

УДК 519.95

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОЦЕНКИ ДЕЯТЕЛЬНОСТИ НАУЧНЫХ РАБОТНИКОВ¹

© А.А. Арзамасцев, К.Г. Троиц, Н.А. Зенкова, А.В. Неудахин

Arzamastsev A.A., Troitzsch K.G., Zenkova N.A., Neudakhin A.V. Use of methods of mathematical modelling and artificial intellect for evaluation of research officers' activity. A methodology of research officers' activity evaluation by means of ANN-models is developed. The methodology includes a complex of ideas, a new device of artificial neural networks (ANN) which is realized in the form of a computer programme, and technology of its use. Nowadays in connection with the entry of Russia in Bologna agreement it becomes necessary to measure work of scientists, teachers, researchers, post-graduate students, students. Therefore the developed methodology can be also used for the purposes of university management.

Key words: mathematical modeling, device of artificial neural networks, artificial intelligence, evaluation of researchers' activity.

ВВЕДЕНИЕ

Оценка деятельности научных работников представляет собой сложную проблему. Это связано с тем, что такая оценка является сложным индексом, который включает большое число различных компонент. Определение таких факторов и индексов часто трудно формализовать. Кроме того, указанные понятия могут иметь не только содержание (под которым следует понимать общее число индивидуальных компонент), но также и некоторую структуру (под которой можно понимать систему взаимосвязи таких компонент).

Необходимость использования математических методов моделирования в этой сфере связана не только со сложностью данного объекта, но также и с необходимостью получения беспристрастных оценок, т.к. именно на их основе часто принимаются различные управленческие решения (например, распределение грантов, выдвижение на различные должности и т. д.).

В настоящее время данная проблема становится особенно актуальной в связи с переходом российских вузов на международную систему вузовского обучения, предполагающую ее вхождение в Болонское соглашение, одной из самых трудных задач организации которого является разработка системы оценок студентов и преподавателей.

В мире данная проблема частично решается путем вычисления различных индексов цитирования и импакт-индексов. Количество ссылок на конкретную статью или иную публикацию конкретного автора, с од-

ной стороны, свидетельствует об общественном признании ее автора, и с другой стороны, необходимо признать, что такая оценка является достаточно грубым критерием интеллектуальной ценности научных идей.

Институциональное и организационное оформление междисциплинарных индексов цитирования связано, в первую очередь, с созданием Института научной информации США (The Institute for Scientific Information – ISI), основанного Ю. Гарфилдом (Eugene Garfield) и занимающегося систематизацией и оценкой научной информации, опубликованной в научных изданиях. В настоящее время эта организация называется Thomson Scientific. Здесь формируются базы данных индекса цитирования по научным дисциплинам. В настоящее время наиболее разработанными индексами цитирования являются SCI, SSCI, A&HCI.

Однако использование индексов цитирования для оценки деятельности научных работников затруднено в связи со следующими обстоятельствами: зависимость от конъюнктуры – «мейнстримные» работы цитируются лучше, легче, чем пионерские или выходящие за рамки парадигмы, правда, последние могут отыгаться со временем, но таких мало – индекс цитируемости зависит не только от научного уровня, но и от PR-активности ученого (конференции, контакты); проблема соавторов; самоцитирование; проблема публикаций с большим числом авторов и т. д.

Имеются и другие проблемы, связанные, например, с тем, что оценка деятельности научного работника может вообще не быть напрямую связана с индексом цитирования или, напротив, в значительной степени детерминируется важностью самой предметной области, в которой ведется исследование.

Немалые проблемы связаны и с тем, что результирующий оценочный показатель обычно представляет собой нелинейную функцию многих переменных, идентификация структуры связей и коэффициентов в которой на основе эмпирического материала (баз дан-

¹ Совместный научный проект «Разработка компьютерных методов оценки активности научных работников на базе математических моделей, использующих аппарат искусственных нейронных сетей (ИНС)» поддержан грантом DAAD, R325, A/08/08610, 2008; и грантом управления образования и науки Тамбовской области «Региональная поддержка научных исследований, проводимых ведущими научными школами Тамбовской области».

ных) представляет собой сложную математическую проблему.

В настоящее время на кафедре компьютерного и математического моделирования Тамбовского государственного университета им. Г.Р. Державина накоплен значительный опыт в идентификации структуры и содержания трудноформализуемых понятий с помощью аппарата искусственных нейронных сетей [1, 2].

В то же время в Институте информатики университета г. Кобленц (Германия) под руководством профессора Клауса Троица ведутся работы в области компьютерного моделирования в социальных системах. Он является автором нескольких книг и большого числа статей в этой области, хорошо известных в мире (например, [3, 4]).

Указанные обстоятельства явились определяющими при выполнении данной совместной работы.

Целью данной работы является разработка методологии оценки деятельности научных работников с помощью математических моделей, построенных на основе аппарата искусственных нейронных сетей.

МАТЕРИАЛЫ И МЕТОДЫ

Первоначально для выполнения данной работы предполагалось использовать следующие эмпирические материалы:

- информацию о рейтингах ученых Тамбовского государственного университета им. Г.Р. Державина за период с 2004 по 2006 гг.;

- информацию о различных индексах цитирования и рейтингах российских ученых (<http://www.scientific.ru/>);

- информацию о публикационной и других видах научной активности лауреатов Нобелевских премий по различным номинациям, которую предполагали получить от известного нобелиста – профессора В.М. Тютюнника;

- данные Института научной информации (ISI), который в настоящее время называется «Thomson Scientific» (<http://scientific.thomsonreuters.com/>).

В дальнейшем оказалось, что базы данных «Thomson Scientific» полностью закрыты, а информация, полученная от профессора В.М. Тютюнника, неполна и, таким образом, не могла быть использована для наших целей непосредственно. Поэтому мы ограничились лишь тем, что приняли информацию из этих источников к сведению. Информация, содержащаяся в первых двух источниках, оказалась хорошо структурированной и использовалась в дальнейшей работе.

Для реализации проекта использовали специальное программное обеспечение, которое разработано нами в ходе предыдущей деятельности [5–7]. Эти программы представляют собой универсальные нейросетевые симуляторы, использующие оригинальные алгоритмы самоорганизации структуры нейронной сети, что позволяет достигать хорошей гибкости и адаптируемости модели к эмпирическим данным. Алгоритм описан ранее в [5]. Программы для последовательного и параллельного вариантов обучения искусственной нейронной сети разработаны на кафедре КММ ТГУ им. Г.Р. Державина и имеют государственные сертификаты Российской Федерации [6, 7].

По рекомендации профессора К. Ван Метера для определения рейтингов научных работников был опро-

бован метод «ключевых слов», реализация которого в настоящее время сильно упрощена в связи с использованием специальных поисковых машин в сети Internet. При этом с целью получения лишь релевантных, по отношению к исследуемым респондентам, ссылок запрос для поисковых машин составлялся таким образом, что это исключало нахождение информации, например, для однофамильцев. Использовались популярные для РУНЕТа поисковые машины yandex.ru и rambler.ru.

