Text Mining Algorithms for Extracting Brand Knowledge. The fashion industry case.

Nina Rizun, Gdańsk University of Technology, Gdańsk, Poland, nrizun@zie.pg.gda.pl
Wioleta Kucharska, Gdańsk University of Technology, Gdańsk, Poland, wkucharska@zie.pg.gda.pl

Abstract

Brand knowledge is determined by customer knowledge. The opportunity to develop brands based on customer knowledge management has never been greater. Social media as a set of leading communication platforms enable peer to peer interplays between customers and brands. A large stream of such interactions is a great source of information which, when thoroughly analyzed, can become a source of innovation and lead to competitive advantage. Semantic analysis is a prominent field of data mining that deals with key contexts, topics, and sentiment of these interactions. The challenge and key to the success are creating a proper searching algorithm to analyze these areas of interest. The purpose of this study is to develop and to test a methodology which will identify principal points of customers' interactions with fashion brands using a set of Text Mining Algorithms. The fashion industry is one of the most successful in the social media environment. Deep understanding of fashion brand communication is interesting from the theoretical and practical point of view. The theoretical value of this study contributes to the social media brand knowledge management by providing a set of gained insights thanks to the implementation of the new methodological approach presented in this study. The practical value is the knowledge about the presence of fashion brands in social media obtained in the course of the study.

Keywords: text mining algorithms methodology, brand knowledge, fashion industry, customer knowledge, semantic analysis, social media

Introduction

Competitive and rapidly changing digital businesses need to generate customer value. To do it, customer knowledge needs to be harvested and subjected to the value generation and evaluation processes. Knowledge management focuses on situated mutual learning as the core of knowledge management (Ferguson et al., 2010). Cepeda-Carrion et al. (2017) pointed that engaging customers in the knowledge-cycle will generate customer-focused synergy in the organization that creates superior customer value. Customer knowledge is a set of information for, from, and about the customers (Khodakarami and Chan, 2014; Trejo, et al., 2015). Garcı´a-Murillo and Annabi (2002), and Winer (2001) stressed that customer knowledge is a critical asset and a source of competitive advantage for organizations. Findings of Fidel et al. (2015) confirm that effects of customer knowledge management (CKM) orientation are reflected by customer collaboration and innovation and finally visible in marketing results. Customer knowledge management practices are mutually beneficial; for the company and for the customer. Mehdibeigi et al. (2016) highlighted that CKM is a good method to capture the knowledge of customers and supply him with tailored knowledge. A good example of such bespoke knowledge is brand knowledge.

Brand knowledge reflects how customers perceive and evaluate brands based on their entire experience of a particular brand (Keller 2003, Pappu et al., 2005). According to Esch et al. (2006), brand knowledge must be considered in the context of a brand-customer relationship that includes brand satisfaction, trust, and attachment to a particular brand. The concept of Brand Knowledge Management focuses on brand-driven organizations, from content to process, and from data to tacit knowledge (Richards et al., 1998). Tacit knowledge as a source of innovation (Kucharska and Kowalczyk, 2016) leads to the desired competitive advantage. Social media development is a rapidly growing trend today. In 2017, Digital Global Overview (wearesocial.com) revealed that more than half of the world's population uses the Internet, and 37% of all people globally are social networks users, which, in 2017, generated 41 billion $ in advertising revenue. The amount went up from 17.85 billion in 2014 (statista.com). The social media era drives new business models creation (Doganova and Eyquem-Renault, 2009) based on the new knowledge elicited from customers. Consumers behavior and value perception are changing very quickly today. Thus, the technology-driven business environment must be more innovative than ever before. The best solution for "keeping up with the customer demand" is first to harvest this brand knowledge from the customer, next analyze it, and finally tailor the value for it. The research problem formulated in this manuscript is: how to extract brand knowledge from social media users?

The fashion industry is one of the most successful in the social media environment. Therefore, understanding rules governing its communication is interesting from the theoretical and practical point of view. The theoretical value of this study is developing a new methodology which will enable the capturing of brand and customer knowledge. The practical value is providing fresh insight on the subject, gained thanks to the new data on the presence of fashion brands in social media obtained during the study. Social media brand performance measurement relies on quantity and quality measures. A recent trend in the digital marketing analytics is to track not only the number of consumer actions but also to analyze consumers' feelings and opinions on brands, products, and services in social media (Hemann and Burbary, 2013). The quantity factor allows us to assess for example the size of particular audience reflected in the number of fans who identify with a particular brand and the level of their engagement reflected in the number of brand references which vary due to the time of day and the day of the week and hour of the day (Dolan et al., 2017). The organic, positive word of mouth (WOM) generated by users is the goal to be achieved by each company. The knowledge on what topics are more frequently discussed and the tonality of these discussions are crucial from the company's point of view. The quality measures allow assessing not only the positive or negative sentiment of comments but also the objective or subjective character of the mentions. The key point of success in this area is to create a proper searching algorithm. The purpose of this study is to develop and test a methodology for identification of key points in customers' interactions with fashion brands using the set of Text Mining Algorithms.

