

Славутская Е.В., Славутский Л.А. Использование искусственных нейронных сетей для анализа гендерных различий младших подростков



English version: [Slavutskaya E.V., Slavutskii L.A. Using artificial neural networks for analysis of gender differences in younger teenagers](#)

Чувашский государственный педагогический университет им. И.Я.Яковлева, Чебоксары, Россия

Чувашский государственный университет им. И.Н.Ульянова, Чебоксары, Россия

[Сведения об авторах](#)

[Литература](#)

[Ссылка для цитирования](#)

Предлагается использование нейронной сети для обработки данных психодиагностики. Главной особенностью алгоритма является его наглядность и максимальная определенность в процессе обучения и использования нейронной сети. Для этого структура нейросети жестко связана с исходными анализируемыми данными, а обучение сети проводится с их же использованием. Построена простая нейросеть с прямой передачей сигнала и обратным распространением ошибки. Обучение и проверка работы сети производится на примере данных психологического тестирования младших подростков (111 учащихся школы). Показано, что алгоритм позволяет эффективно выделять психологические признаки, значимые для оценки гендерных различий.

Ключевые слова: нейронная сеть, психодиагностика, тестирование, гендерные различия

Интеллектуальный анализ данных (в частности, с применением искусственных нейронных сетей) в последние десятилетия находит все более широкое применение в задачах идентификации, прогнозирования, классификации данных, в том числе статистических [Хайкин, 2006; Боровиков, 2008]. Аппарат искусственной нейронной сети (ИНС) основан на принципах когнитивной науки [Hebb, 1961; Розенблатт, 1965; Baxt, 1994; Величковский, 2006] и может использоваться в практической психологии [Grossberg, 1974; Dorrer, 1995; Арзамасцев, 2003]. При этом широкого распространения в задачах обработки психодиагностических данных искусственные нейронные сети пока не получили. Главная причина этого – неопределенность в процессе их обучения и сложность в интерпретации полученных результатов. Предлагаемый в настоящей статье алгоритм оценки гендерных различий младших подростков обеспечивает наглядность и максимальную определенность в процессе обучения и использования нейросети.

Для компьютерной обработки данных психодиагностики повсеместно используются традиционные статистические методы анализа данных, такие как статистическая теория оценивания, корреляционный, факторный анализ и т.д. При этом результаты психодиагностики чаще всего имеют целочисленные значения и бывают неполными (отсутствует часть результатов тестирования некоторых испытуемых). Во многих случаях интерпретировать результаты приходится при большом числе исходных психологических признаков и недостаточном для полноценного статистического анализа числе респондентов. В таких условиях анализ взаимосвязей между психологическими признаками и оценка их значимости [Dorrer, 1995], которым посвящена настоящая работа, оказываются непростой задачей.

Перечисленные проблемы могут быть минимизированы, если для анализа психодиагностических данных используется аппарат искусственных нейронных сетей (ИНС) [Хайкин, 2006], позволяющий работать с ограниченной, неполной выборкой данных. При этом размерность данных не имеет решающего значения [Боровиков, 2008].

В настоящей работе предлагается использование детерминированного нейросетевого алгоритма, при реализации которого структура нейросети жестко связана с исходными анализируемыми данными, а ее обучение производится с их же использованием. Аппарат ИНС применяется для селективной оценки гендерных различий в эмоционально-волевой и интеллектуальной сферах школьников 10–11 лет. Это психологически очень сложный возрастной период, характеризующийся преддверием пубертатного кризиса и переходом из начальной школы в среднее звено [Поливанова, 2000].

Алгоритм обработки данных психодиагностики

На рис. 1 приведен пример результатов психологического тестирования школьников 10–11 лет (5-й класс общеобразовательной школы N 22, Чебоксары) с использованием 12-факторного опросника Р.Б.Кеттела и Р.В.Коэна CPQ (Childrens Personality Questionnaire) для детей 8–12 лет [Cattell, 1990]. Он предназначен для исследования личностных особенностей школьников и содержит 12 шкал для измерения степени выраженности черт личности, которые Р.Б.Кеттел называет конституциональными. На этом же рисунке приведена структура нейронной сети. Оценка коэффициента интеллекта IQ проводилась по методике «Культурно-независимый тест интеллекта Р.Б.Кеттела» [Cattell, 1990].