При анализе и моделировании информации, полученной с сайта scientific.ru, использовали следующие частные критерии: **CI86** – полное число цитирований научных работ ученого с 1986 г. (включительно); **CI_{max}** – число ссылок на самую цитируемую публикацию ученого; **FA** – только для тех публикаций, где автор первый или единственный в списке – **CI**, **CI_{max}**.

ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Как это уже было указано ранее, по рекомендации профессора К. Ван Метера из Ecole Superior de Paris (Париж), являющегося признанным экспертом по оценке важности различных социальных явлений и событий, в первой части работы была предпринята попытка использовать известную технологию «ключевых слов» для получения рейтинга научных работников. Необходимо отметить, что указанная технология является в определенном смысле «стандартом» для оценки значимости различных социальных событий, описана в литературе.

Таблица 1

Различные способы оценок профессоров ТГУ им. Г.Р. Державина

Рейтинг профессора ТГУ им. Г.Р. Державина за 2005 год	Количество ссылок по запросу на yandex.ru (документы)	Количество ссылок по запросу на rambler.ru (сайты)	Количество ссылок по запросу на rambler.ru (документы)
x_1	x_2	x_3	x_4
6,97	34	11	27
23,22	118	30	197
10,14	70	9	27
12,75	158	20	71
9,03	137	29	122
15,53	79	18	52
5,71	18	7	19
7,37	8	5	16
10,54	256	46	182
9,76	24	11	28
18,69	85	20	117
13,08	68	16	35
10,26	71	18	63
11,64	16	5	11
15,52	93	21	114
14,18	20	9	24
24,62	155	14	32
10,49	183	29	106
12,56	152	28	74
12,62	13	4	7
16,37	367	78	177
5,65	37	8	19
11,46	166	40	105

В этой части работы использовали рейтинги 23 профессоров ТГУ им. Г.Р. Державина, представленные в первой колонке в табл. 1. Фамилии участников рейтингов в этой таблице не указаны по этическим соображениям. Для этих же лиц были определены и другие оценки рейтингов, составленные в соответствии с технологией «ключевых слов», которые также показаны в табл. 1.

Считая внутренний рейтинг ТГУ им. Г.Р. Державина объективным на том основании, что во-первых, он отражал позицию ректората университета, а во-вторых, был составлен на основе значительного числа частных критериев по четырем основным разделам (1 – формальные показатели, такие как ученая степень, ученое звание, почетные академические звания и т.д.; 2 – показатели научной работы, такие как число публикаций в различных изданиях, число защищенных диссертаций, количество грантов и т. д.; 3 – показатели педагогической работы, включая число изданных учебников и учебных пособий, курсов лекций и т. д.; 4 – показатели социальной активности преподавателя), были предприняты попытки нахождения корреляций внутреннего рейтинга и числа ссылок, полученного с применением «ключевых слов».

На рис. 1–3 показаны различные парные корреляции указанных критериев. Можно сделать вывод о том, что полученные значения коэффициентов корреляции являются довольно низкими, что не дает оснований для непосредственного использования технологии «ключевых слов».

Необходимо отметить, что в это же время коэффициенты корреляции, числа релевантных ссылок, полу-

ченных с помощью различных поисковых машин достаточно высоки (рис. 4–5). Это может указывать на тот факт, что число ссылок указывает на некоторые черты исследуемого респондента, однако эти черты не имеют безусловного влияния на его профессиональный рейтинг.

Необходимо отметить, что при более тщательном изучении данных рис. 1 (с применением кластерного анализа) заметна их неоднородность, что первоначально было связано нами с присутствием в рейтинге ТГУ им. Г.Р. Державина представителей естественных (точных) и гуманитарных наук. На рис. 6 показаны результаты такого кластерного анализа. Так, рис. 6 *a* дает представление обо всей группе данных, а также о характере зависимостей между показателями для представителей естественных и гуманитарных наук. Рис. 6 *b–e* показывают существование четырех видов кластеров, каждый из которых характеризует существование определенного вида зависимости. Видно, что основная группа данных в одинаковой степени соответствует представителям обоих направлений (рис. 6 *b* и *c*). Что касается верхнего кластера, показанного на частных зависимостях (рис. 6 *d* и *e*) и их обобщения, показанного на рис. 6 *f*, то, как следует из проведенного нами анализа, в него входят ученые, активно использующие различные виды PR-активности в сети Internet. Исключение этих данных из общей совокупности позволило получить лучшую корреляцию между показателями x_1 и x_2 , показанную на рис. 7. Однако, по нашему мнению, такой уровень корреляции также не является достаточным для использования технологии «ключевых слов» для подсчета рейтингов научных работников.

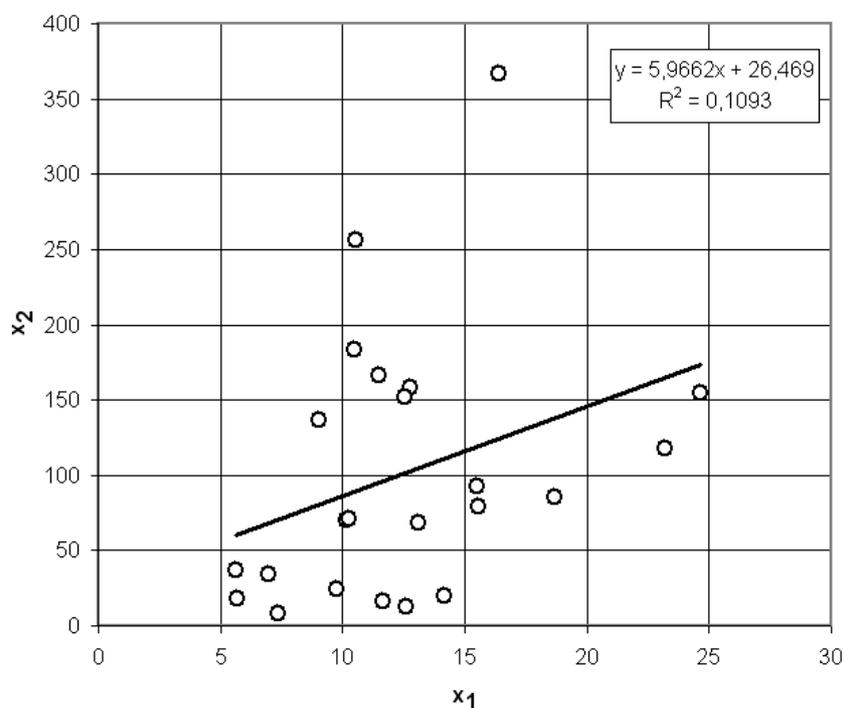


Рис. 1. Корреляция между рейтингом научного работника ТГУ за 2005 г. (x_1) – по оси абсцисс (безразмерный) и числом найденных поисковой машиной yandex.ru релевантных документов (x_2) – по оси ординат. Коэффициент корреляции $r = 0,33$

