

**Бахчевников Валентин Владимирович** – e-mail: bahchevnikov@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский; 44; тел.: +79518289271; кафедра радиотехнических и телекоммуникационных систем; ассистент.

**Бакуменко Алексей Николаевич** – e-mail: baku@sfedu.ru; 347900, г. Таганрог, ул. Петровская, 81; тел.: +79886031853; Инжиниринговый центр приборостроения радио и микроэлектроники; инженер.

**Derkachev Vladimir Aleksandrovich** – Southern Federal University; e-mail: vderkachev@sfedu.ru; 2, Shevchenko street, Taganrog, 347922, Russia; phone +79614154733; the department of radio engineering & telecommunication systems; assistant.

**Bakhchevnikov Valentin Vladimirovich** – e-mail: bahchevnikov@sfedu.ru; 44, Nekrasovsky lane, Taganrog, 347928, Russia; phone: +79518289271; the department of radio engineering & telecommunication systems; assistant.

**Bakumenko Alexey Nikolaevich** – e-mail: baku@sfedu.ru; 81, Petrovskaya street, Taganrog, 347900, Russia; phone: +79886031853; Engineering center of instrument making, radio- and microelectronics; engineer.

УДК 0004.912

DOI 10.18522/2311-3103-2020-4-39-48

**Е.М. Герасименко, В.В. Степенко**

### **ИНТЕЛЛЕКТУАЛЬНЫЙ МЕТОД ИЗВЛЕЧЕНИЯ ЗНАНИЙ НА ОСНОВЕ ОПРЕДЕЛЕНИЯ ТОНАЛЬНОСТИ ОТЗЫВОВ\***

*В этой работе исследуется влияние возраста и пола при анализе тональности отзывов, поскольку эти данные могут помочь ретейлерам электронной коммерции увеличить продажи, ориентируясь на определенные демографические группы, а также увеличить удовлетворение потребностей людей разных возрастных и гендерных групп. Используемый набор данных сформирован путем сбора отзывов о книгах. Был создан вопросник, содержащий информацию о предпочтениях книжных носителей (мнения пользователей об электронных книгах, книгах в мягкой и твердой обложках, изображениях и аудиокнигах), а также данные о возрастной группе и гендерной принадлежности. Помимо этого, вопросник также содержит информацию о положительном либо отрицательном мнении касательно предпочтений, которая послужила основой достоверности для классификаторов. В результате, было получено 900 анкет, которые были разделены на группы по половому признаку и возрасту. Каждая конкретная группа данных была разделена на обучающую и тестовую. Были проанализированы сегментированные данные на предмет настроений в зависимости от каждой возрастной группы и пола. Возрастная группа «старше 50 лет» продемонстрировала лучшие результаты по сравнению со всеми другими возрастными группами во всех классификаторах; данные в женской группе показали более высокую точность по сравнению с данными из групп без информации о гендерной принадлежности. Высокие результаты, показанные этими группами, показывают, что подходы к анализу тональности способны предсказать настроения в этих группах лучше, чем в других. Анализ тональности проводился с использованием различных подходов машинного обучения (ML), включая максимальную энтропию, метод опорных векторов, сверточную нейронную сеть и долговую краткосрочную память.*

*Анализ тональности; машинное обучение; сверточная нейронная сеть; долгая краткосрочная память.*

---

\* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00099 а.

