



Е.Е. Фомина

ПРИМЕНЕНИЕ АППАРАТА НЕЙРОННЫХ СЕТЕЙ ДЛЯ АНАЛИЗА РЕЗУЛЬТАТОВ АНКЕТИРОВАНИЯ

Социологические исследования направлены на изучение и анализ социальной действительности, происходящих в обществе процессов и явлений. Инструментарий исследователя-социолога должен обязательно включать в себя современные методы и технологии обработки больших массивов, как структурированных, так и неструктурированных данных. К таким методам относятся искусственные нейронные сети, которые обладают широкими функциональными возможностями. Актуальность применения искусственных нейронных сетей для обработки социологической информации не вызывает сомнений. Их основное преимущество перед классическими методами многомерной статистики заключается в том, что не накладываются ограничения на исходные данные, т.е. не требуется наличия нормального распределения или линейной зависимости между переменными, как это необходимо при использовании многих статистических методов, таких как факторный или дискриминантный анализ. Кроме того, нейронные сети способны обрабатывать данные, представленные в разных шкалах, что особенно актуально при работе с социологической информацией. Задача настоящей статьи – демонстрация возможностей искусственных нейронных сетей для анализа результатов анкетирования в социологических исследованиях. Искусственные нейронные сети, в частности самоорганизующиеся карты признаков Кохонена, были применены для обработки результатов анкетирования, посвященного теме благотворительности. Построена карта признаков, состоящая из шести кластеров, каждый из которых объединил респондентов в зависимости от их мотивации к благотворительной деятельности. Выявлены социально-демографические характеристики каждого кластера. Далее к полученным результатам был применен метод деревьев классификации, который позволил построить классификационные правила для отнесения новых респондентов к одному из имеющихся кластеров.

Ключевые слова: *нейронные сети, самоорганизующиеся карты признаков Кохонена, деревья классификации, анализ результатов анкетирования.*

Введение. Социологические исследования – один из основных инструментов изучения и анализа процессов, происходящих в обществе. Они, как правило, сопряжены с обработкой большого количества разнородных данных, получаемых в результате анкетирований и опросов; в результате изучения динамики каких-либо показателей исследуемого процесса или явления; в результате анализа документов [1]. Совершенствование инструментов сбора информации приводит к тому, что объем данных постоянно увеличивается [2]. Неоднородную, зачастую неструктурированную, быстро поступающую информацию обработать традиционными методами математической статистики становится затруднительно. Необходимо внедрять в практику со-

© Фомина Е.Е., 2020

Фомина Елена Евгеньевна – канд. техн. наук, доцент кафедры информатики и прикладной математики ФГБОУ ВО «Тверской государственный технический университет», e-mail: f-elena2008@yandex.ru.

циолога новые технологии обработки и анализа больших массивов разнородных данных. В работах [3–5] сформулированы требования, предъявляемые к этим технологиям:

- понятность в использовании социологам, которые не являются специалистами в области математического моделирования, искусственного интеллекта и программирования;
- возможность обработки больших объемов как структурированных, так и слабо структурированных и неструктурированных данных, представленных в разных шкалах;
- возможность поиска в массиве первичной информации, новых выводов и закономерностей, поддающихся интерпретации и имеющих практическое значение для решения задачи, а также выявлять и описывать аномалии.

Таким образом, внедрение современных методов обработки данных в социологическую практику является актуальной задачей. Они позволят повысить как эффективность обработки информации, полученной в результате исследований, так и научную обоснованность получаемых выводов и рекомендаций.

К таким методам относится интеллектуальный анализ, включающий использование искусственных нейронных сетей (ИНС). Основное преимущество нейронных сетей перед классическими методами многомерной статистики заключается в том, их инструментарий не накладывает ограничений на исходные данные. Так, например, не требуется наличия нормального распределения или наличия линейной зависимости между переменными, как это необходимо при использовании многих статистических методов [6, 7]. ИНС способны анализировать разнородные данные, что особенно актуально при обработке социологической информации. Кроме того, ИНС являются эффективным инструментом исследования сложных объектов или процессов, когда формализация задачи затруднена [8, 9].

