International Business Management
MEDWELL PUBLICATIONS

Methods of Marketing and Behavioral Economic Research Based on Automated Text Analytics Tools

[1]Vasily Kashkin and [2]Yuriy Androsik
[1]*State Institute of Macroeconomic Research of Russia, Moscow, Russia*
[2]*Belarusian State Technological University, Minsk, Belarus*

Key words: Text analysis, psychological text analysis, text analytics software, text analysis functions, text classification, text extraction, economic behavioral research, marketing text analysis, customers decision-making process

Corresponding Author:
Vasily Kashkin
State Institute of Macroeconomic Research of Russia, Moscow, Russia

Page No.: 391-397
Volume: 14, Issue 12, 2020
ISSN: 1993-5250
International Business Management

Abstract: The study considers the main text analytics programs and their functions. The researcherd studied the functions of 54 leading text analysis software and identified in which areas of marketing research they can be applied. Authors offer marketing (customers) research methods based on text analytics tools. In this study, we found out that automated text mining and analysis is especially useful to research the decision-making process of target customers which is not the same as their psychological portrait. Study of decision-making process requires specific research design, based on combination and adjusting of different text analytics tools. Understanding of customers decision-making process is relevant for individual work with clients, setting contextual advertising and E-mails customization, designing products, proposals and marketing campaigns for targeted customers groups.

INTRODUCTION

Every day petabytes of information are generated all over the world. People share their opinions on social networks, discuss trends, brands, fashion, food, their problems, friends, colleagues, companies and other life aspects. This unstructured information seems scattered at first glance, useless or of little significance. Nevertheless, this is not so. It contains, so to say, an informational trace or imprint of personalities. The analysis of such textual information opens up wide business opportunities, especially for marketing and marketing research.

Each person is a buyer, consumer or user of products and services. Knowledge of the features of his/her thinking and patterns of behavior allows to closely interact with him and personalize the actions of the business. Obtaining this kind of knowledge is possible by analyzing the text, since, text is a reflection of a person's personality and psychology.

Due to the huge amount of existing data such researches in marketing are advisable to carry out by software. However, the choice of software is also a difficult task, since, the field of natural language processing and analysis of the language as a whole is quite diverse and complex. For this reason, the text analytics software market presents dozens of programs from individual libraries of general-purpose programming languages to off-the-shelf industry solutions for business. Moreover, the range of tasks that this software can solve is significantly different from program to program. This leads to the fact that it is not always clear which marketing task which tool should be addressed.

Research problems: Text analysis is important for marketing. However, today most of the software is divided into its functions, regardless of the marketing tasks. For this reason, programs offer specific functions, often with little or no indirect relation to marketing.

Marketing pursues integral tasks. This is understanding the client and building the right communication with him, finding the target client and knowing his priorities as well as analyzing competitors and customer attitudes towards them.

So far, there is no integrated marketing software because the existing marketing software has narrow tasks. This is either an SMM analysis of own campaigns or an SMM analysis of competitors and their campaigns or what people write on the networks but with the goal of involving them. Or, marketing software is intended for predictive analytics, that is, for statistical processing of data related to marketing and sales.

This leads to the fact that it is not always clear which marketing task it is necessary to solve by which text analytics tool which software functions correspond to which marketing functions and tasks.

Most of marketing research tasks should be solved with a complex use of different text mining and analytics tools.

Research purpose: The main goal of the article is to offer an approach to how we can adapt existing text analysis software to solve marketing research problems, given that most of them are not adapted for these tasks. In particular to define the functions of the text analysis software and suggest the ways how they can be used in marketing and behavioral economic research.

Research tasks: To achieve research goals we have examined over 50 text mining and analytics tools (Table 1). We accomplished an extensive review of previous research on text mining and analytics methods for the purposes of marketing and economic research. Describe functions and methods for identifying customers and determining their parameters, for example, gender, age, education, seniority. Describe functions and methods for analysis of potential customers social relations, for example with which person or company he is corresponding whom he mentions in his posts. Describe functions and methods for analysis of target customers posts, for example, determination of personal interests or what he writes about competitors, brands or specific products. Describe functions and methods for compilation of a target client psychological portrait. Describe functions and methods for the establishment of statistical relationships and patterns in the behavior of target customers.

