

МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ УПРАВЛЕНИИ КАЧЕСТВОМ ПРОДУКЦИИ НА ОСНОВЕ АНАЛИЗА ТЕКСТОВОЙ ИНФОРМАЦИИ

Юсупова Н.И.¹, Богданова Д.Р.¹, Бойко М.В.¹

¹ФГБОУ ВПО «Уфимский государственный авиационный технический университет», Уфа, Россия (450000, Уфа, ул. К. Маркса, 12), e-mail: russian_max@inbox.ru

В работе рассматривается задача поддержки принятия решений при управлении качеством продукции. Для ее эффективного решения предлагается концептуальный подход к исследованию удовлетворенности потребителей на основе глубокого анализа потребительских отзывов, размещаемых в Интернете на естественном языке с применением технологий искусственного интеллекта (ИИ). В качестве инструментов ИИ используются анализ текста, сентимент-анализ, аспектный сентимент-анализ, интеллектуальный анализ данных и машинное обучение. В качестве источников данных используются специализированные интернет-ресурсы для аккумуляции потребительских отзывов, например, tophotels.ru. На основе разработанного подхода реализован прототип системы поддержки принятия решений, который позволяет проводить качественное и количественное исследование удовлетворенности потребителей. Для оценки эффективности предлагаемого подхода был проведен эксперимент по количественному и качественному исследованию удовлетворенности клиентов отелей. Полученные результаты свидетельствуют об эффективности предложенного подхода к поддержке принятия решений при управлении качеством продукции и перспективе его использования вместо классических методов количественного и качественного исследования удовлетворенности потребителей.

Ключевые слова: управление качеством продукции, поддержка принятия решений, сентимент-анализ.

DECISION SUPPORT FOR QUALITY MANAGEMENT BASED ON ARTIFICIAL INTELLIGENCE APPLICATIONS FOR UNSTRUCTURED DATA ANALYSIS

Yussupova N.I.¹, Bogdanova D.R.¹, Boyko M.V.¹

¹Ufa State Aviation Technical University, Ufa, Russia (450000, Ufa, K. Marx's street, 12), e-mail: Russian_max@inbox.ru

The paper describes the application of decision support in product quality management. To effectively complete this application we propose a conceptual approach to research customer satisfaction based on a profound analysis of consumer reviews put on the Internet in natural language with use of Artificial Intelligence (AI) techniques such as Text Mining, Aspect Sentiment Analysis, Data Mining and Machine Learning are used as tools of AI. Special Internet resources for accumulating customer reviews, such as yelp.com, tripadvisor.com, tophotels.ru, are used as data resources. On the basis of the developed approach, a prototype of the Decision Support System that allows to carry out a qualitative and quantitative research of customer satisfaction was realized. To evaluate the efficacy of the proposed approach, we have carried out an experiment on qualitative and quantitative research of hotel client satisfaction. The obtained results prove the efficacy of the proposed approach to decision support in product quality management and the perspective of using it instead of classical methods of qualitative and quantitative research of customer satisfaction.

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

Введение

Для обеспечения качества продукции предприятию необходимо принимать эффективные управленческие решения. Разработка управленческих решений и их принятие должны основываться на знаниях и закономерностях, получаемых в ходе анализа собираемой информации. Для предприятий такой информацией является информация о степени удовлетворенности потребителей (УП), которая выражается в форме потребительских мнений о качестве продукции. Поэтому при управлении качеством

ключевой информацией от которой зависит принятие тех или иных решений является удовлетворенность потребителей.

Для сбора данных и оценки УП международный стандарт качества ИСО 10004 рекомендует использовать следующие методы: личные интервью, интервью по телефону, дискуссионные группы, корреспондентское (почтовая рассылка анкет) исследование и обследование в онлайн-режиме (анкетирование) [1]. Общим недостатком рекомендуемых методов является необходимость выполнения большого объема ручной работы: подготовка вопросов анкеты, подбор респондентской базы, рассылка анкеты и сбор результатов, проведение персональных интервью, подготовка отчета по результатам. Все это приводит к удорожанию исследования. Также из-за характерной им дискретности методы не позволяют вести непрерывный мониторинг за уровнем УП. По этой причине анализ данных ограничивается одним временным срезом и не дает представлений о тенденциях и динамике УП. Также страдает своевременность принятия управленческих решений, которая зависит от частоты поступления актуальных данных о мнениях потребителей.