Публикации последних лет свидетельствуют о том, что нейронные сети все активнее внедряются в набор математических инструментов, используемых социологами. Они применяются для предобработки и подготовки исходных данных. В этой связи можно отметить работы E.-L. Silva-Ramirez [10, 11], S. Yoon и S. Lee [12], O.B. Сташковой и O.B. Шестопал [13], в которых нейронные сети, а именно многослойный персептрон с обратным распространением ошибки, эффективно применяются для решения задачи импутации значений интервальных переменных. ИНС используются для решения задач кластеризации и классификации социологической информации. Исследованиям в этой области посвящены работы O.B. Чудовой [14], B.B. Круглова и M.И. Дли [15], A.Ф. Шакирова [16]. В работах O.K. Яковлевой [17], A. Гомзако [18], П.С. Абросимовой [19] аппарат нейронных был применен для решения задач прогнозирования.

Потенциал использования ИНС в социологии велик. Цель настоящей статьи – продемонстрировать возможности ИНС при обработке результатов анкетирования.

Методы исследования. В качестве метода исследования был использован аппарат самоорганизующихся карт признаков Кохонена, который позволил сформировать и описать группы респондентов, давших сходные ответы по вопросам, касающимся мотивации к участию в благотворительности. Далее, к полученным результатам применялся метод деревьев классификации для составления решающих правил, позволяющих производить классификацию новых респондентов.

Самоорганизующиеся карты признаков Кохонена (СКП Кохонена) – это одна из разновидностей нейронных сетей с обучением без учителя. Они предназначены для решения задач векторной классификации. В качестве исходных данных метода выступает набор объектов, каждому из которых сопоставлен вектор значений признаков. Требуется разбить исходное множество на кластеры, т.е. для каждого объекта найти кластер, которому он принадлежит.

Сеть Кохонена содержит входной слой и слой активных нейронов. Поскольку входной слой просто распределяет входные сигналы по нейронам активного слоя, такую сеть можно считать однослойной [20, 21].

В результате применения метода будет построена двумерная карта, в которой схожие объекты будут расположены рядом, образуя, таким образом, кластеры. Проецирование многомерного пространства на двумерное позволяет визуализировать результат и упростить его интерпретацию.

Метод позволяет решать задачу классификации новых объектов, но не позволяет строить решающих правил классификации. Поэтому в совокупности с СКП Кохонена могут быть использованы деревья классификации.

Метод деревьев классификации (ДК) позволяет изучать статистическую взаимосвязь между одной зависимой и группой независимых переменных, а также определять принадлежность объектов тому или иному классу в зависимости от значений переменных, характеризующих объекты [22, 23]. Метод не накладывает каких-либо ограничений на тип исходных данных, позволяет визуализировать результаты классификации в виде дерева, а также формулировать набор правил, позволяющих отнести объект к тому или иному классу, что существенно упрощает интерпретацию [24].

На сегодняшний день разработано большое число алгоритмов, позволяющих реализовывать деревья классификаций. Самые популярные из них это алгоритмы CART, C4.5, QUEST, CHAID [25–27].

Методика комбинация СКП Кохонена и ДК позволит выработать эффективный алгоритм структурирования и анализа информации, который включает в себя два этапа: разбиение множества объектов на кластеры и применение ДК для построения наглядного решающего правила распределения объектов.

Материалы исследования. *Применение самоорганизующихся карт признаков Кохонена и деревьев классификации для анализа результатов анкетирования.* В качестве примера рассмотрим анкетирование, посвященное

теме благотворительности в России. Одной из задач анкетирования являлась задача формирования групп респондентов в зависимости от их мотивации к участию в благотворительности, а также описание этих групп.

База данных с результатами анкетирования включала в себя более 140 переменных (вопросов) различного типа (категориальные, логические, интервальные, порядковые и другие) и более 5000 записей. В качестве инструментария исследования использовался пакет STATISTICA [28].

На первом этапе из базы данных были отобраны переменные, включающие в себя вопросы, касающиеся сформулированной выше задачи исследования, и социально-демографические характеристики респондентов. Далее проводилась процедура предобработки данных, состоящая из удаления записей, содержащих ошибки, записей, содержащих пропуски по вопросам, относящимся к задаче исследования. В результате была сформирована выборка из 2300 записей.

