



Leszek Ziara

Politechnika Częstochowska
Wydział Zarządzania
Katedra Informatyki Ekonomicznej
ziara@zim.pcz.pl

THE SENTIMENT ANALYSIS AS A TOOL OF BUSINESS ANALYTICS IN CONTEMPORARY ORGANIZATIONS

Summary: Nowadays contemporary organizations apply different business analytics tools such as sentiment analysis for the purpose of business functionality improvement and support of decision making processes. Sentiment analysis also called opinion mining allow for gathering and analysis of opinion concerning particular product or service. The aim of the paper is to present the notion of sentiment analysis and its areas of application in contemporary organizations. It also presents practical examples and case studies concerning sentiment analysis application in different areas of business activity.

Keywords: sentiment analysis, business analytics, machine learning, data mining, big data.

Introduction

Contemporary organizations in its business activity apply many methods, techniques and tools of business analytics. Such solutions allow to improve its general functionality and particularly to improve its business and decision making processes at all levels of management. Sentiment analysis solution is part of business analytics and especially text analytics as well as it belongs to the areas of natural language processing and computational linguistics. It may have direct impact on gaining profits and achievement of competitive advantage of contemporary enterprise on the worldwide market.

1. The notion and benefits of sentiment analysis application

B. Liu presents the notion of sentiment analysis which is also called opinion mining as “the field of study that analyzes people’s opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes” [Liu, 2012, p. 7]. B. Agarwal and N. Mittal claim that “the automatic analysis of online contents to extract the opinion requires deep understating of natural text by the machine. Sentiment analysis research can be categorized among document level, sentence-level, and aspect-feature-level sentiment analysis” [Agarwal et al., 2016, p. 1]. They further state that “document-level sentiment analysis classifies a review document as containing positive or negative polarity. It considers a document as a single unit. Sentence-level sentiment analysis takes a sentence to extract the opinion or sentiment expressed in that sentence” [Agarwal et al., 2016, p. 1].

Medhat et al. conducted a survey of which an aim was to review 54 articles from recent years concerning the area of sentiment analysis. Their survey provided a categorization of techniques applied in the area of sentiment analysis and is presented in Fig. 1. According to this categorization the sentiment analysis embrace multiple of approaches, methods, techniques and tools such as lexicon-based approach which can further be divided into dictionary-based approach and corpus-based approach which is then subdivided into statistical and semantic one. Another branch of sentiment analysis division embraces machine learning approach which is consequently divided into supervised and unsupervised learning. Supervised learning includes decision tree, linear (including support vector machines and neural network), rule-based and probabilistic classifiers which may be subdivided into naive Bayes, Bayesian network and maximum entropy techniques.

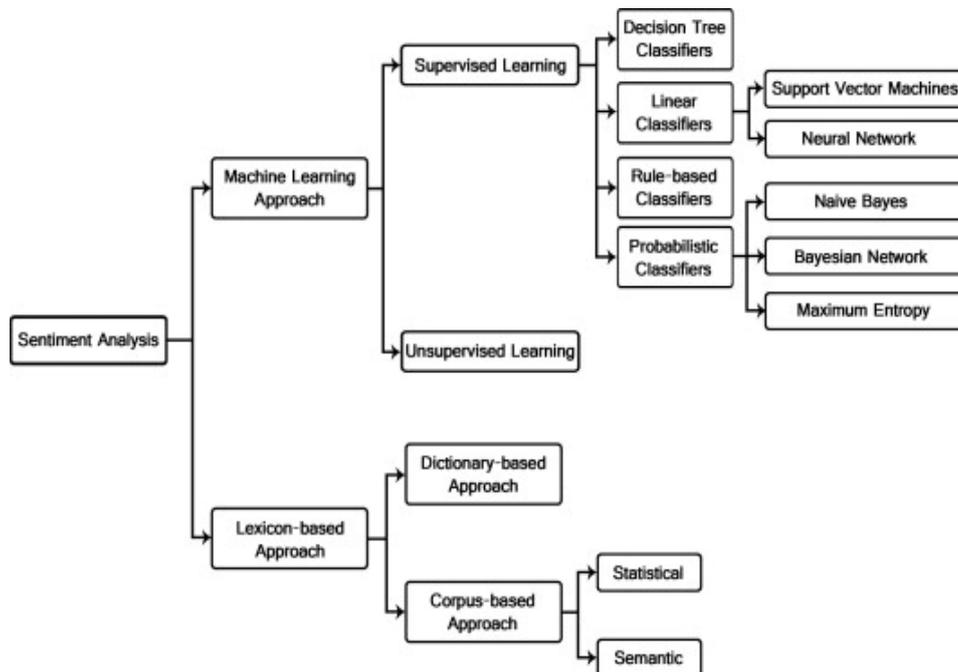


Fig. 1. Techniques of sentiment classification

Source: [Medhat et al., 2014].

