The Model "Information Gatekeepers" for Sentiment Analysis of Text Data

Nataliia Kunanets[0000-0003-3007-2462][1], Yurii Oliinyk[2], Dmytro Kobylynskyi[2], Antonii Rzheuskyi[0000-0001-8711-4163][1], Khristina Shunevich[2], Valentyn Tomashevskyi[2]

[1]Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine
[2]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
nek.lviv@gmail.com, oliyura@gmail.com, kobylynskiy.d@outlook.com, antonii.v.rzheuskyi@lpnu.ua, krishirak@gmail.com, simtom@i.ua

Abstract. The approach for application of the model "information gatekeepers" for filtering messages, taking into account the tonality of text data and making a decision on further dissemination of information for the implementation of socially necessary restrictions on the processes of dissemination and obtaining information is considered in the paper. The peculiarities of the classical models of "information gatekeepers" are analyzed and the approach to clustering of messages in social networks is proposed based on the function of evaluation of the message using sentiment data analysis for the implementation of the procedures of "semantic content filtering". The specifics of the sources of information, the ways of its presentation, the peculiarities of the formation of target communities, which are information oriented, socio-political, religious, ethnic, cultural, legal, age, and other aspects of social life have been taken into account. The main actor of the system based on the use of this model is the "information gatekeeper", which, before the broadcast of the message via specific communication channel, implements an evaluation function for the message. The problem of sentiment analysis of custom text data is considered. An algorithm for classifying custom texts is introduced. The rules for creating a Thesaurus of basic words reflecting the evaluation of a particular object and methods for classifying texts are described. The algorithm of allocation of problem statements is proposed. To solve the problem of sentiment text analysis, an approach based on knowledge is used that involves the use of additional expert resources in the form of thesaurus of indicative words and expressions, composed manually or automatically, and writing rules that reflect the structure of fragments of text data. The advantage of this approach is the ability to ensure the effectiveness of the classification of texts without loss of quality of work for various subject areas. Two methods of classification of problem statements are proposed.

Keywords: Information Gatekeeper, Semantic Analysis, Sentiment Analysis, Tonality Analysis, Thesaurus of Keywords, Clustering.

1 Introduction

The modern information society is rapidly transforming and adapting to new technological challenges and quite often gets into a situation where it is necessary to adopt system decisions and implement them in practice in order to preserve and protect the civilizational, moral and ethical values of the established norms, rules and laws that have been generated during the previous long period of development.

One of these challenges, which is generated by the current revolutionary information and technological impetus, is the problem of implementation of socially necessary restrictions on the processes of dissemination and obtaining of information. In general, by maintaining the principles of free and democratic creation and dissemination of information, enshrined in international treaties and agreements, it is necessary to systematically apply procedures and rules that do not contradict basic established social norms.

In the context of this, the principle of necessity of implementing the procedures of "semantic content filtering" is generated, taking into account the specifics of the sources of information, the ways of its presentation, the features of the target communities to which information is focused, socio-political, religious, ethnic, cultural, legal, age, and other aspects of social life. "Unfiltered" information flows in some cases can play not only a destructive, but also a generally powerful devastating role in relation to social values.

The aim of the paper is an analysis of developed methods and software for evaluating expressions related to problem situations, taking into account the features of unstructured texts of users of social networks. The development of new original methods for filtering messages in information flows is one of the most relevant areas of research in the context of providing information security both to individual actors and to the country as a whole. This will allow to deal with real and potential threats, in particular, the dissemination of harmful, in some cases dangerous, unreliable and biased information by individual entities. Practice shows that information can be successfully processed regardless of its content [1]. Therefore, for modeling the information flows it can be sufficiently adequately described by the classical theory of information, which is interpreted as a mathematical theory of message transmission, developed by C. Shannon [2], and is substantially supplemented and advanced by the works of N. Wiener, V. Kotelnikov, A. Kolmogorov. At the same time, in many cases, for problems arising in the analysis of information flows, there is much in common with the tasks of statistical physics and hydrodynamics, which in our opinion can be intuitively solved by similar methods [3, 4]. Network technologies change the traditional linear flow of information in society and connect communicators at different levels of presentation and use of information. American researchers in the area of communication sciences, J. Bryant and S. Thompson note: "features of new technologies are forced to go beyond traditional communication. This new area can be called transactional media communication. It means the change of roles – the transition to interpersonal communication relations, in which each side can in turn act as a sender, receiver or transmitter of information. Media communication means that these technologies

include media". In this context, we should turn to the theory of "information gatekeepers".