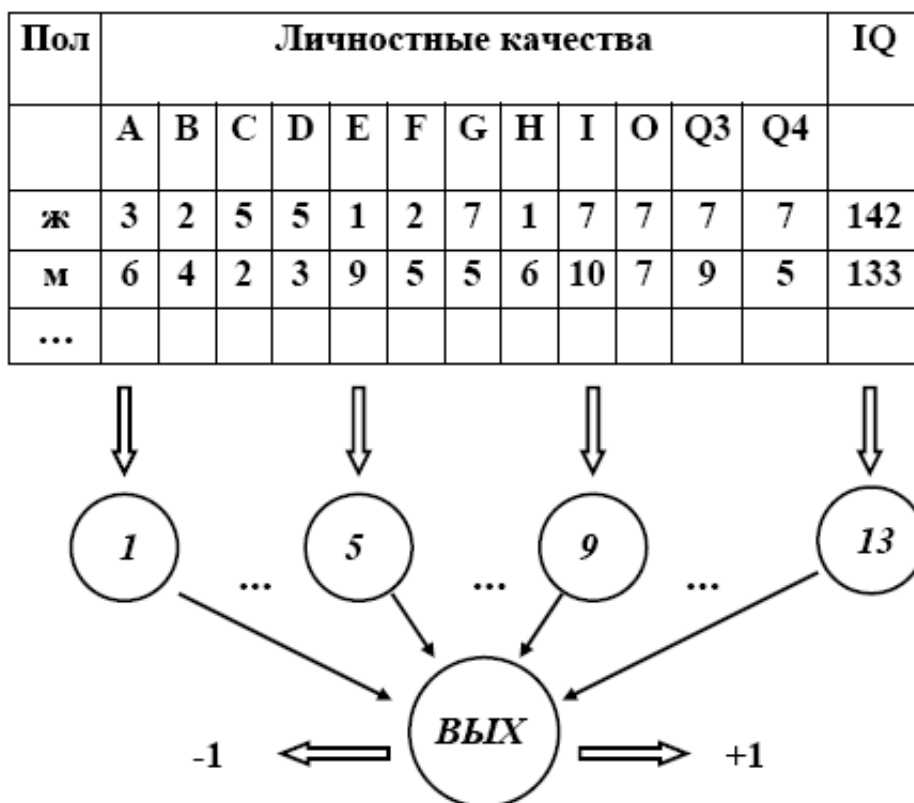


Рис. 1. Пример данных психодиагностики и структура нейросети.

Примечания. Интерпретация личностных качеств в 12-факторном опроснике Р.Б.Кеттела и Р.В.Коэна (CPQ): А – общительность–замкнутость; В – абстрактное-конкретное мышление; С –

эмоциональная стабильность – неустойчивость; D – возбудимость–уравновешенность; E – независимость–покорность; F – беспечность–озабоченность; G – высокая–низкая дисциплинированность; H – смелость–робость; I – мягкость–твердость; O – тревожность–спокойствие; Q3 – высокий–низкий самоконтроль; Q4 – напряженность–расслабленность.

Структура используемой нейросети, реализованной в среде MatLab [Медведев, 2002], соответствует анализируемым данным. Тип сети – двухслойная сеть с прямой передачей сигнала и обратным распространением ошибки (feed-forward backprop).

Построенная нейронная сеть (НС) имеет последовательную структуру, когда выход первого слоя служит входом второго слоя и т.д. Предполагаемое количество слоев – не более трех. В отличие от аппарата корреляционного и факторного анализа такая архитектура может воспроизводить нелинейные зависимости между входными и выходными данными сети. Каждый слой сети имеет свою матрицу весов, которая состоит из весовых функций нейронов. Количество нейронов в первом слое соответствует количеству исходных психологических признаков, и на вход каждого нейрона в первом слое последовательно подаются численные значения соответствующих результатов тестирования по этому признаку. Количество выходов сети соответствует количеству индивидуальных характеристик испытуемых, взятых за основу анализа. Например, при изучении значимости исходных признаков для гендерных различий на выходе получаем два целочисленных значения «-1» и «+1» в соответствии с половой принадлежностью тестируемого (мальчик или девочка).

