

# ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ ИССЛЕДОВАНИЯ ПСИХОЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ СТУДЕНТОВ

© 2020 Е.П. Бобкова<sup>1</sup>, С.В. Зыкин<sup>1,2</sup>, А.Н. Полуянов<sup>2</sup>

<sup>1</sup>Омский государственный технический университет  
(644050 Омск, пр. Мира, д. 11),

<sup>2</sup>Институт математики им. С.Л. Соболева СО РАН  
(630090 Новосибирск, пр. ак. Коптюга, д. 4)

E-mail: lissqa@mail.ru, szykin@mail.ru, andrey.pohuyanov@gmail.com

Поступила в редакцию: 30.07.2020

В связи со сложностью объекта исследования анализ данных в медицине является основным инструментом поиска закономерностей и проверки гипотез. Прежде всего, это относится к психологии, в том числе, к анализу поведения субъектов в тех или иных ситуациях. Для выявления высокотревожного состояния студентов, анализа склонности к депрессии или суициду ежегодно в Омском промышленно-экономическом колледже проводится исследование психоэмоционального состояния студентов. Традиционно для этого используются стандартные тесты, основанные на методике «Шкалы тревоги» Спилбергера—Ханина. Целью данной работы является снижение трудоемкости стандартных тестов. Значительные и слабо мотивированные усилия приходится прилагать студентам при заполнении тестов, затем преподавателям при обработке и анализе тестов. Для решения указанной проблемы предлагается сделать тест компактным за счет применения стандартных и оригинальных методов анализа данных с минимизацией потери точности тестирования. Основным новым результатом данной работы является диагностическая шкала, положенная в основу экспресс-оценки психоэмоционального состояния студентов. Расчет диагностической шкалы был выполнен с использованием графических процессоров на суперкомпьютере ИМ СО РАН. Исследования ориентированы на старшие классы общеобразовательных школ и младшие курсы учебных заведений среднего профессионального образования.

*Ключевые слова:* уровень тревожности, корреляционный анализ, дискриминантный анализ, диагностическая шкала.

## ОБРАЗЕЦ ЦИТИРОВАНИЯ

Бобкова Е.П., Зыкин С.В., Полуянов А.Н. Применение технологий интеллектуального анализа данных для исследования психоэмоционального состояния студентов // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2020. Т. 9, № 3. С. 64–76. DOI: 10.14529/cmse200304.

## Введение

Методы математического анализа используются для обработки и исследования данных с целью обнаружения зависимостей различной природы в исследуемых объектах. Это позволяет, в том числе, построить дискриминацию объектов и впоследствии найти эффективные способы управления проблемными ситуациями. Значительное применение методов анализа имеет место в психологической науке. Одной из актуальных проблем этого направления [1–5] является выявление психоэмоционального состояния молодежи.

Согласно Ч.Д. Спилбергеру [1], персональное беспокойство подразумевает наличие тенденции восприятия довольно широкого спектра жизненных ситуаций как угрозу и реагирование на каждую из них. Так персональное беспокойство становится активным, когда возникают определенные ситуации, интерпретируемые человеком как опасные для самоуважения. Ситуационное или реактивное беспокойство определено Ю.Л. Ханиным как «состояние, которое характеризуется субъективно переживаемыми эмоциями: напряжением, беспокойством, озабоченностью, нервозностью» [2]. Это является эмоциональной реакцией на ситуацию стресса и может различаться интенсивностью во времени.

Лица, принадлежащие группе высокотревожных, могут видеть опасность в большом разнообразии ситуаций и показывать сильно выраженное состояние беспокойства. Если

психологический тест показывает высокий индекс персонального беспокойства, то это дает право допускать, что он отображает состояние беспокойства в различных ситуациях. Особенно когда ситуации имеют отношение к оценке знания, опыта и общественного статуса личности.

Психологи используют ряд методов, чтобы оценивать персональное беспокойство: «лабиринт», «тест анаграммы», «метод Роршаха» и многое другое. Для определения уровня персонального беспокойства Л.М. Костина рекомендует методику «Шкала тревоги» Спилбергера—Ханина, метод экспертных интервью учителей и родителей студентов [3]. Оценивание тревожности дает возможность сделать рекомендации для изменения поведения субъекта. Например, когда оценка показывает высокий уровень тревожности, то рекомендуется мотивировка снижения субъективного значения ситуации и смещение акцентов на понимание конечных целей и средств их достижения. Наоборот, при низком уровне тревожности необходима дополнительная мотивировка к процессу достижения результата.