2 The Survey of "Information gatekeepers" Models

The theory of "information gatekeepers" is used in the following areas: information science, information systems management, management, political science, communication science, jurisprudence, public relations, sociology.

Information gatekeeper is a person who controls access and decides whether the message will be disseminated by mass media. The mass media means broadcasting information to a wide audience via traditional channels, as well as websites, news portals, blogs.

The "information goalkeepers" model can also have an individual form of implementation, deciding what message will be posted on the website or broadcasted via e-mail. The term "gatekeeper" was introduced by Kurt Lewin in 1947.

Kurt Lewin noted that the movement of news by certain communication channels depends on the fact that certain areas in the middle of the channels function as "gates". Continuing the opinion, Levin pointed out that the sections of the gates are managed by mutual rules or "information gatekeepers" who are authorized to make a decision on **acceptance** or rejection of the message.

Kurt Levin identified key parts of the "information gatekeeper" model:

- Information moves step by step through channels. The number of channels varies, and the amount of time a message stays in each channel can be different.
- Information must pass the "gates" to go from one channel to another.
- There may be a number of psychological barriers that turn into a conflict, which creates resistance to the movement of the message through the channel.
- The number of channels that lead to the end result may be unlimited.
- Different actors can control channels and act as information gatekeepers at different times.

Wilbur Schramm expressed his own observation on the theory of "information gatekeepers." He noted: "there is no aspect of communication that is so impressive as the large number of choices and deviations that have been made between the formation of meaning in the imagination of the communicator and the probability of the appropriate meaning in the recipient's imagination."

There is an opinion that in the narrowest sense the theory of control at entrance is based on the selection mechanism. Different approaches to application of the theory of "information gatekeepers" were distinguished:

- studying decision-making processes by journalists and editors regarding the broadcast or rejection of news;
- studying media content intended for broadcasting by channels of information dissemination;

- use of the control mechanism at entrance and information gatekeepers in the communication system.

David White proposed a successful combination of Schramm's "source–message–receiver" approach in mass communication to the theory of control at the entrance.

Feedback among the participants in the communication process is a key component, missing in previous models of mass communication, formed under the influence of the H. Lasswell's model. Lasswell's model also known as Lasswell's communication model (1948) describes the act of communication, identifying the following questions: who said, what was said, what channel it was transmitted, to whom it was said, and what effect it would have.

The existing control theories at entrance distorted the concept of "information gatekeeper" in the context of information networks, where the proposed theory is fully applied. A new theory is necessary because hybrid interpretations of the concepts of control at entrance and gatekeeper are not sufficiently used and investigated in the context of information society and the Internet.

This makes it necessary to revise the means of terminology of "information gatekeepers" theory, moving from selection process (source, channel), dissemination and protection of information, choice of information mediator (science of management) to a more flexible design of management and processing of information through networks.

According to the traditional concept, control at entrance is carried out by the sender-receiver model. Mediators (editors, collectors) are considered senders, and "locking" (newspaper readers, community members) played the role of receivers. The sending and receiving of information can change depending on the context: news, technological development, etc. Traditionally, information gatekeeper is responsible for editing, translation, creation and dissemination of information messages.

Thanks to Web 2.0 capabilities, users started to play an important role in creation and dissemination of online news through social networks such as Twitter and Facebook. P. Shoemaker and T. Vos suggested such practice as a control at entrance of audience. According to them, control at entrance of audience is a process in which users appeal to existing news and comment them based on their own set of criteria about value of news. The functions of the information gatekeeper are to use the technology of personal mediator, mediator between groups and communities, and as well as controller for access to information.

The information gatekeeper before broadcasting a message through defined communication channel carries out an evaluation function of the message.

Notification about recommended resources (group information) does not allow the exclusion of information gatekeepers from communication process. This branch of verbal communication model illustrates location of information gatekeepers who control the channels of information transmission: social networks (area of activities of curators of content), video hostings. As soon as the generated message falls into the social network, the users of the information gateway acquire the responsibilities of retransmitting the message, while the number of recipients increases with geometric progression. Feedback is carried out on two levels: a) the user - the group administrator;

b) the user is a user. Another vector of a branch of a model is the opening by the information gatekeepers of the channel cloud storage - social networks, since these platforms are closely correlated with each other.