Вопросы вызывают существующие шкалы для оценки УП и субъективность их восприятия респондентами. Получаемые оценки УП выражаются в виде абстрактных индексов удовлетворенности, которые сложны для понимания, сравнения и интерпретации результатов. Методы анализа собираемых данных рекомендуемых стандартом ИСО 10004 позволяют выявлять только линейные зависимости.

Целью исследования является повышение эффективности процесса выработки и принятия решений при управлении качеством товаров и услуг. В данной работе для повышения эффективности управления качеством продукции предлагается подход к поддержке принятия решений на основе исследования УП с применением технологий ИИ.

Предлагаемый подход к принятию решений

Предлагаемый подход поддержки принятия управленческих решений при управлении качеством продукции реализуется посредством объединения методов и алгоритмов сбора и обработки данных об УП в единую интеллектуальную систему поддержки принятия решений (ИСППР). Схема управления качеством продукции, разработанной с помощью ИСППР, представлена на Рисунке 1. Основные действующие лица в процессе управления качеством: потребители продукции и лицо принимающее решения (ЛПР). Под объектом управления понимается качество предоставляемых услуг, от которого зависит эффективность деятельности предприятия. Под субъектом управления понимается ЛПР, принимающее решения по улучшению качества услуг. Под управляющим воздействием понимаются управленческие решения, оказывающие влияние на качество услуги, например, модернизация товара, выбор контрагентов, определение ценовой политики, подбор

персонала заданной квалификации, внедрение инноваций, повышение уровня обслуживания, обеспечение ресурсами и др. Процесс выработки и принятия решений обеспечивается посредством интеллектуальной системы поддержки принятия решений (ИСППР). В соответствии со схемой, предприятие производит товары или услуги и поставляет их потребителям. После их использования, потребители опубликовывают свои мнения о качестве продукции в сети Интернет в виде текстовых отзывов. ИСППР производит автоматический сбор отзывов, их очистку и загрузку в подсистему хранения данных. В подсистеме анализа данных ИСППР производится обработка отзывов, оценка УП и извлечение знаний. Результаты анализа отзывов визуализируются в подсистеме взаимодействия с пользователем. На основе полученных результатов ЛПР принимает управленческие решения по повышению качества обслуживания.

