Sentiment Analysis of Web Trends for the Antisocial Behaviour Detection

Kristína Machová[1], Ján Birka[1]
[1]*Department of Cybernetics and Artificial Intelligence, Technical University of Košice, Letná 9, Košice, Slovakia*
kristina.machova@tuke.sk, jan.birka@student.tuke.sk

Keywords: Sentiment Analysis, Web Trends, Antisocial Behaviour, Online Discussion, Lexicon Approach.

Abstract: The paper presents an approach to extraction of current web trends for research into automated recognition of antisocial behaviour in online discussions. Antisocial behaviour is a drawback of online discussions as compared to their advantages such as wisdom of crowds and collective intelligence. The first step to recognition of antisocial behaviour is the identification of web trends connected with it. These are studied in dynamic conditions using sentiment analysis as a webometric. A new sentiment analysis method based on a lexicon was developed. Two modifications of the lexicon sentiment analysis method were designed and tested involving NLP (natural language processing) and an original technique for negations and intensifications processing. The most effective sentiment classification method was used for the extraction of web trends. Extracted web trends were analysed in a dynamic way and findings of this analysis were compared to known historical events.

1 INTRODUCTION

Social web platforms enable web users to share their knowledge or ideas and express their opinions and attitudes to various themes. Online media are an inevitable part of modern life. They have a lot of positive, but unfortunately, also many negative effects on our lives. Examples of positive use of social media are: social connectivity, education, getting help or up-to-date information, helping to prevent crime, building communities, etc. However, other positive influences of the online space are starting to attract users' attention, namely wisdom of crowds and collective intelligence. We can say that so called "discourse content" (created in online discussions) represents an instantiation of wisdom of crowds in a "rough" form of data suitable for extraction of useful knowledge from the summarized opinions on an important event and monitoring of current web trends.

We have decided to use sentiment analysis for extraction of summarized opinions and consequently for capturing of current web trends. We have also used sentiment analysis for extraction of information about changes of crowd opinions throughout time. So, our aim is a dynamic analysis of web trends. Our sentiment analysis approach is based on a lexicon. To improve results, we used a technique of natural language processing for recognition of word classes (parts of speech) and for the same purpose we have experimented with various approaches to processing negation and intensification.

Examples of the negative impact of social media on our modern society are various forms of antisocial behaviour like trolling, fake news, hoaxes, hacking, rumours, social spamming, hate speech etc. (Ahmad, 2010). Because of the existence of antisocial content in online discussions, some forms of regulation of online media posting should be introduced. Our research is oriented on new methods for the detection of antisocial behaviour in online media. A component part of this research is capturing current web trends, because these web trends are probably influenced by antisocial posts.

2 ASPECTS OF ONLINE DISCUSSIONS

Nowadays, the Web offers a wide spectrum of applications, which enable us to share knowledge and experiences of a whole community of users and in this way to create collective intelligence and wisdom of crowds. They are highly successful and popular within information technologies. But they also have drawbacks namely connected with security problems and the protection of web user privacy.

2.1 Collective Intelligence

Collective intelligence is the shared intelligence of a whole group. It arises in competition or cooperation of many people during the process of searching for a solution or consensus in complex problems. We assume that no individual knows everything but every individual knows something and after a suitable aggregation of knowledge we can obtain a form of extensive collective intelligence (Malone, 2006).

It is clear, that collective intelligence existed before the existence of information technologies, for example in communities such as families, nations, armies, etc. Typical examples of systems based on collective intelligence are Wikipedia, Google products, web discussions and forums, blogs, etc. Using these platforms, human communities manage to act with much higher intelligence as before (Lévy, 1997).

The same principles, which enable collective intelligence, can enable collective stupidity in the case when people blindly believe in the opinions of some users and follow their antisocial behaviour.

2.2 Wisdom of Crowds

According to (Surowiecki, 2004) wisdom of crowds represents a process of aggregation of anonymous data to find a wisdom, which results from an opinion estimation of a great number of people without any mutual influence among them. Four basic principles of wisdom of crowds are the following: diversity of opinions, independence of evaluations of individuals, decentralization (nobody will dictate his/her own opinion) and aggregation into a collective decision. In real life, it is not possible to ensure the principle of independence of evaluations of community members. The mutual influence of members of community can leads to group thinking and tolerance of antisocial behaviour.