E.M. Gerasimenko, V.V. Stetsenko

## INTELLIGENT METHOD OF KNOWLEDGE EXTRACTION BASED ON SENTIMENT ANALYSIS

*The paper explores the impact of age and gender in sentiment analysis, as this data can help e-commerce retailers increase sales by targeting specific demographic groups. The data set used was created by collecting book reviews. A questionnaire was created containing questions about preferences in books, as well as age groups and gender information. The article analyzes segmented data on the subject of moods depending on each age group and gender. Sentiment analysis was performed using various machine learning (ML) approaches, including maximum entropy, support vector method, convolutional neural network, and long short-term memory. This paper investigates the impact of age and gender in sentiment analysis, because this data can help e-commerce retailers to increase sales by targeting specific demographic groups, as well as increase the satisfaction of the needs of people of different age and gender groups. The dataset used is generated by collecting book reviews. A questionnaire was created containing questions about preferences in books (user opinions of e-books, paperbacks, hardbacks, images and audiobooks), as well as data on age group and gender. In addition, the questionnaire also contains information on a positive or negative opinion regarding preferences, which served as the basis for reliability for the classifiers. As a result, 900 questionnaires were received, which were divided into groups according to gender and age. Each specific group of data was divided into training and test one. Segmented data were analyzed for sentiment analysis depending on age group and gender. The age group "over 50 years old" showed the best results in comparison with all other age groups in all classifiers; data in the female group performed higher accuracy compared to data from the groups without gender information. The high scores shown by these groups indicate that sentiment analysis approaches are able to predict moods in these groups better than in others. Sentiment analysis was performed using a variety of machine learning (ML) approaches, including maximum entropy, support vector machines, convolutional neural networks, and long short term memory.*

*Sentiment analysis; machine learning; convolutional neural network; long short-term memory.*

**Введение.** Рост популярности Интернета привел к огромному притоку данных, содержащих обширную и ценную информацию об общественном мнении. Каждый пользователь интернета, выражающий свое мнение, становится частью информационной цепи, где другие пользователи, изучая публичные обзоры, могут принимать обоснованные решения. Для определения общественного мнения (положительного, отрицательного или нейтрального) используются данные (обзоры, посты, комментарии), собранные из различных социальных сетей, таких как Facebook, Twitter, Amazon, Goodreads, IMDb или блогов. Процесс определения полярности общественного мнения называется анализом тональности. Анализ тональности обычно проводится на основе обзоров фильмов [1, 2], обзоров ресторанов или продуктов питания [3, 4], а также данных из микроблогов [5, 6], что помогает различным организациям улучшить бизнес-стратегии за счет привлечения новых клиентов. Разделение клиентов по возрасту и полу позволяет получить важную информацию, использование которой поможет эффективнее удовлетворить потребности людей разных возрастных и гендерных групп.

Анализ тональности как направление интеллектуального анализа данных [7–10] развивался в течение многих лет, при этом использовались различные методики на основе словарей и машинного обучения. С появлением методов глубокого обучения [11–13] предшествующая информация также сыграла большую роль в адекватном выражении полярности мнений.

Ли и соавторы [14] предложили структуру, обеспечивающую определение полярности мнений при помощи анализа тональности. В предложенном подходе авторы приняли во внимание субъективность мнения и доверие пользователей. Локенхоф и соавторы [15] проанализировали, как разные возрастные группы вы-

ражают свои эмоции. Авторы обнаружили, что пожилые люди описывают положительные эмоции лучше, чем отрицательные, по сравнению с более молодыми людьми. Циммерман и соавторы [16] исследовали тональность в зависимости от возраста и пола [17]. Результаты показали, что люди в подростковом возрасте демонстрировали наименьшую эмоциональную регуляцию по сравнению с другими возрастными группами. Гендерные различия проявлялись как при занижении, так и при переоценке конкретной эмоции. Ох и соавторы [18] изучили сегментацию рынка по полу и возрасту пользователей в целях определения потенциальных направлений путешествий для разных групп в зависимости от их доходов и свободного времени. Кешари и соавторы [19] проанализировали эффективность рекламных объявлений для разных гендерных и возрастных групп на основе того, как потребители реагируют на эти рекламные объявления. Таким образом, основываясь на этих установленных психологических различиях между людьми разного возраста и пола, мы постараемся выяснить, существуют ли эти различия в мнениях, которые люди выражают на онлайн-платформах.

На рис. 1 показана основная блок-схема структуры, используемой для анализа тональности отзывов. Отзывы, собранные из социальных сетей, используются для составления двух наборов данных на основе возраста и пола. Затем в полученных наборах данных выполняется анализ тональности с применением различных подходов анализа тональности.

