2013) state, it is in the interest of companies to take notice to customers "not just when they call, but also during online transactions and when they write about the company in their blogs, tweets, consumer forums, and review websites so that they can immediately know whether the customers are happy with, dissatisfied with, losing trust in, or angry with their product or a particular feature of the product" (p. 439). In this way, they can take corrective action when necessary, and accentuate the most positively associated features in communication.

Traditional approaches for studying consumer behavior, such as marketing survey and focus groups, require a large amount of time and resources. Alternative methods are upcoming, with the growing availability and popularity of opinion-rich resources like online review sites and social media. For instance, Zhu et al. (2011) studied customers' opinions about certain aspects of services from customer reviews, without requiring them to answer questions. With these new sources, opportunities arise. The use of publicly available data to monitor and analyze customer opinions can reduce costs, efforts and time compared to traditional large-scale surveys (Bollen, Mao & Pepe, 2011). Moreover, Ren and Quan (2012) argue that the growing popularity and high use of social media result in a large number of real opinions from customers, and they are easy to acquire. However, social media texts also have specific characteristics that impose challenges in determining the correct emotions. They often have terms not seen in dictionaries such as misspellings, creatively spelled words, hash tagged words, emoticons and abbreviations (Mohammad, 2015).

There is already a vast amount of research using social media for text analysis. The existing literature is mainly focusing on retrieving sentiment from various online text sources. Sentiment analysis represents a computer-based analysis of written text for extracting the attitude of the author or speaker about specific identities or topics. It provides an analysis that aims to establish the overall orientation (positive or negative) and intensity (weak or strong) of the sentiments expressed by statements (Pang & Lee, 2008). Next to sentiment analysis, there is the field of emotion analysis. As discussed in the section before, there is no consensus amongst researchers about so called “basic emotions”. As a result, different standards are used in emotion analysis from text as well. There is currently less work on emotion analysis in text compared with sentiment analysis. Several studies have used the Ekman “basic emotions” in relation with text analysis. Holzman and Pottenger (2003) annotated 1201 chat messages for the six emotions. Alm, Roth, and Sproat (2005) studied 22 Grimm fairy tales for distinguishing the Ekman emotions. Strapparava and Mihalcea (2007) interpreted newspaper headlines with scores for each of the Ekman emotions. However, there is no existing research yet that uses this method to study customer emotions related to companies on social media. In this thesis, we aim to contribute to the field of text analysis by implementing this method to a data set with Dutch social media messages.

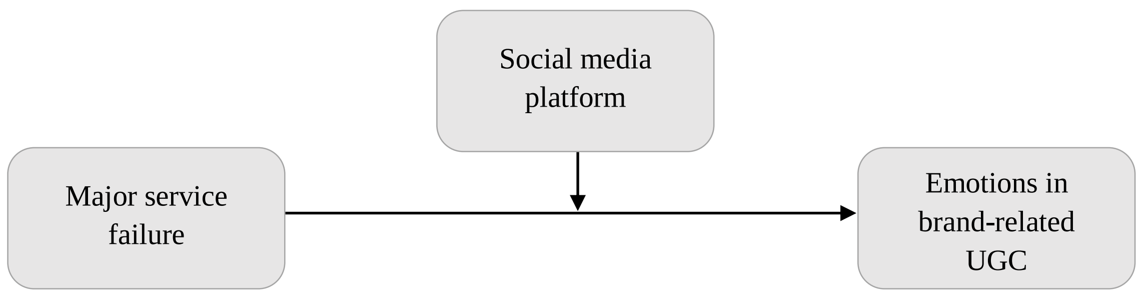
Emotion analysis in text has already been studied to some extent, but is improving because of emerging text analysis techniques nowadays. The paper of Kao et al. (2009) presents an overview of the emerging field of emotion detection, and describes the techniques that are divided into three central categories: keyword-based, learning-based, and hybrid recommendation approaches. In this thesis, a supervised machine-learning technique is performed. This approach lets the model “learn” from a set of manually labeled texts with the correct emotion, and automatically identifies what the characteristics are for texts with this emotion. This set of example instances is called the training set. The model is then able to predict the emotion of new text, using the training set as example. In order to determine the prediction accuracy, the model is evaluated on a held-out portion of the labeled data, called the test set. A more extensive explanation is given in Chapter 4, the Data section.

In conclusion, emotions change the way people think, feel and behave. Online emotional expression of customers is both influential and contagious to other social media users, so service failure situations have to be carefully managed by companies in order to diminish negative outcomes. UGC can be used to gather valuable feedback from customers.

Chapter 3. Conceptual framework

The goal of this thesis is to provide insights into how emotions prevail over social media after a major service failure, using the output of an emotion classification method. We investigate a major service failure of KPN that lasted 26 hours, from November 20th 2015 10:00 until November 21st 2015 12:00. The event caused a steep rise in brand-related UGC volume in November (177% higher than the average volume per month for KPN), but what is more important is which emotions the service failure elicits. Next to the rise in volume, it is the channeled emotion of customers that we are interested in. We want to study which emotions were present in which period during and after the service failure, and how the observed emotions differ according to which one of the social media platforms is measured. In this way we aim to provide an insight in how emotions prevail on the different social media platforms we investigated. Effects on emotions differ depending on the social media platform, hence the moderating effect. The included social media platforms are Twitter, Facebook, blogs and forums. Emotions are the basic emotions as defined by Ekman (1992) and represented by the output of the emotion classification model. The supervised machine-learning method is described in detail in Chapter 4. We use a differences-in-differences regression approach to estimate the effects of the service failure. Models are estimated for four social media sources and seven categories (i.e. the six basic emotions and a neutral group).

Graphic 1: Overview of the studied effects



Based on existing literature and intuition, we expect the following effects as a result of the service failure.

(i) Effects of the service failure on individual emotions

a. Negative effect on neutrality

Because a large part of UGC is caused by the service failure (illustrated by the rise in volume), we expect that most messages are not standard and therefore emotional rather than neutral. This effect is likely to restore back to normal levels together with the decline in volume of messages related to KPN.

b. Negative effect on happiness

Naturally, levels of happiness are expected to be lower during the failure. We expect happiness levels to recover after the failure, or even increase, as this might be a relief after a negative experience (Scherer & Ekman, 1984).

c. Positive effect on surprise in the beginning of the failure

We expect the emotion of surprise to be present in the beginning of the service failure, as an event like the one in this study is sudden and unexpected (Izard, 2013). The effect on surprise is likely to decrease short after the failure when the biggest surprise of the situation is over.

d. Positive effect on fear in the beginning of the failure

We also expect that fear is likely to be present in the beginning of the service failure. Due to the uncertainty around the failure, people might express insecurity about the threatening of their customs (Izard, 2013), like the possibility to call people or to watch television, as fear. The effect on fear is expected to decrease short after the failure, when more is known about the cause and implications of the service failure.

e. Positive effect on sadness during and after the service failure

In later stadia, disappointment about the long duration of the service failure or the handling of the problem by the company might cause sadness to be more present on social media (Shaver et al., 1987). We expect these effects to last longer after the failure, as people could still complain about the processing of the problem.

f. Positive effect on anger in the beginning of the failure

We expect anger, as a reaction to the discomfort caused by the outage, to be present from the start of the failure. People who are confronted with an interference in their daily life could be expressing anger, since frustration caused by external forces to the self is a cause for anger (Chipperfield et al., 2009). Customer anger is an intense emotional reaction to disappointing features of the service experience (McColl-Kennedy & Smith, 2006). These anger levels are expected to decline when the situation of discomfort is accepted by customers.

g. Positive effect on disgust in the beginning of the failure

The emotion disgust might also be more present during the service failure, as it can be a reason to speak out prior negative experiences or a more persistent disgust toward a company (Izard, 2013). Disgust is not expected to be present after the failure since there is no direct reason anymore to express this emotion.

(ii) Differences in emotions between social media platforms

The social media platforms we included in our study have different functions and utilizations for both customers and companies. As Smith, Fischer and Yongjian (2012) found, overall sentiment of UGC varies between different types of social media, so it is highly likely the same will hold for a broader range of emotions. Blogs and forums are expected to be less emotional in general. This is due to their characteristics of sharing and asking information, instead of channeling thoughts and daily life which is more common on Facebook and Twitter. We expect Twitter to be the most emotional and the most real-time in reactions to the failure because of the short time to use. Twitter is enhancing its users to express thoughts in real-time, which is often resulting in more emotional statements than might normally occur (Roberts et al., 2012).

