A Decision Support Approach for Quality Management Based on Artificial Intelligence Applications

Nafissa Yussupova, Maxim Boyko, Diana Bogdanova
Faculty of informatics and robotics
Ufa State Aviation Technical University
Ufa, Russian Federation
Emails: {yussupova@ugatu.ac.ru, maxim.boyko87@gmail.com, dianochka7bog@mail.ru}

Andreas Hilbert
Faculty of Economics
Dresden University of Technology
Dresden, Germany
andreas.hilbert@tu-dresden.de

***Abstract*—This paper describes the application of a novel domain-independent decision support approach for product quality management. It is based on customer satisfaction research through deep analysis of consumer reviews posted on the Internet in natural language. Artificial Intelligence (AI) techniques, such as Text Mining and Data Mining, are used for realization of consumer reviews analysis. In paper, specific Internet resources (such as yelp.com, tripadvisor.com, tophotels.ru) are used for accumulating customer reviews as a data source is considered. This is performed in accordance with the quality standard ISO 10004 proposed decision support approach, which allows for both qualitative and quantitative customer satisfaction surveys to be carried out. The output of the quantitative survey are values of customer satisfaction by product and each product's aspect. The output of the qualitative survey are significance values of products aspect for customers and identified latent relations between satisfaction by product and satisfaction by products' aspects. The proposed approach is performed as a prototype of a Decision Support System. To evaluate the efficacy of the proposed approach, an experiment on hotel's customer satisfaction has been carried out. The obtained results prove the efficacy of the proposed decision support approach for quality management and the concept of using it instead of classical methods of qualitative and quantitative research of customer satisfaction.**

Keywords-quality management; decision support system; sentiment analysis.

I. INTRODUCTION

In order to provide product quality, a company should make effective managerial decisions. In the modern world, the efficacy of managerial decision-making process depends on the information available to the person that makes decisions and the depth of information analysis. Therefore, a company should develop processes of automated collection of information and its further processing and analysis. Decision-making should be based on the knowledge and principles obtained during the analysis of the collected data.

Quality assurance is currently attained through a process approach based on the model of a quality management system [1]. It describes the interaction of the company and the customer during the process of product production and consumption. To correct the parameters of a product's quality in order to improve it for the customer, the model has feedback. For companies, feedback during the process of quality management is the information about the level of customer satisfaction, which is expressed in the form of customer reviews about a product's quality. That is why customer satisfaction is key information for quality management that influences decision-making.

To collect data and evaluate customer satisfaction, International Quality Standards ISO 10004 (International Organization for Standardization) recommends using the following classical methods: face to face interviews, telephone interviews, discussion groups, mail surveys (postal questionnaires), on-line research and surveys (questionnaire surveys) [2]. However, these methods of collection and analysis of customer opinions have a number of significant drawbacks.

A general drawback of these methods is a large amount of manual work: preparing questions, creating a respondent database, mailing questionnaires and collecting results, conducting a personal interview, and preparing a report. All of these procedures make a research expensive. These methods cannot monitor customer satisfaction continuously. For this reason, monitoring is limited by a one time period. There is no possibility for monitoring trends of customer satisfaction. It also has a negative influence on lengthiness of managerial decision making.

Another problem regards various scales for measuring customer satisfaction and their subjectivity perception. Value of customer satisfaction is estimated by abstract satisfaction indices that are difficult to understand, hard to compare and interpret. Furthermore, methods for data analysis recommended by ISO 10004 [2] allow detection of only linear dependencies and relations in data.

 ISBN: 978-1-61208-358-2

The aim of this paper is the development of a decision support approach for quality management based on the research of customer satisfaction with use of AI technologies.

The remainder of this paper is organized as follows: in Section 2, we focused on overview of recent solutions and frameworks for analysis of user generated content and their drawbacks. In Section 3, we described architecture and workflow of proposed decision support system. In Section 4, we described using AI techniques for qualitative and quantitative customer satisfaction surveys. In Section 5, we provide experiment with researching customer satisfaction of two hotels and whole resort. The obtained results could be used for decision making.

II. RELATED WORK

Applying Text Mining tools for analyzing customers' reviews posted on the Internet is not novel. There are many studies concerning models and methods for data collection, sentiment analysis and information extraction. Recent studies show acceptable accuracy of methods for sentiment classification. Gräbner et al. [3] proposed a system that performs the sentiment classification of customer reviews on hotels. Lexicon-based method [27] allowed the correct classification of reviews with a probability of about 90%. These achievements make sentiment analysis applicable for an application on quality management.

Jo and Oh [4] and Lu et al. [5] considered the problems of automatically discovering products' aspects and sentiments estimation for these aspects which are evaluated in reviews. For solving these problems, they suggested methods based on Latent Dirichlet Allocation [28] and its modifications.

A lot of social monitoring systems and frameworks have been developed for automatic analysis of reviews and topics. Liu et al. [6] presented framework called Opinion Observer for analyzing and comparing consumer opinions of competing products. This prototype system is able to visualize the strengths and weaknesses of each product in terms of various product features. Kasper and Vela [7] presented a web based opinion mining system for hotel reviews and user comments that supports the hotel management called BESAHOT. The system is capable of detecting and retrieving reviews on the web, classifying and analyzing them, as well as generating comprehensive overviews of these comments. Blair-Goldensohn et al. [8] proposed a system that summarizes the sentiment of reviews for a local service, such as a restaurant or hotel. In particular, they focus on aspect-based summarization models. Ajmera et al. [9] developed a Social Customer Relationship Management (SCRM) system that mines conversations on social platforms to identify and prioritize those posts and messages that are relevant to enterprises. Bank [13] proposed interactive Social Media monitoring system to extract related information from user generated content. One of the important contribution of this work was the proposition of new quality indeces.

In some related work, authors pay attention to relations between overall ratings of products, and ratings of products' aspects evaluated in the review. Wang et al. [10] formulated a novel text mining problem called Latent Rating Analysis (LARA). LARA aims at analyzing opinions expressed in each review at the level of topical aspects to discover each individual reviewer's latent rating on each aspect as well as the relative importance weighted on different aspects when forming the overall judgment. De Albornoz et al. [11] aimed to predict the overall rating of a product review based on the user opinion about the different product features that are evaluated in the review. For experiments, authors used reviews on hotels.

