

УДК 681.327.12.001.362

© 2008 г. **Н.С. Безруков,**  
**Е.В. Колосова**

(Амурский государственный университет, Благовещенск)

## **СПОСОБЫ РЕГИОНАЛЬНОЙ КЛАСТЕРИЗАЦИИ ПО ПАРАМЕТРАМ ЧЕЛОВЕЧЕСКОГО КАПИТАЛА НА ОСНОВЕ САМООБУЧАЮЩИХСЯ НЕЙРОННЫХ СЕТЕЙ**

Рассматривается применение нейронных сетей в региональной кластеризации по параметрам человеческого капитала. Предложенный вариант классификации расчета может быть использован для совершенствования целевых программ регионального развития.

### **Введение**

Применение нейронных сетей в различных отраслях человеческой деятельности позволяет строить алгоритмы решения многих сложных задач на принципиально новом уровне. Нейросетевые технологии успешно применяются в медицине, биологии, введением статистики и финансовых расчетов, для управления роботами и непрерывными производствами, для понимания и синтеза речи, для диагностики технических неполадок в машинах и приборах, для предсказания курсов валют и т.п.

В последнее время возрастает роль такой экономической категории как человеческий капитал. А.И. Добрынин понимает под человеческим капиталом имеющийся у человека запас знаний, навыков, способностей, мотиваций, которые содействуют росту его производительности труда и влияют на рост доходов (зарботков). В данном контексте нельзя не отметить роль оценки человеческого капитала для разработки мероприятий по сохранению и наращиванию этого ресурса. Существуют различные методологические подходы к оценке человеческого капитала, основанные на натуральных и стоимостных показателях. Для комплексной оценки существенный интерес представляет индекс развития человеческого потенциала ИРЧП (Human Development Index), введенный ООН в 1991 г., который отражает величину благосостояния и уровень жизни населения. Однако высокая популярность данного индекса обусловила и активную его критику, связанную с включением в него разнотипных показателей, ошибками вычислений и статистическими неточностями.

Также одним из способов исследования человеческого капитала является применение нейронных сетей ввиду широкого спектра их использования. В связи

с этим целью данной работы является рассмотрение нескольких способов классификации регионов России по параметрам человеческого капитала на основе самообучающихся нейронных сетей.

### **Типовые и «нетиповые» измерители человеческого капитала**

Так как в ИРЧП объединяются стоимостные и натуральные показатели, схема расчета индекса предполагает индексировать каждый показатель в пределах от 0 до 1. В расчетном органе статистической комиссии ООН принято считать, что ИРЧП меньше 0,5 означает низкий уровень жизни, в пределах от 0,5 до 0,8 свидетельствует о среднем уровне жизни, более 0,8 соответствует высокому уровню жизни населения.

Однако такая классификация стран по ИРЧП хорошо описывает мир в целом со слабо развитыми странами Африки, развивающимися – Азии и развитыми – Европы, в то время как при классификации областей в рамках одного региона предлагаются другие подходы, также основанные на количественной оценке ИРЧП.

Так, в докладе о развитии человеческого потенциала в Российской Федерации за 2006/2007 гг. «Регионы России: цели, проблемы, достижения» при классификации регионов России предлагалось считать регион с ИРЧП менее 0,73 – плохо развитым, в диапазоне от 0,73 до 0,78 – средне развитым, в диапазоне от 0,78 до 0,8 – хорошо развитым, а более 0,8 – отлично развитым [4]. А в работе Н.П.Тарасовой, Е.Б. Кручина, С.А. Судакова, С.Н. Мясоедова, М.П. Беляевой предлагалось разбивать группу стран на два класса посредством кластерного анализа [3].

Задача классификации регионов по параметрам человеческого капитала крайне актуальна по следующей причине: в случае реализации программы развития человеческого капитала в рамках какого-либо региона высока вероятность того, что данная программа будет эффективна в рамках другого региона, принадлежащего к тому же классу, что и первый.