2.3 Antisocial Behaviour in Online Communities

Antisocial behaviour is connected with disinformative content and may be of a dual nature. First, it can represent information that will affect and manipulate its recipients (fake news, hoax). Second, it is misinformation, which is caused by misunderstanding without manipulation (Kumar, 2016). The first is based on disseminating propaganda, which tries to make reality relative by generating arguments that distort the truth. Sometimes this truth distortion could be generated automatically using algorithms based on similarity measures (Wang, 2013).

There are various disinformation techniques. These techniques are discussed in (Řimnáč, 2018), which presents a probability approach to the detection of relativized statements.

Opinion sharing by product reviews is a part of online purchasing. This opinion sharing is often manipulated by fake reviews. The paper (Dematis, 2018) presents an approach which integrates content and usage information in fake reviews detection. The usage information is based on reviewers' behaviour trails. In this way, a reviewer's reputation is formed.

3 USED METHODS

3.1 Text Mining

The extraction of knowledge from texts is a complex problem. Its complexity stems from the fact, that texts of the discourse content are unstructured and uncertain. From such texts a new piece of knowledge should be extracted. The new knowledge has to be unknown until now, potentially useful, and valid in the statistical meaning. There are a great number of methods for text mining, for example statistic methods, methods of supervised machine learning, cluster techniques and also techniques of natural language processing (Jurafsky, 2017).

3.2 Natural Language Processing

Natural language processing (NLP) is, together with expert systems, one of the most advanced applications of artificial intelligence. It can be applied on a written as well as a verbal form of language. Our work focuses on the written form to analyse expressed opinions.

The techniques of natural language processing represent a different approach to the techniques of mining knowledge from texts. According to (Kao, 2005) the differences are the following:

- The techniques of NLP are oriented on a language. A text is analysed using information about the formal grammar and dictionaries.
- The mining of knowledge from texts uses techniques of information retrieval, statistics and machine learning methods. The goal of it is not to understand the meaning of a text but to

extract important patterns from great number of documents.

Nowadays, the utilization of NLP in computer systems grows in many domains. In this work, NLP is used to increase the effectiveness of our method for sentiment analysis.

3.3 Sentiment Analysis

Today, there is a growing interest in sentiment analysis (SA), not only because it has a wide and perspective potential in real applications but also because it can solve more drawbacks of NLP. SA is beneficial for marketing, research, artificial intelligence, computer linguistics and also for social psychology.

We consider sentiment analysis to be the most important webometric. Other webometrics like social networks analysis (SNA) and mention analysis (MA) are not used in this work, because SNA works with graphs of communication instead of with texts of posts and MA is too simplistic. Webometrics analysis is quite a new research discipline. It uses statistical methods for research in the area of World Wide Web (Thelwall, 2005).

During SA, it is important to take into account information about the kind of users whom SA targets. Individual users and societies could have a slightly different view on the text data. For example, the sentence "Prices of mobiles are decreasing lately." has a negative meaning for companies specializing on mobiles marketing. But the same sentence has a positive meaning for users planning to buy a new mobile. SA can be helpful for a common user because it can evaluate a great amount of information, reviews and opinions on a product in an automatic way.

The methods of SA can be divided into two main groups: lexicon approaches and machine learning approaches. More about sentiment analysis methods can be found in (Machova, 2018).

The lexicon based sentiment analysis hypothesizes, that sentiment estimation is a function of an algorithm, data sample and external knowledge, for example in the form of a lexicon. A lexicon contains words together with weights for positive or negative evaluations of each word. Such a lexicon can be created manually, but there are also some semi-automatic approaches to generation of it.

The algorithm of SA searches for words from the lexicon in the analysed text. If a given word is presented in the text, the value of its weight (negative or positive) is extracted from the lexicon. If more words from the lexicon are presented in the text, all weights of these words are inputs to a function for computing a result value of sentiment for the whole text.