Проверка функционирования предлагаемого алгоритма проводится с разными группами респондентов и разным количестве исходных психологических признаков. Например, сначала обучение проводится по трем группам учащихся, полученные результаты по каждой из групп фиксируются, а потом эти результаты сравниваются с соответствующими данными, полученными при обучении сети отдельно по каждой группе. Другой вариант – обучение сети по данным, полученным у респондентов в начале учебного года, а анализируются результаты психодиагностики, полученные с теми же учащимися в конце учебного года (через несколько месяцев). Принципиального отличия в работе предложенного алгоритма не наблюдалось, однако окончательно, для получения адекватных результатов и их корректного сравнения, исследования проводятся по следующей единой процедуре.

Пусть имеется таблица данных результатов тестирования N признаков по M респондентам. На каждый из N нейронов первого слоя нейросети подаются данные первого респондента, а на выходе фиксируется целевая реакция сети («-1» или «+1») в зависимости от гендерной принадлежности тестируемого. Сеть запоминает необходимые значения весовых функций нейронов. Процедура повторяется для данных каждого учащегося (M раз), и таким образом завершается первый цикл обучения. Процесс обучения прекращается, когда различие Δ между реакцией сети и целевыми функциями достигает некоторого малого значения или перестает уменьшаться. После окончания обучения, когда матрица весовых функций нейронов сформирована, на вход сети снова последовательно подаются исходные данные таблицы результатов тестирования и фиксируются выходные значения. При этом данные по одному или нескольким исходным признакам обнуляются. Основным результатом является количество ошибок на выходе нейросети при определении пола испытуемого в зависимости от того, данные какого из психологических признаков на входе сети оказались искажены. Результаты представляются в виде гистограмм или таблиц. Таким образом анализируется значимость того или иного психологического признака в общей совокупности гендерных различий тестируемых.

Результаты

Приведем результаты, полученные при совместной обработке данных 12-факторного опросника с данными культурно-независимого теста Р.Б.Кеттла при использовании простейшей двухслойной архитектуры сети (см. рис. 1). Общее количество исходных признаков $N = 13$ (12 качеств + IQ).

Учащиеся пятых классов были разделены на три группы: К-1 – с высоким и выше среднего уровнем интеллектуального развития, К-2 – со средним и низким уровнем интеллектуального развития, Э – экспериментальная группа учащихся с различным уровнем интеллектуального развития (по методике «Культурно-независимый тест интеллекта Р.Б.Кеттла»). На рис. 2 показаны графики обучения нейросети для 86 школьников (46 девочек и 40 мальчиков), а также отдельно – по контрольным группам К-1 (16 мальчиков и 19 девочек), К-2 (11 мальчиков и 11 девочек) и экспериментальной группе Э (13 мальчиков и 16 девочек).