Wachsmuth et al. [12] formulated and validated an important hypothesis that the global sentiment score of a hotel review correlates with the ratio of positive and negative opinions in the review's text and that the global sentiment score of a hotel review correlates with the polarity of opinions on certain product features in the review's text.

The main drawback of these considered systems is that they can provide entirely only a quantitative survey of customer reviews, i.e., provide measurement of the degree of customer satisfaction by a product and its aspects. Qualitative survey were usually only conducting the extraction of a product's aspects. Estimation of the significance of each product's aspects for the customer is missed. The information about products' aspects that influence satisfaction and their relative importance for the customer is missing, as well as an insight into customer expectations and perceptions.

The most related work to this problem is [24]. It is dedicated to the topic of aspect ranking, which aims to automatically identify important aspects of product from online consumer reviews. Most proposals used a probabilistic model with a large number of parameters that lead to low robustness of the model. Total weighting values of aspects are calculated as the average of the weighting values by each review. Finally, significance values of aspects are estimated independently of an opinion's sentiment, e.g., in real life, we can discuss in review about bad "signal connection", but we usually omit comments in case of good "signal connection", because it must be in phone. In this manner, it is possible to use the Kano's model of customer satisfaction [25], which classifies customer preferences into four categories.

In this paper, for qualitative survey is used a novel approach based on transformation results of Sentiment Analysis into binary data. After that, binary data is processed with a Data Mining tool – Decision Tree (see Section IV). Qualitative survey aims to identify how the sentiment of reviews depends on the sentiment of different products' aspects. In other words, how overall customer satisfaction by product depends on the customer satisfaction by a product's aspects. Decision Tree performs this aim and identifies latent relations between the sentiment of reviews and sentiment of a product's aspects. We estimate the significance of aspects from the constructed Decision Tree. Output of qualitative survey are significance values of positive and negative mentions about a product's aspects for customers and identified latent relations extracted by Decision Tree. The availability of both quantitative and qualitative surveys allows realizing Decision Support System for Quality Management in accordance with quality standard ISO 10004.

III. THE PROPOSED DECISION SUPPORT APPROACH

The suggested approach to decision making in product quality management accomplished through unification of methods for collecting and processing text data into Intelligent Decision Support System (IDSS). The architecture (subsystems and contained modules) of the obtained IDSS is presented in Figure 1. The subsystem of monitoring and data collection fills the warehouse with customer reviews and other relevant information. It also supports the actuality of data via automated monitoring of Internet resources and carries out data cleansing. The data storage subsystem provides safe-keeping and integrity of collected reviews and results of data processing. In the subsystem of data analysis are realized methods of sentiment analysis, aspect extraction, aspect sentiment analysis, and decision tree. In subsystem of user interaction is visualized results of analysis.

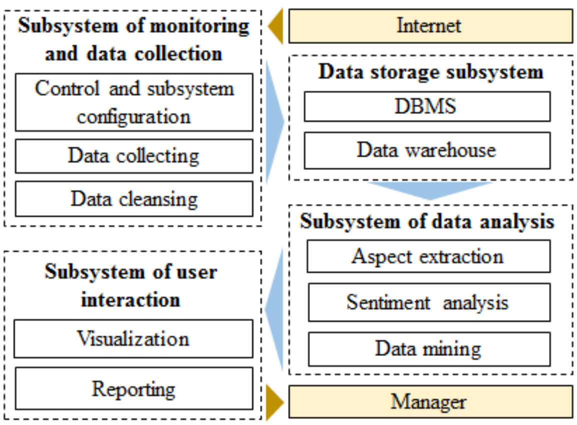

Figure 1. The architecture of Intelligent Decision Support System for product quality management.

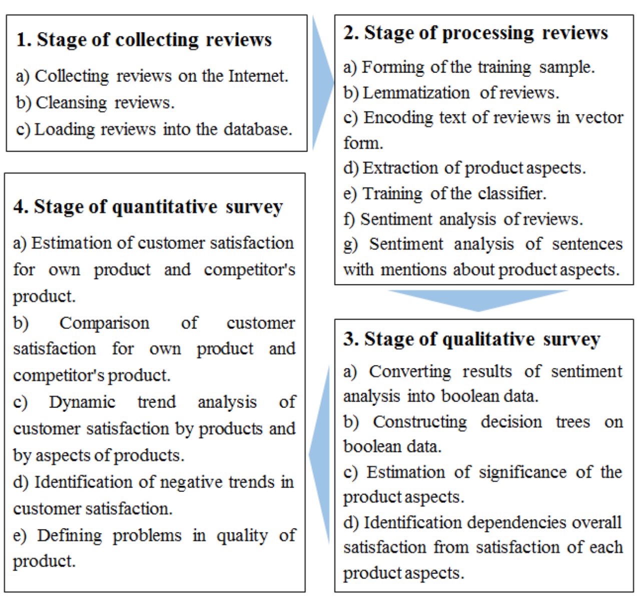

Figure 2. Working algorithm of Intelligent Decision Support System

In Figure 2, the algorithm of the IDSS is presented. It consists of four stages. The first stage includes collection of reviews from Internet resources, data cleansing and loading reviews into the database. The second stage performs processing collected reviews. It includes preprocessing procedures, such as preparing training samples of reviews, text lemmatization, and encoding text of reviews in vector form. Processing procedures include extraction of a product's aspects, training of the classifier and sentiment analysis.

The third stage is the quantitative survey. The quantitative survey is based on sentiment analysis of reviews entirely, and aspect sentiment analysis of sentences with mentions of a product's aspects. Sentiment classification is attained through binary scale – positive and negative sentiments. As a measure of the customer satisfaction is used a ratio of positive reviews (or positive sentences with mentions of a product's aspect) to the sum of positive and negative reviews (or sum of positive and negative sentences with mentions of a product's aspect). The output of the quantitative survey is values of customer satisfaction by a product and each product's aspect.