Zhai and Zhang mention the bag of words (BoW) models where “in machine learning, documents usually represented as bag of words reduces a piece of text with random length to a fixed length vector and claim that bags of words remain the dominant representation in many applications including text classification” [Zhai and Zhang, 2015]. In the bag of words the text used for document classification is divided into separate words without taking into consideration grammar rules or word order. An example of such a model may be the sentence “He participated in the creativity conference” which is separated into particular words such as: “he”, “participated”, “in”, “the”, “creativity”, “conference”. In the Fig. 2 there is presented an exemplary sentiment analysis for the mentioned above sentence with the application of TheySay solution. Conducted analysis indicated the polarity of the sentence as positive in 84,2 %, neutral in 15,8% and negative in 0%. Emotion detection that was provided here indicated positive aspects in the following categories: calm, happy, like, sure. The other categories which were analyzed include named entity recognition, language detection, gender detection, text summarization, part-of-speech tagging, and dependency parsing.

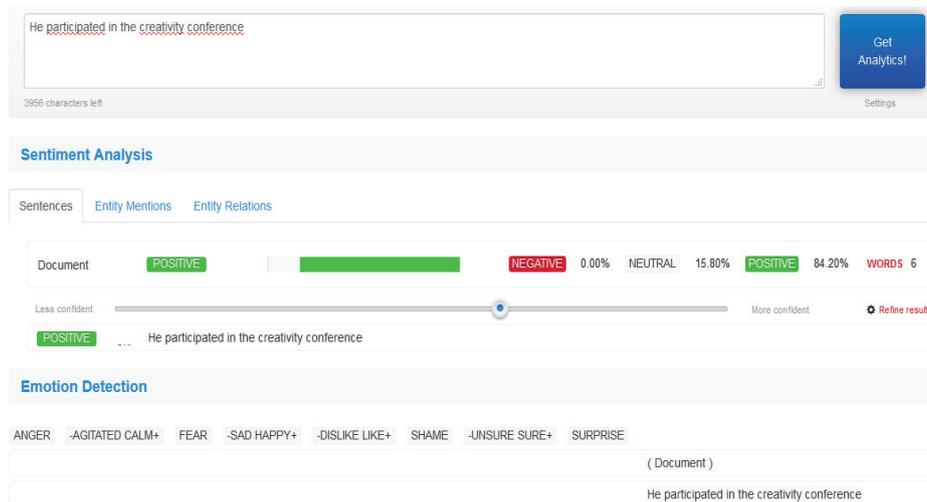


Fig. 2. An example of sentiment analysis

Source: <http://apidemo.thesay.io/#sentimentSentencesTab>.

As far as the benefits resulting from application of sentiment analysis in contemporary company is concerned it is worth noticing the study of Ribeiro et al. who mention applications of sentiment analysis in such areas as: “monitoring the reputation or opinion of a company or a brand with the analysis of reviews of consumer products or services, provision of analytical perspectives for financial investors who want to discover and respond to market opinions, its application in politics, where marketing campaigns are interested in tracking sentiments expressed by voters associated with candidates” [Ribeiro et al., 2016]. SAS Company on the basis of its solution mentions such benefits as: “evaluation of sentiment and monitoring of changes over time, continuous improvement of customer experience and competitive position, identification of feedback sources in order to define new targets, promotion of discovery with a closed-loop, integrated analysis environment” [SAS, 2016]. Sentiment analysis can be used in multiple areas such as economics, finance and marketing. In economics e.g. allows to respond to the question of how supervised learning methods can be used to learn the association between polarity of financial news and key financial indicators [Ahmad, 2011]. It can also be applied in sociology for instance regarding the issue of how ideas and innovations diffuse [Pang and Lee, 2008]. Exemplary tools used in sentiment analysis may include: Google analytics, meltwater, people browser, hootsuite, tweetstats, marketing grader, social mention and pagelever.

