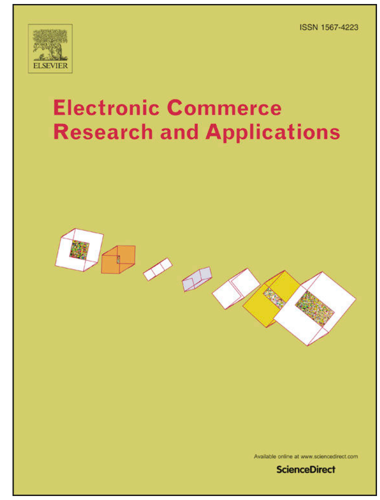
Journal Pre-proofs




Electronic Word of Mouth Analysis for New Product Positioning Evaluation

Andreas Gregoriades, Maria Pampaka

PII: S1567-4223(20)30063-6
DOI: https://doi.org/10.1016/j.elerap.2020.100986
Reference: ELERAP 100986

To appear in: *Electronic Commerce Research and Applications*

Received Date: 31 August 2019
Revised Date: 31 May 2020
Accepted Date: 2 June 2020

Please cite this article as: A. Gregoriades, M. Pampaka, Electronic Word of Mouth Analysis for New Product Positioning Evaluation, *Electronic Commerce Research and Applications* (2020), doi: https://doi.org/10.1016/j.elerap.2020.100986

# Electronic Word of Mouth Analysis for New Product Positioning Evaluation

Andreas Gregoriades
Cyprus University of Technology
Limassol, Cyprus
andreas.gregoriades@cut.ac.cy

Maria Pampaka
The University of Manchester
Manchester, UK
maria.pampaka@manchester.ac.uk

**Abstract**
People increasingly choose to express themselves online through electronic word of mouth (eWOM), generating large amounts of data, making eWOM a valuable source of information through big data analytics. This enables organizations to gain insights directly from customers' opinions for better decision making. This work presents a new methodology for evaluating an organisation's product-positioning strategy through eWOM analytics. A product's mispositioning has significant negative effects and there is strong interest in identifying ways to avoid it. Current methods that utilize eWOM for product positioning evaluation mostly use post-product release reviews and do not statistically evaluate the effect of time on the product positioning; nor do they provide any means to diagnose the cause of mispositioning. The temporal aspect of positioning, however, provides valuable insights into which product features are more time-invariant and accordingly makes it possible to plan for product redesign or repositioning to maximize profitability. A case study is presented in the context of smartphones using design science research, utilizing Twitter data regarding the release of a new product, collected using a custom Android application. The research questions addressed in this paper are: (1) How do consumers' preferences change over time with regards to the product's positioning? (2) Which product features positively influence product positioning and which negatively? To answer these questions, we compared the product-positioning strategy and consumers' opinions before and after the release of a new product to identify possible discrepancies between expected and actual positioning of the product. This work constitutes a methodological contribution with demonstrated implications for new product positioning strategy evaluation using tweet analysis.

**Keywords**
Micro-blogs, Sentiment Analysis, Product positioning, Topic modelling.

## 1. Introduction

The recent information explosion from the proliferation of data in large corporate databases, mobile apps, websites, social media, and sensor networks is yielding new benefits for companies that can discover information patterns hidden in big data sources through effective data mining (Khade, 2016). The rise of social media and the active role of consumers in evaluating products and businesses are changing organizations' reputation (Etter et al., 2019) and sales performance (Babić Rosario et al., 2016), with many practical applications in the area of marketing and e-commence. The diffusion of consumer opinions in social media is often linked with strong emotions such as anger, frustration, excitement or joy (Berger and Milkman, 2012; Pfeffer et al., 2014). The need to analyze these communications has become an essential component of e-marketing, as evidence suggests that anger and high arousal can make a message viral (Pfeffer et al., 2014). This is in line with results indicating that sentiment is a stronger predictor of new product success than the volume of online messages alone (Nguyen and Chaudhuri, 2019). Therefore, social media analytics has become a mainstream activity in e-commerce, used reciprocally to diffuse and collect information from consumers (Jansen et al., 2009), assisting in producing the right product for the right customer at the right time, place and price; this is also known as the marketing mix (Chandrasekar, 2011).

This work focuses on product positioning, performed in the early stages of new product development where a decision should be made regarding the specification of the product's features in accordance with customers' needs (Kotler and Armstrong, 2018; Xiao et al., 2016) and competitors' products. The importance of product positioning is well established in the marketing and e-commerce literature (Kotler and Armstrong, 2018; Lee et al., 2016); during this stage companies establish their markets and address the needs of a specific market segment with specific products, using strategies that promote their product's unique features, benefits to consumers, price difference and quality (Chandrasekar, 2011). This involves analysis of the relationships between the new product's features, competing alternatives and diverse consumers' needs (Kwong et al., 2011; Lilien and Rangaswamy, 2003; Petiot and Grognet, 2006). Hence, positioning as an activity is linked with branding by promoting features which differentiate the product and exactly match the needs of the target customer. Successful positioning improves customer satisfaction, and creates enhanced consumer loyalty and it can thus improve the company's competitive advantage, performance and profitability (Hooley et al., 2001). Positioning errors, such as under-positioning, over-positioning, confused positioning and doubtful positioning (Chandrasekar, 2011) can, however, jeopardize an organisation's performance. Therefore, evaluating the effectiveness of a new product's positioning is key to maximizing the objectives of a firm and assists in repositioning and product redesign (Albers, 1979). This work presents a method for evaluating the performance of a product-positioning strategy and provides diagnostic means to identify causes of successful or unsuccessful positioning.

Optimal product positioning refers to the specification of product features while implementing a set of tactics in designing an image of a product to occupy a distinctive place in the target consumers' minds (Lilien and Rangaswamy 2003; Kotler and Armstrong 2010). Product positioning is thus not only about how companies design a product, but it also involves activities for shaping the image of a product in the minds of prospective customers. This is part of the e-marketing strategy and involves the expression of the company's overall strategy in precise statements in advertising campaigns that convey the difference with other products and the relevance of the said product to consumer needs. Positioning thus highlights the key benefits of a product to the target customers, with the aim of acquiring market share using the appropriate advertising theme, with catchy taglines that do not confuse people. Companies constantly monitor their product positioning by gathering online and offline data and accordingly adjust their communication strategy (i.e. messaging).

Word of mouth and viral advertisement are of key importance in product positioning since they relate to consumer engagement with a new product. Social networks and media are therefore valuable tools in this endeavor, with micro-blogging having an important role complementing traditional marketing avenues (Chamlertwat et al., 2012). Micro-blogs, a form of eWOM, are small messages communicated via social media platforms such as Twitter; they have gained popularity as they allow people to express their views on a wider scale and constitute one type of big data with unstructured information. They emerged as part of Web 2.0 that made consumers co-producers of content and value (de Chernatony, 2001) by providing their opinions freely through micro-blogs and "reviews". Micro-blogs and other forms of eWOM are thus crucial for the evaluation of costumers' perceptions of a brand or product, considering that a brand is no longer what the company tells a customer it is – but, rather, what customers tell each other it is (Nayab et al., 2016). Therefore they are directly linked with product positioning with companies analyzing eWOM as part of their e-marketing strategy (Jansen et al., 2009) to better position their products against customer needs and opinions (Jung, 2008).

Twitter and other social media tools have become valuable sources of eWOM and with natural language processing techniques they can provide valuable insights of consumers' sentiments and opinions. According to Jansen et al. (2009), around 20% of micro-blogs mention a brand name, allowing companies to benefit if they exploit information from these platforms to inform their e-marketing strategy and management of brand perception. Several studies have investigated the use of Twitter and other social media to mine consumer-sentiment for customer behaviour analysis (Moon and Kamakura, 2017) and positioning (Lee et al., 2019), and provided evidence that eWOM sentiment is a stronger predictor of new product success than eWOM volume and that early product reviews through tweets influence consumers' early adoption behavior to an economically relevant degree (Nguyen and Chaudhuri, 2019). Deshpande and Rokne (2018) additionally showed that Twitter provides more critical, objective and instant consumer views, and earlier than online reviews at websites for products feedback. Such information is more likely to be spontaneous and honest compared to rationalized reviews published on a website at later stages (Hennig-Thurau et al., 2015).

To the best of our knowledge no effort has been reported evaluating a new product position against its planned strategy and what causes product mispositioning using tweets. Most studies have focused on comparing products' positions against those of competitors, or on post-product release evaluations (Chen et al., 2015) using product reviews and topic analysis (Wang et al., 2018) with little attention given to the critical effect of time on a product's positioning. Strategic product positioning however, extends beyond the pre-product release optimization or the post-release assessment and aims to build and maintain a sustainable competitive advantage over time by evaluating the effect of time on consumers' opinions. To achieve this, longitudinal information on target market opinions through eWOM analytics is required to improve product positioning, given the significant proportion of consumers who express their opinions through eWOM. Therefore, to address this gap in product positioning studies, we propose a framework that is based on dynamic topic analysis (Tirunillai and Tellis, 2014), to evaluate a product-positioning strategy by comparing the planned and actual strategies, through longitudinal evaluation of consumers' sentiment and opinions over a period of time. Unlike other research (e.g. Tirunillai and Tellis, 2014), the proposed method applies inferential statistics on eWOM's sentiment and discussion topic/theme to identify significant discrepancies in consumers' perceptions of the product's features over time, along with diagnostic capabilities to identify possible causes of mispositioning. A case study applying and validating the method is presented in the context of the positioning strategy of Huawei's P20 and P20-pro products, before their initial release (March 2018) and up to 18 months after consumers' engagement with the products.

The proposed method combines contemporary data analytics techniques for the design of a novel method of evaluating product positioning. To ensure its validity and acceptance, the design science research approach (Hevner et al. 2004; Hevner and Chatterjee 2010) is adopted to guide the design of the method (artefact) that aim to solve the problem of how to evaluate a new product's positioning strategy over time, utilizing eWOM analytics. The artefact of interest in this case is the proposed product positioning evaluation method.

The paper is organized as follows. The next section elaborates on the concepts and related literature on product positioning techniques, sentiment analysis and natural language processing. Subsequently, the method proposed in this study is presented, followed by an implementation case study in the domain of a smartphone's positioning strategy. Data collection and preprocessing stages are illustrated next, while the final sections deal with the analysis of the data, artefact evaluation and the conclusions drawn from the results.

## 2. Literature Review

Provided that positioning aims to identify the optimal new product position in terms of features that can satisfy a consumer segment, Wang (2015) defines the implementation of product positioning through a series of steps: (1) visualizing competitive alternatives and product features, (2) constructing a forecasting model to estimate how potential buyers will react to marketing stimuli, (3) specifying the optimal position of new product(s) and identifying niche segment(s). Several techniques exist in assisting the product positioning process. The role of online communities, particularly in the context of new product development and positioning, has been discussed in many studies (e.g. Franke and Piller 2003). Pang and Lee (2008) highlighted that online product reviews enable marketers and manufacturers to gain a deeper understanding of customers. Zhu and Zhang(2010) showed that online customer reviews can be a good proxy for communicating customer experience, and can shape customers' awareness and perceptions about products (Dai et al., 2017; Le and Mikolov, 2014; Xiao et al., 2016). Similarly, eWOM about products reflects customers' opinions and can influence consumers before they make purchasing decisions (Zhao et al., 2018). It is also considered as a main driver for future product sales (Krouska et al., 2017) and a reliable source of information about product quality (Dai et al., 2017).

According to Kwong et al. (2011) the main research activities in product positioning focus on strategy, modelling and algorithms, and simulation. Studies in positioning strategy specification and its optimization (Gavish et al., 1983) have used game theoretic models (Hadjinicola, 1999), genetic algorithms (Rhim and Cooper, 2005), neural networks (Lee et al., 2019), fuzzy logic (Hsieh and Chen 1999; Gruca and Klemz 2003), dynamic programming and simulation (Kohli and Sukumar, 1990) or other data mining techniques (Lei and Moon, 2015) along with techniques to support market segmentation (Kim and Street, 2004), and product design (Moon et al., 2010). Another stream of work focuses on techniques for determining vacant slots in terms of consumers' perception for the positioning of products in a market, and how to attain or sustain this. Efforts include ways to compare product positioning against those of competitors using techniques such as perceptual maps.

Other studies utilized social media analytics and generated 2D perceptual maps from online product reviews to express consumer perceptions, emotions and topics discussed in reviews (Tirunillai and Tellis 2014; Lee and Bradlow 2011; Netzer et al. 2012; Moon and Kamakura 2017). Perceptual maps are popular tools for visualizing the relationships between competing products and their associated features in a comprehensible way, and they can therefore help firms assess the strengths and weaknesses of competing alternatives and assist in product positioning. Three ways have been mainly used to construct a perceptual map: (a) multi-dimensional scaling; (b) conjoint analysis; and (c) correspondence analysis (Hair et al., 2010). Multi-dimensional scaling involves a series of statistical techniques used for identifying the key dimensions underlying customers' evaluations of products. Each product is represented as a point in the multi-attributed perceptual space, with consumers also represented in the same space according to their preferences (Hsieh and Chen, 1999). Conjoint analysis is a survey-based method which aims to identify how people value different products' features. Correspondence analysis can use categorical variables to project on a plot the relationships between benchmarks and associated products' features, without using multiple regression. Due to its simplicity, correspondence analysis is adopted more often in products' positioning activities by managers, and in the analysis of online product reviews into a product positioning map using psychometric mapping techniques (Moon and Kamakura, 2017).

Despite the wide use of multi-dimensional scaling, conjoint analysis and perceptual maps, these techniques have some important limitations that stem from the fact that they use traditional approaches for identifying consumer behaviour, such as marketing surveys, interviews, focus groups and experiments. These require a great amount of time and resources and can suffer from bias, subjectivity, and preference-change over time (Verma, 2013). Since most electronic products, such as smartphones, have a short-term product life cycle, these approaches are not appropriate due to the time required to collect the data and perform relevant analyses. Specifically, in the smartphone market, the advance of technology is so fast that new versions may be released in less than a year. According to HTC Ltd, the average shelf life for smartphones decreased from three years in 2007 to around six to nine months in 2011 (Ferreira, 2011). According to reported statistics (Ng, 2019), the average duration for people to change their smartphones does not exceed two years (≈22 months). Therefore, producers have very little time to research the market using traditional techniques. According to the Technology Adoption Life Cycle, smartphones have limited time to prove their product adoption due to fast-pace market changes. Work on eWOM-based product positioning includes constructing a positioning map by parsing the perceptions and preferences for competing brands taken from online reviews (Aggarwal et al., 2009; Lee and Bradlow, 2011; Netzer et al., 2012) and translating them in aggregate brand-by-attribute scores on a 2D map or multi-dimensional maps (Lee et al., 2016). Twitter and other eWOM platforms provide the means for micro-blog analysis due to the vast amount of data they handle (Chamlertwat et al., 2012). Natural Language Processing techniques such as sentiment analysis and topic extraction have been very active research fields recently due to the availability of vast amounts of textual data. These methods enable the automatic categorization of users' emotions, opinions and topics of interest based on consumer-generated information (Al-Obeidat et al., 2018). Therefore, these methods could be more efficient in analyzing consumers' perceptions in rapidly changing product domains such as the smartphone market. However, the main limitation of these approaches is the lack of inferential analysis on the effect of time on product positioning. This work utilizes the textual content generated by consumers through eWOM to define topics and sentiments determining the positioning map of a product before and after its initial release. Table 1 lists all described e-WOM-based product positioning studies comparatively to the proposed method. These approaches are evaluated and compared on three main aspects: (1) their analytical techniques; (2) whether and how they consider temporal aspects; and (3) their prediction/diagnostic capability through inferential analysis.