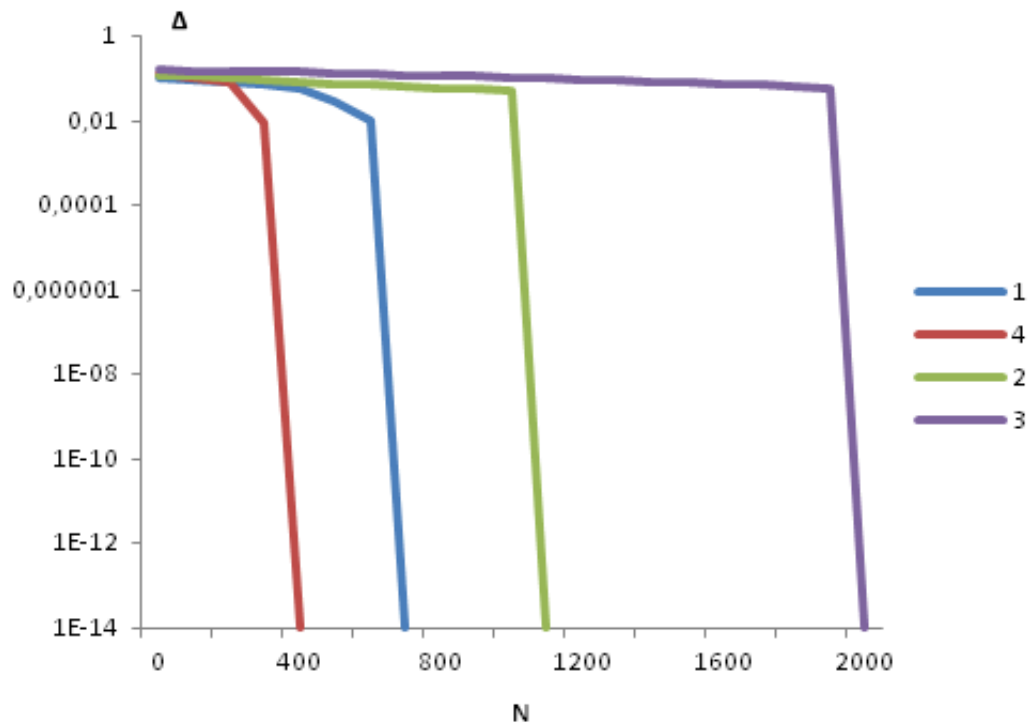


Рис. 2. Обучение нейросети, включая IQ.

Примечания. Зависимость ошибки Δ от числа циклов обучения N : 1 – все дети (670 циклов); 2 – группа К1 (1147 циклов); 3 – группа Э (2109 циклов); 4 – группа К2 (357 циклов). Описание групп см. в тексте.

Значение IQ имеет отличную от 12 остальных признаков размерность, поэтому возникает необходимость оценить работу нейросетевого алгоритма в этих условиях. Для сравнения строилась аналогичная НС с 12 нейронами в первом слое (по числу личностных качеств, исключая IQ). Графики обучения в этом случае имеют похожую форму, но время обучения (количество циклов) оказывается в десятки раз меньше: группа Э – 40 циклов; группа К1 – 21 цикл; все дети – 37 циклов; группа К2 – 23 цикла. Таким образом, добавление тринадцатого признака с другой размерностью в десятки раз увеличивает количество необходимых циклов обучения нейросети, но общие закономерности сохраняются:

1. Наибольшее число циклов для обучения нейросети потребовала группа Э как наиболее неоднородная с психологической точки зрения [Славутская, 2011].

2. Количество циклов обучения не зависит от числа испытуемых.

Для приведения к одинаковой размерности исходные данные психодиагностики нормировались. Пусть имеется матрица признаков Y с элементами y_{ij} . Матрица исходных признаков Y нормируется

следующим образом и переходит в матрицу Z с элементами:

$$z_{ij} = \frac{y_{ij} - \bar{y}_i}{S_i},$$

где $\bar{y}_i$ – среднее значение показателя в столбце i ; S_i – среднеквадратичное отклонение (СКО):

$$S_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}.$$

На рис. 3 приведены соответствующие графики обучения сети по 13 исходным признакам после нормировки исходных данных.

