

Ахметова С. Т., Ыскак Назира Пернебек кызы

МАТЕМАТИЧЕСКИЕ МОДЕЛИ И АЛГОРИТМЫ ПСИХОДИАГНОСТИКИ

S.T. Akhmetova, YskcikNazir Pernebek kuzu

MA THEMATIC MODELS AND ALGORITHMS PSYCHODIAGNOSTICS

УДК: 535/341(548-04)

В этой статье рассматривается серия экспериментов, направленных на проверку гипотезы о применимости нейронных сетей к задачам психодиагностики.

Проводится исследование возможности применимости нейронных сетей, как аппарат психодиагностики - при определении и оптимизации структуры психологических тестов.

Исследуется влияние структуры психологических тестов на диагностическую интуицию искусственной нейронной сети.

This article discusses a series of experiments designed to test hypotheses about the applicability of neural networks to problems in psycho-diagnostics.

We study the possible applicability of neural networks as a tool of psycho - in defining and optimizing the structure of psychological tests.

The influence of the structure of psychological tests for diagnostic intuition of an artificial neural network.

Психодиагностика (от греч ψυχή — душа, и греч. διαγνωστικός — способный распознавать) — отрасль психологии, разрабатывающая теорию, принципы и инструменты оценки и измерения индивидуально-психологических особенностей личности [1].

В XX веке психодиагностика и ее приложения интенсивно разрабатывались, и в настоящее время образуют разветвленную структуру методов и методик. Средства современной психодиагностики разделяются на две группы: строго формализованные методики, и методики мало формализованные [2].

К строго формализованным методикам относятся тесты, опросники, некоторые методики проективной техники и психофизиологические методики. Для строго формализованных методик характерны детальная регламентация, стандартизация (установление единообразия проведения обработки и представления результатов диагностических экспериментов), объективизация процедуры обследования или испытания (точное соблюдение инструкций, строго определенные способы предъявления стимульного материала, невмешательство исследователя в деятельность испытуемого и др.), надежность и валидность. Многие строго формализованные методики доведены до уровня компьютерной реализации [3].

Мало формализованные методики — это наблюдения, беседы и интервью, анализ продуктов деятельности. Они дают ценные сведения об испытуемом, особенно когда предметом изучения выступают такие психические процессы и явления, которые мало поддаются объективизации и формализации (например, плохо осознаваемые субъективные переживания, личностные смыслы) или являются чрезвычайно изменчивыми по содержанию (динамика целей, состояний, настроений и т. д.). Мало- формализованные методики очень трудоемки (например, наблюдения за обследуемым осуществляются иногда в течение нескольких месяцев) и требуют большого профессионального мастерства и опыта псих диагноста.

Мало формализованные и строго формализованные диагностические методики дополняют друг друга и должны использоваться в комплексе. Развитие новых интеллектуальных методов анализа данных постепенно расширяет область применения компьютеров в психодиагностике. Новые технологии позволяют использовать имитацию психологической интуиции компьютерными системами.

В работе исследователя по конструированию психодиагностического теста принято выделять три этапа.

На первом этапе конструируется «черновой» вариант теста. В него включаются задания, ответы на которые, по мнению экспериментатора, должны отражать индивидуально-психологические различия испытуемых по данному конструкту.

На втором этапе исследователь выбирает диагностическую модель и определяет ее параметры. Под диагностической моделью понимается способ компоновки (преобразования, агрегирования) исходных диагностических признаков (вариантов ответов на задания теста) в диагностический показатель.

На третьем этапе проводится стандартизация и испытание построенной диагностической модели.

Наиболее употребляемой в психодиагностике является линейная диагностическая модель. Без применения эмпирико-статистического анализа не обходится ни одна серьезная попытка конструирования или адаптации тестов [4]. Исходным материалом для такого анализа служат результаты экспериментального обследования репрезентативной выборки испытуемых с помощью «чернового» варианта психодиагностического теста. Из полученных данных формируется таблица экспериментальных данных (см. табл. 1).