Table 1: Focus and techniques of existing studies on eWOM and comparison to the proposed approach

| **EWOM-based Product Positioning Methods** | **Focus** | **Analytical Technique** | | | **Temporal Comparison** | | | **Inferential Analysis** |
|---|---|---|---|---|---|---|---|---|
| | | **Text Mining (Lexical/text)** | **Sentiment Analysis (Emotion)** | **Topic Modelling** | **After product release** | **Before and after product release** | **Atemporal** | |
| Aggarwal et al. 2009 | Detergent Brands | ✓ | | | | | ✓ | |
| Lee and Bradlow 2011 | Digital Cameras | ✓ | | | ✓ | | | |
| Netzer et al. 2012 | Vehicles and Diabetes Drugs | | | | ✓ | | | |
| Tirunillai and Tellis 2014 | Computer, Footwear and Toys | ✓ | ✓ | ✓ | ✓ | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Markets | | | | | | | |
| Moon and Kamakura 2017 | Hotel and Wines Markets | ✓ | ✓ | ✓ | | | ✓ | |
| Chamlertwat et al. 2012 | Mobile Phones | ✓ | ✓ | | | | ✓ | |
| Lee et al. 2016 | Mobile phones | ✓ | ✓ | ✓ | ✓ | | | |
| **Proposed method** | Mobile Phone | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ Diagnosis of product positioning problem |

### 2.1 Natural Language Processing and Thematic Analysis

Text mining is the process of automating the information extraction from unstructured textual data generated by users and is a popular natural language processing activity. The two main frameworks for text mining are statistical modelling and machine learning. The bag of words[1] technique, belongs to the first category and involves the representation of text as vectors that describe the occurrence of words within documents (D'Andrea et al., 2015). A more sophisticated technique, the term-frequency-inverse document frequency[2] (Salton and McGill, 1983), counts the number of occurrences of each word in each document and applies appropriate normalization to the frequencies of words. The word embeddings technique (Tang et al., 2015) is an alternative to bag of words (Le and Mikolov, 2014), representing each word by a low-dimensional vector. These techniques[3] must be trained using a large amount of data (Dai et al., 2017). However, word embedding models miss the context of documents and neglect sentiment information. Therefore, probabilistic techniques such as Latent Dirichlet Allocation (LDA) and Latent Semantic Indexing were introduced to overcome these limitations. LDA is an unsupervised three-level hierarchical Bayesian model applicable to discrete data based on random sampling (Blei et al., 2003) capable of discovering latent topics in large volumes of data. LDA is currently one of the most widely applied topic modelling techniques. Each document in LDA is a random mixture of topics, characterized by a probability distribution over a set of words. LDA is a bag of words model, so it ignores the order of words in a document. The model assumes that in a total number of documents and a total number of words in a document, each document features $K$ topics. LDA considers a document as a collection of topics and each word in the document as part of one or more topics. LDA clusters words under their respective topics and provides the probability of each word belonging to a topic and the probability of each topic belonging to each document. An LDA model, however, might miss topics from short text in non-specific contexts and ignore the relationships between topics and documents when word probabilities are low. Ontologies that correspond to structured representations of information



---

[1] Bag of words is the oldest model for representing documents (Le and Mikolov, 2014) as a dictionary containing all the words occurring in the document. The model is easy to implement, works fast, and achieves good results with little data. However, it ignores word order and semantics and can yield word vectors in high dimensionality.

[2] The result of term-frequency-inverse document frequency is a term-document-matrix whose columns contain the normalized frequency of words in each document

[3] Word2vec, Doc2vec, and Global Vectors for Word Representation

regarding a specific domain are considered an alternative approach that might enhance the performance of LDA when used in combination to find appropriate topics.

## 2.2 Sentiment Analysis

Sentiment analysis (SA) and opinion mining have been studied for more than two decades with several techniques emerging during this time for analyzing emotions and opinions from eWOM (Martin-Domingo et al., 2019). The majority of work in SA, however, considers it as a classification problem; nevertheless, SA may involve many other natural language processing activities such as sarcasm detection (Sulis et al., 2016), subjectivity classification (Lu and Tsou, 2010), polarity classification (Pang et al., 2002), word sense disambiguation (Kageb and Salomonsson, 2016), opinion summarization (Ku et al., 2006), and opinion spam detection (Ren and Ji, 2017), which highlight the complexity of SA. Of these, polarity detection and subjectivity classification are the most popular SA activities that focus on extracting data patterns and discovering dynamic trends (Lei and Moon, 2015) based on textual concepts from eWOM. Polarity detection is useful for online opinions analysis due to its ability to automatically measure emotion in online content using algorithms to detect sentiment in eWOM (Pang and Lee, 2008). Subjectivity classification refers to the task of classifying a sentence as objective or subjective, indicating opinionated or not opinionated content (Gambhir and Gupta, 2017), using the presence of certain terms such as adjectives, adverbs and group of verbs and nouns. Subjectivity detection is an essential subtask of subjectivity classification since most polarity detection techniques are optimized by distinguishing positive from negative text, while subjectivity detection ensures factual information is filtered out and only opinionated information is used for polarity detection (Liu, 2010) making the latter more relevant to eWOM opinions. Some polarity detection algorithms assign an overall polarity to a text ranging from negative to positive, while others identify topics and their associated polarity estimate (Gamon et al., 2005). Three common SA approaches are: Machine Learning Lexicon-based Methods and Linguistic Analysis techniques.

Machine learning techniques can be either supervised or unsupervised (Witten et al., 2016), the main difference being that supervised techniques use labelled opinions that have been pre-evaluated as negative, positive or neutral to train models. Such techniques include, Support vector machine, Naïve Bayes, Logistic regression, Multilayer perceptron, K-Nearest Neighbours and Decision Trees (Krouska et al., 2017). Machine learning approaches need a training and testing dataset to learn the model from corpus data and verify its performance (Krouska et al., 2017; Song et al., 2017). The labelled data can either be annotated by human coders in terms of polarity or through automated techniques that approximate micro-blogs' polarity using either user-rating or emoticons. These datasets are then used to train a model to detect features associated with positive, negative or neutral emotions of micro-blogs. The trained model can look for the same patterns in new texts to predict their polarity (Thelwall et al., 2011). Sets of words (n-grams, with n denoting the number of items which can be words, letters, etc.) can also be used during model training.

Lexicon-based approaches use lists of words that are pre-coded for polarity and process their occurrence within texts to predict the overall polarity of a paragraph (Thelwall et al., 2011). The most popular lexicon in the field of sentiment analysis is SentiWordNet (Gatti et al., 2016). However, this approach may not yield satisfactory results in some domains due to differences in word meanings. Therefore, domain-dependent lexicons are used in such cases. Moreover, lexicon-based methods are unable to capture the underlying structure of grammar in a sentence.

Finally, the linguistic SA analysis method exploits the grammatical structure of texts to predict its polarity, using a lexicon. Therefore, linguistic algorithms could identify context and idioms as part of the polarity prediction process (Thelwall et al., 2011).

Of the above three categories, machine learning techniques are considered the most effective and simplest to use, with Naïve Bayes and Support Vector machines being the most popular. As these are supervised learning techniques, it is important to train the classifiers prior to their use. A training dataset contains labelled cases, and the quality of this data affects the model's accuracy. This constitutes the main limitation of machine learning techniques. Manually labelling massive datasets such as tweets is time consuming (Lee and Kim, 2017), while domain-specific data is important to achieve satisfactory classifier accuracy in the given domain (Aue and Gamon, 2005). Therefore, in most cases the training is performed using either labelled datasets related to the domain of interest, or emotion vocabularies (Read, 2005). Naive Bayes is the simpler and more efficient classifier, and has been shown to work well in many domains, and in large datasets (Abellán and Castellano, 2017); it produces better results than other Machine learning techniques and in micro-blog polarity evaluation (Sailunaz and Alhajj, 2019). Therefore, Naïve Bayes was utilized in this study.

## 3. Methodology

The design and development of new artefacts in design science research[4], such as the evaluation method proposed in this work, requires a systematic approach towards (artefact) design, development and evaluation to guarantee that the new artefact contributes to the problem domain. The work presented in this paper is based on design science and offers an approach for guiding the evaluation of product-positioning strategy for organizations. Here, the artefact is a new evaluation method for product positioning. The evaluation of an artefact in design science research can be achieved through analytical, case, experimental, field study, or simulation methods (Hevner et al., 2004). Pries-Heje et al. (2008) added that evaluation can be ex-ante or ex-post, naturalistic or artificial, with the ex-ante perspective offering the possibility of evaluation prior to developing the artefact and ex-post referring to the evaluation of the developed artefact using concrete terms. Naturalistic evaluation uses real artefacts to solve real problems (Sun and Kantor, 2006), while artificial evaluation methods aim to test an artefact using artificial means such as simulation (Gregoriades and Sutcliffe, 2018), lab experiments or mathematical proofs. The technology acceptance model is considered the most influential evaluation framework in information systems (Davis, 1989); however, despite its popularity it has been criticized as leading to biased results due to use of subjective data for measuring intended system use in an ex-ante method (Bagozzy, 2007). Therefore, the method proposed in this paper is based on objective (volume of reviews) combined with subjective measures (eWOM) quantified using topic and sentiment analysis, and therefore is classified as ex-post. The evaluation of the artefact is achieved through a case study implementation utilizing eWOM opinions along with comparative analysis with results from Google Trends, reported to be linked to consumers' adoption (Nguyen and Chaudhuri, 2019) and product sales (Schaer et al., 2019), that indicate successful product positioning.

Based on relevant recommendations (Jones and Gregor, 2007) that suggest eight critical components of design science research, we develop our design artefact (the product positioning evaluation methodology) as detailed in Appendix 1. The main steps of the proposed method for evaluating the effectiveness of a product-positioning strategy using online opinion data are depicted in Figure 1. The details of each step are elaborated through a case study implementing the method, presented in subsequent sections.

The first step is the examination of the product's marketing strategy and associated product positioning. An important element of the marketing strategy is the target customer segment and the specification of

---

[4] Design science, as inspired by Simon (1996), is conceptualized as generic, validated and actionable knowledge that can be used to design and implement actions, processes or systems, to bring about the desired change in organizations.

the product-positioning statement that explicitly specifies what customer needs are supported by the new product, and in what way this is better than competitor products. This is specified by the marketing team and is used prior to new product release in the marketing campaign.

The next step of the methodological framework focuses on collecting relevant micro-blogs regarding the target customers of the product. In this step an automated technique is used to collect the data based on specific criteria that relate to the product under investigation. Twitter is considered the most appropriate platform for this task since it enables users to freely express their opinions in an unconstrained manner, unlike other product review platforms that expect users to respond to specific rating scales. In addition, the limit in characters forces users to articulate their opinion in a concise manner which eases the analysis process. The collected data needs to undergo pre-processing, which involves data cleansing, dimensionality reduction and elimination of irrelevant data. This is a necessary step to enable data analysis and the extraction of meaningful insights regarding the success of the positioning strategy. The next step involves analysis of consumers' sentiments and the topics of eWOM through polarity detection and topic/thematic analysis.

Given the enormous amount of data generated by Twitter users every day, categorizing the content into topics is a challenging task. Twitter tried to alleviate this problem using the #hash-tags. However, according to Mazzia and Juett (2010), only 16% of all tweets are hash-tagged, possibly because the limit on the number of characters in each tweet includes the hashtag. Therefore, to correctly analyze a discussion topic (i.e. the product of interest in our case), it is essential to retrieve as many of the remaining 84% of untagged tweets and use automated topic modelling and classification techniques to categorize them. In the proposed framework the LDA topic modelling approach is utilized due to its popularity and proven results. In LDA, a topic is a probability distribution function over a set of words, used as a type of text summarization. LDA is a Bayesian version of probabilistic Latent Semantic Indexing and leads to better generalization.

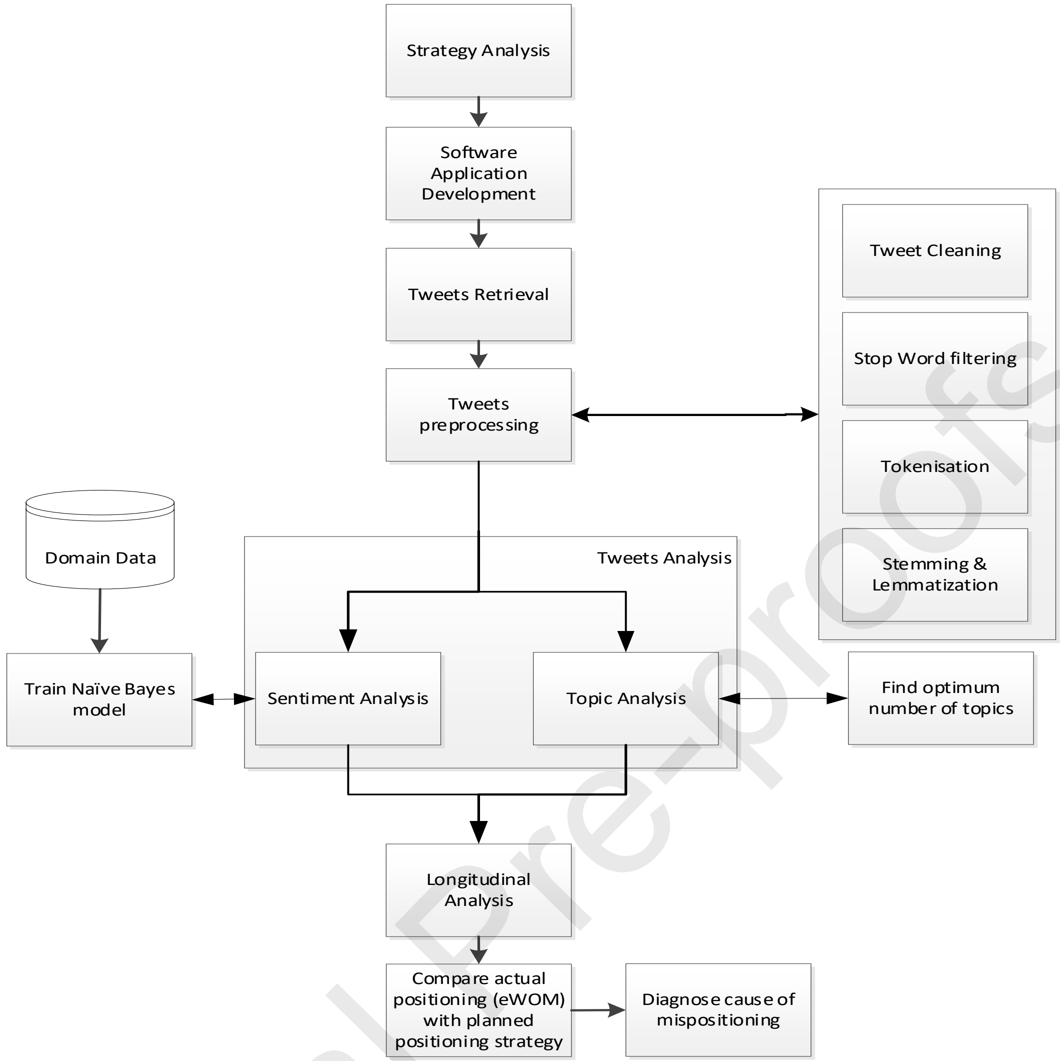



Fig. 1. Overview of Methodology

For these tasks a python algorithm was developed to identify the main topics in the dataset and to assign the most likely topic to each tweet. A Naïve Bayes model was utilized that evaluated the polarity and subjectivity of each tweet.

