УДК 004.89:004.932:159.9

ПРИМЕНЕНИЕ VOW-СЛОВАРЕЙ НА БАЗЕ KAZE-ДЕСКРИПТОРОВ ДЛЯ ПРОГНОЗИРОВАНИЯ ЛИЧНОСТНЫХ ФАКТОРОВ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ

¹Латышев А.В., ²Фраленко В.П.

¹ООО «РИТЕХ», Москва, e-mail: andrey.latyshev@gmail.com;

²ФГБУН Институт программных систем им. А.К. Айламазяна Российской академии наук
(ИПС им. А.К. Айламазяна РАН), Вельково, e-mail: alarmod@pereslavl.ru

В статье в качестве инструментальных средств оценки личностных факторов «Большой пятерки» предлагается применять современные методы автоматической обработки изображений: извлекать из изображений в профилях пользователей социальной сети «ВКонтакте» VoW-словари на базе инвариантных к повороту, смещению и изменению освещения KAZE-дескрипторов с дальнейшей их обработкой на искусственных нейронных сетях с целью построения инструмента для прогнозирования личностных факторов пользователей. Для формирования обучающей и тестовой выборки была использована база данных изображений и анкет пользователей социальной сети, изготовленная при непосредственном участии Психологического института Российской академии образования, анкетирование осуществлялось по опроснику «NEO-FFI». Нейронные сети реализованы с помощью библиотеки «Microsoft Cognitive Toolkit» (CNTK). Исследованы различные архитектуры нейронных сетей, разнообразные активационные функции и оптимизаторы обучения. Проводится сравнение с методами прямой нейросетевой обработки изображений. Выделены личностные факторы, коррелирующие с размещаемыми изображениями. Предложенный метод оказался лучше в задаче прогнозирования личностных факторов, чем подход с двумя нейронными сетями (прямого распространения и сверточной), где первая сжимает данные, а вторая уже обучается на них.

Ключевые слова: личностные факторы, «Большая пятерка», социальная сеть, дескрипторы, изображения, анкетирование

APPLICATION OF BOW-DICTIONARIES ON THE BASIS OF KAZE-DESCRIPTORS FOR FORECASTING PERSONAL FACTORS OF SOCIAL NETWORKSUSERS

¹Latyshev A.V., ²Fralenko V.P.

¹Ritech, Moscow, e-mail: andrey.latyshev@gmail.com;

²Ailamazyan Program Systems Institute of Russian Academy of Sciences,
Veskovo, e-mail: alarmod@pereslavl.ru

The article proposes to apply modern methods of automatic image processing as a tool for assessing «Big Five» personal factors: offered extract BoW-dictionaries based on invariant to rotation, displacement, and changes in the lighting of KAZE-descriptors from the images in the profiles of users of the «VKontakte» social network with their processing on artificial neural networks in order to build a tool for predicting the personal factors of users. To form training and test samples we used a database of images and profiles of a social networkusers, made with the direct participation of the Psychological Institute of Russian Academy of Education, questioning was carried out according to the «NEO-FFI» questionnaire. Neural networks are implemented using the Microsoft Cognitive Toolkit (CNTK) library. Various neural network architectures, various activation functions, and learning optimizers have been investigated. A comparison is made with methods of direct neural network image processing. Identified personal factors that correlate with the posted images. The proposed method turned out to be better in the task of predicting personality factors than the approach with two neural networks (direct distribution and convolutional), where the first compresses the data, and the second is already trained on them.

Keywords: personality factors, Big Five, social network, descriptors, images, questioning

В «Большую пятерку» личностных факторов (черт) входят открытость к опыту, интеллект («Openness»); сознание, самосознание, добросовестность («Consciousness»); экстраверсия, энергичность поведения, склонность к контактам («Extraversion»); доброжелательность, приятность, способность прийти к согласию («Agreeableness»); нейротизм, эмоциональная неустойчивость, тревожность, низкое самоуважение («Neuroticism») – сокращенно OCEAN [1]. Настоящая работа продолжает исследование применимости нейросетей для прогнозирования личностных факторов [2, 3].

Предложен принципиально новый подход, предполагающий уход от оценки отдельных изображений к формированию специальных VoW-словарей (BagsOfWords) с дальнейшей нейросетевой обработкой соответствующих дескрипторов.

