

**ПОДХОД К АВТОМАТИЧЕСКОМУ РАСПОЗНАВАНИЮ ЭМОЦИЙ
В ТРАНСКРИПЦИЯХ РЕЧИ**А. А. ДВОЙНИКОВА^{1*}, К. О. КОНДРАТЕНКО²¹*Санкт-Петербургский Федеральный исследовательский центр РАН,
Санкт-Петербург, Россия
*dvoynikova.a@iias.spb.su*²*Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия*

Аннотация. Исследован актуальный в различных областях вопрос распознавания эмоций в транскрипциях речи. Проанализировано влияние методов предобработки (удаление стоп-слов, лемматизация, стемминг) на точность распознавания эмоций в текстовых данных на русском и английском языках. Для проведения экспериментальных исследований использовались орфографические транскрипции диалогов из многомодальных корпусов RAMAS и CMU-MOSEI на русском и английском языке соответственно. Аннотирование этих корпусов выполнялось по следующим эмоциям: радость, удивление, страх, злость, грусть, отвращение и нейтральное состояние. Предобработка текстовых данных включала в себя удаление знаков пунктуации и стоп-слов, токенизацию, лемматизацию и стемминг. Векторизация полученного материала была осуществлена при помощи методов TF-IDF, BoW, Word2Vec. В качестве классификаторов выступили метод опорных векторов и логистическая регрессия. Разработан подход автоматического распознавания эмоций в текстовых данных, представляющий собой комбинацию методов. Для русского языка достигнута наибольшая точность распознавания эмоций по взвешенной F-мере = 92,63 %, для английского языка — 47,21 %. Кроме того, проведены исследования по выявлению количества удаленных стоп-слов для эффективного распознавания эмоций по текстовым данным. Результаты экспериментов показывают, что сохранение стоп-слов в исходном тексте позволяет достичь наиболее высокой точности классификации текстов.

Ключевые слова: распознавание эмоций, методы предобработки текстовых данных, удаление стоп-слов, многоклассовая классификация, анализ текстовых данных

Благодарности: работа выполнена в рамках проекта Российского научного фонда (раздел „Подход к классификации текстовых данных по эмоциям“ выполнен в рамках проекта № 22-11-00321), остальные исследования выполнены частично в рамках ведущей научной школы РФ (грант № НШ-17.2022.1.6) и бюджетной темы СПб ФИЦ РАН (№ FFZF-2022-0005).

Ссылка для цитирования: Двойникова А. А., Кондратенко К. О. Подход к автоматическому распознаванию эмоций в транскрипциях речи // Изв. вузов. Приборостроение. 2023. Т. 66, № 10. С. 818—827. DOI: 10.17586/0021-3454-2023-66-10-818-827.

APPROACH TO AUTOMATIC RECOGNITION OF EMOTIONS IN SPEECH TRANSCRIPTIONSA. A. Dvoynikova^{1*}, K. K. Kondratenko²¹*St. Petersburg Federal Research Center of the RAS, St. Petersburg, Russia
dvoynikova.a@iias.spb.su*²*St. Petersburg State University, St. Petersburg, Russia*

Abstract. The issue of recognizing emotions in speech transcriptions, which is relevant in various fields, is studied. The influence of preprocessing methods (stop word removal, lemmatization, stemming) on the accuracy of emotion recognition in text data in Russian and English is analyzed. To conduct experimental studies, orthographic transcriptions of dialogues from the multimodal corpora RAMAS and CMU-MOSEI in Russian and English, respectively, are used. These corpora are annotated for the following emotions: joy, surprise, fear, anger, sadness, disgust and neutral. Preprocessing of text data includes removal of punctuation marks and stop words, tokenization, lemmatization and stemming. Vectorization of the resulting material is carried out using the TF-IDF, BoW, Word2Vec methods. The used classifiers are support vector machines and logistic regression. An approach is developed that is a combination of the above methods. For the Russian language, the highest accuracy of emotion recognition achieved using a weighted F-measure is

92.63 %, for the English language – 47.21 %. In addition, studies are conducted to identify the number of remote stops for effective emotion recognition from text data. Experimental results show that storing stop words in the source text allows to achieve the highest accuracy of text classification.

Keywords: emotion recognition, text data preprocessing methods, stop-words removal, multiclass classification, text data analysis

Acknowledgments: The work was carried out within the framework of a project of the Russian Science Foundation (section “An approach to classifying text data by emotions” was carried out within the framework of project No. 22-11-00321), the rest of the research was carried out partially within the framework of the leading scientific school of the Russian Federation (grant No. NSh-17.2022.1.6) and the budget theme of the St. Petersburg Federal Research Center of the Russian Academy of Sciences (No. FFZF-2022-0005).

For citation: Dvoynikova A. A., Kondratenko K. K. Approach to automatic recognition of emotions in speech transcriptions. *Journal of Instrument Engineering*. 2023. Vol. 66, N 10. P. 818—827 (in Russian). DOI: 10.17586/0021-3454-2023-66-10-818-827.