The next step addresses the longitudinal effect of consumers' perceptions of the product, based on their sentiment during the periods before and after official release. This is evaluated using ordinal logistic regression. The final step in the process involves a comparison between the planned product-positioning strategy and the perceptions of consumers as disseminated through eWOM. During this step, a comparison is made between how consumers perceived the product and the way the company planned the product positioning. Results from the polarity and topic analysis provide valuable insights to marketing teams regarding the effectiveness of their campaigns.

It should be noted that user privacy and ethical considerations for accessing and using online data are adhered to at all times in the use of the method described in this paper. We have used only public Twitter accounts for which users gave permission for anyone to access their eWOM. The use of the OAuth (open standard for web access delegation) process also addresses the need for privacy and data protection ensuring no personal account details are accessible.

## 4. Case Study

To validate the proposed method a case study is used, similar to that of Chen et al. (2015), utilizing eWOM data to evaluate the positioning strategy of the Huawei P20 and P20pro smartphones. The proposed method further utilizes regression modelling in addition to topic modelling and sentiment analysis to investigate the effect of time on product positioning. To illustrate the application of the method it was essential to isolate data relating to these new products and to monitor the eWOM generated about the products before and after their official release date (27 March 2018). P20 products were selected since they were new and were expected to yield a significant volume of eWOM, given the customers they targeted. Below, we illustrate the steps followed within the proposed method, in the context of the P20 case study. The assumption made here is that if the company is successful in satisfying the needs of the targeted market segment, this will be reflected in its consumers' eWOM. In this study we explore the effect of the strategy on eWOM, before and after the release of the products, to evaluate consumers' anticipation of their qualities and perceived features before and after engaging with the product.

### 4.1 Strategy analysis

The analysis of Huawei's marketing strategy was necessary to examine if the consumers' perception of the new products was in line with the expectations of the company prior to the smartphone's release, based on the firm's dissemination material. To analyze the strategy it was essential to first identify the company's competitive advantage for a chosen market, and classify its components into one or more of the three generic categories: differentiation, cost leadership, or focus strategy (Porter, 1980). In differentiation strategy, a company aims to provide superior product-value to the customer; in cost leadership it attempts to optimize business processes to realize cost-saving; and focus strategy combines the above two in a niche market. In product positioning, these strategies are expanded to product features, benefits to consumers (value), price difference, quality, category, competitor, user groups and application domain (Chandrasekar, 2011).

A qualitative analysis of the firm's disseminated documents revealed that Huawei adopted a value-based and cost-differentiation strategy for the positioning of both of its new products. The company aimed to target consumers who look for a high-specification smartphone, usually provided by competitor brands such as the major competitors - iPhone and Samsung - at a lower price. Based on this information, Huawei's marketing strategy aimed to promote the unique and competitive features of quality photography of the new P-series smartphones. The company emphasized the camera's high resolution to differentiate itself by focusing on features such as the triple-camera introduced in the smartphone market for the first time, and the intelligent photography using AI. To promote these features the firm used the See More slogan, referring to the Leica triple-lens system, which was presented in all promotional campaigns. To verify that this positioning was successful, and that consumers perceived the product as planned, the proposed methodology is utilized.

### 4.2 Data collection

A data-gathering mobile application was developed to collect relevant tweets using the Android platform, which emailed the tweets to the researcher. The reason for developing an app was the convenience of initiating a weekly tweets collection. To obtain access to Twitter's APIs, the mobile app was registered on the Twitter platform. This is a typical procedure required by Twitter to secure access to its database by authorized users. To filter the data obtained from the platform, specific keywords relating to the product-positioning strategy and product features were used. Twitter offers two options for retrieving tweets, the Rest API and the Streaming API. For this app the Rest API was used, which allows for searching specific tweets using certain criteria such as keywords, location, etc. The timeline of the data gathering was limited by the API to seven days. Figure 2 illustrates the main screen of the developed Android app.


Fig. 2. Screenshot of the developed Android application for retrieving tweets relating to specific keywords, locations and dates

The specification of the tweets query keywords was based on the analysis of the two products' features and the company's positioning strategy. Therefore, terms such as #HuaweiP20, #HuaweiP20Pro, and #Huawei, were used. The slogan #SeeMore related to the products' photography, referring to the product's positioning strategy as a smartphone with sophisticated mobile photography, the key feature of its marketing campaign. Collected tweets were restricted to English language, and the target market was Western Europe, for its ranking in active Twitter users worldwide, with emphasis on the countries where the product was first released. According to descriptive statistics of this sample of tweets, 45% were from the UK, 20% from Germany, 15% from France, and 20% from the Netherlands. It was not possible to obtain the age distribution of the micro-bloggers, although given the characteristics of the target market they were assumed to be between 25 and 50 years old.

Data collection was performed on a weekly basis, within the time limitations of Tweeter API and it covered a period from 6 months before release to 18 months after the product was released. This was required in order to evaluate the sustainability of the firm's product positioning. On 27 March 2018 the Huawei smartphones were officially introduced during an event that took place in Paris. Hence, this date was the reference point for our data analysis.

Although the app enabled the collection of tweets up to seven days before, to have full coverage of older relevant tweets for specific locations an additional python script was prepared utilizing a library (i.e. GetOldTweets)[5], that enabled the gathering of past tweets based on the same keyword criteria. The final joint dataset contained 5,994 relevant tweets from September 2017 to August 2019.

[5] https://pypi.org/project/GetOldTweets3/

In order to address the problem of customers referring to product features using different linguistic terms, it was necessary to create an ontology to enable the grouping of similar terms under appropriate categories. This was realized through a python script utilizing ontology and natural language processing libraries (OWL API and NLTK libraries)[6] to manually develop an ontology based on the products' feature categories obtained from Huawei's website and frequent terms from the tweet corpus. The ontology was used prior to the topic analysis to minimize the variability of terms referring to similar concepts.

### 4.3 Data pre-processing

Pre-processing refers to the procedure of cleansing and preparing the tweets before analysis. Unstructured information on the Internet, like tweets, contains significant amounts of noise, such as data with no useful information for the analysis at hand. Filtering tweets for irrelevant information preceded the analysis, to eliminate useless metadata, and keep tweets in a simple format.

| |
|---|
| *"# GirlTalkZA # KingTalkZA I â€™ took a picture on the beach during my time with my parents with the Huawei p20 pro, and the image is amazing 😊, Iâ€™ve zoomed in from the original pic @ HuaweiMobile # Autumn # HuaweiP20Propic.twitter.xxx/xxxx"* |
| *"I took a picture on the beach during my time with my parents with the Huawei p20 pro, and the image is amazing, zoomed in from the original pic"* |

Fig.3. Sample tweets before and after cleansing of irrelevant content

The pre-processing involved the following steps, to facilitate sentiment analysis and topic extraction. An example of a tweet showing the result of this stage can be seen in Figure 3.

***Tweet cleaning***

In order to achieve minimum redundancy in tweets, duplicate tweets, i.e., tweets with the same tweet ID, possibly retrieved from different searches, were eliminated using the MongoDB manipulation. Non-English tweets that occasionally emerged due to keyword similarities were also eliminated. Subsequently, tweets were cleaned by applying a regular expression filter, extracting only the text from each tweet, and ignoring all irrelevant metadata such as user-id, tweeter-id, etc. Hashtags were not eliminated since they contained valuable keywords. Therefore, they were transformed into words without the hashtag symbol. Additionally, punctuation, irrelevant characters, links, URLs and protocols were eliminated since they contain no valuable information regarding the sentiment of the tweet. Further cleaning steps involved replacing common symbols and numbers with their word equivalents (e.g. $ to dollar), replacing abbreviations with their full-text equivalents (e.g. "BTW" becomes "by the way"), converting contractions back to their base words (e.g. "wouldn't" becomes "would not"), and converting text to lowercase for improved term aggregation.

***Stop-word removal***

*Stop-words* are words providing little or no useful information to text analysis, and are considered as noise. Common stop-words include articles, conjunctions, prepositions, pronouns, etc. Other stop-words are those typically appearing very frequently in sentences, or in specific contexts. In this work, we employ an augmented version of the stop-word list to include common words that occur frequently such as "*and", "is*", and "*the",* and words referring to irrelevant abbreviations of text.

***Tokenization***

*Tokenization* transforms a stream of strings into a stream of processing units, referred to as *tokens*. Thus, during this step tweets were converted into a sequence of tokens, by choosing *n*-grams (phrases of n words in length) as tokens after removing punctuation marks and special symbols.

[6] https://pypi.org/project/Owlready2/

***Stemming and lemmatization***
*Stemming* is the process of converting a word to its root form and is typically required in dealing with fusional languages, like English. *Lemmatization* uses a vocabulary and morphological analysis of words, to return the base-form of a word, known as the lemma. Lemmatization, unlike stemming, reduces the word to its lemma, ensuring that the root word belongs to the language and context of interest. Stemming usually employs a heuristic process that eliminates endings of words and often results in the removal of derivational affixes. This process is sometimes called word normalization in natural language processing, and consists of reducing each token to its *stem*, to group words with closely related semantics; for instance, “playing”, ”plays” and “played” become “play”. Stemming and lemmatization derive the families of related words with similar meanings such as “liked”, “liking”, “likability” => “like”.

***Filtering***
*Stem filtering* removes words (stems) considered as irrelevant. Thus, each tweet is cleaned of items not belonging to the set of relevant stems.

## 4.4 Feature extraction

Feature extraction is required during model training. Machine learning models are not able to process words directly, so these are converted to numerical vectors of features that represent words or sentences. There are different feature extraction techniques[7]. An initial descriptive analysis of features is performed using bag of words and Term-Frequency-Inverse-Document-Frequency, yielding the results shown in Figure 4 depicting a consolidated log-tail chart with the ten most popular terms. From these graphs it is evident that the terms “pic” and “camera”, for the timeframe of the study, are the most frequent, indicating that the slogan of the product-positioning statement is well perceived by consumers and is used during eWOM.

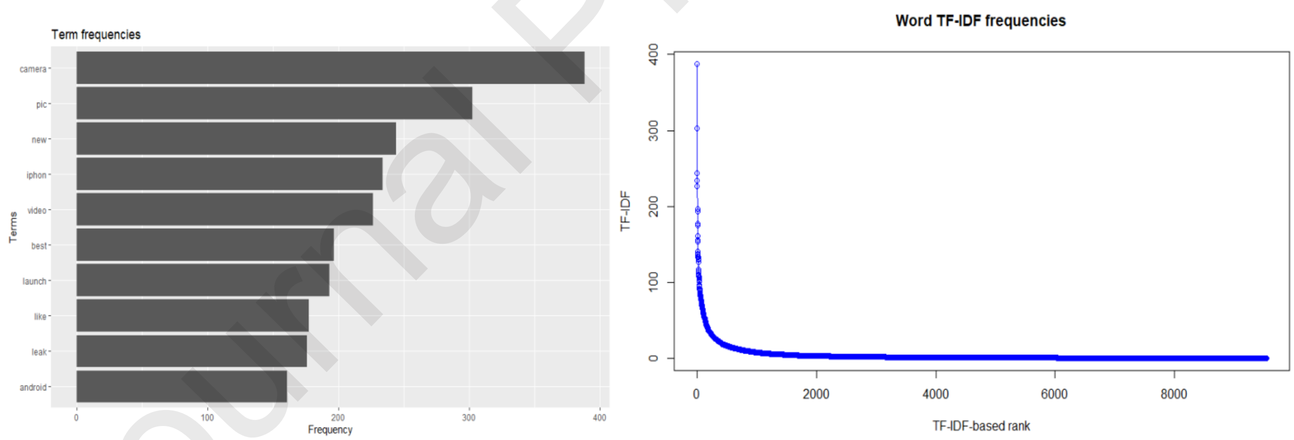



Fig.4. The ten most frequent features and the associated Term-Frequency-Inverse-Document-Frequency

[7] Count Vectors, n-grams, bigrams, Bag of Words, Term-Frequency-Inverse-Document-Frequency and word embeddings (Xia et al., 2011). Bag of Words and Term-Frequency-Inverse-Document-Frequency are the most popular, that convert raw text to features coded in numeric form, then can be fed to machine learning algorithms for sentiment analysis/topic modelling

## 5. Analysis and Results

Tweets underwent pre-processing and feature extraction prior to polarity and LDA model training. The trained models were used for topic classification and polarity detection for each tweet.

### 5.1 Topic analysis

Thematic analysis aimed to identify the main topics on the collected tweets.

LDA is employed in this study, with each tweet representing a distribution of a finite set of topics, and each topic assumed as a multinomial distribution of the words in the corpus developed from all tweets. LDA examines a collection of tweets and learns which words tend to be used in similar tweets. This is a classic clustering problem with the derived clusters corresponding to the topics. A main element of LDA is that clusters are not distinct.