The fourth stage is the qualitative survey of customer satisfaction. It is based on transformation results of sentiment analysis into binary data and following by constructing of decision tree on it. The qualitative survey aims to identify how sentiment of review depends on the sentiment of different aspects of a product. Decision tree performs this aim and identifies latent relations between sentiment of a review and sentiment of a product's aspects. The output of the qualitative analysis is significance values of a product's aspects for customers and identifying latent relations extracted by decision tree. Managerial decision development and making is carried out on the basis of the performed quantitative and qualitative surveys.

IV. APPLIED ARTIFICIAL INTELLIGENCE TECHNIQUES

In this Section are described implemented AI techniques for customer satisfaction surveys and support decision making.

A. Data collection

Nowadays there are a large number of Internet resources where users can leave their opinions about products and services. The most popular examples of review sites are tophotels.ru (635 thousand reviews), yelp.com (53 million reviews), tripadvisor.com (130 million reviews). Similar resources continue to gain popularity. As opposed to social networking services, the advantage of review sites lies in their purpose - accumulation of customer reviews. One more advantage is that many of such resources have moderators of reviews and confirmation of author's objectivity, e.g., registration procedure.

There are two main types of collecting data from the Internet resources of customer reviews: 1) by using API (Application Programming Interface), and 2) by web data extraction. API is a set of ready-to-use tools - classes, procedures, and functions - provided by the application (Internet resource) for using in an external software product.

Unfortunately, only a few resources that accumulate reviews have API.

In this paper is used the second method for data collection – web data extraction. It is a process of automated content collection from HTML-pages of any Internet resource using special programs or script. Related work is presented in [14][15]. Scheme of reviews collection is presented in Figure 3. Web pages of review sources use HTML (HyperText Markup Language) that sets the structure typical for a review. Such structure includes separate blocks with the name of a product or company with a review, and other blocks with additional information. Therefore, all reviews are clearly identified in relation to the review object. It significantly simplifies the process of data collection in contrast to collecting messages from social networking services.

Input: review site

1. Gathering links on pages with reviews or generating links by using a template.
2. Setting boundaries for a content by using a HTML structure.
3. Collection data on the set of links.
4. Cleaning data and removing duplicates.
5. Loading data into the database.

Output: set of consumer reviews *D*

Figure 3. Algorithm of reviews collection

B. *Sentiment Analysis*

After data collection, it is possible to process review data with Text Mining tools. In this paper automatic sentiment analysis of reviews is used to evaluate product satisfaction. Sentiment stands for the emotional evaluation of author's opinion about a product that is referred to in the reviews.

There are three main approaches to sentiment analysis: 1) linguistic, 2) statistical, and 3) combined. The linguistic approach is based on using rules and vocabularies of emotionality words [16][17]. This approach is quite time-consuming due to the need of compiling vocabularies, patterns, and making rules for identifying sentiments. However, the main drawback of this approach is the impossibility to get a quantitative evaluation of the sentiment. The statistical approach is based on the methods of supervised and non-supervised machine learning (ML) [18][19]. The combined approach presupposes a combined use of the first two approaches.

In present work the methods of supervised machine learning is used - a Bayesian classifier and Support Vector Machines. Their realization in IDSS is based on techniques described by Pang and Lee [18][19]. More detailed information about implemented methods of sentiment analysis used in this paper can be found in [20][21]. In Figure 4 algorithms of learning and classification for naive Bayes classifier based on Multinomial model are presented. An advantage of these ML methods that they are quite easy in software implementation, and do not require making linguistic analyzers or sentiment vocabularies. They are able to evaluate sentiment quantitatively. For sentiment classification is used binary scale - positive and negative tonality. We use vector representation of review texts with help of the bag-of-words model. As attributes, we consider bit vectors - presence or absence of the word in the review text, and frequency vectors – a number of times that a given word appears in the text of the review. Lemmatization is also used.

Learning of naive Bayes classifier
Input: training set of reviews $D' = \{(d_1,c_1), \ldots, (d_m,c_m)\}$, set of classes $C = \{positive, negative\}$.

1. Extract all words from D' to the vocabulary V
2. For each $c \in C$ do
   3. Count documents N^c in each class c
   4. Calculate probability $p(c) = N^c / N$
   5. For each $w_t \in V$ do
      6. Count number of occurrences $K^c_{w_t}$ of word w_t in each class
      7. Calculate prob. $p(w_t \mid c) = (K^c_t + 1) / \sum_{t=1}^{|V|} (K^c_t + 1)$

Output: V, $p(c)$, $p(w_t \mid c)$

Classification with naive Bayes classifier
Input: review *d* from set *D, V,* $p(c)$, $p(w_t \mid c)$

1. Extract all words from *d* to the vocabulary V_d
2. For each $c \in C$ do
   3. Calculate $score[c] = \ln p(c)$
   4. For each $w_t \in V_d$ calculate $score[c] \mathrel{+}= \ln p(w_t \mid c)$
5. If $score[pos] > score[neg]$ than $d \in positive$ else $d \in negative$

Output: sentimental label *Sent* of review *d* (positive/negative)

Figure 4. Algorithm of naive Bayes classifier

C. *Aspect Sentiment Analysis*

Sentiment Analysis of reviews allows the evaluation of overall customer product satisfaction. However, it does not clearly show what customers like about a product and what they don't like. To answer this question, it is necessary to perform an aspect sentiment analysis. An aspect means characteristics, attributes, and properties that characterize the products, e.g., a "phone battery" or "delivery period". However, one sentiment object can have a great number of aspects. Furthermore, aspects in the text can be expressed by words-synonyms, e.g., "battery" and "accumulator". In this case, it makes sense to combine aspects into aspect groups.

An Aspect Sentiment Analysis of the review is a more difficult task and consists of two stages – identifying aspects and determining the sentiment of the comment on them. To complete the task of the Aspect Sentiment Analysis, we developed a simple and effective algorithm (see Figure 6). Aspects extraction based on the frequency of nouns and noun phrases mentioned in reviews based [22].

A frequency vocabulary [23] (created on text corpus) that helps to compare the obtained frequencies from reviews with frequencies from corpus is used to identify aspects. The nouns with maximum frequency deviations are claimants to be included into aspect groups. Clustering of the nouns into aspect groups was carried out manual. It should be noted, that if a sentence includes nouns from several aspect groups, then it would refer to opinion about each aspect group.