Структура таблицы экспериментальных данных

Объекты (испытуемые)	Исходные признаки					
	x1	x2	...	xi	...	xp
X1	x11	x12	...	x1j	...	x1p
...
Xi	xi1	xi2	...	xij	...	xip
...
XN	xN1	xN2	...	xNi	...	xNp

В табл. ... N - общее количество объектов (испытуемых), p - общее количество признаков, xj - j-й признак, xij - значение j-го признака, измеренное у i-го объекта, X=(x1, ... , xp)T - вектор признаков, Xi=(xi1, ... , xip)T - i-й объект, X={ Xi} - множество объектов.

Исходные признаки xj, как правило, измерены в номинальных и порядковых (ординальных) шкалах [5]. Для большинства объективных методик нельзя априорно установить ни количественных отношений ни отношений порядка, поскольку их признаки представляют собой номинальные измерения. Зачастую при формализации тестовых методик применяют «дихотомизацию» - процедуру преобразования исходных показателей в набор признаков с двумя градациями.

Для ординальных признаков существенен лишь порядок градаций на шкале, и для них считаются допустимыми любые монотонные преобразования не нарушающие этот порядок. Методически строгим является применение к ординальным признакам методов обработки, результат которых инвариантен относительно допустимых преобразований порядковой шкалы.

Далее, после формирования таблицы экспериментальных данных, производится построение диагностической модели. Считается, что модель должна в определенной форме выражать зависимость между вектором входных признаков и тестируемым свойством (значение выраженности свойства далее будет обозначаться y). Модель должна отражать механизм преобразования y=y(x).

Предварительным этапом в построении диагностических моделей является как правило выяснение структуры таблицы экспериментальных данных. На этом этапе производится оценка корреляции между факторами и близости между объектами. Набор математических моделей и алгоритмов, используемых для этого, определяется исходя из специфики экспериментальных данных в психодиагностике.

Для определения степени связи между признаками используются [6]:

-Кoeffициент корреляции Пирсона, являющийся мерой линейной связи двух переменных:

$$r_{kj} = \frac{s_{kj}}{\sqrt{s_{kk}s_{jj}}}$$

где

$$s_{kj} = \frac{1}{N-1} \sum_{i=1}^N (x_{ik} - m_k)(x_{ij} - m_j)$$

и

$$m_j = \frac{1}{N} \sum_{i=1}^N x_{ij}$$

В рамках этого же подхода сконструированы коэффициент ранговой корреляции Спирмена, точечный бисериальный коэффициент корреляции и тетракорический коэффициент корреляции.

- Кoeffициент φ , предназначенный для измерения связи двух дихотомических признаков. Кoeffициент вычисляется на базе таблиц сопряженности признаков (см. табл. 2) по формуле

$$\varphi = \frac{bc - ad}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

Таблица 2

Таблица сопряженности дихотомических признаков

Признак X_i	Признак X_j		Итого
	1	0	
1	a	b	a+b
0	c	d	c+d
Итого	a+c	b+d	

- Кoeffициент ранговой корреляции «тау» Кендалла, основанный на подсчете числа несовпадений в ранжировке объектов по сопоставляемым переменным. Данный коэффициент разработан исходя из задачи истолкования процесса измерения связи между переменными без помощи принципа произведения моментов. Рассматриваются

два признака X_i и X_j , на каждый из которых N объектов отображаются в N последовательных

рангов. Из N объектов формируется $\frac{N(N-1)}{2}$ пар.

Тогда коэффициент вычисляется по формуле $\tau = \frac{P-Q}{N(N-1)/2}$, где P - количество совпадений

порядка на признаке X_i с порядком на признаке X_j , Q - количество несовпадений.

Степень связи между признаками может быть использована для оценки избыточности набора признаков «черновой» модели, для взаимоконтроля шкал и т.п.

Для определения близости объектов используются различные меры расстояния:

* Евклидово расстояние

$$d_{ij}^{(E)} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

* Взвешенное евклидово расстояние

$$d_{ij}^{(wE)} = \sqrt{\sum_{k=1}^p w_k (x_{ik} - x_{jk})^2}$$

* Расстояние Махаланобиса

$$d_{ij}^{(M)} = \sqrt{(X_i - X_j)S^{-1}(X_i - X_j)}$$

где S – ковариационная матрица генеральной совокупности, из которой извлечены объекты X_i и X_j .