Методы на основе VoW-словарей весьма популярны в том числе для фильтрации нежелательного контента на веб-страницах. Например, работа [4] посвящена подходу к обнаружению порнографических изображений. В качестве шага предобработки предложен метод вычисления инвариантных к масштабу SIFT-дескрипторов. Далее

они используются в качестве входных данных BoW-алгоритма. В результате получается выборка с обучающими векторами очень большой размерности, для работы с которыми предложен алгоритм ARCX4 на основе мультиномиального наивного байесовского классификатора [5]. Полнота обнаружения нежелательного контента полученным классификатором составляет от 87,93 % до 91,75 % (в зависимости от размера обучающей выборки). В работе [6] представлена экспериментальная оценка классификации изображений с использованием BoW-дескрипторов и нейронных сетей. Все изображения масштабировались без искажений пропорций до одного и того же масштаба. F1-мера при работе с дескрипторами достигла 78 %. Наше же исследование опирается на новые инвариантные к повороту, смещению и изменениям освещения KAZE-дескрипторы [7].

Исследование посвящено созданию современных инструментальных средств анализа данных социальных сетей для проведения научных исследований в предметной области. Предлагается применять современные методы машинного обучения по графической информации пользователей и экспертной информации. Цель исследования состоит в упрощении и автоматизации работы психологов с существенными объемами данных социальных сетей, в том числе – в облегчении работы с сырыми (необработанными) данными. При этом инструментальные средства помогают в извлечении статистических данных и в их обработке на высокопроизводительной вычислительной технике с целью извлечения и приобретения новых знаний.

Материалы и методы исследования

Описание предложенного подхода

На этапе подготовки обучающей выборки отобраны альбомы пользователей социальной сети «ВКонтакте» с числом изображений от 5 штук. Всего были использованы данные 84 пользователей: 42 альбома для обучения и 42 для тестирования (база данных изображений и анкет пользователей социальной сети, прошедших специальное анкетирование по опроснику «NEO-FFI» и разрешивших использовать эти данные). База сформирована при участии Психологического института Российской академии образования с целью проведения корректного сравнительного анализа полученных результатов с результатами, описанными в работе [2], указанная работа посвящена применению сверточных искусственных нейронных сетей (ИНС) для оценки связи

изображений, размещаемых пользователями, с их личностными факторами.

Из альбомов последовательно считываются изображения до тех пор, пока не будет достигнут лимит по максимальному количеству обрабатываемых изображений пользователя (для сокращения времени обучения/тестирования). В экспериментах использовалось ограничение в 160 штук, что обеспечило порядка 100 изображений на одного пользователя социальной сети. Обрабатываемые изображения приводились к фиксированному масштабу, так, чтобы количество пикселей составляло около 300 тысяч, что приблизительно соответствует размеру 548x548 пикселей. Каждому профилю пользователя сопоставлен json-файл со сведениями об уровне факторов «Большой пятерки», сведения получены с помощью опросника «NEO-FFI» [8]. Низкий уровень: 0–20, средний: 21–32, высокий: 32–48.

Далее из каждого изображения пользователя с порогом чувствительности 0,005 (KAZE-Response) вычисляются KAZE-дескрипторы, каждый дескриптор представляет собой вектор с числами с плавающей точкой. Затем выполняется кластеризация KAZE-дескрипторов с использованием метода k-средних [9]. При этом на старте задается ожидаемое число кластеров на выходе алгоритма кластеризации. Для того чтобы отсеять KAZE-дескрипторы, наиболее далекие от выделенных кластеров, для каждого дескриптора обучающей выборки находится кластер, наиболее близкий к нему по расстоянию Евклида. Затем по всей обучающей выборке считается distVAG, среднее расстояние от дескрипторов до их кластеров-победителей. Полученное расстояние в дальнейшем используется для отсева тех дескрипторов, что далеки от кластеров (считаем их случайными выбросами).