At the same time, it should be noted that the procedures for semantic content filtering of information flows can be realized both for the purpose of selecting valuable, useful for certain semantic features of messages, as well as for the removal of semantically useless and harmful messages from information flows by certain communities, social groups and social categories from the information flows. Such a binary nature of semantic filtering of messages in information flows makes it possible to formulate complexes of direct and inverse problems, the basic tool of solution of which, according to the authors of work, is a model approach with the use of mathematical models based on physical analogies and similarities that are inherent processes of filtration in material flows.

The implementation of semantic content filtering procedures for the purpose of removing semantically useless and harmful messages from information flows becomes particularly relevant in the context of providing information security to a country in a state of hybrid warfare, one of the areas of which is information confrontation.

Therefore, it is important to implement highly productive means of evaluating user communications before being distributed to social networks, video hosting and media.

It should take into account the need to address the following tasks:

- conduct a classification of user messages for different emotional load;
- create thesaurus for key indicators and evaluative words;
- develop classification methods based on rules and thesaurus, as well as on the grammatical structure of complex sentences;
- develop a method for identifying problem phrases in relation to objects for which the problem phrase is expressed and related to the subject domain, based on the public thesaurus.

3 An Approach Using Sentiment Analysis of Text Information and Analysis of Anomalies in Data Flows

In order to take into account the sentiment loading of messages in social networks that form huge data streams, it is proposed to add to the model "information gatekeeper" the tasks of preliminary processing of data:

- determination of emotional colour of messages;
- definition of anomalies in data flows.

In the proposed approach, the emotional colour is taken into account as follows:

- the message with a neutral colour is presented for the next step.
- messages with a pronounced negative or positive colour require further analysis, as they are the element of dissemination of manipulation information, such as propaganda or misinformation.

Defining abnormalities in text data streams will allow you to define non-standard flow characteristics, such as increasing the concentration of message specific topics, concentration of messages by geographic affiliation, and regulating the further dissemination of such information.

The quality of the proposed methods is higher in comparison with existing models, it is adaptation for working with texts of different lengths and in different languages.

4 The Task of Sentiment Analysis of Texts

Let’s consider sentiment analysis on an example analysis of user comments on certain events [15]. An analysis of user expressions can serve as an effective means of monitoring and evaluating user opinions expressed on social networks to evaluate the absence of provocative and harmful judgments. Sentiment analysis will contribute to the formation of ratings in public surveys, analysis of past and future events, in promotional tools (targeting, services that recommend products), customer consultation and technical support, etc. In article [16] sentiment analysis is used for filtering "positive" or "negative" text with combination with convolutional neural network. In scientific research, the following tasks of sentiment analysis of user feedback text are described:

- sentiment analysis of texts in user's statements regarding aspects;
- the allocation of evaluative phrases and words;
- classification of texts at the level of documents and sentences.

To solve the problem the prevailing comments are highlighted, which have a problematic vocabulary or incorrect submission of information about the object. A plurality of user content in social networks is used and a body of texts for analysis is created. The task is divided into the following subtasks, which correspond to the peculiarities of the tasks of sentiment analysis of thoughts [17]:

- Allocation of reviews containing problematic statements;
- To distinguish users from problems related to the subject area;
- Identification of target objects of a certain topic to determine the problems described in the set of reviews of the relevant subject area[18,19].

After analyzing sentences from the text of comments or messages in Ukrainian and English four phrases are highlighted:

- Explicit mention of the problem. The type contains a direct indication of dissatisfaction with the object, such as: "constant problems with ...", "not working properly", etc.
- Implicit mention of the problem. The type of these phrases does not mention the problem, but contains auxiliary words and implies a problem that results in dissatisfaction or hostility to the user. Examples: "It's all an imitation of work", "absurd", "confusing situation."

- Denial of the existence of the problem. When using this type of phrase, the user denies the previously mentioned or expected problems. For example: "we managed to reach ...", "without complaints".
- Lack of problems. The opinion of the user does not contain references to the expected or actual failure, dissatisfaction. Examples: "excellent result", "good thing happened".