(iii) An influence of the service recovery attempt on emotions

Additionally, as is common after a major service failure, a recovery attempt is undertaken by the company to try to counter the negative experience for customers. In our studied failure, KPN’s recovery strategy was to supply all customers with three free movies to watch. This was communicated on social media five days after the failure. We expect to see a reaction in emotions, as customers are involved and more conscious of recovery attempts than when the service is normal (Berry & Parasuraman, 1991). Bitner, Booms and Tetreault (1990) argue that customers could even be more dissatisfied by an organization's attempt to recover than by the service failure itself, so we expect to see the service recovery back in the data. This could be positive or negative, depending on the evaluation of the recovery effort by customers.

Chapter 4. Data

In this chapter, the analyses and results of the research are outlined. First of all, the raw data collection and the data transformation towards the cleaned data set is described in paragraph 4.1. We present our data cleaning steps, variable operationalization steps and the development of new variables from data. In paragraph 4.2, the process is described in which we used the cleaned data to develop an emotion classification method with machine-learning techniques. We discuss the performance and the validation of the method in paragraph 4.3. Paragraph 4.4 gives an overview of the final data set we obtained, including the newly created emotion variable. The distribution of emotional messages over time, brands and online sources are displayed. We also provide an example of how the output of the method can be used for competitive analysis between the brands we investigated. Finally, we give an overview of the subset of data that we use to study the effects of the service failure in paragraph 4.5. At the end, we provide a conclusion of this chapter.

4.1 Data collection and transformation

4.1.1 Study context and raw data collection

The main interest of this thesis is in providing insights from UGC that is related to brands. In order to do so, we have to collect, analyze and compare messages from customers on various social media platforms. A message is considered to be brand-related if either one of the seven selected telecommunication providers in the Netherlands is mentioned in the text. Our first objective was therefore to create a data set with online messages that are related to telecommunication providers. The companies that are included in the data set are *KPN*, *Vodafone*, *T-Mobile*, *Telfort*, *Tele2*, *HollandsNieuwe* and *Simyo*. We obtained messages from four different types of social media sources: a microblog (Twitter), a social network (Facebook), blogs and forums. UGC on Twitter and Facebook origin from their respective websites. For UGC from forums, we bundle all messages from sources that accommodate threaded discussions on various topics (which sometimes requires membership to comment) in a typical question and answer style. Our blog category includes individual blogs, company blogs (e.g. as part of a website) and comments on company blogs.

The raw data used in this research is collected in cooperation with Underlined. First, we mined the data by extracting historical UGC from the aforementioned social media sources. We used the program Radian6, a social media analysis tool designed to track customer conversations about brands across social media channels. By applying text filters we were able to retrieve 2,289,020 relevant messages from Dutch social media users in the period January 1st 2014 until February 29th 2016.

4.1.2 Data preparation and variable operationalization

The data mining procedure resulted in a data set with variables that showed for each message: an identity number, the date and time of posting, the author of the message, the online source and the associated text. The following data transformation is done in SAS Guide.

First, we merged data from 2014, 2015 and 2016 into one data set. The next step considered removing incomplete or irrelevant data. From the original raw data set, the type of messages in Table 2 were left out of further research. We removed observations with missing or incomplete data as well. Furthermore, the scope of this research is restricted to emotions in user-generated content, so messages from providers and their social media webcare are taken out by applying text filters. In a similar way, we removed messages from stores, advertising and other branded content. An explorative search through the data delivered the appropriate filters. Irrelevant messages for this research are messages from social media webcare teams (e.g. "kpnwebcare"), job offers (e.g. "vacature") and promotional or professional messages (e.g. "GSM4You"). The complete text filter can be seen in Appendix A.

We created a new variable that shows which of the brands is mentioned in the message. The variable can either display one of the brands, the term "No Brand" or the term "Multiple Brands". "No Brand" is assigned when none of the companies of interest are mentioned in the text, "Multiple Brands" appears when more than one brand is mentioned in the text. We filter messages with "Multiple Brands" and "No Brand" out, as they have no further purpose in this research.

Table 2: Data cleaning steps, in order of implementation

Message characteristic	Reason for exclusion	N
	Raw data	*2,289,020*
Empty/hidden data	Not all messages are readable due to privacy settings	127,630
No brand	None of the brands are mentioned in the text	1,010,271[1]
Multiple brands	Multiple brands are mentioned in the text	51,964
Webcare messages	Content generated by webcare teams are not relevant	66,106
Commercial messages	Promotion and commercial messages are not relevant	160,949
Job offers	Vacancies and job offers are not relevant	50,870
	Cleaned and filtered data set	*1,035,358*

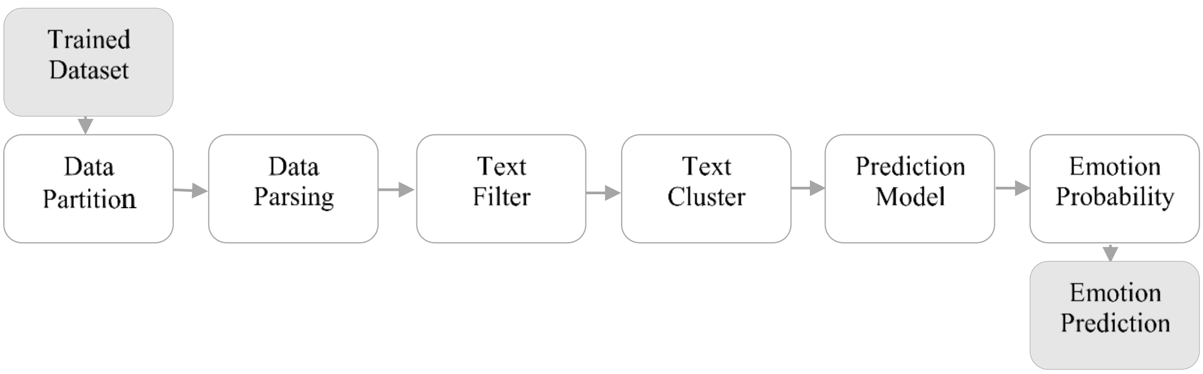
[1] A large quantity of messages with "No Brand" are messages from webcare teams

After we filtered all irrelevant and incomplete data out, the set contained 1,035,358 messages. The final data set consists of messages regarding telecommunication providers written by social media users. From this data, we generate a random sample to collect training data as input for the emotion classification algorithm.

4.2 Emotion classification using supervised machine learning

We trained an algorithm with a humanly annotated, or labeled, random sample to recognize and classify emotions from social media messages. In this paragraph we describe how we developed this method. We used a supervised machine learning technique, conducted in SAS Enterprise Miner. The algorithm (also called the learner) is best described as a technique that: "automatically builds a classifier for a category *x* by observing the characteristics of a set of documents manually classified under *x* by a domain expert. From these characteristics, the inductive process gleans the characteristics that a new unseen document should have in order to be classified under *x*" (Sebastiani, 2002, p. 8). We explain the procedure step by step in this paragraph, after which the performance and the validation of the algorithm are discussed in paragraph 4.3. A visual overview is given in the graphic below.

Graphic 2: Overview of the emotion classification process



4.2.1 Creation of the training data set

A random sample is taken out of the final data set in order to train the algorithm, as the first part of the supervised machine-learning technique that is used. We train the algorithm with labeled cases, so input is needed of which the desired output, the observed emotion, is known. The label can be one of the six basic emotions or "neutral". For instance, a message has the label "anger" when it is perceived as predominantly angry. The algorithm then uses the text in the message as input to learn what words, punctuation and other characteristics determine what makes a message "anger". It modifies the model by comparing its output with the labeled outputs to find flaws and adjusts it accordingly.

Judgement of emotions in messages was done after gathering a considerable amount of definitions of what basic emotions are, resulting in the definition overview table in Chapter 2 (p.10). During the training data gathering process we found that the expected performance of the model[2] improved by adding more training data, so we put more effort in collecting additional training data to train the algorithm. However, the problem of overfitting arises when too much training data is used in machine learning techniques. As Sebastiani (2002) states, algorithms that "overfit" are good at reclassifying the data they have been trained on, but much worse at classifying previously unseen data. We avoided this problem by testing the algorithm on different test data sets. A different training set composition is generated by running the Data Partition step again, as explained in paragraph 4.2.2. Eventually, we selected 175 messages for each of the seven categories (i.e., the six emotions and a neutral group) from the random sample. This resulted in a total data set of 1,225 labeled messages.

4.2.2 Data Partition

The labeled messages are then divided into a "training part" containing 60% of the 1,225 messages, a "validation part" containing 20% and a "test part" containing the remaining 20% of the data set. This data partitioning step delivers mutually exclusive data sets, so the three data sets share no identical observations. We use the training part to find the best possible model structure. The validation part is used to monitor and adjust the model during training and is also used for model evaluation. The test part is a supplementary hold-out set that we used for assessment of the model during the search for the best adjustments.