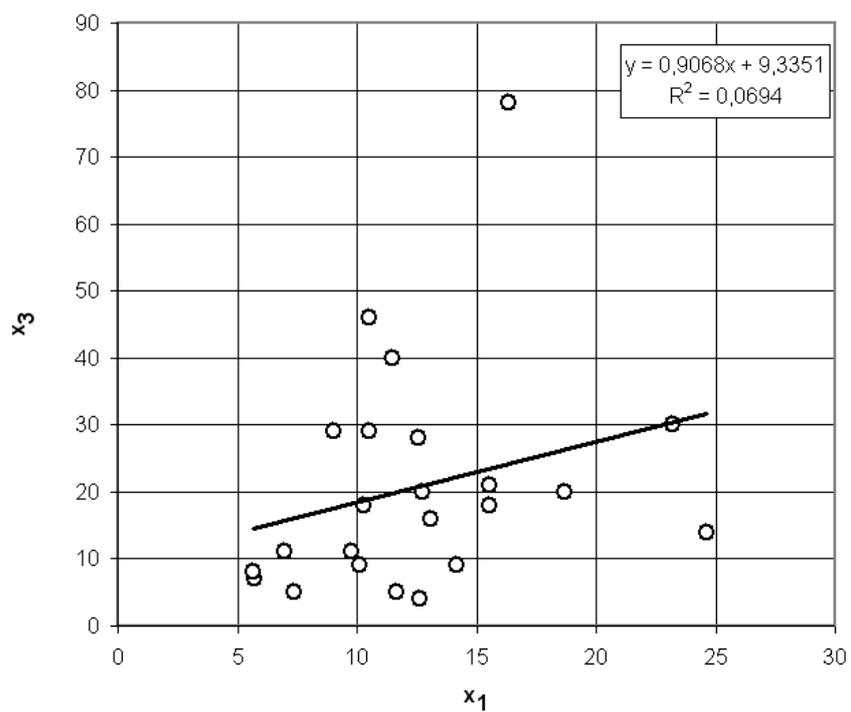


Рис. 2. Корреляция между рейтингом научного работника ТГУ за 2005 г. (x_1) – по оси абсцисс (безразмерный) и числом найденных поисковой машиной gamblеr.ru релевантных сайтов (x_3) – по оси ординат. Коэффициент корреляции $r = 0,263$

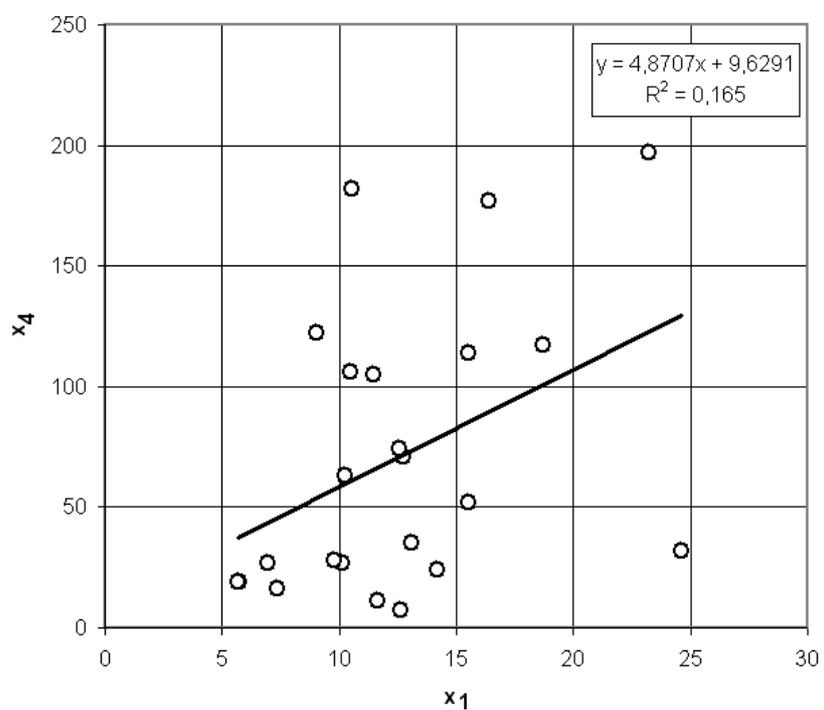


Рис. 3. Корреляция между рейтингом научного работника ТГУ за 2005 г. (x_1) – по оси абсцисс (безразмерный) и числом найденных поисковой машиной gamblеr.ru релевантных документов (x_4) – по оси ординат. Коэффициент корреляции $r = 0,4$

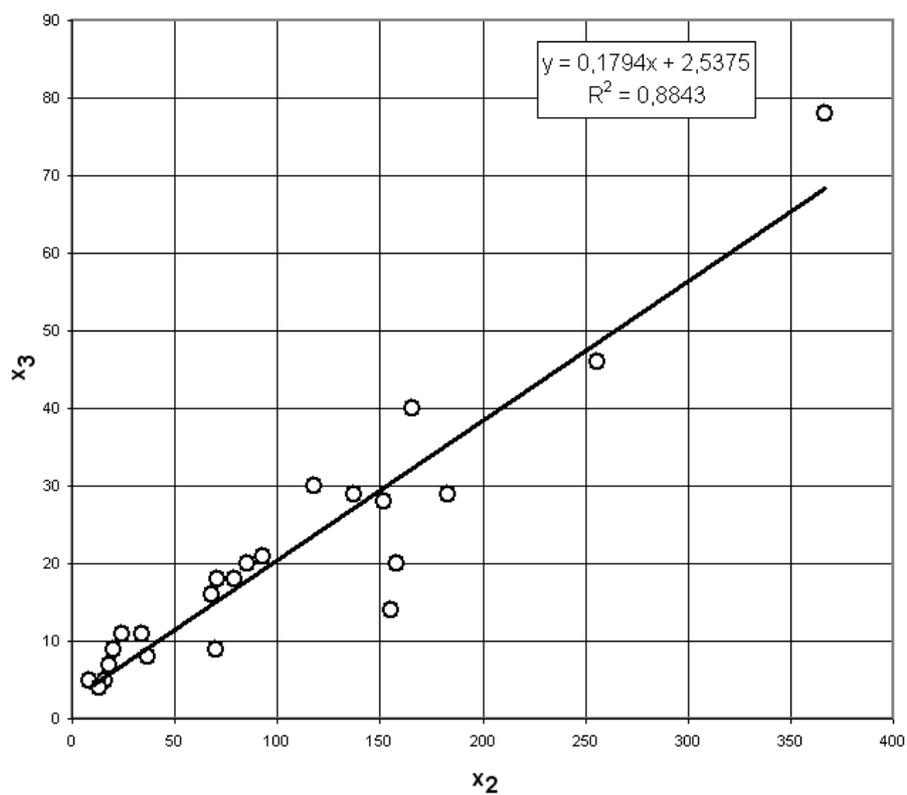


Рис. 4. Корреляция между числом релевантных документов найденных поисковой машиной yandex.ru (x_2) – по оси абсцисс и числом релевантных сайтов, найденных поисковой машиной Rambler.ru (x_3) – по оси ординат. Коэффициент корреляции $r = 0,94$