Введение. Распознавание эмоций является одной из отраслей sentiment-анализа, изучающих вопросы выделения эмоционально окрашенных сегментов и их анализа [1]. Распознавание эмоций играет ключевую роль в областях, для которых важна реакция человека: бизнес, торговля, образование, сфера развлечений [2]. Кроме того, распознавание эмоций облегчает взаимодействие человека с машиной. В частности, с его помощью можно определить качество товаров и услуг, проанализировать состояние человека, осуществить прогноз политической ситуации на основе новостных лент или оценить взаимоотношения внутри учебного или рабочего коллектива. В связи с этим разрабатываются системы и методы анализа проявления эмоций. Человек может выражать эмоции с помощью мимики, жестов, интонаций, речи, поэтому большинство систем распознавания эмоций анализируют визуальные (изображения и видео) [3], акустические (аудиозаписи речи) [4] и лингвистические данные [5—7]. Некоторые исследования подтверждают, что текстовая модальность более репрезентативна при sentiment-анализе [8]. В связи с этим анализ методов обработки текстовой модальности является актуальным исследованием. В настоящей работе рассмотрены методы распознавания эмоций в тексте.

Анализ эмоциональной окраски представляет собой сложную задачу: в связи с субъективностью восприятия эмоций отсутствует единый принцип их классификации [9]. При решении этой задачи выделяют категориальный и многомерный подходы [10].

В первом случае предполагается, что можно выделить отдельные категории эмоций. К категориальным моделям эмоций относятся модели П. Экмана [11], К. Изарда [12]. На данный момент в области искусственного интеллекта используют распознавание базовых эмоций (модель Экмана) — радость, грусть, гнев, страх, отвращение, удивление, — а также нейтрального состояния [7].

Многомерный подход основан на представлении о том, что эмоции являются точками n -мерного пространства, оси которого определяются полюсами характеристик эмоционального состояния [13]. В области автоматического анализа модель Дж. Рассела [14] является наиболее используемой для описания эмоций. Рассел предложил расположить эмоции в двумерном пространстве — на осях полярности и интенсивности эмоций. Проецирование эмоций на ось полярности часто сводится к задаче распознавания тональности (негативная, нейтральная, позитивная) в текстовых данных [5, 6].

Целью настоящей статьи является определение эффективной комбинации методов анализа текстовых данных на русском и английском языках. Под эффективностью понимается достижение наибольшей точности распознавания эмоциональных состояний человека. Научная новизна статьи заключается в нахождении оптимального числа удаляемых стоп-слов в тексте для обеспечения максимальной точности классификации эмоций. Результаты, полученные в ходе экспериментальных исследований (наиболее эффективная комбинация

методов анализа текста), могут быть применены в задачах оценки качества товаров и услуг, распознавания токсичных сообщений в социальных сетях и др.

Методы анализа текстов. В общем виде распознавание эмоциональной окраски текста состоит из следующих этапов: предобработка исходных данных, извлечение признаков (векторизация), построение классификаторов машинного обучения и оценка качества классификаторов [15]. Первый этап необходим для стандартизации текста: данные приводятся к единой форме (заглавные буквы заменяются строчными и разные формы слова сводятся к одной) и удаляются знаки пунктуации. Этап извлечения признаков предназначен для представления текста в числовом виде. Все методы векторизации можно разделить на две группы: алгебраические (статистические и матричные) и нейросетевые. К традиционным методам векторизации относятся: мешок слов (Bag of Words, BoW), частотный метод (Term Frequency – Inverse Document Frequency, TF-IDF), прямое кодирование (One-hot encoding), сингулярное разложение (Singular Value Decomposition, SVD) [16]. К методам на основе нейронных сетей относят: Glove [17], FastText [18], Word2Vec [19], BERT [20], ELMo [21]. Они применяются в работах [6—8, 13] для анализа текстов. Далее извлеченные текстовые признаки подаются на вход классификатору для многоклассового распознавания эмоций. Классификаторы можно также разделить на две группы: традиционные (метод опорных векторов, логистическая регрессия, наивный Байес, деревья решений и др.) и нейросетевые (RNN, LSTM, Transformers). Все они применяются для распознавания сентимента и эмоций в текстовых данных в работах [5—8, 22].

При анализе текстов большую роль играет предобработка данных. Одной из проблем, связанных с подготовкой текстовых данных для машинного обучения, является большой объем информации, которую необходимо обработать. Таким образом, многие исследователи [5—7, 13, 22, 23] при предобработке удаляют из текста стоп-слова — слова, не содержащие смысловой нагрузки (например, предлоги, союзы, частицы и т.п.), чтобы выделить релевантную информацию и уменьшить размерность вектора [23].

Задача настоящей работы — определить влияние наличия стоп-слов в тексте и метода нормализации (лемматизация, стемминг) текстовых данных на точность распознавания эмоциональной окраски текста на русском и английском языках.

Подход к классификации текстовых данных по эмоциям. На рис. 1 представлены этапы распознавания эмоций в текстовых данных. Разработанный подход к классификации текстовых данных представляет собой комбинацию из следующих методов: сохранение стоп-слов в тексте, стемминг в качестве нормализации, TF-IDF в качестве векторизации, логистическая регрессия (для русскоязычных текстов) и метод опорных векторов (для англоязычных текстов) в качестве классификации. В рамках настоящей работы предлагается подход к классификации текстовых данных по их эмоциональной окраске, представленный на рис. 1.

На рис. 1 жирным шрифтом выделены методы, относящиеся к базовому подходу [15]. Синим цветом выделены блоки, в которых проводились эксперименты с различными методами. Цель исследования — найти оптимальную комбинацию методов для эффективного распознавания эмоций в текстовых данных.