Оставшиеся дескрипторы отдельных пользователей социальной сети используются для формирования усредненных векторов встречаемости KAZE-дескрипторов на изображениях, то есть для непосредственной подготовки так называемых мешков визуальных слов, или обобщающих BoW-дескрипторов пользователей (Bag-of-visual Words). Размер BoW-дескриптора определяется числом используемых кластеров. На первом шаге вычисления BoW-дескрипторов для отдельных пользователей предлагается в виде вектора записать количество выявленных KAZE-дескрипторов с расстоянием до ближайшего центра кластера $\leq \text{distVAG}$. На втором шаге для обеспечения относительной нечувствительности к разному количеству изображений в со-

ответствующих профилях пользователей осуществляется нормирование полученных векторов путем деления на количество изображений в данных профилях. На основе полученных данных фиксируется дополнительный вектор из максимальных нормированных значений, вектор используется на заключительном, третьем шаге для получения значений векторов VoW-дескрипторов в диапазоне от 0,15 до 0,85 или в диапазоне от -0,85 до 0,85 (для разных активационных функций на выходном слое ИНС). Например, для получения значений в указанном диапазоне все нормированные значения делятся на соответствующие им максимумы, далее полученные величины умножаются на 0,70, и к ним прибавляется 0,15. Данная предобработка необходима для нормализации работы искусственной нейронной сети, в том числе за счет ухода от нулей в VoW-дескрипторах обучающей выборки.

Обучение нейронных сетей

Далее происходит обучение искусственных нейронных сетей прямого распространения, реализация на основе библиотеки «Microsoft Cognitive Toolkit» (CNTK) версии 2.6 [10]. В экспериментах размер словаря варьировался от 64 до 16384 элементов, чтобы подобрать лучший вариант, обеспечивающий высокую точность про-

гнозирующих нейронных сетей. Изначально обучалось два варианта сетей, с одним и тремя слоями. Далее при обобщении полученных результатов был сформирован третий вариант. Выбор лучших конфигураций осуществлялся с помощью тестирования разных конфигураций ИНС, изменялось число нейронов в соответствующих слоях, был выполнен перебор по сетке разных сочетаний числа нейронов, активационных функций и иных настроек. Исследовалась работа активационных функций, поддерживаемых CNTK [11], дополнительно проверены вейвлет-функции PolyWog [12]. Список оптимизаторов обучения был ограничен набором тех, что поддерживается в CNTK [13].

Первый вариант архитектуры ИНС:

- входной слой, 125–1000 нейронов;
- dropout-слой с вероятностью срабатывания 0,01 [14];
- выходной слой 5 нейронами.

Наилучший результат показал вариант с 2048 элементами словаря и 250 нейронами входного слоя с ReLU-активационной функцией $f(x) = \max(0, x)$, для выходного слоя активационная функция не использовалась ($f(x) = x$), нейросеть обучалась с помощью Adam-оптимизатора [15] с коэффициентами L1- и L2-регуляризации 0,0001 и 0,01 соответственно (табл. 1, 2).

Таблица 1

Совместное прогнозирование пяти факторов, обучающая выборка
(первый вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	0,29	0,32	0,35	0,32	0,27
Усредненное значение СКВО по отдельным факторам	0,31	Среднее значение СКВО по отдельным пользователям		0,28	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	1,00	Точность выделения наименее четко выраженного фактора (от 0 до 1)		1,00	

Таблица 2

Совместное прогнозирование пяти факторов, тестовая выборка
(первый вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	12,54	14,36	12,40	11,65	15,99
Усредненное значение СКВО по отдельным факторам	13,39	Среднее значение СКВО по отдельным пользователям		12,43	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	0,24	Точность выделения наименее четко выраженного фактора (от 0 до 1)		0,29	

Второй вариант архитектуры ИНС:
 – входной слой, 64–1000 нейронов;
 – dropout-слой с вероятностью срабатывания 0,01;
 – скрытый слой, 50–800 нейронов;
 – dropout-слой с вероятностью срабатывания 0,01;
 – скрытый слой, 35–600 нейронов;
 – dropout-слой с вероятностью срабатывания 0,01;
 – выходной слой с 5 нейронами.

Наилучший результат показал вариант с 1536 элементами словаря, нейронная сеть в своих слоях содержала 250, 200 и 150 нейронов с ReLU-активационной функцией, использовался Adam-оптимизатор с коэффициентами L1- и L2-регуляризации 0,0001 и 0,01 соответственно, выходной слой также без активационной функции (табл. 3, 4).