The results of sentiment Analysis and Aspect Sentiment Analysis can be presented in the form of text variables $Obj = (\mathrm{Re}\,v_i, Sent_i, Neg_{i1},..., Neg_{im}, Pos_{1i},..., Pos_{im})$, where *Obj* – a object or a product, Rev_i – text of the *i* review, $Data_i$ – date of *i* review publication, $Sent_i$ – sentiment of *i* review, Neg_{ij} – negative sentences with mention about the *j* aspect group in the *i* review, Pos_{ij} – positive sentences with mention about the *j* aspect group in the *i* review, *i* – number of review, *j* – number of aspect group, *m* – amount of aspect groups .

Aspect extraction
Input: set of reviews *D*

1. Extract all nouns *S* from the set of reviews *D*.
Count the frequency of words $\forall t = 1,|S| : f_t = N_t / N$ in the whole set of reviews D, where N – number of appearances of all words, N_t – number of appearances of the t noun.
2. Count the difference $\forall t : \Delta_t = f_t - f_t^v$ between the counted frequencies f_i and vocabulary frequencies f_t^V.
3. Sort the set of nouns *S* in descending order Δ_t.
4. Divide the set of nouns *S* from $\Delta_t > 0$ into aspect groups.

Output: set of aspect groups and aspect words

Aspect sentiment classification
Input: sentiment classifier, set of aspect groups and aspect words

1. Divide a set of reviews into set of sentences.
2. Perform sentiment classification for each sentence.
3. Check each sentence for the condition: if a sentence has a sentiment value (negative or positive) greater than a threshold *h* and contains at least one noun from any aspect group, then this sentence is labeled as an opinion (negative or positive) about the given product's aspect.

Output: positive and negative sentences with mentions about product's aspects $\{Neg_{i1},..., Neg_{im}, Pos_{i1},..., Pos_{im}\}$

Figure 5. Algorithm of Aspect Sentiment Analysis

D. Data Mining

The present paragraph suggests an algorithm of the following processing of results of sentiment analysis. The aim of the developed algorithm is to discover latent knowledge that can be used for decision support in product quality management. To realize this algorithm we use the Data Mining tool – Decision Tree, since this tool is easy to understand and interpret results. It also can explain relations between overall sentiment of review and sentiment of each aspects by means of Boolean logic.

The developed algorithm of knowledge discovery in results of sentiment analysis includes procedures presented in Figure 6. The described algorithm allows understanding of which sentiment sentences about a product's aspects influence the overall sentiment of review or, in other words, what product aspects influence customer satisfaction and in what way. The constructed Decision Tree model allows the consideration of the influence of not only separate sentiment sentences on aspects, but also their mutual presence (or absence) in the text on overall satisfaction. The Decision Tree model also allows the detection of the most significant product's aspects that are essential for the customer.

Input: positive and negative sentences with mentions about product's aspects $\{Neg_{i1},..., Neg_{im}, Pos_{i1},..., Pos_{im}\}$, vector of sentimental labels *Sent* of reviews.

1. Convert a text data $Obj = (\mathrm{Re}\,v_i, Sent_i, Neg_{i1},..., Neg_{im}, Pos_{1i},..., Pos_{im})$ into a boolean type by the following rules:
2. If $Sent_i = negative$, then $newSent_i = 1$, else $newSent_i = 0$
3. If $Neg_{ij} \neq null$, then $newNeg_{ij} = 1$, else $newNeg_{ij} = 0$
4. If $Pos_{ij} \neq null$, then $newPos_{ij} = 1$, else $newPos_{ij} = 0$
5. Creating a decision tree where the variable $newSent_i$ is a dependent variable from $\{newNeg_{i1},..., newNeg_{im}, newPos_{i1},..., newPos_{im}\}$
6. Estimation significances of aspect groups and interpretation of extracted rules

Output: significance values of product's aspects, latent relations between satisfaction by product and satisfaction by aspects

Figure 6. Algorithm of knowledge discovery

In Figure 7 an example of Decision Tree model is presented. Nodes of the Decision Tree are the aspect variables, i.e., presence or absence in the review tonality sentences (positive or negative) with mention about some aspect from aspect group. Edges of the tree are the values of aspect variables, i.e. 1 is presence, 0 is absence. Leaves present overall sentiment of review, i.e., each branch leads to either a positive review or a negative review that meets customer satisfaction or dissatisfaction. The Decision Tree model can be expressed both in the form of Boolean functions (see Eq. 1) in a disjunctive normal form, and in natural language as a rules. Each rule is characterized by measures of reliability and support. The reliability shows what percentage of reviews containing conditions of some rule, has the same sentiment corresponding to this rule. The support shows percentage of reviews that contain conditions of some rule regarding the entire number of reviews.

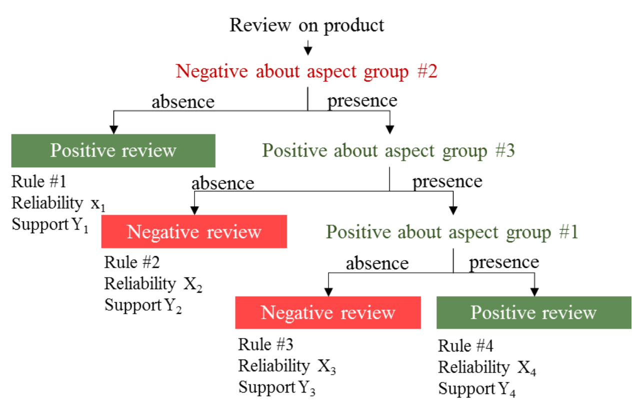

Figure 7. Example of the Decision Tree model

$$\begin{aligned}
&Rule\ \#1: \ \overline{Neg.a.g.\#2} \rightarrow Pos.\ review \\
&Rule\ \#2: \ Neg.a.g.\#2 \cap \overline{Pos.a.g.\#3} \rightarrow Neg.\ review \\
&Rule\ \#3: \ Neg.a.g.\#2 \cap Pos.a.g.\#3 \cap \overline{Pos.a.g.\#1} \rightarrow Neg.\ review \\
&Rule\ \#4: \ Neg.a.g.\#2 \cap Pos.a.g.\#3 \cap Pos.a.g.\#1 \rightarrow Pos.\ review
\end{aligned} \quad (1)$$

E. Measures for customer satisfaction

To measure customer satisfaction by products (or product's aspect group) we use a ratio of amount positive reviews (or positive sentences containing mentions about a product's aspect group) to all reviews (or all sentences containing mentions about product's aspect). The score of customer satisfaction *CS* by product is calculated by the formula:

$$CS = \frac{Z^{pos}}{Z^{pos} + Z^{neg}} \cdot 100\%, \quad (2)$$

where Z^{pos} – the number of positive reviews, Z^{neg} – the number of negative reviews.