Literature review: The topic of the text analysis significance and importance in the literature has been discussed for a long time and the results of such analysis are widely used. Text analytics tools are used most successfully in areas such as psychology and sociology. Marketing partly use these methods, especially when it comes to consumer behavior and the impact on it. As a rule, only certain marketing methods and functions are considered[1]. A holistic approach to the study of consumer behavior is not proposed.

The scientific works that are the starting point of our research relate to the analysis of consumer comments and reviews which in particular involves the study of customer feedback and behavioral reactions[2-5]; branding and interconnected processes but as a rule, from the standpoint of the text tonality and its emotional coloring[6-8]; analysis of consumer behavior, in particular personality differences, linguistic and psychological characteristics of the individual but taking into account the types of personality[9-12]; customer lifecycle but mainly in terms of predictive analytics[13]; content marketing which involves the use of the widest possible list of existing text analysis functions[14-16]; social media marketing which is the main direction of text analysis, including not only text analysis but also video, audio and graphic information[17, 18]; segmentation, usually based on text classification or categorization[19]; positioning[20]; targeting[21]; impact analysis of the text on the target audience which is mainly used in advertising[22]; research on consumer preferences and needs, usually based on expressed interests and opinions[23]; customer service level analysis and loyalty[24]. In this case, we undertake to expand the possibilities of using text analytics and apply them to a wider range of marketing tasks.

Table 1: List of text analytics software that were studied in this research

LIWC	SAS text miner	IBM watson	Meaning cloud	MEM	Lexalytics
SPSS modeler	Amazon	Google cloud	Rosette text	Cat coding	Angoos
text analytics	comprehend	natural language	analytics platform	analysis toolkit	knowledge reader
Twinword	MonkeyLearn	NVivo	Luminoso	DiscoverText	MaxQDA
Aylien	Atlas.ti	Intellexer	GATE	Alceste	Clustify
Eagle online	Knime	Full text mapper	Cogito discover	Chinese text analyzer	Leximancer
Linguamatics	Loop Q	QDA miner	Yoshikoder	Keatext	Voyant tools
Textometrie	Tagtog	Wordle	Ikanow	S-EM	LingPipe
VisualText2.0	Wmatrix	TagCrowd	Power text solution	HyperPo	Tisane
Sketch engine	KH coder	TAMS analyzer	Textable	Orange kanvas	TokenX

It should be noted that the narrower the marketing task is, the less scientific publications are there. In some applied areas of marketing, text analytics tools are not used and adapted options are also not developed.

MATERIALS AND METHODS

The task of text analysis was facilitated by the massive introduction of information technology in research practice. This allowed to create applications for automatic text analysis. One of the first applications was Linguistic Inquiry and Word Count (LIWC, pronounced the same as "Luke"). The main idea of the application is that if a person often uses certain words and speaks on a specific topic, this reflects his psychological characteristics and characterizes him as a person. For example, if a person is unhappy, then most of the words he uses will refer to the topic of unhappiness. The program counts the frequency of words that relate to a topic or category. But the application does not take into account the context and requires the presence of certain sets of words to draw conclusions. This was partly resolved using the Meaning Extraction Method (MEM). The meaning of the method is that it automatically determines the words that are used together and naturally make up a certain topic. The algorithm compares them with categories and dictionaries of the program or forms a new category. The method was implemented in the program Meaning Extraction Helper (MEH).

At the moment, there are many programs for automatic text analysis (we will not consider libraries and modules Python, R, Java as they require programming skills). The implementation of these programs includes many text processing algorithms. At the announcement of each program, developers indicate a set of functions and a range of tasks that it performs. Despite the differences in functionality, one trend can be traced a synthesis of effective statistics algorithms, computer linguistics and machine learning (including deep learning and neural networks).