Обобщая полученные результаты, следует отметить, что добавление дополнительных слоев в ИНС улучшило показатель СКВО на тестовой выборке (с 12,43–13,39 до 11,40–12,34). При этом точность выделения факторов колеблется незначительно. Оптимальное число нейронов в первом слое обеих сетей – 250 штук. В связи с этим было принято решение увеличить число слоев, сохранив архитектурные особенности ИНС и иные условия эксперимента. Для третьей ИНС взято 400,

350, 300, 250, 200 и 150 нейронов в слоях (вперемежку с dropout-слоями), использовалось 1536 элементов словаря. Число слоев подобрано экспериментально. Результаты тестирования показаны в табл. 5, 6.

Результаты исследования и их обсуждение

Для сравнения полученных результатов с предыдущими достижениями обратимся к работе [2]. В первом эксперименте там была исследована возможность использования сверточных нейронных сетей в качестве аппарата для оценки взаимосвязи изображений, размещаемых пользователями, с их личностными чертами. При этом никаких вспомогательных инструментов не использовалось, нейронная сеть выполняла прогноз личностных факторов за счет собственной архитектуры. Во втором эксперименте проверялось использование предварительно обученных сверточных нейросетей для получения признаков высокого уровня с дальнейшим их использованием для обучения второй нейронной сети прогнозу личностных факторов. При обработке пользовательских профилей вектор прогнозируемых личностных черт получался усреднением отдельных нейросетевых прогнозов по всем изображениям, размещенным в соответствующих профилях.

Таблица 3

Совместное прогнозирование пяти факторов, обучающая выборка
(второй вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	0,62	0,56	0,56	0,69	0,51
Усредненное значение СКВО по отдельным факторам	0,59	Среднее значение СКВО по отдельным пользователям		0,52	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	0,98	Точность выделения наименее четко выраженного фактора (от 0 до 1)		1,00	

Таблица 4

Совместное прогнозирование пяти факторов, тестовая выборка
(второй вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	10,89	13,73	11,11	10,87	15,11
Усредненное значение СКВО по отдельным факторам	12,34	Среднее значение СКВО по отдельным пользователям		11,40	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	0,21	Точность выделения наименее четко выраженного фактора (от 0 до 1)		0,31	

Таблица 5

Совместное прогнозирование пяти факторов, обучающая выборка
(третий вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	1,00	0,77	0,83	0,98	0,68
Усредненное значение СКВО по отдельным факторам	0,85	Среднее значение СКВО по отдельным пользователям		0,75	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	0,98	Точность выделения наименее четко выраженного фактора (от 0 до 1)		0,95	

Таблица 6

Совместное прогнозирование пяти факторов, тестовая выборка
(третий вариант архитектуры)

	О	С	Е	А	Н
Средние значения СКВО по отдельным факторам «Big Five»	8,74	12,64	8,97	9,92	13,85
Усредненное значение СКВО по отдельным факторам	10,82	Среднее значение СКВО по отдельным пользователям		9,98	
Точность выделения наиболее четко выраженного фактора (от 0 до 1)	0,21	Точность выделения наименее четко выраженного фактора (от 0 до 1)		0,38	

В указанном исследовании на тестовой выборке точность выделения наиболее четко выраженного фактора составила 0,19–0,21, что фактически есть случайный выбор, при использовании BoW-словарей и дескрипторов достигнута точность 0,21. Однако точность выделения наименее четко выраженного фактора в предыдущей работе составила 0,33–0,45, тут достигнута точность 0,38. Новый метод оказался лучше в задаче прогнозирования личностных факторов, чем подход с двумя нейронными сетями (сверточной и прямого распространения), где первая сжимает данные, а вторая уже обучается на них, однако он уступает сверточной нейронной сети, прогнозирующей личностные факторы своим выходным слоем нейронов.

Заключение

Очевидно, что в этой области добиться хороших результатов проблематично ввиду нечеткости и неоднозначности экспертных оценок личностных факторов. Предложенный метод сжатия информации на основе BoW-словарей сохраняет больше информации, пригодной для обучения ИНС прямого распространения, чем специализированная сверточная ИНС, при этом наблюдается

рост СКВО по отдельным факторам и отдельным пользователям.