Prior to generating an LDA model, an additional step is needed to improve its results: the specification of the optimal number of topics (k) to be used for tweets' classification, since different numbers of topics will also affect the outcome of the analysis. Popular goodness-of-fit methods include likelihood metrics such as log-likelihood (Griffiths and Steyvers, 2004), and non-likelihood measures. The intruder test (Chang and Gerrish, 2009) - also known as coherence test - automatically identifies uncommon words (intruders) in the corpus. Divergence metrics such as Kullback-Leibler (Deveaud et al., 2014) estimate the number of latent concepts, based on the word distributions that maximize information divergence between all pairs of topics. Divergence assesses how similar (or different) two probability distributions are. Alternatively, the coherence score can be used, representing a probabilistic measure of a topic-term's coherence, and evaluating whether words in the same topic make sense when they are put together. This is expressed in the form of probability density and dissipation of vectors around the centre of the distribution (Korenčić et al., 2018). The higher the score for the specific number of k, the more meaningful the topic. For instance, in the topics T1:{picture, camera, photography} vs T2:{phone, holiday, revolution}, T1 will yield a higher coherence score than T2 since its words are more closely related. This method can be implemented through an R package (i.e. "ldatuning")[8] and is used in this case study to find the optimum number of topics that would yield the optimum topic coherence. The result indicated that around ten topics should be used by the LDA model to achieve maximum coherence, as illustrated by the arrow in Figure 5. The distribution of coherence probability indicates that the maximum coherence for the trained model lies in the region of 11%; however, given the variability of terms characterizing smartphones, "battery" and "picture" for instance have little to no coherence in English but are related in a smartphone ontology. Therefore, the score was used as an indication of how to maximize coherence while minimizing the number of topics, to ease comprehensibility.

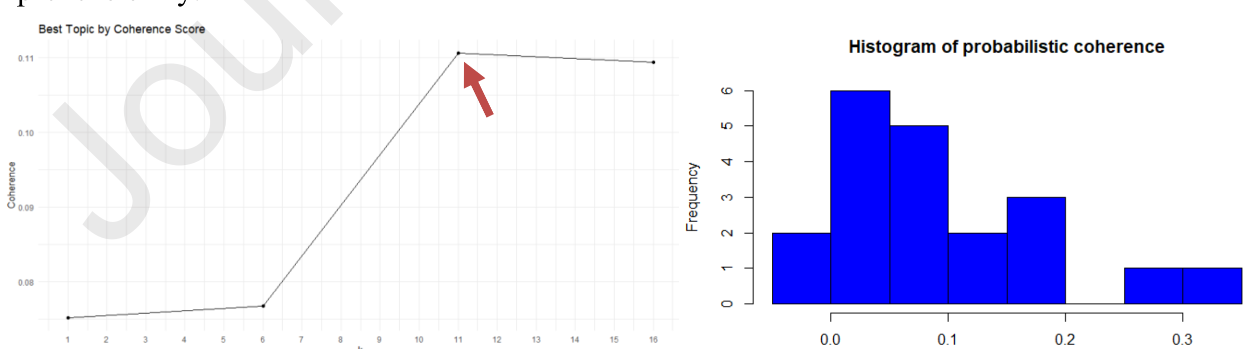



Fig.5: Identification of the optimum number of topics for LDA analysis using coherence score

[8] https://cran.r-project.org/web/packages/ldatuning/ldatuning.pdf

The LDA model learns using the Gibbs sampling technique that essentially performs a random walk in a way that reflects the characteristics of the desired distribution, starting at a random initial point. To improve the comprehension of the generated model, the terms in each topic are ranked by their frequency. This is expressed by the beta values that are the Dirichlet priors for tokens over topics. Extracted topics were inspected based on prior domain knowledge, drawing on expertise in the field under investigation to make the necessary connections. The refined number of topics for the final LDA model was eight, after evaluating results of various topic sizes based on the estimated number of k. These provided fairly distinct topics that could be used to evaluate consumers' eWOM. Results from the LDA model learning yielded the topics depicted in Figure 6, represented by the distribution of words in each topic. From these topics, the following super-topics were identified: phone's photography and main characteristics (covering topics 5, 7 and 8), phone's video capabilities (topic 6), the release of the product (topics 3 and 4), product features such as battery and looks (topic 1), alternative products such as other brands and their capabilities (topic 2). These clusters of topics were formed manually and used subsequently for the interpretation of the eWOM before and after the product release. The main steps performed during topic analysis are shown in Appendix 2.

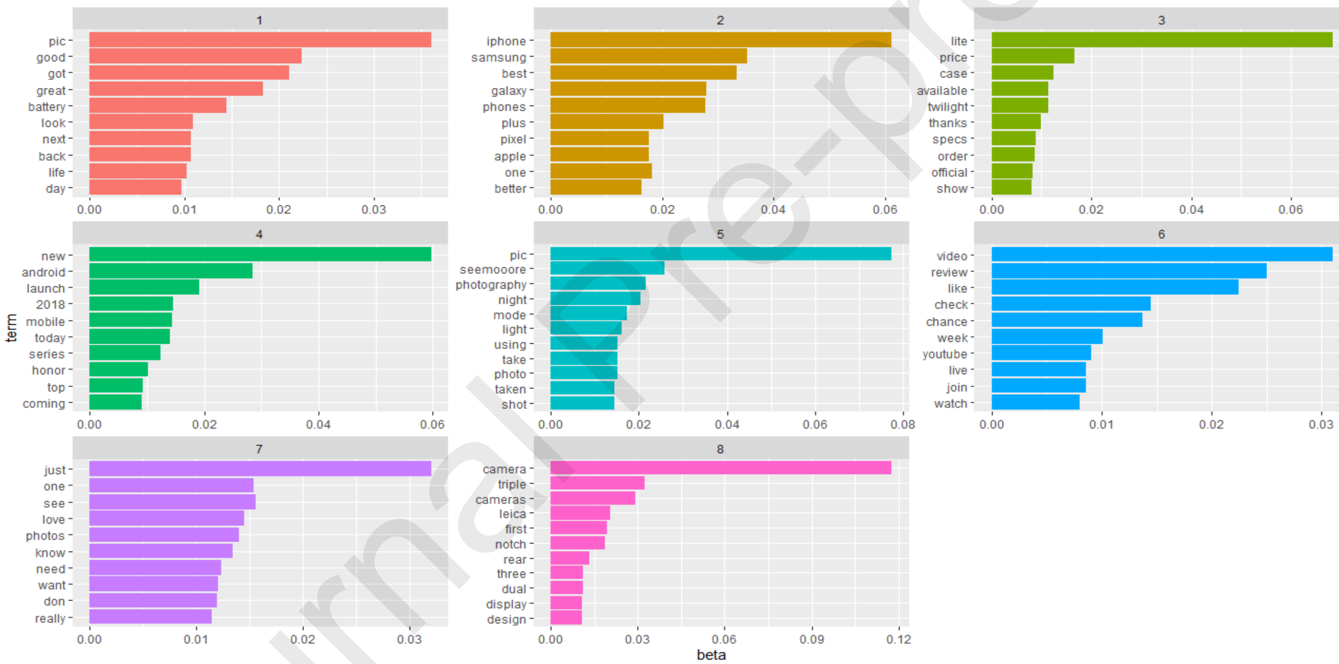



Fig.6. Identified topics from the tweets' dataset and the associated terms' probability distribution

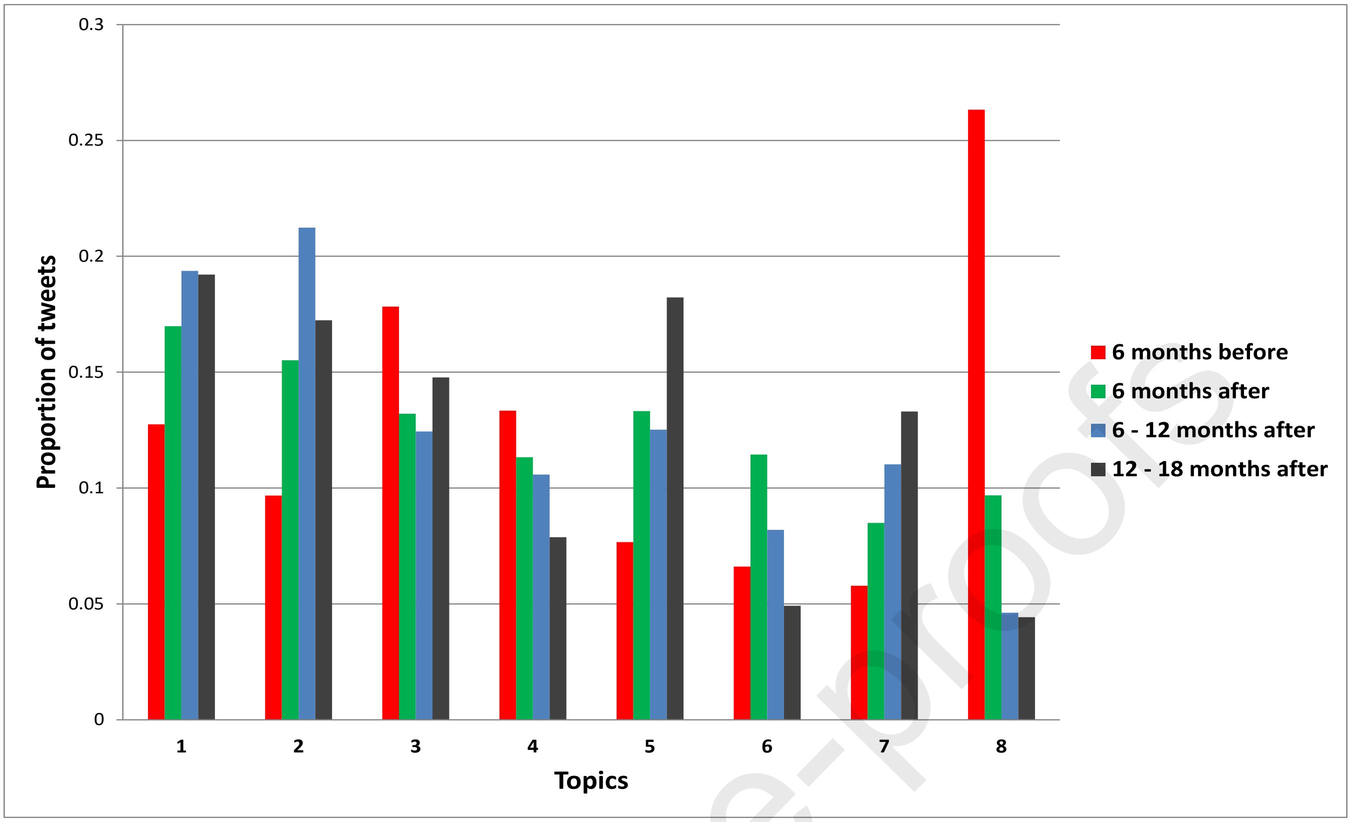



Fig.7: Distribution of tweets across topics within the different product lifecycle timeslots

Following the learning of the LDA model and the identification of the main topics, each tweet was associated with the most prevalent topic based on the trained model and the result was saved in a new matrix which was normalized and subsequently used to visualize the distribution of tweets across topics within each timeslot (6 months before product release, 6 months after, 6-12 months after and from 12 months to 18 months after), as depicted in Figure 7. From this analysis it is evident that consumers' eWOM and associated topics changed over the period of 24 months covered within this analysis. Specifically, topics which refer to the initial release of the product, its price and camera (topics 3, 4 and 8) demonstrate higher eWOM activity before product release (up to 6 months before). The discussions of photographic features (topic 8) decline significantly in timeslots 6-12 and 12-18 months after; in the same time period there is an increase in discussions of alternative products from other vendors (topic 2). In the same vein, product features such as battery (topic 1) are discussed more 6 and 12 months after its release, which might indicate an issue or satisfaction with battery life. Overall the discussion regarding the main product's differentiating features, photographic quality and the triple camera, remained relatively stable, indicating that the product positioning was successful. On the other hand, discussions regarding the video quality (topic 6) of the product declined after 6 months, which could indicate some issues or interest in the topic.

Based on the results from the topic analysis and Huawei's positioning strategy (product differentiation through value benefit and reduced cost), there is an alignment between the planned and actual reaction from consumers. The firm's product positioning was to promote the new smartphone as a cost-efficient high-quality phone with improved camera features, such as the triple-camera. From the tweets' analysis and the LDA model, the camera features and the price of the product are discussed more frequently after the initial release, which indicates that the product positioning was aligned with the consumers' opinions.

### 5.2 Sentiment analysis

SA was a central aspect of this work, aiming to automatically evaluate the polarity of each tweet and to gauge how the public opinion fluctuated over time. However, despite their popularity, SA and associated assessment techniques have been criticized for their lack of accuracy, and specifically with regard to the detection of polarity of plain text content. For example, in certain domains the accuracy of sentiment analysis was lower than 50% with less precision in detecting negative sentiments (Jongeling et al., 2015). Research suggests that machine learning classifiers perform better than lexicon-based solutions; however, their superiority is often limited to the domain they have been trained for. To alleviate this problem, we first trained a classifier based on tweets from a similar domain using features identified in the feature extraction step. We also constrained the search criteria to specific keywords and hence eliminated irrelevant tweets that could have influenced the results.

A python algorithm was developed to train a Naïve Bayes classifier (using the "NLTK" library)[9], to evaluate the polarity of tweets in three categories: positive, negative and neutral. The dataset (i.e. SemEval)[10] used to train the Naïve Bayes classifier consists of tweets (around 65K) that were manually labelled (Nakov et al., 2013) and found to be 26% positive, 18% negative, and 56% neutral. However, this model was not explicitly trained on mobile phones data, and since the performance of a sentiment model is dependent on the similarity of its training data to the applied data, the results of the classifier when applied to the Huawei dataset were not satisfactory, achieving an accuracy level of 65%. Therefore, an alternative approach was evaluated using two pre-trained models, Textblob and Vader (Hutto and Gilbert, 2014), which are popular alternatives with satisfactory precision and recall scores. Textblob's similarly with Vader is based on bag of words but the former includes subjectivity analysis estimates. The metric of subjectivity is in the range of [0-1] with 1 referring to subjective and 0 to objective content. Both classifiers were used in an ensemble manner (merging the results from the two classifiers) to improve our confidence in the results. To verify the predictive accuracy of the combined classifier's results on the case-study's dataset, a sample of 100 tweets was manually examined for true-positive (TP), true-negative (TN), false positive (FP) and false negative (FN), to estimate accuracy and precision. The model's accuracy is expressed by the percentage of correctly classified tweets and is equal to (TP + TN)/(TP + TN + FP + FN). This test-sample corresponds to a margin of error of 9.7% for a confidence level of 95% and a population of 6K tweets. Classifiers' results indicated an overall accuracy of 71% which was deemed satisfactory based on a 9.7% margin of error, which translates into tweets correctly classified in the range of [3857-4673] out of 6K. The python algorithm automatically preprocessed all tweets and through the use of the two aforementioned SA models, iteratively assessed polarity and subjectivity of each tweet, and in an ensemble fashion averaged their estimates before saving the results in a new (csv) file.