В психологии используются методы для установления аналитической зависимости параметров исследуемых объектов [6–8]. Наиболее распространены методы, основанные на параметрической статистике. В основе этих методов используется предположение о нормальном распределении значений параметров, например,  $t$ -критерий Стьюдента и  $F$ -критерий Фишера. Указанное предположение не всегда является обоснованным. В этих случаях используются непараметрические методы, например, критерий Манна—Уитни и критерий Вилкоксона.

Статья организована следующим образом. В разделе 1 представлены результаты анализа данных, полученные по традиционной методике расчета «Шкалы тревоги» Спилбергера—Ханина. В разделе 2 рассматривается оригинальная методика расчета неравномерных диагностических шкал. Раздел 3 посвящен анализу данных с целью минимизации исходного множества параметров и построению диагностической шкалы на их основе. В заключении приводятся выводы, полученные в результате проведенных исследований.

## 1. Исходные данные для анализа

В данной работе в качестве обучающей выборки были использованы экспериментальные данные, которые в колледже были классифицированы по «Шкале тревоги» Спилбергера—Ханина. В исследовании приняли участие 149 студентов первого курса. Сокращенный перечень параметров представлен в табл. 1. Всего в базисной таблице присутствует 40 параметров.

Таблица 1

Перечень параметров для эксперимента

№ параметра	Параметр
1	Я спокоен
2	Мне ничто не угрожает
3	Я нахожусь в напряжении
4	Я внутренне скован
5	Я чувствую себя свободно
	...
36	Я бываю доволен
37	Всякие пустяки отвлекают и волнуют меня
38	Бывает, что я чувствую себя неудачником
39	Я уравновешенный человек
40	Меня охватывает беспокойство, когда я думаю о своих делах и заботах

Для поиска новых зависимостей было решено в анализе использовать двенадцать дополнительных параметров, приведенных в табл. 2. Значения параметров 50–52 были получены за счет дополнительного тестирования. Параметрам 41, 43, 45, 48, 49 были поставлены в соответствие целочисленные значения: чем хуже показатель, тем больше значение. Далее на параметры будем ссылаться по их номерам. Каждый параметр исходного теста имеет четыре значения: 1 — «нет, это не так»; 2 — «пожалуй, так»; 3 — «верно»; 4 — «совершенно верно». Наличие всего четырех дискретных значений параметров снижает качество анализа традиционными методами. Непрерывная шкала оценок могла бы дать более качественный результат, в том числе в тесте Спилбергера—Ханина. Есть несколько вариантов электронной версии этого теста в сети Internet [9–11]. Эти приложения позволяют провести тестирование без регистрации на сайтах и обрабатывают данные в онлайн-режиме. У приложений отсутствует способность сбора статистики по группам пользователей.

**Таблица 2**

Дополнительный перечень параметров

№ параметра	Параметр
41	Тип семьи
42	Количество детей в семье
43	Условия проживания в городе Омске
44	Средний балл аттестата при поступлении
45	Место проживания до поступления в колледж
46	Средний балл оценок за первый семестр обучения в колледже
47	Средний балл оценок за второй семестр обучения в колледже
48	Тип общеобразовательного учреждения до поступления
49	Состав группы
50	Склонность к аддиктивному поведению
51	Склонность к делинквентному поведению
52	Склонность к суицидальному поведению

Интерпретация результатов по методике определения тревожности Спилбергера—Ханина позволила разделить испытуемых на 3 группы по уровню выраженности личностной тревожности и на 3 группы по уровню выраженности ситуативной тревожности. Результаты распределения представлены в табл. 3. Группа 1 — испытуемые с низким уровнем тревожности, группа 2 — со средним уровнем тревожности, группа 3 — с высоким уровнем тревожности. Далее в эксперименте использовался результат интегрированной оценки тревожности Спилбергера—Ханина.

**Таблица 3**

Распределение результатов исследования по уровням и видам тревожности

Уровень тревожности	Ситуативная тревожность	Личностная тревожность
Высокий	1	41
Средний	14	97
Низкий	134	11

## 2. Метод расчета диагностической шкалы

Методика предварительной обработки параметров описана в экспериментальной части работы, что необходимо для поиска информативных параметров. Целевым этапом эксперимента является построение диагностической шкалы [12, 13] по следующей методике.

Для расчета шкалы применяется решающая функция [14] в виде линейной комбинации  $N$  значимых параметров:

$$F(x) = a_1x_1 + a_2x_2 + \dots + a_Nx_N,$$

где  $x = (x_1, x_2, \dots, x_N)$  — вектор значений выделенных параметров (координат в пространстве параметров),  $a = (a_1, a_2, \dots, a_N)$  — веса выделенных параметров (коэффициенты).