4 WEB TRENDS ANALYSIS

A concept trend is defined as a development direction or a tendency to change something interesting for people over time. Trends can appear in fashion or economics but also in technologies, particularly in web technologies. Our goal is to study some web trends over time and analyse its development. We say that some product or event can be considered a trend, when many comments and posts about it can be found in online discussions. We were interested in the nature and polarity of these opinions and their change in time and so we have selected sentiment analysis as the method for web trend analysis, because sentiment analysis naturally involves an opinion analysis. The process of web trends analysis consisted of the following steps.

- Selection of a web discussion
- Extraction of text data from the web discussion
- Text pre-processing
- Lexicon building
- Implementing the lexicon approach to SA
- Testing the lexicon approach to SA
- Improvement of the lexicon approach to SA involving possibilities of NLP
- Testing the improved approach
- Experiments with various methods of negation and intensification processing
- Selection of the best way of negation and intensification processing
- The second improvement of the lexicon approach with NLP involving the selected methods for negation and intensification processing
- Testing of the second improvement
- Results analysis

Within our lexicon approach to sentiment analysis (DASA), a new lexicon was generated. The lexicon has to contain a value of opinion polarity for each word in the lexicon. The values of word polarity can be generated automatically (Mikula, 2017) or the lexicon can be derived from a lexical resource (Baccianella, 2010).

Our lexicon for DASA was derived from SentiWord Net 3.0. The lexicon contains sets of synonyms and their values of opinion polarity. The total number of words included is 117659. It also needed to be cleaned of needless words and information using Pars.py. Table 1 illustrates the form of the final derived lexicon.

Table1: Illustration of the lexicon derived from SentiWordNet 3.0.

Word	Positive weight	Negative weight
conceptual	0,375	0,25
easy	0,25	0,625
unacceptable	0,125	0,375
Too-bad	0,222	0,778

4.1 Experiments with DASA

The DASA algorithm searches for words from the analysed text which are presented in the derived lexicon. In the case of a match, the value of its weight (negative or positive) is extracted from the lexicon and added to a sum of opinion polarity values. If the resulting sum after processing all words from the text is positive (negative) the opinion polarity of the whole text is positive (negative).

At first, experiments with the basic DASA method were performed. The results are presented in Table 2.

Table 2: Results of tests of DASA approach in known measures Precision, Recall, F1-measure and Accuracy.

Opinion	Precision	Recall	F1	Accuracy
Positive	0,554	0,593	0,573	-
Negative	0,562	0,523	0,542	-
Average	0,558	0,558	**0,558**	**0.558**

All experiments with DASA and other modifications of it (DASA Involving NLP and a new method of negation and intensification processing) were performed on Movie Reviews data (csfd.cz). The data obtained texts of reviews on movies. The data was pre-processed according to the CRISP-DM methodology (Paralič, 2010). The data were manually annotated. Using a confusion matrix, several indicators of binary classification efficiency were quantified in the process of testing. The results of testing were poor and the processing time was too high.

To improve the DASA approach, we decided to utilize possibilities of natural language processing (NLP).

4.2 DASA Involving NLP

In NLP, the analysed text was partitioned into words or morphemes according to rules of morphologic analysis. Our work was oriented on English, because English has quite simple and regular morphology. We assumed that relations between words are represented above all by word-order. A key part of the systems of natural language processing and also the sentiment analysis system is a module for parsing.

The parser decomposes a sentence into words and consequently assigns a word class to each word. It enables us to generate a parsing tree – a structure for extraction of the meaning of a sentence. A reoccurring problem was shape homonymy, which appeared when one word could have multiple different word classes. In our approach to sentiment analysis, these homonyms were processed in the following way. The final value of the polarity of a homonym was computed as an average of the values of all occurrences (shapes) of the word.

During sentiment analysis, the most important words are adjectives, adverbs, nouns and verbs. These word classes usually express the polarity of an opinion in the best way. So in our improvement of DASA, the analysed texts were decomposed into sentences and sentences were decomposed into words. Consequently, word classes were assigned to each word and then only adjectives, adverbs, nouns and verbs were taken into account during sentiment analysis.

The implemented approach was tested. Results in the measures of precision, recall and F1 measure are presented in Table 3.

Table 3: Results of tests of DASA+NLP approach in Precision, Recall, F1-measure and Accuracy.