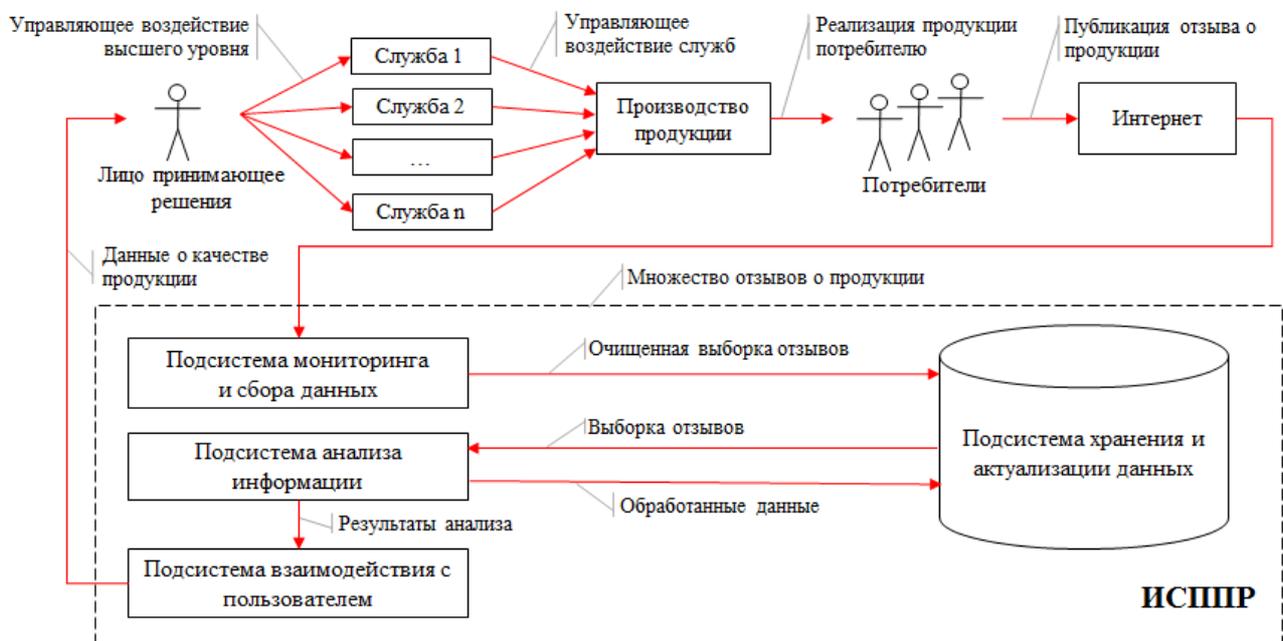


Рисунок 1 – Схема управления качеством услуг с использованием ИСППР

На Рисунке 2 представлен алгоритм работы ИСППР. Он состоит из четырех основных этапов. Первый этап содержит процедуры сбора отзывов с интернет-ресурсов, очистку данных и загрузки в базу данных. На втором этапе производится обработка и анализ собранных отзывов. Она включает разметку отзывов по их эмоциональной тональности (например, негативные и позитивные), идентификацию аспектов продукции, определения тональности отдельных высказываний об аспектах. После этапа обработки данных с помощью средств визуализации проводится количественное исследование УП. Качественное исследование УП проводится с помощью построения моделей на основе деревьев решений, где тональность отзыва выступает в качестве зависимой переменной, а тональные высказывания по аспектам продукции являются независимыми переменными. На основе проведенных исследований производится разработка и принятие управленческих решений.



Рисунок 2 – Алгоритм работы интеллектуальной системы поддержки принятия решений

Применяемые технологии искусственного интеллекта

Сбор данных. Сегодня существует большое число интернет-ресурсов, где пользователи могут оставлять свои отзывы о товарах и услугах. Наиболее популярными примерами являются tophotels.ru (635 тыс. отзывов), yelp.com (53 млн. отзыва), tripadvisor.com (путешествия, 130 млн. отзывов). Их достоинство как источника данных для оценки УП заключается в их назначении-аккумуляции потребительских отзывов. В отличие от социальных сетей, страницы интернет-баз используют XML-разметку, которая задает структуру, свойственную отзыву. Такая структура содержит отдельные блоки с названием продукции или предприятия, с отзывом и другие блоки с дополнительной информацией. Это значительно упрощает процесс сбора данных и исключает проблему многозначности ключевых слов. Еще одним достоинством является то, что на многих подобных ресурсах существует модерация отзывов и подтверждения объективности автора.

Существует два основных способа сбора данных с интернет-ресурсов потребительских отзывов: 1) использование API (интерфейс программирования приложений) и 2) web-парсинг. API представляет собой набор готовых инструментов – классов, процедур, функций, предоставляемых приложением (интернет-ресурсом) для использования во внешних программных продуктах. К сожалению, немногие ресурсы, аккумулирующие отзывы, имеют свой API. В этом случае для сбора отзывов можно применить второй способ сбора данных – web-парсинг. Под web-парсингом понимается процесс автоматизированного анализа и сбора контента с html-страниц какого-либо интернет-ресурса с помощью специальных программ.

Анализ тональности отзывов. Когда данные собраны и очищены, можно приступить к процессу их обработки с помощью инструментов анализа текста. Для оценки удовлетворенности автора продукцией в работе используется сентимент-анализ. Под тональностью или сентиментом понимается эмоциональная оценка мнения автора по отношению к объекту, о котором идет речь в тексте.

Можно выделить три основных подхода к анализу тональности: 1) лингвистический, 2) статистический, 3) комбинированный. Лингвистический подход основан на применении правил и тональных словарей [3, 6]. Он достаточно трудоемок из-за необходимости составления тональных словарей, шаблонов и построения правил определения тональности. Но главным недостатком подхода является невозможность получения количественной оценки тональности. Статистический подход основан на методах машинного обучения с учителем и без него [7, 8].

В данной работе используется подход на основе методов машинного обучения с учителем – наивный байесовский классификатор и метод опорных векторов. Они достаточно просты в программной реализации, не требуют построения лингвистических анализаторов или словарей тональности. Оценка тональности текста может выражаться количественно. Для применения данных методов строилась обучающая выборка. Для описания признакового пространства использовалось векторное представление текста отзывов с помощью модели «мешка слов» (bag-of-words). В качестве признаков рассматривались бинарные векторы – наличие или отсутствие слова в тексте отзыва и частотные векторы – число вхождений слова в текст отзыва. Также использовалась процедура лемматизации, приводящая все слова отзыва к начальной форме. Более подробней с используемыми в данной работе методами машинного обучения можно ознакомиться в работах [4, 5].