The score of customer satisfaction cs_j by j product's aspect group is calculated by the formula:

$$cs_j = \frac{z_j^{pos}}{z_j^{pos} + z_j^{neg}} \cdot 100\%, \quad (3)$$

where z_j^{pos} – number of positive sentences containing mention about the j product's aspect group, z_j^{neg} – number of negative comments containing mention about the j product's aspect group.

Significance of aspects group shows how much the sentiment of a review depends on the aspect group in positive and negative sentences. Let the number of aspect groups is *g/2*, then the number of independent variables *g* (positive and negative statements for each group of aspect). According to the methodology described in [26] the formula for calculating the significance of variable *m* is:

$$Sign_m = \frac{\sum_{j=1}^{k_m} \left(E_{m,j} - \sum_{i=1}^{q_{m,j}} E_{m,j,i} \cdot \frac{Q_{m,j,i}}{Q_{m,j}} \right)}{\sum_{l=1}^{g} \sum_{j=1}^{k_l} \left(E_{l,j} - \sum_{i=1}^{q_{l,j}} E_{l,j,i} \cdot \frac{Q_{l,j,i}}{Q_{l,j}} \right)} \cdot 100\%, \quad (4)$$

where k_l – number of nodes that were split by attribute l, $E_{l,j}$ – entropy of the parent node, split by attribute l, $E_{l,j,i}$ – subsite node for j, which was split by attribute l, $Q_{l,j}$, $Q_{l,j,i}$ – number of examples in the corresponding nodes, $q_{l,j}$ – number of child nodes for j parent node.

V. Experiment

Efficacy evaluation of the developed IDSS was performed on the data obtained from 635 824 reviews about hotels in the Russian language. The reviews have been collected from the popular Internet resource tophotels.ru for the period of 2003-2013. The initial structure of the collected data consisted of the following fields: hotel name; country name; resort name; visit date; review's text; author's ratings of placement, food, and service. The data was preprocessed and loaded into the database SQL Server 2012.

Classifying sentiment used a binary scale (negative and positive) on the hypothesis that the absence of negative is positive for the product. A training set of positive and negative reviews was formed using the collected data on an author's ratings of placement, food, and service. The review site tophotels.ru uses a five-point grading scale. A review can have a maximum total rating of 15 points, and minimum total rating of 3 points. The training set included 15790 negative reviews that have 3 and 4 total points, and 15790 positive reviews that have 15 total points. We did not use the author's ratings for further data processing. Classification of another 604 244 reviews was carried out using a trained classifier.

TABLE I. COMPARISON OF METHODS FOR SENTIMENT CLASSIFICATION

#	Machine learning methods	Vector	Accuracy	
			Test No.1	*Test No.2*
1	SVM (linear kernel)	Frequency	94,2%	83,1%
2	SVM (linear kernel)	Binary	95,7%	84,1%
3	NB	Binary	96,1%	83,7%
4	NB	Frequency	97,6%	92,6%
5	NB (stop-words)	Frequency	97,7%	92,7%
6	Bagging NB	Frequency	97,6%	92,8%
7	NB (negations)	Frequency	98,1%	93,6%

For the purpose of training an effective sentiment classifier, the accuracy of classification was evaluated for machine learning methods and some peculiarities of their realization (see Table 1). The measure accuracy as a ratio of the number of correctly classified reviews to total number of reviews was used to estimate classification accuracy. Accuracy estimation was performed on two sets of data. The first set (Test No.1) represented a training set of strong positive (15 790) and strong negative reviews (15 790). Classifiers were tested by using cross validation by dividing the first set into 10 parts. The second set (Test No. 2) included random reviews from initial set of reviews (635 824) with different points (3-15 points) and was labeled manual (497 positive and 126 negative). It was used only for accuracy control of the classifier that had been trained on the first data set.

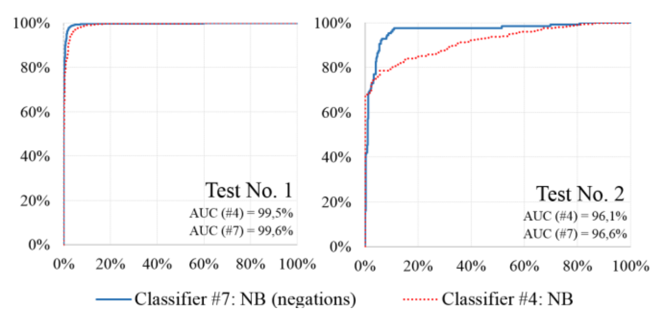

Figure 8. Comparison of ROC-curves classifiers number 7 and number 4

To estimate influence of negative particles “not” and “no”, the tagging technique was used; for example, the phrase “not good” was marked as “not_good”, and was regarded by the classifier as one word. This technique allowed the increasing of sentiment classification accuracy. Accuracy values are presented in Table 1. The most efficient ML method was naive Bayes classifier with negation techniques (#7). In Figure 8 are presented ROC-curves classifiers #4 and #7. The classifier #7 was trained on the training set and was used for further sentiment analysis.

Using the aspect extraction algorithm (Section III), we extracted the nouns that were divided into seven basic aspect groups (see Figure 9): "beach/swimming pool", "food", "entertainment", "place", "room", "service", "transport". The following step was extracting and sentiment classification sentences with words from aspect groups using classifier #7. However, not all sentences with aspects had a clearly expressed sentiment; therefore, the sentences with poorly expressed sentiment using threshold *h* were filtered out.