2. Practical examples

Sentiment analysis can be applied in many areas of business activity of contemporary organizations. Its selected areas of application include product and services review, movies and books reviews, opinion on restaurants and blogs. Hamdan et al. present its application in scholarly book reviews where “supervised classifiers extract the opinion target expressions and detect their polarities” [Hamdan et al., 2016]. The authors as they claim “have constructed a corpus of book reviews, segmented each review into sentences and asked three annotators to extract the opinion targets and their polarities in each sentence. They trained a conditional random field (CRF) model for the purpose of opinion target extraction and a logistic regression one for sentiment polarity” [Hamdan et al., 2016]. In order to perform this analysis the authors of the study “automatically segmented each review into sentences in order to annotate each sentence using Talismane syntax analyzer. In the analysis they listed such aspects as: book presentation, problematic, scientific context, scientific method, author’s arguments, book organization, judgment about the book” [Hamdan et al., 2016].

Thakkar and Patel mention the conduct of sentiment analysis of Twitter posts by such techniques as lexical analysis, machine learning based analysis and hybrid analysis. They compared all the mentioned approaches and state that lexical analysis “use a dictionary consisting of pre-tagged lexicons. The input text is converted to tokens by the Tokenizer. Every new token encountered is then matched for the lexicon in the dictionary. If there is a positive match, the score is added to the total pool of score for the input text. The classification of a text depends on the total score it achieves” [Thakkar and Patel, 2015]. In case of machine learning the authors present the following steps involved in it such as: data collection e.g. from various sources like Twitter, blogs and social networks, pre-processing where acquired data is cleansed and prepared for feeding into classifier; training data which will be fed to the algorithm for learning purpose; classification where support vector machines and naive Bayes are applied for analysis purpose. The classifier after training completion can be deployed to the real time tweets/text for sentiment extraction purpose; where results are provided in the form of graphs or charts. The hybrid analysis combines those two approaches. According to the authors the “best results have been observed from machine learning approaches, and least by lexical approaches” [Thakkar and Patel, 2015].

The sentiment analysis of tweets was conducted by Fornacciari et al. where authors in their research “built a system for social network and sentiment analy-

sis operating on Twitter data. Twitter allows writing messages of 140 characters in length containing opinions, thoughts, facts and links to images and other multimedia files. The authors in their analysis collected three types of data: the user type representing users' profiles such as: *user_id*, *name*, *location*, *num_followers*, *num_tweets*; posted messages data: *tweet_id*, *user_id*, *message*, *date* and the friend type representing the "follow" relationships among users. After applying different filters the authors analyzed tweets for discriminating the basic classes of objectivity and polarity at two levels. The training set for classifiers included in their case study 86000 instances (polarity) and 32000 instances (subjectivity) and they were obtained by exploration of over 60 channels. The obtained results by the classifiers showed an accuracy of: 77,45% (polarity classifier) and 79,50% (subjectivity classifier). These results show that the model of the classifiers contains effective features for the recognition of the sentiment of a message. The case study considered in this work is the social network of the #SamSmith channel (the singer who won four awards at the Grammy Awards 2015)" [Fornacciari, 2015]. Another worth mentioning application of sentiment analysis is its application on YouTube. Asghar et al. presented "sentiment analysis using YouTube in order to find the polarity of user comments. Three issues included in their survey framework were event classification, detection of sentiment polarity, and predicting YouTube comments. The analyzed comments may include short syllable comments, advertisements of any kind, negative criticism, rambling arguments [Asghar et al., 2015].

3. Review of the case study

One of the examples of sentiment analysis application in contemporary organization is its application in Unilever company which is the corporation selling "products ranging from personal care and homecare items to packaged foods in more than 150 countries and controls subsidiaries in at least 90 countries worldwide" [Unilever..., 2016]. The goal of conducted in this company sentiment analysis was "accuracy improvement of customer sentiment analysis and efficiency maintenance of this automated process" [Unilever..., 2016]. The authors of the report state that "Unilever needed to monitor online feedback associated with the launch of its Dove Men+Care line of products. The applied solution to the problem was that a round-the-clock team of human reviewers with the appropriate language skills analyzed sentiment of customers around the globe. What is more the statistical quality control technology and automated workflow management streamlined project delivery. The result of sentiment analysis ap-

plication was that the percentage of content in which positive or negative sentiment was identified rose from 15% to > 95% and the accuracy rate rose from less than < 30% to over > 90%” [Unilever..., 2016].