После этапа предобработки данных решалась задача кластеризации с использованием самоорганизующихся карт признаков Кохонена и задача классификации с использованием ДК.

В ходе решения задачи кластеризации проводился сравнительный анализ результатов при различном числе кластеров. Наилучшую интерпретацию решения удалось получить при числе кластеров, равном шести:

первый кластер – респонденты, основной мотив участия в благотворительности которых – это попытка самоутверждения и стремление заслужить одобрение окружающих (10 % от общего числа записей в выборке);

второй кластер – респонденты, для которых участие в благотворительности связано со стремлением облегчить свою совесть (11 % от общего числа записей в выборке);

третий кластер (самый многочисленный) – респонденты, для которых основной мотив участия в благотворительности – это проявления чувств сострадания, милосердия и сочувствия (34 % от общего числа записей в выборке);

четвертый кластер – респонденты, для которых стремление почувствовать себя нужными людьми является доминирующим фактором (17 % от общего числа записей в выборке);

пятый кластер – респонденты, основной мотив участия в благотворительности которых – чувство долга (11 % от общего числа записей в выборке);

шестой кластер – респонденты, для которых чувство долга и чувства сострадания, милосердия и сочувствия в равной степени являются определяющими факторами (17 % от общего числа записей в выборке).

Анализ карт Кохонена позволил описать каждый кластер и выявить его особенности.

Так, *первый кластер* составляют респонденты, возраст которых 20–39 лет, преобладают мужчины (60 %), свой уровень дохода характеризуют как низкий

или средний, их участие в благотворительности проявляется в виде передачи вещей нуждающимся в ответ на обращение по телевидению или радио и безвозмездной сдачи крови.

Второй кластер объединил приблизительно равное количество мужчин и женщин, возраст респондентов – 40–54 лет, проживают в городах с населением до 100 тыс. жителей или селах, свой уровень дохода характеризуют как низкий или средний. Участие в благотворительности связано с переводом небольших денежных средств на счет благотворительных организаций, передачей вещей нуждающимся, работой на добровольных началах.

Третий кластер – респонденты в возрасте 55 лет и старше; преобладают женщины (61 %), проживают как в крупных и средних, так и в малых городах, участие в благотворительности связано, в основном, с передачей вещей в ответ на обращение по телевидению и радио.

Четвертый кластер включает в себя приблизительно равное количество мужчин и женщин всех возрастных категорий, свой уровень дохода характеризуют как средний или высокий, участие в благотворительности связано с передачей вещей и переводом денежных средств в ответ на обращение по телевидению и радио.

Пятый кластер включает в себя также примерно равное количество мужчин и женщин всех возрастных категорий, свой уровень дохода характеризуют как средний или высокий, участие в благотворительности связано в основном с переводом крупных денежных средств на благотворительные цели.

Шестой кластер – в основном женщины, которые свой уровень дохода характеризуют как средний или высокий; участие в благотворительности связано в основном с работой волонтером, безвозмездной сдачей крови, переводом небольших денежных средств на счет благотворительных фондов.

Модуль кластеризации с использованием самоорганизующихся карт признаков Кохонена в программе STATISTICA позволяет наряду с задачей разбиения объектов на группы решать задачу классификации новых объектов. Однако он не позволяет сформулировать классификационные правила. Для решения этой задачи удобно использовать аппарат деревьев классификации.

Построение дерева осуществлялось с использованием метода CART. В качестве условия прекращения ветвления была выбрана прямая остановка по методу FACT. На рисунке представлено дерево решений, на котором для примера выделена ветка и классификационные правила, позволяющие отнести респондента к шестому кластеру: *если* {[пол = *женский*] и [размер населенного пункта = *малый город*] и [(образование = *среднее*) или (образование = *средне-специальное*)] и [(доход = *средний*) или (доход = *высокий*)]} *то* респондент относится к *кластеру 6*.

При V-кратной кросс-проверке средняя доля ошибочно классифицированных наблюдений составила 5,82 %, что говорит о хорошем качестве модели.