Аспектный сентимент-анализ. В работе используется тональный анализ высказываний об аспектах или аспектный сентимент-анализ. Под аспектом тональности понимаются характеристики, атрибуты, качества, свойства, характеризующие продукцию, например, аккумулятор телефона или время доставки и т.п. При большом количестве аспектов имеет смысл объединение отдельных аспектов в аспектные группы. Пример подобных аспектных групп представлен на Рисунке 3. Аспектный анализ тональности отзыва является более сложной задачей и состоит из двух этапов – идентификации аспектов и определения тональности высказывания о них. Для решения задачи аспектного анализа тональности был разработан простой и эффективный алгоритм:

Первый этап.

1. Извлечь все существительные S на множестве отзывов D .

2. Посчитать частоты употреблений существительных $\forall i = 1, \overline{|S|} : f_i = N_i / N$ на всем множестве отзывов D , где N – число употреблений всех слов, N_i – число употреблений i -го существительного.

3. Посчитать разность $\forall i : \Delta_i = f_i - f_i^v$ между посчитанными частотами f_i и словарными частотами f_i^v .

4. Отсортировать множество существительных S по убыванию Δ_i .

5. Разбить множество существительных S с $\Delta_i > 0$ на аспектные группы.

Второй этап.

1. Разбить множество отзывов на множество предложений.

2. Для каждого предложения выполнить классификацию тональности.

3. Для каждого предложения выполняется проверка: если оценка тональности предложения (негативная или позитивная) превышает некоторый порог h и содержит хотя бы одно существительное из какой-либо аспектной группы, то предложение помечается как позитивное или негативное мнение о данной аспектной группе.

Результаты sentiment-анализа отзывов и аспектного sentiment-анализа представляют собой набор текстовых данных $\langle Rev_i, \dots, Sent_i, Neg_{i_1}, \dots, Neg_{i_n}, Pos_{i_1}, \dots, Pos_{i_n} \rangle$, где Rev_i – текст i -го потребительского отзыва, $Sent_i$ – тональность i -го отзыва, Neg_{ji} – негативные мнения о j -й аспектной группе в i -м отзыве, Pos_{ji} – позитивные мнения о j -й аспектной группе в i -м отзыве, i – номер отзыва, j – номер аспектной группы.

Последующая обработка данных с помощью деревьев решений. Для проведения качественного анализа был разработан оригинальный метод последующей обработки результатов sentiment-анализа на основе деревьев принятия решений. Построение деревьев решений производилось при помощи алгоритма C4.5. Результаты, получаемые с его помощью, позволяют понять, какие аспектные группы продукции и каким образом влияют на УП. Преимущество разработанного метода анализа – в том, что он позволяет выявлять нелинейные зависимости между общей удовлетворенностью продукцией и удовлетворенностью ее отдельными аспектными группами. Также метод позволяет выявить значимые аспекты продукции и получить ее количественные оценки. Метод состоит из следующих процедур:

1. Преобразование набора текстовых данных $\langle Rev_i, \dots, Sent_i, Neg_{i_1}, \dots, Neg_{i_n}, Pos_{i_1}, \dots, Pos_{i_n} \rangle$ в логический тип данных по следующим правилам:

1.1. Если $Sent_i = negative$, то $newSent_i = 1$, иначе $newSent_i = 0$;

1.2. Если $Neg_{ji} \neq null$, то $newNeg_{ji} = 1$, иначе $newNeg_{ji} = 0$;

1.3. Если $Pos_{ji} \neq null$, то $newPos_{ji} = 1$, иначе $newPos_{ji} = 0$.

2. Построение дерева принятия решений, в которой переменная $newSent_i$ является зависимой переменной от $\langle newNeg_{i_1}, \dots, newNeg_{i_n}, newPos_{i_1}, \dots, newPos_{i_n} \rangle$.

3. Расчет значимостей аспектных групп и интерпретация результатов.

