

А. А. БРАНИЦКИЙ, Е. В. ДОЙНИКОВА, И. В. КОТЕНКО
Санкт-Петербургский институт информатики и автоматизации Российской академии наук
(СПИИРАН), Санкт-Петербург

МЕТОДИКА КЛАССИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНОЙ СЕТИ ПО ПСИХОЛОГИЧЕСКИМ ШКАЛАМ ТЕСТА АММОНА НА ОСНОВЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Разработана методика классификации пользователей социальной сети по психологическим шкалам теста Аммона. В качестве классификаторов были использованы искусственные нейронные сети нескольких видов (машина опорных векторов, линейная регрессия, многослойная нейронная сеть и сверточная нейронная сеть). Область применения методики – выявление отклонений в психологическом состоянии пользователей социальных сетей и мониторинг этого изменения. Проведен эксперимент, в результате которого было получено, что наилучшими показателями точности обладает многослойная нейронная сеть с функцией активации типа ReLU.

Введение. Социальные сети представляют собой Интернет-ресурсы, предназначенные для общения, поиска друзей и обмена цифровой информацией. Их популярность в кругу молодого поколения является высокой, и распространение вредоносной информации в социальных сетях может оказывать деструктивное влияние на поведение этой группы пользователей. В частности, пропаганда насильственных действий и вредных привычек может отрицательно сказываться на психологическом состоянии и мировоззрении Интернет-пользователей. Для обнаружения такого влияния в современной литературе предложено множество подходов.

В [1] рассматриваются сверточные нейронные сети, предназначенные для обнаружения отклонений в психологическом состоянии пользователей сети Интернет. С этой целью обучение нейронных сетей выполнялось с использованием данных, которые были собраны из сервисов социальных микроблогов. Были выделены два вида признаков, которые описывали анализируемый нейронными сетями поток данных: (1) данные внутри сообщения ограниченной длины (твита), которые содержали текст или изображение, (2) статистические данные, которые включали количество сообщений от пользователя, число комментариев к сообщению (эти данные собирались в рамках заданного промежутка времени). В проведенных авторами экспериментах наилучшую точность (более 78.5 %) показала четырехслойная нейронная сеть с методом субдискретизации mean-over-time.

В [2] предлагается решение задачи определения психологического портрета и индивидуальных особенностей характера человека. С этой целью в качестве входных данных использовались фотографии, собранные из соответствующих профилей в социальной сети Facebook, а в качестве метода анализа и классификатора – анализ коэффициентов корреляции Спирмена и логистическая регрессия. В результате выполненного авторами исследования было отмечено, что фотографии, размещенные в профилях коммуникабельных людей, содержат светлые области, а в фотографиях необщительных людей наблюдаются замкнутые области.

В [3] используется глубокая нейронная сеть, построенная для прогнозирования настроения отправителя сообщений. Вектор признаков, подаваемый на вход нейронной сети, формируется иерархически: часть признаков извлекается из отдельных символьных последовательностей, другая часть – из предложений.

В [4] применяется наивный байесовский классификатор, обученный для выявления уровня эмоционального воздействия, которое производится на человека в результате просмотра им заранее подобранных изображений. При формировании вектора признаков использовались различные параметры, касающиеся цветовых палитр изображения.

Анализ рассмотренных работ подтверждает актуальность задачи построения систем, направленных на обнаружение подверженности пользователей деструктивному воздействию. В отличие от этих работ, исследование, представленное в настоящем докладе, направлено на разработку комплексной методики, охватывающей анализ разнородных данных (скалярных величин, текста, изображений) и сочетающей в себе искусственные нейронные сети.

Методика. В рамках разработанной методики выделяются три шага:

1. На первом шаге собираются данные, описывающие профили пользователей социальной сети;
2. На втором шаге формируется вектор признаков, подлежащий обработке при помощи искусственных нейронных сетей;
3. На третьем шаге выполняется обучение нейронных сетей или классификация пользователей по шкалам теста Аммона.

Первый шаг методики. Собранные данные о пользователях социальной сети были размечены по 18 психологическим шкалам теста Аммона (6 Я-функций с компонентом конструктивности, 6 Я-функций с компонентом деструктивности, 6 Я-функций с компонентом дефицитарности). Каждая шкала имеет значение Т-балла (T): если $0 \leq T \leq 39$, то значение Т-балла низкое; если $40 \leq T \leq 60$, то значение Т-балла среднее, если $61 \leq T \leq 110$, то значение Т-балла высокое. В анализируемом наборе данных представлено 460 уникальных записей о профилях пользователей социальной сети.