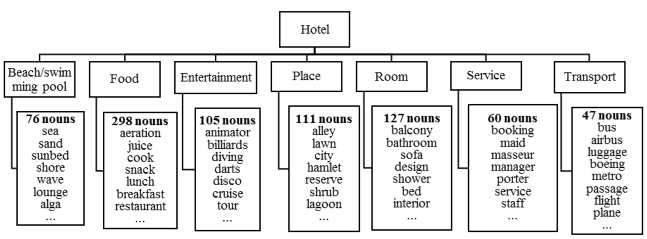

Figure 9. Aspect groups of object “hotel”

In present work we give an example of qualitative and quantitative surveys for two 5-star hotels “A” (1692 reviews) and “B” (1300 reviews) located on the resort Sharm el-Sheikh (63 472 reviews) in Egypt. Firstly, we will make a quantitative survey, measure customer satisfaction, compare it with average satisfaction in the whole resort, detect negative trends by each hotel's aspect group, and identify problems in the quality of hotels.

The dynamics of customer satisfaction is presented in Figure 10. Concerning the hotel “A”, there is a positive upward satisfaction trend from 2009, and it fixes on the average-resort level in 2013. Concerning the hotel “B”, in 2012 there was a sharp satisfaction decline and the same sharp increase in 2013. For the hotel “B”, satisfaction decrease started in June 2012, and stopped in October 2012. Then, customer satisfaction grew to the level that was higher than the average resort level being ahead of its competitor – hotel “A”.

To find reasons of hotel “B” satisfaction decrease, we will examine the diagrams in Figure 11. We can see that in 2012, the hotel “B” on average was second to the hotel “A” in such aspects as “room” (Δ12%), “place” (Δ8%), “service” (Δ5%), “beach/swimming pool” (Δ3%) and “entertainment” (Δ3%). Besides, in 2012, the Hotel “B” had more registered cases of intoxication, as well as cases of theft in August 2012. We should also note that one of the reasons of customer dissatisfaction with the hotel “B” was the initiated repair of hotel rooms and buildings which, however, paid off in 2013. Customer satisfaction with the hotel “A” aspects conforms with the average resort level.

In 2013, customer satisfaction with the hotel “B” exceeded the average level in all aspects (see Figure 12). Customer satisfaction with the hotel “A” dropped lower than average values in such aspects as “service” (Δ3%), “food” (Δ3%), “beach/swimming pool” (Δ3%) and “transport” (Δ4%). For hotel “A” manager arise questions like which aspects are the most significant for the customer and that should be improved in the first place, is it possible to “substitute” the dissatisfaction with the service, e.g., by tasty food or employ new entertainer?

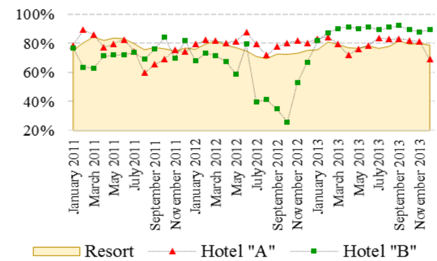

Figure 10. Dynamics of the customer satisfaction by months.

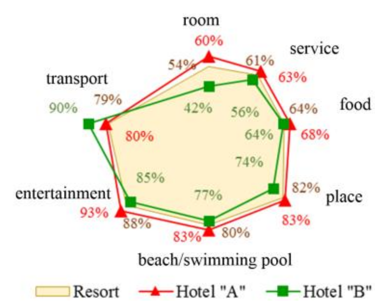

Figure 11. Comparison of the consumer satisfaction by aspects in 2012.

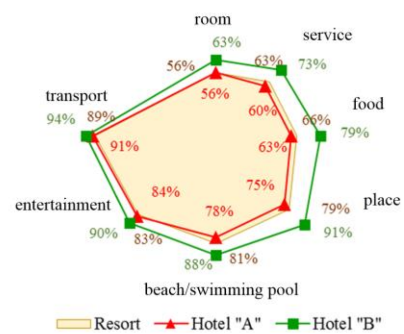

Figure 12. Comparison of the consumer satisfaction by aspects in 2013.

Decision Trees were constructed using algorithm C4.5. At the first step was constructed a tree for the all hotels of resort. Extracted rules are represented in Table 3. At the second step trees were constructed for hotel “A” and hotel “B”. Significance values of a product’s aspect groups are represented in Table 2. In Figure 13, there are the decision trees created for hotel “A” and hotel “B”. Due to the large size of the produced decision tree of the whole resort we omitted it, but in Table 3 its rules are presented that have reliability >80% and support >5% (5 rules from 27).

By analyzing significance values (see Table 2), we can say that the main factors of consumer dissatisfaction are a low service level (34,8%), problems with food (16%), and complaints about the hotel rooms (4%). The most critical aspect group for the hotel "B" is "room" (57,3%). In absence of negative opinions on the aspect group "room", the review would be positive with a reliability of 95,5% (see Table 3, rule #10). That is why the repair that was performed facilitated to a significant increase of consumer satisfaction. The most critical aspect group for the Hotel "A" is "service" that corresponds with the resort in a whole.

TABLE II. SIGNIFICANCE OF PRODUCT'S ASPECT GROUPS

Aspect group	Kano's model category	Sentiment of mention	Significance values		
			Resort	*Hotel "A"*	*Hotel "B"*
Service	Must-be quality	Negative	34,8%	60,2%	-
		Positive	0,7%	-	-
Food	One-dimensional quality	Negative	30,3%	27,2%	30,3%
		Positive	16%	-	-
Entertainment	Attractive quality	Negative	-	-	-
		Positive	8,5%	12,7%	12,4%
Room	One-dimensional quality	Negative	4%	-	57,3%
		Positive	2,1%	-	-
Beach/swimming pool	Attractive quality	Negative	0,2%	-	-
		Positive	2,5%	-	-
Territory	Attractive quality	Negative	-	-	-
		Positive	1%	-	-
Transport	Indifferent quality	Negative	-	-	-
		Positive	-	-	-

Using significance values, we can relate each aspect group with Kano's model categories [25]. Aspect group "service" has high significance on customer satisfaction in a "negative" case, but significance is near zero in a "positive" case (34,8% vs. 0,7%). It relates to "Must-be quality" of Kano's categories. That's why positive sentences with mentions about such aspect groups as "service" do not have an influence on sentiment of review, i.e., on overall satisfaction with hotel. That means the consumer a priori awaits a high-level service as a matter of course. For such aspect groups as "food" and "room" significance values are comparable (30,3% vs. 16% for "food" and 4% vs. 2,1% for "room"), that relates to "One-dimensional quality" of Kano's categories. Aspect groups "beach/swimming pool", "entertainment" and "territory" relates to "Attractive quality" because they have a significance on customer satisfaction in "positive" case only.