Значимость аспектной группы показывает, насколько сильно тональность отзыва зависит от тональности данной аспектной группы. Рассчитывается после построения дерева классификационных правил. Пусть число аспектных групп равно $g/2$, тогда число независимых переменных равно g (негативные и позитивные высказывания по каждой аспектной группе). Формула для расчета значимости m -й переменной будет иметь вид:

$$\text{Значимость}_m = \frac{\sum_{j=1}^{k_m} \left(E_{m,j} - \sum_{i=1}^{n_{m,j}} E_{m,j,i} \cdot \frac{N_{m,j,i}}{N_{m,j}} \right)}{\sum_{l=1}^g \sum_{j=1}^{k_l} \left(E_{l,j} - \sum_{i=1}^{n_{l,j}} E_{l,j,i} \cdot \frac{N_{l,j,i}}{N_{l,j}} \right)} \cdot 100\% , \quad (1)$$

где k_l – количество узлов, которые были разбиты по атрибуту l , $E_{l,j}$ – энтропия родительского узла, разбитого по атрибуту l , $E_{l,j,i}$ – дочерний узел для j -ого, который был разбит по атрибуту l , $N_{l,j}$, $N_{l,j,i}$ – количество примеров в соответствующих узлах, $n_{l,j}$ – количество дочерних узлов для j -ого родительского.

Оценка УП потребителей продукцией рассчитывается по формуле:

$$\text{УП} = \frac{N^{pos}}{N^{pos} + N^{neg}} \cdot 100\% , \quad (2)$$

где N^{pos} – число позитивных отзывов, N^{neg} – число негативных отзывов.

Оценка УП j -й аспектной группой продукции рассчитывается по формуле:

$$\text{УП}_j = \frac{n_j^{pos}}{n_j^{pos} + n_j^{neg}} \cdot 100\% , \quad (3)$$

где n_j^{pos} – число отзывов содержащих позитивное упоминание по j -й аспектной группой, n_j^{neg} – число отзывов содержащих негативное упоминание по j -й аспектной группой.

Эксперимент

Оценка эффективности разработанного прототипа ИСППР проводилась на наборе данных из 635 824 отзывов на русском языке, посвященных курортно-гостиничному бизнесу. Отзывы были собраны с популярного интернет-ресурса tophotels.ru за период 2003-2013 гг. [8]. Данные были предварительно обработаны (удалены дубликаты, фрагменты html-разметки и отзывы длиной менее 30 символов) и загружены в БД SQL Server 2012.

Для классификации тональности отзывов использовалась бинарная шкала (негатив и позитив). Обучающая выборка позитивных и негативных отзывов формировалась с

помощью собранной информации об авторских оценках размещения, питания и сервиса. Интернет-ресурс tophotels.ru использует 5-ти бальную оценочную шкалу по которым оцениваются питание, размещение и сервис. В обучающую выборку попали 15790 негативных отзывов, имеющих по 3 и 4 суммарных балла и 15790 позитивных отзывов, имеющих 15 суммарных баллов. При дальнейшей обработке данных авторские оценки не использовались. Разметка оставшихся 604 244 отзывов осуществлялась с помощью обученного классификатора.

С целью построения эффективного классификатора тональности была проведена оценка точности классификации алгоритмов машинного обучения и некоторых особенностей их построения (Таблица 1). Для оценки точности классификации используется соотношение числа правильно классифицированных позитивных и негативных отзывов к их общему количеству. Оценка точности производилась на двух наборах данных. Первый набор (Тест №1) представлял сформированную обучающую. Его тестирование производилось с помощью перекрестной проверки с разбивкой данных на 10 частей. Вторым набор (Тест №2) содержал отзывы с разными количествами баллов и был размечен вручную (497 позитивных и 126 негативных). Вторым набором использовался для контроля точности классификаторов, обученных на первом наборе данных.

Таблица 1 – Сравнение точности классификации тональности

№	Методы машинного обучения	Вектор	Тест №1	Тест №2
1	SVM (линейное ядро)	Частотный	94,2%	83,1%
2	SVM (линейное ядро)	Бинарный	95,7%	84,1%
3	NB	Бинарный	96,1%	83,7%
4	NB	Частотный	97,6%	92,6%
5	NB (слова-исключения)	Частотный	97,7%	92,7%
6	NB (тегирование частиц «не» и «ни»)	Частотный	98,1%	93,6%

Для разметки отзывов и анализа тональности был выбран классификатор №6 на основе метода NB, частотными векторами в качестве признакового пространства и с использованием техник лемматизации и теггирования отрицаний «не» и «ни». С использованием разработанного алгоритма из всего набора отзывов были извлечены аспектные существительные, которые были разбиты на семь основных аспектных групп (Рисунок 3). Далее были извлечены и размечены по тональности предложения с упоминанием слов из аспектных групп.