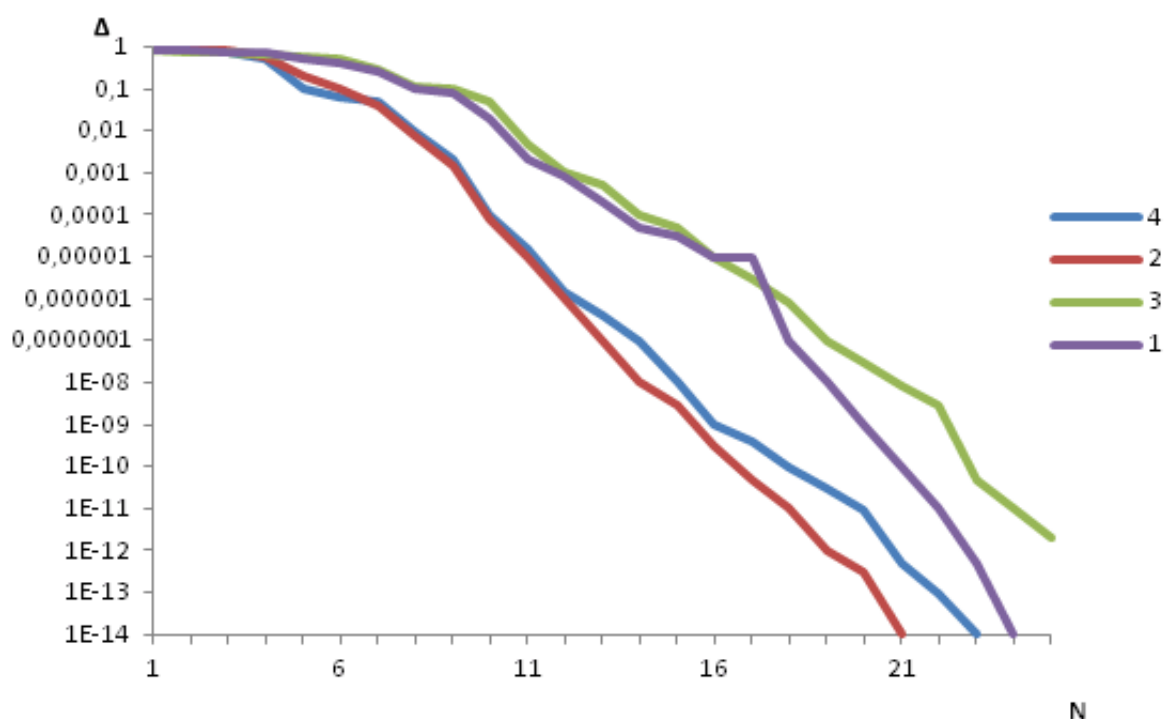


Рис. 3. Обучение нейросети, включая IQ после нормировки исходных данных.

Примечания. Зависимость ошибки Δ от числа циклов обучения N : 1 – все дети (24 цикла); 2 – группа К1 (22 цикла); 3 – группа Э (26 циклов); 4 – группа К2 (23 цикла). Описание групп см. в тексте.

Как видно из рисунка, после нормировки исходных данных время обучения сети сокращается до нескольких десятков циклов, но общее поведение графиков обучения в целом сохраняется, что может служить подтверждением устойчивой работы алгоритма.

В соответствии с предложенным алгоритмом для оценки значимости каждого из 13 признаков (включая IQ) для гендерных особенностей тестируемых каждый из признаков на входе сети последовательно обнулялся, результаты реакции сети фиксировались. Соответствующие гистограммы приведены на рис. 4. Как видно из рисунка, полученные результаты крайне неоднородны, как по 13 исходным качествам, так и по гендерной принадлежности. Обнуление некоторых признаков не влечет за собой ошибок в определении пола ребенка (нулевая высота столбцов гистограмм, рис. 4). С другой стороны, для некоторых из 13 качеств ошибка в определении пола ребенка при исключении (обнулении) входных данных по одному признаку влечет 100%-ную ошибку в определении половой принадлежности. Примером может служить качество

«С» в обработке данных тестирования по группе Э. В этом случае сеть выдала ошибку в определении пола всех 16 девочек. Согласно результатам обработки (рис. 4), для девочек наиболее значимыми являются качества «С» и «Q4», для мальчиков – «Е» и «Н», что соответствует выводам, сделанным по результатам факторного анализа [Славутская, 2008].