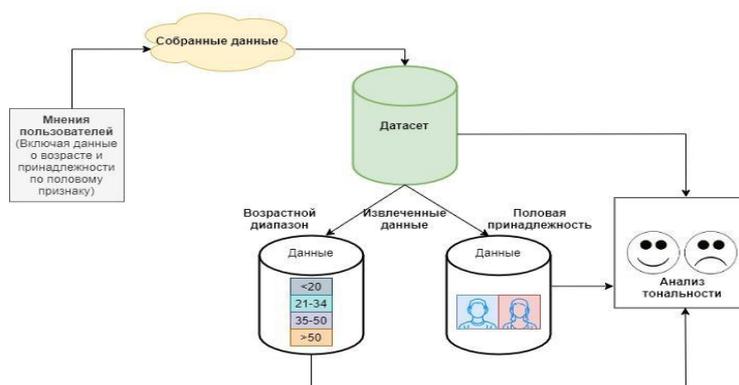


Рис. 1. Блок-схема структуры анализа

**Обзор источников.** До появления интернета было проведено множество исследований, посвященных тому, как разные люди справляются с разными эмоциями и выражают их. Авторы [20] исследовали гендерные различия при проведении исследования 400 студентов из колледжей, разделенных на пять возрастных групп: от дошкольников до взрослых. Исследование было связано со стереотипами половой и возрастной эмоциональной выразительности. Стоунер и соавторы [21] рассматривали людей обоих полов и разных возрастных групп с целью изучения их способности выражать гнев. Исследование показало, что в группе молодых людей гнев был выражен больше, чем в старших возрастных группах. В данном исследовании автор не обнаружил больших различий по половому признаку в этом аспекте.

Исследование Дэвиса [22] по гендерным различиям в выражении отрицательных эмоций показало, что мальчики проявляли больше отрицательных эмоций, когда были разочарованы, по сравнению с девочками. Броди и соавторы [23] более детально исследовали гендерную и эмоциональную экспрессию и показали, что гендерные различия в эмоциональной выразительности были культурно специфичны среди азиатских иностранных студентов. В исследовании Кринга и соав-

торов [24] группе студентов были продемонстрированы эмоциональные видео, и ученые получили подтверждение того факта, что женщины, как правило, более выразительны, чем мужчины, даже в случае пережитых эмоций. В исследовании Бердитта [25] изучались возрастные и гендерные различия при описании эмоциональных реакций. В эксперименте приняли участие 185 человек (из них – 85 мужчин и 100 женщин) в возрасте от 13 до 99 лет. Результат исследования показал, что подростки и молодые взрослые чаще сообщали о гневе и давали более интенсивные отрицательные ответы, чем группа взрослых мужчин

**Методология.** Для обработки обзоров выполняются следующие шаги, отображенные на рис. 2.



Рис. 2. Блок-схема, представляющая шаги, предпринятые для анализа тональности

Изначальный набор данных разделен на два подмножества, называемых датасетами (от англ. data set – набор данных), а затем каждый из них дополнительно разделен на группы, в зависимости от конкретного возраста и пола. Каждая конкретная группа данных разделяется на обучающие данные и данные для тестирования. Алгоритм классификации реализуется в конце фазы обучения после предварительной обработки данных и извлечения характеристик и используется на фазе тестирования для получения окончательных результатов. Обзоры пользователей проходят предварительную обработку и извлечение характеристик на этапе тестирования, а затем проходят этап классификации.

Обзоры предварительно обрабатываются с целью удаления ненужной информации, которая не влияет на полярность предложения [26], с помощью этапов, представленных ниже:

1. Удаление знаков препинания, цифр и символов, поскольку они не несут смысловой нагрузки.
2. Удаление стоп-слов, поскольку они не влияют на оценку настроения высказанного мнения.
3. Замена аббревиатур слова реальным словом.
4. Преобразование текста в нижний регистр.
5. Замена смайликов эмоциями, которые выражает смайлик.
6. Токенизация.

Далее происходит извлечение характеристик. И, наконец, алгоритм классификатора предсказывает метку, которая при сравнении с основополагающей достоверностью определяет точность классификатора.

Были собраны данные о предпочтениях людей в отношении книг (твердый переплет, электронные книги или аудиокниги), а также информация о возрасте и поле. В подходе реализуются различные алгоритмы анализа тональности для каждого набора данных в отдельности, а затем результаты сравниваются, и выявляются соответствующие различия среди групп. Также, в собранном наборе данных был реализован подход анализа тональности на основе словаря.