Было проведено исследование УП для двух 5-ти звездочных отелей – отеля «А» (1692 отзыва) и отеля «Б» (1300 отзывов), расположенных на курорте Шарм-эль-Шейх (63 472 отзыва) в Египте. Результаты количественного анализа представлены ниже. Рисунок 4 показывает динамику УП. Рисунок 5 показывает удовлетворенность по аспектным группам. В категорию «Курорт» входят отзывы обо всех отелях данного курорта.

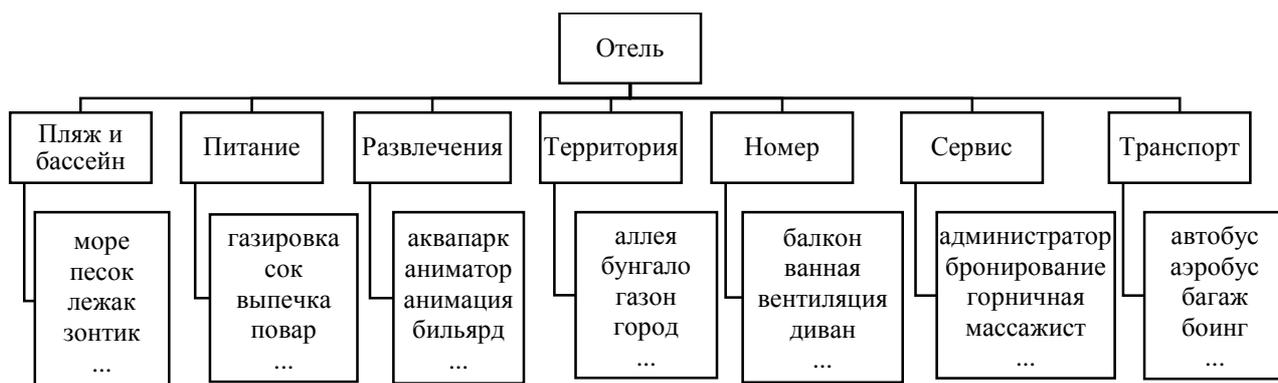


Рисунок 3 – Аспектные группы объекта тональности «отель»