To obtain meaningful results, we tested >50 text analysis software products. Below are the most popular commercial and non-commercial software (or companies developing a software product) that we studied:

For more information, see the list of software on the resources of Digital Research Tools Wiki, TAPoR, KDNuggets.

These programs generally perform similar functions or complement each other. The difference between them as a rule, lies in the details and nuances of work as well as inherent algorithms. Different software solutions can perform the same function but work on algorithms of different efficiency. For this reason as a rule, researchers choose those software solutions that are known, proven effective, have significant scientific results and belong to large companies or open source projects. Important is the presence in the program of a marked-up and empirically substantiated dictionary.

RESULTS AND DISCUSSION

Based on the research we prepared full list of the text analytics features. These functions are a software implementation of text processing methods and algorithms. Functions that relate solely to text preprocessing, such as stemming, lemmatization or tokenization were excluded from consideration (Table 2).

The study of software and its functions led us to the conclusion that there are highly specialized functions of text analysis and there are quite general ones. Therefore, for the initial analysis of the text it is necessary to choose a program that covers the widest possible range of functions and can necessarily carry out the following functions. For convenience, we have presented them in the form of a Fig. 1. As a rule, this is enough to understand in which direction you need to move on.

However, if there is a need to use text analytics to solve marketing problems, the list of functions should be supplemented with such mandatory functions as psychometric text analysis, predictive analysis and categorization.

Thus, software for performing marketing tasks should have such functions as text classification including sentiment analysis, topic modeling, language detection, categorization and predictive analysis and text extraction including keyword extraction, entity analysis, summarization as well as psychometric text analysis.

We gave examples of text analysis functions use in marketing. This list is not exhaustive and will be constantly expanded. However, summarizing the data in table (regarding the application of software functions in marketing), we can conclude that the above functions bring the business closer to understanding the client and his behavior. These functions are ideally suited to the following areas of research in marketing: targeting, positioning, segmentation, advertising, promotion, distribution, branding, customer lifecycle management, consumer behavior management. This is due to the fact that, knowing the features of various consumer groups, corrections and a more accurate impact on a person are possible in the process of decision making which essentially results in demand management through text analysis. Separately, studies of competitors and their marketing activities through text analysis should be highlighted. As well as such private research areas as digital marketing, social listening, influencer marketing, SEO, evaluating marketing impact and performance, customer service, media marketing, event marketing, content marketing.

Table 2: Key functions summary of text analytics software (based on examination of 54 Apps listed in Table 1)