Conceptual Framework

According to the theory of Text Mining Algorithms, to implement opinion mining it is usually necessary to perform the following tasks: entity (product, service, person, event, organization, or topic) extraction and grouping; entity sentiment classification; entity aspects (keywords) of meaning and sentiment extraction and grouping, and visualization/presentation of results (Liu and Zhang, 2012; Junaid et al., 2017). At the same time, sentiment analysis and opinion mining are mainly based on supervised learning (machine learning, naive Bayesian classification, and support vector machines), unsupervised methods (text categorization, text clustering, topic modelling), and hybrid approach. Most of the current developments are aimed at improving the quality of the evaluation of the polarity as well as sentence subjectivity of the analyzed opinions (based on machine learning, Natural Language Processing and hybrid approaches) taking into account the specifics of languages and dialects (Kumar and Reddy, 2012; Glucker et al., 2014; Hemalatha et al., 2013). Less attention is paid to the development of qualitative algorithms and techniques for identifying the main aspects (topics, problems) described in the analyzed opinions, as well as finding the effective instruments to support the process of interpreting the results obtained not only in the context of their polarity, but also the cumulative positioning (rating) of the objects of these opinions.

Methodology

Due to their global position, three fashion brands, i.e. Adidas, Nike, and Zara have been selected for the purposes of the study. Adidas and Nike sell sportswear, whereas Zara offers casual clothes. The choice of brands was intentional. Its purpose was to observe insights provided by the particular brands' knowledge and the particular brands' category knowledge. The following research questions were formulated:

Q1: What topics are most often discussed on Facebook pages of the chosen brands?
Q2: What kind of tonality characterizes the most important topics of the chosen brands?
Q3: Are there prerequisites for the clustering of the chosen brands? (visible similarities)

Data were gathered from January to November 2017. All textual data in English (posts, comments, etc.) were grouped according to selected brands and next analyzed. The datasets were collected via Facepager software which is usually used to extract public existing data from Twitter, Facebook, and other social media based on API. It collects URLs from query setup. Next, the extracted data was stored in a local database and exported to a CSV file. The structure of the sample is presented in Table 1.

Table 1: Structure of the sample

Brand	Number of posts	Number of comments
Adidas	4 300	438 300
Nike	5 403	504 987
Zara	4 844	403 986

Source: authors' own study.

The text mining procedure of the obtained representative sample used the principle of revealing the Latent-Semantic Relationships (LSR) in non-structural (textual) data. The aim of the LSR analysis is to extract "semantic structures" from the text collection and to automatically expand them onto an underlying topic. To increase the accuracy of the LSR Modelling, the advantages of two different mathematical frameworks were used (Rizun et al., 2017):

- Latent Dirichlet Allocation (LDA) is a generative model that uses a Bayesian model that treats each document as a mixture of latent underlying topics, where each topic is modelled as a mixture of word probabilities from a vocabulary.
- Latent Semantic Analysis (LSA) is a discriminant theory and method for extracting context-dependent word meanings by statistical processing of large sets of text data. It uses a "bag-of-words" for modelling. This procedure starts from transforming text corpora into term-document frequency matrices to reduce the high dimensional term spaces of textual data to a user-defined number of dimensions by singular value decomposition (SVD) and next generates weighted term lists for each concept or topic. The concept or topic content is weighted for each document, then outputs that can be used to compute document relationship measures are produced.

As was proved by Rizun and Waloszek (2017), the synergy effect of using the advantages of LSA and LDA methods consists in applying them for each brand. This process was carried out on two levels:

- Level of decomposition – which defined a Contextual Summary for each Facebook Page as a set of latent probabilistic topics with information about most probable (significant) words assigned to this topic (LDA algorithm). For obtaining the optimal combination – a number of topics/number of terms in the topic – the values of Perplexity were used. The optimum value of the Perplexity Index is achieved at the moment when further changes in the parameters do not lead to its significant decrease. Restriction: due to the inability to qualitatively identify the topic in a shorter message, only textual messages not shorter than 30 characters can be subject to this analysis.