Индекс развития человеческого потенциала является ключевым при сравнительном анализе социальных процессов, в т.ч. развития этого показателя в различных странах и регионах. Он рассчитывается как композитный показатель по следующим абсолютным значениям (рис. 1): ВВП на душу населения (скорректированный по паритету покупательской способности и выражаемый в долларах США); ожидаемой продолжительности жизни (для населения от 25 до 85 лет); уровню образования населения, который измеряется долей грамотного населения и долей учащихся от 7-24 лет в учебных заведениях всех видов (школы, начальные, средние и высшие профессиональные учебные заведения). Под грамотностью понимается способность людей читать и писать.

Ожидаемая продолжительность жизни измеряется отдельно для мужчин и женщин и рассчитывается на основе условного поколения, которое составляется из совокупности людей различных возрастов, умерших в данном году.

На основе данных абсолютных значений рассчитываются относительные индексы: индекс ВВП, индекс продолжительности жизни и индекс образования.

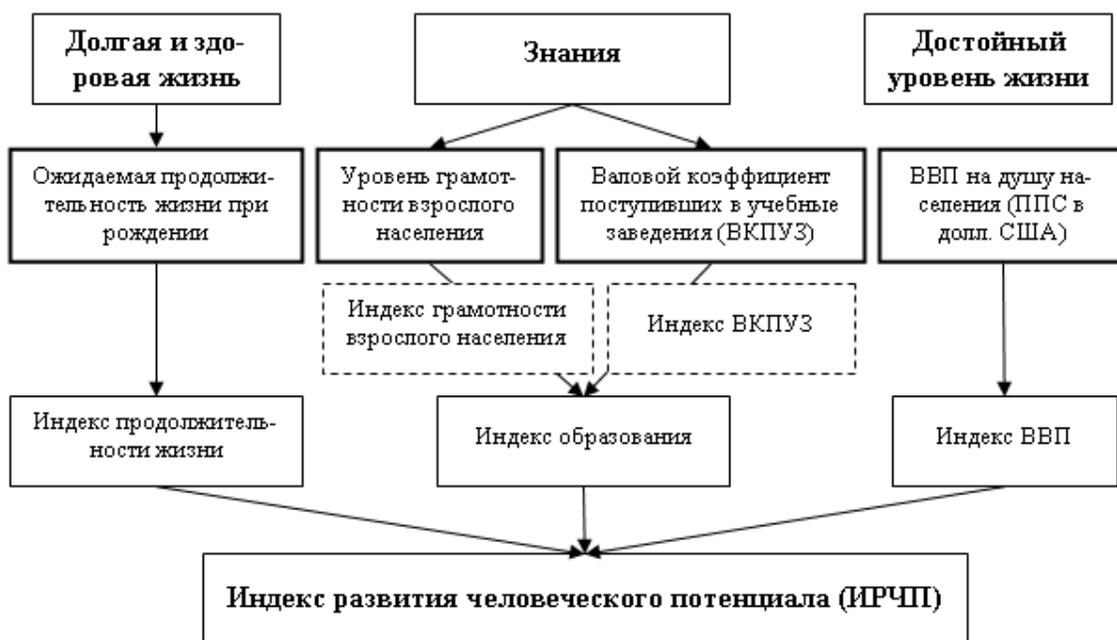


Рис. 1. Структурный состав индекса развития человеческого потенциала (ИРЧП).

Методология расчета индекса развития человеческого потенциала постоянно претерпевает изменения. В настоящее время ИРЧП вычисляется следующим образом [6]:

$$I_0 = \frac{I_1 + I_2 + I_3}{3}, \quad (1)$$

где  $I_1$  – индекс ожидаемой продолжительности жизни;  $I_2$  – индекс достигнутого уровня образования;  $I_3$  – индекс скорректированного реального объема ВВП на душу населения.

Расчет показателей, входящих в состав ИРЧП, может осуществляться двумя способами: отдельно для каждого показателя или по одной общей формуле для всех показателей.

Для расчета указанных выше трех индексов используют следующие формулы [6]:

$$I_1 = \frac{x - x^*}{x^* - x^*}, \quad (2)$$

где  $x$  – ожидаемая продолжительность жизни;  $x^*$  – минимальная продолжительность жизни;  $x^*$  – максимальная продолжительность жизни.

$$I_2 = \frac{2 \times I_{21} - I_{22}}{3}, \quad (3)$$

где  $I_{21}$  – индекс грамотности взрослого населения;  $I_{22}$  – индекс охвата населения образованием первой, второй и третьей ступени.

$$I_3 = \frac{\log y - \log y_{\min}}{\log y_{\max} - \log y_{\min}}, \quad (4)$$

где  $y$  – фактический доход на душу населения;  $y_{\min}$  – минимальный доход на душу населения;  $y_{\max}$  – максимальный доход на душу населения.