5 Thesaurus of Appraisal Vocabulary in Ukrainian and English

The described types of statements contain information on the existence of problems on the basis of structures that clearly identify them. One of the key tasks that underlies the development of methods for sentiment analysis of the views of users of text data is the creation of vocabulary indicative words. The Thesaurus highlights the main array of the most characteristic expressions. Statistics for thesaurus generated manually.

Table 1 Statistics of the size of thesaurus generated manually

Thesaurus	Thesaurus size	
	For Ukrainian language	For English language
Action	7863	7886
ProblemWord	942	190
NotProblemWord	69	42
NegativeWord	1476	4169
PositiveWord	1078	2323
AddWord	30	15
ImperativePhrases	26	6
Words-Denial	14	22

6 Approach and Classification Methods

To determine the sentiment load of texts using a wide range of methods, the leading place in which is allocated to those that are intended for automated detection of "subjective" information (thoughts, judgmental judgments, emotions, feelings, etc.). The analysis of the sentiment load is to find the thoughts in the text and determine their properties. Their choice depends on the tasks to be solved, as well as the context that needs to be shaped: personalization of the content; b) the subject of messages; c) evaluation of the object ("positive" or "negative") [1].

The most commonly used methods are the classification of sentiment texts. However, it does not automatically determine the emotional colour (positive, negative, neutral) of text data. This is due to ambiguous statements, in particular the style of texts in social networks can vary from slang to literary or scientific. At the same time, the sentiment load will be different in terms of ambiguity, uncertainty, sarcasm, which does not contribute to a clear assignment of evaluation.

Frequency methods, which involve the establishment of weight coefficients, can be used to evaluate the importance of words. The weight of a single word is defined as the product of the frequency of its use in a specific document (TF) and the degree of importance of the word in the context of the collection (IDF – inverse document frequency):

$$TF = \frac{n_t}{\sum_k n_k}$$

where n_t is quantitative indicator of the use of the word t in the document, $\sum_k n_k$ is the total number of words in this document [6].

IDF needs to reduce the weight of widely used words. For a unique word within a specific collection of documents, only one IDF value is formed.

$$\mathrm{DF} = \log \frac{|D|}{|\{d_i \in D | t \in d_i\}|},$$

where $|D|$ is a number of documents in the collection; $|\{d_i \in D | t \in d_i\}|$ is a number of documents from the collection D, in which it meets t (when $n_t \neq 0$) [4].

The frequency of word use becomes the basis for determining its importance for this document. Mostly it is defined as the ratio of the quantitative index of the use of the given word to the total number of words of the document.

In this way, the linking of the evaluation criterion to a particular document occurs because the weight of words with a high frequency of use in a particular document increases, however, the low frequency of their use in other documents may increase.

V. Purto proposed a hypothesis to evaluate semantic significance of sentences. The author tested at automatic reflexion of texts and used the frequency analysis of the text regarding the presentation of important terms in it. The researcher noted the regularity as the importance of a certain period of the text affects the frequency of its use in it. Therefore, for quasi-abstract V. Purto considered it necessary to select such sentences containing the largest number of terms, which is often repeated in this document. However, the methods discussed do not make it possible to clearly determine the sentiment load of short post texts that are predominantly inherent in the message in social networks.

To achieve the objectives of the sentiment analysis, an approach based on knowledge has been applied. This approach involves the use of additional expert resources in the form of thesaurus of indicative words and expressions, composed manually or automatically, and writing rules that reflect the structure of fragments of text data. The advantage of this approach is the ability to ensure the effectiveness of the classification of texts without loss of quality of work for various subject areas. Two methods of classifying problematic propositions are proposed:

- a method that takes into account the conditions for entering words or phrases from thesaurus;
- a method that performs the analysis of the grammatical structure of complex sentences on conjunctions.

Thus, a class of problem statements (problem class) and a class of statements without problems (no-problem class) are distinguished.

7 The Algorithm of Allocation of Subject-Oriented Problem Statements and Target Objects

When applying the method based on the rules and the grammatical structure of the sentences, the first grammatical part of the sentence (to the conjugate) has a positive tone, while the second part (after the conjugate) differs by the tonal estimation. The first grammatical part of the sentence (to the connecting connector) confirms the existence of a problem or difficulty in use, but the second part of the sentence (after the conjugate) denies the problem or the negative situation.