Новые результаты, полученные в настоящей работе, показывают, что прогнозирование личностных факторов по BoW-дескрипторам уступает прогнозированию с прямой обработкой изображений с помощью нейронных сетей в части выбора наиболее и наименее выраженных факторов. Получено отличное запоминание обучающей выборки, результаты работы с тестовой выборкой могут быть улучшены путем существенного расширения обучающей выборки. На тестовых выборках хорошо прогнозируются факторы «Openness» и «Agreeableness», что подтверждает полученные ранее выводы о связи изображений, размещаемых пользователями социальной сети, с этими личностными факторами. Хуже всего прогнозируется «Neuroticism».

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 18-29-22003-мк «Разработка методов нейросетевого анализа визуального интернет-контента пользователей социальных сетей с целью автоматизированного определения выраженности личностных черт, связанных с психологическим неблагополучием»).

Список литературы

1. Tupes E.C., Christal R.E. Recurrent personality factors based on trait ratings. *Journal of Personality*. 1992. vol. 60. no. 2. P. 225–251. DOI: 10.1111/j.1467-6494.1992.tb00973.x.
2. Кисельникова Н.В., Куминская Е.А., Латышев А.В., Фраленко В.П., Хачумов М.В. Инструментальные средства анализа депрессивного состояния и личностных черт человека // Программные системы: теория и приложения. 2019. № 3. С. 129–159. DOI: 10.25209/2079-3316-2019-10-3-129-159.
3. Fralenko V., Khachumov V., Khachumov M. Correlation analysis and prediction of personality traits using graphic data collections. *SHS Web of Conferences*. 2019. vol. 72. Proceeding of the International Scientific Conference «Achievements and Perspectives of Philosophical Studies» (APPSCONF-2019). DOI: 10.1051/shsconf/20197201012. [Электронный ресурс]. URL: https://www.shs-conferences.org/articles/shsconf/pdf/2019/13/shsconf_appsconf2019_01012.pdf (дата обращения: 05.03.2020).
4. Do T.-N. Detection of pornographic images using Bag-Of-Visual-Words and ARCX4 of random multinomial naive bayes. *Journal of Science and Technology*. 2011. vol. 49. no. 5. P. 13–24.
5. Deselaers T., Pimenidis L., Ney H. Bag-of-visual-words models for adult image classification and filtering. *Proceeding of the 19th International Conference on Pattern Recognition*. 2008. P. 1–4.
6. Fralenko V.P., Suvorov R.E., Tikhomirov I.A. Automatic Image Classification for Web Content Filtering: New Dataset Evaluation. *Recent Developments and the New Direction in Soft-Computing Foundations and Applications (Selected Papers from the 6th World Conference on Soft Computing, May 22–25, 2016, Berkeley, USA)*. 2018. P. 351–360.
7. Alcantarilla P.F., Bartoli A., Davison A.J. KAZE Features. *Proceedings of the European Conference on Computer Vision (ECCV)* (October 2012, Firenze, Italy). 14 p. [Электронный ресурс]. URL: <http://www.robots.ox.ac.uk/papers/Alcantarilla12eccv.pdf> (дата обращения: 05.03.2020).
8. Costa P.T., McCrae R.R. Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological assessment*. 1992. vol. 4. no. 1. P. 5–13. DOI: 10.1037/1040-3590.4.1.5.
9. Celebi M.E., Kingravi H.A., Vela P.A. A comparative study of efficient initialization methods for the k-means clustering algorithm. 2013. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1209.1960> (дата обращения: 05.03.2020).
10. Microsoft Cognitive Toolkit. 2017. [Электронный ресурс]. URL: <https://www.microsoft.com/en-us/cognitive-toolkit/> (дата обращения: 05.03.2020).
11. Cntk.ops package – Python API for CNTK 2.6 documentation. 2017. [Электронный ресурс]. URL: <https://cntk.ai/pythondocs/cntk.ops.html> (дата обращения: 05.03.2020).
12. Ван Лянпэн. Программные средства с использованием методов и алгоритмов кратномасштабного вейвлет анализа для обработки и поиска изображений: дис. ... канд. тех. наук. Москва, 2019. 138 с.
13. Cntk.learners package – Python API for CNTK 2.6 documentation. 2017. [Электронный ресурс]. URL: <https://cntk.ai/pythondocs/cntk.learners.html> (дата обращения: 05.03.2020).
14. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*. 2014. vol. 15. P. 1929–1958.
15. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. *Proceedings of International Conference on Learning Representations (ICLR)*. 2015. 13 p. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1412.6980.pdf> (дата обращения: 05.03.2020).