На первом этапе исследований была выполнена предобработка исходного текста. Она включала в себя токенизацию, т.е. разбиение текста на отдельные значимые единицы (токены) — слова; приведение всех слов к нижнему регистру; удаление знаков пунктуации, нормализацию (лемматизация, стемминг или одновременно лемматизация и стемминг). Опционально выполнялось удаление стоп-слов (слов, не содержащих смысловой нагрузки, например, союзов, предлогов, частиц). Из полученного в библиотеке nltk языка Python списка стоп-слов были удалены слова, которые могут влиять на определение эмоций (например, „хорошо“, „не“, „даже“, „больше“, „лучше“, „again“, „all“, „each“, „few“, „no“, „not“). Форма слова не всегда несет в себе полезную информацию, поэтому при анализе текста производилась

нормализация слов, т.е. их приведение к единой форме. Нормализация может осуществляться двумя способами: лемматизацией и стеммингом. При лемматизации слово преобразуется к его начальной форме (лемме), а при стемминге — к его основе (отбрасываются окончания, суффиксы, приставки). Для экспериментов были использованы оба метода, а также их сочетание: лемматизация и последующий стемминг.

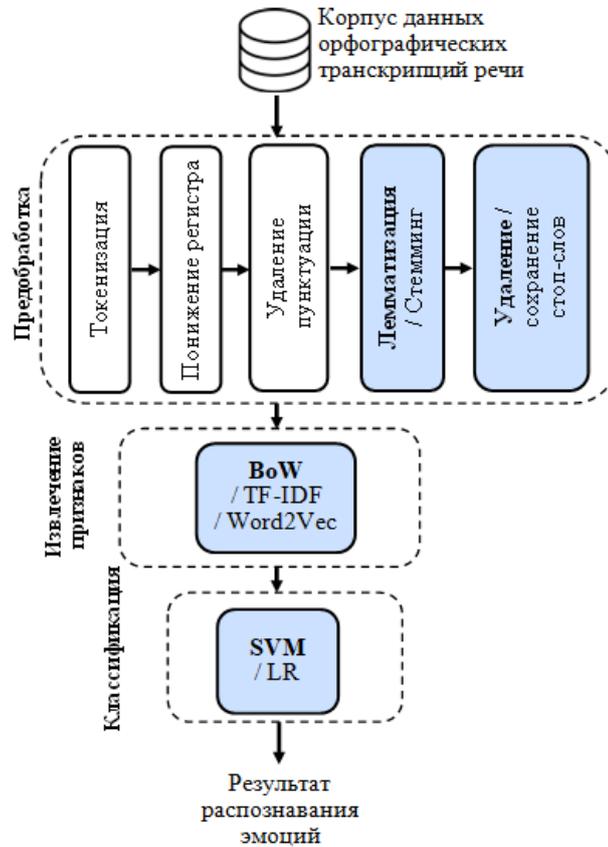


Рис. 1

Далее была произведена векторизация, поскольку перед тем, как использовать машинный классификатор, необходимо представить полученный текст в числовом виде. В настоящей работе применены методы Bag of Words, TF-IDF и Word2Vec [19].

На следующем этапе исследования в качестве классификаторов выступили метод опорных векторов (Support Vector Machine, SVM) и логистическая регрессия (Logistic Regression, LR). Преимущества данных классификаторов заключаются в простоте их использования и высокой скорости обучения. Для каждого классификатора производился подбор гиперпараметров.

Для оценки точности предсказаний использовалась взвешенная F-мера, поскольку она наиболее репрезентативна при неравномерном распределении данных в классах.

Корпусы данных. Исследование проводилось на материале двух корпусов. Для русского языка использовались транскрипции (орфографические) аудиозаписей из многомодального корпуса RAMAS [24]. Он содержит записи спонтанных диалогов и монологов (для нейтральной речи) десяти актеров. При создании корпуса использовались диадические сценарии, в рамках которых говорящие импровизировали на заданную тему, воспроизводя определенную эмоцию. Данные размечены одной из базовых эмоций: радость, гнев, удивление, страх, грусть, отвращение и нейтральное состояние. Общий объем корпуса — 581 аудиозапись. Транскрипции диалогов корпуса RAMAS были получены нами с помощью автоматической системы распознавания речи (Automatic Speech Recognition, ASR) для 535 записей. Подробно

подход к извлечению лингвистической информации из аудиоданных корпуса RAMAS описан в работе [25].

Для английского языка использовались субтитры видеозаписей из многомодального корпуса CMU-MOSEI [26]. Он включает фрагменты речи более чем 1000 человек, полученные с платформы YouTube видеозаписи были подготовлены таким образом, чтобы в видео присутствовала речь только одного диктора. Так, часть записей представляет собой монологи, а часть — реплики из диалогов. Поскольку некоторые материалы данного корпуса имеют разметку нескольких эмоций одновременно, для эксперимента более чем из 23 000 фраз были выбраны только те, которые включали одну из перечисленных выше базовых эмоций. Так, объем исследуемого материала составил 14 802 фразы. Распределение фраз M корпусов RAMAS и CMU-MOSEI представлено на рис. 2.