- Level of synthesis – identifying latent semantic relationships between topics within a defined Contextual Summary in order to identify top clusters that characterize each fashion brand of each set of Facebook pages. The result of this phase – the Contextual Framework (sets of latent semantic topics) for the whole sample is based on the presence of latent semantic relationships between the elements of Contextual Summary (LSA algorithm). Restriction: with the objective of unifying the results for the whole sample, 5 topics were selected for each brand's Contextual Summary. This particular amount provided the maximum average level of semantic cosine similarity of topics within the clusters for each brand's sample.
- Python programming language and natural language processing (NLTK) package were used as technical tools.

Identifying the structure of Contextual Framework of fashion brands enabled formulating the Hierarchical Semantic Corpora as a structure of the top sub-samples of each fashion brand. In this step, the LSA algorithm is used in the following interpretation: each element of the Contextual Framework in the format of a set of topical keywords and their weights was added to appropriate sample of fashion brands as a query. As a result of LSA algorithm realization, we receive a Hierarchical Semantic Corpora as a structured set of textual messages (posts, comments) semantically close to topics from the Contextual Framework (Table 2).

Table 2: The topical structure of Hierarchical Semantic Corpora of fashion brands' sample

Adidas		Nike		Zara	
Topics	**comments [%]**	**Topics**	**comments [%]**	**Topics**	**comments [%]**
Person_Product	12.94%	**Customer_Service**	14.03%	**Collection**	9.88%
Product	24.18%	**Online_Shop**	8.91%	**Products**	30.29%
Saling_Buying	14.74%	**Product**	20.19%	**Customer_Service**	10.55%
Sport	21.11%	**Collection**	22.08%	**People/Events**	10.68%
Customer Service	13.51%	**Zara Look**	10.02%	**Sport**	7.63%
Other	13.52%	**Other**	24.79%	**Other**	30.98%

Source: authors' own study.

The next step involved conducting a sentiment analysis of the Hierarchical Semantic Corpora content. The technical tool which we used for this stage was open-source solutions: AYLIEN Text Analysis for Google Sheets. The threshold value for valid polarity confidence equaled 0.7. Extracting sentiment from textual messages was provided by revealing the author's emotions and perspective. Was the tone of the message positive, neutral or negative? Was the text subjective (reflecting the author's opinion) or objective (expressing a fact)? For interpretation of the received results, Multidimensional Scaling (MDS) found similarities among groups of topics. MDS technique creates a map displaying relative positions of a number of objects, given only a table of the distances between them. R software was employed for this purpose. The results obtained in this step were visualized on MDS maps, next analyzed and discussed. The results of the methodology of the sentiment analysis by topics have been summarized in Table 3.

Table 3: Sentiment Analysis by Topics (SAT) Methodology scheme

STEP 1	The problem statement, objectives formulation, sample identification and data gathering. *tool: Facepager*
STEP 2	Latent-Semantic Relationships (LSR) in non-structural (textual) data analysis based on the combination of: ▪ Latent Dirichlet Allocation (LDA) ▪ Latent Semantic Analysis (LSA)

	a. Revealing of Latent-Semantic Topical Structure based on: ▪ Latent Probabilistic Topics definition (decomposition) ▪ Latent Semantic Relationships identification (synthesis) *tool: Python*
	b. Hierarchical Semantic Corpora creation *tool: Python*
STEP 3	Sentiment Analysis in the context of the key topics *tool: AYLIEN Text Analysis for Google Sheets*
STEP 4	Quantitative estimating of similarity among groups of Topics using Multidimensional Scaling (MDS) *tool: R software*
STEP 5	Determining similarities between fashion brands using LSA and MDS algorithms *tool: R software*

Source: authors' own study.

Results

Semantic analysis is one of the prominent fields of data mining that deals with key contexts, topics, and sentiment identification. The subjectivity and objectivity of obtained results also matter. Tables 4a-c present the sentiment in the context of key topics identified in Table 2, also including subjectivity/objectivity analysis of opinions for the chosen brands. Results show that the majority of analyzed comments are subjective, whereas the sentiment varies respectively for different topics and brands. Neutral sentiment dominates for Nike, except for the "product" topic where negative comments prevail. The topic with the majority of negative comments for Adidas is "sport". The most positive is "people and events" and "customer service" in a sharp contrast to Zara where "customer service", next to "online shop" and "product", is the most negative subject for comments and conversations. Sentiment analysis identifies and analyses sentimental contents generally available in social media (Dhaoui et al., 2017). It can be achieved by using a lexicon of weighted words (Taboada et al., 2011). It is an approach used widely for sentiment analysis in the marketing research community (Bolat and O'Sullivan, 2017) as it does not require any pre-processing or training of the classifier. It is important to highlight that each social media platform has a specific set of methods which make it possible to express sentiment, e.g., emoticons, emojis, hashtags, abbreviations, slang language, etc. Moreover, semantic analysis based on language texts allows studying sentiment directly from the text (comments) by analyzing not only the meaning of words but also the context of a particular opinion. These methods require careful validation of each language separately before being used by marketers on the social media data. A good example of a professional tool for sentiment analysis is Sentione.com which has been successfully validated and enables monitoring and analyzing online brand mentions in over 26 languages. In fact, this makes analyzing the sentiment of a particular brand globally available.