Features	Descriptions	Suggested application in marketing/customers research
Boolean queries	A search type that allows you to combine words with AND or, NO logical operators to create more relevant queries. This limits the search results to only those documents that contain a logical expression, that is two or more keywords	It has the widest application as it is designed to create logical search conditions and filter information in databases. It is mainly used to reduce the processed amount of data and processing time
Document filtering	Document filtering refers to the process by which the system monitors the flow of incoming documents, classifies them according to their content. Then selects those that are considered relevant to a particular user or topic. This function allows to filter out irrelevant information and organize relevant information into relevant categories	The function is similar to the previous one with the difference that it filters and classifies documents based on their content or thematic focus. It has the widest application. It is mainly used to reduce the processed amount of data and processing time
Language detection	Identification of the language in which the text is written. Usually using text classification algorithms	It is relevant in cross-cultural marketing research when working with documents in various languages. Defining the language, allows to choose the necessary dictionary for further research. Also used in cross-cultural research in advertising, promotion and distribution
Sentiment analysis	An analysis of the text tonality which is intended for the automated identification of emotionally colored vocabulary in texts and the assessment of authors in relation to the objects discussed in the text	Allows to determine the relationship of the writer or speaker to products, companies or events. In part, to clarify the psychological profile of a person. Allows to set the reaction of the consumer to the action. Used to develop strategies in advertising, promotion, targeting, segmenting, positioning
Summarization	The process of reducing the size of the text without losing meaning or finding a subset of the data that contains the information of the entire set, a kind of creating a representative sample or resume	It has the widest application. Usually used to obtain a brief description (set of keywords) of customers, competitors, influencers, reference groups, communities, products or brands
Tagging	The process of labeling words or phrases. A tag or label is a kind of metadata that helps to describe an element and then find it by browsing or searching	The process of matching any word to any marketing category. For example, creating a dictionary for assigning a target client according to its description or parameters to a specific segment; creation of a dictionary for assigning competitors to different groups by description or keywords
Classification	Sometimes called text tagging or text categorization. It is the process of dividing text into organized groups. Text classifiers can automatically analyze text. Then assign a set of predefined tags or categories based on its contents. The classification of text includes topic detection, sentiment analysis, language detection. When classifying text, a document or piece of text can be assigned to one or more classes or categories	It is used to find target segments or groups of consumers by signs or classify customers by interests and reaction patterns
Topic clustering (Topic modelling)	The process of grouping the contents of a document or documents by topics or subtopics as a result of which a thematic cluster is formed, showing closely related content	It is used to form groups of users or clients who discuss related topics, express similar thoughts or have the same interests, to manage the client's life cycle, branding and control consumer behavior
Entity analysis	Analysis of entities or recognition of named entities (Named-entity recognition (NER)). The task of extracting information that seeks to find and classify references to named objects in unstructured text into predefined categories, such as names of people organizations, locations, various codes, dates, monetary values, percentages, etc.	Used to search for company names, products, places, events, dates, names of individuals, etc., their references in social networks or the media. Also for monitoring competitors
Graphical data presentation	The ability to visualize data obtained during the analysis of the text (such as a word cloud or occurrence frequency of words)	Used for a comparative frequency analysis of the brands or individual products mentions. Also used for comparative consumer analysis of attitudes towards products, competitors and marketing activities of companies. Word clouds are convenient for setting keywords with which consumers describe events, promotions, actions, products
Categorization	The general direction of text categorization is also called the task of clustering. The process is very similar to classification with the difference that the boundaries of the categories are fuzzy compared to the boundaries of the classes and are established not by formal signs but by comparing the categories with each other	It is used to form generalizing features in order to create categories, for example, to create and specify categories or types of customers/buyers; definitions of signs and patterns of behavior that are common to several consumer groups; finding common signs in communities on social networks
Extraction	The name of the general direction for extracting text or words from a document.	Extract any relevant marketing information from unstructured text
Predictive modeling	The function makes it possible to create predictive analytics models. This means predicting, for example, whether the tag/label or text belongs to a particular topic, whether the	Usually used to correlate a client (usually a new one) or a consumer, as well as interests, posts and opinions with a previously known group or category

Table 2: Continue

Features	Descriptions	Suggested application in marketing/customers research
	extracted word or part of the text belongs to a specific label whether the extracted text matches the query or whether the content matches the query	
Aspect-based sentiment analysis	It determines not the general tonality of the text but the various aspects of tonality for each part of the text. Provides a more detailed analysis of the text. In other words, gives out positive or negative opinions on various topics or aspects of something	Used to establish the tonality (emotionality) of the text/post in relation to any aspect, for example, positive or negative emotions in the framework of providing an additional service, adding a new product function; in relation to a promotion, event with competitors or business partners
Entity-level sentiment analysis	This analysis does not determine the overall tonality of the text but the tonality of the text relative to a specific named object	Used to identify tonality or emotions regarding names, for example, brands, trademarks, company names or events
Document-level sentiment analysis	Analysis of tonality at the document level, that is, the ability to determine the tonality of the whole document	It is used to determine the tonality of any document, for example, a file with business correspondence, a file with a competitor description or a description of a client, a file with reviews or comments on a product
Hashtags suggestion	The function of automatically offering relevant hashtags for publishing content on social networks	Used to analyze posts and automatically create hash tags
Image tagging	Tags not only text but also images found on web pages	Used to analyze posts and automatically create hash tags for images
Semantic similarity	Search not by keywords but by meaning. A keyword search returns what they said and not what they actually meant. If the words are ambiguous, this causes certain difficulties. Semantic search allows to search not by words but by meaning. It is used mainly for interlanguage search, when there may not be full lexical translatability of words. Also, to search for relevant terms and concepts or to generate them in other languages, to search for repetitions in documents, to search for plagiarism	It is used in international marketing research to search for information within the meaning as a rule, for cross-language search, when there is no complete lexical translatability of words
Subjectivity analysis	The task associated with the analysis of sentiment, the main purpose of which is to designate opinion as subjective or objective	Used to determine the subjectivity or objectivity of an opinion which is expressed in comments or posts. Convenient when working with objections and complaints, to improve the level of customer service
Psychometric text analysis	Determination of the text psychometric properties, that is, the main psychological/cognitive characteristics of the author, in particular, the degree of analytical thinking, level of leadership, degree of honesty and openness and emotional background	Used to compile a person's psychological profile, determine his personality traits and establish the parameters of target groups and segments of users or clients