Из формулы видно, что при расчете относительных индексов, которые входят в состав ИРЧП, необходимо знать не только абсолютные параметры человеческого капитала на данный момент, но и то, как эти параметры вели себя в прошлом, что не всегда удобно на практике.

Для любого показателя, входящего в состав ИРЧП, индексы могут быть рассчитаны по следующей формуле [1]:

$$I_i = \frac{x_i - x_{i \min}}{x_{i \max} - x_{i \min}}, \quad (5)$$

где  $x_i, x_{i \min}, x_{i \max}$  – фактическое, минимальное и максимальное значения показателя соответственно.

Высокое применение ИРЧП обусловило и его критику. Главной проблемой вычисления индекса является недостаток и качество информации. По данным ООН, для 29 стран, для определения индекса развития человеческого потенциала не хватает базовых данных [1]. Для имеющихся данных серьезной проблемой является их качество, в результате чего они не могут давать достоверной информации для анализа.

Наряду с вышеописанным ИРЧП, регионы можно классифицировать на основе как абсолютных значений параметров оценки человеческого капитала, так и относительных индексов, используя нейронные сети.

В данном исследовании предлагается классифицировать регионы, используя слой Кохонена и самоорганизующиеся сети.

### **Нейросетевой подход при классификации данных**

Нейронные сети Кохонена, или самоорганизующиеся карты Кохонена (Kohonen's Self-Organizing Maps) предназначены для решения задач автоматической классификации, когда обучающая последовательность образов отсутствует [5], т.е. отсутствует априорно определенная информация о выходе изучаемого объекта. Соответственно нет и фиксации ошибки между выходом сети и объекта, на минимизации которой основаны алгоритмы обучения, – например, алгоритм обратного распространения ошибки. Сеть Кохонена – это двухслойная нейронная сеть, содержащая входной слой (слой входных нейронов) и слой Кохонена (слой активных нейронов). На рис. 2 приведен пример такой сети с тремя входами и четырьмя активными нейронами.

В силу отсутствия обучающей последовательности образов, для каждого из которых известна от учителя принадлежность к тому или иному классу, определение весов нейронов слоя Кохонена основано на использовании алгоритмов классической классификации (кластеризации или самообучения) [5].

Слой Кохонена классифицирует входные векторы в группы схожих векторов. При обучении данного слоя на вход подается входной вектор  $\{x_1, x_2, x_3\}$  и вычисляется его скалярные произведения с векторами весов всех нейронов  $\{w_{j1}, w_{j2},$

$w_{j3}$ . Нейроны «конкурируют» друг с другом, нейрон с максимальным значением скалярного произведения является «победителем» и его веса при обучении подстраиваются. В ходе такого обучения получается сетка, сгущающаяся там, где густы исходные данные. Сетка может рассматриваться как средство классификации новых данных – этим данным можно приписывать тот же класс, который ранее был приписан ближайшему к вектору данных узлу карты Кохонена.

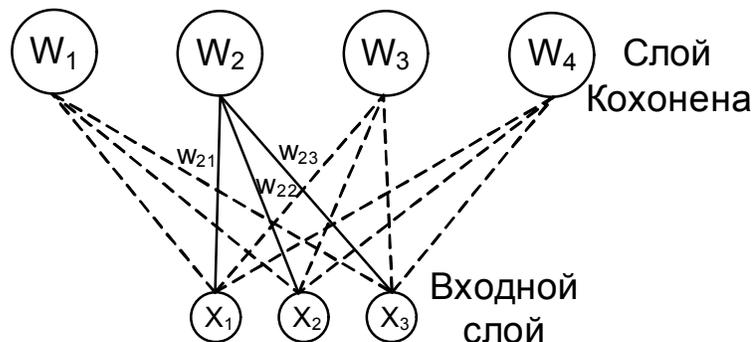


Рис. 2. Сеть Кохонена.