4.2.3 Data Parsing

As a third step, we parsed the sentences in the message. In text analysis, data parsing is the process of analyzing text to determine its grammatical structure with respect to a given grammar (Russell et al., 2003). This enables the algorithm to properly analyze the information in the text. The text is separated into nouns, verbs, multiword terms, punctuation marks and other types of information. These are called "parts of speech" (POS) and it is possible to select certain POS to be excluded from analysis to shift focus to the more informative parts in text. By default, the program filters out some POS. For instance, pronouns (e.g. "hun") and numbers are left out of the analysis. Caution is needed here, because the strength of machine learning is recognizing previously unnoticed characteristics to be useful for analysis. We also filtered out "entities" like names of people and companies because brands should not be determinants for

[2] Calculated by SAS Enterprise Miner. Models and adjustments were evaluated using the hold-out test data set.

the emotionality of a message. SAS Enterprise Miner uses the "bag-of-words" approach to explain texts. This means that each text is represented by a vector that contains the frequency in which a term or symbol occurs.

4.2.4 Text Filter

In the fourth step, we further filter the parsed text. Text filters are used to reduce the total number of parsed terms that are analyzed. Therefore, it enables us to exclude irrelevant information and correct erroneous information so only valuable and relevant information is used as input. The use of text filters is especially beneficial in the context of this thesis, since the text in the analysis is coming from social media. An important characteristic of social media text is the high amount of misspellings (Kiritchenko, Zhu & Mohammad, 2014). Because of the short time in which it is written, a lot of social media users do not take their spelling and grammar all too seriously when writing a message. This is why we need to take misspellings and other characteristics into consideration. For instance, “telfoon” is treated as “telefoon” and “abo” as “abbo”. Another issue in our analysis is the use of abbreviations. For instance, the word “abbo”. This is an abbreviation for “abonnement”, the Dutch word for subscription. We selected these words to be treated as synonyms in the method. After we compared the results, a minimum amount of four appearances is set for terms or symbols in order to be included in the analysis. Since we consider topics (e.g. “wifi”, “internet”, “telefoon”) to be of undesirable influence on the perceived emotion in a message, they are filtered out as well. In a similar way we consider brands to be undesirable for determining emotions, so an extra filter is added to account for the included brands (e.g. “kpn”, “telfort”).

4.2.5 Text Cluster

The fifth step involves clustering of the text. The algorithm segments data by grouping terms and symbols that are statistically similar for determining the output. Similar observations tend to be in the same cluster, and observations with different associations tend to be in other clusters. The main purpose of clustering is to optimize the structure of the data for input in a statistical model, as further explained in paragraph 4.2.6. We used hierarchical clustering for our algorithm since it yielded the best results in combination with subsequent statistical models. Hierarchical clusters generate decision trees to divide terms and symbols.

4.2.6 Optimal Model Selection

The sixth step consists of deciding on the learning algorithm that performs best in predicting emotions. In contrast to various other statistical modeling methods, which can value interpretation over prediction, the focus of machine learning is predictive accuracy (Breiman, 2001). The highest predictive accuracy is achieved by training complex models, which involve using different optimization routines. The tools of SAS Enterprise Miner for prediction and classification include high-performance Bayesian networks, neural networks, forests, and support vector machines. Some of those models are called “black box techniques” because their internal mechanisms are not easy to interpret by humans (Sebastiani, 2002). Our approach was to train different learning algorithms and assess them by comparing their performance on the hold-out test data. Ultimately, our best performing model was a “High Performance Forest” that used the output of the hierarchical text cluster (see paragraph 4.2.6). The HP Forest combines different kinds of decision trees to automatically find the best fitting model (Breiman, 2001). The output of the method was a newly created data set with columns having probabilities (between 0 and 1) for each of the seven dimensions (i.e., the six emotions and a neutral group).

4.2.7 Emotion prediction

In the last step we interpreted the predictions of the algorithm and assigned decision rules for individual dimensions. In this way, we can assign a dominant emotion label to the messages in the data set. After we analyzed the output generated by the model, different cut-off rules appeared to be giving the best results for the prediction. For instance, the probability for “anger” has to be at least 0.21 and the probability for “neutral” has to be lower than 0.15 for a message in order to be judged as a predominantly angry message.

4.3 Validation of the model

In order to assess the performance of the method, 50 messages per dimension are checked by two independent raters. The results can be seen in Table 3.

Table 3: Validation of sample by two independent raters

		Misclassifications 1		Misclassifications 2	
Emotion	n in sample	n missed	% missed	n missed	% missed
Neutral	50	8	16.0%	8	16.0%
Happiness	50	8	16.0%	10	20.0%
Anger	50	18	36.0%	16	32.0%
Fear	50	18	36.0%	22	44.0%
Surprise	50	19	38.0%	23	46.0%
Sadness	50	21	42.0%	23	46.0%
Disgust	50	22	44.0%	23	46.0%
		Average	32.6%	Average	35.7%

The average misclassification rate is 34.15%, which implies an average precision rate of 65.85%. As can be derived from the table, there are differences in performance between the dimensions. The categories “neutral” and “happy” perform the best, while “sad” and “disgusted” are rather more difficult to predict for the model.

The raters could either agree or disagree in their judgment. The diagnoses in agreement are located on the main diagonal of Table 4. We used Cohen’s kappa to measure the reliability of the performance by measuring the agreement between the two independent raters. This measure takes into account the fact that raters agree with each other a certain percentage of the time, simply based on chance (Cohen, 1960). The integer approaches 1 when coding is perfectly reliable and moves to 0 when there is no agreement besides what would be expected by chance only. A Cohen’s Kappa of 0.817 is calculated which is, according to the benchmarks of Landis and Koch (1977), an “almost perfect” strength of agreement.

Table 4: Overview of Inter Annotator Agreement

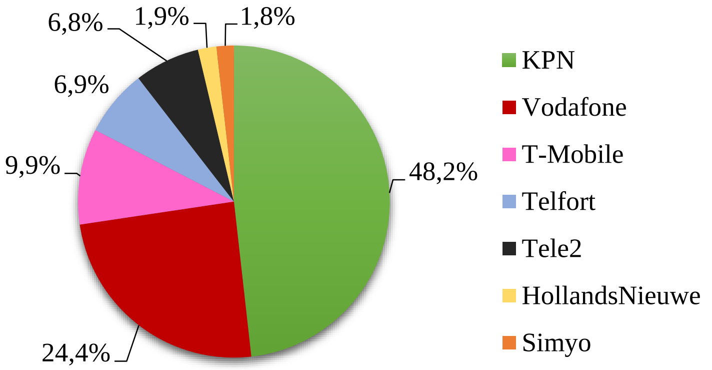
	Rater 2							
Rater 1	Neutral	Happiness	Anger	Fear	Surprise	Sadness	Disgust	Total
Neutral	113	3		4	4	4	5	133
Happiness	2	43		1				46
Anger	3		33	1			3	40
Fear	1			31				32
Surprise	4		2	1	25		1	33
Sadness	2		1			24	5	32
Disgust	1		2		1		30	34
Total	126	46	38	38	30	28	44	350
Agreement	113	43	33	31	25	24	30	299
By chance	47.88	6.05	4.34	3.47	2.83	2.56	4.27	71.41
Cohen's Kappa	**0.817**							

Following the output of the emotion classification method, the final data set is completed with the new variable “Emotion”. This variable shows the prediction for the emotion of a message as calculated by the classification method. According to this variable, the text in the social media message mainly contains mainly anger, sadness, happiness, disgust, fear, surprise or neutrality.

4.4 Overview of the final data set

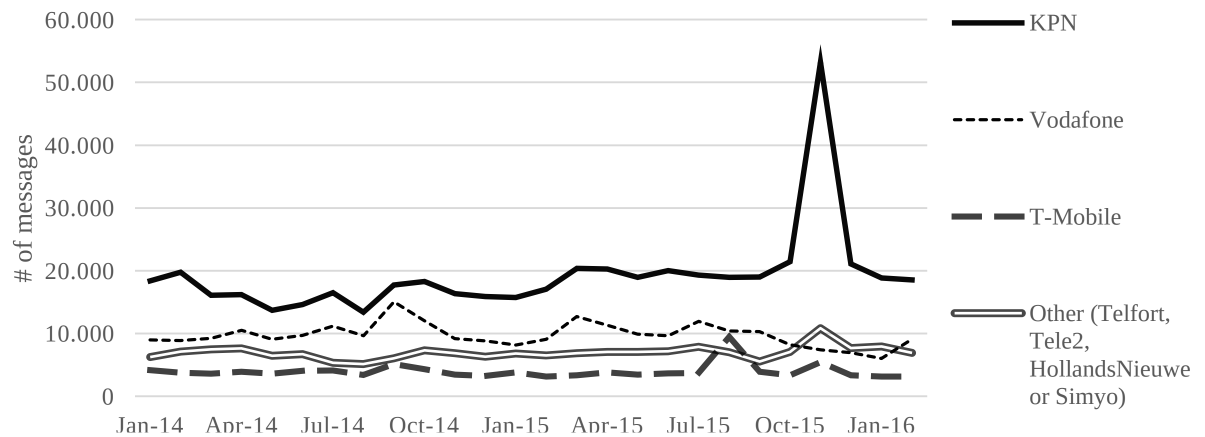
In this paragraph we present an overview of the final data set, and how the output might be used to compare social media performance of brands. Of all retrieved data, 79.8% (N=826,377) is from Twitter, while 11.7% (N=120,818) is from forums. Facebook (4.7%, N=49,074), blogs (2.4%, N=25,016) and mainstream news messages (1.4%, N=14,072) account for the remaining volume. Graphic 3 shows the distribution of the online messages over the different brands. Almost half (48.2%) of the messages are about KPN, while almost a quarter (24.4%) is mentioning Vodafone. The other five brands have lower volumes.