Table 4a: Adidas' sentiment in the context of key topics

ADIDAS	**positive**	**Neutral**	**Negative**	**subjective**	**objective**
Collection	24.39%	**58.28%**	17.33%	85.71%	14.29%
Market/Products	24.68%	**55.10%**	20.22%	94.60%	5.40%
Customer_Service	**60.59%**	35.96%	3.45%	93.60%	6.40%
People/Events	**48.24%**	29.65%	22.11%	74.13%	25.87%
Sport	31.28%	29.19%	**39.53%**	94.96%	5.04%
Other	14.10%	**76.36%**	9.54%	87.83%	12.17%

Source: authors' own study.

Table 4b: Nike’s sentiment in the context of key topics

NIKE	positive	Neutral	negative	subjective	objective
Person_Product	21.36%	**47.84%**	30.80%	96.19%	3.81%
Product	30.19%	29.38%	**40.43%**	98.82%	1.18%
Purchase	30.17%	**36.94%**	32.89%	97.58%	2.42%
Sport	22.93%	**56.35%**	20.72%	76.67%	23.33%
Customer Service	33.54%	**36.06%**	30.40%	97.97%	2.03%
Other	33.54%	**36.06%**	30.40%	94.55%	5.45%

Source: authors’ own study.

Table 4c: Zara’s sentiment in the context of key topics

ZARA	positive	Neutral	negative	subjective	objective
Customer_Service	21.08%	20.46%	**58.46%**	98.35%	1.65%
Online_Shop	27.14%	28.57%	**44.29%**	99.04%	0.96%
Product	26.54%	25.31%	**48.15%**	96.52%	3.48%
Collection	**68.00%**	14.29%	17.71%	99.51%	0.49%
Zara Look	**53.70%**	20.37%	25.93%	97.17%	2.83%
Other	13.30%	**44.14%**	42.56%	96.10%	3.90%

Source: authors’ own study.

The obtained plots allow visualizing the tonal similarity between topics within each fashion brand using the two-dimensional scatter. The main characteristics of interest are the relative positions of points and groups of topical opinions that are apparent. Dimension titles result from the underlying factors that create these tonal topical similarities. Figures 1a-c show visualized effects of this step. Tables 5a-c summarize the tonality according to key topics.

Table 5a: Adidas’ tonality of key topics

Adidas Group Description	Topics of Opinions
Positive opinions prevail, partly – neutral	Customer Service
Diverse opinions	People/Events, Sport
Neutral opinions prevail	Market/Products/Collection

Source: authors’ own study.

Table 5b: Summary of Nike’s tonality of key topics

Nike Group Description	Topics of Opinions
Neutral opinions prevail	People/Product, Sport
Diverse opinions	Customer Service, Sailing/Buying
Negative opinions prevail	Product

Source: authors’ own study.

Table 5c: Summary of Zara’s tonality of key topics

Zara Group Description	Topics of Opinions
Positive opinions prevail	Collection, Zara Look
Neutral opinions prevail	Online Shop, Product
Negative opinions	Customer Service

Source: authors’ own study.

Tonality analysis makes it possible to assess the tonal similarity between topics for each brand.