Для значений  $F(x)$  при расчете определяются границы (оценочная шкала):

$$g_0, g_1, \dots, g_K,$$

где  $K$  — количество групп объектов  $O_1, O_2, \dots, O_K$ . При условии, что  $g_0 < g_1 < \dots < g_K$ , для определения принадлежности произвольного объекта  $o$ , представленного вектором значений параметров  $x'$ , к группе  $O_j$  выполняется проверка следующего неравенства:

$$g_{j-1} < F(x') < g_j.$$

Введем понятие функционала риска  $E$  — суммарного количества ошибок при отнесении объекта к группе. Тогда задача построения оценочной шкалы может быть записана в следующем виде:

$$E \rightarrow \min, g_0 < g_1 < \dots < g_K, -1 \leq a_i \leq 1, i = 1, 2, \dots, N.$$

Рассмотрим в общем виде схему реализации алгоритма построения шкалы [13]:

- 1) задаем начальное значение вектора весов  $a^0 = (a^0_1, a^0_2, \dots, a^0_N)$ ;
- 2) методом покоординатного спуска с заданным шагом перебираем возможные значения вектора весов  $a = (a_1, a_2, \dots, a_N)$ , выполняя на каждом шаге пункты 3–6;
- 3) расчет значений решающей функции  $F_{ij} = F(x_{ij})$  по всем объектам для текущего набора весов;
- 4) сортировка  $F_{ij}$  по возрастанию;
- 5)  $g_0 = \min(F_{ij}) - \varepsilon$ ,  $g_K = \max(F_{ij}) + \varepsilon$ , где  $\varepsilon$  — малая величина;
- 6) подбор  $g_1, \dots, g_{K-1}$  с учетом выполнения неравенства  $g_0 < g_1 < \dots < g_K$  и минимизации функционала риска.

Для ускорения работы алгоритма используется технология CUDA [13], предназначенная для разработки приложений, исполняемых на графических процессорах (GPU). Программная реализация алгоритма выполнена на языке C++ с использованием технологии CUDA. Код приложения разбивается на последовательную и параллельные части. Последовательная часть выполняется на CPU, а параллельная часть, оформленная в виде функции ядра (kernel function), выполняется GPU.

В ходе анализа последовательного варианта алгоритма построения диагностической шкалы выявлено, что максимальное время затрачивается на выполнение шага 6 — 96% от общего времени выполнения программы, следовательно, согласно закону Амдаля, максимально возможный прирост производительности при распараллеливании вычислений на данном шаге — 25-кратный.

Функция ядра для шага 6 алгоритма (выполняемая на GPU) на входе получает массив со значениями  $F_{ij}$  для каждого объекта выборки. Каждая нить GPU (аналог потока на CPU) рассчитывает значение функционала риска  $E$  для определенного набора границ  $g_0, \dots, g_K$  (набор зависит от номера нити). Полученные результаты записываются в массив в памяти GPU, массив передается из памяти GPU в память CPU, далее CPU определяет минимальное значение функционала риска для текущего набора весов и выполняется следующий шаг алгоритма. При использовании параллельного алгоритма получено в среднем десятикратное ускорение вычислений по сравнению с последовательным алгоритмом.

### 3. Результаты анализа данных

Основным показателем будем считать количество ошибок, поскольку он лежит в основе большинства методов анализа данных. Для организации экспресс-тестирования необходимо выделить наиболее значимые параметры в предложенной модели и выявить характер их взаимодействия между собой и с другими параметрами. Оригинальный метод ранжирования параметров применялся Н.Г. Загоруйко [15, 16]. Параметры обучающей выборки предварительно ранжируются по значимости: вычисляется Евклидова мера для всех данных, затем аналогичная мера для данных без одного параметра. В итоге, чем больше убывание меры, тем существеннее влияние параметра, хотя характер этого влияния пока не известен. Результаты расчета в 52-мерном пространстве приведены в табл. 4, в которой указаны параметры с наибольшим убыванием меры.