Graphic 3: Message distribution over brands



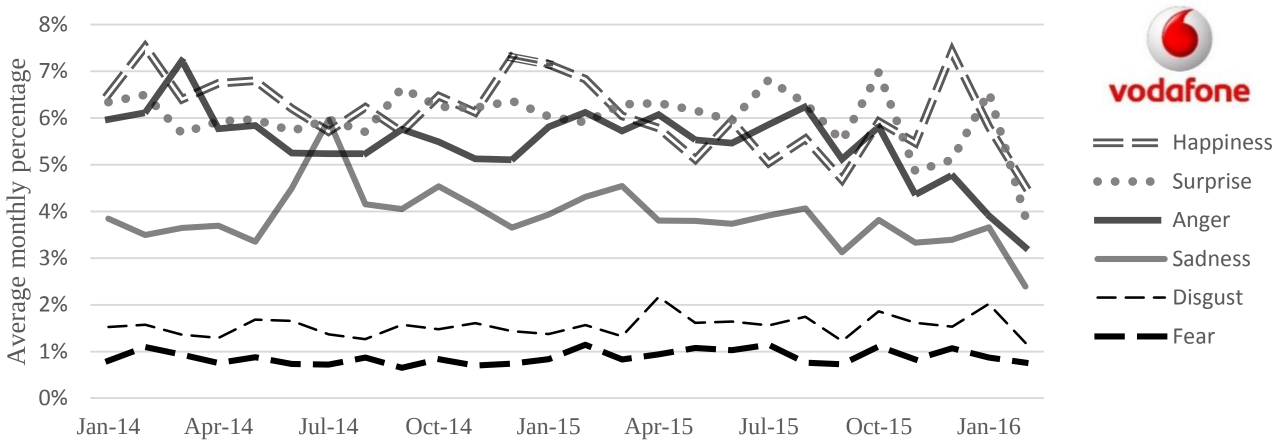
Furthermore, the monthly amount of messages per brand is graphed below. As can be seen, the service failure of KPN in November 2015 generated a huge peak in UGC about the brand. An overview of the monthly amount of messages for all companies can be found in Appendix B.

Graph 1: Volume of messages per provider over time



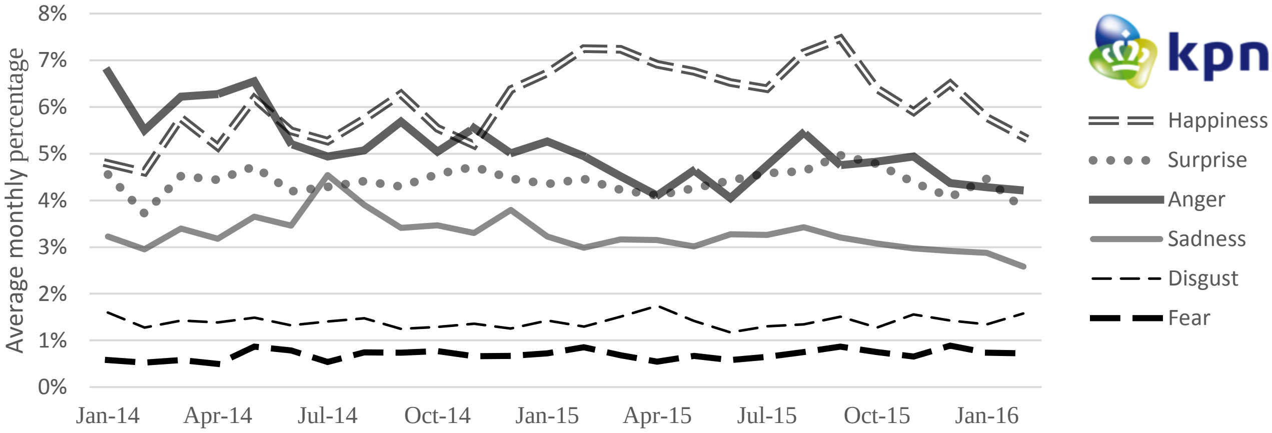
Now we can take a deeper look into the observed emotions in UGC for individual brands. When we plot our observations over time for the different brands we see some interesting insights. Graphs are plotted to visualize the average monthly amount of emotions in the messages of the two largest brands in the Netherlands, KPN and Vodafone. From the Vodafone graph, two months particularly stand out. A peak in sadness can be seen in July 2014. Looking at the data implies that this seems to be related to the launch of a new version of a mobile application, which was disappointing for many customers. Furthermore, the peaks in happiness in December 2014 and 2015 seem to be related to the winter holiday activities on social media by Vodafone, which generated a lot of positive reactions.

Graph 2: Average amount of emotions in Vodafone UGC



For KPN, an interesting trend can be seen in the graph. In the year 2014, happiness, anger and surprise are distributed quite similarly. Then, from the beginning of the year 2015, happiness is substantially more present in UGC compared to the other emotions. An overall downward trend of anger in KPN–related UGC can be seen as well, from an average of 6.8% in January 2014 to 4.2% in February 2016. These results therefore suggest that, on average, UGC related to KPN is becoming more happy and less angry.

Graph 3: Average amount of emotions in KPN UGC



4.5 Overview of the service failure data set

The goal of this thesis is to provide insights into how the six emotions diffuse over social media after a major service failure. We analyze the service failure of KPN in November 2015, when technical problems caused a major outage for KPN customers. The outage took 26 hours to fix all problems, from 10:00 on November 20th 2015 until 12:00 on November 21th 2015. We created a subset of the final data set, consisting of data from one week until one week after the failure. The data ranges from November 13th 2015 0:00 until November 28th 2015 0:00. In total, there are 54,403 messages in this subset of two weeks. We summarize the distribution of messages in Table 5. Emotions in KPN-related UGC are compared with competition as control group, so other providers are grouped together. On average, UGC related to KPN is 19,432 per month so a volume of 41,057 can be considered very high.

Table 5: Distribution of data in the service failure subset

	Twitter	Facebook	Blogs	Forums	Total
KPN	27,826	10,132	703	2,396	41,057
Other providers	11,787	703	215	641	13,346
Total	39,613	10,835	918	3,037	54,403

In conclusion of this chapter, we developed an algorithm to recognize emotions in brand-related UGC using historical social media data from the Dutch telecommunications market. The average prediction accuracy of the method is 65.85%. We can use its output for competitive analysis and for studying emotional behavior on the included social media platforms, as will be carried out in the next chapter.

Chapter 5. Model

We used a differences-in-differences regression approach to provide the insights into the spreading of emotions over social media after the service failure of KPN. With this method, the observed emotions for KPN are evaluated against emotions in other companies' UGC in the different stadia before, during and after the service failure. This takes out biases in comparing the period of failure of KPN with the control group, which consists of UGC related to the competition combined. We have to take permanent differences between those groups into account, as well as biases in time periods that result from general time trends. The regression models use the output from the automated emotion classification method (as described in Chapter 4) as input for the dependent variable. This data, the expectancy of presence of an emotion, is aggregated and standardized by brand and platform to account for brand-specific and platform-specific dissimilarities. The regression models are estimated for eight periods after the service failure to measure the effects for the short run and longer run. Models of the following type are estimated:

$$\begin{aligned} \mathrm{E}_j^K = \; & \mathrm{c}_j^K + \delta^K \mathrm{I}(\mathrm{KPN}) + \alpha_{j1}^K \mathrm{I}(\mathrm{T}_1) + \alpha_{j2}^K \mathrm{I}(\mathrm{T}_2) + \alpha_{j3}^K \mathrm{I}(\mathrm{T}_3) + \alpha_{j4}^K \mathrm{I}(\mathrm{T}_4) + \alpha_{j5}^K \mathrm{I}(\mathrm{T}_5) + \alpha_{j6}^K \mathrm{I}(\mathrm{T}_6) \\ & + \alpha_{j7}^K \mathrm{I}(\mathrm{T}_7) + \alpha_{j8}^K \mathrm{I}(\mathrm{T}_8) + \beta_{j1}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_1) + \beta_{j2}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_2) + \beta_{j3}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_3) \\ & + \beta_{j4}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_4) + \beta_{j5}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_5) + \beta_{j6}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_6) + \beta_{j7}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_7) \\ & + \beta_{j8}^K \mathrm{I}(\mathrm{KPN} \cap \mathrm{T}_8) + \varepsilon_j^K \end{aligned}$$

where E_j^K is the dependent variable. c_j^K is the constant that differs per emotion *K* and social media platform *j*. The indicator variable $\mathrm{I}(\mathrm{KPN})$ is 1 if messages are related to KPN, or 0 if messages are related to one of the other brands (the control group). The indicator variables $\mathrm{I}(\mathrm{T})$ are 1 if the expressed emotions occurred in the given time period and 0 if not. Indicator variables $\mathrm{I}(\mathrm{KPN} \cap \mathrm{T})$ are 1 if messages are both related to KPN and occur in the given time period, 0 if not. ε_j^K represents the error term of the models. Effects are specified for the different time periods *T* after the service failure occurred. They consist of four 12-hour periods for the short term effects and four 24 hour-periods for the longer term effects. Due to the included time periods, we can study the effects for different stadia during the failure, as well as after the failure.