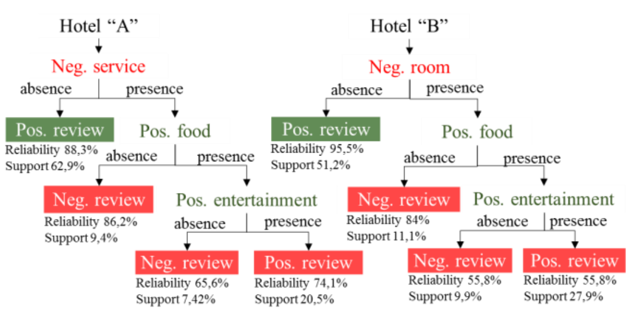

Figure 13. Decision trees for hotels

In some cases, positive mentions about "food" and "entertainment" simultaneously in a review could substitute negative mentions about "services" and provide a positive review. That's why the hotel's aspects which are contributing to customer satisfaction and important for both the resort and for the hotels are good food (30,3%) and amusing entertainment activities (8,5%). Customer satisfaction with these aspect groups can overlap dissatisfaction with "service" or "rooms" and make the customer overall satisfied (see Table 3, rules #5, #7, #11).

TABLE III. RULES EXTRACTED BY USING DECISION TREES

#	Rules	S[a]	R[b]
	Extracted rules on resort reviews		
1	$\text{Food}^{+} \cap \overline{\text{Service}^{-}} \cap \overline{\text{Food}^{-}} = \text{Positive review}$	37,2%	97,4%
2	$\text{Food}^{+} \cap \overline{\text{Service}^{-}} \cap \overline{\text{Food}^{-}} \cap \text{Beach}^{+} = \text{Positive review}$	11%	86,2%
3	$\overline{\text{Food}^{+}} \cap \overline{\text{Service}^{-}} \cap \overline{\text{Service}^{-}} \cap \overline{\text{Room}^{-}} = \text{Positive review}$	10,6%	83,9%
4	$\overline{\text{Food}^{+}} \cap \text{Service}^{-} \cap \overline{\text{Entertainment}^{+}} = \text{Negative review}$	6,9%	92,3%
5	$\text{Food}^{+} \cap \text{Service}^{-} \cap \text{Food}^{-} \cap \text{Entretainment}^{+} = \text{Positive review}$	5,8%	88,4%
	Extracted rules on Hotel "A" reviews		
6	$\overline{\text{Service}^{-}} = \text{Positive review}$	62,9%	88,3%
7	$\text{Food}^{+} \cap \text{Service}^{-} \cap \text{Entertainment}^{+} = \text{Positive review}$	20,5%	74,1%
8	$\overline{\text{Food}^{+}} \cap \text{Service} = \text{Negative review}$	9,4%	86,2%
9	$\text{Food}^{+} \cap \text{Service}^{-} \cap \overline{\text{Entertainment}^{+}} = \text{Negative review}$	7,2%	65,6%
	Extracted rules on Hotel "B" reviews		
10	$\overline{\text{Room}^{-}} = \text{Positive review}$	51,2%	95,5%
11	$\text{Food}^{+} \cap \text{Room}^{-} \cap \text{Entertainment}^{+} = \text{Positive review}$	27,9%	81%
12	$\overline{\text{Food}^{+}} \cap \text{Room}^{-} = \text{Negative review}$	11,1%	84%
13	$\text{Food}^{+} \cap \text{Room}^{-} \cap \overline{\text{Entertainment}^{+}} = \text{Negative review}$	9,9%	55,8%

a. Support. b. Reliability

The performed qualitative survey allowed the detection of the main ways to increase customer satisfaction for hotel "A". The problem aspect groups identified through quantitative survey correspond to the most significant aspects detected during the qualitative research. Hotel "A" manager should firstly increase service quality, and then increase the quality of "food" and "beach/swimming pool" maintenance. "Transport" problems – concerning flights, early check-in, and baggage storage – are not significant for customers and can be solved in the frames of service improvement. The process of service quality increase can take much time; that is why organizing entertainment and animated programs together with enhancement of restaurant service could be immediate measures for increasing customer satisfaction. Specification of managerial decisions can be performed on the basis of the information on existing problems contained in negative reviews. The extracted sentences on aspects can be directed to the appropriate hotel services.

VI. CONCLUSION AND FUTURE WORK

Poor quality of products and services contributes to a decrease of customer satisfaction. On the other hand, under the conditions of stiff competition, there are no barriers for

 ISBN: 978-1-61208-358-2

the consumer to change the supplier of goods and services. All these things can cause loss of clients and a decrease of a company's efficiency indexes. Therefore, maintaining high-quality standards should be provided by effective managerial decisions and based on opinion mining as a feedback.

The suggested conception of decision support based on the developed approach of text data processing and analysis allows performing quantitative and qualitative surveys of customer satisfaction using computer-aided procedures, and making effective managerial decisions on product quality management. The present conception allows effective reduction of labor intensity of customer satisfaction research that makes it available for use by a wide range of companies. A prototype of IDSS was developed on the basis of the suggested conception. The performed experiment has proved its efficacy for solving real problems of quality management and consistency of the results obtained. Future research on the given topic can be devoted to automatic annotating of text data, representing text amount of review in the form of a summary, and extracting useful and unique information.

REFERENCES

[1] ISO 9000-2008 The quality management system. Fundamentals and vocabulary.

[2] ISO10004:2010 Quality management. Customer satisfaction. Guidelines for monitoring and measuring.