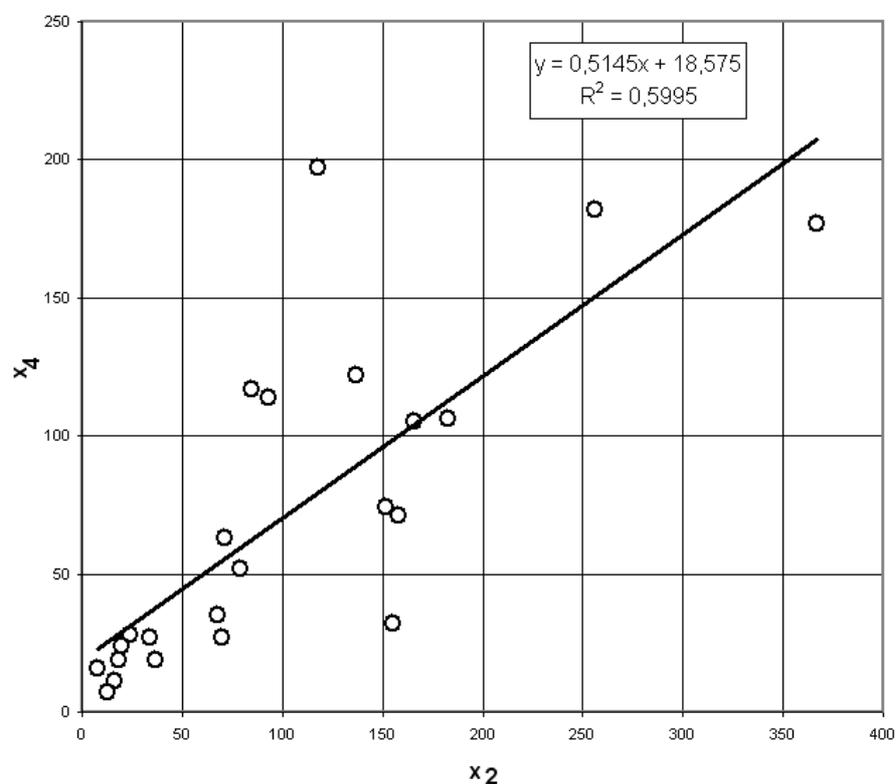


Рис. 5. Корреляция между числом релевантных документов найденных поисковой машиной yandex.ru (x_2) – по оси абсцисс и числом релевантных документов, найденных поисковой машиной Rambler.ru (x_4) – по оси ординат. Коэффициент корреляции $r = 0,77$

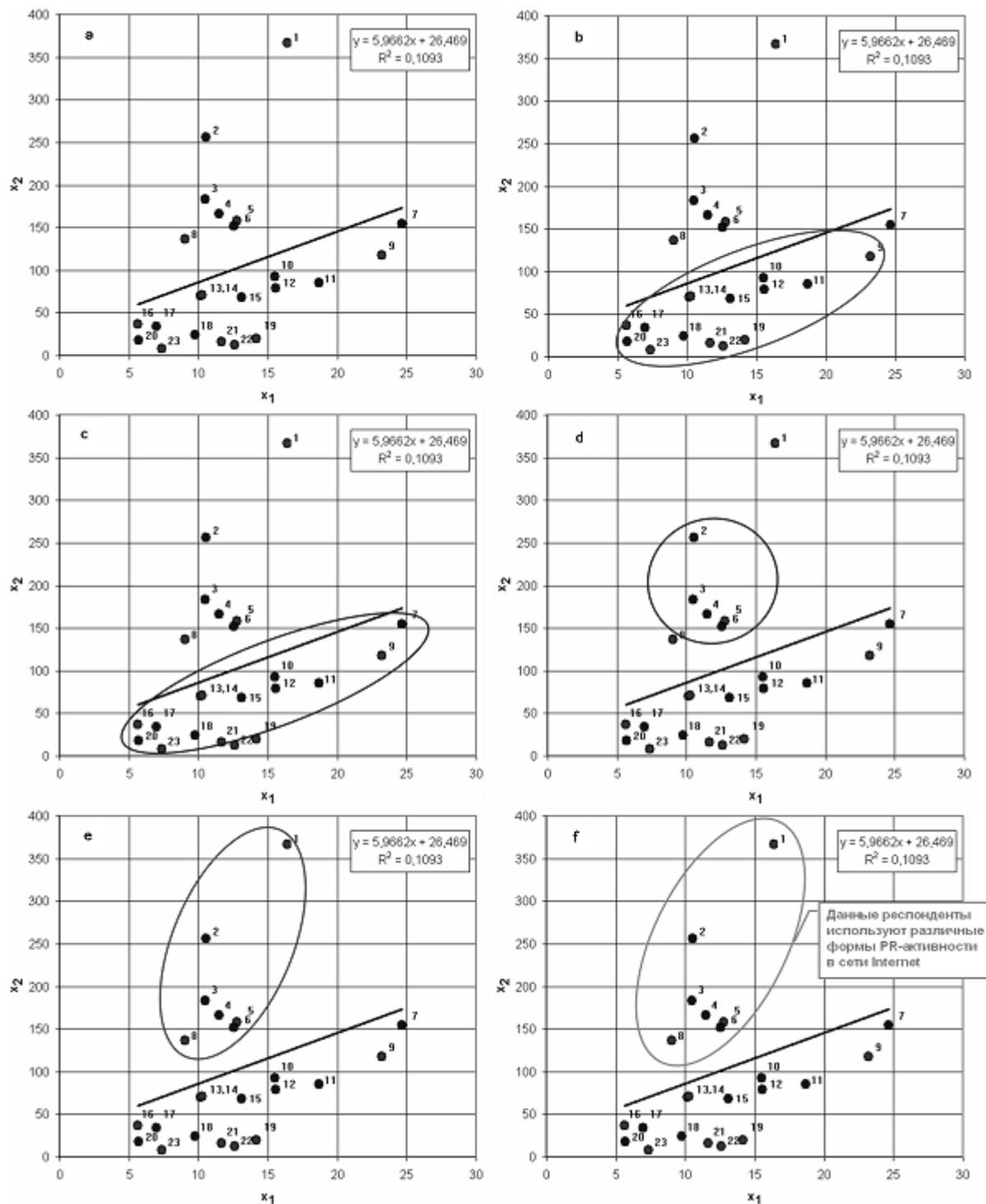


Рис. 6. Кластерный анализ данных рис. 1. Красные точки и линии соответствуют представителям точных наук, синие – представителям гуманитарных наук, а также экономики и права