Для извлечения характеристик применяется «мешок слов» [27] (используется в методах NB, ME и SVM), а word2vec создает вектор признаков, используя либо непрерывный пакет слов, либо модель Skip-gram, которая в дальнейшем используется в LSTM и CNN. Методы описаны ниже.

#### 1. «Мешок слов» (*Bag of Words*)

Bag of Words – очень гибкая и простая модель, используемая для извлечения характеристик. Эта модель отслеживает количество вхождений, также называемых частотой каждого слова, используемого в предложении. Кроме того, каждому слову в предложении присваивается конкретная оценка субъективности. Общая оценка тональности предложения складывается из суммы оценок каждого слова.

#### 2. *Word2Vec*

Модель Word2Vec используется для формирования встраивания слов. Эта модель очень полезна, поскольку обнаруживает сходство слов в векторной форме, а не в текстовом формате. Эти сходства обнаруживаются на основе значения слова, угаданного по его прошлым появлениям и ассоциации с другими словами.

#### 3. *Словарь на основе классификатора*

VADER (Valence Aware Dictionary and sEntiment Reasoner) – это лексикон и инструмент анализа настроений на основе правил, который специально настроен на чувства, выраженные в социальных сетях. Каждому слову, присутствующему в предложении, присваивается оценка по значению этого слова в словаре. Рассчитывается итоговая сложная оценка предложения, которая варьируется от -1 до 1. Эта оценка показывает, является ли предложение положительным или отрицательным. Составной балл для каждого предложения в наборе данных объединяется, и для всего документа анализируется средний балл. Чтобы сравнить его с другими подходами машинного обучения, мы преобразуем среднюю оценку в точность, разделив оценку всего документа на общее количество проверок в этом конкретном наборе данных.

**Классификаторы на основе машинного обучения.** Мы применили пять алгоритмов, основанных на машинном обучении, для определения точности набора данных.

1. Наивный байесовский метод.
2. Максимальная энтропия.
3. Метод опорных векторов (SVM).
4. Долгая краткосрочная память (LSTM).
5. Сверточная нейронная сеть (CNN).

Остановимся подробнее на LSTM и сверточной нейронной сети.

На первом этапе определяется информация, которая будет удалена из ячейки памяти (уравнение 1). Сигмовидный слой выполняет это решение после просмотра предыдущей информации  $i_{t-1}$  и текущего ввода  $c_t$ . Этот сигмовидный слой выводит число от 0 до 1, которое определяет количество информации, которое необходимо сохранить, основываясь на весе  $W_0$ .  $0_t$  представляет собой выходной сигнал текущей ячейки, а  $b_0$  представляет собой смещение для этой конкретной ячейки.

$$0_t = \sigma(W_0 * [i_{t-1}, c_t] + b_0). \quad (1)$$

Затем определяется новая информация, которая должна быть обновлена в ячейку памяти. Этот процесс происходит в два этапа: сначала задействуется сигмовидный слой для определения значений для обновления и затем коричневый слой для создания вектора новых значений (уравнение 2).  $n_t$  обозначает информацию, которая должна быть обновлена на основе веса  $W_n$  и смещения  $b_n$ , а  $\tilde{V}_t$  – это данные, которые должны быть включены в информацию о текущем состоянии (уравнение 3):

$$n_t = \sigma(W_n * [i_{t-1}, c_t] + b_n), \quad (2)$$

$$\tilde{V}_t = \tanh(W_V * [i_{t-1}, c_t] + b_V), \quad (3)$$

Далее эта информация обновляется в следующую ячейку  $V_t$  путем умножения старого состояния на  $0_t$  согласно (4).

$$V_t = 0_t * V_{t-1} + n_t * \tilde{V}_t, \quad (4)$$

На заключительном шаге мы снова используется сигмовидный слой, чтобы найти  $f_t$ , которая обозначает информацию, которая будет предоставлена в качестве выходных данных на основе веса  $W_f$  и смещения  $b_f$  (уравнение 5). Слой  $\tanh$  обновляет необходимые части и выдает их в качестве выходных данных ячейки (уравнение 6).

$$f_t = \sigma(W_f * [i_{t-1}, c_t] + b_f), \quad (5)$$

$$i_t = f_t * \tanh(V_t), \quad (6)$$

Окончательный вывод из этой ячейки будет служить в качестве предварительной информации для следующей ячейки, чтобы определить ее последующее состояние.