**Таблица 4**

Результаты расчетов по Евклидовой мере

Параметр	Значение меры
49	8,221
48	8,355
25	8,373
29	8,387
31	8,437
10	8,471
7	8,478
45	8,484
32	8,485

Альтернативой Евклидовой мере при определении значимости параметра является использование методов дискриминантного анализа. В его основе лежат методы решения задач «различения (дискриминации) объектов наблюдения по определенным признакам» [7, 17]. На данном этапе был использован одномерный анализ в предположении, что области значений отдельного параметра могут пересекаться для отдельных групп субъектов. Другими словами, часть ответов на один и тот же вопрос могут совпадать для различных психоэмоциональных состояний студентов. Задачей реализованного нами алгоритма является минимизация этой области. Итоговый список значимых параметров представлен в табл. 5.

**Таблица 5**

Итоговый список параметров при дискриминации

Параметр	Суммарное количество ошибок
29	13
31	16
23	21
32	21
25	22
28	30
37	31
40	31
35	33
38	35
11	36
10	37

На следующем этапе необходимо выяснить наличие и характер зависимости между параметрами. Зависимости могут ухудшить результаты моделирования за счет доминирования таких параметров. Простейшую зависимость, близкую к линейной, можно выяснить за счет вычисления корреляции [18]. В [7] рассматриваются две основные задачи корреляционного анализа: а) определение формы связи в виде математической формулы; б) измерение меры связи между признаками для установления степени влияния данного параметра на результат за счет решения корреляционного уравнения.

В данной работе влияние параметров на результат исследования определяется за счет построения диагностической шкалы, для которой целесообразно отсутствие коррелированных параметров. В табл. 6 представлены коэффициенты корреляции для значимых параметров.

Согласно табл. 6, значимый коэффициент корреляции (0,74) между параметрами 31 и 29. Относительно этих параметров необходимо дополнительное исследование их влияния на шкалу, если они и далее будут находиться в списке значимых параметров.

Заметим, что отсутствие корреляционной зависимости не гарантирует отсутствие иной зависимости между параметрами. Эта зависимость может иметь нелинейный характер. Например, значения параметра  $x$  принадлежат множеству  $\{1, 2, \dots, 11\}$ , а параметр  $y = (x - 6)^2$ . Между параметрами  $x$  и  $y$  квадратичная зависимость, однако корреляция между ними равна нулю.

Традиционно для построения многомерной дискриминантной функции используется кусочно-линейная аппроксимация [19, 20]. Нами разработан новый метод построения нелинейной дискриминантной функции. Основная идея метода заключается в построении функции в  $(n+1)$ -мерном пространстве, где  $n$  количество параметров. Функция аппроксимирует положительные значения с весами для первой группы объектов и отрицательные значения с весами для другой группы объектов. Граница дискриминации образуется пересечением аппроксимации с пространством параметров. Для выбора весов используется метод штрафных функций с целью получения минимального количества ошибок дискри-

Таблица 6

Корреляция параметров

Параметры	10	11	23	25	28	29	31	32	35	37	38	40
10	1											
11	0,49	1										
23	-0,34	-0,3	1									
25	-0,29	-0,2	0,65	1								
28	-0,4	-0,4	0,47	0,48	1							
29	-0,29	-0,3	0,66	0,65	0,62	1						
31	-0,28	-0,3	0,66	0,65	0,5	0,74	1					
32	-0,44	-0,5	0,44	0,47	0,45	0,52	0,48	1				
35	-0,34	-0,4	0,46	0,42	0,4	0,39	0,37	0,47	1			
37	-0,19	-0,3	0,52	0,5	0,47	0,6	0,51	0,35	0,47	1		
38	-0,38	-0,5	0,44	0,45	0,46	0,46	0,46	0,57	0,5	0,4	1	
40	-0,22	-0,2	0,32	0,35	0,42	0,51	0,42	0,28	0,33	0,41	0,31	1