Второй шаг методики. После сбора данных, касающихся профилей пользователей социальной сети, было выполнено вычисление параметров, формирующих векторы признаков. Данные векторы используются для обучения нейронной сети и вычисления значения шкалы теста Аммона. Исходные данные, описывающие вектор признаков, были сгруппированы по трем типам: (I) численные параметры (например, количество фотографий и друзей в профиле пользователя), (II) набор слов (например, наиболее часто используемые слова внутри каждой психологической шкалы), (III) параметры, полученные в результате обработки изображений.

Третий шаг методики. Из трех компонентов психологических шкал теста Аммона был выбран компонент деструктивности, в качестве классификаторов использовались машина опорных векторов (МОВ), линейная регрессия (ЛР), многослойная нейронная сеть (МНС) с функцией активации ReLU и сверточная нейронная сеть (СНС). Первые два классификатора обучались только на численных параметрах (параметрах первого типа), для МНС вектор признаков включал параметры первого, второго и третьего типов, а также посты, обучение СНС осуществлялось только с использованием постов, размещенных пользователем социальной сети.

Эксперимент. В таблице представлена точность классификаторов, вычисленная для деструктивного компонента шести Я-функций. При проведении эксперимента использовалась 10-блочная перекрестная проверка, в результате применения которой было выявлено, что наилучшей точностью обладает классификатор, построенный на основе МНС.

Таблица

Точность классификаторов

Классификатор						
МНС (I тип)	МНС (I, II типы)	МНС (I, II, III типы)	МНС (I, II, III типы, посты)	МОВ (I тип)	ЛР (I тип)	СНС (посты)
Деструктивный компонент первой Я-функции						
68,74%	69,62%	70,1%	70,15%	68,49%	30,89%	64,48%
Деструктивный компонент второй Я-функции						
57,87%	62,18%	59,36%	59,76%	55,4%	41,84%	46,94%
Деструктивный компонент третьей Я-функции						
48,96%	49,89%	52,32%	52,82%	48,35%	39,21%	44,51%
Деструктивный компонент четвертой Я-функции						
66,23%	66,47%	66,65%	66,12%	65,07%	30,41%	58,99%
Деструктивный компонент пятой Я-функции						
50,43%	48,51%	51,39%	51,59%	48,57%	41,12%	43,85%
Деструктивный компонент шестой Я-функции						
67,39%	67,62%	67,8%	67,6%	67,79%	47,74%	59,06%
Среднее значение						
59,94%	60,71%	61,27%	61,34%	58,95%	38,54%	52,97%

Заключение. В докладе была предложена методика классификации пользователей социальной сети по психологическим шкалам теста Аммона. В результате эксперимента были исследо-

ваны несколько типов нейронных сетей, среди которых наибольшей точностью обладает МНС. Направление дальнейших исследований – расширение набора вычисляемых признаков.

Работа проводилась при поддержке гранта РФФИ 18-29-22034 мк

ЛИТЕРАТУРА

1. Lin H., Jia J., Guo Q., Xue Y., Li Q., Huang J., Cai L., Feng L. User-level psychological stress detection from social media using deep neural network. Proceedings of the 22nd ACM international conference on Multimedia. 2014. P. 507–516.
2. Segalin C., Celli F., Polonio L., Kosinski M., Stillwell D., Sebe, N., Cristani M., Lepri B. What your Facebook profile picture reveals about your personality. Proceedings of the 2017 ACM on Multimedia Conference. 2017. P. 460–468.
3. dos Santos C., Gatti M. Deep convolutional neural networks for sentiment analysis of short texts. Proceedings of COLING 2014, the 25-th International Conference on Computational Linguistics: Technical Papers. 2014. P. 69–78.
4. Machajdik J., Hanbury A. Affective image classification using features inspired by psychology and art theory. Proceedings of the 18th ACM international conference on Multimedia. 2010. P. 83–92.

A.A. Branitskiy, E.V. Doynikova, I.V. Kotenko, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), St. Petersburg

Technique of Classification of Social Network Users by the Psychological Scales of the Ammon Test Based on Artificial Neural Networks

The technique of classification of social network users by the psychological scales of the Ammon test has been developed. Artificial neural networks of several types (support vector machine, linear regression, multilayer neural network and convolutional neural network) were used as classifiers. The scope of the technique is to identify deviations in the psychological state of social network users and monitor this change. Experiment has been carried out, as a result of which it was found that a multilayer neural network with an activation function of the ReLU type has the best accuracy.