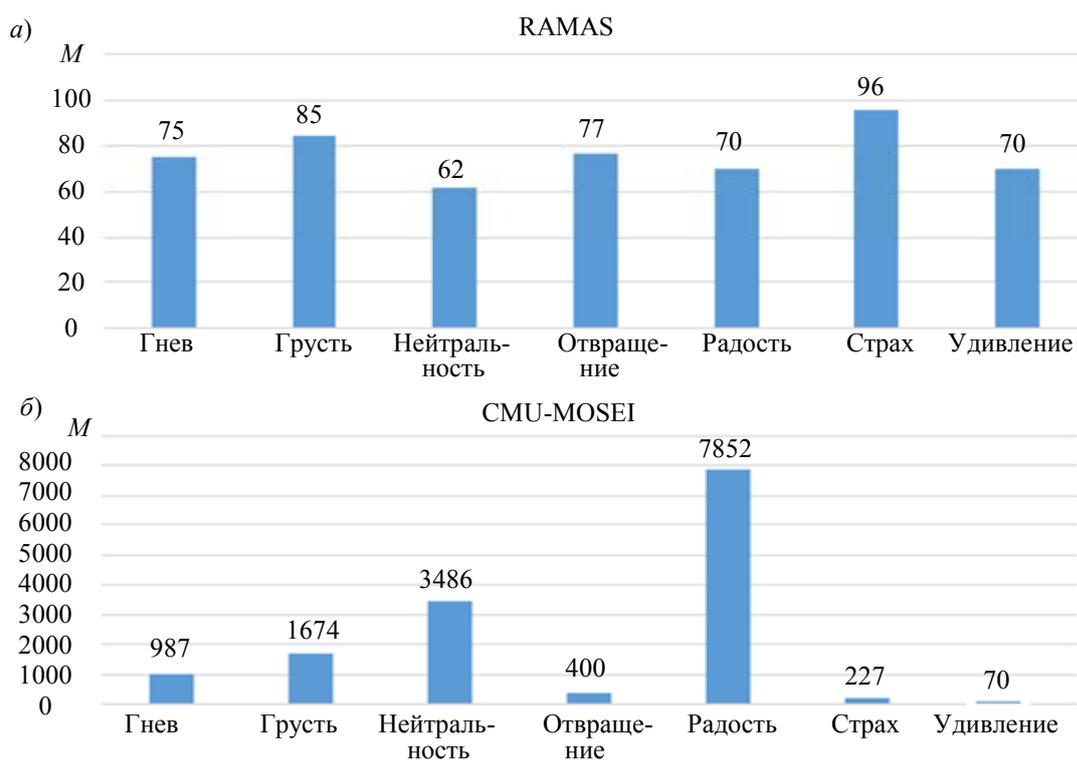


Рис. 2

Как видно из рис. 2, тексты корпуса RAMAS более сбалансированы, а в корпусе CMU-MOSEI существенно не сбалансированы по классам эмоций.

Результаты экспериментальных исследований. Всего было обработано 535 транскрипций для русского языка (RAMAS) и 14 802 текста субтитров для английского языка (CMU-MOSEI). Для каждого языка проведены эксперименты с целью нахождения наиболее эффективной (позволяющей достичь наибольшей точности распознавания) комбинации методов предобработки, нормализации, векторизации и классификации на качество распознавания эмоций в текстовых данных.

Результаты экспериментов, проведенных на материале корпусов RAMAS [24] и CMU-MOSEI [26], представлены в таблице (лем — лемматизация, стем — стемминг, лем+стем — лемматизация и последующий стемминг. Жирным шрифтом выделена наибольшая достигнутая точность для различных классификаторов. Курсивом выделены значения, полученные с помощью базового подхода. Нижним подчеркиванием отмечен лучший результат по методу предобработки.

Результаты классификации эмоций, взвешенная F-мера, %

Стоп-слова	Метод нормализации	Tf-idf		BoW		Word2Vec	
		SVM	LR	SVM	LR	SVM	LR
RAMAS							
Удалены	Лем	89,36	88,71	80,34	87,37	75,30	77,64
Удалены	Стем	88,69	89,55	81,94	88,43	75,69	78,60
Удалены	Лем+Стем	88,74	88,22	80,56	86,78	76,04	77,36
Сохранены	Лем	92,24	92,63	82,43	89,84	82,34	83,20
Сохранены	Стем	92,26	91,89	<u>82,77</u>	89,68	81,36	81,65
Сохранены	Лем+Стем	91,50	92,06	82,64	<u>89,93</u>	<u>82,57</u>	<u>83,67</u>
CMU-MOSEI							
Удалены	Лем	45,33	44,42	44,36	44,21	43,94	41,88
Удалены	Стем	45,97	45,48	44,50	44,56	<u>43,95</u>	40,68
Удалены	Лем+Стем	44,55	44,55	44,80	43,88	43,23	40,61
Сохранены	Лем	46,01	45,75	45,22	<u>45,15</u>	<u>43,95</u>	42,08
Сохранены	Стем	47,21	46,39	45,28	45,06	43,63	<u>42,22</u>
Сохранены	Лем+Стем	46,79	45,79	<u>45,35</u>	44,80	43,27	41,92

На рис. 3 представлены все используемые в экспериментах методы предобработки, векторизации и классификации. На оси абсцисс отражен показатель E процентного соотношения эффективности методов. Данный показатель высчитывался как отношение количества экспериментов, в которых метод достиг наибольшей точности, к общему количеству всех экспериментов.