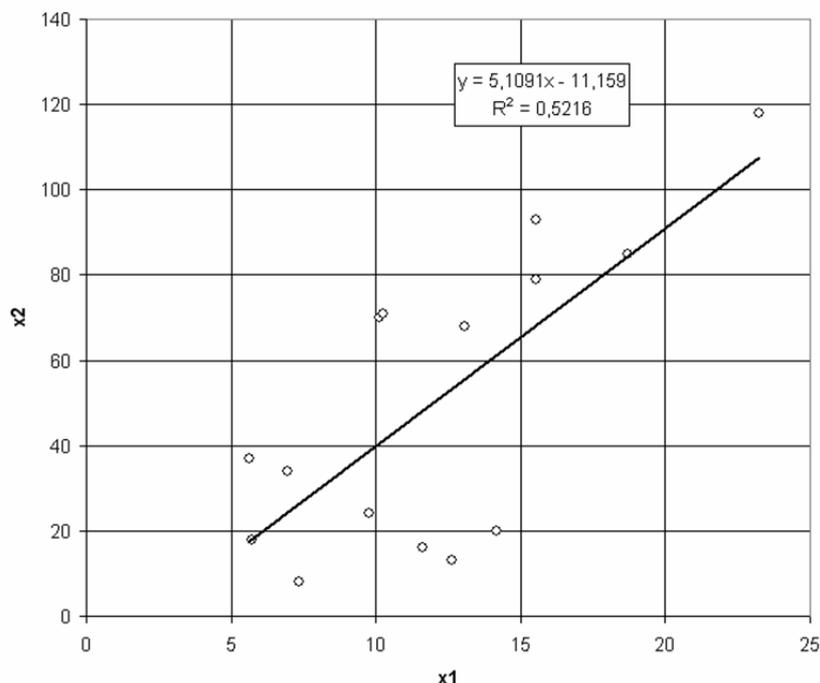


Рис. 7. Корреляция между рейтингом научного работника ТГУ за 2005 г. (x_1) – по оси абсцисс (безразмерный) и числом найденных поисковой машиной yandex.ru релевантных документов (x_2) – по оси ординат без учета лиц, использующих различные виды PR-активности в сети Internet. Коэффициент корреляции $r = 0,72$

Вторая часть работы посвящена выяснению возможности использования аппарата искусственных нейронных сетей (ИНС) для построения математических моделей рейтингов научных работников.

Отметим, что аппарат ИНС, а также технология его использования описаны в специальной литературе (например [1, 3, 4]). В данной статье мы не будем касаться этих вопросов.

В этом разделе работы использовали данные о рейтингах для 566 преподавателей ТГУ им. Г.Р. Державина за 2004 г. Использовали простую структуру ИНС, показанную на рис. 8. Обучение ИНС проводили по методу Монте-Карло, что позволяло надеяться на «проскок» локальных минимумов и нахождение глобального минимума при обучении сети. Для обучения ИНС использовали соответственно 10, 20 и 40 % от всей генеральной совокупности данных о рейтингах. Обученную модель использовали для вычислений рейтингов на основе частных критериев: x_1 – формальные показатели, такие как ученая степень, ученое звание, почетные академические звания и т. д.; x_2 – показатели научной работы, такие как число публикаций в различных изданиях, число защищенных диссертаций, количество грантов и т. д.; x_3 – показатели педагогической работы, включая число изданных учебников и учебных пособий, курсов лекций и т. д.; x_4 – показатели социальной активности преподавателя.

Рис. 9 и 10 суммируют результаты обучения ИНС с линейной и нелинейной внутренними структурами. Видно, что как в первом, так и во втором случаях сеть прекрасно обучается на выборке, включающей от 10 до 40 % всех данных. При этом приведенная погрешность обучения практически не зависит от выбора ИНС-модели и составляет от 0,02 до 0,03 %. Однако прогностические способности моделей различаются

очень существенно: если линейная модель показывает среднюю погрешность прогноза около 1 %, что является вполне приемлемым, то погрешность нелинейной модели может достигать 71% при средних значениях порядка 10 %. Полученные результаты означают:

- ИНС-модели хорошо обучаются на эмпирических данных, представляющих собой рейтинги ученых в совокупности с частными критериями;

- значительные погрешности, возникающие на определенном этапе при усложнении структуры сети (и/или введении в нее новых нелинейных элементов) сопряжены с переобучением сети; этого можно легко избежать, если внимательно следить за погрешностями прогноза; в частном случае (рис. 9 и 10) такие результаты свидетельствуют, что при расчете рейтинга ТГУ им. Г.Р. Державина была использована линейная комбинация частных критериев.

В следующем разделе работы также проверяли возможности использования ИНС-моделей для расчета активности научных работников на примере данных, полученных с сайта www.scientific.ru. Так, на рис. 11, 12 и 13 показаны частные корреляции показателей **CI86**, **CImax** и **FA**. Из их анализа можно сделать вывод, что значения коэффициентов корреляции между этими показателями не велики.

Так, на рис. 11 показана зависимость **CImax**, числа ссылок на самую цитируемую публикацию ученого, от **CI86**, полного числа цитирований научных работ ученого с 1986 г. для 860 российских ученых, специализирующихся в области естественных наук. Большинство точек на этом графике локализовано в малой области $0 \leq \text{CImax} \leq 900$, $1000 \leq \text{CI86} \leq 5000$. По мере удаления от начала координат их плотность падает. Поскольку рейтинг ученого зависит одновременно и от **CImax**, и от **CI86**, на этом рисунке тонкими линиями показаны линии одинаковых рейтингов.

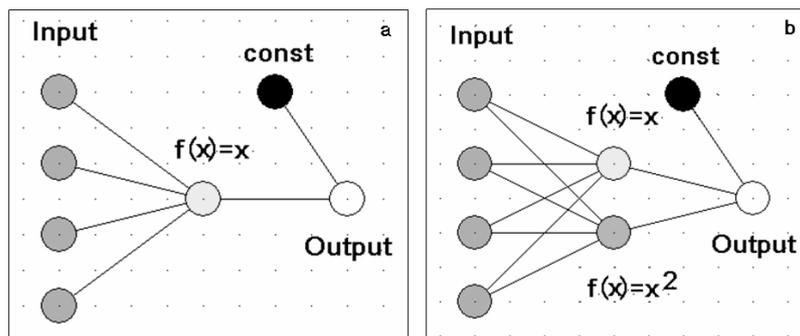


Рис. 8. Линейная – *a* и нелинейная – *b* версии ИНС-моделей для расчета рейтингов ученых ТГУ им. Г.Р. Державина. Линейная модель: общее число нейронов – 7; функциональных нейронов – 1; средняя квадратическая ошибка при обучении модели – 0,29 (0,03 %). Нелинейная модель: общее число нейронов – 8; функциональных нейронов – 2; средняя квадратическая ошибка при обучении модели – 0,2 (0,02 %)