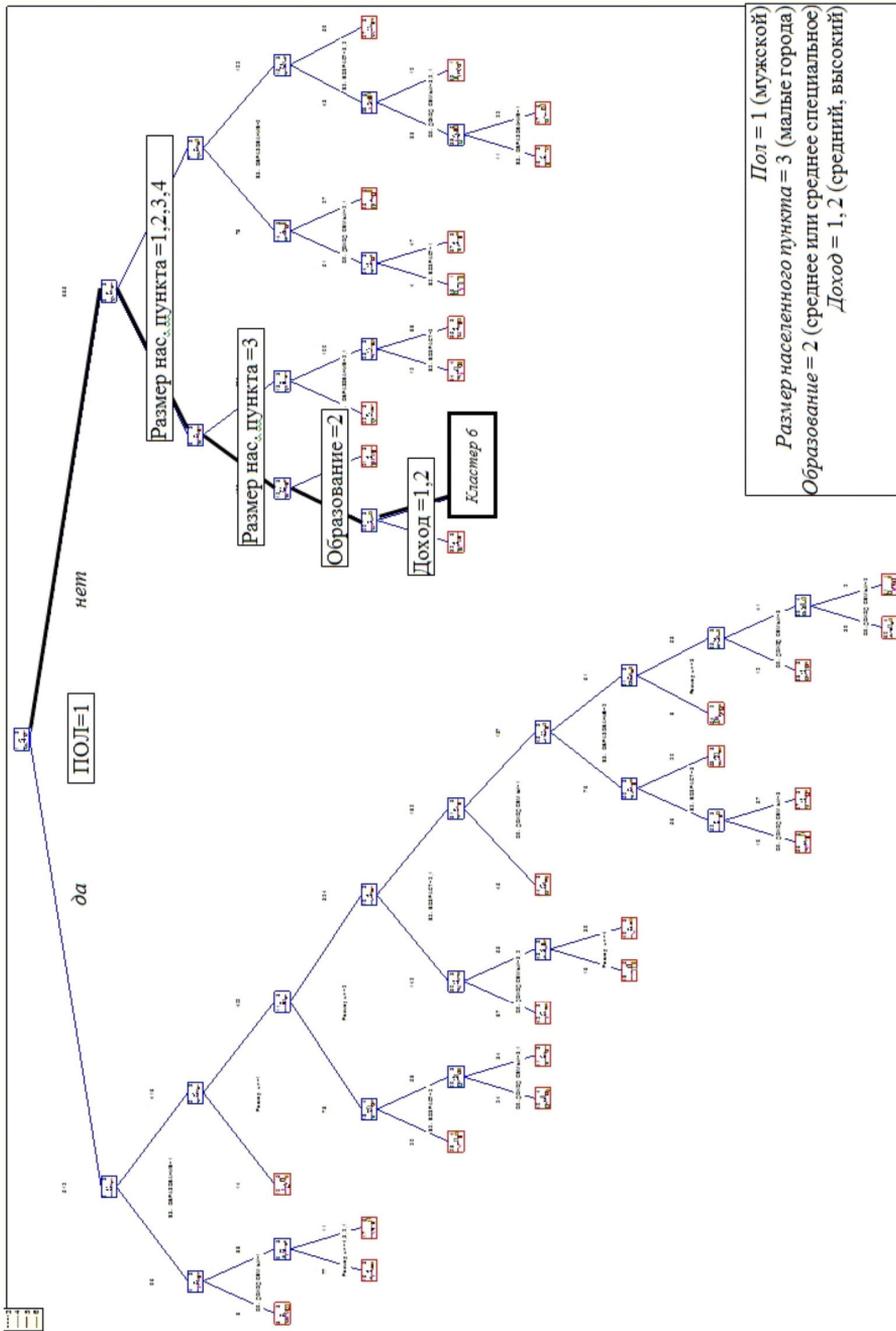


Рис. Дерево классификации

Выводы. Задачей статьи являлась демонстрация возможностей нейронных сетей при анализе социологической информации. Нейронные сети, в частности самоорганизующиеся карты признаков Кохонена, были применены для анализа результатов анкетирования, посвященного теме благотворительности. Было выделено и описано шесть групп респондентов в зависимости от их мотивации к благотворительной деятельности. Для классификации новых объектов применен аппарат деревьев решений. Получены классификационные правила.