All grammatical parts of the sentence contain similar information about the existence of certain problems. The first grammatical part of the sentence contains the condition of the problem, while the second part does not indicate a difficult situation. Examples of rules for a method based on the grammatical structure of sentences

$\triangleright clause1 \rightarrow P - IP - DP,\ conj \rightarrow$ but;
$clause2 \rightarrow A - DP;\ S \rightarrow PS;$

$\triangleright clause1 \rightarrow AW - IP,\ conj \rightarrow$ but;
$clause2 \rightarrow \neg\ DP;\ S \rightarrow \neg\ ;\ \triangleright$
$clause1 \rightarrow DP - IP,\ conj \rightarrow$ but;
$clause2 \rightarrow \neg\ DP \mid \neg\ IP;\ S \rightarrow \neg\ PS;$
$clause1 \rightarrow IP \mid DP,\ conj \rightarrow$ though;
$clause2 \rightarrow \neg\ DP;\ S \rightarrow PS;$

The algorithm uses the results of the analysis of the text statement by the methods previously proposed: a method based on a number of conditions, and a method based on the analysis of complex sentences. The general description of the algorithm consists of several steps:

Step 1. Extract from the statement s_{ij} indicators entry $\{pw_{i1}, pw_{i2}, \ldots, pw_{in}\}$, $n \leq |s_{ij}|$ depending on the related objections from the thesaurus of verbs, problem words, words with negative tonality, additional words, command phrases using the method analyzing on a number of conditions;

Step 2. For each pw_{ij} determine the set of possible target objects $\{t_1, t_2, \ldots, t_k\}$, if the target object t_k syntactically related to w_{ij}, that is, there is a direct or indirect relationship between t_k and pw_{ij} in statements s_{ij}; if the set of objects is empty, w_{ij} excluded from the set of indicators;

Step 3. For each t_k to determine whether the object is subject-oriented on the basis of measures of connectedness of terms t_k and the terms of the subject area in the linguistic resource;

Step 4. Classify the statement s_{ij} as a statement that points to a problem situation about a subject-oriented target object, if there is at least one combination (pw_{ij}, t_k) and

$r(s_{ij}) \neq 0$ according to the results of the analysis by the method based on the analysis of complex sentences; otherwise, to classify the statement sij as having no problem.

Morphological processing of the text was carried out using the Misto7 library for the Ukrainian language: at the stage of preprocessing the texts were made the lepretization of all words in the Ukrainian language.

The following algorithm of the method of the Expression of words, indicating problem situations with products, is formed on the basis of user feedback:

Algorithm 1: Algorithm for obtaining subject-oriented problem statements and target objects

```
Function lookupForRelatedTargets (pw, DRs)
Input: pw is found problem indicator, DRs is dependent
set between words
Output: Ts is set of target objects
Ts ← ∅
foreach d in DRs do
if dr.contains (pw) then
/ * Search for target target objects directly dependent
on the indikator * /
if dr.matches (direct_type_of_relations) then
target = getTargetFromDep (dr)
target = getAddWordsForTarget (target, DRs)
Ts = Ts∪ {target}
else
/ * Search for target target objects directly dependent
on the word-intermediary * /
successor = theOtherWordFromDep (dr, pw)
Ts = Ts ∪ lookupForRelatedTargets (successor, DRs \ {dr})
return Ts
Function lookupForProblemsWithTargets (s, domain_terms,
common_terms)
Input: s is - Original sentence, domain_terms -
Subject-oriented Terms, common_terms - Background Terms
that define a large group of goods
Output: PWTs is set of pairs (problem indicator, object)
PWTs← ∅
    / * Search annotations from the Thesaurus in the
sentence * /
PWs = lookupForPW (s);
/ * Sentence analysis using grammatical parsing * /
DRs = (getGrammStructure (s)). TypedDependenciesCollapsed
(true)
foreach in PWs do
targets = lookupForRelatedTargets (pw, DRs)
foreach in targets do
```

```
/ * Calculation of semantic connectivity between the
target object and the terms of the region domain_terms
and a wide range of products common_terms * /
if relScore (domain_terms, ti) relScore (common_terms,
ti)
then
PWTs = PWTs ∪ {pair (pw, ti)}
return PWTs
```

The classification results for the class of statements about problem situations and the classification results obtained by macroaveraging are presented in Table 2.

Table 2. Results of classification.