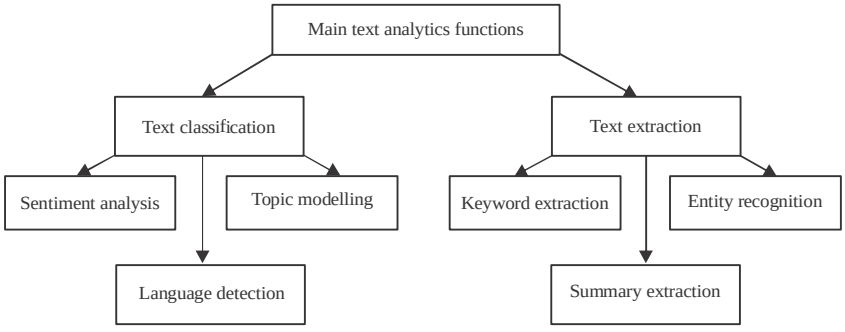

Fig. 1: Main text analytics software functions

Marketing purposes of using text analytics: We suggest the following structure of use text analytics tools for different purposes of marketing (Fig. 2).

Suggested design of comprehensive customer's decision-making research: In this part, we suggest own design of comprehensive customer's behavioral research based on the automated text analytics tools.

This research helps to find out behavioral patterns and significant factors of target customers decision-making process (Table 3).

Table 3: Suggested by authors design of marketing consumer research based on text analytics and mining tools

Research parameters: customer profile and behavior	Automated text analysis tools
Target customers identification according to the topics discussed/shared. The topics discussed is not the only one way of identification target customers. Additionally the profile parameters are important, such as location, education, gender, job occupation etc. But text analysis makes target customer identification much more precise	Classification, Topic clustering; Predictive modelling Profile analysis tools (not only text analysis); Entity analysis
Experience of customers in use of targeted products or services: how customers evaluate their experience? how customers evaluate particular brands (competitors) and particular services? are the customers interested to continue use of these products/services?	Content analysis; Entity-level sentiment analysis; Summarization; Entity analysis
Customer personal goals and values. Customer values are very important in decision-making process. Therefore they are important for marketers to adapt services/proposals to customer goals and values. The goals and values may include: Cultural/religious/political; values Individualism vs. collectivism; Stability vs. risk/new opportunities; Personal goals and ambitions; Importance of career/income/family/friends etc.; Especially for individuals sales this study help to customize marketing proposals and prepare negotiations properly	Subjectivity analysis; Psychometric analysis; Sentiment analysis; Summarization; Entity analysis
Educational, cultural and intellectual level. This parameter is important in marketing communication. People with different educational and intellectual level consider different arguments as significant in their decision-making process	Grammar check; Summarization; Clustering/Classification; Sources of citations/shared content
Persistent interests and hobbies: Shared content topics mostly interesting for target customers; Statements about hobbies and interests in discussions	Classification, Topic clustering; Sentiment analysis Summarization; Entity analysis; Clustering/Classification
SELF-concept analysis. Self-concept in psychology is a self-positioning of personality, how a person perceives and positions her/himself. This is another important factor in consumer decision-making process	Psychometric text analysis; Entity-level sentiment analysis; Summarization
Psychological patterns of reactions to the news, events etc. How the person comment news: optimistic/pessimistic/friendly/arrogant/emotional/quietly/conformist/non-conformist etc.	Subjectivity analysis; Psychometric text analysis Entity-level sentiment analysis
Explicit patterns of thinking/behavior. People often describe their behavioral/cognitive decision-making patterns explicitly, for example: Opinion of my friends is very important for me/only my own decision does matter; Optimistic/pessimistic; Rational/emotional attitude; Conformism/non-conformism; Leader/follower, etc.	Subjectivity analysis; Summarization; Sentiment analysis; Psychometric analysis
Psychological traits of personality. Not all of the psychological traits can be useful in marketing. Besides automated identification of psychological traits requires further research and verification. We consider this problem more detailed in our next research. Some of the psychological text analysis parameters can be useful for understanding customer begavioral patterns and consumer decision-making process	Psychometric text analysis; “Big 5” tests based on text analysis; MBTI test based on text analysis