An aggregate evaluation of the product eWOM sentiment was found to be positive after the initial product release, with a ratio of positive over negative of 6 to 1, as indicated in Figure 8 which depicts the distribution of polarity across different times from the product release date. Neutral tweets were not considered on the aggregate level due to their limited information regarding consumers' satisfaction. It is evident that the consumers' polarity shifted from neutral to positive after the initial product release with the sentiment weakening after 12 months and slightly shifting towards the negative scales along the way. Subjectivity scores for the periods before and after the release also differ. Consumers used less opinionated content in their eWOM before product release, possibly due to the lack of experience with it, while after the product was introduced into the market there was an increase in subjectivity, indicating that consumers shaped an opinion about the product. This is more evident in the interval 6-12 months

[9] https://www.nltk.org/

[10] http://alt.qcri.org/semeval2017/task4/

after the release, while in the interval 12-18 months the opinions of consumers start to decline, as illustrated in Figure 8. In the same vein, Figure 9 illustrates the rate of information diffusion that defines the process of dissemination of product information from user to user. The figure uses unique user IDs (usernames) to calculate the frequency of eWOM at different time slots after product release. The information diffusion rate shows a peak after the first 6 months that deteriorates significantly after 12 months.

To validate the results of sentiment analysis, the sentiment viz tool[11] was used to obtain the polarity of consumers regarding the products' reputation for sub-periods of the analysis (Figure 10). Tweets are shown by circles positioned by their sentiment. Negative tweets are drawn as blue circles on the left, and positive tweets as green circles on the right of the x-axis. The figure illustrates the overall sentiment for the "HuaweiP20Pro" keyword. This represents the distribution of tweets across the sentiment scale. The tweets lean towards the positive side of the spectrum, indicating that consumers perceived positively the new product and its differentiating features. Similarly, the "camera" keyword shows a positive sentiment as also seen in Figure 6. Both observations confirm the results of our sentiment analysis.

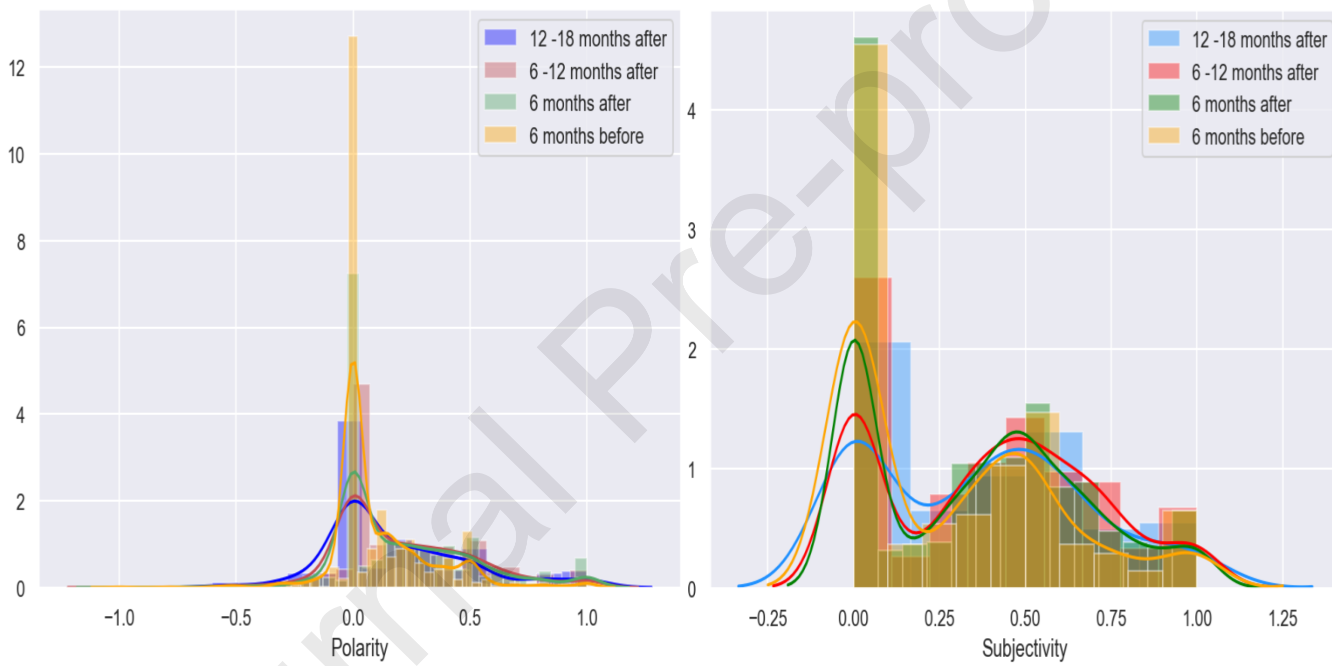


Fig.8. Distribution of sentiment polarity (left) and subjectivity (right) for the four periods under study

[11] https://www.csc2.ncsu.edu/faculty/healey/tweet_viz/tweet_app/

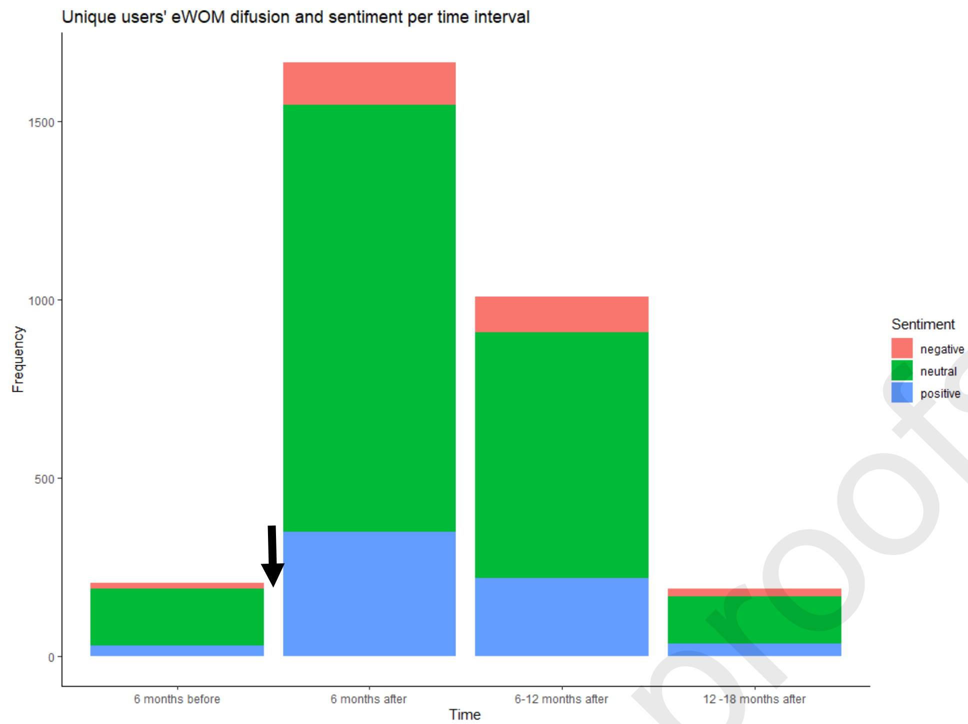
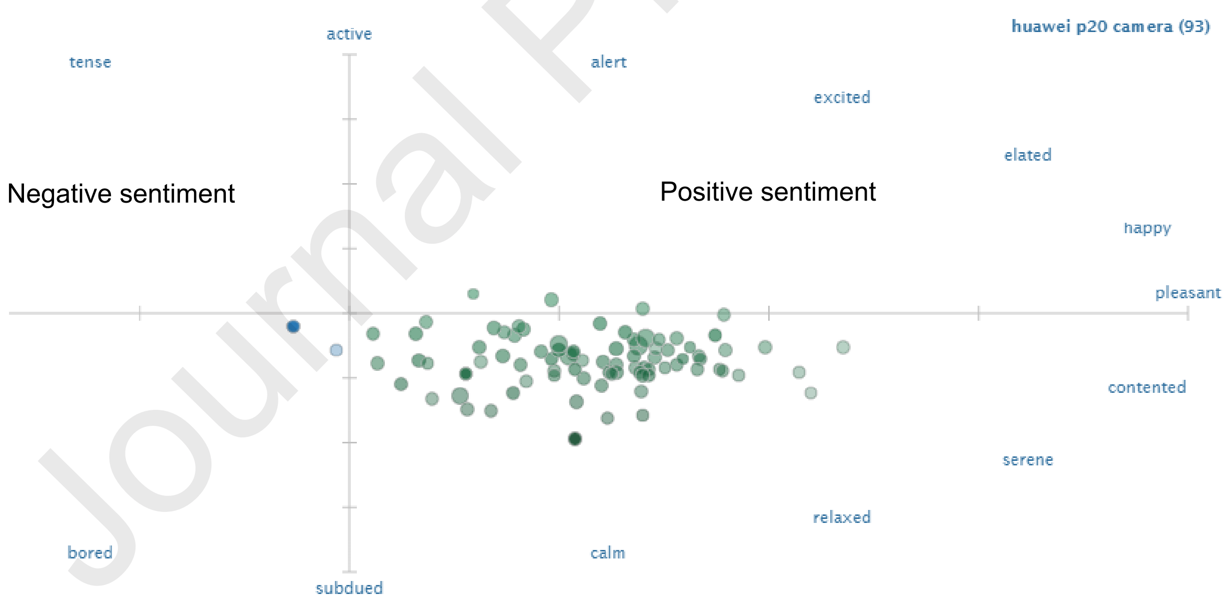



Fig.9. Information diffusion against consumers' sentiment and time of product release, indicated with an arrow



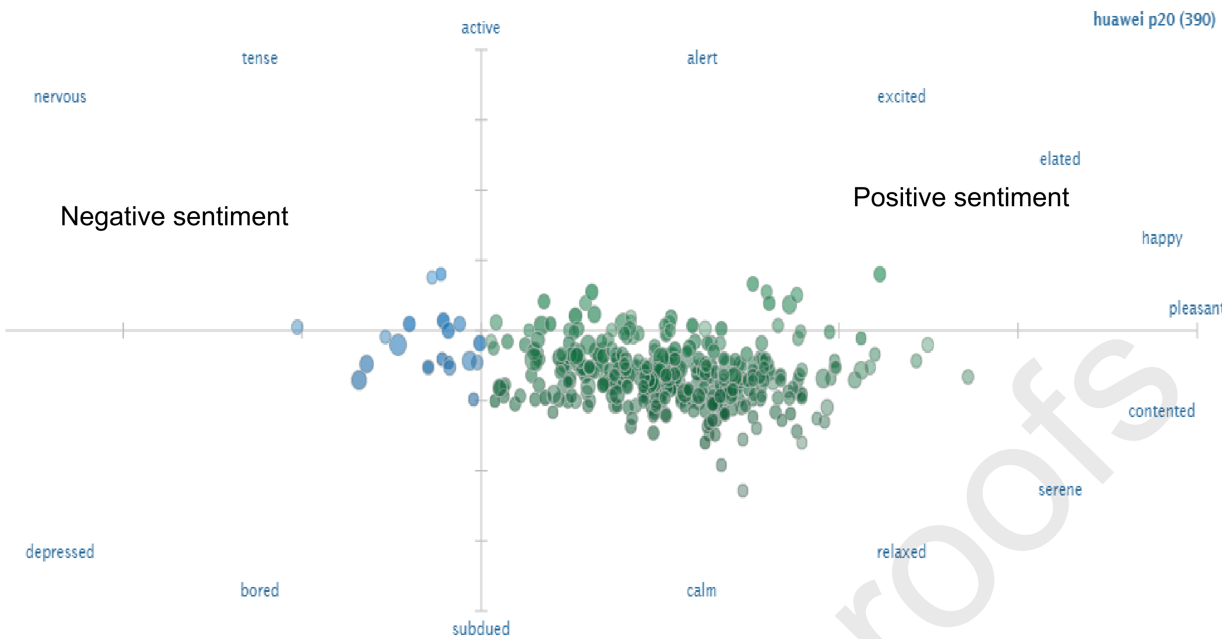



Fig.10. Sentiment analysis for the keywords: “HuaweiP20Pro” (bottom) and “HuaweiP20Pro Camera” (top) visualized with sentiment Viz tool

### 5.3 Longitudinal and diagnostic analysis

So far, we have presented descriptive analysis and visualized trends regarding consumers’ perceptions about the product. To evaluate the persistence of the consumers’ perceptions about the products and the reputation of the firm, it was important to perform a longitudinal analysis_based on tweets collected before and after product release. Therefore, all tweets collected throughout this period were categorized into four groups based on their timing, as also shown with previous descriptive analysis: 6 months before release, up to 6 months after release, 6-12 months after release and 12-18 months after release. The tweets were initially evaluated against their topics and sentiment and subsequently time stamped and saved in a dataset with information regarding the date of their publication. The sentiment of each tweet was estimated by the sentiment analysis algorithm, and was scored as a continuous number in the range [-1,1]. This was subsequently re-coded into positive, neutral and negative sentiment using threshold values of >0.4 for positive, less than zero for negative and between values [0-0.4] for neutral polarity state. Two analyses were performed utilizing two datasets. The first analysis examined the effect of “time-passed” on consumers’ sentiment and the second diagnosed causes of customers’ dissatisfaction/satisfaction with the new features of the product. For the latter, the topics of interest were those referring to the product’s photographic capabilities (topics 5 and 8 in Figure 6) since these were the differentiating features in the product-positioning campaign. The research question answered by this analysis was whether consumers’ expectations with regards to the product’s differentiating feature before its release were met after its release. The assumption here is that if consumers’ expectations were met, the sentiment of their eWOM regarding the differentiating product features should have been similar before and after the release. The dataset in this case included only data referring to super-topics linked to topics 5, 7 and 8, while for the first analysis no filtering was performed. The new datasets were analyzed in Rstudio to identify statistical associations between the customers’ sentiment and time-passed after product release. An ordinal logistic regression was used for this purpose (significant effect $p<0.005$) with the tweets’ sentiment coded to “neg”, “neu”, “pos”. The dependent (outcome) variable was the consumers' sentiment and the explanatory (independent) variable was the timing with state "6 month before" used as the reference category.