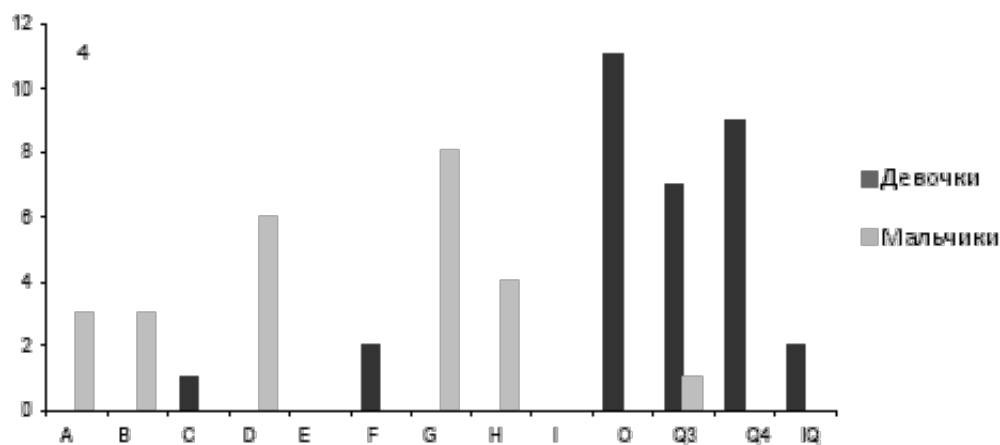
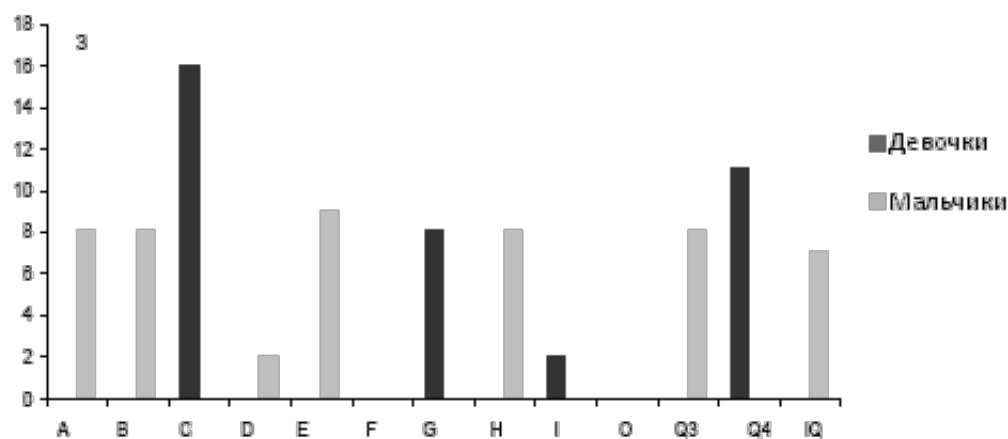
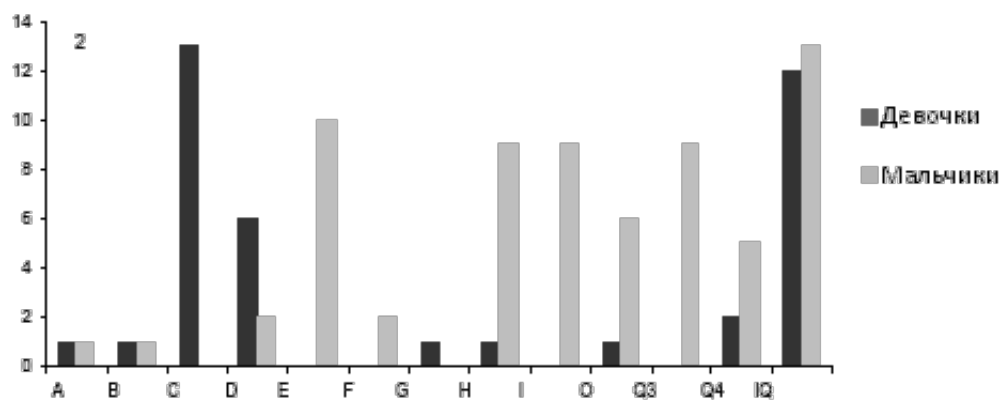
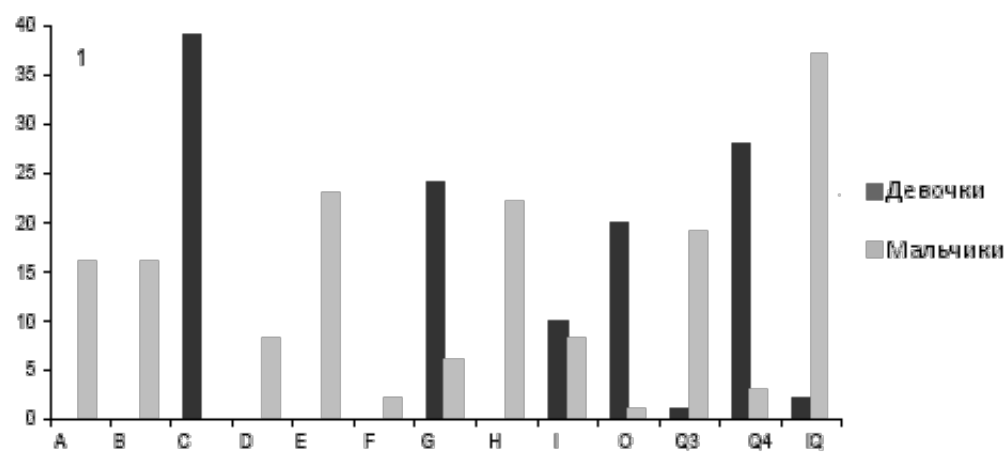


Рис. 4. Число ошибок в определении гендерной принадлежности испытуемых при последовательном

обнулении входных признаков.

Примечания. 1 – все дети, 2 – группа К1, 3 – группа Э, 4 – группа К2. Описание групп см. в тексте.