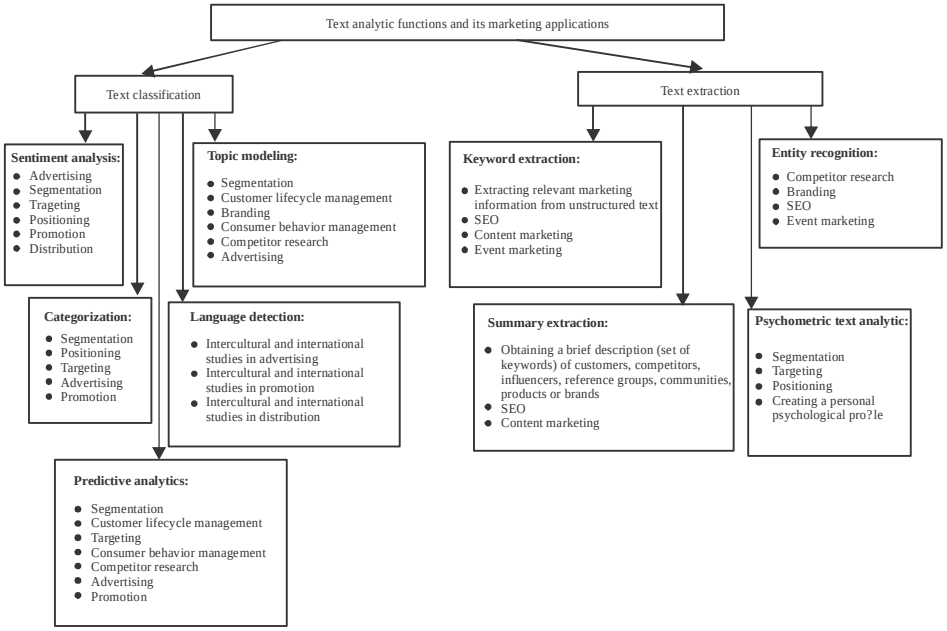

Fig. 2: Text analytics functions and its application in marketing

CONCLUSION

This research helps to identify individual patterns of decision-making process. Based on the research results individual proposals can be prepared for customers as well as customized negotiation plan can be designed. In this research, we found out that text analysis is useful to understand decision-making process of customers which is not the same as their psychological portrait. Decision-making process includes goals, values, interests, relationships. Automated text analysis allows to research this wide range of decision-making aspects. In our further research, we are going to conduct customers research according to this design and evaluate how very valuable are different decision-making patterns for marketing.