Таким образом, искусственные нейронные сети могут быть эффективно применены в социологической практике.

Список литературы

1. Шилина М.Г., Левченко В.Ю. BIG DATA, OPEN DATA, LINKED DATA, метаданные в PR: актуальные модели трансформации теории и практики // Медиаскоп. – 2014. – № 1. – С. 16.
2. Качалов Д.Л., Мишустин А.В., Фархадов М.П. Современные методы обработки больших данных в крупномасштабных системах // Математические модели современных экономических процессов, методы анализа и синтеза экономических механизмов. Актуальные проблемы и перспективы менеджмента организаций в России: материалы XI Всерос. науч.-практ. конф. / Самар. науч. центр РАН. – Самара, 2017. – С. 65–71.
3. Шайтура С.В. Интеллектуальный анализ геоданных // Перспективы науки и образования. – 2015. – № 6 (18). – С. 24–30.
4. Tsvetkov V.Ya. Information field // Life Science Journal. – 2014. – № 11 (5). – P. 551–554.
5. Tsvetkov V.Ya. Information interaction as a Mechanism of Semantic Gap Elimination // European Researcher. – 2013. – Vol. 45, № 4–1. – P. 782–786.
6. Фомина Е.Е. Методы многомерной статистики в социологических и социально-экономических исследованиях. – Тверь, 2019. – 112 с.
7. Буреева Н.Н. Многомерный статистический анализ с использованием ППП «STATISTICA»: учебно-методический материал по программе повышения квалификации «Применение программных средств в научных исследованиях и преподавании математики и механики» / Нижегород. гос. ун-т им. Н.И. Лобачевского. – Н. Новгород, 2007. – 112 с.
8. Касторнова В.А., Можяева М.Г. Искусственные нейронные сети как современные средства информатизации // Информационная среда образования и науки. – 2012. – № 7. – С. 18–34.
9. Борзилова Ю.С. Применение нейросетевых информационных технологий в социально-экономической сфере // Актуальные направления научных исследований XXI века: теория и практика. – 2015. – Т. 3, № 7–3 (18–3). – С. 390–393.

10. Missing value imputation on missing completely at random data using multilayer perceptrons / E.-L. Silva-Ramírez [et al.] // *Neural Networks*. – 2011. – Vol. 24, iss. 1. – P. 121–129.

11. Single imputation with multilayer perceptron and multiple imputation combining multilayer perceptron and k-nearest neighbours for monotone patterns / E.-L. Silva-Ramírez [et al.] // *Applied Soft Computing*. – 2015. – № 29. – P. 128–132.

12. Yoon S., Lee S. Training algorithm with incomplete data for feed-forward neural networks // *Neural Processing Letters*. – 1999. – № 10(3). – P. 171–179.

13. Сташкова О.В., Шестопад О.В. Использование искусственных нейронных сетей для восстановления пропусков в массиве исходных данных // *Изв. вузов. Северо-Кавказский регион. Технические науки*. – 2017. – № 1 (193). – С. 37–42.

14. Чудова О.В. Применение нейронных сетей в социологии // *Социология в современном мире: наука, образование, творчество*. – 2009. – № 1. – С. 82–87.

15. Круглов В.В., Дли М.И. Применение аппарата нейронных сетей для анализа социологических данных // *Социологические исследования*. – 2001. – № 9. – С. 112–114.

16. Шакирова А.Ф. Особенности применения нейросетевого и пространственного анализа в социологических науках на примере построения индекса социальных настроений жителей города Казани // *История, политология, социология, философия: теоретические и практические аспекты: сб. ст. по материалам XXIII междунар. науч.-практ. конф.* – Новосибирск: СибАК, 2019. – № 8(16). – С. 47–50.

17. Яковлева О.К. К вопросу использования нейросетевых технологий в социологических исследованиях // *Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века: сб. ст. по материалам Третьей всерос. науч.-практ. конф.* / Перм. гос. нац. исслед. ун-т. – Пермь, 2018. – С. 149–155.