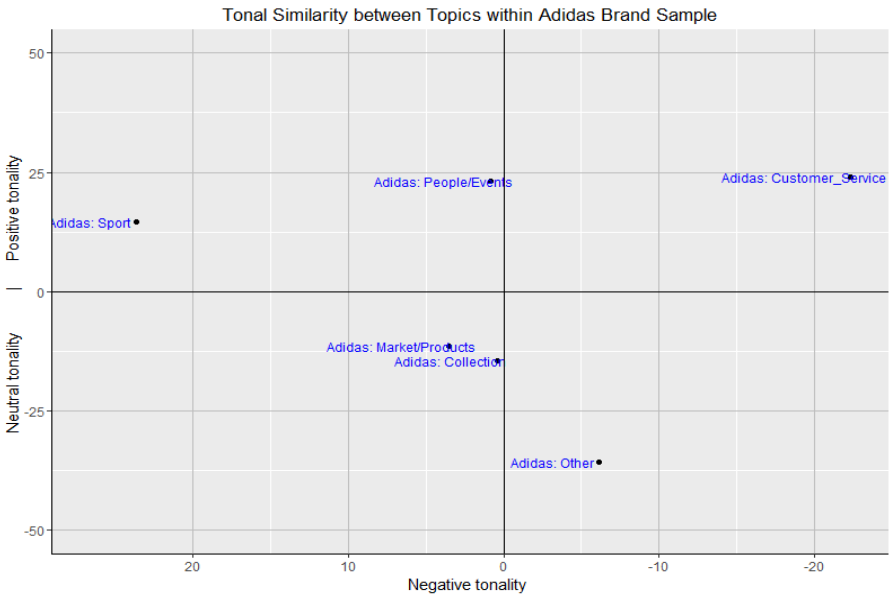

Fig 1a. Visualization for Adidas' tonality of key topics.
Source: authors' own study developed with R software environment.

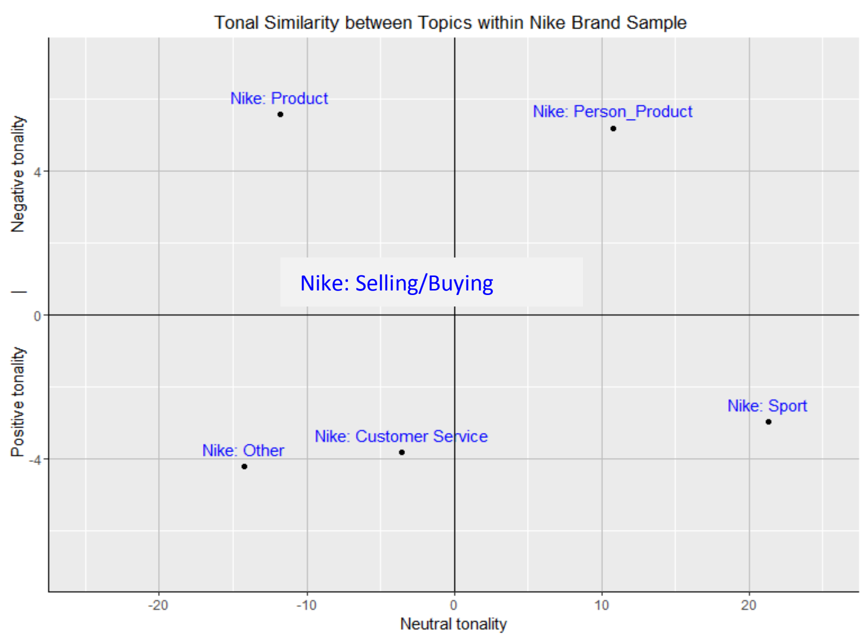

Fig 1b. Visualization for Nike's tonality of key topics.
Source: authors' own study developed with R software environment.

Results for Nike and Adidas, brands which offer sports clothes, show that topics such as People/Product/Sports dominate; however, Nike achieves more positive tonality. The Adidas sentiment is rather neutral in this area. It means that Nike performs much better with regard to these topics, which makes it an area of brand expertise and a source of strength, a good platform for market advantage creation. In conversations about the product, Nike achieved a rather neutral sentiment, whereas Adidas sentiment is diversified. According to Erickson and Rothberg (2017) and Kucharska's (2018) findings, neutral sentiment usually dominates in social media. Thus, the described situation leads to a conclusion that Nike brand is more “mainstream” than Adidas. Nike is more often spoken about with a neutral sentiment, whereas Adidas sentiment is less spoken about but more expressive. The main similarity between these two-sport fashion brands is that People/Product/Sport /Events topics are the main interest areas for their fans. The main difference is the sentiment which reflects the bond character. Figure 1c presents Zaras’ tonality of key topics visualization.