The estimated ordered logit model investigated whether time-passed after the first product release can predict the sentiment of consumers, in other words it explains if time has any impact on the liking of the product. Ordinal logistic regression, also known as proportional odds model, is one of the most popular methods in data analytics for problems of this type. For the first analysis, the estimated model showed that the predictor accounted for a significant amount of variance in the outcome (likelihood ratio chi square (3) = 44.5, $p < 0.001$). From all timing categories, "6 months" and "6-12 months after" were significant in predicting the outcome (sentiment) as compared to the reference category "6 months before" ($b = 0.53$, $SE=0.07$ $p <0.001$, and $b= 0.45$ $SE=0.08$ $p < 0.001$, respectively).

The ordinal logistic regression model results shown in Figure 11 are depicted in terms of the effect of the independent variable (time) on consumers' sentiment and the predicted probabilities from the logistic model. This plot shows a change in positive sentiment, indicating that consumers were more inclined to evaluate the product highly six months after initial release in contrast to 12-18 months after. It is also evident that the probability of sentiment being neutral increases after 12 months. This indicates an initial decline in consumers' sentiment regarding the product after a period of 12 months.

The model also shows a positive effect of time on sentiment degradation with significant effects for 6-months and 6-12 months and 12-18 months after product release. So, for 6 up to 12 months after product release, there is a 0.45 increase in the odds of sentiment being negative instead of neutral, which indicates the negative effect of time on product sentiment. This is attributed to issues that might emerge with regards to product features, or the effect from the introduction of new smartphones in the market, and discussions around them denoted by topic 2 (product rivals).

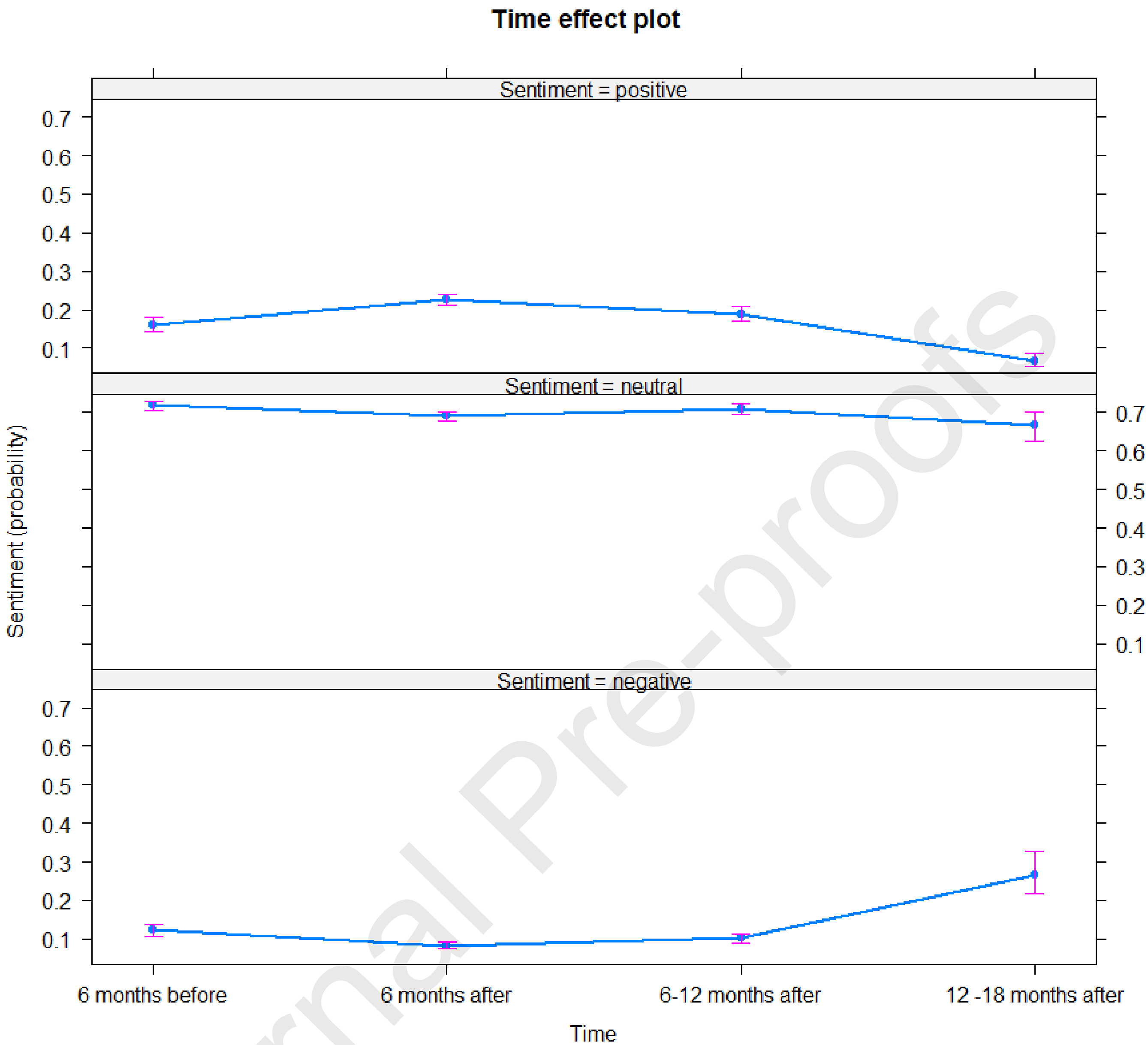



Fig**.** 11. Predicted probabilities of consumers' sentiment in relation to months after initial product release, from ordinal logistic regression. Negative sentiment initially decreases marginally and increases from 6 to 12 months after. Positive sentiment decreases 12 months following product release, after an initial increase in the first 6 months.

Diagnostic drill-down analysis (Figure 12) of sentiment using specific topics over time enables the identification of the cause of potential dissatisfaction or satisfaction with product features or qualities. Therefore, given the topics identified from the dataset (Figure 6), the analyst can evaluate the sentiment of each topic in relation to time after release. The results from the temporal topic sentiment analysis for the theme relating to photography (topic 8) were positive, neutral and finally positive, six months, 6-12 months, and 12-18 months after product release respectively, indicating that consumers were more inclined to evaluate the product's photographic feature highly, six months and 12-18 months after initial release, in contrast to 6-12-months after, where the sentiment was similar to pre-product release. This suggests that consumers' initial expectations were met with regards to photography, and engagement with the product's features was positively associated with satisfaction. This provides additional evidence that the product's positioning was successful, with the new photographic capabilities matching consumers' needs and expectations.

Subsequent topic sentiment analysis for themes relating to "price", "battery" and "night mode" features (topics 1, 3, 5 respectively in Figure 6) revealed that the consumers' perceptions were not significantly different after product release for each theme except for the topic "price", indicating that the product positioning was satisfactory for the "battery" and "night mode" features while highlighting an area of concern with regards to price. This could be attributed to the product price not being competitive after product release, with gradually improving sentiment as the price drops.

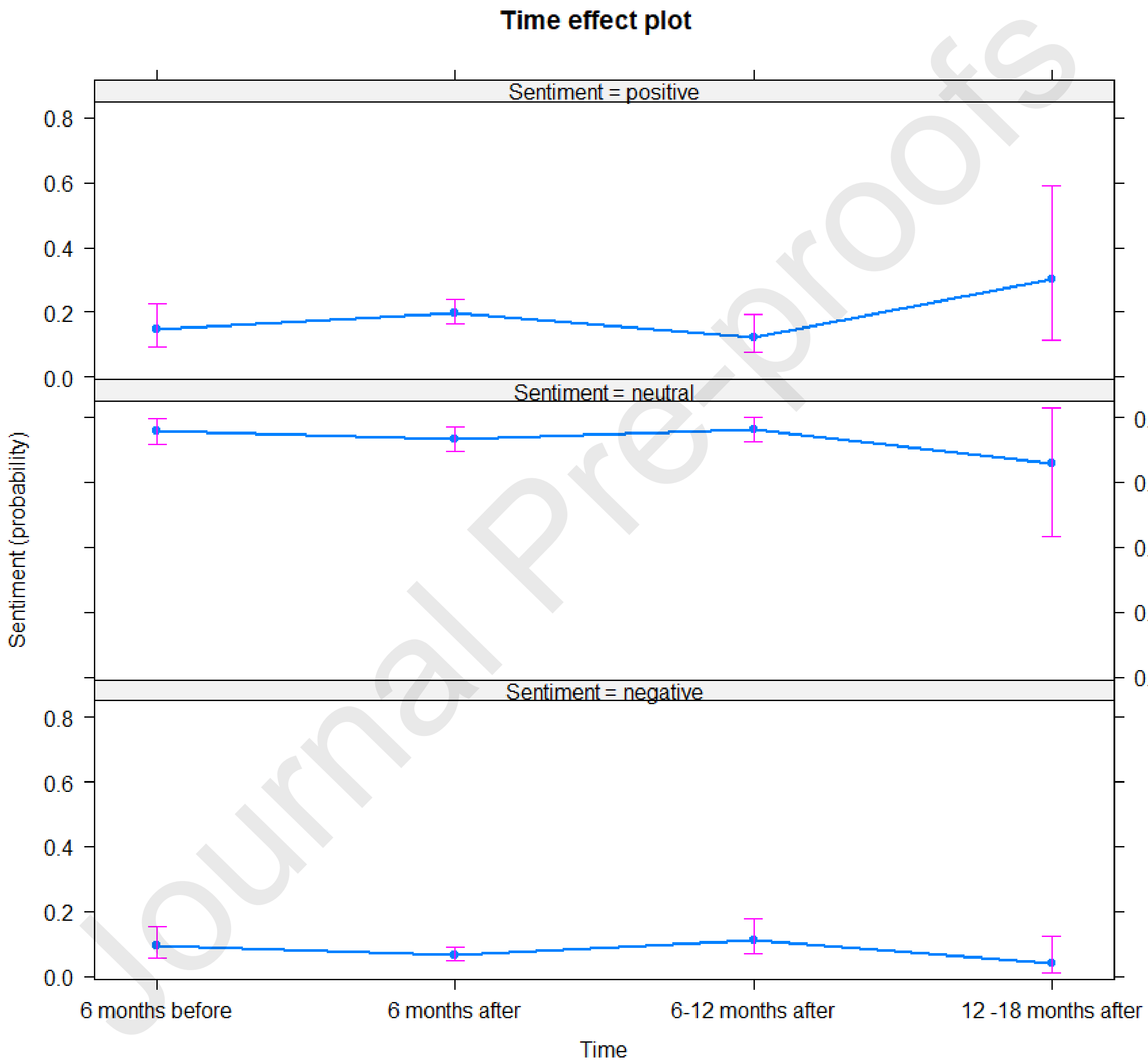



Fig. 12. Predicted probabilities of consumers' sentiment in relation to topic "photography" from ordinal logistic regression. The sentiment of consumers before and after product release is similar. The overall negative sentiment of consumers decreases marginally and increases from 6 to 12 months after.

### 5.4 Artefact evaluation

To further examine the validity of the proposed evaluation method (the artefact), online keyword search data were collected and analyzed to assess the level of similarity between keyword trend and sentiment, given evidence that sentiment is linked to product success (Nguyen and Chaudhuri, 2019), while online search traffic is the most widely used source of information for sales forecasting and market analysis (Schaer et al., 2019). Google Trends is a tool that enables the collection of historical keyword searches relative to the overall number of searches during a specific time interval. Past studies showed that Google Trends data can explain and forecast certain trends with regards to consumers' interest in a product, marketing activities, purchase behaviours, adoption of products and product performance (Nguyen and Chaudhuri, 2019). Choi and Varian (2012) demonstrated the application of Google Trends to forecast near-term retail sales of a wide range of products; Carrière-Swallow and Labbé (2013) further examined the case of automobile sales in an emerging market. Similar work also showed that Google Trends is a better indicator of consumers' consumption compared to surveys (Vosen and Schmidt, 2011) while Chumnumpan and Shi (2019) showed that the combination of diffusion models and Google Trends can explain fluctuations in market growth better than conventional diffusion models.

We thus, compared results from the longitudinal sentiment analysis with consumers' online search behaviour using Google Trends. The assumption is that since Google Trends results could be linked to consumers' adoption, this is an indication of successful product positioning. Therefore, results from the longitudinal analysis should reflect this pattern. However, Google Trends alone cannot provide justification for the cause of change in consumers' trend nor for its statistical significance in contrast to our method. To that end, data was collected for the critical pre and post product release period from January 2018 to April 2019 regarding the Huawei P20 product series. Several keywords were used for data collection including "Huawei p20", "P20pro", "P20+Leica", "P20 Leica", "P20", "P20 pro" and variations of these to reflect the differentiating features of the product as articulated in the company's marketing slogan. The same keywords were used for all target market countries. Subsequently, the Twitter sentiment data for the same period underwent analysis, to collate average scores of sentiment per month for the same period. To enable the comparison between keyword search and average sentiment per month, normalization of the scores of each time series was performed prior to visualizing the results shown in Figure 13, indicating significant correlation between the two ($r=0.855, p<0.00, n=16$).

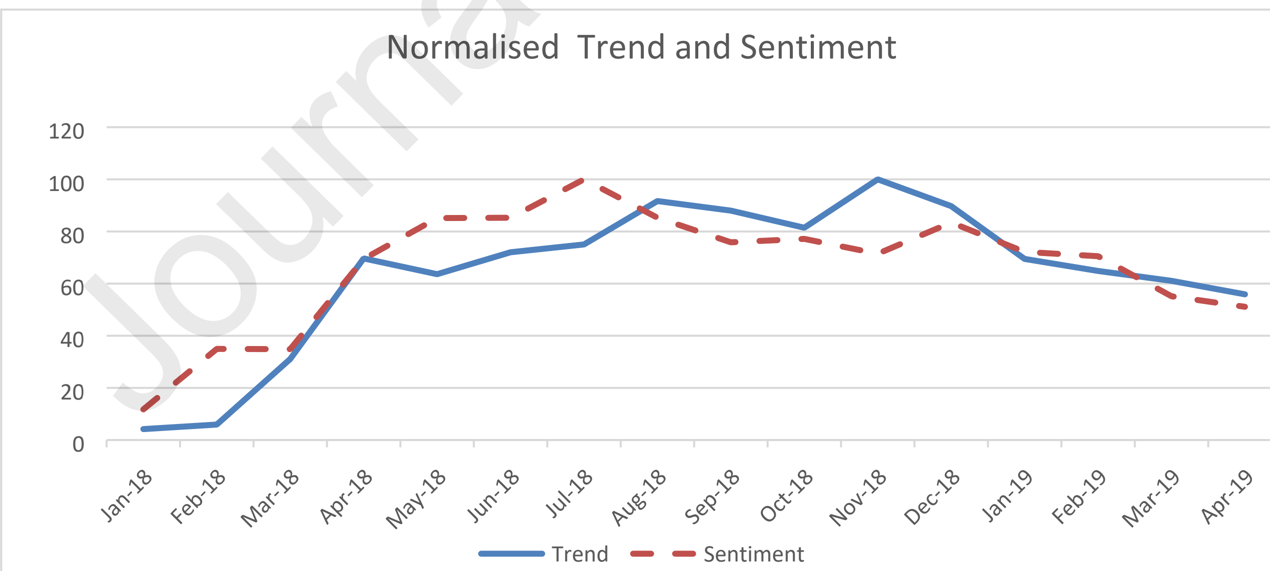



Fig.13: P20 related keywords search trend against proposed method's sentiment evaluation

## 6. Discussion

### Key findings

The study showcased a new method which integrates techniques from data analytics and statistics into the timely domain of textual big data, and in particular tweets, for evaluating new products' positioning strategy. Firms already use eWOM analytics to gain insights of customer opinions and sentiments regarding their products and to consequently inform their product positioning (Chandrasekar, 2011; Zeng et al., 2010). Existing product positioning methods mainly use data from online product reviews (Wang et al., 2018) hence are ex-post, or atemporal (Stieglitz and Dang-Xuan, 2013) relying heavily on summary counts from a predefined list of words, product features (Aggarwal et al., 2009), opinions (Moon and Kamakura, 2017; Tirunillai and Tellis, 2014), or ontologies specified manually or inferred from data (Moon and Kamakura, 2017). There is very limited research on methods to evaluate new products-positioning strategy (before/after product release) and diagnose product dissatisfaction due to mispositioning using tweets.