Анализ результатов

Результаты, аналогичные показанным на рис. 4, получены для 111 учащихся. Особое внимание при обработке результатов психодиагностики уделялось экспериментальной группе Э учащихся пятых классов, с которой проводилась коррекционно-развивающая работа и тестирование которой осуществлялось несколько раз в течение года по разным методикам [Славутская, 2011].

Сформулируем основные особенности, обнаруженные при нейросетевой обработке психодиагностических данных.

Нейросетевой алгоритм устойчиво работает при неоднородных данных. Примером могут служить приведенные выше результаты совместной обработки данных тестирования пятиклассников по двум методикам (12-факторного опросника и Культурно-независимого теста Р.Кеттела). При этом добавление в исходные данные значений коэффициента интеллекта IQ, имеющих другую размерность, не меняют общей картины распределения значимости 12 личностных качеств для гендерных различий учащихся.

Наиболее значимыми с гендерной точки зрения признаками для школьников 10–12 лет являются качества С и Q4 – для девочек, Е и Н – для мальчиков.

Минимальную гендерную нагрузку по результатам тестирования и обработки данных всех обследованных несут признаки F и D.

Личностные качества учащихся, характеризующиеся признаками О, G, I и Q3, по результатам нейросетевой обработки по-разному проявили свою значимость для определения пола ребенка в разных группах и в разные моменты времени.

Так же изменчиво в разных группах и в разное время выглядели результаты влияния на гендерные различия коэффициента интеллекта IQ. В начале пятого класса IQ несет гендерную нагрузку в экспериментальной группе Э и контрольной группе К1 (дети с высокими показателями интеллекта, см. рис. 4).

Проанализируем динамику изменения результатов в экспериментальной группе Э в течение 5-го класса.

В начале учебного года максимальные ошибки в определение пола девочек вносят признаки С и Q4 (как и для всех девочек в целом) и признак G, обнуление которого приводит к 60%-ной ошибке в оценке гендерной принадлежности.

К концу учебного года результаты обработки данных психодиагностики девочек экспериментальной группы принципиально меняются. Все качества личности (включая С и Q4), кроме признака I, слабо влияют на определение пола.

У мальчиков экспериментальной группы в начале пятого класса кроме IQ выраженную факторную нагрузку несет только признак А, что подтверждает связь интеллекта с их общительностью. Этот результат согласуется с данными факторного анализа [Славутская, 2009].

К концу года, по результатам тестирования мальчиков, значимых факторов для их гендерных

особенностей не выявлено.

Таким образом, обработка данных при помощи аппарата искусственных нейронных сетей дает результаты, в целом согласующиеся с данными факторного анализа [Славутская, 2008, 2009, 2011].

Для девочек этого возраста важна эмоциональная стабильность, и для них характерна повышенная фрустрация. Для мальчиков важны факторы «независимость–покорность» и «смелость–робость» в общении.

Если в начале учебного года для девочек важно соблюдение правил поведения (дисциплинированность), то в конце года проявляется их мягкость, потребность во внимании окружающих.

В конце пятого класса вербальный (качество В) и невербальный (IQ) интеллект не оказывают существенного влияния на определение пола ребенка. Этот результат также согласуется с данными, полученными другими средствами: целью и следствием коррекционно-развивающей работы в экспериментальной группе явился тот факт, что интеллектуальные показатели детей стали меньше зависеть от их эмоционально-волевой сферы [Славутская, 2011].

Заключение

Предложенный нейросетевой алгоритм не вызывает традиционных проблем с недостаточным (или избыточным) обучением сети, что подтверждается эффективной работой схемы с разными группами респондентов. Алгоритм позволяет достаточно эффективно выделять наиболее значимые психологические признаки, определяющие гендерные различия испытуемых. При этом результаты могут быть персонифицированы. Предлагаемая схема обработки психодиагностических данных обладает селективностью не только по исходным психологическим признакам, но и по индивидуальным характеристикам испытуемых. С ее помощью может осуществляться, например, разделение респондентов по тому, насколько значимыми для каждого из них являются те или иные психологические качества. Такая процедура может осуществляться по малой выборке данных, в разные моменты времени и, соответственно, описывать временные изменения значимости психологических признаков в динамике (например, в переходные или кризисные для испытуемых периоды). Исключающая неопределенность схема обучения сети и обработки данных в этом случае будет определять высокую достоверность результатов.