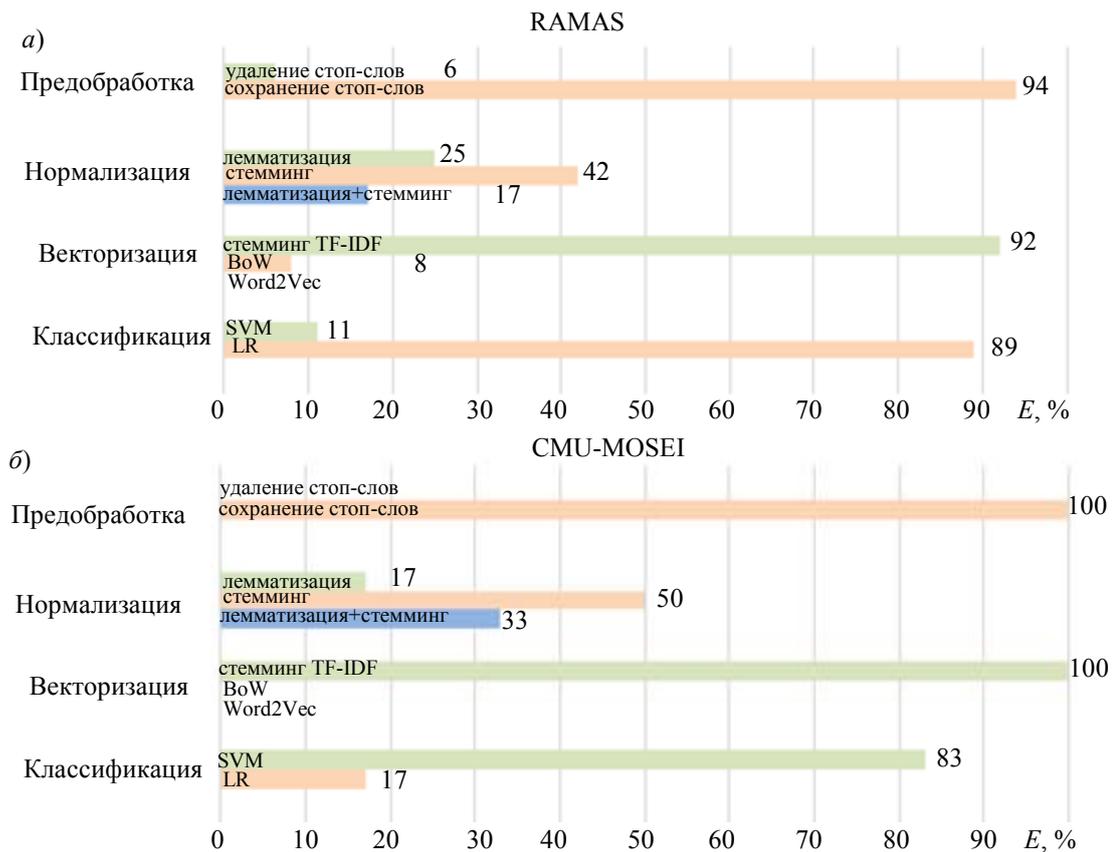


Рис. 3

На основании данных, приведенных выше, можно заметить, что для русского языка эксперименты с сохранением стоп-слов во всех случаях показывали большую точность, чем с удалением. Для корпуса на английском языке только в одном случае при удалении стоп-слов точность оказалась выше — в этом эксперименте для нормализации был использован стемминг,

и классификация была осуществлена методом опорных векторов. Во всех остальных случаях сохранение стоп-слов позволило достичь большей точности.

Поскольку на точность классификации могло повлиять удаление какого-то определенно-го слова или комбинации слов, было принято решение провести эксперимент с удалением разного числа стоп-слов для данных корпуса RAMAS. Для каждого числа стоп-слов (от 1 до 119) выбиралось по пять случайных комбинаций и оценивалась точность модели. Полученные в результате данные представлены на рис. 4, где на оси абсцисс указано число N удаляемых стоп-слов, на оси ординат — средняя точность классификации по показателю „взвешенная F-мера“.

Поскольку с удалением большего числа стоп-слов точность постепенно снижалась, можно сделать вывод, что в общем случае их удаление негативно влияет на определение эмоциональной окраски текста.