Мы обучили модель CNN [28] с использованием одного слоя на характеристиках, извлеченных из предложений с использованием модели word2vec. Первый слой – это сверточный слой, в котором мы перемещаем несколько фильтров разных размеров по 128 измерениям вложения слов, чтобы создать карту объектов на основе конкретного фильтра. Затем идет слой максимального пуллинга, который сворачивает результаты предыдущего слоя в один длинный вектор признаков. Слой максимального пуллинга находит наиболее заметный вектор объектов из карты объектов, принадлежащий каждому фильтру, который затем полностью передается подключенному слою softmax. Регуляризация отсева выполняется от того, как мы используем слой softmax для классификации результата. Регуляризация случайным образом выбрасывает некоторые скрытые единицы из слоя, чтобы предотвратить совместную адаптацию обучающих данных, которая может привести к перегрузке.

**Ход исследования и результаты.** В качестве набора данных для этого исследования были использованы обзоры 900 человек. Обзоры представлены в виде вопросника, содержащего несколько вопросов, касающихся предпочтений книжного носителя (мнения пользователя об электронных книгах, книгах в мягкой и твердой обложках, изображениях и аудиокнигах). Кроме того, в вопроснике обсуждается, считает ли пользователь, что цифровые носители, такие как Kindle или электронные книги, могли бы заменить книги в твердом переплете или в мягкой обложке. Вопросы уточняли, нравятся ли пользователю аудиокниги больше других форматов, а также предлагалось кратко изложить собственное мнение. Вопросник содержит мнение пользователя, а также пол и возрастную группу опрошиваемого. Наряду с мнением пользователей, они также содержат информацию о положительном / отрицательном мнении касательно предпочтений, которая служит основой достоверности для классификаторов.

60 % обзоров являются положительными, а остальные 40% – отрицательными. Этот набор данных мы разделили на отдельные группы: сначала по полу (70 % принадлежит женщинам, 30 % – мужчинам), затем по возрастному признаку (40 % принадлежат к возрастной группе до 20 лет, 30 % – возрастной группе 21–34 года, 20 % – возрастной группе 35–50 лет, 10 % – возрастной группе – старше 50 лет).

Мы получили результаты с использованием классификаторов на основе машинного обучения и основанных на словаре. Результаты этих классификаторов выражены в терминах точности [29].

$$\text{Точность} = \frac{\text{количество правильно предсказанных наблюдений}}{\text{общее количество наблюдений}}$$

#### Влияние возрастного фактора

Извлеченный набор данных, основанный на возрасте, подразделяется следующие группы: первая группа – младше 20 лет, вторая – от 21 до 34 лет, третья – от 35 до 50 лет и последняя – старше 50 лет. Таким образом, всего создано четыре группы, содержащие положительные и отрицательные отзывы от людей каждой конкретной возрастной группы. Другая группа (без информации о возрасте), содержащая обзоры из всех возрастных групп, сформирована для сравнения ее результатов с другими группами, как показано на рис. 3.

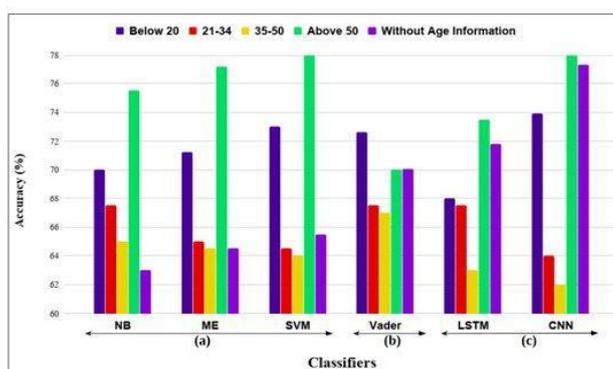


Рис. 3. Сравнение различных методов извлечения характеристик на основе возраста: ((a) – классификаторы, основанные на машинном обучении; (b) – классификаторы на основе словаря; (c) – подходы машинного обучения с использованием метода извлечения признаков word2vec)