The proposed method contributes to filling this gap using contemporary techniques in topic modelling and sentiment assessment in combination with inferential statistics. Unlike existing methods on eWOM product positioning it uses tweets rather than product reviews given evidence (Deshpande and Rokne, 2018) that consumers provide feedback of their experience with a product firstly on twitter and then at review websites, which also affects consumers early product adoption behavior (Hennig-Thurau et al., 2015). Substantially, the method enables the analysis of perceived and actual product value and benefit to consumers before its release and after when consumers start to use it, and provides the means to diagnose product features causing dissatisfaction. The case study implemented the method to validate the product-positioning strategy of Huawei and demonstrated that the firm's strategy was successful and had a positive impact on its target users. Thus, consumers' expectations with regards to product's differentiating feature before its release (planned positioning) were met after its release (actual positioning). The case study constitutes an instantiation of the method (artefact) and contributes towards its design science evaluation through a comparative analysis between results from the proposed method and results from secondary online consumer keyword search patterns from Google Trends. This comparison highlighted similarities between estimated adoption by consumers and product success trends, due to satisfactory product positioning; this way it supported the validity of the method, and contributed to its evaluation.

The first step of the method is firm strategy analysis, to identify target consumers, the product's differentiating features (linked to perceived consumers' needs) and the marketing promotion slogan using explicit messages highlighting the product's unique selling propositions. The firm's product-positioning plan in this case was to differentiate on price and technological features such as the use of the triple-camera for enhanced photography. Our results show that consumers embraced the new triple-camera feature and referred to it in their tweets with positive polarity. The new camera triggered people's interest as identified by the longitudinal topic analysis, with consumers discussing this feature in a positive way especially in the first 6 and between 6-12 months after the product release. Similarly, longitudinal sentiment analysis of consumers' perceptions from eWOM one year after the product release showed signs of maturity, with sentiments converging towards neutral polarity. This timeframe could be an indication of the maturity stage in the product's lifecycle, as eWOM is linked with consumer behaviour and product sales (Fan et al., 2017).

Topic modelling enabled the identification of the main themes discussed by consumers throughout the study's test-period, and helped to visualize topics across time. Clusters of topics expressed as super-topics across time reduced the dimensionality of the data and helped simplify the visualizations regarding consumers perceptions on an aggregate level. The identified super-topics addressed aspects pertaining to products' photography capabilities, video, battery and looks, and competing products of other brands. Sentiment analysis helped to evaluate topics valence and provided another dimension to product positioning evaluation. An aggregate eWOM data analysis, revealed a positive consumer's sentiment after the initial product release, with a weakening effect after 12 months of release. Subjectivity scores before product release showed less opinionated content with an increase in subjectivity after product release, when consumers shaped an opinion about the product after using it. The rate of product information diffusion from user to user shows a peak in the first 6 months and subsequently deteriorates significantly after 12 months that shows the buzz produced from the new product which is another indication of a successful product positioning.

Additional analysis of the above effects using inferential statistics enabled the identification of links between consumer dissatisfaction/satisfaction using associated sentiment at different time periods. Application of ordinal logistic regression showed that 12 months following the product release consumers' sentiment has significantly decreased (towards neutral). This could indicate a product saturation point when firms consider revising their pricing strategy. Firm's price change is verified by market price data 12 months after product release from pricespy.co.uk that validates the predicted point of product saturation.

Diagnostic analysis of product features combining topics filtering with sentiment polarity against time showed that product's photographic capabilities satisfied consumers' initial expectations, who became more satisfied with more engagement with this product features. This also verifies that the product's positioning was successful, since the product's photographic capabilities was the differentiating feature in the firm's planned positioning which matched consumers' needs and expectations after release (actual positioning). Additional analyses regarding the product's "price", "battery" and "night mode" features revealed that, the price of the product received negative comments after 6-12 months of its release, possibly attributed to the product-price not being competitive at that period. The steep drop of the product's price after 12 months (as indicated by pricespy.co.uk) matches the estimated effect of feature (price) on satisfaction which improves after 12 months and aligns with price reduction.

**Theoretical and practical implications**

This work makes a number of theoretical and practical contributions. Firstly, the substantive application of this method provides valuable insight of the success or failure of a firm's product positioning strategy and guides on what features of a product cause dissatisfaction and need attention in subsequent releases. Secondly, the study showed that Twitter is a valuable tool for analyzing and predicting consumer behaviour with regards to the positioning of new products in contrast to product review data analysis, since there are no online reviews posted by consumers for products that are not released yet. Thirdly, it fills a gap in previous work on product positioning which has been relying on secondary eWOM data suffering from limited generalizability due to the low data quality and quantity. This work provided the tools for primary data collection using a tailor-made application and scripts to gather and analyse data for specific products of interest using data filtering criteria, making the method more generic and addressing the data sparsity problem of secondary sources. Therefore, collected data requires less pre-processing and minimizes the risk of bias from irrelevant content. Fourthly, the use of topic and sentiment analysis, combined with inferential statistics, enabled the longitudinal assessment of online consumers' opinions and the diagnosis of product's mispositioning based on statistical evidence. This allows the identification of discrepancies in planned and actual product positioning and the diagnosis of mispositioning's causes using filtered topic analysis. Finally, the proposed evaluation method is based on the design science framework to ensure that the final artefact produces results that are valid and have a significant impact on the problem domain.

**Management Implications**

This study also has valuable implications to managerial practice. Traditional approaches to product positioning, such as positioning maps, define in advance the features that express the dimensions of a product in a 2D map, and then ask a sample of consumers to rate competing brands on these product features. This, however, constrains consumers' viewpoints within the boundaries of the specified feature scales creating researcher/manager bias. The proposed method is grounded on the big data paradigm and hence is bottom up. It, thus enables managers to observe how new products positioning performs over time against topics emerging from data rather than being specified a-priori. This could assist in the identification of issues not foreseen ahead of time and could inform the specification of new product's revisions through the identification of new user requirements.The method support managers identify what could causes a product's mispositioning by evaluating the distribution of identified topics and their sentiment against time, through longitudinal analysis of tweets

**Limitations and further work**

Limitations of this work lie in the use of publicly available data from Twitter users, which constitutes a subset of all eWOM activity. This may have implications with regards to the generalizability of the findings, depending on the volume of the remaining non-accessible data and how representative of the relevant consumer populations are the users who post their opinions. However, this work focuses on a novel method for product-position evaluation that can be implemented with any dataset. Another limitation is the use of mainstream algorithms for sentiment and topic analysis. Further studies could investigate alternative techniques that might improve the accuracy of sentiment analysis and topic classification. However, as the aim of this study was to evaluate the overall trend in sentiment fluctuation and topics discussed online, and to evaluate the effectiveness of eWOM analytics in validating product positioning, the use of these specific algorithms was sufficient for this study.

Lessons learned from this study point to the importance of optimizing the underlying topic model used to produce coherent topics that are interpretable and describe well the main themes discussed by consumers. Tweets cleansing is a necessary precondition for this process, since they are contaminated with irrelevant third-party advertising content. This can influence the results and needs to be identified and filtered out. Moreover, when investigating specific products comparatively, few valuable tweets can be obtained, and hence the study needs to span over a long period of time, making it more challenging. For future research work, the authors aim to improve further the validity of the method by comparing its topic analysis performance against two other topic modelling techniques that also model how topics prevail across time, the dynamic LDA and structural topic modelling using metadata (time). Additionally, gender is considered an interesting dimension worth exploring further when considering sentiments and online eWOM analysis, so further studies should address this issue as well as other relevant demographics. Finally, we aim to expand the methodology to analyse consumers' emotions with product's features to refine target consumer requirements and inform the product's redesign.

## 7. Conclusions

The cost of product mispositioning is high since it jeopardizes the strategy of a company and could influence sales performance. Evaluating positioning strategies early in the product release cycle constitutes a vital process for effective sales performance.

This work contributes to resolving this problem by comparing firm's actual against planned positioning strategy using eWOM analytics. The method assists management in identifying components of the positioning strategy that are effective or ineffective and to identify the product's saturation point when firms need to consider revising their pricing. A case study is presented that applies the method to evaluate the positioning strategy of Huawei's new P20 products. The method combines machine learning and statistical modelling techniques to identify key themes discussed online by consumers regarding the product of interest, and to assess their sentiment before and after the product release. Longitudinal analysis through original logistic regression and topic filtering enables the evaluation of the effect of time on consumer sentiment and assists in diagnosing causes of dissatisfaction over time. The visualization of the results improves comprehension and communication of the findings with management and other relevant stakeholders for more efficient and effective decision making.

## Appendices

### Appendix 1: The eight components of design science theory as implemented in this study

The list below illustrates how the eight components of design science theory introduced by Gregor and Jones (2007) are used for the design and evaluation of the artefact(method for evaluating new product positioning).

1. Purpose and scope of the artefact in assisting marketing managers to evaluate their new product-positioning strategy;
2. Constructs are the theories and techniques for evaluating eWOM content using theories in NLP such as sentiment and topic analysis;
3. Principles of form and function include the architecture of the proposed evaluation method that utilizes eWOM data collection through the Twitter API, and sentiment analysis using Naïve Bayes.
4. Artefact mutability is linked to the generalizability of the method since it can be adapted to online content from various markets;
5. Testable propositions draw on the assessment of a prospective positioning strategy against evidence of its implementation from eWOM and online keyword searches;
6. Justification knowledge draws on theories of our design such as LDA topic analysis and sentiment analysis methods;
7. Principles of implementation describe the implementation of the artefact using online reviews from Twitter in the context of the smartphone industry;
8. Expository instantiation is realized through a realistic case study of a new smartphone product release.

### Appendix 2: Topics analysis approach

The main steps followed to discover topics in the dataset:

1. Cleaning the data.
2. Creating a Term Document Matrix (TDM).
3. Calculating the Term Frequency Inverse Document Frequency for all the words in TDM.
4. Excluding words with TD-IDF<= 0.1, less frequent.
5. Calculating the optimal number of topics (K) in the corpus using topic coherence and likelihood estimates
6. Applying LDA method using the optimum K to train an LDA classification model
7. Assigning topics to each tweet using the trained LDA model
8. Visualizing distribution of topics per product lifetime (timeslots)

## References

Abellán, J., Castellano, J. G., 2017. Improving the Naive Bayes classifier via a quick variable selection method using maximum of entropy. Entropy, 19(6), 247. https://doi.org/10.3390/e19060247

Aggarwal, P., Vaidyanathan, R., Venkatesh, A., 2009. Using Lexical Semantic Analysis to Derive Online Brand Positions: An Application to Retail Marketing Research. Journal of Retailing, 85(2), 145–158. https://doi.org/10.1016/j.jretai.2009.03.001

Al-Obeidat, F., Spencer, B., Kafeza, E., 2018. The Opinion Management Framework: Identifying and addressing customer concerns extracted from online product reviews. Electronic Commerce Research and Applications, 27, 52–64. https://doi.org/10.1016/j.elerap.2017.11.003

Albers, S. , 1979. Extended algorithm for optimal product positioning. European Journal of Operational Research, 3(3), 222–231.

Aue, A., Gamon, M., 2005. Customizing Sentiment Classifiers to New Domains: A Case Study. Recent Advances in Natural Language Processing (RANLP), 207–218.

Babić Rosario, A., Sotgiu, F., De Valck, K., Bijmolt, T. H. A., 2016. The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. Journal of Marketing Research, 53(3), 297–318. https://doi.org/10.1509/jmr.14.0380

Berger, J., Milkman, K. L., 2012. What Makes Online Content Viral? Journal of Marketing Research, 49(2), 192–205. https://doi.org/10.1509/jmr.10.0353

Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. Latent dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.

Carrière-Swallow, Y., Labbé, F., 2013. Nowcasting with Google trends in an emerging market. Journal of Forecasting. https://doi.org/10.1002/for.1252

Chamlertwat, W., Bhattarakosol, P., Rungkasiri, T., Haruechaiyasak, C., 2012. Discovering Consumer Insight from Twitter via Sentiment Analysis. Journal of Universal Computer Science, 18(8), 973–992.

Chen, K., Kou, G., Shang, J., Chen, Y., 2015. Visualizing market structure through online product reviews: Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches. Electronic Commerce Research and Applications, 14(1), 58–74. https://doi.org/10.1016/j.elerap.2014.11.004

Chumnumpan, P., Shi, X., 2019. Understanding new products' market performance using Google Trends. Australasian Marketing Journal, 27(2), 91–103. https://doi.org/10.1016/j.ausmj.2019.01.001

D'Andrea, E., Ducange, P., Lazzerini, B., Marcelloni, F., 2015. Real-Time Detection of Traffic from Twitter Stream Analysis. IEEE Transactions on Intelligent Transportation Systems, 16(4), 2269–2283. https://doi.org/10.1109/TITS.2015.2404431

Dai, X., Bikdash, M., Meyer, B., 2017. From social media to public health surveillance: Word embedding based clustering method for twitter classification. SoutheastCon17. https://doi: 10.1109/SECON.2017.7925400

Davis, F. D., 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319. https://doi.org/10.2307/249008

de Chernatony, L , 2001. Succeeding with brands on the Internet. Brand Management, 8(3), 186–195. https://doi.org/10.1057/palgrave.bm.2540019

Deshpande, G., Rokne, J., 2018. User Feedback from Tweets vs App Store Reviews: An Exploratory Study of Frequency, Timing and Content. Proceedings - 2018 5th International Workshop on Artificial Intelligence for Requirements Engineering, AIRE 2018. https://doi.org/10.1109/AIRE.2018.00008

Deveaud, R., SanJuan, E., Bellot, P., 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. Document Numérique, 17(1), 61–84. https://doi.org/10.3166/dn.17.1.61-84

Etter, M., Ravasi, D., Colleoni, E., 2019. Social Media and the Formation of Organizational Reputation. Academy of Management Review, 44(1), 28–52. https://doi.org/10.5465/amr.2014.0280

Fan, Z. P., Che, Y. J., Chen, Z. Y., 2017. Product sales forecasting using online reviews and historical sales data:

A method combining the Bass model and sentiment analysis. Journal of Business Research, 74, 90–100. https://doi.org/10.1016/j.jbusres.2017.01.010

Ferreira, A., 2011. Android OS changes smartphone life cycle. The Vista.