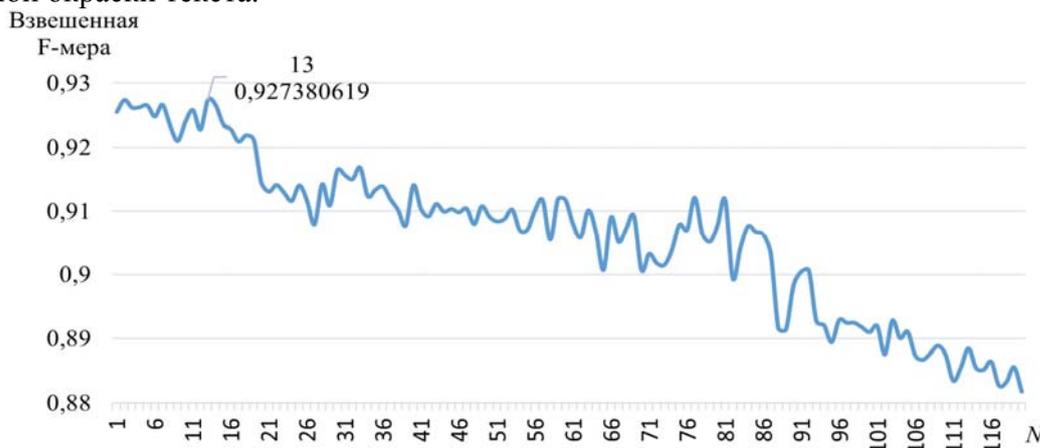


Рис. 4

Обсуждение результатов экспериментальных исследований. При изучении данных для методов нормализации текстовых данных удалось выяснить, что и для русского языка (в 50 % случаев), и английского (в 42 % случаев) точность классификации была выше при использовании стемминга. Что касается остальных методов, в экспериментах на материале русского языка большая точность распознавания эмоций была достигнута для лемматизации с последующим стеммингом, а в экспериментах на материале английского языка — для лемматизации.

Выбор в качестве метода векторизации TF-IDF позволил достичь более высоких результатов классификации эмоций во всех экспериментах для русского языка и в большей части экспериментов — для английского. Лишь в 8 % случаев точность была выше при векторизации статистическим методом Bag of Words. Кроме того, можно заметить, что TF-IDF обычно лучше работает в сочетании со стеммингом, а Bag of Words и Word2Vec — со стеммингом или лемматизацией с последующим стеммингом.

Классификация методом опорных векторов позволила достичь большей точности распознавания эмоциональной окраски текстов на английском языке, а классификация методом логистической регрессии — на русском. Для обоих классификаторов лучше справиться с задачей распознавания эмоций помогает векторизация методом TF-IDF.

Для русского языка наибольшая точность предсказания по взвешенной F-мере = 92,63 % была достигнута при сохранении стоп-слов в тексте, лемматизации при векторизации методом TF-IDF и классификации методом логистической регрессии. Данный результат превосходит результаты базового подхода (удаление стоп-слов, лемматизация, BoW и логистическая регрессия) на 12,29 %. Несмотря на то что для русского языка при стемминге результат классификации был выше в большей части случаев, именно применение лемматизации в сочетании с другими методами предобработки позволило достичь наибольшей точности определения эмоциональной окраски.

Для английского языка наибольшая точность предсказания по F-мере = 47,21 % была достигнута при сохранении стоп-слов в тексте и стемминге и векторизации методом TF-IDF и классификации методом опорных векторов. При базовом подходе удалось достичь F-меры = 44,36 %, что на 2,85 % ниже наилучшего результата.

Экспериментальные исследования показали, что количество удаленных из текста стоп-слов влияет на точность распознавания эмоций в текстовых данных. Из рис. 4 можно заметить, что чем больше стоп-слов удаляется из исходного текста, тем ниже точность классификации эмоции. Оптимальное число удаленных стоп-слов в тексте для корпуса RAMAS — 13, в таком случае достигается F-мера = 92,74 %. Данное значение на 0,11 % выше, чем при сохранении всех стоп-слов в тексте. Составление оптимального списка стоп-слов для классификации эмоций является нерешенной и актуальной задачей. Стоит также предположить, что для каждой задачи и каждого типа текстовых данных оптимальный список стоп-слов будет различным.

Заключение. На основании проведенных исследований можно сделать следующие выводы о результатах распознавания эмоций по текстовым данным с помощью различных методов предобработки:

- 1) удаление стоп-слов снижает точность классификации эмоций в текстовых данных;
- 2) для русского и английского языков точность распознавания выше при использовании стемминга в качестве метода нормализации текстовых данных;
- 3) для русского и английского языков точность распознавания выше при использовании метода TF-IDF для векторизации;
- 4) в качестве классификаторов наиболее эффективна логистическая регрессия для русского языка, метод опорных векторов — для английского.

В качестве направления дальнейшего исследования можно предложить экспериментальное определение стоп-слов, удаление которых снижает эффективность модели классификации. Кроме того, дальнейшие работы могут быть связаны с улучшением результатов распознавания путем применения более сложных методов машинного обучения, в том числе нейросетевых архитектур.

СПИСОК ЛИТЕРАТУРЫ

1. *Acheampong F. A., Wenyu C., Nunoo-Mensah H.* Text-based emotion detection: Advances, challenges, and opportunities // *Engineering Reports*. 2020. Vol. 2, N 7. P. e12189. DOI: 10.1002/eng2.12189.
2. *Dzedzickis A., Kaklauskas A., Bucinskas V.* Human emotion recognition: Review of sensors and methods // *Sensors*. 2020. Vol. 20, N 3. P. 592. DOI: 10.3390/s20030592.
3. *Рюмина Е. В., Карпов А. А.* Аналитический обзор методов распознавания эмоций по выражениям лица человека // *Научно-технический вестник информационных технологий, механики и оптики*. 2020. Т. 20, № 2. С. 163—176. DOI: 10.17586/2226-1494-2020-20-2-163-176.
4. *Мубаракишина Р. Т., Яковенко Р. Т.* Обзор подходов к проблеме распознавания эмоций по параметрам устной речи // *Системный анализ в проектировании и управлении*. 2019. Т. 23, № 1. С. 392—397.
5. *Богданов А. Л., Дуля И. С.* Сентимент-анализ коротких русскоязычных текстов в социальных медиа // *Вестник Томского государственного университета. Экономика*. 2019. № 47. С. 220—241. DOI: 10.17223/19988648/47/17.
6. *Дюличева Ю. Ю.* Учебная аналитика MOOK как инструмент анализа математической тревожности // *Вопросы образования*. 2021. № 4. С. 243—265. DOI: 10.17323/1814-9545-2021-4-243-265.
7. *Adoma A. F., Henry N. M., Chen W.* Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition // 2020 17th Intern. Computer Conf. on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). 2020. P. 117—121. DOI: 10.1109/iccwamtip51612.2020.9317379.
8. *Verkholyak O., Dvoynikova A., Karpov A.* A Bimodal Approach for Speech Emotion Recognition using Audio and Text // *J. Internet Serv. Inf. Secur.* 2021. Vol. 11, N 1. P. 80—96.