REFERENCES

01. Amado, A., P. Cortez, P. Rita and S. Moro, 2018. Research trends on big data in marketing: A text mining and topic modeling based literature analysis. Eur. Res. Manage. Bus. Econ., 24: 1-7.
02. Calheiros, A.C., S. Moro and P. Rita, 2017. Sentiment classification of consumer-generated online reviews using topic modeling. J. Hospitality Marketing Manage., 26: 675-693.
03. Gan, Q., B.H. Ferns, Y. Yu and L. Jin, 2017. A text mining and multidimensional sentiment analysis of online restaurant reviews. J. Qual. Assurance Hospitality Tourism, 18: 465-492.
04. Buschken, J. and G.M. Allenby, 2016. Sentence-based text analysis for customer reviews. Marketing Sci., 35: 953-975.
05. Ordenes, F.V., B. Theodoulidis, J. Burton, T. Gruber and M. Zaki, 2014. Analyzing customer experience feedback using text mining: A linguistics-based approach. J. Serv. Res., 17: 278-295.
06. Koltringer, C. and A. Dickinger, 2015. Analyzing destination branding and image from online sources: A web content mining approach. J. Bus. Res., 68: 1836-1843.
07. Moro, S., P. Rita and B. Vala, 2016. Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. J. Bus. Res., 69: 3341-3351.
08. Arvidsson, A. and A. Caliandro, 2016. Brand public. J. Consum. Res., 42: 727-748.
09. Leung, R., J. Rong, G. Li and R. Law, 2013. Personality differences and hotel web design study using targeted positive and negative association rule mining. J. Hospitality Marketing Manage., 22: 701-727.
10. Erevelles, S., N. Fukawa and L. Swayne, 2016. Big data consumer analytics and the transformation of marketing. J. Bus. Res., 69: 897-904.
11. Seret, A., V.S.K. Broucke, B. Baesens and J. Vanthienen, 2014. A dynamic understanding of customer behavior processes based on clustering and sequence mining. Expert Syst. Appl., 41: 4648-4657.
12. Humphreys, A. and R.J.H. Wang, 2018. Automated text analysis for consumer research. J. Consum. Res., 44: 1274-1306.
13. Moro, S., P. Cortez and P. Rita, 2015. Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns. Neural Comput. Appl., 26: 131-139.
14. Paschen, J., M. Wilson and K. Robson, 2020. #BuyNothingDay: Investigating consumer restraint using hybrid content analysis of Twitter data. Eur. J. Marketing, 54: 327-350.
15. Tellis, G.J., D.J. MacInnis, S. Tirunillai and Y. Zhang, 2019. What drives virality (sharing) of online digital content? The critical role of information, emotion and brand prominence. J. Marketing, 83: 1-20.
16. Kietzmann, J. and L.F. Pitt, 2020. Computerized content analysis of online data-opportunities for marketing scholars and practitioners. Eur. J. Marketing, 54: 473-477.
17. Salloum, S.A., M. Al-Emran, A.A. Monem and K. Shaalan, 2017. A survey of text mining in social media: Facebook and Twitter perspectives. Adv. Sci. Technol. Eng. Syst. J., 2: 127-133.
18. Appel, G., L. Grewal, R. Hadi and A.T. Stephen, 2020. The future of social media in marketing. J. Acad. Marketing Sci., 48: 79-95.
19. Pitt, C.S., A.S. Bal and K. Plangger, 2020. New approaches to psychographic consumer segmentation: Exploring fine art collectors using artificial intelligence, automated text analysis and correspondence analysis. Eur. J. Marketing, Vol. 1, 10.1108/EJM-01-2019-0083
20. Moon, S. and W.A. Kamakura, 2017. A picture is worth a thousand words: Translating product reviews into a product positioning map. Int. J. Res. Marketing, 34: 265-285.
21. Provost, F., B. Dalessandro, R. Hook, X. Zhang and A. Murray, 2009. Audience selection for on-line brand advertising: Privacy-friendly social network targeting. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, June 28-July 1, 2009, ACM, Paris, France, pp: 707-716.
22. Berger, J., A. Humphreys, S. Ludwig, W.W. Moe, O. Netzer and D.A. Schweidel, 2020. Uniting the tribes: Using text for marketing insight. J. Marketing, 84: 1-25.
23. Timoshenko, A. and J.R. Hauser, 2019. Identifying customer needs from user-generated content. Marketing Sci., 38: 1-20.
24. Muller, O., I. Junglas, S. Debortoli and J.V. Brocke, 2016. Using text analytics to derive customer service management benefits from unstructured data. MIS Q. Executive, 15: 243-258.