18. Гомзяков А. Применение нейронных сетей для прогнозирования результатов президентских выборов // *Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века: сб. ст. по материалам Третьей всерос. науч.-практ. конф.* / Перм. гос. нац. исслед. ун-т. – Пермь, 2018. – С. 200–206.

19. Абросимова П.С. Использование методов искусственного интеллекта в прогнозировании среднего возраста населения страны // *Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века: сб. ст. по материалам Третьей всерос. науч.-практ. конф.* / Перм. гос. нац. исслед. ун-т. – Пермь, 2018. – С. 213–217.

20. Kohonen T. *Self-organization and associative memory*. – Berlin: Springer, 1987. – 352 с.

21. Шматов Г.П., Фомина Е.Е. Нейронные сети и генетический алгоритм. – Тверь: Изд-во ТвГТУ, 2019. – 200 с.
22. Бова А. Деревья решений как техника добычи данных // Социология: теория, методы, маркетинг. – 2002. – № 1. – С. 128–136.
23. Толстова Ю.Н. Анализ социологических данных. Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Научный мир, 2000. – 352 с.
24. Фомина Е.Е. Использование методов многомерной статистики для анализа социальной и экономической информации // Экономика. Социология. Право. – 2018. – № 2 (10). – С. 61–67.
25. Фомина Е.Е. Возможности метода деревьев классификации при обработке социологической информации // Гуманитарный вестник. – 2018. – № 11 (73). – С. 5.
26. Фомина Е.Е. Подготовка и анализ результатов анкетирования с применением математических методов // Социосфера. – 2018. – № 2. – С. 194–198.
27. Генрихов И.Е., Дюкова Е.В. Усовершенствование алгоритма c4.5 на основе использования полных решающих деревьев // Математические методы распознавания образов. – 2009. – Т. 14, № 1. – С. 104–107.
28. Боровиков В. STATISTICA. Искусство анализа данных на компьютере. – СПб.: Питер, 2003. – 688 с.

References

1. Shilina M.G., Levchenko V.Iu. Big Data, Open Data, Linked Data, metadannye v PR: aktual'nye modeli transformatsii teorii i praktiki [Big Data, Open Data, Linked Data, Metadata in PR: Current models in the transformation of theory and practice]. *Mediascope*, 2014, no. 1, p. 16.
2. Kachalov D.L., Mishustin A.V., Farkhadov M.P. Sovremennye metody obrabotki bol'shikh dannykh v krupnomasshtabnykh sistemakh [Modern methods of Big Data analysis in large-scale systems]. *Matematicheskie modeli sovremennykh ekonomicheskikh protsessov, metody analiza i sinteza ekonomicheskikh mekhanizmov. Aktual'nye problemy i perspektivy menedzhmenta organizatsii v Rossii*. Proceedings of XI All-Russ. Sci-Pract. Conf. Samara, Samara Scientific Centre of RAS, 2017, pp. 65–71.
3. Shaitura S.V. Intellektual'nyi analiz geodannykh [Geodata mining]. *Perspektivy nauki i obrazovaniia*, 2015, no. 6(18), pp. 24–30.
4. Tsvetkov V.Ya. Information field. *Life Science Journal*, 2014, no. 11(5), pp. 551–554.
5. Tsvetkov V.Ya. Information interaction as a mechanism of semantic gap elimination. *European Researcher*, 2013, vol. 45, no. 4–1, pp. 782–786.

6. Fomina E.E. *Metody mnogomernoi statistiki v sotsiologicheskikh i sotsial'no-ekonomicheskikh issledovaniiax* [Methods of multidimensional statistics in sociological and socio-economic studies]. Tver, 2019, 112 p.

7. Bureeva N.N. *Mnogomernyi statisticheskii analiz s ispol'zovaniem PPP "STATISTICA"* [Multivariate statistical analysis with the software package "STATISTICA"]. *Uchebno-metodicheskii material po programme povysheniia kvalifikatsii "Primenenie programnykh sredstv v nauchnykh issledovaniiax i prepodavanii matematiki i mekhaniki"*. Nizhny Novgorod, Lobachevsky State University of Nizhny Novgorod, 2007, 112 p.