Opinion	Precision	Recall	F1	Accuracy
Positive	0,586	0,612	0,598	-
Negative	0,491	0,568	0,527	-
Average	0,539	0,590	**0,563**	**0,536**

The testing was provided on 1000 positive and 1000 negative reviews. The achieved results were only slightly better and still insufficient. But the time of processing was cut short by half.

4.3 DASA and NLP Involving New Methods of Negation and Intensification Processing

The previous testing results have confirmed that the results of sentiment analysis cannot be satisfied, when only separate words are processed. An important part of sentiment analysis is also the processing of groups of words, for example for negation ("insufficiently functional") or intensification ("very nice graphic design"). The negation and intensification represents derivatives of a language. Both negation as well as intensification can change the polarity of a connected word (consequently also polarity of the whole text) and in this way they can increase the precision of sentiment analysis.

The processing of negation and intensification is based on information from an external source – the classification lexicon. Our classification lexicon contains a special part for negation processing (with words as no, not, never, neither, nobody, none, nothing, etc.) and a special part for intensification processing (very, highly, too, most, extremely, etc.).

There are two known main methods for negation processing: switch negation and shift negation (Taboada, 2011). Within our approach four various methods for negation processing were tested:

- modification of an opinion polarity by a direct reverse turn
- shifting opinion polarity of related word
- change of opinion polarity using a constant value
- change of opinion polarity using a percentage value.

All these methods were implemented and tested. These tests have shown, that the most suitable and precise possibility is the method of change of opinion polarity using a constant value. We have also made experiments with various constant values. According to the experiments the best values are PosValue = 0,5 and NegValue = 0,65.

The second improvement of the "lexicon approach involving NLP and the selected methods for negation and intensification processing" works in the following way. It detects words in the analysed text, which are carriers of opinion (adjectives, adverbs, nouns and verbs). Consequently, it checks the existence of possibility of occurrences of negations or intensifications in the neighbourhood of the processed words. In the case of positive matching, the value of the polarity of the processed word is recomputed. The results of testing of this approach are presented in the Table 4. The achieved results in this case were significantly better and sufficient for using this approach in the dynamic analysis of a web trend.

Table 4: Results of tests of DASA+NLP involving the negation and intensification processing (DASA+NLP+NandI).

Opinion	Precision	Recall	F1	Accuracy
Positive	0,786	0,848	0,816	-
Negative	0,835	0,769	0,800	-
Average	0,811	0,809	**0,810**	**0,809**

Figure 1 illustrates the results of testing of the following implementations (from left to right): DASA, DASA+NLP and DASA+NLP+NandI. There is a gradual improvement of effectiveness. The Recall has gradually increased and the best implementation from this point of view is DASA+NLP+NandI.

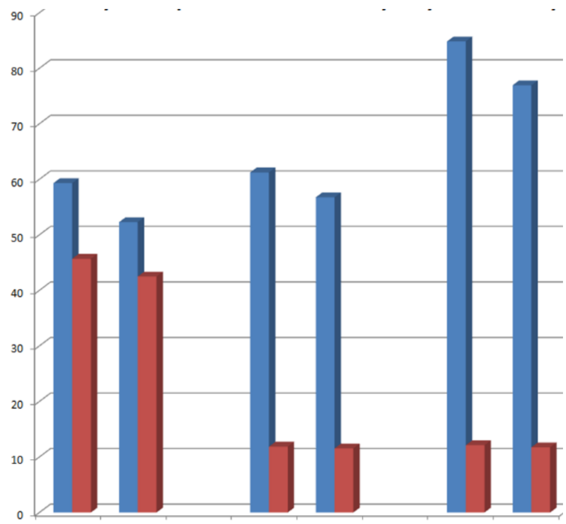

Figure 1: An overall evaluation of all tested implementations (blue for recall and red for processing time).

The processing time has rapidly dropped after involving NLP (DASA+NLP) to one quarter of the original time. But the last implementation very slightly increased the processing time by adding negation and intensification processing.

The novelty of our approach in comparison with work (Taboada, 2011) is in the DASA + NLP + NandI method, which uses original processing of negation and intensification and also the NLP technique.