Method	Machines (ukr.)							Applications (ukr.)						
	Acc.	P	R	F	macroaver			Acc.	P	R	F	macroaver		
						R	F					P	R	F
NaiveBayes	.754	.380	.470	.420	.624	.645	.634	.791	.809	.834	.821	.786	.783	.784
NRC+Dicts	.847	.621	.496	.552	.754	.712	.732	.831	.841	.874	.857	.829	.824	.826
GU+Dicts	.852	.694	.391	.501	.782	.675	.725	.833	.843	.874	.858	.831	.826	.829
KLUE+Dicts	.853	.715	.380	.496	.792	.672	.727	.832	.843	.870	.856	.829	.825	.827
DbA	.814	.507	.636	.564	.708	.746	.726	.806	.829	.837	.833	.802	.803	.802
CbA	.814	.508	.649	.571	.709	.751	.730	.820	.842	.846	.845	.816	.815	.816

A method based on a number of conditions, denoted as DbA; A method based on the analysis of complex sentences, designated as CbA. Classics based on "bag of words" are marked with c "1gr."; classifiers, taught in words and phrases, are marked with "2gr." The classification results allow a number of subsequent observations. First, depending on the subject area, among the base models (DecisionTrees, MaxEnt, SVM), the best results for the macro-F scale are shown by different classifiers: SVM for texts on electronics, children's products, applications, tools and machines (ukr.); MaxEnt for machine texts (English). Within the framework of the models that show the best results in the tonal analysis (NRC, GU, KLUE), the best results show: NRC for machine texts (ukr. and eng.), electronics, tools; GU and KLUE for text about attachments (ukr.); KLUE for children's goods texts. Qualitative analysis of classification results. In Fig. 1 the results of the classification error analysis are presents, which identify the following types of most common Errors when analyzing vehicle reviews:

- the error associated with the definition of related objections, conditions and rules;
- insufficient completeness of coverage of texts with the help of created thesaurus and rules;
- superfluous thesaurus;
- the problem situation arose in the specific (certain) conditions;

- request / requirement of the functional or recommendation to change;
- questions to developers on use;
- spelling Errors;
- meaningless for the developers of a statement or statement about another product;
- errors related to the individual benefits of the user.

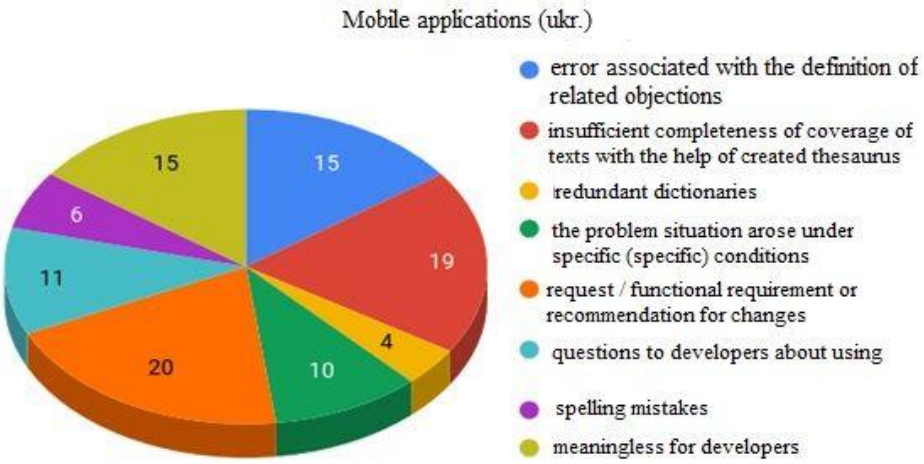

Fig.1. Errors discovered when analyzing machine reviews (ua.)

8 Define Anomalies in Text Data Streams

Isolation Forest algorithm was used to determine the anomalies [14].The method by which the algorithm constructs a partition initially creates an isolation tree of random decision tree. Tree is builded based on a extracted keywords from tweets using RAKE algorithm. Those keywords are translated to feature vectors with a help of word2vec predefined models of Matlib before they can be used in DecisionTrees.

For realtime data stream analysis used MLlib library, which is the part of the Apache Spark Server. And for data processing used Apache Kafka, RDD, MongoDB technologies. There, using the embedded Spark Streaming module, data analysis using the MLlib module is performed. A model is being built and taught of machine learning. Output of posts and notifications about anomalies detection is implemented in the interface. Spark is very well integrated with with the HBase database, where all the publications and instances attributes are stored. This software architecture has several advantages – realtime mode, scalability, high performance and big support AI modules and library.