8. Kastornova V.A., Mozhaeva M.G. *Iskusstvennye neironnye seti kak sovremennye sredstva informatizatsii* [Artificial neural networks as modern means of informatization]. *Informatsionnaia sreda obrazovaniia i nauki*, 2012, no. 7, pp. 18–34.

9. Borzilova Iu.S. *Primenenie neirossetevykh informatsionnykh tekhnologii v sotsial'no-ekonomicheskoi sfere* [Application of neural network information technologies in the socio-economic sphere]. *Aktual'nye napravleniia nauchnykh issledovaniy XXI veka: teoriia i praktika*, 2015, vol. 3, no. 7–3(18–3), pp. 390–393.

10. Silva-Ramírez E.-L. [et al.]. *Missing value imputation on missing completely at random data using multilayer perceptrons*. *Neural Networks*, 2011, vol. 24, no. 1, pp. 121–129.

11. Silva-Ramírez E.-L. [et al.]. *Single imputation with multilayer perceptron and multiple imputation combining multilayer perceptron and k-nearest neighbours for monotone patterns*. *Applied Soft Computing*, 2015, no. 29, pp. 128–132.

12. Yoon S., Lee S. *Training algorithm with incomplete data for feed-forward neural networks*. *Neural Processing Letters*, 1999, no. 10(3), pp. 171–179.

13. Stashkova O.V., Shestopal O.V. *Ispol'zovanie iskusstvennykh neironnykh setei dlia vosstanovleniia propuskov v massive iskhodnykh dannykh* [Use of artificial neural networks for restoration of initial data array]. *Izvestiia vysshikh uchebnykh zavedenii. Tekhnicheskie nauki*. North Caucasus region, 2017, no. 1(193), pp. 37–42.

14. Chudova O.V. *Primenenie neironnykh setei v sotsiologii* [Application of neural networks in sociology]. *Sotsiologiya v sovremennom mire: nauka, obrazovanie, tvorchestvo*, 2009, no. 1, pp. 82–87.

15. Kruglov V.V., Dli M.I. *Primenenie apparata neironnykh setei dlia analiza sotsiologicheskikh dannykh* [Application of neural network apparatus for analysing sociological data]. *Sotsiologicheskie issledovaniia*, 2001, no. 9, pp. 112–114.

16. Shakirova A.F. *Osobennosti primeneniia neirossetevogo i prostranstvennogo analiza v sotsiologicheskikh naukakh na primere postroeniia indeksa sotsial'nykh nastroenii zhitelei goroda Kazani* [Features of the use of neural network and spatial analysis in the sociological sciences on the example of building an index of social moods of residents of the city of Kazan]. *Istoriia, politologiya, sotsiologiya*,

filosofia: teoreticheskie i prakticheskie aspekty. Proceedings of XXIII Int. Sci.-Pract. Conf., no. 8(16), Novosibirsk SibAK, 2019, pp. 47–50.

17. Iakovleva O.K. K voprosu ispol'zovaniia neirosetevykh tekhnologii v sotsiologicheskikh issledovaniiaxh [To the question of the use of neuro network technologies in sociological researches]. *Iskusstvennyi intellekt v reshenii aktual'nykh sotsial'nykh i ekonomicheskikh problem XXI veka*. Proceedings of the 3rd All-Russ. Sci.-Pract. Conf. Perm State National Research University, Perm, 2018, pp. 149–155.

18. Gomziakov A. Primenenie neironnykh setei dlia prognozirovaniia rezultatov prezidentskikh vyborov [Application of neural networks for predicting results of presidential elections]. *Iskusstvennyi intellekt v reshenii aktual'nykh sotsial'nykh i ekonomicheskikh problem XXI veka*. Proceedings of the 3rd All-Russ. Sci.-Pract. Conf. Perm State National Research University, Perm, 2018, pp. 200–206.

19. Abrosimova P.S. Ispol'zovanie metodov iskusstvennogo intellekta v prognozirovanii srednego vozrasta naseleniia strany [Using methods of artificial intelligence in forecasting the middle age of country's population]. *Iskusstvennyi intellekt v reshenii aktual'nykh sotsial'nykh i ekonomicheskikh problem XXI veka*. Proceedings of 3rd All-Russ. Sci.-Pract. Conf. Perm State National Research University, Perm, 2018, pp. 213–217.