There are some other approaches based on lexicons. For example, (Mohammad, 2016) presents sentiment lexicons for Arabic social media. They present several large sentiment lexicons that were automatically generated using supervision techniques on Arabic tweets, and translation English sentiment lexicons into Arabic. The approach is not comparable to our work.

Another approach in (Labille, 2017) generates a domain-specific lexicon using probabilities and information theoretic techniques. Their results are better than our results. But we used a general lexicon and usually general lexicons cannot be more precise in some given domain than the lexicon generated specifically for this domain.

In the paper (Cambria, 2016), the SenticNet4 is presented. Authors achieved better results than we, probably because they used semantically enriched approaches to sentiment analysis.

Nielsen presents a labelled word list – a new ANEW lexicon, where each word has been scored for valence, a 'sentiment lexicon' or 'affective word lists' in (Nielsen, 2011). This interesting approach cannot be compared to our approach, because they used unusual efficiency measure.

5 DYNAMIC ANALYSIS

A dynamic analysis of sentiment plays an important role in solving real problems. It is more important for companies, because of the decrease in the cost of the analysis, its higher precision and wider possibilities for utilization. So the companies can effectively obtain a feedback from users.

5.1 Web Forum Selection

As a data source, the discussion on a new mobile Apple iPhone X was selected. Reviews on this new product were extracted from the web forum Gsmarena (Gsmarena, 2018). This mobile was introduced and brought to marked recently and specialists consider it to be a revolution among modern mobiles. The discussion forum Gsmarena (see Figure 2) was selected because the owner of this forum is the society which:

- does not sell any mobiles
- does not offer any recommendations for clients which mobile to buy
- doesn't have any preferences connected with mobiles
- doesn't have any profit from mobile sales.

Thus, this web discussion about mobiles is not influenced by the owner of this discussion forum and opinions are not modified, preferred or deleted. Another reason for the selection of Gsmarena forum was the fact, that millions of unique users are active on it each day and a great majority of them express their opinions. Rules were defined for contributing to the Gsmarena web forum to guarantee valuable and real reviews.

5.2 Results of the Web Trend Analysis

The text data from Gsmarena were extracted together with information about time and date of comment posting and processed using the implementation DASA+NLP+NandI. During this processing all reviews were analysed from the point of view of positivity/negativity of texts. All polarities of all posts's texts on the given theme were summarized in the form of an unweighted normalized sum.

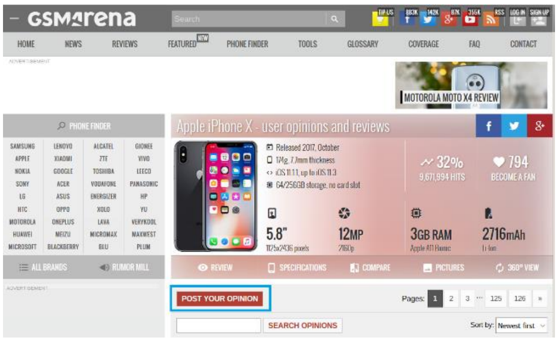

Figure 2: An illustration of Gsmarena web forum.

We have obtained information needed for future dynamic analysis of the web trend connected with reviews on mobile Apple iPhone X. There are some details on the dataset presented in the Table 5.

Table 5: The results of application DASA+NLP+NandI on data from Gsmarena on the mobile Apple iPhone X.

	N	PN	NN	PS	NS
September	833	443	390	0,5318	0,4682
October	361	182	179	0,5042	0,4958
November	543	288	255	0,5304	0,4696
December	252	154	098	0,6111	0,3889
January	157	085	072	0,5414	0,4586
February	136	065	071	0,4779	0,5221
March	025	017	008	0,6800	0,3200

Where:

N is the number of comments extracted from Gsmarena during the given period

PN is the number of comments from the given period classified as a positive opinion by application DASA+NLP+NandI

NN is the number of comments from the given period classified as a negative opinion by application DASA+NLP+NandI

PS is PN/N (ratio of the amount of positive opinions to all opinions)

NS is NN/N (ratio of the amount of negative opinions to all opinions)

The total number of posts in the dataset was 2307 (1234 positive and 1073 negative posts). The results from Table 5 were transformed into a graphical form and illustrated in Figure 3.