Возрастная группа «старше 50 лет» демонстрирует лучшие результаты по сравнению со всеми другими возрастными группами во всех классификаторах с максимальной точностью 78 % в классификаторах CNN и SVM. Возрастная группа «до 20 лет» имеет более высокую точность по сравнению с двумя другими группами среднего возраста, где возрастная группа «21–4» показывает лучший результат, чем другие возрастные группы во всех случаях, хотя разница между этими двумя возрастными группами незначительна. Высокие показатели старших возрастных групп показывают, что подходы анализа тональности способны предсказать настроения в этой возрастной группе лучше, чем в других группах. Группа данных без какой-либо информации о возрасте работает лучше в LSTM и CNN.

**Влияние фактора гендерной принадлежности.** Исходный набор данных разделен на две группы в зависимости от пола (мужскую и женскую) и содержит положительные и отрицательные отзывы. Данные также проходят этапы предварительной обработки, извлечения характеристик и применения классификаторов. Результаты представлены на рис. 4.

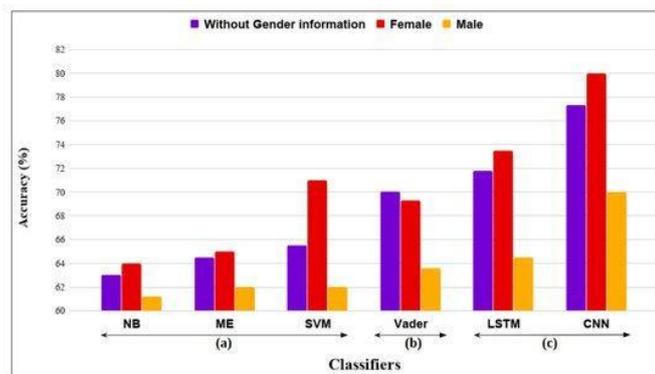


Рис. 4. Сравнение результатов по половому признаку среди различных методов извлечения характеристик и классификаторов (a) – классификаторы, основанные на машинном обучении; (b) – классификаторы на основе словаря; (c) – подходы машинного обучения с использованием метода извлечения признаков word2vec

Хорошо видно, что данные в женской группе имеют более высокую точность по сравнению с данными из групп без гендерной информации и мужской. Данные из женской группы имеют лучшую точность в классификаторе CNN – 80 %, нежели в других классификаторах. Этот результат согласуется с психологическими исследованиями, согласно которым, женщины выражают свое мнение лучше, чем мужчины. Настроения в женских группах легче предсказать, что дает большую точность результата.

**Заключение.** В этой статье мы сравнили различные методы анализа настроений в наборе данных, состоящим из 900 анкет, содержащих информацию о возрасте и поле пользователя. Чтобы проанализировать влияние возраста и пола на то, как пользователи выражают свое мнение, набор данных был разделен на четыре группы. Для определения тональности отзывов использовались методы машинного обучения и словарные технические приемы. В итоге, женская группа показала лучшую точность по сравнению с мужской, а возрастная группа старше 50 лет имеет более высокую точность, по сравнению со всеми другими возрастными группами.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Manek A.S., Shenoy P.D., Mohan M.C., Venugopal K. Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier, *Worldw. Web.*, 2017, No. 20, pp. 135-154.
2. Dos Santos C., Gatti M. Deep convolutional neural networks for sentiment analysis of short texts, *Proceedings of the COLING, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014*, pp. 69-78.
3. Kiritchenko S., Zhu, X., Cherry C., Mohammad S. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews, *Proceedings of the 8th International Workshop on Semantic Evaluation, Dublin, Ireland, 23–24 August 2014*, pp. 437-442.
4. Pontiki M., Galanis D., Papageorgiou H., Manandhar S. Androutopoulos I. Semeval-2015 task 12: Aspect based sentiment analysis, *Proceedings of the 9th International Workshop on Semantic Evaluation, Denver, CO, USA, 4–5 June 2015*, pp. 486-495.
5. Cao D., Ji R., Lin D., Li S. A cross-media public sentiment analysis system for microblog, *Multim. Syst.*, 2016, No. 22, pp. 479-486.
6. Ghosh R., Zhang L., Dekhil M.E., Liu B. Performing sentiment analysis on microblogging data, including identifying a new opinion term therein. US Patent 9,275,041, 1 March 2016.