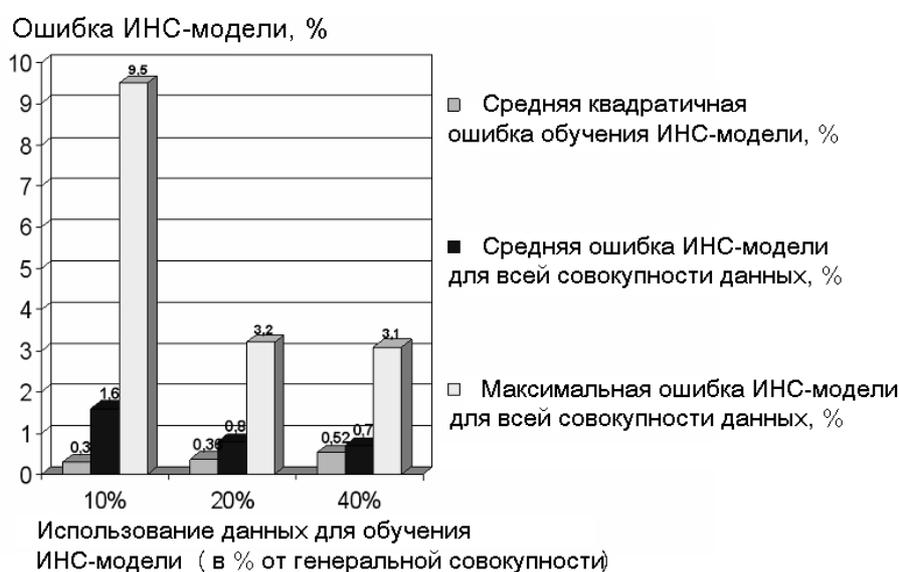


Рис. 9. Сравнительные характеристики линейных ИНС-моделей, полученных при моделировании рейтингов ТГУ им. Г.Р. Державина за 2004 г.

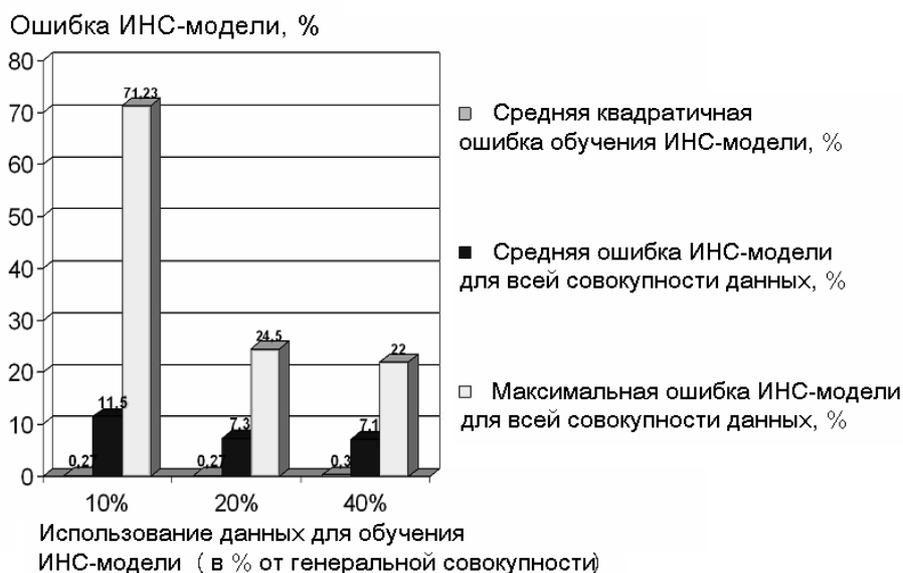


Рис. 10. Сравнительные характеристики нелинейных ИНС-моделей, полученных при моделировании рейтингов ТГУ им. Г.Р. Державина за 2004 г.

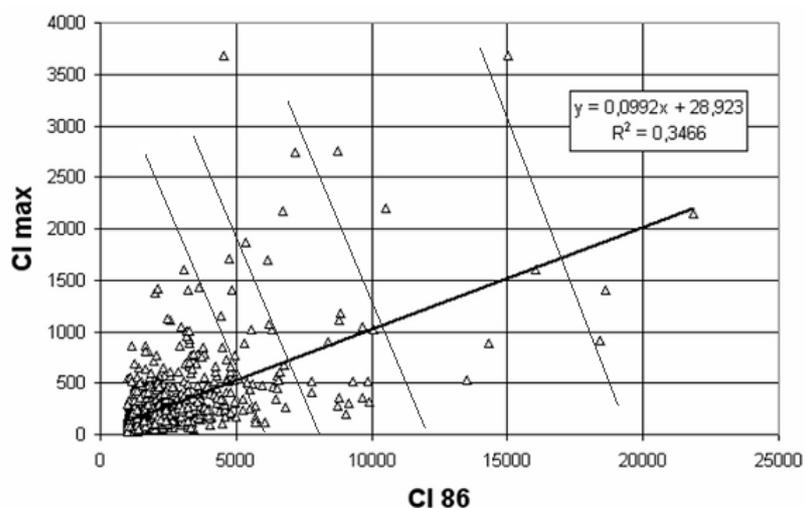


Рис. 11. Корреляция между индексами цитирования CI86 и CI max; коэффициент корреляции $r = 0,59$. Тонкие линии – линии равных рейтингов

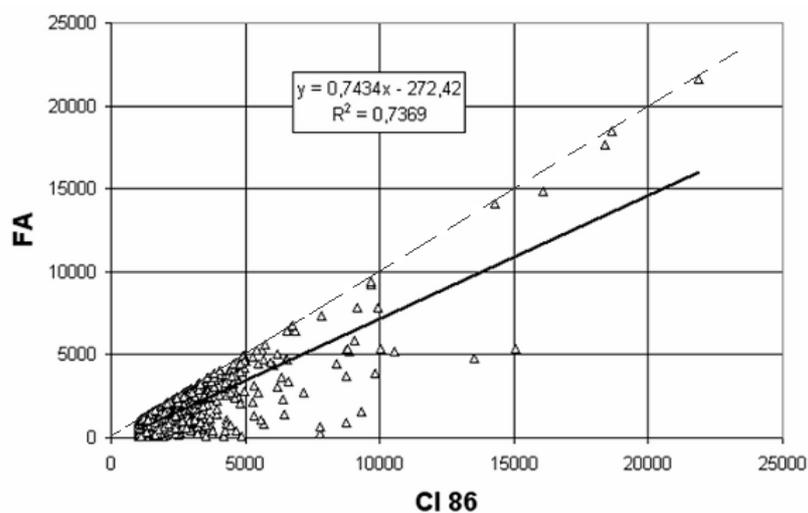


Рис. 12. Корреляция между индексами цитирования CI86 и FA; коэффициент корреляции $r = 0,86$

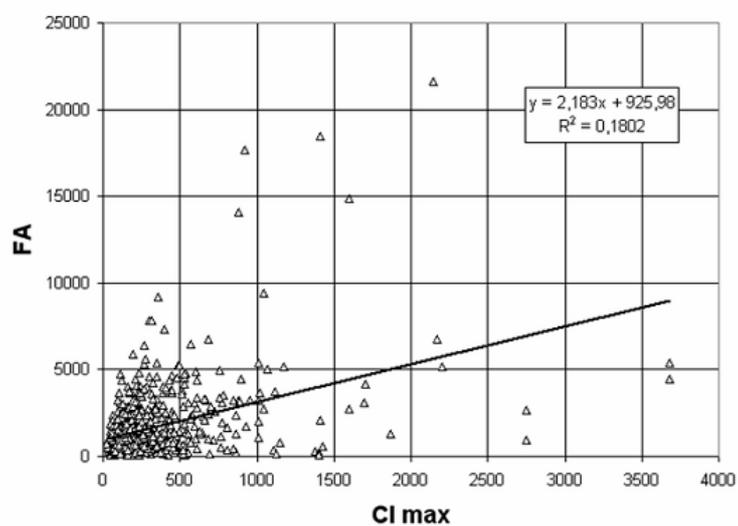


Рис. 13. Корреляция между индексами цитирования CI max и FA; коэффициент корреляции $r = 0,42$

На рис. 12 показана зависимость **FA** – числа ссылок – только для тех публикаций, где автор первый или единственный в списке – **CI**, **CI_{max}** от **CI₈₆** – полного числа цитирований научных работ ученого с 1986 г. В этом случае коэффициент корреляции значительно выше, т. к. $0 \leq FA \leq CI_{86}$. Тонкая пунктирная линия на графике соответствует **FA = CI₈₆**.