20. Kohonen T. Self-organization and associative memory. Berlin. Springer, 1987, 352 p.

21. Shmatov G.P., Fomina E.E. Neironnye seti i geneticheskii algoritm [Neural networks and the genetic algorithm]. Tver, Tver State Technical University, 2019, 200 p.

22. Bova A. Derev'ia reshenii kak tekhnika dobychi dannykh [Decision trees as a data mining technique]. *Sotsiologiya: teoriia, metody, marketing*, 2002, no. 1, pp. 128–136.

23. Tolstova Iu.N. Analiz sotsiologicheskikh dannykh [Analysis of sociological data]. *Metodologiya, deskriptivnaia statistika, izuchenie svyazei mezhdu nominal'nymi priznakami*. Moscow, Nauchnyi mir, 2000, 352 p.

24. Fomina E.E. Ispol'zovanie metodov mnogomernoi statistiki dlia analiza sotsial'noi i ekonomicheskoi informatsii [Use of multivariate statistics for the analysis of social and economic information]. *Ekonomika. Sotsiologiya. Pravo*, 2018, no. 2(10), pp. 61–67.

25. Fomina E.E. Vozmozhnosti metoda derev'ev klassifikatsii pri obrabotke sotsiologicheskoi informatsii [The potential of the method of classification trees in the sociological information processing]. *Gumanitarnyi vestnik*, 2018, no. 11(73), p. 5.

26. Fomina E.E. Podgotovka i analiz rezultatov anketirovaniia s primeneniem matematicheskikh metodov [Preparation and analysis of results of the questionnaire using mathematical methods]. *Sotsiosfera*, 2018, no. 2, pp. 194–198.

27. Genrikhov I.E., Diukova E.V. Usovershenstvovanie algoritma c4.5 na osnove ispol'zovaniia polnykh reshaiushchikh derev'ev [Improvement of c4.5 algorithm using full decision trees]. *Matematicheskie metody raspoznavaniia obrazov*, 2009, vol. 14, no. 1, pp. 104–107.

28. Borovikov V. STATISTICA. Iskusstvo analiza dannykh na komp'yutere [STATISTICA. The art of computer data analysis]. St. Petersburg, Piter, 2003, 688 p.

Оригинальность 80 %

Получено 05.12.2019 Принято 09.01.2020 Опубликовано 29.06.2020

E.E. Fomina

APPLICATION OF NEURAL NETWORKS FOR ANALYSIS OF THE SURVEY RESULTS

Sociological studies aim at studying and analyzing social reality, processes and phenomena taking place in society. The tools of a researcher in sociology must include modern methods and technologies for processing large arrays of both structured and unstructured data. Such methods include artificial neural networks which have wide functional capabilities. The relevance of using artificial neural networks for processing sociological information is doubtless. Their main advantage over the classical methods of multidimensional statistics is that their tools do not impose restrictions on the source data, so a normal distribution or a linear relationship between variables is not required, as is necessary when using many statistical methods, for example, such as factor or discriminant analysis. In addition, neural networks are able to process data presented on different scales, which is especially important when analyzing sociological information. The purpose of this article is to demonstrate the capabilities of artificial neural networks for analyzing survey results in sociological research. Artificial neural networks, in particular Kohonen self-organizing maps, were used to process the results of a questionnaire on the topic of charity. A sign map was constructed consisting of six clusters, each grouping respondents depending on their motivation for charity work. The socio-demographic characteristics of each cluster are revealed. Further, the classification tree method was applied to the obtained results, which made it possible to construct classification rules for assigning new respondents to one of the available clusters.

Keywords: neural networks, Kohonen self-organizing maps, classification trees, analysis of questionnaire results.

Elena E. Fomina – Candidate of Technical Sciences, Associate Professor, Department of Applied Mathematics and Informatics, Tver State Technical University, e-mail: f-elena2008@yandex.ru.

Received 05.12.2019 Accepted 09.01.2020 Published 29.06.2020