Самоорганизующиеся сети являются усовершенствованной модификацией слоя конкурирующих нейронов (слоя Кохонена). В этой сети нейроны распределяются некоторым пространственным образом (по одному из трех возможных вариантов: в узлах прямоугольной решетки, гексагональной решетки или случайно).

### Пример расчета человеческого капитала на основе нейронных сетей

Для выявления различий в оценке человеческого капитала авторами предлагается использовать данные из доклада ПРООН о развитии человеческого потенциала в Российской Федерации за 2006/07 г. [4].

Используя исходные абсолютные параметры, с помощью слоя Кохонена и самоорганизующейся сети, определяется рейтинг каждого региона РФ, затем полученные результаты сравниваются с расчетом человеческого капитала с помощью ИРЧП. Аналогичная кластеризация производится для относительных индексов. Результат расчетов индекса развития человеческого потенциала в регионах РФ в 2004 г. представлен в табл. 1, где графа *A* – классификация регионов по ИРЧП, *B* – классификация регионов с помощью слоя Кохонена, графа *C* – классификация регионов при помощи самоорганизующейся сети,  $x_1$  – классификация по относительным индексам,  $x_2$  – классификация по абсолютным параметрам.

Таблица 1

Регион	A	B		C		Регион	A	B		C	
		$x_1$	$x_2$	$x_1$	$x_2$			$x_1$	$x_2$		
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
Москва	4	4	4	4	4	Хабаровский край	2	2	3	2	2
Тюменская обл.	4	4	2	4	4	Калужская обл.	2	2	1	2	1
Санкт-Петербург	4	4	3	4	4	Ставропольский рай	2	3	1	1	1
Респ. Татарстан	4	4	4	4	4	Респ. Северная Осетия-Алания	2	3	1	1	1

Продолжение табл. 1

1	2	3	4	5	6	1	2	3	4	5	6
Томская обл.	3	4	4	4	4	Респ. Дагестан	2	3	1	1	1
Липецкая обл.	3	4	4	4	4	Ульяновская обл.	2	3	1	1	1
Омская обл.	3	4	4	4	4	Чукотский автономный округ	2	2	4	4	4
Республика Саха (Якутия)	3	4	4	4	4	Кабардино-Балкарская Респ.	2	3	1	1	1
Белгородская обл.	3	3	3	3	3	Пензенская обл.	2	3	1	1	1
Самарская обл.	3	4	3	4	3	Респ.Калмыкия	2	3	1	1	1
Респ. Башкортостан	3	4	3	4	3	Иркутская обл.	2	2	3	2	2
Волгодонская обл.	3	4	4	4	4	Карачаево-Черкесская Респ.	2	3	1	1	1
Красноярский край	3	4	4	4	4	Камчатская обл.	2	2	1	2	1
Оренбургская область	3	4	3	4	3	Респ. Карелия	2	2	3	2	2
Респ.Коми	3	4	4	4	4	Ленинградская обл.	2	2	3	4	4
Новосибирская обл.	3	2	3	3	2	Респ. Хакасия	2	2	1	2	1
Челябинская обл.	3	4	3	4	3	Новгородская обл.	2	2	3	2	2
Ярославская обл.	3	4	3	4	3	Кировская обл.	2	2	1	1	1
Респ.Мордовия	2	3	1	3	1	Смоленская обл.	2	2	1	2	1
Архангельская обл.	2	4	4	4	4	Тульская обл.	2	2	1	2	1
Орловская обл.	2	2	3	3	2	Приморский край	2	2	1	1	1
Свердловская обл.	2	2	3	4	3	Алтайский край	2	3	1	1	1
Курская обл.	2	2	3	3	2	Тверская обл.	2	2	1	2	1
Удмуртская Респ.	2	2	3	3	2	Костромская обл.	2	2	1	2	1
Волгоградская обл.	2	3	1	3	2	Владимирская обл.	2	2	1	1	1
Магаданская обл.	2	2	3	4	3	Респ.Марий Эл	2	2	1	1	1
Краснодарский край	2	3	1	3	1	Брянская обл.	1	2	1	1	1
Мурманская обл.	2	2	3	4	3	Курганская обл.	1	2	1	1	1
Чувашская Респ.	2	3	1	1	1	Респ. Бурятия	1	2	1	2	1
Пермская обл.	2	2	3	4	3	Амурская обл.	1	2	1	2	1
Саратовская обл.	2	2	1	3	1	Респ.Адыгея	1	3	1	1	1
Астраханская обл.	2	2	3	3	2	Калининградская обл.	1	2	1	2	1
Рязанская обл.	2	2	3	2	2	Псковская обл.	1	2	1	2	1
Нижегородская обл.	2	2	3	2	2	Ивановская обл.	1	2	1	1	1
Кемеровская обл.	2	2	3	4	3	Респ.Алтай	1	2	1	1	1
Сахалинская обл.	2	2	3	4	4	Читинская обл.	1	2	1	2	1
Воронежская обл.	2	3	1	1	1	Респ.Ингушетия	1	3	1	1	1
Ростовская обл.	2	3	1	1	1	Респ.Тыва	1	2	1	1	1
Московская обл.	2	2	3	3	2	Коэф.корреляции		0,7 07	0,6 29	0,6 40	0,7 13
Тамбовская обл.	2	3	1	1	1						