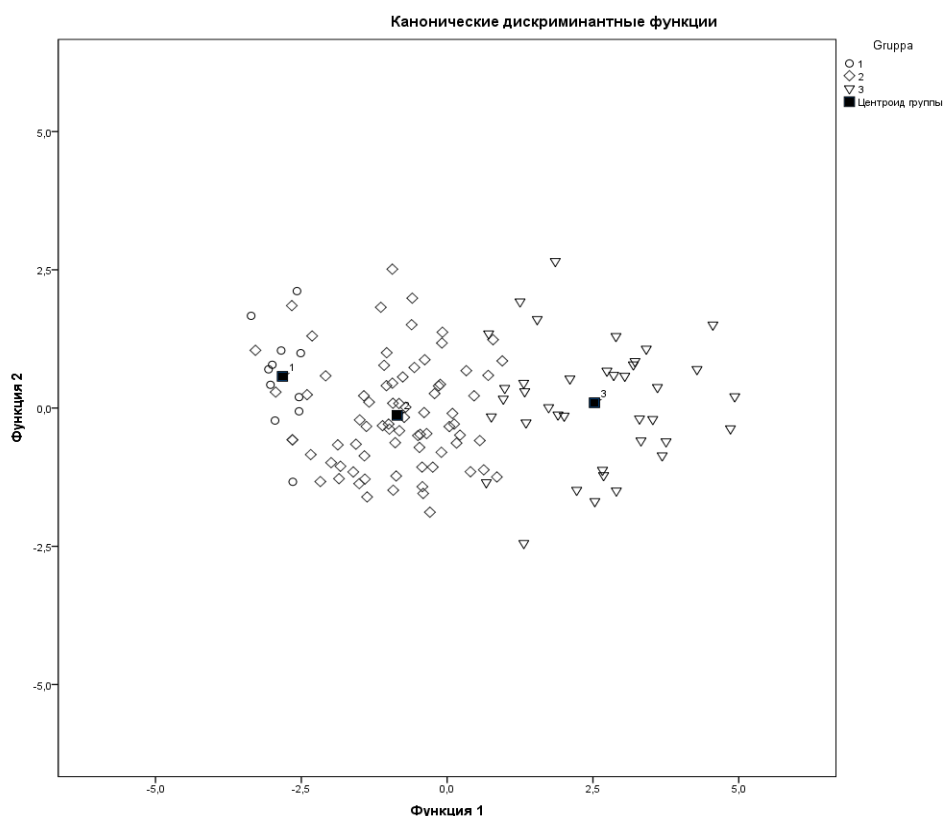


Рис. 2. Результаты анализа с использованием пакета SPSS Statistics

Заключительным этапом эксперимента явилось построение диагностической шкалы [12, 14], адаптированной для расчетов на графических процессорах [13]. Идея и реализация метода описаны в разделе 2 работы. Итоговый список параметров, прошедших предварительный отбор, представлен в табл. 7. В итоговом результате вычислительного эксперимента (см. рис. 3) осталось только 4 значимых параметров для диагностики: 29, 31, 32, 38. При этом достигнут уровень ошибок — 4.

Таблица 7

Сводный список параметров для построения шкалы

№ параметра	Параметр
10	Я испытываю чувство внутреннего удовлетворения
11	Я уверен в себе
23	Я легко расстраиваюсь
25	Я сильно переживаю неприятности и долго не могу о них забыть
28	Меня тревожат возможные трудности
29	Я слишком переживаю из-за пустяков
31	Я все принимаю близко к сердцу
32	Мне не хватает уверенности в себе
37	Всякие пустяки отвлекают и волнуют меня
38	Бывает, что я чувствую себя неудачником
40	Меня охватывает беспокойство, когда я думаю о своих делах и заботах
44	Средний балл аттестата при поступлении
47	Средний балл за 2 семестр в колледже



```

Version 4
<<<<<<<check result>>>>>>>>>>>>
-----
ВЕСА ПАРАМЕТРОВ:
№1 10 0
№2 11 0
№3 23 0
№4 25 0
№5 28 0
№6 29 0.1
№7 31 0.1
№8 32 0.1
№9 37 0
№10 38 0.1
№11 40 0
№12 44 0
№13 47 0

ГРАНИЦЫ ГРУПП:
0.3 (Низкая) 0.4 (Средняя) 0.9 (Высокая) 1.6

КОЛИЧЕСТВО ОШИБОК: 4
-----
FINAL_RESULT. RASCHET OKONCHEN.
Time clocked by CUDA events: milliseconds
1.13183999061584
Time clocked by clock() CPU milliseconds
0
    
```

**Рис. 3.** Результаты расчета шкалы психоэмоционального состояния

Пусть  $x_1$  — значение параметра 29,  $x_2$  — 31,  $x_3$  — 32,  $x_4$  — 38, подставим значения этих параметров в выражение:

$$y = 0,1 \cdot x_1 + 0,1 \cdot x_2 + 0,1 \cdot x_3 + 0,1 \cdot x_4.$$

Если  $0,3 \leq y \leq 0,4$ , то тестируемый принадлежит к группе с низким уровнем тревожности,  $0,4 \leq y \leq 0,9$  — средняя тревожность, если  $0,9 \leq y \leq 1,6$  — высокая тревожность. При попадании на границу интервала необходимо провести полное тестирование по исходной шкале. Ранее было получено, что параметры 29 и 31 имеют значительный уровень корреляции. Расчет шкалы без параметра 29 дал 6 ошибок, без параметра 31 — 5 ошибок, что дает основание не учитывать корреляцию.