Chapter 6. Results

6.1 Model fit

The goal of the research is to find out whether the emotions elicited by a major service failure change over time and if this varies across social media platform. In order to test this, 28 models are estimated (i.e., four social media sources and seven categories, for the six basic emotions and the neutral group). The coefficients for the emotions and the R square, Adjusted R Square and significance probability of the models can be found in Table 6 (p. 30). The effects on each of the emotions are observed for eight time periods during and after the service failure.

The statistical significance of models and the individual coefficients vary highly. Where some models are statistically favorable, others show no notable significance at all. The models that are based on the smallest amount of data, forums and blogs, score lowest in significant results. The insights we present here are therefore mainly based on Twitter and Facebook. On the one hand, they confirm some of the main expectations we had before the analysis. Also, surprising and possibly valuable trends to understand the effects of a service failure on customers' emotions are derived. On the other hand, the results seem to be supporting the potential of the developed emotion classification method for further use.

In the next paragraph, the coefficients are displayed in a table. The spreading of all emotions over the social media platforms are graphed and explained in paragraph 6.3.

6.2 Estimated coefficients

Table 6 (p. 30) gives an overview of the output from the regressions models. Coefficients (i.e. for the eight time periods in hours after the service failure occurred), R Square, Adjusted R Square and significance probability are shown in separate columns for each of the models, which vary on social media platform and emotion. Significant coefficients are marked.

Table 6: Emotion coefficients for time periods after the service failure

	Neutral				Happiness				Surprise				Fear			
Hours since failure	Twitter	Face-book	Blogs	Forums	Twitter	Face-book	Blogs	Forums	Twitter	Face-book	Blogs	Forums	Twitter	Face-book	Blogs	Forums
0	1.049	-1.765	1.242	-1.171	-1.176	.609	-1.916	.538	-.610	**2.298***	1.297	.446	-1.627	.524	1.732	-.107
0-12	.088	-.054	.655	.085	.065	.077	1.246	1.263	-.760	.682	-1.466	-.817	-.178	-.892	-.461	-.193
12-24	-.097	**2.230*****	1.538	1.682	1.122	1.825	.107	-.276	-.405	-.658	-.697	-1.789	-.412	**2.251****	-1.585	-.808
24-36	.609	.161	-1.182	.944	.247	**-1.923***	1.411	-1.178	-.914	-1.567	-.305	-.120	.281	**1.597****	-.602	.548
36-60	.259	1.053	-.650	-.311	.157	-.728	1.269	.648	**-1.577*****	-.259	.674	.754	-.226	-.101	-.774	.275
60-84	-.305	.395	.486	.224	.259	.932	-.171	1.108	-.332	1.172	-.668	-.885	.452	-.923	.481	-.182
84-108	.987	**1.576*****	-.478	-.930	**-1.970****	-.480	.300	.910	-.213	-.048	-.128	-.578	-.414	-1.240	-1.221	-1.318
108-132	-.250	-.377	**1.730****	-.585	.758	.283	**1.895****	1.183	-.479	**-1.734****	-.047	.993	.368	-.167	-1.038	-.700
R^2	.575	.370	.347	.255	.313	.332	.286	.162	.235	.410	.139	.269	.275	.431	.430	.240
Adj. R^2	.403	.096	.070	-.046	.035	.041	-.017	-.177	-.075	.153	-.227	-.027	-.018	.184	.188	-.067
Sig.	*.001*	*.215*	*.272*	*.634*	*.363*	*.354*	*.535*	*.949*	*.727*	*.113*	*.983*	*.569*	*.539*	*.076*	*.068*	*.703*

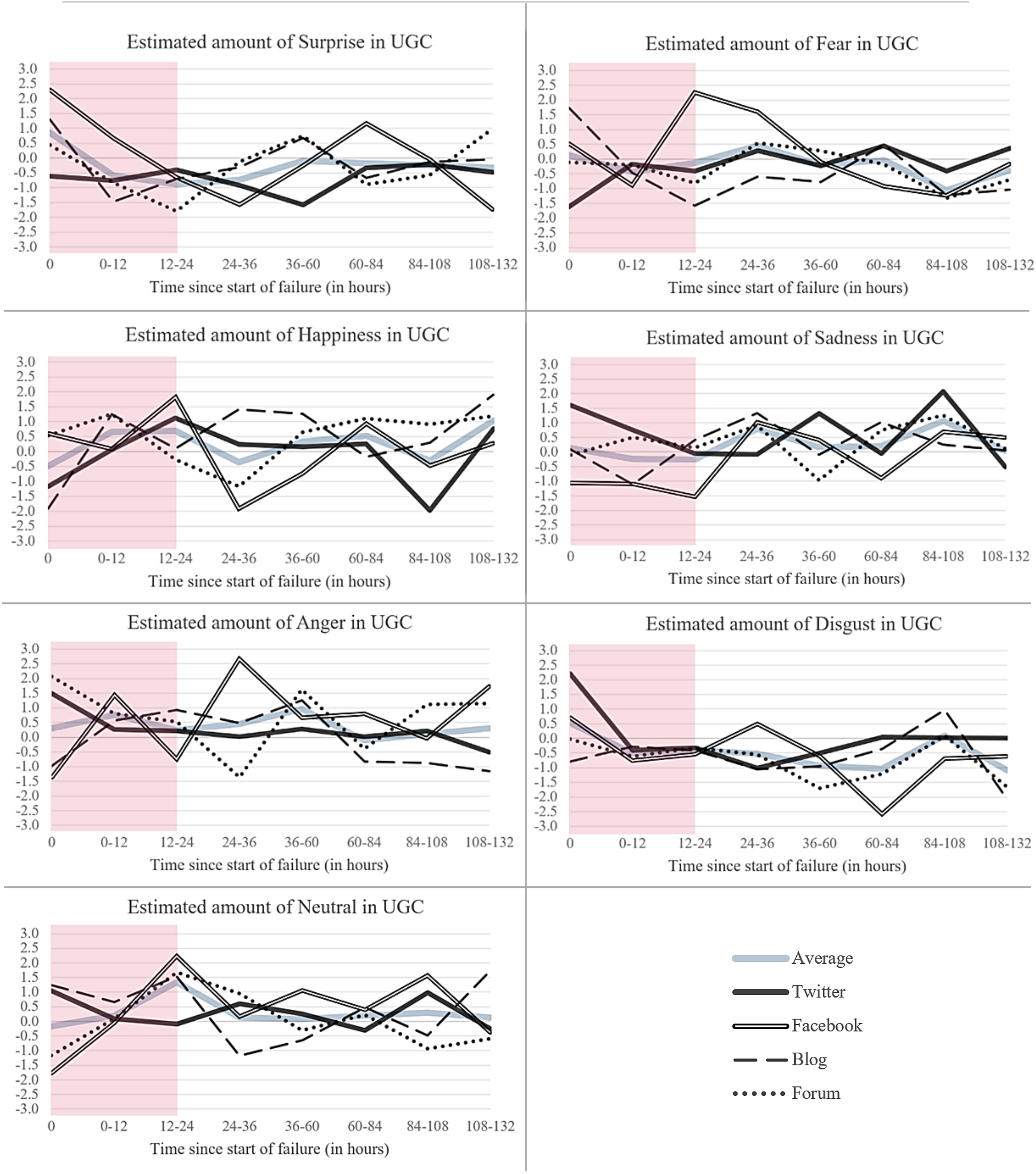
	Sadness				Anger				Disgust			
Hours since failure	Twitter	Face-book	Blogs	Forums	Twitter	Face-book	Blogs	Forums	Twitter	Face-book	Blogs	Forums
0	1.607	-1.053	.097	-.107	1.496	**-1.371***	-.982	2.077	**2.215*****	.706	-.797	-.018
0-12	.751	-1.090	-1.115	.501	.265	1.438	.570	.800	-.406	-.752	-.286	-.615
12-24	-.052	**-1.527***	.415	.147	.210	-.739	.931	.534	-.316	-.542	-.417	-.326
24-36	-.081	1.025	1.339	.918	.012	**2.672***	.493	-1.360	-1.013	.490	-1.056	-.556
36-60	1.326	.415	-.095	-.965	.288	.667	1.261	**1.631*****	-.496	-.615	-.947	**-1.695*****
60-84	-.052	-.889	1.017	.758	.013	.803	-.828	-.385	.033	**-2.585***	-.363	-1.215
84-108	**2.081***	.692	.249	1.268	.223	-.038	-.868	1.111	.019	-.688	.975	.088
108-132	-.505	.500	.049	.110	-.507	**1.739****	-1.155	1.146	.008	-.602	**-2.077***	**-1.659*****
R^2	.392	.513	.327	.323	.108	.490	.284	.220	.141	.275	.287	.175
Adj. R^2	.146	.300	.041	.049	-.253	.268	-.020	-.096	-.207	-.041	-.016	-.159
Sig.	*.110*	*.012*	*.351*	*.321*	*.995*	*.021*	*.543*	*.789*	*.977*	*.609*	*.531*	*.925*