Наконец, на рис. 13 показана зависимость **FA** от **CI_{max}**. В этом случае картина похожа на рис. 11.

Таким образом, из анализа рис. 11–13 можно сделать вывод о несущественности корреляций между некоторыми частными показателями, используемыми в рейтинге научных работников.

Следующий раздел данной работы посвящен возможности построения ИНС-модели рейтинга ученого. В качестве рейтинга использовали специальный показатель, который вычисляли следующим образом. Сначала данные по всем научным работникам были ранжированы в порядке убывания общего числа ссылок. При этом на первое место попал ученый, на работы которого с 1986 г. было сделано 21879 ссылок. Число, соответствующее рейтингу, вычисляли по формуле: $R = 1000 - n$ (номер места ученого в ранжированном списке). Таким образом, все 860 ученых расположи-

лись в порядке по мере убывания рейтинга от 999 (у первого) до 142 (у последнего).

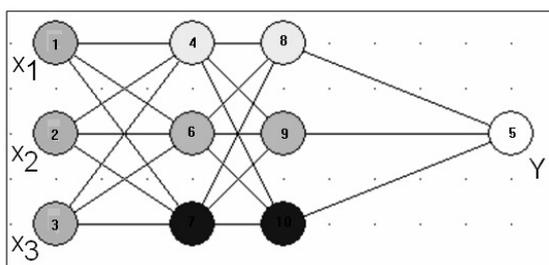
Матрица для обучения ИНС-модели была составлена следующим образом. В качестве входных параметров были взяты: **CI₈₆** – x_1 , **CI_{max}** – x_2 , **FA** – x_3 . В качестве единственного выходного параметра использовали рассчитанный рейтинг ученого **Y=R**.

Для обучения ИНС-модели использовали свободнораспространяемую программу NNC и наши разработки [6, 7]. В процессе обучения ИНС-модели структуру сети изменяли в соответствии с конструктивным алгоритмом, описанным в статье [5].

Окончательная структура ИНС-модели показана на рис. 14 вместе с коэффициентами связей между нейронами. На рис. 15 показано, что корреляция между модельными и эмпирическими данными высокая, что означает хорошую прогностическую способность модели.

Таким образом, показано, что ИНС-модели можно использовать для формирования рейтинга ученых. Поскольку процесс накопления данных в хранилище информации, оценки деятельности ученых, обучения ИНС-модели могут происходить не одновременно, приведем здесь схематично одну из возможных технологий оценки рейтинга ученых на основе ИНС-моделей (рис. 14, 15).

Схема ИНС-модели для подсчета рейтинга ученого на базе индексов цитирования: **CI₈₆** – x_1 , **CI_{max}** – x_2 , **FA** – x_3 . **Y** – рейтинг ученого



Цвета соответствуют:
серый – входные нейроны;
желтый – $f(x)=x$; зеленый – $f(x)=x^2$; синий – $f(x)=x^3$; белый – выходной нейрон.

Таблица связей между нейронами ИНС-модели

#	X	Y	TYPE	DESCRIPTION	LinkTo 4	LinkTo 5	LinkTo 6	LinkTo 7	LinkTo 8	LinkTo 9	LinkTo 10
1	6	2	Input		0,83951		-0,31856	-2,26E-4			
2	6	4	Input		0,01644		-0,0059	0,00564			
3	6	6	Input		0,00679		2,45E-4	0,10284			
4	9	2	$f=x$						1,21994	0,72867	-0,1145
5	16	4	OUT								
6	9	4	$f=x^2$						-2,29896	0,29494	-0,04503
7	9	6	$f=x^3$						0,19841	-0,18618	0,04569
8	11	2	$f=x$				1				
9	11	4	$f=x^2$				0,21723				
10	11	6	$f=x^3$				1				

Рис. 14. Параметры нелинейной ИНС-модели для расчета рейтингов ученых за 2008 г. на базе индексов цитирования. Номера нейронов показаны на схеме их соединений; коэффициенты связи между нейронами, соответствующие их номерам, показаны в таблице в нижней части рисунка

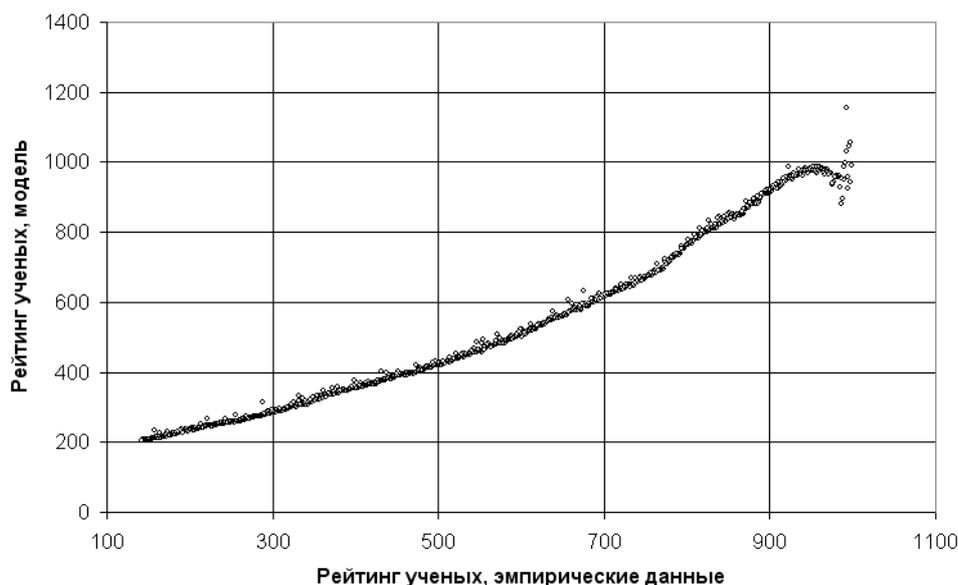


Рис. 15. Корреляция между экспериментальными (эмпирическими) данными и модельными расчетами рейтингов ученых России за 2008 г. Коэффициент корреляции $r = 0,98$

Общая методология определения рейтингов с помощью ИНС-моделей может быть представлена последовательностью следующих шагов.