Полученные результаты позволяют рекомендовать применение рассмотренного метода для построения диагностических шкал. Окончательный выбор применяемого метода остается за экспертом, аргументом в пользу выбора того или иного метода является количество ошибок на обучающей выборке и количество параметров, используемых для построения шкалы.

## Заключение

Диагностическая шкала позволила свести до четырех количество ошибочного диагностирования на обучающих выборках, в то время как в пакете SPSS Statistics количество ошибок на тех же данных — 14 при использовании полного набора параметров. В рассмотренном материале не использовались какие-либо статистические гипотезы, поскольку это не влияет на конечный результат.

Применение рассмотренной «Шкалы экспресс-оценки» психоэмоционального состояния позволило ускорить сбор и обработку исходных данных, поскольку при ответе на четыре вопроса у студентов меньше мотивации отложить тестирование либо исказить его результаты. По предложенной методике был протестирован 501 студент, среди которых выявлено 15 высокотревожных студентов.

Рассмотренный материал предполагает дальнейшее развитие методики тестирования. Прежде всего, необходимо отказаться от дискретных значений параметров, привлечь дополнительную информацию из других тестов для повышения достоверности ответов, разработать адаптивное тестирование [22], собрать и обработать статистику по прецедентам.

*Работа выполнена при поддержке программы фундаментальных научных исследований СО РАН № 1.5.1., проект № 0314-2019-0020.*

## Литература

1. Spielberger, C.D. Assessment of state and trait anxiety: Conceptual and methodological issues // *Southern Psychologist*. 1985. Vol. 2, no 4. P. 6–16.
2. Ханин Ю.Л. Краткое руководство к применению шкалы реактивной личностной тревожности Ч.Д. Спилбергера. Ленинград: ЛНИИФК, 1976. 18 с.
3. Костина Л.М. Методы диагностики тревожности. СПб.: Изд-во «Речь», 2005. 198 с.
4. Прихожан А.М. Тревожность у детей и подростков: психологическая природа и возрастная динамика. Воронеж: Изд-во НПО «МОДЭК», 2000. 304 с.
5. Карелин А.П. Большая энциклопедия психологических тестов. М.: Изд-во «Проспект», 2015. 420 с.
6. Титкова Л.С. Математические методы в психологии. Владивосток: Изд-во Дальневосточного университета, 2014. 140 с.
7. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: Изд-во «Речь», 2015. 350 с.
8. Герасевич В.А., Аветисов А.Р. Современное программное обеспечение для статистической обработки биомедицинских исследований // *Белорусский медицинский журнал*. 2015. № 1. С. 12–25.
9. Тест Спилбергера—Ханина. URL: <https://psyttests.org/psystate/spielberger-run.html> (дата обращения: 27.07.2020).
10. Тест исследование тревожности (опросник Спилберга). URL: <https://onlinetestpad.com/ru/test/714-test-issledovanie-trevozhnosti-oprosnik-spilberga> (дата обращения: 27.07.2020).
11. Тест на уровень тревожности Спилбергера—Ханина. URL: [https://psychojournal.ru/tests\\_online/127-test-na-uroven-trevozhnosti-spielbergera-hanina.html](https://psychojournal.ru/tests_online/127-test-na-uroven-trevozhnosti-spielbergera-hanina.html) (дата обращения: 27.07.2020).
12. Зыкин С.В., Редреев П.Г., Чернышев А.К. Формирование представлений данных для построения медицинских диагностических шкал // *Омский научный вестник*. 2011. № 2(100). С. 190–193.
13. Полуянов А.Н. Расчет диагностической шкалы на графических процессорах // *Омский научный вестник*. 2012. № 3(113). С. 277–279.
14. Лбов Г.С., Неделько В.М., Неделько С.В. Метод адаптивного поиска логической решающей функции // *Сибирский журнал индустриальной математики*. 2009. Т. XII, № 3(39). С. 66–74.
15. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. Новосибирск: Изд-во Института математики СО РАН, 1999. 270 с.
16. Загоруйко Н.Г., Кутненко О.А., Борисова И.А., Дюбанов В.В., Леванов Д.А., Зырянов О.А. Выбор информативных признаков для диагностики заболеваний по генетическим данным // *Вавиловский журнал генетики и селекции*. 2014. Т. 18, № 4–2. С. 898–903.
17. Дюк В.А. Компьютерная психодиагностика. СПб.: Изд-во «Братство», 1994. 364 с.
18. Буров А.В. Корреляционный анализ взаимосвязей между индивидуально-психологическими особенностями детей и показателями адаптации к школьному обучению // *Научно-методический электронный журнал «Концепт»*. 2014. № 7. С. 101–105. URL: <http://e-koncept.ru/2014/14193.htm> (дата обращения: 27.07.2020).
19. Glen J.J. Mathematical Programming Models for Piecewise-Linear Discriminant Analysis // *The Journal of the Operational Research Society*. 2005. Vol. 56, no. 3. P. 331–341. DOI: 10.1057/palgrave.jors.2601818.