* Расстояние Минковского

$$d_{ij}^{(M)} = \sum_{k=1}^p I_k(X_i, X_k)$$

(городская метрика), применяющаяся для измерения расстояния между объектами, описанными ординальными признаками. $I_k(X_i, X_k)$ равно разнице номеров градаций по k -му признаку у сравниваемых объектов X_i и X_j .

* Расстояние Хэмминга

$$d_{ij}^{(H)} = \sum_{k=1}^p |x_{ik} - x_{jk}|$$

, которое используется для определения различий между объектами, задаваемыми дихотомическими признаками и интерпретируется как число несовпадений значений признаков у рассматриваемых объектов X_i и X_j .

Полученная на основе какой-либо метрики информация о степени близости объектов может быть использована для выделения их группировок.

Представление информации о структуре экспериментальных данных служит промежуточным звеном в построении диагностической модели. Независимо от типа модели ее создание может опираться на два подхода:

1. Стратегия, основанная на автоинформативности экспериментальных данных. Высокая степень близости между группой признаков может свидетельствовать о том, что признаки, вошедшие в группу, отражают эмпирический фактор, соответствующий диагностическому конструкту. Выделение геометрических группировок в пространстве объектов может свидетельствовать о различии изучаемых объектов по тестируемому свойству, что

позволяет строить диагностический алгоритм. Для стратегий, основанных на автоинформативности экспериментальных данных, важной категорией является согласованность заданий теста. Согласованность измеряемых реакций испытуемых на тестовые стимулы означает, что они должны иметь статистическую направленность на выражение общей, главной тенденции теста.

На стратегии, основанной на автоинформативности экспериментальных данных, строятся конструирование диагностического алгоритма при помощи метода главных компонент, факторного анализа и метода контрастных групп.

2. Стратегия, основанная на критериях внешней информативности. Внешняя информация может быть представлена в виде привязки к объектам значений «зависимой» переменной, измеренной в количественной шкале, в виде номера однородного по тестируемому свойству класса, в виде порядкового номера (ранга) объекта в ряду всех объектов, упорядоченных по степени проявления диагностируемого свойства или в виде совокупности значений набора внешних (не включенных в таблицу экспериментальных данных) признаков, характеризующих тестируемый психологический феномен.

Методы, основанные на внешней информативности признаков принято подразделять на экспертные, экспериментальные и жизненные. К числу экспертных критериев относят оценки, суждения, заключения об испытуемых, вынесенные одним экспертом или их группой.

Экспериментальными критериями служат результаты одновременного и независимого исследования испытуемого другим тестом, который считается апробированным и измеряющим то же свойство, что и конструируемый тест. В качестве жизненных критериев используются объективные социально - демографические и биографические данные. На стратегии, основанной на внешней информативности экспериментальных данных, строятся конструирование диагностического алгоритма при помощи регрессионного анализа, дискриминантного анализа и типологического подхода.

Наиболее широко в настоящее время употребляются линейные диагностические модели. Однако в условиях неоднородности обучающей выборки они обладают практической успешностью не выше 70-80%.

Таким образом, можно сказать что построенная диагностическая модель может считаться психодиагностическим тестом только после прохождения ею испытаний на предмет проверки психометрических свойств - надежности и валидности.

Надежность теста - характеристика методики, отражающая точность психодиагностических измерений, а также устойчивость результатов теста к воздействию посторонних случайных факторов.

Валидность - мера соответствия тестовых оценок представлениям о сущности свойств или их роли в той или иной деятельности.

Литература:

1. Бурлачук Л. Ф. Психодиагностика: Учебник для вузов. - СПб: Питер, 2006. -351 с: ил. - (Серия «Учебник нового века»), с. 104
2. Титкова Л. С., Психодиагностика: Учебное пособие. - Издательство Дальневосточного университета, Владивосток, 2002.
3. Дюк В. А. Компьютерная психодиагностика. - СПб: «Братство», 1994-364 с.
4. Шмелев А.Г., Похилько В.И. Анализ пунктов при конструировании и применении тест - опросников: ручные и компьютерные алгоритмы // Вопросы психологии – 1985 №4 - с. 126-134.
5. Пфанцгаль И. Теория измерений. - М.: Мир, 1976 с. 248.
6. Миркин Б.Г. Анализ качественных признаков и структур. - М.: Статистика, 1980. - с. 319.

Рецензент: д.т.н., профессор Ескендилов Ш.З.