[3] D. Gräbner, M. Zanker, G. Fliedl, and M. Fuchs, "Classification of Customer Reviews based on Sentiment Analysis", Proceedings of the International Conference in Helsingborg, Springer Vienna, Jan. 2012, pp. 460-470, ISBN 978-3-7091-1141-3.

[4] Y. Jo, A. Oh, "Aspect and Sentiment Unification Model for Online Review Analysis", Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11), ACM New York, Feb. 2011, pp. 815-824, ISBN: 978-1-4503-0493-1.

[5] B. Lu, M. Ott, C. Cardie, and B. Tsou, "Multi-aspect Analysis with Topic Models", Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW '11), Dec. 2011, pp. 81-88, ISBN: 978-0-7695-4409-0.

[6] B. Liu, M. Hu, and J. Cheng, "Opinion Observer: Analyzing and Comparing Opinions on the Web", Proceedings of the 14th international conference on World Wide Web (WWW '05), ACM New York, May 2005, pp. 342-351, ISBN: 1-59593-046-9.

[7] W. Kasper, M. Vela, "Sentiment Analysis for Hotel Reviews.", Proceedings of the Computational Linguistics-Applications Conference (CLA-2011), Oct. 2011, pp. 45–52.

[8] S. Blair-Goldensohn, K. Hannan, and R. McDonald, "Bilding a Sentiment Summarizer for Local Service Reviews", Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL '09), March 2009, pp. 514-522.

[9] J. Ajmera, H. Ahn, M. Nagarajan, A. Verma, D. Contractor, S. Dill, and M. Denesuk, "A CRM system for Social Media", Proceedings of the 22nd international conference on World Wide Web (WWW '13), May 2013, pp. 49-58, ISBN: 978-1-4503-2035-1.

[10] H. Wang, Y. Lu, and C. Zhai, "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach", Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '10), ACM New York, July 2010, pp. 783-792, ISBN: 978-1-4503-0055-1.

[11] J. C. de Albornoz, L. Plaza, P. Gervás, and A. Díaz, "A Joint Model of Feature Mining and Sentiment Analysis for Product Review Rating", Proceedings of the 33rd European Conference on Information Retrieval (ECIR '11), April 2011, pp. 55-66, ISBN: 978-3-642-20160-8.

[12] H. Wachsmuth, M. Trenkmann, B. Stein, G. Engles, and T. Palakarska, "A Review Corpus for Argumentation Analysis", Proceedings of the 15th International Conference on Computational Linguistics and Intelligent Text Processing, Springer (Kathmandu, Nepal), LNCS, April 2014, pp. 115-127, ISBN 978-3-642-54902-1.

[13] M. Bank, "AIM – A Social Media Monitoring System for Quality Engineering", PhD thesis, Universität Leipzig, June 2013, p. 235.

[14] J. Thomsen, E. Ernst, C. Brabrand, and M. Schwartzbach, "WebSelF: A Web Scraping Framework", Proceedings of the 12th international conference on Web Engineering (ICWE 2012), July 2012, pp. 347-361, ISBN: 978-3-642-31752-1.

[15] R. Penman, T. Baldwin, and D. Martinez, "Web scraping made simple with sitescraper", 2009. [Online]. Available from: http://sitescraper.googlecode.com.

[16] J. Yi, T. Nasukawa, W. Niblack, and R. Bunescu, "Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques", In Proceedings of the 3rd IEEE international conference on data mining, ICDM 2003, pp. 427-434.

[17] A. G. Pazelskaya, A. N. Soloviev, "Method of the determination emotions in the lyrics in Russian", Computer program linguistics and intellectual technologies, Issue 10 (17), 2011, pp. 510-522.

[18] B. Pang, L. Lee, "Thumbs up? Sentiment Classification using Machine Learning Techniques", Proceedings of the Conference on Empirical Methods in Natural. Language Processing (EMNLP), Philadelphia, July 2002, pp. 79-86.

[19] C. Manning, P. Raghavan, and H. Schuetze, "An Introduction to Information Retrieval", Cambridge University Press. Cammbridge, England, 2009, pp. 1-544.

[20] N. Yussupova, D. Bogdanova, and M. Boyko, "Algorithms and software for sentiment analysis of text messages using machine learning", Vestnik USATU, T. 16-6(51), 2012, pp. 91-99.

[21] N. Yussupova, D. Bogdanova, and M. Boyko, "Applying of Sentiment Analysis for Texts in Russian Based on Machine Learning Approach", Proceedings of the 2nd International Conference on Advances in Information Mining and Management (IMMM2012), Venice, Italy, 2012, pp. 8-14. ISBN: 978-1-61208-227-1.

[22] C. Scaffidi, K. Bierhoff, E. Chang, M. Felker, H. Ng, and C. Jin, "Red Opal: Product-Feature Scoring from Reviews", Proceedings of the 8th ACM conference on Electronic commerce, 2007, pp. 182-191. ISBN: 978-1-59593-653-0.

[23] O. N. Ljashevskaja, S. A. Sharov, "The new frequency vocabulary of Russian lexic based on the Russian National Corpus", RAS, Institut of Russian language named of V. V. Vinogradov, 2009. [Online]. Available from: http://dict.ruslang.ru/freq.php.

[24] J. Yu, Z.-J. Zha, M. Wang, and T.-S. Chua, "Aspect Ranking: Identifying Important Product's aspects from Online Consumer Reviews", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT '11), June 2011, pp. 1496-1505, ISBN: 978-1-932432-87-9.

[25] N. Kano, N. Seraku, F. Takashi, and S. Tsuji, "Attractive quality and must-be quality", In The Journal of the Japanese Society for Quality Control, 1984, V. 14, pp. 39-48.

[26] Deductor. The Algorithm Manual (ver. 5.2.0), BaseGroup Labs, 2010. [Online]. Available from: http://www.basegroup.ru/download/guide_algorithm_5.2.0.pdf

[27] M. Taboada, J. Brooke, M. Tofiloski, K. D. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," Computational Linguistics, vol. 37(2), pp. 267-307, June 2011, doi:10.1162/COLI_a_00049.

[28] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," Journal of Machine Learning Research, vol. 3 (4–5), pp. 993-1022, Jan. 2003, doi:10.1162/jmlr.2003.3.4-5.993.