9. Liu Y., Fu G. Emotion recognition by deeply learned multi-channel textual and EEG features // *Future Generation Computer Systems*. 2021. Vol. 119. P. 1—6. DOI: 10.1016/j.future.2021.01.010.
10. Овсянникова В. В. К вопросу о классификации эмоций: категориальный и многомерный подходы // *Финансовая аналитика: проблемы и решения*. 2013. Т. 37, № 175. С. 43—48.
11. Ekman P. Basic emotions // *Handbook of cognition and emotion*. 1999. P. 45—60.
12. Изард К. Э. Психология эмоций. СПб: Питер, 1999. 464 с.
13. Sogancioglu G., Verkholyak O., Kaya H., Fedotov D., Cadée T., Salah A. A., Karpov A. Is Everything Fine, Grandma? Acoustic and Linguistic Modeling for Robust Elderly Speech Emotion Recognition // *INTERSPEECH*. 2020. P. 2097—2101. DOI: 10.21437/interspeech.2020-3160.
14. Russell J. A. Culture and the categorization of emotions // *Psychological bulletin*. 1991. Vol. 110, N 3. P. 426—450. DOI: 10.1037/0033-2909.110.3.426.
15. Двойникова А. А., Карпов А. А. Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных // *Информационно-управляющие системы*. 2020. № 4(107). С. 20—30. DOI:10.31799/1684-8853-2020-4-20-30.
16. Henry E. R., Hofrichter J. Singular value decomposition: Application to analysis of experimental data // *Methods in enzymology*. Academic Press, 1992. Vol. 210. P. 129—192. DOI: 10.1016/0076-6879(92)10010-B.
17. Pennington J., Socher R., Manning C. D. Glove: Global vectors for word representation // *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*. 2014. P. 1532—1543. DOI: 10.3115/v1/d14-1162.
18. Bojanowski P., Grave E., Joulin A., Mikolov T. Enriching word vectors with subword information // *Transactions of the association for computational linguistics*. 2017. Vol. 5. P. 135—146. DOI: 10.1162/tacl_a_00051.
19. Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. Distributed representations of words and phrases and their compositionality // *Advances in neural information processing systems*. 2013. Vol. 26. P. 1—9.
20. Devlin J., Chang M. W., Lee K., Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding // *arXiv preprint arXiv:1810.04805*. 2018. DOI: 10.48550/arXiv.1810.04805.
21. Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L. Deep contextualized word representations // *Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2018. Vol. 1. P. 2227—2237.
22. Halim L. R., Suryadibrata A. Cyberbullying Sentiment Analysis with Word2Vec and One-Against-All Support Vector Machine // *IJNMT (International Journal of New Media Technology)*. 2021. Vol. 8, N 1. P. 57—64. DOI: 10.31937/ijnmt.v8i1.2047.
23. Duong H. T., Nguyen-Thi T. A. A review: preprocessing techniques and data augmentation for sentiment analysis // *Computational Social Networks*. 2021. Vol. 8, N 1. P. 1—16. DOI: 10.1186/s40649-020-00080-x.
24. Perepelkina O., Kazimirova E., Konstantinova M. RAMAS: Russian multimodal corpus of dyadic interaction for affective computing // *Intern. Conf. on Speech and Computer*. Springer, Cham, 2018. P. 501—510. DOI: 10.1007/978-3-319-99579-3_52.
25. Двойникова А. А., Верхоляк О. В., Карпов А. А. Сентимент-анализ разговорной речи при помощи метода, основанного на тональных словарях // *Альманах научных работ молодых ученых Университета ИТМО*. 2020. Т. 3. С. 75—80.
26. Zadeh A. B., Liang P. P., Poria S., Cambria E., Morency L. P. Multimodal language analysis in the wild: Cmumosei dataset and interpretable dynamic fusion graph // *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics*. 2018. Vol. 1: Long Papers. P. 2236—2246. DOI: 10.18653/v1/p18-1208.

Сведения об авторах

Анастасия Александровна Двойникова

— Санкт-Петербургский Федеральный исследовательский центр РАН, лаборатория речевых и многомодальных интерфейсов; мл. научный сотрудник; E-mail: dvoynikova.a@iias.spb.su

Кристина Олеговна Кондратенко

— бакалавр; Санкт-Петербургский государственный университет, кафедра фонетики и методики преподавания иностранных языков; E-mail: st076959@student.spbu.ru

Поступила в редакцию 29.05.23; одобрена после рецензирования 07.06.23; принята к публикации 28.08.23.