Franke, N., Piller, F. T., 2003. Key research issues in user interaction with user toolkits in a mass customisation system. International Journal of Technology Management, 26(5–6), 578–599. https://doi.org/10.1504/ijtm.2003.003424

Gambhir, M., Gupta, V., 2017. Recent automatic text summarization techniques: a survey. Artificial Intelligence Review, 47(1), 1–66. https://doi.org/10.1007/s10462-016-9475-9

Gamon, M., Aue, A., Corston-Oliver, S., Ringger, E., 2005. Pulse: Mining Customer Opinions from Free Text. Lecture Notes in Computer Science 3646, 121–132. https://doi.org/10.1007/11552253_12

Gatti, L., Guerini, M., Turchi, M., 2016. SentiWords: Deriving a High Precision and High Coverage Lexicon for Sentiment Analysis. IEEE Transactions on Affective Computing, 7, 409–421. https://doi.org/10.1109/TAFFC.2015.2476456

Gavish, B., Horsky, D., Srikanth, K., 1983. An Approach to the Optimal Positioning of a New Product. Management Science, 29(11), 1277–1297. https://doi.org/10.1287/mnsc.29.11.1277

Gregoriades, A., Sutcliffe, A., 2018. Simulation-based evaluation of an in-vehicle smart situation awareness enhancement system. Ergonomics, 61(7), 947–965. https://doi.org/10.1080/00140139.2018.1427803

Griffiths, T. L., Steyvers, M., 2004. Finding scientific topics. Proceedings of the National Academy of Sciences, 101(1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Gruca, T. S., Klemz, B. R., 2003. Optimal new product positioning: A genetic algorithm approach. European Journal of Operational Research, 146(3), 621–633. https://doi.org/10.1016/S0377-2217(02)00349-1

Hadjinicola, G. C., 1999. Product Positioning and Pricing Under Production Cost Considerations. Decision Sciences, 30(3), 849–864. https://doi.org/10.1111/j.1540-5915.1999.tb00909.x

Hennig-Thurau, T., Wiertz, C., Feldhaus, F., 2015. Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. Journal of the Academy of Marketing Science, 43(3), 375–394. https://doi.org/10.1007/s11747-014-0388-3

Hevner,A.R., March, S.T., Park, J., Ram.,S., 2004. Design Science in Information Systems Research. MIS Quarterly, 28(1), 75. https://doi.org/10.2307/25148625

Hooley, G., Greenley, G., Fahy, J., Cadogan, J., 2001. Market-focused Resources, Competitive Positioning and Firm Performance. Journal of Marketing Management, 17(5–6), 503–520. https://doi.org/10.1362/026725701323366908

Hsieh, C. H., Chen, S. H., 1999. Model and algorithm of fuzzy product positioning. Information Sciences, 121, 61–82. https://doi.org/10.1016/S0020-0255(99)00050-X

Hutto, C. J., Gilbert, E., 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. Eighth International AAAI Conference on Weblogs and Social Media (ICWSM), 216–225.

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. https://doi.org/10.1002/asi.21149

Jones, D., Gregor, S., 2007. The Anatomy of a Design Theory. Journal of the Association for Information Systems, 8(5), 312–335. https://doi.org/10.17705/1jais.00129

Jongeling, R., Datta, S., Serebrenik, A., 2015. Choosing your weapons: On sentiment analysis tools for software engineering research. 2015 IEEE 31st International Conference on Software Maintenance and Evolution, ICSME 2015 - Proceedings. https://doi.org/10.1109/ICSM.2015.7332508

Jung, J. J., 2008. Taxonomy alignment for interoperability between heterogeneous virtual organizations. Expert Systems with Applications, 34(4), 2721–2731. https://doi.org/10.1016/j.eswa.2007.05.015

Kageb, M., Salomonsson, H., 2016. Word Sense Disambiguation using a Bidirectional LSTM. Workshop on

Cognitive Aspects of the Lexicon at International Conference on Computational Linguistics.

Khade, A. A., 2016. Performing Customer Behavior Analysis using Big Data Analytics. Procedia Computer Science, 79, 986–992. https://doi.org/10.1016/j.procs.2016.03.125

Kim, Y., Street, W. N., 2004. An intelligent system for customer targeting: a data mining approach. Decision Support Systems, 37(2), 215–228. https://doi.org/10.1016/S0167-9236(03)00008-3

Kohli, R., Sukumar, R., 1990. Heuristics for Product-Line Design Using Conjoint Analysis. Management Science, 36(12), 1464–1478. https://doi.org/10.1287/mnsc.36.12.1464

Korenčić, D., Ristov, S., Šnajder, J., 2018. Document-based topic coherence measures for news media text. Expert Systems with Applications, 114, 357–373. https://doi.org/10.1016/j.eswa.2018.07.063

Kotler, P., Armstrong, G., 2018. Principles of Marketing. Pearson.

Krouska, A., Troussas, C., Virvou, M., 2017. Comparative evaluation of algorithms for sentiment analysis over social networking services. Journal of Universal Computer Science, 23, 755–768.

Ku, L., Liang, Y., Chen, H., Lun-Wei, K., Yu-Ting, L., Hsin-Hsi, C., 2006. Opinion Extraction, Summarization and Tracking in News and Blog Corpora. American Association on Artificial Intelligence, 100–107. https://doi.org/citeulike-article-id:2913694

Kwong, C. K., Luo, X. G., Tang, J. F., 2011. A methodology for optimal product positioning with engineering constraints consideration. International Journal of Production Economics, 132(1), 93–100. https://doi.org/10.1016/j.ijpe.2011.03.014

Le, Q., Mikolov, T., 2014. Distributed representations of sentences and documents. 31st International Conference on International Conference on Machine Learning, 1188–1196.

Lee, A., Yang, F.-C., Chen, C. H., Wang, C.-S., Sun, C.-Y., 2016. Mining perceptual maps from consumer reviews. Decision Support Systems, 82, 12–25. https://doi.org/10.1016/j.dss.2015.11.002

Lee, H.-C., Rim, H.-C. Lee, D. G., 2019. Learning to rank products based on online product reviews using a hierarchical deep neural network. Electronic Commerce Research and Applications, 36, 100874. https://doi.org/10.1016/j.elerap.2019.100874

Lee, S., Kim, W., 2017. Sentiment labeling for extending initial labeled data to improve semi-supervised sentiment classification. Electronic Commerce Research and Applications, 26, 35–49. https://doi.org/10.1016/j.elerap.2017.09.006

Lee, T. Y., Bradlow, E. T., 2011. Automated Marketing Research Using Online Customer Reviews. Journal of Marketing Research, 48(5), 881–894. https://doi.org/10.1509/jmkr.48.5.881

Lei, N., Moon, S. K., 2015. A Decision Support System for market-driven product positioning and design. Decision Support Systems, 69, 82–91. https://doi.org/10.1016/j.dss.2014.11.010

Lu, B., Tsou, B. K., 2010. Combining a large sentiment lexicon and machine learning for subjectivity classification. International Conference on Machine Learning and Cybernetics (ICMLC), 3311–3316. https://doi.org/10.1109/ICMLC.2010.5580672

Martin-Domingo, L., Martín, J. C., Mandsberg, G., 2019. Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). Journal of Air Transport Management, 78, 106–115. https://doi.org/10.1016/j.jairtraman.2019.01.004

Moon, S. K., Simpson, T. W., Kumara, S. R. T., 2010. A methodology for knowledge discovery to support product family design. Annals of Operations Research, 174(1), 201–218. https://doi.org/10.1007/s10479-008-0349-7

Moon, S., Kamakura, W., 2017. A picture is worth a thousand words: Translating product reviews into a product positioning map. International Journal of Research in Marketing, 34(1), 265–285. https://doi.org/10.1016/j.ijresmar.2016.05.007

Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., Wilson, T., 2013. Task 2: Sentiment analysis in twitter. Proceedings of the Seventh International Workshop on Semantic Evaluation, SemEval'13.

Nayab, G., Bilal, M., Shrafat, A. S., 2016. A brand is no longer what we tell the customer it is - it is what customers tell each other it is: Validation from Lahore. Science International (Lahore), 28(3), 2725–2729.

Netzer, O., Feldman, R., Goldenberg, J., Fresko, M., 2012. Mine Your Own Business: Market-Structure Surveillance Through Text Mining. Marketing Science, 31(3), 521–543. https://doi.org/10.1287/mksc.1120.0713

Ng, A., 2019. Smartphone users are waiting longer before upgrading - here's why. CNBC. https://www.cnbc.com/2019/05/17/smartphone-users-are-waiting-longer-before-upgrading-heres-why.html

Nguyen, H. T., Chaudhuri, M., 2019. Making new products go viral and succeed. International Journal of Research in Marketing, 36(1), 39–62. https://doi.org/10.1016/j.ijresmar.2018.09.007

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

Pang, B., Lillian, L., Shivakumar, V., 2002. Thumbs up?: Sentiment classification using machine learning techniques. Conference on Empirical Methods in Natural Language Processing (EMNLP), 79–86.

Petiot, J. F., Grognet, S., 2006. Product design: A vectors field-based approach for preference modelling. Journal of Engineering Design, 17(03), 217–233. https://doi.org/10.1080/09544820500275032

Pfeffer, J., Zorbach, T., Carley, K. M., 2014. Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. Journal of Marketing Communications, 20(1–2), 117–128. https://doi.org/10.1080/13527266.2013.797778

Read, J., 2005. Using emoticons to reduce dependency in machine learning techniques for sentiment classification. 43nd Meeting of the Association for Computational Linguistics (ACL), 43–48. https://doi.org/10.3115/1628960.1628969

Ren, Y., Ji, D., 2017. Neural networks for deceptive opinion spam detection: An empirical study. Information Sciences, 385, 213–224. https://doi.org/10.1016/j.ins.2017.01.015

Rhim, H., Cooper, L. G., 2005. Assessing potential threats to incumbent brands: New product positioning under price competition in a multisegmented market. International Journal of Research in Marketing, 22(2), 159–182. https://doi.org/10.1016/j.ijresmar.2004.06.005

Sailunaz, K., Alhajj, R., 2019. Emotion and sentiment analysis from Twitter text. Journal of Computational Science, 36, 101003. https://doi.org/10.1016/j.jocs.2019.05.009

Schaer, O., Kourentzes, N., Fildes, R., 2019. Demand forecasting with user-generated online information. International Journal of Forecasting, 35(1), 197–212. https://doi.org/10.1016/j.ijforecast.2018.03.005

Song, J., Kim, K. T., Lee, B., Kim, S., Youn, H. Y., 2017. A novel classification approach based on Naïve Bayes for Twitter sentiment analysis. KSII Transactions on Internet and Information Systems, 11(6), 2996–3011. https://doi.org/10.3837/tiis.2017.06.011

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. https://doi.org/10.2753/mis0742-1222290408

Sulis, E., Irazú Hernández Farías, D., Rosso, P., Patti, V., Ruffo, G., 2016. Figurative messages and affect in Twitter: Differences between #irony, #sarcasm and #not. Knowledge-Based Systems, 108, 132–143. https://doi.org/10.1016/j.knosys.2016.05.035

Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., Qin, B., 2015. Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification. 52nd Annual Meeting of the Association for Computational Linguistics, 1555–1565. https://doi.org/10.3115/v1/p14-1146

Thelwall, M., Buckley, K., Paltoglou, G., 2011. Sentiment in Twitter events. Journal of the Association for Information Science and Technology, 62(2), 406–418.

Tirunillai, S., Tellis, G. J., 2014. Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. Journal of Marketing Research, 51, 463–479.

https://doi.org/10.1509/jmr.12.0106

Verma, J. P., 2013. Multidimensional Scaling for Product Positioning. In Data Analysis in Management with SPSS Software (pp. 443–460). Springer India. https://doi.org/10.1007/978-81-322-0786-3_14

Vosen, S., Schmidt, T., 2011. Forecasting private consumption: survey-based indicators vs. Google trends. Journal of Forecasting, 30(6), 565–578. https://doi.org/10.1002/for.1213

Wang, W., Feng, Y., Dai, W., 2018. Topic analysis of online reviews for two competitive products using latent Dirichlet allocation. Electronic Commerce Research and Applications, 29, 142–156. https://doi.org/10.1016/j.elerap.2018.04.003

Xia, R., Zong, C., Li, S., 2011. Ensemble of feature sets and classification algorithms for sentiment classification. Information Sciences, 181(6), 1138–1152. https://doi.org/10.1016/j.ins.2010.11.023

Xiao, S., Wei, C. P., Dong, M., 2016. Crowd intelligence: Analyzing online product reviews for preference measurement. Information and Management, 53, 169–182. https://doi.org/10.1016/j.im.2015.09.010

Zeng, D., Chen, H., Lusch, R., Li, S. H., 2010. Social media analytics and intelligence. IEEE Intelligent Systems, 25(6), 13–16. https://doi.org/10.1109/MIS.2010.151

Zhao, K., Stylianou, A. C., Zheng, Y., 2018. Sources and impacts of social influence from online anonymous user reviews. Information and Management, 55(1), 16–30. https://doi.org/10.1016/j.im.2017.03.006

**23/12/2019**

Dear Sir/Madam,

I confirm that the manuscript has not been previously published, is not currently submitted for review to any other journal and will not be submitted elsewhere before a decision is made by this journal.

Sincerely
Andreas Gregoriades

**Manuscript Highlights**

The manuscript makes the following contributions:

- Proposes new method that utilizes eWOM analytics for evaluating a new product positioning strategy utilizing contemporary techniques in topic modelling and sentiment analysis
- Uses longitudinal analysis to identify the effect of time on product's sentiment
- Uses ordinal logistic regression to evaluate the effect of time on the product positioning and diagnose the cause of mispositioning
- Employs design science as a framework for the development and evaluation of the proposed method
- Apply the method in a case-study from the smartphone industry