20. Chang D.S., Kuo Y.C. An approach for the two-group discriminant analysis: An application of DEA // *Mathematical and Computer Modelling*. 2008. Vol. 47, no. 9–10. P. 970–981. DOI: 10.1016/j.mcm.2007.05.010.
21. Дубнов П.Ю. Обработка статистической информации с помощью SPSS. М.: Изд-во «АСТ», 2004. 221 с.
22. Томашев М.В., Авдеев А.С., Краснова М.В. Адаптивное тестирование как средство управления качеством образования // *Информатика и образование*. 2018. № 9. С. 27–33. DOI: 10.32517/0234-0453-2018-33-9-27-33.

Бобкова Елена Петровна, магистрант кафедры ПМиФИ Омского государственного технического университета, преподаватель Омского промышленно-экономического колледжа (Омск, Российская Федерация)

Зыкин Сергей Владимирович, д.т.н., профессор, заведующий лабораторией МППИ Института математики им. С.Л. Соболева СО РАН (Новосибирск, Российская Федерация), профессор кафедры ПМиФИ Омского государственного технического университета (Омск, Российская Федерация)

Полуянов Андрей Николаевич, к.т.н., старший научный сотрудник лаборатории МППИ Института математики им. С.Л. Соболева СО РАН (Новосибирск, Российская Федерация)

---

DOI: 10.14529/cmse200304

## APPLICATION OF INTELLIGENT DATA ANALYSIS TECHNOLOGIES FOR STUDENTS PSYCHO-EMOTIONAL STATE STUDY

© 2020 E.P. Bobkova<sup>1</sup>, S.V. Zykin<sup>1,2</sup>, A.N. Poluyanov<sup>2</sup>

<sup>1</sup>*Omsk State Technical University (pr. Mira 11, Omsk, 644050 Russia),*

<sup>2</sup>*Sobolev Institute of Mathematics SB RAS*

*(pr. Acad. Koptyug 4, Novosibirsk, 630090 Russia)*

*E-mail: lissqa@mail.ru, szykin@mail.ru, andrey.poluyanov@gmail.com*

Received: 30.07.2020

Due to the complexity of the research object, data analysis in medicine is the main tool for finding patterns and testing hypotheses. First of all, this applies to psychology, including the analysis of the behavior of subjects in certain situations. In order to identify the high-anxiety state of students, to analyze the tendency to depression or suicide, a study of the psycho-emotional state of students is conducted annually at the Omsk Industrial and Economic College. Traditionally, standard tests based on the technique of “Anxiety Scale” of Spielberger–Hanin's are used for this. The purpose of this work is to reduce the complexity of the standard tests. Significant and poorly motivated efforts have to be made by students in completing the tests, and then by teachers in processing and analyzing the tests. To solve this problem, it is proposed to make the test compact by applying standard and original data analysis methods while minimizing the loss of test accuracy. The main result of this work is a diagnostic scale, which forms the basis for the rapid assessment of the psycho-emotional state of students. The diagnostic scale was calculated using graphics processors on a supercomputer of IM SB RAS. Target audience: senior classes of secondary schools and junior courses of educational institutions of secondary vocational education.

*Keywords: anxiety level, correlation analysis, discriminant analysis, diagnostic scale.*

### FOR CITATION

Bobkova E.P., Zykin S.V., Poluyanov A.N. Application of Intelligent Data Analysis Technologies for Students Psycho-emotional State Study. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*. 2020. Vol. 9, no. 3. P. 64–76. (in Russian) DOI: 10.14529/cmse200304.