Figure 3 represents the dynamic analysis of the sentiment of the selected web trend. The dynamic change of the positive opinion is drawn in blue (upper curve) and the dynamics of the negative opinion in red (lower curve). Figure 4 represents the same dynamic analysis, but in this graph only the differences between positive and negative polarity values are shown. In Figure 3 and Figure 4, we can see three turning points. The beginning of the graph represents the 12 September 2017, when iPhone X was officially announced. It was the day when a web discussion about the iPhone started.

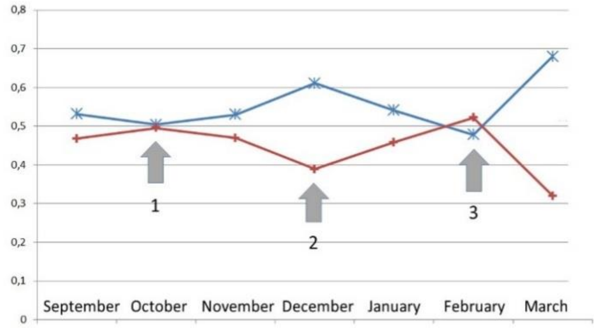

Figure 3: Dynamic analysis of a web trend - mobile Apple iPhone X with separate representation of positive (blue – upper curve) and negative (red – lower curve) opinion.

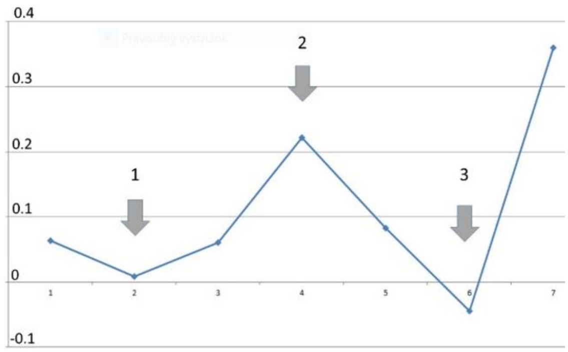

Figure 4: Dynamic analysis of a web trend (mobile Apple iPhone X) with value of resulting opinion in the form of difference between positive and negative values.

The *first turning point* is connected with the date 3 November 2017, when iPhone X was released which could explain the increase in sentiment from October to November. From then the summarized polarity started growing until the *second turning point*, when a new actualisation of iPhone X was rolled out and the Face ID function caused many errors. Consequently, the summarized polarity decreased until the *third turning point*, when all mistakes in new actualisation were corrected. After this point the summarized polarity stated to increase again.

6 CONCLUSIONS

We designed a new approach to web trend analysis. This design was based on the most important webometric – sentiment analysis. Our new approach to sentiment analysis based on a lexicon was designed and improved, to improve precision. This solution was used for the dynamic analysis of the selected web trend, which was the new mobile Apple iPhone X. The dynamic analysis showed a trend which corresponds to the real life events in the life of this mobile. We can say that our goals were fulfilled and for the future we would like to use our experience with web trends analysis in recognition of antisocial behaviour in online posting of reviews.

One limitation of this work is that only one webometric is being considered, which is sentiment analysis. For future studies other webometrics such as social networks analysis (SNA) and mention analysis (MA) might also be used. Another improvement could be to use the information about an authority or trolling of the given reviewer to increase the effectivity of the sentiment analysis (Mikula, 2018). Our approach could be a core of a future recommender system (Tarnowska, 2019).

ACKNOWLEDGEMENTS

The work was supported by the Slovak Research and Development Agency under the contract No. APVV-16-0213 and the contract No. APVV-17-0267 “Automated Recognition of Antisocial Behaviour in Online Communities”.

REFERENCES

Ahmad, B., 2010. 10 Advantages and Disadvantages of Social Media for Society. `https://www.techmaish.com/advantages-and-disadvantages-of-social-media-for-society/`, Accessed 22 July, 2019.

Baccianella, S., Esuli, A., Sebastiani F., 2010. SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *Proc. of the Conference on Language Resources and Evaluation, LREC 2010*, Valletta, Malta, 17-23.