Another case study of sentiment analysis application is Beyond the Arc Company which is dealing with e.g. data science, customer experience, strategic marketing, brand strategy, content marketing and social media. The data team of mentioned company “uses text mining tools to monitor negative and positive comments and tracks acceptable ranges for each over a period of time. They state that “if the percentage of total comments falls outside of set ranges, they look for themes and categorizations by analyzing comments captured over this timeframe. Using this method, there can be taken advantage of sentiment analysis to: gain insights from social media conversations, build a sentiment index to measure company and product reputation, develop a strategy for responding to negative sentiment to positively impact perception and identify and engage top influencers to reinforce desired messages” [Taylor, 2011].

Summary

Sentiment analysis as a part of business analytics and machine learning solutions brings multiple advantages when applied in different areas of enterprise’s business activity. It allows for performing analyses concerning different reviews of products or services, monitor online feedback and so on. It is a tool which is still being developed and new areas of its application still arise.

Literature

- Agarwal B., Mittal N. (2016), *Prominent Feature Extraction for Sentiment Analysis*. Springer, “Socio-Affective Computing”, Vol. 2, Switzerland.
- Ahmad K. (ed.) (2011), *Affective Computing and Sentiment Analysis. Emotion, Metaphor and Terminology*, Springer, Dublin.
- Asghar M., Ahmad S., Afsana M., Kundi F. (2015), *Sentiment Analysis on You Tube: A Brief Survey*, <http://arxiv.org/ftp/arxiv/papers/1511/1511.09142.pdf>.
- Fornacciari P., Mordonini M., Tomaiuolo M. (2015), *A Case-Study for Sentiment Analysis on Twitter*, Proceedings of the 16th Workshop “From Object to Agents” (WOA15), June 17-19, Naples, Italy, <http://ceur-ws.org/Vol-1382/paper8.pdf>.
- Hamdan H., Bellot P., Bechet F. (2016), *Sentiment Analysis in Scholarly Book Reviews*, <http://arxiv.org/abs/1603.01595v1> (access: 4.03.2016).
- Liu B. (2012), *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>.

- Medhat W., Hassan A., Korashy H. (2014), *Sentiment Analysis Algorithms and Applications: A Survey*, "Ain Shams Engineering Journal", <http://www.sciencedirect.com/science/article/pii/S2090447914000550>.
- Pang B., Lee L. (2008), *Opinion Mining and Sentiment Analysis*, "Foundations and Trends in Information Retrieval", Vol. 2, No. 1-2, pp. 1-135, <http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>.
- Ribeiro F., Araujo M., Goncalves P., Goncalves M., Benevenuto F. (2016), *A Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods*, <http://arxiv.org/abs/1512.01818v3> (access: 1.02.2016).
- SAS Sentiment Analysis. Align with Your Customers' Desires* (2016), http://www.sas.com/en_us/software/analytics/sentiment-analysis.html (access: 04.2016).
- Taylor S. (2011), *Using Sentiment Analysis to Improve Customer Experience*, <http://beyondthearc.com/blog/2011/customer-experience/using-sentiment-analysis-to-improve-customer-experience#sthash.pKXKjoak.zN9NnJKg.dpuf>.
- Thakkar H., Patel D. (2015), *Approaches for Sentiment Analysis on Twitter: A State-of-Art study*, <http://arxiv.org/ftp/arxiv/papers/1512/1512.01043.pdf> (access: 3.12.2015).
- Unilever Case Study Sentiment-Analysis* (2016), <https://success.crowdfunder.com/hc/en-us/articles/203510355-Case-Study-Unilever-Sentiment-Analysis> (access: 04.2016).
- Zhai S., Zhang Z. (2015), *Semisupervised Autoencoder for Sentiment Analysis*, <http://arxiv.org/abs/1512.04466v1> (access: 14.12.2015).

ANALIZA OPINII JAKO NARZĘDZIE ANALITYKI BIZNESOWEJ WE WSPÓŁCZESNYCH ORGANIZACJACH

Streszczenie: Obecnie współczesne organizacje stosują różne narzędzia analityki biznesowej, takie jak analiza opinii w celu ulepszenia funkcjonalności biznesowej i wspierania procesów decyzyjnych. Analiza opinii pozwala na zebranie i analizę opinii dotyczących poszczególnych produktów i usług. Celem niniejszego artykułu jest prezentacja znaczenia analizy opinii i jej obszarów zastosowań, w tym korzyści wynikających z tych aplikacji we współczesnych organizacjach. Artykuł przedstawia także praktyczne przykłady i studia przypadków dotyczące zastosowań tejże analizy w różnorodnych obszarach działalności biznesowej.

Słowa kluczowe: analiza opinii, analityka biznesowa, data mining, big data, uczenie maszynowe.