Шаг 1. Определяем целевую группу научных работников, для которых будет производиться оценка их деятельности. В качестве такой группы могут выступать научные работники, для которых известны основные показатели их деятельности, такие как количество научных работ за различные периоды, число публикаций, выступлений на конференциях, количество и суммы грантов, защиты диссертаций и т. д., а также независимые оценки респондентов этой группы.

Шаг 2. Определяем входные и выходные показатели (поля базы данных), на основании которых в дальнейшем будет производиться обучение искусственной нейронной сети. Поскольку на данном этапе нет возможности принимать какие-либо решения по поводу того, насколько значим или незначим данный показатель и как (каким образом) он может влиять на общий рейтинг научного работника, то в дальнейшем будем использовать максимально возможное число входных показателей, которые удастся извлечь из базы данных. Незначимые и малозначимые входные показатели будут удалены после завершения этапа первичной идентификации ИНС-модели. Необходимо отметить также, что в качестве выходного показателя могут быть использованы различные формальные достижения, уже полученные ученым и показывающие его признание научным сообществом – награды, гранты, членство в академиях и т.д.

Шаг 3. Готовим данные для обучения искусственной нейронной сети. Данный этап является формальным, т. к. он включает лишь конверсию таблиц с эмпирическими показателями в формат DBF или Microsoft Excel, т.к. используемое программное обеспечение работает именно с такими форматами данных (см. пункт б).

Шаг 4. Задаем первичную структуру ИНС и проводим ее обучение эмпирическим данным. На этом этапе

выбираются: количество входных нейронов (как правило, это число должно быть равно числу различных показателей, имеющихся в базе данных); количество выходных нейронов (это число должно быть равно числу независимых оценок); количество слоев и количество нейронов в слое (выбираются в соответствии с решаемой задачей); система связей между нейронами (выбирается в соответствии с решаемой задачей, первоначальная конструкция может быть разработана по принципу «каждый с каждым»); активационные функции нейронов (выбираются из известного набора – линейная, параболическая, сигмоид, ступенчатая функция и т.д.). Реализация этого этапа является одним из самых тонких мест в указанной технологии. От того, насколько удачно выбрана структура ИНС, в значительной степени зависит ее способность к обучению, адекватность модели и количество итераций, которое предстоит выполнить на пути реализации технологии. Необходимо отметить, что последующие шаги гарантируют адекватность ИНС-модели даже в том случае, если первоначальная структура была выбрана неверно. На этом шаге получаем первичную ИНС-модель оценки труда ученого, которая должна быть уточнена в дальнейшем.

Шаг 5. Осуществляем удаление несущественных связей ИНС и проводим окончательную настройку модели. Несущественными считаются связи, весовые коэффициенты которых ниже заданного, определяемого пользователем уровня. Такие связи практически не оказывают никакого воздействия на качество работы всей модели в целом. Тем не менее после их удаления для окончательной настройки модели необходимо еще раз провести обучение сети. Изменение структуры сети осуществляем на основе созданного нами алгоритма [5]. В результате получаем вторичную модель оценки труда ученого, по которой возможно проводить анализ структуры и содержания этого понятия.

Шаг 6. Проверка адекватности модели и коррекция ее параметров осуществляются в сравнении с реальными

ми эмпирическими данными. После проверки адекватности ИНС-модели должен быть выполнен ее анализ: определены структуры и содержание самого понятия «оценка деятельности научного работника» или «рейтинг научного работника». Алгоритмы такого анализа базируются на оценках чувствительности определяемого фактора, в данном случае это «оценка деятельности научного работника» к составляющим факторам, представляющим собой отдельные компоненты.

Шаг 7. После анализа модели может быть выполнено ее упрощение и, возможно, получено выражение для расчета рейтинга научного работника.

ЗАКЛЮЧЕНИЕ

Таким образом, в данной работе показано, что аппарат искусственных нейронных сетей может быть использован в задачах определения рейтингов научных работников. Приведена возможная технология использования ИНС-моделей для решения данных задач. Указанная технология может быть использована в качестве ядра информационной системы для оценки деятельности научных работников.

ЛИТЕРАТУРА

1. *Арзамасцев А.А., Зенкова Н.А.* Система психологического тестирования на основе аппарата искусственных нейронных сетей // Искусственный интеллект. 2004. № 2. С. 237–242.
2. *Арзамасцев А.А., Зенкова Н.А.* Моделирование в психологии на основе искусственных нейронных сетей. Тамбов, 2003.
3. *Gilbert N., Troitzsch K.* Simulation for the Social Scientist. N. Y., 2006.
4. Representing Social Reality. Third Conference of European Social Simulation Association / ed. by K. Troitzsch. Koblenz, 2005. September 5-9.
5. *Арзамасцев А.А. и др.* Реализация проекта TEMPUS TACIS “System modernization of University Management” в Тамбовском государственном университете им. Г.Р. Державина // Information Technologies in University Management. Int. Conf. Tambov, 2006. October 17-20. С. 9–44.

6. *Арзамасцев А.А., Крючин О.В., Королев А.Н., Зенкова Н.А.* Многофункциональный программный комплекс для компьютерного моделирования на основе искусственной нейронной сети с самоорганизацией структуры. Свидетельство об официальной регистрации программы для ЭВМ. № 2007610622. Заявка № 2006614383. Дата поступления 15.12.2006. Зарегистрировано в Реестре программ для ЭВМ 8.02.2007.
7. *Арзамасцев А.А., Крючин О.В., Королев А.Н., Суспицына М.А., Вязовова Е.В., Семенов Н.О.* Универсальный симулятор, базирующийся на технологии нейронных сетей, способный работать на параллельных машинах. Свидетельство об официальной регистрации программы для ЭВМ. № 2008610860. Заявка № 2007615431. Дата поступления 28.12.2007. Зарегистрировано в Реестре программ для ЭВМ 20.02.2008.

БЛАГОДАРНОСТИ: Авторы благодарят: профессора ТГУ им. Г.Р. Державина А.А. Слеткова за предоставленную возможность работать с базой данных рейтингов ученых университета за 2004–2006 гг.; профессора В.М. Тютюника за возможность ознакомиться с материалами, численно характеризующими вклад в науку лауреатов Нобелевской премии по различным номинациям; профессора Карла Ван Метера (Karl Van Meter, Ecole Superior de Paris) за обсуждение возможных подходов и полезные советы.

Поступила в редакцию 17 сентября 2008 г.

Ключевые слова: математическое моделирование, аппарат искусственных нейронных сетей (ИНС), искусственный интеллект, оценка деятельности научных работников.

Разработана методология оценки деятельности научных работников с помощью ИНС-моделей. Методология включает комплекс идей, новый аппарат искусственных нейронных сетей (ИНС), который реализован в виде компьютерной программы, и технологию ее использования. В настоящее время из-за вхождения России в Болонское соглашение имеется потребность измерения работы ученого, преподавателя, исследователя, аспиранта, студента. Поэтому развитая методология может использоваться и для целей университетского управления.