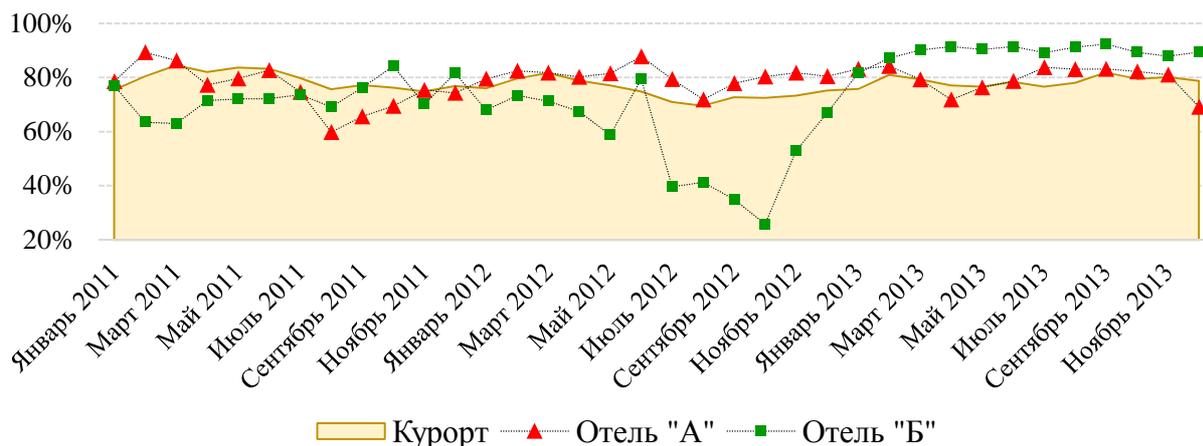


Рисунок 4 – Динамика индикаторов удовлетворенности потребителей по месяцам

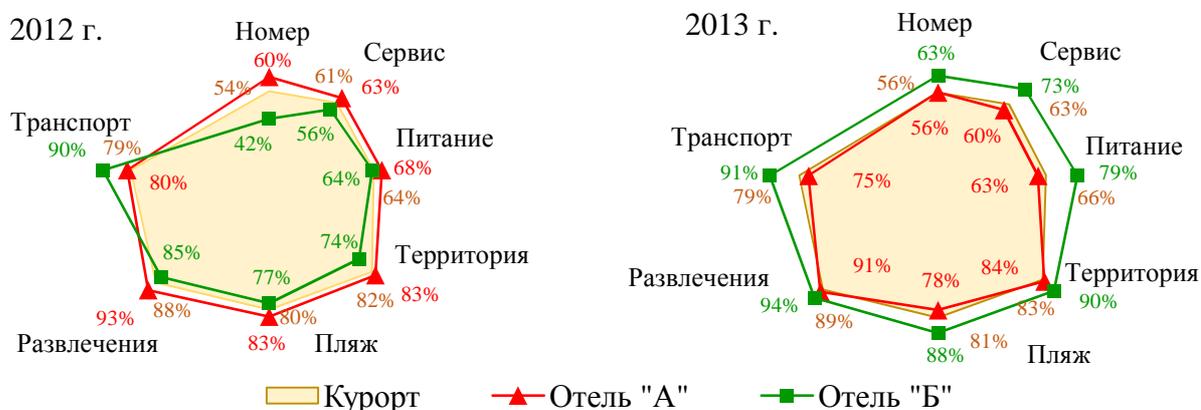


Рисунок 5 – Удовлетворенность потребителей по аспектным группам в 2012 и 2013 годах

Для проведения качественного исследования по разработанному методу были построены деревья решений, по отзывам всех отелей курорта и отдельно по отелям «А» и «Б». Извлеченные правила деревьев решений представлены в Таблице 2. Значимости аспектных групп представлены в Таблице 3.

Проведенные количественные и качественные исследования УП для отеля «А» позволили определить проблемные аспектные группы и выявить те из них, которые оказывают наибольшее влияние на УП, с учетом этого определен приоритет принимаемых решений и пути их разрешения в долгосрочной и краткосрочной перспективе.

Таблица 2 – Извлеченные правила с помощью деревьев решений

№	Правила	Тональность	Поддержка ¹	Достоверность ²
Извлеченные правила по всем отзывам курорта				
1	Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u>	Позитив	37,2%	97,4%
2	Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u> ∩ <u>Пляж⁺</u>	Позитив	11%	86,2%
3	<u>Питание⁺</u> ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u> ∩ <u>Номер⁻</u>	Позитив	10,6%	83,9%
4	<u>Питание⁺</u> ∩ <u>Сервис⁻</u> ∩ <u>Развлечение⁺</u>	Негатив	6,9%	92,3%
5	Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u> ∩ <u>Развлечение⁺</u>	Позитив	5,8%	88,4%
Извлеченные правила по отзывам отеля «А»				
6	<u>Сервис⁻</u>	Позитив	62,9%	88,3%
7	Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Развлечение⁺</u>	Позитив	20,5%	74,1%
8	<u>Питание⁺</u> ∩ <u>Сервис⁻</u>	Негатив	9,4%	86,2%
9	Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Развлечение⁺</u>	Негатив	7,2%	65,6%
Извлеченные правила по отзывам отеля «Б»				
10	<u>Номер⁻</u>	Позитив	51,2%	95,5%
11	Питание ⁺ ∩ <u>Номер⁻</u> ∩ <u>Развлечение⁺</u>	Позитив	27,9%	81%
12	<u>Питание⁺</u> ∩ <u>Номер⁻</u>	Негатив	11,1%	84%
13	Питание ⁺ ∩ <u>Номер⁻</u> ∩ <u>Развлечение⁺</u>	Негатив	9,9%	55,8%

¹ Поддержка показывает долю отзывов от исходной выборки, содержащих данное правило.
² Достоверность показывает какая доля отзывов, содержащих правило, имеют данную тональность.