*Notes: * where p< 0.05, ** where p< 0.10, *** where p< 0.15*

6.3 Estimated patterns

When the coefficients are plotted, patterns in emotions can be derived from the analysis. The following seven graphs show the spreading of the studied basic emotions as a result of the service failure of KPN. The lines show how emotions in KPN-related UGC compare to UGC of its competitors. The highlighted area shows the period of outage. Lines are plotted for each of the four social media platforms, the "average" line shows the trend for all of them combined.

Graphs 4-10: Visualizations of estimated emotions



Although the overall significance is small, we can derive valuable insights from looking at the plots and into the data. We repeat our previous expectations and summarize the main findings.

(i) Effects of the service failure on individual emotions

a. Negative effect on neutrality

We expected a negative effect on neutrality. However, as the data shows, neutrality on Facebook, blogs and forums increases from the beginning of the failure until the end of the failure. In the increase in UGC volume are a lot of neutral messages from people informing about the outage. People want to know what is happening and later on, how long it will take to fix the problem in their region. Importantly, effects are not observed for Twitter, where the company shares frequent updates about the situation.

b. Negative effect on happiness

Happiness was expected to be lower during the failure and was expected to demonstrate feelings of relief after the failure was recovered. What we see in the graph of happiness might be surprising at first. Confirming our expectations, happiness is low at the start of the service failure. In the first twelve hours after though, overall happiness is increasing. Looking at the data gives a possible explanation for this pattern; loyal customers are mentioning the company in happy social media messages about a fun night without television or telephone connection.

c. Positive effect on surprise in the beginning of the failure

We expected the emotion of surprise to be present in the beginning of the service failure, as an event like the one in this study is sudden and unexpected. Confirming to these expectations, surprise is present in the beginning of the failure. People were not expecting a failure of this size from their telecommunications provider and are surprised by the impact on their daily life. If we look at differences between platforms, we see that there is hardly any effect on Twitter. Surprise among people might be low here due to the high amount of updates that KPN shared on its Twitter pages from the beginning of the service failure.

d. Positive effect on fear in the beginning of the failure

We also expected that fear was most likely to be present in the beginning of the service failure. This early effect is only observed for Facebook and blogs. On Twitter, less fear is present in the beginning of the failure. We might infer that the uncertainty around the failure, where people might express their insecurities with fear, are lower on Twitter because of the information density here. Noteworthy, we see that on Facebook, there is another peak in expressed fear in the last period of the failure. This might be related to people who were worried about the duration and the lack of information from the company on this platform specifically.

e. Positive effect on sadness during and after the service failure

Disappointment about the long duration of the service failure, or the handling of the problem by the company, was expected to cause sadness to be more present on social media. We see that expressions of sadness develop in a different way between Twitter and Facebook. There is a strong impact on Twitter in the beginning and then declines to normal levels, whereas an increase in sadness is observed on Facebook in time periods after the failure.

f. Positive effect on anger in the beginning of the failure

We expected anger, as a reaction to the discomfort caused by the outage, to be present from the start of the failure. Anger also seems to spread out differently on the social media platforms. As with most of the other emotions, angry reactions at the start of the failure are the strongest on Twitter. Similar to sadness, anger is observed on Facebook in time periods after the failure as well. Furthermore, if we look at the data we see that a large quantity of the angry expressions on Facebook seem to be from people who are angered with other KPN customers who complain about the situation. They turn their anger towards those who apparently cannot live without a night of television.

g. Positive effect on disgust in the beginning of the failure

As we expected, people who want to express disgust about the company do so during the beginning of the failure but levels immediately decline in the periods after. A rise in disgust is only observed for the first time period, indicating that disgust is indeed related to a long-lasting feeling of hate towards a company and is not suddenly initiated by one event.

(ii) Differences in emotions between social media platforms

As can be inferred from the previously explained effects there are indeed differences in how emotions spread out, depending on the social media platform. Overall, UGC from Twitter is more emotional in the beginning of the failure and UGC from Facebook displays more emotions in later time periods. Emotions on blogs and forums spread out quite similar and tend to be less emotional.

(iii) An influence of the service recovery attempt on emotions

The company's service recovery strategy was to supply all customers with three free movies to watch. This was communicated on social media five days after the failure. A sudden decline in happiness and a peak in both sadness and disgust on this day, 84 to 108 hours after the service failure, is observed. This is likely to be related to the offered compensation by the company. Many people are expressing their disappointment about the offered compensation by the company. Moreover, it turns out that not all of the problems are fixed for all customers and they are triggered by the communication of the compensation on social media to react either with sadness or disgust. The emotional effects are the strongest on Twitter.

Chapter 7. Discussion

7.1 Summary

With the use of supervised machine learning we were able to identify the expression of emotions in consumers' social media messages with an average precision of 65.85%. We used the developed algorithm to study the effects of a major service of KPN in November 2015. We were able to see in which time periods during the service failure people were surprised, when they became sad, when they were angry and at which point they were relieved. The observed patterns show how some of our preconceptions are confirmed, while others might provide useful new insights into the effect of a major service failure on customers' emotions on social media, and how the actions of the company influence these emotions.

7.2 Discussion

As Forgas (1990) argues, emotions "appear to influence what we notice, what we learn, what we remember, and ultimately the kinds of judgments and decisions we make" (p. 273). The role of emotion is also attracting acceptance and interest from the field of marketing, in the search for better understanding of customer behavior and of the experience of the service (Mattila & Enz, 2002; Oliver, Rust & Varki, 1997). Still, customer emotions are considered messy and therefore dangerous for business strategies (Bagozzi, 1992). Because of its characteristics and its measurement difficulty, emotions are hard to be quantified and put into a spreadsheet. However, substantial investments are made by companies to try to increase satisfaction of their customers and fix problems when they arise, since happy customers are more likely to stay with the company, tell others about their positive experience and recommend the organization to their friends and acquaintances (Reichheld, 2003). Our study provides a quantifiable measure of how the service of a service company is evaluated by their customers in terms of emotions. Moreover, we can derive more information from the data about the causes of these emotions. Liljander and Strandvik (1997) call for more understanding of customer emotions through profound analysis, identification of events that trigger negative emotions and investigation of how effective service recovery efforts can minimize the impact of negative emotions on customer satisfaction and loyalty. If customers express negative emotions and the company can successfully decode them, the service company can use this feedback to improve its performance in search of a higher customer satisfaction.

Overall, we observed differences between social media platforms in how emotions were affected by the service failure. Twitter is overall more emotional in the beginning, where emotional Facebook messages are mostly observed in later time periods. This might be confirming to the belief that Twitter is used as a real-time instrument to channel thoughts, in this context emotional messages about a company. Emotional reactions on Facebook start later, but hold on for a longer time. Social media platforms have different functionalities, they can focus more on personal identity, conversations, sharing, relationships, personal reputation, or on groups (Kietzmann et al., 2011). The impact of the service failure therefore differs as well across platforms. In contrast with other emotions, fear and surprise occur more in the beginning of the failure on Facebook. This in turn could be related to differences in the communication of the company; on Twitter are frequent updates whereas on Facebook, no information is provided about the failure.