Cambria, E., Poria, S., Bajpai, R., Schuller, B., 2016. SentiNet4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives. In *Proc. of COLING 2016, the 26th International Conference on Computational Linguistics*, Technical Papers, Osaka, Japan, December, 2016, 11-17.

Dematis, I., Karapistoli, E., Vakali, A., 2018. Fake Review Detection via Exploitation os Spam Indicators and Reviewer Behaviour Characteristics. In *Proc. of the 44th International Conference on Current Trends in Theory and Practice of Computer Science beyond Frontiers* - SOFSEM 2018, Krems an der Donau, Springer LNCS, 2018, 1-14.

Gsmarena, 2018. Gsmarena. `http://www.gsmarena.com/`, Accessed 22 July, 2019.

Jurafsky, D., Martin, J.H., 2017. *Speech and Language Processing.* An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 3rd edition, University of Colorado at Boulder, 2017, 1-499.

Kao, A., Poteet, S., 2005. Text Mining and Natural Language Processing. *ACM Exploration Newsletter*, Vol 7, No 1(2005), ACM Digital Library, 1-2.

Kumar, S., West, R., Leskovec, J., 2016. Disinformation on the Web. In *Proc. of the 25th International on Word Wide Web – WWW16*, Association for Computing Machinery (ACM), 2016, 591-602.

Labillle, K., Gauch, S., Alfarhood, S., 2017. Creating Domain-specific Sentiment Lexicons via Text Mining. In *Proc. of WISDOM'2017*, Halifax, Canada, Association for Computing Machinery, 2017, 1-8.

Lévy, P., 1997. C*ollective intelligence: Mankind's Emerging World in Cyberspace.* Perseus Books Cambridge, Massachusetts, 1997, 1-277, ISBN 0-7382-0261-4.

Machová, K., Mikula, M., Szabóová, M., Mach, M., 2018. Sentiment and Authority Analysis in Conversational Content. *Computing and Informatics*, Vol.37, No.3(2018), 737-758, ISSN 1335-9150, IF 0,410.

Malone, T.W., 2006. *What is collective intelligence and what we will do about it? Collective intelligence: creating a prosperous world at peace.* MIT centre for Collective Intelligence, 2006, 1-610.

Mikula, M., Machová, K., 2018. Combined Approach for Sentiment Analysis in Slovak Using a Dictionary Annotated by Particle Swarm Optimization. *Acta Elektrotechnica et Informatica*, Vol.18, No.2(2018), ISSN 1335-8243, 27-34.

Mikula, M., Machová, K., 2017. Annotation of the Dictionary for Sentiment Analysis Using PSO. In *Proc. of Data and Knowledge*, Plzeň, 2017, 151-156.

Mohammad, S., Salameh, M., Kiritchenko, S., 2016. Sentiment Lexicons for Arabic Social Media. In. *Proc. of the Tenth International Conference on Language Resources and Evaluation*, Portorož, Slovenia, LREC, 2016, 33-37.

Nielsen, F.A., 211. A new ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs. In *Proc. of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, 2011, 93-98.

Paralič, J., Furdík, K., Tutoky, G., Bednár, P., Sarnovský, M., Butka, P., Babič, F., 2010. *Mining of knowledge from texts*," Košice: Equilibria, 2010.

Řimnáč, M., 2018. Detection of a Disinnformation Content – Case Study Novičok in CR. In *Proc. of the Conference Data a znalosti & WIKT 2018*, 11.-12.10.2018, Brno, VUT, 2018, 65-69.

Surowiecki, K., 2004. *The Wisdom of Crowds - Why the many are smarter than the few*. Abacus, Doubleday, 2004, 1-6.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M, 2011. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, Vol.37, No.2(2011), 267-307.

Tarnowska, K., Ras, Z.W., Daniel, L., 2019. Recommender System for Improving Customer Loyalty. *Studies in Big Data*, Vol. 55, Springer, 2019.

Thelwall, M., 2005. Webometrics. *Encyclopaedia of Library and Information Science*, Vol.1, No.1(2005), Taylor and Francis Group, 1-8.

Wang, S.E., Garcia-Molina, H., 2013. Disinformation Techniques for Entity resolution. In *Proc. of the 22nd ACM International Comnference on Information and Knowledge Management, New York, USA,* 2013, 715-720.