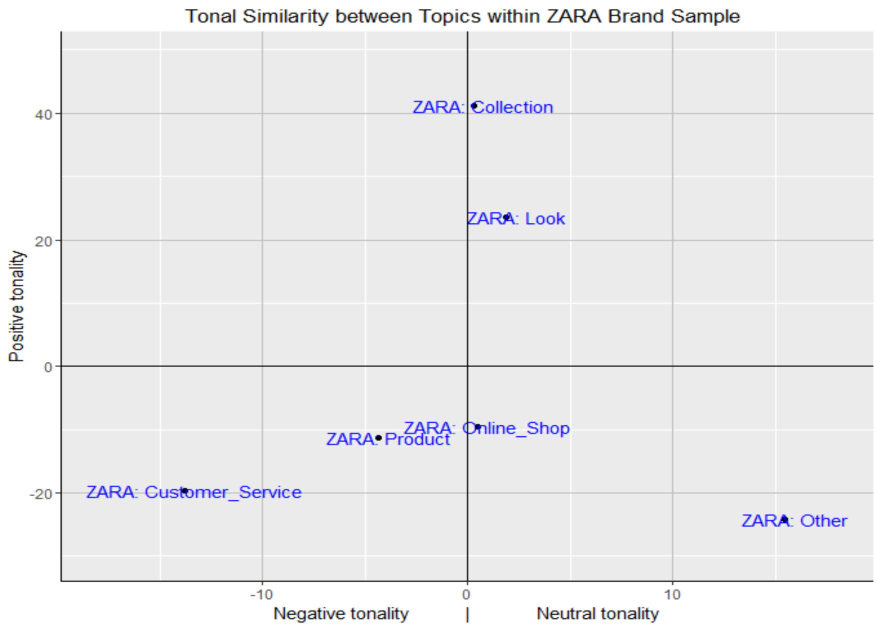

Fig 1c. Visualization for Zara’s tonality of key topics.
Source: authors’ own study developed with R software environment.

In reference to Zara, a casual fashion brand, the positive sentiment is connected with the company’s fashion collection, whereas the negative one with the customer service. Figure 2 presents Zara, Adidas, and Nike's social media performance. Leading sport fashion brands vary. It is reflected in their brand image and expertise area. Zara’s image is of a casual, “mainstream” fashion expert. Nike and Adidas' image of expertise in the sports area is consistent with the identified topics (People/Product/Sport /Events) which sustain the brand expertise and are interesting for the social media audience. It is much easier to draw the attention of brand followers using content related to up-to-day news and which is far away from forceful sales tactics, which does not generate the organic WOM (worth of mouth). The mentions (WOM) of Zara, an expert in casual clothing, relate to the brand utility such as collection and customer service.

Based on the presented cases, the answer to the research question asked at the beginning of the study is that fashion-related topics most often discussed on Facebook depend on the fashion sub-category to which they belong. Positive and neutral tonality is observed for “narrow expertise”, whereas “mainstream” brands are more exposed to negative comments. For further analysis, brands should be clustered in sub-categories.

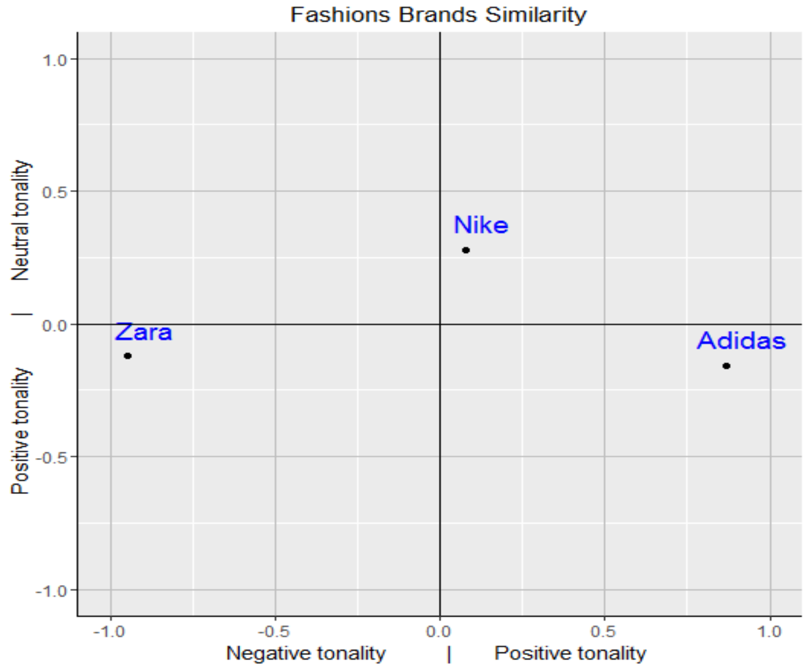

Fig 2. Fashion brands similarity assessment
Source: authors’ own study developed with R software environment