REFERENCES

1. Acheampong F.A., Wenyu C., Nunoo-Mensah H. *Engineering Reports*, 2020, no. 7(2), pp. e12189, DOI: 10.1002/eng2.12189.
2. Dzedzickis A., Kaklauskas A., Bucinskas V. *Sensors*, 2020, no. 3(20), pp. 592, DOI: 10.3390/s20030592.
3. Ryumina E.V., Karpov A.A. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, no. 2(20), pp. 163–176, DOI: 10.17586/2226-1494-2020-20-2-163-176. (in Russ.)
4. Mubarakshina R.T., Yakovenko R.T. *Sistemnyy analiz v proyektirovani i upravlenii* (System Analysis in Design and Management), 2019, no. 1(23), pp. 392–397. (in Russ.)
5. Bogdanov A.L., Dulya I.S. *Tomsk State University Journal of Economics*, 2019, no. 47, pp. 220–241, DOI: 10.17223/19988648/47/17. (in Russ.)
6. Dyulicheva Yu. *Voprosy obrazovaniya* (Educational Studies), 2021, no. 4, pp. 243–265, DOI: 10.17323/1814-9545-2021-4-243-265. (in Russ.)
7. Adoma A.F., Henry N.M., Chen W. *17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2020. pp. 117–121, DOI: 10.1109/iccwamtip51612.2020.9317379.
8. Verkholyak O., Dvoynikova A., Karpov A. *J. Internet Serv. Inf. Secur.*, 2021, no. 1(11), pp. 80–96.
9. Liu Y., Fu G. *Future Generation Computer Systems*, 2021, vol. 119, pp. 1–6, DOI: 10.1016/j.future.2021.01.010.
10. Ovsyannikova V.V. *Financial Analytics: Science and Experience*, 2013, no. 175(37), pp. 43–48. (in Russ.)
11. Ekman P. *Handbook of cognition and emotion*, 1999, pp. 45–60.
12. Izard C.E. *The psychology of emotions*, NY, London, Plenum Press, 1991.
13. Sogancioglu G., Verkholyak O., Kaya H., Fedotov D., Cadée T., Salah A. A., Karpov A. *INTERSPEECH*, 2020, pp. 2097–2101, DOI: 10.21437/interspeech.2020-3160.
14. Russell J.A. *Psychological bulletin*, 1991, no. 3(110), pp. 426–450, DOI: 10.1037/0033-2909.110.3.426.
15. Dvoynikova A.A., Karpov A.A. *Information and Control Systems*, 2020, no. 4(107), pp. 20–30, DOI: 10.31799/1684-8853-2020-4-20-30. (in Russ.)
16. Henry E.R., Hofrichter J. *Methods in enzymology*, Academic Press, 1992, vol. 210, pp. 129–192, DOI: 10.1016/0076-6879(92)10010-B.
17. Pennington J., Socher R., Manning C.D. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543, DOI: 10.3115/v1/d14-1162.
18. Bojanowski P., Grave E., Joulin A., Mikolov T. *Transactions of the association for computational linguistics*, 2017, vol. 5, pp. 135–146, DOI: 10.1162/tacl_a_00051.
19. Mikolov T., Sutskever I., Chen K., Corrado G.S., Dean J. *Advances in neural information processing systems*, 2013, vol. 26, pp. 1–9.
20. Devlin J., Chang M.W., Lee K., Toutanova K. *arXiv preprint arXiv:1810.04805*, 2018, DOI: 10.48550/arXiv.1810.04805.
21. Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018, vol. 1, pp. 2227–2237.
22. Halim L.R., Suryadibrata A. *IJNMT (International Journal of New Media Technology)*, 2021, no. 1(8), pp. 57–64, DOI: 10.31937/ijnmt.v8i1.2047.
23. Duong H.T., Nguyen-Thi T.A. *Computational Social Networks*, 2021, no. 1(8), pp. 1–16, DOI: 10.1186/s40649-020-00080-x.
24. Perepelkina O., Kazimirova E., Konstantinova M. *International Conference on Speech and Computer*, Springer, Cham, 2018, pp. 501–510, DOI: 10.1007/978-3-319-99579-3_52.
25. Dvoynikova A.A., Verkholyak O.V., Karpov A.A. *Almanac of Scientific Works of Young Scientists of ITMO University*, 2020, vol. 3, pp. 75–80. (in Russ.)
26. Zadeh A.B., Liang P.P., Poria S., Cambria E., Morency L.P. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, 2018, vol. 1, Long Papers, pp. 2236–2246, DOI: 10.18653/v1/p18-1208.

Data on authors

- | | | |
|---------------------------------|---|--|
| Anastasia A. Dvoynikova | — | St. Petersburg Federal Research Center of the RAS, Speech and Multimodal Interfaces Laboratory; Junior Researcher;
E-mail: dvoynikova.a@iias.spb.su |
| Khrystyna O. Kondratenko | — | Bachelor; St. Petersburg State University, Department of Phonetics and Methods of Teaching Foreign Languages;
E-mail: st076959@student.spbu.ru |

Received 29.05.23; approved after reviewing 07.06.23; accepted for publication 28.08.23.