7. Ullah M.A., Islam M.M., Azman N.B., Zaki Z.M. An overview of multimodal sentiment analysis research: Opportunities and difficulties, *Proceedings of the 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition, Himeji, Japan, 1–3 September 2017*, pp. 1-6.
8. Cambria E. Affective computing and sentiment analysis, *IEEE Intell. Syst.*, 2016, No. 31, pp. 102-107.
9. Liu B. Sentiment analysis and subjectivity, *Handb. Nat. Lang. Proc.*, 2010, No. 2, pp. 627-666.
10. Kumar S., Yadava M., Roy P.P. Fusion of eeg response and sentiment analysis of products review to predict customer satisfaction, *Inf. Fus.*, 2019, No. 52, pp. 41–52.
11. Kim J.H., Kim B.G., Roy P.P. Jeong D.M. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure, *IEEE Access*, 2019, No. 7, pp. 41273-41285.
12. Yoo S.M., Cho C., Lee K.H., Park J., Jin, S., Lee Y., Kim B.G. Structure of deep learning inference engines for embedded systems, *Proceedings of the IEEE 2019 International Conference on Information and Communication Technology Convergence, Kuala Lumpur, Malaysia, 24–26 July 2019*, pp. 920-922.
13. Kim J.H., Hong G.S., Kim B.G., Dogra D.P. Deepgesture: Deep learning-based gesture recognition scheme using motion sensors, *Displays*, 2018, No. 55, pp. 38-45.
14. Li Y.M., Li T.Y. Deriving market intelligence from microblogs, *Decis. Support Syst.*, 2013, No. 55, pp. 206-217.
15. Lockenhoff C.E., Costa P.T., Lane R.D. Age differences in descriptions of emotional experiences in oneself and others, *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.*, 2008, No. 63, pp. 92-99.
16. Zimmermann P., Iwanski A. Emotion regulation from early adolescence to emerging adulthood and middle adulthood, *Int. J. Behav. Dev.*, 2014, No. 38, pp. 182-194.
17. Kaur B., Singh D., Roy P.P. Age and gender classification using brain–computer interface, *Neural Comput. Appl.*, 2019, No. 31, pp. 5887-5900.
18. Oh H., Parks S.C., Demicco F.J. Age-and gender-based market segmentation: A structural understanding, *Int. J. Hosp. Tour. Adm.*, 2002, No. 3, pp. 1-20.
19. Keshari P., Jain S. Effect of age and gender on consumer response to advertising appeals, *Paradigm*, 2016, No. 20, pp. 69-82.
20. Fabes R.A., Martin C.L. Gender and age stereotypes of emotionality, *Personal. Soc. Psychol. Bull.*, 1991, No. 17, pp. 532-540.
21. Stoner S.B., Spencer W.B. Age and gender differences with the anger expression scale, *Educ. Psychol. Meas.*, 1987, pp. 487-492.
22. Davis T.L. Gender differences in masking negative emotions: Ability or motivation?, *Dev. Psychol.*, 1995, No. 31, pp. 660-667.
23. Brody L.R. Gender and emotion: Beyond stereotypes, *J. Soc. Issues*, 2010, No. 53, pp. 369-393.
24. Kring A.M., Gordon A.H. Sex differences in emotion: Expression, experience, and physiology, *J. Personal. Soc. Psychol.*, 1998, Vol. 74, pp. 686-703.
25. Birditt K.S., Fingerman K.L. Age and gender differences in adults' descriptions of emotional reactions to interpersonal problems, *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.*, 2003, No. 58, pp. 237-245.
26. Kharde V.A., Sonawane S.S. Sentiment analysis of Twitter data: A survey of techniques, *Int. J. Comput. Appl.*, 2016, No. 139, pp. 5–15.
27. Saini R., Kumar P., Roy P.P., Pal U. Trajectory classification using feature selection by genetic algorithm, *Proceedings of the 3rd International Conference on Computer Vision and Image Processing*. Springer: Berlin/Heidelberg, Germany, 2020, pp. 377-388.
28. Kim Y. Convolutional neural networks for sentence classification, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014*, pp. 1746-1751.
29. Saito T., Rehmsmeier M. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets, *PLoS ONE*, 2015, No.10. e0118432.