Limitations, Discussion and Further Research

The main limitation of the study, next to the window of data collection period, is the fact that it is based solely on three fashion brands. Such a number of brands belonging to the same industry is enough for the SAT methodology testing (the purpose of the study), but it is not enough to draw definite conclusions about fashion brands. A broader study is required to accomplish this goal. Geurin and Burch (2017) examined a user-generated content (UGC) of six sport fashion brands on Instagram. They found out that Nike, a brand with the largest number of followers and leading global popularity, which posts mostly brand-generated content, does not engage fans as much as Adidas does which has more UGC than Nike. Their findings are complementary to the findings presented in this study regarding Facebook. Having a broader picture of the social media brand strategy in mind, and despite the fact that one study is on sentiment and the other on UGC, and each of them examined a different social media platform (Instagram and Facebook), it is stressed in both of them that neutral sentiment and moderated engagement dominates Nike's mentions. Nike's content on Facebook and Instagram does not engage its followers as much as Adidas whose content is more often generated by users than it is in Nike's case. The sentiment of Adidas' content on Facebook is more diversified and more positive than Nike's. It proves that the presented SAT method if effective as a tool for gaining knowledge. In light of findings by Geurin and Burch (2017) and Burmann and Arnhold (2009), the results of this study confirm that user-generated-branding (UGB) is viewed by consumers as more authentic and engaging than the content generated by companies. The emergence of social media turned Internet users from content readers into content publishers (Chua and Banerjee, 2013), which would be impossible without the knowledge about both customers and brands. The SAT sentiment extraction model proposed in this study is a useful tool that helps to generate knowledge which then serves as a platform for further development of a social media strategy for a particular brand.

Conclusion and Implications

The purpose of this study was to develop and test a methodology for the identification of key points of customers' interactions with fashion brands, using a set of Text Mining Algorithms. From the theoretical point of view, the novelty of the proposed SAT methodology is: increasing the quality of mining knowledge about the analyzed brands by improving algorithms for determining topics of comments; providing opportunities for knowledge-based decision support by applying algorithms for interpreting and visualizing the relative position of a brand in the light of social media reputation. As a practical result, the SAT methodology proposed in this study allowed us to identify key topics which drive social media brands' WOM, their subjective or objective character, and the tonality. The organic, positive WOM generated by users is something that all companies aim for. Social media strategy should reflect the general strategy of the brand and complement it. Knowing the key, frequently discussed topics and the tonality of these discussions is a good starting point to achieve this goal. The proposed SAT methodology is a ready, step by step guide how to gain the knowledge about brands and their customers in the social media. The study which included three brands belonging to one category was enough to test and present the SAT methodology. From the practical point of view, a broader, deeper study of a particular industry, category, and brand made it possible to gain a full understanding of the sentiment, useful for effective management of a brand in social media. The scope of the paper is limited by the window of data collection period, one brand category, and the number of analyzed cases (only three brands). The preliminary character of the research, a rapidly changing digital environment, and applying the SAT methodology to other industries and categories are the basis for this method's further development. The main value of this study is creating a tool which will facilitate the acquisition of knowledge necessary for effective brand management in the social media environment.

References

Burmann, C. and Arnhold, U. (2009), User-generated branding: State of the art research, Wiesbaden, Germany: Gabler Verlag.

Cepeda-Carrion, I., Martelo-Landroguez, S. and Leal-Rodríguez, A.L. (2017), 'Critical processes of knowledge management: An approach toward the creation of customer value,' *European Research on Management and Business Economics*, 23, 1–7.

Chua A.Y.K and Banerjee, S. (2013), 'Customer knowledge management via social media: the case of Starbucks,' *Journal of Knowledge Management*, 17(2), 237-249.

Doganova, L. and Eyquem-Renault, M. (2009), 'What do business models do?: innovation devices in technology entrepreneurship,' *Research Policy*, 38(1),1559-1570.

Dolan, R., Conduit, J., Fahy, J. and Goodman, S. (2017), 'Social media: communication strategies, engagement and future research directions,' *International Journal of Wine Business Research*, 29(1), 2-19.

Erickson S. and Rothberg H. (2017), 'Digital Media and Relational Capital", In: Proceedings of the 18th European Conference on Knowledge Management, 1, 303-309.

Esch F- R., Langner T., Schmitt B.H. and Geus P. (2006), 'Are brands forever? How brand knowledge and relationships affect current and future purchases,' *Journal of Product & Brand Management*, 15(2), 98-105.

Ferguson, J., Huysman, M. and Soekijad, M. (2010), 'Knowledge Management in Practice: Pitfalls and Potentials for Development,' *World Development*, 38(12), 1797–1810.

Fidel, P., Schlesinger, W. and Cervera, A. (2015), 'Collaborating to innovate: Effects on customer knowledge management and performance,' *Journal of Business Research*, 68, 1426–1428.