Another insight our analysis provides, is the emotional impact of the service recovery by the investigated company. In a failure and recovery context, customers' recovery expectations are their beliefs about the appropriate level of repair after a service failure (Zeithaml, Berry, & Parasuraman, 1993), which implies understanding recovery as another service performance. The evaluation of customers towards service recovery efforts by the company can reinforce negative reactions to the service (Smith, Bolton, & Wagner, 1999) or improve overall satisfaction, repatronage intentions, and positive word of mouth (Blodgett, Hill, & Tax, 1997; Maxham & Netemeyer, 2003). We observe a strong decline in happiness and a rise in sadness and disgust in the time of the service recovery. The results suggest that overall, the service recovery is not positively evaluated by customers.

Concluding, happiness and sadness tend to move in opposite directions according to the graphs. This is contributing to the belief that these two basic emotions are each other's natural opposites (Ekman, 1992). Besides, it is an indication that the emotion classification method seems to be performing well on social media data, as these emotions should not simultaneously occur.

7.3 Theoretical and managerial take-aways

When things go wrong with a service, and during attempts by the company to fix the problem, customers can experience emotions such as frustration, annoyance, anger, and sometimes even rage (McCollough, Berry & Yadav, 2000). In this case as well, which might be useful for the company to evaluate how they handle a service failure and how their compensation strategy turns out. Not only the failure itself, but mainly KPN's reaction and compensation elicits the emotional response of consumers. A better understanding of basic emotions can help service companies to improve the way they handle a service failure. A different approach may be necessary depending on the social media platform to minimalize negative online word of mouth.

This study is providing insights into customer emotions, which is typically difficult to measure with other techniques. Our method gives a proof of concept for further development of measuring emotional behavior through social media. The experiential nature of services and the nature of emotion make investigations through traditional research methodologies difficult, as customers are often not able to identify their emotions and they tend to rationalize their emotions (Scherer, 2005). Social media analysis has its drawbacks as well but is providing genuine and real-time insights that are hard to collect in other ways. We recommend service providers to make use of multiple analysis methods to form a thorough image of how its service is evaluated.

7.4 Limitations and future research

It would be interesting to work together with a service company in future research. In this way, we might be able to find more explanation from their perspective to the results. In our study, it would be interesting to find out whether the company made a change in policy in social media from January 2015 onwards, to explain the transition to more positive emotions. We are also curious about the firms' own evaluation of the service recovery strategy after the failure.

Our analysis is based on one company in the Dutch telecommunications market. Other service categories or other companies would likely yield different results. For instance a market like the finance industry, where customers are not easily favored towards the companies.

Further research is needed to provide insights into the effective identification and management of customer emotions. In our basic emotions approach by Ekman (1990) for instance, only one positive emotion is present. It would be useful to investigate a broader spectrum of emotions to provide more nuance. Moreover, the results of our study should be compared with traditional market research methodologies to find parallels or differences.

Concluding, the people that are active on the different social media platforms determine which emotions they channel. The results are therefore biased by the people who are active on the social media platform and to what extent they express their experiences there. It would be interesting to see what the demographics are of social media users, to analyze target groups more in-depth.

References

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of Twitter data. In *LSM '11 Proceedings of the workshop on languages in social media* (pp. 30-38). Stroudsburg, PA: Association for Computational Linguistics.

Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 579-586). Stroudsburg, PA: Association for Computational Linguistics.

Alm, C.O. (2009). *Affect in Text and Speech*. Saarbrucken, Germany: VDM Verlag.

Averill, J. R. (2012). *Anger and aggression: an essay on emotion*. New York, NY: Springer Science & Business Media.

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions, and behavior. *Social Psychology Quarterly, 55*(2), 178-204.

Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science, 27* (2), 184-206.

Batson, C., Shaw, L. & Oleson , K. (1992). Differentiating affect, mood, and emotion: toward functionally based conceptual distinctions. In Clark, M. (Ed.), *Emotion. Review of personality and social psychology* (pp. 294-326). Newbury Park, CA: Sage.

Berry, L. L. & Parasuraman, A. (1991). *Marketing services: competing through quality*. New York: The Free Press.

Bitner, M. J. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *Journal of Marketing*, *54*(2), 69-82.

Blodgett, J. G., Hill, D. J., & Tax, S. S. (1997). The effects of distributive, procedural, and interactional justice on postcomplaint behavior. *Journal of Retailing*, *73*(2), 185-210.

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *ICWSM, 11*, 450-453.

Boyd, D. M. & Ellison, B.E. (2008). Social Network Sites: definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*, 210–230.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32.

Bruhn, M., Schoenmueller, V., & Schäfer, D. B. (2012). Are social media replacing traditional media in terms of brand equity creation? *Management Research Review*, *35*(9), 7710-7790.

Calix, R. A., Mallepudi, S. A., Chen, B., & Knapp, G. M. (2010). Emotion recognition in text for 3-D facial expression rendering. *IEEE Transactions on Multimedia*, *12*(6), 544-551.

Chipperfield, J. G., Perry, R. P., Weiner, B., & Newall, N. E. (2009). Reported causal antecedents of discrete emotions in late life. *The International Journal of Aging and Human Development*, *68*(3), 215-241.

Clark, J. A. (2013). Intersections between development and evolution in the classification of emotions. *Developmental Psychobiology*, *55*(1), 67-75.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Cova, B., & Dalli, D. (2009). Working consumers: the next step in marketing theory? *Marketing Theory*, *9*(3), 315-339.

Das, S., Martínez-Jerez, A., & Tufano, P. (2005). eInformation: A clinical study of investor discussion and sentiment. *Financial Management*, *34*(3), 103-137.

DeWitt, T., & Brady, M. K. (2003). Rethinking service recovery strategies: the effect of rapport on consumer responses to service failure. *Journal of Service Research*, *6*(2), 193-207.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169-200.

Feick, L. F., & Price, L. L. (1987). The market maven: a diffuser of marketplace information. *The Journal of Marketing*, *51*(1), 83-97.

Forgas, J. P. (1990). Affective influences on individual and group judgments. *European Journal of Social Psychology*, *20*(5), 441-453.

Gillin, P., & Moore, G. A. (2009). *The new influencers: A marketer's guide to the new social media*. St. Fresno, CA: Linden Publishing.

Gronroos, C. (1988). Service quality: The six criteria of good perceived service. *Review of Business*, *9*(3), 10-23.

Gupta, N., Gilbert, M., & Fabbrizio, G. D. (2013). Emotion detection in email customer care. *Computational Intelligence, 29*(3), 489-505.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: a case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464-472.

Hennig- Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word- of- mouth via consumer- opinion platforms: what motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38-52.

Heskett, J. L., Sasser, W. E., & Hart, C. W. L. (1990). *Breakthrough Service*. New York, NY: The Free Press.

Holzman, L. E., & Pottenger, W. M. (2003). Classification of emotions in internet chat: an application of machine learning using speech phonemes. *Technical Report, Department of Science and Engineering, Leigh University.*

Izard, C.E. & Buechler, S. (1980). Aspects of consciousness and personality in terms of differential emotions theory. In Plutchik, R. & Kellerman, H. (Eds.), *Emotion: Theory, research and experience* (pp. 165-188). London, UK: Academic Press.

Izard, C. E. (2013). *Human emotions*. New York, NY: Springer Science & Business Media.

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.

Kao, E. C. C., Liu, C. C., Yang, T. H., Hsieh, C. T., & Soo, V. W. (2009). Towards text-based emotion detection a survey and possible improvements. In *International conference on information management and engineering.* (pp. 70-74). Kuala Lumpur, Malaysia: IEEE.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59-68.

Kassinove, H., & Sukhodolsky, D. G. (1995). Anger disorders: Basic science and practice issues. *Issues in Comprehensive Pediatric Nursing, 18*(3), 173-205.

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *The Journal of Marketing*, *59*(2), 71-82.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241-251.

Kiritchenko, S., Zhu, X., & Mohammad, S. M. (2014). Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, *50*, 723–762.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790.

Krcadinac, U., Pasquier, P., Jovanovic, J., & Devedzic, V. (2013). Synesketch: An open source library for sentence-based emotion recognition. *IEEE Transactions on Affective Computing*, *4*(3), 312-325.

Kumar, A., Olshavsky, R. W., & King, M. F. (2001). Exploring alternative antecedents of customer delight. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, *14*, 14-26.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Laros, F. J., & Steenkamp, J. B. E. (2005). Emotions in consumer behavior: a hierarchical approach. *Journal of Business Research*, *58*(10), 1437-1445.

Lee, S., & Lang, A. (2009). Discrete emotion and motivation: Relative activation in the appetitive and aversive motivational systems as a function of anger, sadness, fear, and joy during televised information campaigns. *Media Psychology*, *12*(2), 148-170.

Lee, M., Rodgers, S., & Kim, M. (2009). Effects of valence and extremity of eWOM on attitude toward the brand and website. *Journal of Current Issues & Research in Advertising*, *31*(2), 1-11.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*(5), 834-855.