*This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 3.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.*

## References

1. Spielberger C.D. Assessment of State and Trait Anxiety: Conceptual and Methodological Issues. *Southern Psychologist*. 1985. Vol. 2, no 4. P. 6–16.
2. Khanin Yu.L. A Brief Guide to the Use of the Scale of Reactive Personal Anxiety by Ch.D. Spielberger. Leningrad, LNIIFK, 1976. 18 p. (in Russian)
3. Kostina L.M. Methods for Diagnosing Anxiety. St. Petersburg, Rech, 2005. 198 p. (in Russian)
4. Prikhozhan A.M. Anxiety in Children and Adolescents: the Psychological Nature And Age Dynamics. Voronezh, NPO “MODEK”, 2000. 304 p. (in Russian)
5. Karelin A.P. Great Encyclopedia of Psychological Tests. Moscow, Prospect, 2015. 420 p. (in Russian)
6. Titkova L.S. Mathematical Methods in Psychology. Vladivostok, Publishing of the Far Eastern State University, 2014. 140 p. (in Russian)
7. Sidorenko Ye.V. Methods of Mathematical Processing in Psychology. St. Petersburg, Rech, 2015. 350 p. (in Russian)
8. Gerasevich V.A., Avetisov A.R. Modern Software for Statistical Processing of Biomedical Research. *Belarusian Medical Journal*. 2015. no. 1. P. 12–25. (in Russian)
9. Spielberger–Khanin Test. Available at: <https://psyttests.org/psystate/spielberger-run.html> (accessed: 27.07.2020). (in Russian)
10. Test Anxiety Study (Spielberg questionnaire). Available at: <https://onlinetestpad.com/ru/test/714-test-issledovanie-trevozhnosti-oprosnik-spilberga> (accessed: 27.07.2020). (in Russian)
11. Test on The Level of Anxiety Spielberger–Hanina. Available at: [https://psychojournal.ru/tests\\_online/127-test-na-uroven-trevozhnosti-spilbergera-hanina.html](https://psychojournal.ru/tests_online/127-test-na-uroven-trevozhnosti-spilbergera-hanina.html) (accessed: 27.07.2020). (in Russian)
12. Zykina S.V., Redreyev P.G., Chernyshev A.K. Formation of Data Representations for the Construction of Medical Diagnostic Scales. *Omsk Scientific Bulletin*. 2011. no. 2(100). P. 190–193. (in Russian)
13. Poluyanov A.N. Calculation of the Diagnostic Scale on Graphics Processors. *Omsk Scientific Bulletin*. 2012. no. 3(113). P. 277–279. (in Russian)
14. Lbov G.S., Nedel'ko V.M., Nedel'ko S.V. The Method of Adaptive Search for a Logical Solving Function. *Journal of Applied and Industrial Mathematics*. 2009. Vol. 12, no. 3(39). P. 66–74. (in Russian)
15. Zagoruyko N.G. Applied Methods of Data and Knowledge Analysis. Novosibirsk, Sobolev Institute of Mathematics SB RAS, 1999. 270 p. (in Russian)
16. Zagoruyko N.G., Kutnenko O.A., Borisova I.A., Dyubanov V.V., Levanov D.A., Zyryanov O.A. The Choice of Informative Features for the Diagnosis of Diseases According to Genetic Data. *Russian Journal of Genetics: Applied Research*. 2014. Vol. 18, no. 4–2. P. 898–903. (in Russian)
17. Dyuk V.A. Computer Psychodiagnostics. St. Petersburg, Bratstvo, 1994. 364 p. (in Russian)
18. Burov A.V. Correlation Analysis of the Relationship Between the Individual Psychological Characteristics of Children and Indicators of Adaptation to School Education. Scientific-

- Methodical Electronic Journal “Koncept”. 2014. no. 7. P. 101–105. Available at: <http://e-koncept.ru/2014/14193.htm> (accessed: 27.07.2020). (in Russian)
19. Glen J.J. Mathematical Programming Models for Piecewise-Linear Discriminant Analysis. The Journal of the Operational Research Society. 2005. Vol. 56, no. 3. P. 331–341. DOI: 10.1057/palgrave.jors.2601818.
  20. Chang D.S., Kuo Y.C. An approach for the two-group discriminant analysis: An application of DEA. Mathematical and Computer Modelling. 2008. Vol. 47, no. 9–10. P. 970–981. DOI: 10.1016/j.mcm.2007.05.010.
  21. Dubnov P.Yu. Statistical information processing using SPSS. Moscow, ACT NT Press, 2004. 221 p. (in Russian)
  22. Tomashev M.V., Avdeyev A.S., Krasnova M.V. Adaptive Testing as a Tool for Managing Quality of Education. Informatics and Education. 2018. no. 9. P. 27–33. (in Russian) DOI: 10.32517/0234-0453-2018-33-9-27-33.