Garcı´a-Murillo, M. and Annabi, H. (2002), 'Customer knowledge management,' *Journal of the Operational Research Society*, 53, 875–884.

Geurin, A.N. and Burch, L.M. (2017), 'User-generated branding via social media: An examination of six running brands,' *Sport Management Review*, 20, 273-284.

Glucker, H., Burghardt, M. and Wolff, Ch. (2014), ' Sentilyzer – A Mashup Application for the Sentiment Analysis of Facebook Pages,' Proceeding of KONVENS 2014, 111-116.

Hemann, C. and Burbary, K. (2013), Digital Marketing Analytics: Making Sense of Consumer Data in a Digital World, Pearson Education.

Hemalatha, I., Saradhi Varma, G.P and Govardhan, A. (2013), 'Sentiment Analysis Tool using Machine Learning Algorithms", International Journal of Emerging Trends & Technology in Computer Science, 2(2), 105-109.

Junaid, S.M., Jaffry, S.W., Yousaf, M.M., Aslam, L. and Sarwar S. (2017), 'Sentiment Analysis and Opinion Mining - A Facebook Posts and Comments Analyzer,' *Technical Journal, University of Engineering and Technology (UET) Taxila*, 22, 98-104.

Keller, K.L. (2003), Strategic Brand Management: Building, Measuring, and Managing Brand Equity, 2nd ed., Prentice Hall, Upper Saddle River, NJ.

Khodakarami, F. and Chan, Y.E. (2014), 'Exploring the role of customer relationship management (CRM) systems in customer knowledge creation,' *Information & Management*, 51, 27–42.

Kucharska, W. and Kowalczyk, R. (2016), 'Tacit Knowledge Sharing and Creativity. How to Derive Innovation from Project Teams? ,' Proceedings of the 11th European Conference on Innovation and Entrepreneurship: ECIE 2016, 444-452.

Kucharska, W. (2018), 'Personal brand value and social media. The top football players' case,' Proceedings of the 31st International Business Information Management Conference IBIMA 2018, Milan, Italy 25-26 April.

Kumar S.B. and Reddy B.S.D. (2016), 'An Analysis on Opinion Mining: Techniques and Tools,' *Paripex - Indian Journal of Research*, 5(8), 489-492.

Liu, B. and Zhang, L. (2012), 'A survey of opinion mining and sentiment analysis,' In: *Mining Text Data*, Aggarwal, Ch.C., Zhai, Ch.X. (eds.),415-463.

Mehdibeigi, N., Dehghani, M. and Yaghoub, N.M. (2006), 'Customer Knowledge Management and Organization's Effectiveness: explaining the mediator role of Organizational Agility,' *Procedia - Social and Behavioral Sciences*, 230, 94 – 103.

Pappu, R., Quester, P. and Cooksey, R.W. (2005), 'Consumer-based brand equity: improving the measurement & empirical evidence,' *Journal of Product & Brand Management*, 14(3), 143-54.

Richards, I., Foster, D. and Morgan, R. (1998), 'Brand Knowledge Management: Growing Brand Equity,' *Journal of Knowledge Management*, 2(1), 47-54.

Rizun, N. and Waloszek, W. (2017), 'The Algorithm of Building the Hierarchical Contextual Framework of Textual Corpora,' Proceeding of Eighth IEEE International Conference on Intelligent Computing and Information System ICICIS 2017, Cairo, Egypt, 366-372.

Rizun, N., Taranenko, Y. and Waloszek, W. (2017), 'The Algorithm of Modelling and Analysis of Latent Semantic Relations: Linear Algebra vs. Probabilistic Topic Models,' Knowledge Engineering and Semantic Web. Knowledge Engineering and Semantic Web, Publisher: Proceedings of the 8th International Conference (KESW 2017), 53-68.

Statista.com (2017), 'Social network advertising revenue from 2014 to 2017,' available at: https://www.statista.com/statistics/271406/advertising-revenue-of-social-networks-worldwide/ (1.02.2018).

Trejo, J.M., Gutiérrez, J.S. and Guzman, G.M. (2016), 'The customer knowledge management and innovation,' *Contaduría y Administración*, 61, 456–477.

Wearesocial.com (2017), 'Digital in 2017 Global Overview Report 2017,' available from https://wearesocial.com/special-reports/digital-in-2017-global-overview (31.01.2018).

Winer, R.S. (2001), 'A framework for customer relationship management,' *California Management Review*, 43, 89–105.

Zhang, Z.J. (2011), 'Customer knowledge management and the strategies of social software,' *Business Process Management Journal*, 17(1), 82-106.