Если считать, что классификация регионов, предложенная в работе [3] и представленная в графе А (табл. 1), «оптимальная», то результат работы слоя Кохонена менее пригоден, чем самоорганизующейся сети. Слой Кохонена при классификации по относительным индексам опустил один класс. Коэффициенты корреляции, рассчитанные относительно заданной классификации и «оптимальной», также это подтверждают.

Классификации самоорганизующейся сетью для абсолютных параметров дала наиболее близкий результат – для крайних классов (1, 4) результат получился идентичный и корреляция у данного примера самая высокая (0,713). Таким образом, применение нейронных сетей в оценке человеческого потенциала позволяет определять группы регионов со схожими закономерностями в базовых показателях человеческого потенциала.

### Заключение

Проведенная работа позволила проанализировать различные подходы к оценке человеческого капитала. Можно сделать вывод о целесообразности применения самоорганизующейся сети в оценке человеческого капитала ввиду близости ее решения к общепринятому. Одним из основных преимуществ применения нейронных сетей является возможность быстро и объективно получить картину положения каждого региона в общем рейтинге. Использование нейронных сетей в региональной кластеризации по параметрам человеческого потенциала позволяет выявлять скрытые закономерности развития в регионах со схожими показателями, а также определять эффективные направления развития человеческого потенциала в регионах, принадлежащих к одной группе. Предложенный вариант классификации по параметрам расчета человеческого капитала может быть использован для совершенствования целевых программ регионального развития.

Таким образом, нейронные сети в оценке человеческого капитала выступают как инструмент для мониторинга процессов формирования регионального человеческого капитала.

### ЛИТЕРАТУРА

1. *Айвазян С.А.* Интегральные индикаторы качества жизни населения: их построение и использование в социально-экономическом управлении межрегиональных сопоставлениях. – М.: ЦЭМИ РАН, 2000.
2. *Загородникова Т.* Об уровне жизни населения // Вопросы статистики. – 1998. – № 6.
3. *Тарасова Н.П., Кручина Е. Б., Судаков С.А., Мясоедов С.Н., Беляева М.П.* Оценка вклада базовых показателей в динамику комплексного показателя человеческого потенциала при формировании механизмов устойчивого развития экономических систем // Менеджмент в России и за рубежом. – 2006. – № 2. – С.17-28.
4. Доклад о развитии человеческого потенциала в Российской Федерации за 2006/07 г. «Регионы России: цели, проблемы, достижения», подготовленный коллективом независимых экспертов и консультантов по инициативе Программы развития ООН (ПРООН) // Общество и экономика. – 2007. – № 5-6. – С.181-202.
5. *Круглов В.В., Борисов В.В.* Искусственные нейронные сети. Теория и практика. – М.: Горячая линия; Телеком, 2001.
6. *Докторович А.* Смысл и методика расчета индекса развития человеческого потенциала // РЭЖ. – 2001. – № 8. – С. 89-91.

*Статья представлена к публикации членом редколлегии Е.А. Ереминым.*