Статью рекомендовал к опубликованию д.т.н., профессор А.В.-Боженюк.

**Герасименко Евгения Михайловна** – Южный федеральный университет; e-mail: egerasimenko@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

**Стеценко Валентина Витальевна** – e-mail: stecenکو@sfedu.ru; кафедра систем автоматизированного проектирования; магистрант.

**Gerasimenko Evgeniya Michailovna** – Southern Federal University; e-mail: egerasimenko@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia, phone: +78634371621; the department of computer-aided design; associate professor.

**Stetsenko Valentina Vitalievna** – e-mail: stecenکو@sfedu.ru; the department of computer-aided design; master student.

УДК 519.712.2

DOI 10.18522/2311-3103-2020-4-48-59

**Л.А. Гладков, Н.В. Гладкова****ЭВОЛЮЦИОНИРУЮЩИЕ МНОГОАГЕНТНЫЕ СИСТЕМЫ  
И ЭВОЛЮЦИОННОЕ ПРОЕКТИРОВАНИЕ\***

*Статья посвящена обсуждению проблем построения эволюционирующих мультиагентных систем на основе использования принципов эволюционного проектирования и гибридных моделей. Рассмотрено понятие агента. Представлен набор базовых свойств агента. Рассмотрены аналогии между многоагентными и эволюционными системами. Рассмотрены принципы построения и организации мультиагентных систем. Отмечены сходства между основными определениями теории агентов и теории эволюции. Отмечено, что основные модели эволюции и эволюционные алгоритмы, могут быть с успехом использованы при проектировании многоагентных систем. Проведен анализ существующих методов и методологий проектирования агентов и многоагентных систем. Отмечены существующие различия в подходах к проектированию многоагентных систем. Описаны основные типы моделей и приведены их важнейшие характеристики. Представлена модель взаимодействия агентов, включающая описание услуг (сервисов), взаимосвязей и обязательств, существующих между агентами. Описана модель отношений (контактов), которая задает коммуникационные связи между агентами. Отмечена важность и перспективность использования агентно-ориентированного подхода к проектированию многоагентных систем. Предложена концепция проектирования агентов и многоагентных систем, согласно которой процесс проектирования включает в себя базовые компоненты самоорганизации, в том числе процессы взаимодействия, скрещивания, адаптации к среде и т.д. Рассмотрены различные подходы к эволюционному проектированию искусственных систем. Предложена Эволюционная модель формирования агентов и агентств, как основной компонент эволюционного проектирования. Предложены модифицированные эволюционные операторы кроссингвера для реализации процесса проектирования агентов.*

*Мультиагентная система; агент; свойства агентов; агентство; эволюционное проектирование; эволюционные операторы; методология проектирования; гибридные методы.*

**L.A. Gladkov, N.V. Gladkova****EVOLUTIONARY DESIGN AS A TOOL FOR DEVELOPING MULTI-AGENT  
SYSTEMS**

*The article is devoted to the discussion of the problems of constructing evolving multi-agent systems based on the use of the principles of evolutionary design and hybrid models. The concept of an agent is considered. A set of basic properties of the agent is presented. The analogies between multi-agent and evolutionary systems are considered. The principles of construction and organization of multi-agent systems are considered. The similarities between the main definitions of the theory of agents and the theory of evolution are noted. It that the main evolution models and evolutionary algorithms can be successfully used in the design of multi-agent systems is noted. The analysis of existing methods and*

\* Исследование выполнено при поддержке РФФИ в рамках научного проекта № 19-01-00715.