Таблица 3 – Значимости аспектных групп

Аспект	Курорт	Отель «А»	Отель «Б»	Аспект	Курорт	Отель «А»	Отель «Б»
<u>Сервис⁻</u>	34,8%	60,2%	-	Сервис ⁺	0,7%	-	-
Питание ⁺	30,3%	27,2%	30,3%	<u>Пляж⁻</u>	0,2%	-	-
<u>Питание⁻</u>	16%	-	-	Транспорт ⁺	-	-	-
Развлечение ⁺	8,5%	12,7%	12,4%	Воровство	-	-	-
<u>Номер⁻</u>	4%	-	57,3%	Отравления	-	-	-
Пляж ⁺	2,5%	-	-	Развлечения ⁻	-	-	-
Номер ⁺	2,1%	-	-	Территория ⁻	-	-	-
Территория ⁺	1%	-	-	Транспорт ⁻	-	-	-

Таблица 4 – Применение результатов для выработки управленческих решений для отеля «А»

Проблемные аспектные группы	Значимость	Правила с итогом Позитив	Примеры рекомендуемых управленческих решений
1. <u>Сервис⁻</u>	Отель	№6: <u>Сервис⁻</u> №7: Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Развлечение⁺</u>	Обучить и мотивировать обслуживающий персонала, провести проверку качества ресторанного обслуживания, провести увеселительные мероприятия.
	Курорт	№5: Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u> ∩ <u>Развлечение⁺</u>	
2. <u>Питание⁻</u>	Отель	-	Разнообразить ассортимент блюд, организовать сбор мусора на пляже.
	Курорт	№2: Питание ⁺ ∩ <u>Сервис⁻</u> ∩ <u>Питание⁻</u> ∩ <u>Пляж⁺</u>	
3. <u>Пляж⁻</u>	Отель	-	См. выше
	Курорт	-	
4. <u>Транспорт⁻</u>	Отель	-	Не существенны или вне зоны компетенции.

Заключение

Предложенная концепция поддержки принятия решений на базе разработанной методологии обработки и анализа текстовых данных позволяет в автоматизированном режиме проводить количественное и качественное исследование удовлетворенности потребителей и принимать эффективные управленческие решения по управлению качеством продукции. Данная концепция позволяет значительно снизить трудоемкость исследования удовлетворенности потребителей, что делает ее доступной для применения широким кругом предприятий.

На базе предложенной концепции был разработан прототип ИСППР. Проведенный эксперимент показал эффективность подхода при решении реальных задач управления качеством продукции, удовлетворительную точность алгоритмов обработки анализа текста, логичность полученных результатов. ИСППР позволяет принимать решения по управлению качеством на основе аналитической обработки текстовых отзывов из Интернета, содержащих в неявном виде информацию об удовлетворенности потребителей.

Список литературы

1. ГОСТ Р 54732-2011/ISO/TS 10004:2010 Менеджмент качества. Удовлетворенность потребителей. Руководящие указания по мониторингу и измерению. – М.: Стандартинформ, 2012. — 28 с.
2. Интернет-ресурс посвященный курортно-гостиничному бизнесу [Электронный ресурс] — Режим доступа: <http://www.tophotels.ru>.
3. Пазельская А., Соловьев А. Метод определения эмоций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии. Сб. научных статей / Вып. 10 (17). — М.: Изд-во РГГУ, 2011. — С. 510-522.
4. Юсупова Н. И., Богданова Д. Р., Бойко М. В. Подход к применению анализа настроений в текстах на русском языке на основе машинного обучения // ИМММ 2012: Вторая Международная конференция «Передовые технологии извлечения и управления информацией», Венеция, Италия. — 2012. — С. 8-14.
5. Юсупова Н. И., Богданова Д. Р., Бойко М. В. Алгоритмическое и программное обеспечение для анализа тональности текстовых сообщений с использованием машинного обучения // Вестник УГАТУ. — 2012. — Т. 16, №6(51). — С. 91-99.
6. Nasukawa T., Yi J. Sentiment analysis: capturing favorability using natural language processing // In Proceedings of the 2nd international conference on Knowledge capture, Florida, USA, October 23–25, 2003. — С. 70–77.

7. Pang B., Lee L. Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval. — Vol. 2, Nos. 1-2. — 2008. —135 с.
8. Pang B., Lee L. Thumbs up? Sentiment Classification using Machine Learning Techniques // Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). — Philadelphia. — 2002. — С. 79-86.

Рецензенты:

Черняховская Л.Р., д.т.н., профессор, профессор, ФГБОУ ВПО «Уфимский государственный авиационный технический университет», г. Уфа.

Картак В.М., д.ф.-м.н., профессор, зав. кафедрой прикладной информатики, ФГБОУ ВПО «Башкирский государственный педагогический университет имени М. Акмуллы», г. Уфа.