Leventhal, H. (1982). The integration of emotion and cognition: a view from the perceptual–motor theory of emotion. In Clarke, M.S. & Fiske, S.T. (Eds.), *Affect and Cognition* (pp. 121-56). Hillsdale, NJ: Erlbaum.

Liljander, V., & Strandvik, T. (1997). Emotions in service satisfaction. *International Journal of Service Industry Management*, *8*(2), 148-169.

Malthouse, E. C., Haenlein, M., Skiera, B., Wege, E., & Zhang, M. (2013). Managing customer relationships in the social media era: introducing the social CRM house. *Journal of Interactive Marketing*, *27*(4), 270-280.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, *52*(4), 357-365.

Mattila, A. S., & Enz, C. A. (2002). The role of emotions in service encounters. *Journal of Service Research*, *4*(4), 268-277.

Maxham, J. G., & Netemeyer, R. G. (2003). Firms reap what they sow: the effects of shared values and perceived organizational justice on customers' evaluations of complaint handling. *Journal of Marketing*, *67*(1), 46-62.

McAlexander, J. H., Schouten, J. W., & Koenig, H. F. (2002). Building brand community. *Journal of Marketing*, *66*(1), 38-54.

McColl-Kennedy, J.R., & Smith, A.K. (2006), Customer emotions in service failure and recovery encounters. In Zerbe, W.J., Ashkanasy, N.M. & Härtel, C.E.J. (Eds.), *Research on emotion in organizations: individual and organizational perspectives on emotion management and display* (pp. 237-268). Oxford, UK: Elsevier.

McCollough, M. A., Berry, L. L., & Yadav, M. S. (2000). An empirical investigation of customer satisfaction after service failure and recovery. *Journal of Service Research*, *3*(2), 121-137.

Muniz, A. M., & O'Guinn, T. C. (2001). Brand community. *Journal of Consumer Research, 27*(4), 412-432.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436-465.

Mohammad, S. M. (2015). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In H.L. Meiselman (Ed.), *Emotion Measurement* (pp. 201-237). Cambridge, MA: Woodhead.

Morris, J. D., Woo, C., Geason, J. A., & Kim, J. (2002). The power of affect: Predicting intention. *Journal of Advertising Research*, *42*(3), 7-17.

Nielson (2009). *Nielson Global Online Consumer Survey: Trust, Value and Engagement in Advertising.* In Bruhn, M., Schoenmueller, V., & Schäfer, D. B. (2012). Are social media replacing traditional media in terms of brand equity creation? *Management Research Review*, *35*(9), 771.

Oliver, R. L., Rust, R. T., & Varki, S. (1997). Customer delight: foundations, findings, and managerial insight. *Journal of Retailing*, *73*(3), 311-336.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315-331.

Ortony, A., Clore, G. L., & Foss, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*(3), 341-364.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *Language Resources and Evaluation Conference*, 10, 1320-1326.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1-135.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, *81*(12), 46-55.

Ren, F., & Quan, C. (2012). Linguistic-based emotion analysis and recognition for measuring consumer satisfaction: an application of affective computing. *Information Technology and Management*, *13*(4), 321-332.

Risius, M., & Beck, R. (2015). Effectiveness of corporate social media activities in increasing relational outcomes. *Information & Management*, *52*(7), 824-839.

Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (2012). EmpaTweet: Annotating and Detecting Emotions on Twitter. *Language Resources and Evaluation Conference*, 5, 3806-3813.

Roos, I., & Friman, M. (2008). Emotional experiences in customer relationships: a telecommunication study. *International Journal of Service Industry Management*, *19*(3), 281-301.

Russell, S. J., Norvig, P., Canny, J. F., Malik, J. M., & Edwards, D. D. (2003). *Artificial Intelligence: a Modern Approach*. Upper Saddle River, NJ: Prentice Hall.

Scherer, K. R. & Ekman, P. (1984). On the nature and function of emotion: A component process approach. In Scherer, K. R. & Ekman, P. (Eds.), *Approaches to emotion* (pp. 293-317). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695-729.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, *34*(1), 1–47.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6), 1061-1086.

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research*, *36*(3), 356-372.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of Interactive Marketing*, *26*(2), 102-113.

Sparks, B. A., & Bradley, G. L. (2014). A "Triple A" typology of responding to negative consumer-generated online reviews. *Journal of Hospitality and Tourism Research*, *6*, 38-65.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, *29*(4), 217-248.

Strapparava, C., & Mihalcea, R. (2007). SemEval-2007 Task 14: Affective text. In *Proceedings of the 4th international workshop on semantic evaluations* (pp. 70-74). Stroudsburg, PA: Association for Computational Linguistics.

Tirunillai, S., & Tellis, G. (2014). Mining marketing meaning from online chatter: strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, *51*(4), 463–479.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, *22*(12), 2864-2885.

Westbrook, R. A. (1987). Product/consumption-based affective responses and postpurchase processes. *Journal of Marketing Research*, *24*(3), 258-270.

Wierzbicka, A. (1999). *Emotions across languages and cultures: diversity and universals*. Cambridge, UK: Cambridge University Press.

Williams, R. L., & Cothrel, J. (2000). Four smart ways to run online communities. *Sloan Management Review*, *41*(4), 81-91.

Wirtz, J., & Bateson, J. E. (1999). Consumer satisfaction with services: integrating the environment perspective in services marketing into the traditional disconfirmation paradigm. *Journal of Business Research*, *44*(1), 55-66.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science*, *21*(1), 1-12.

Zhu, J., Wang, H., Zhu, M., Tsou, B., & Ma, M. (2011). Aspect-based opinion polling from customer reviews. *IEEE Trans Affect Computing*, *2*(1), 37–49.

Appendices

Appendix A: Text filters in the data cleaning procedure

Column name	“Commercial” filter	“Webcare” filter	“Vacancy” filter
author	gsm	vodafone	vacatur
author	telecom	kpn	banen.nu
author	simkaart	telfort	sollicit
author	4you	simyo	zoektwerk
author	telefoon	hollandsnieuwe	jouwbaan
author	techmania	tele2	bijbaan
author	stoozy	tmobile	jobbird
author	storing	webcare	
author	ondh		
author	mobicom		
author	decom		
author	tweakers		
author	media		
author	provider		
author	ict		
author	b2b		
message_text			vacature

Appendix B: Volume of messages per provider per month

Month	Year	KPN	Vodafone	T-Mobile	Telfort	Tele2	Hollands Nieuwe	Simyo	Total
January	2014	18.349	8.979	4.116	2.596	2.045	750	816	37.651
February	2014	19.768	8.846	3.753	2.888	2.843	628	723	39.449
March	2014	16.079	9.212	3.572	3.787	2.405	697	559	36.311
April	2014	16.166	10.469	3.868	3.614	2.504	796	671	38.088
May	2014	13.686	9.042	3.592	1.837	2.562	958	1.082	32.759
June	2014	14.590	9.657	4.063	1.787	2.932	1.115	863	35.007
July	2014	16.474	11.133	4.079	1.575	2.120	729	884	36.994
August	2014	13.377	9.629	3.397	1.995	1.654	513	951	31.516
September	2014	17.705	15.047	5.104	1.981	2.698	564	808	43.907
October	2014	18.247	11.996	4.271	2.736	3.359	477	724	41.810
November	2014	16.353	9.167	3.440	2.338	3.483	328	708	35.817
December	2014	15.883	8.794	3.228	2.501	2.535	386	844	34.171
January	2015	15.713	8.139	3.784	3.247	2.365	560	607	34.415
February	2015	17.052	9.071	3.137	3.268	2.107	449	639	35.723
March	2015	20.365	12.671	3.324	3.011	2.648	483	678	43.180
April	2015	20.236	11.308	3.759	2.720	2.877	774	653	42.327
May	2015	18.950	9.872	3.443	2.793	2.447	765	1.050	39.320
June	2015	20.015	9.644	3.619	3.245	2.322	875	729	40.449
July	2015	19.275	11.911	3.690	2.986	3.148	827	911	42.748
August	2015	18.933	10.376	9.468	2.827	2.827	768	555	45.754
September	2015	19.002	10.296	3.909	2.499	1.797	717	489	38.709
October	2015	21.407	8.219	3.320	2.896	2.811	684	659	39.996
November	2015	53.200	7.389	5.395	4.833	4.592	855	531	76.795
December	2015	21.069	6.929	3.310	2.627	2.959	1.613	474	38.981
January	2016	18.811	5.989	3.107	2.695	2.912	1.810	513	35.837
February	2016	18.524	9.076	3.138	2.385	3.345	672	504	37.644
Total		499.229	252.861	102.886	71.667	70.297	19.793	18.625	1.035.358
Monthly average amount		19.201	9.725	3.957	2.756	2.704	761	716	39.821