9 Conclusions

The approach of information gatekeepers, which takes into account the tone of text data for making a decision on further dissemination of information is considered in the article. The model of "information gatekeeper" proposes inclusion of the task of preliminary data processing to determine the emotional colour of messages and identify anomalies in data flows.

The emotional colouring takes into account: the message with neutral colour, which is submitted for the next step, and the messages with a pronounced negative or positive colour will be further analyzed as they are the element of the distribution of not desirable information (propaganda or disinformation). The method of sentiment analysis of texts is proposed on the basis of analysis of user comments about certain events. Analysis of the results of sentiment analysis confirmed that further improvement of the results is possible due to the creation of highly specialized dictionaries and the development of conditions for the entry of lexical units, depending on the thematic category of selected text fragment.

Defining anomalies in text data streams will allow us to define non-standard flow characteristics and regulate the further distribution of unwanted information.

References

1. Haken, G.: Information and self-organization. Macroscopic approach to complex systems. 2nd edn. Librokom (2005).
2. Shannon, K.: Works on the theory of information and cybernetics. Izd. foreign lit. (1963).
3. Lande, D. V. :Fundamentals of information flow integration. Engineering, Kyiv (2006).
4. Lande, D.V.: Modeling the Dynamics of Information Flows. Fundamental Research 6-3, 652-654 (2012).
5. Krasnoyarova, O.: Modern transformation and traditional modeling mass communication, http://www.slideshare.net/tuesdaytalks/media-gatekeeping-theory?related=1.
6. Lewin, K.: Forces behind food habits and methods of change. Bulletin of the National Research Council 108, 35–65 (1943).
7. Schramm, W.: MassCommunications. Urbana, IL: University of Illinois Press (1949).
8. Roberts, C.: Gatekeeping theory: an evolution, http://www.reelaccurate.com/about/gatekeeping.pdf.
9. Shoemaker P. J. GatekeepingTheory / Pamela J. Shoemaker, Tim P. Vos. – London : Taylor&FrancisLtd, 2008. – P. 113.
10. Shoemaker, P. J.: How to build social science theories. Thousand Oaks: Sage Publications (2004).
11. Rzheuskyi, A., Kunanets, N.: The model "information gatekeepers" in the system of social communication. In: Information, communication, society 2014: materials of the 3rd Mizhnar. sciences conf. ICS-2014, May 21-24. 2014, Ukraine, Lviv, Slavske. Lviv. Polytechnic, Lviv, pp. 304-305 (2014).
12. Chalaya, L. E., Shevyakova Yu., Shafronenko A.: Measures of the importance of concepts in the semantic network of ontological knowledge base. In: materials of the second intern. Sci.-Tech. conf. "Modern trends in the development of information and communication technologies and management tools".KDAVT, Kyiv, p. 51 (2011).

13. Medicovsky, M., Shunevich, O.: Investigation of the effectiveness of determining weighting factors of importance. Bulletin of the Khmelnitsky National University 5, 176-182 (2011).
14. Tomashevskii, V., Oliynik Y., Yaskov V., Romanchuk V.: Realtime text stream anomalies analysis system. Visnyk of Kherson National Technical University 66 (3), 361-366 (2018)
15. Huffman, E., Prentice, S.: Social media's new role in emergency management (No. INL/CON-07-13552). Idaho National Laboratory (INL) (2008).
16. Gavrilenko O., Oliinyk Y., Khanko H. Analysis of propaganda elements detecting algorithms in text data. In: Hu Z., Petoukhov S., Dychka I., He M. (eds) Advances in Computer Science for Engineering and Education II. ICCSEEA 2019. Advances in Intelligent Systems and Computing, 938 (2019).
17. Pang, B., Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1–2), 1-135 (2008).
18. Rzheuskyi, A., Matsuik, H., Veretennikova, N., Vaskiv, R.: Selective dissemination of information – technology of information support of scientific research. Advances in Intelligent Systems and Computing III, 871, 235-245 (2019).
19. Rzheuskyi, A., Kunanets, N., Stakhiv, M.: Recommendation System "Virtual Reference". In 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), vol. 1, pp. 203-206 (2018).