Литература

Арзамасцев А.А., Зенкова Н.А. Моделирование в психологии на основе искусственных нейронных сетей. Тамбов: ИМФИ ТГУ им. Г.Р.Державина, 2003.

Боровиков В.П. Нейронные сети. Методология и технологии современного анализа данных / под. ред. В.П.Боровикова. М.: Горячая линия-Телеком, 2008.

Величковский Б.М. Когнитивная наука: Основы психологии познания: в 2 т. М.: Академия, 2006.

Медведев В.С. Нейронные сети. MATLAB 6 / под. ред. В.Г.Потемкина. М.: Диалог-МИФИ, 2002.

Поливанова К.Н. Психология возрастных кризисов: учеб. пособие для студ. высш. пед. учеб. заведений. М.: Академия, 2000.

Розенблатт Ф. Принципы нейродинамики: М.: Мир, 1965.

Славутская Е.В. Факторный анализ психологических качеств, определяющих «дезадаптационный синдром пятого класса» // Психология обучения. 2008. No. 12. С. 103–112.

Славутская Е.В. Экспериментальное изучение гендерных различий в развитии эмоционально-волевых и интеллектуальных свойств младших подростков // Вестник Томского государственного педагогического университета. 2009. No. 9(87). С. 127–129.

Славутская Е.В. Исследование интеллектуальных показателей младших подростков при развивающем воздействии на их эмоционально-волевою сферу // Известия Саратовского университета. Новая серия, Серия: Философия. Психология. Педагогика. 2011. Т. 11, No. 2. С. 76–81.

Хайкин С. [Haykin S.] Нейронные сети: полный курс: пер. с англ. 2-е изд. М.: И.Д.Вильямс, 2006.

Baxt W.G. Complexity, chaos and human physiology: the justification for non-linear neural computational analysis // Cancer Letters. 1994. Vol. 77, No. 2–3. P. 85–93.

Cattell R.B. Advanced in Cattellian Personality Theory. Handbook of Personality. Theory and Research. New York: Guilford Press, 1990.

Dorrer M.G., Gorban A.N., Kopytov A.G. Zenkin V.I. Psychological intuition of neural networks // Proceedings of the World Congress on Neural Networks'95, Washington DC, July 1995. P. 193–196.

Grossberg S. Classical and instrumental learning by neural networks: Progress in theoretical biology. New York: Academic Press, 1974. Vol. 3. P. 51–141.

Hebb D. Organization of behavior: New York: Science Edition, 1961.

Поступила в редакцию 5 октября 2011 г. Дата публикации: 29 июня 2012 г.

Сведения об авторах

Славутская Елена Владимировна. Кандидат психологических наук, доцент (ученое звание), доцент кафедры психологии и социальной педагогики, Чувашский государственный педагогический университет им. И.Я.Яковлева, ул. К.Маркса, д. 38, 428000 Чебоксары, Россия.
E-mail: slavutskayaev@gmail.com

Славутский Леонид Анатольевич. Доктор физико-математических наук, профессор (ученое звание), профессор кафедры управления и информатики, Чувашский государственный университет им. И.Н.Ульянова, Московский проспект, д. 15, 428015 Чебоксары, Россия.
E-mail: las_co@mail.ru

Ссылка для цитирования

Стиль psystudy.ru

Славутская Е.В., Славутский Л.А. Использование искусственных нейронных сетей для анализа гендерных различий младших подростков. Психологические исследования, 2012, 5(23), 4.
<http://psystudy.ru.0421200116/0028>.

ГОСТ 2008

Славутская Е.В., Славутский Л.А. Использование искусственных нейронных сетей для анализа гендерных различий младших подростков // Психологические исследования. 2012. Т. 5, № 23. С. 4.

URL: <http://psystudy.ru> (дата обращения: чч.мм.гггг). 0421200116/0028.

[Последние цифры – номер госрегистрации статьи в Реестре электронных научных изданий ФГУП НТЦ "Информрегистр". Описание соответствует ГОСТ Р 7.0.5-2008 "Библиографическая ссылка". Дата обращения в формате "число-месяц-год = чч.мм.гггг" – дата, когда читатель обращался к документу и он был доступен.]

[К началу страницы >>](#)