



DOI: 10.19181/4m.2023.32.1.5

EDN: SKANLQ

О.Р. Чепьюк, О.Ю. Ангелова, А.Л. Сочков, Т.О. Подольская
(*Нижний Новгород*)

ТИПОЛОГИЗАЦИЯ ПРОФЕССИОНАЛЬНЫХ ТРАЕКТОРИЙ ОДАРЕННЫХ ЛИЧНОСТЕЙ С ПОМОЩЬЮ НЕЙРОСЕТЕВОГО АНАЛИЗА¹

На основе массива данных (100 биографий), сформированного авторами по результатам контент-анализа биографического материала о выдающихся ученых XIX и XX вв. в гуманитарной и естественно-научных сферах, проведена кластеризация профессиональных траекторий одаренных личностей. Методом кластеризации стал нейросетевой анализ на основе самоорганизующихся карт Кохонена. Сами траектории были сформированы в рамках поведенческой модели линейно-стадиального подхода в исследовании жизненных циклов. В рамках этого подхода карьера и профессиональная самореализация человека понимаются как

Ольга Ростиславовна Чепьюк – доктор философских наук, профессор кафедры управления человеческими ресурсами, Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия. Email: cheruuko@yandex.ru.

Ольга Юрьевна Ангелова – кандидат экономических наук, доцент кафедры информационных технологий и инструментальных методов в экономике, Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия. Email: oangelova@mail.ru.

Андрей Львович Сочков – кандидат технических наук, доцент кафедры информационных технологий и инструментальных методов в экономике, Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия. Email: sochkov@iee.unn.ru.

Татьяна Олеговна Подольская – кандидат социологических наук, доцент кафедры управления человеческими ресурсами, Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия. Email: podolskaya@iee.unn.ru.

последовательность этапов эволюции человека, фиксированных в порядке наступления. Каждый из этапов был закодирован, а биографии преобразованы в систему векторов. В свою очередь задача кластеризации заключалась в разбиении массива объектов из сотни векторов на типовые группы, имеющие несколько вещественных интервальных координат. Критериями качества кластеризации стали показатели минимальной суммы ошибок квантования, а также коэффициент силуэта. По итогам исследования были выделены и интерпретированы семь профессиональных траекторий одаренных личностей. Анализ траекторий проводился с точки зрения скорости достижения успеха (среднего возраста успеха) и тех факторов и условий жизненного пути, которые могли повлиять на более быстрое или медленное достижение профессиональных целей и самореализацию. На этом примере были показаны возможности и ограничения использования нейросетевого анализа для решения сходных исследовательских задач, в особенности когда требуется работать со сложными формами кластеров и находить их оптимальное число.

Ключевые слова: нейросетевой анализ, одаренность, одаренная личность, профессиональная траектория, машинное обучение, нейронная сеть, «социальный лифт», карты Кохонена

Введение

Управление талантами и поддержка социальных лифтов декларируются как одна из важнейших задач для государства (макроуровень)¹, корпораций (мезоуровень) и отдельной личности (микроуровень). Изучение возможностей для самореализации одаренной личности находится на стыке наук: педагогики, психо-

¹ Создание в России «возможностей для самореализации и развития талантов» является одной из приоритетных целей развития Российской Федерации до 2030 г. (национальных целей). См.: Указ Президента РФ от 21 июля 2020 г. № 474 «О национальных целях развития Российской Федерации на период до 2030 года» // Официальное опубликование правовых актов [site]. URL: <http://publication.pravo.gov.ru/document/0001202405070015> (дата обращения: 20.01.2023).

логии, социологии, философии, экономики. На фундаментальном уровне вопрос может быть сформулирован следующим образом: каким образом образование, рынок труда и в целом хозяйственный организм обеспечивают человека возможностями и условиями для развития его задатков и самореализации в профессиональном плане? На прикладном уровне задача заключается в выявлении повторяющихся паттернов (траекторий), описывающих социальные условия, в которых личность достигает успеха в кратчайшие сроки. Появление многочисленных структурированных источников биографической информации (WikiData, Pantheon), а также возможностей их автоматической обработки дает возможность проводить такие исследования: имеются значительные объемы цифровых следов, отражающих основные этапы жизненного пути человека¹. Методологи отмечают, что ученые, по сравнению с коммерческими компаниями, предпринимают осторожные попытки в применении численных методов анализа биографий [1]. В то же время специалисты в сфере компьютерного анализа активно создают структурированные базы биографий [2; 3; 4; 5], а также инструменты для визуализации [6; 7].

Российская и советская научные школы традиционно фокусировались на изучении жизненного пути человека. Так, Л.С. Выготский делал акцент на процессах становления личности: человек – субъект жизни, а только затем – субъект поведения. В одаренности важно не столько наличие задатков, сколько их развитие и применение [8]. Исторические исследования одаренности, начиная с работ Ф. Гальтона в XIX в. и продолжая трехкольцевой моделью Дж. Рензулли, подчеркивают важность так называемых «земных» факторов. В России вклад в исследование одаренности внесли научные школы А.Н. Леонтьева [9], В.С. Мерлина, Я.А. Пономарева, Б.М. Теплова, В.Д. Шадрикова [10]. Особое внимание в контексте

¹ См., напр.: Wikidata [site]. URL: https://www.wikidata.org/wiki/Wikidata:Main_Page (date of access: 20.01.2023).

исследования заслуживают работы американского исследователя Д. Галенсона [11; 12]. Он выделил два паттерна успеха: *экспериментальные* инноваторы достигают результатов методом проб и ошибок, на закате профессиональной жизни; *концептуальные* ученые и деятели искусства совершают прорыв на самых ранних этапах карьеры.

Опираясь на традицию российской научной школы, главной задачей исследования стал поиск новых методов, которые бы позволили изучать процессы профессионального становления личности. Ключевым критерием стал срок достижения успеха у одаренного с учетом социальных условий его жизни (стимулов и препятствий к успеху).

Обоснование метода

На методологическом уровне задача выявления устойчивых (повторяющихся) траекторий развития одаренной личности связана с вопросами кластеризации. Кластеризация используется в современной социологической науке для группировки объектов или наблюдений на основе их сходства. Этот метод помогает выявлять паттерны и структуры в данных, что может быть полезным при изучении траекторий одаренных. Тема кластеризации в социологическом исследовании является одной из часто обсуждаемых: неоднократно описаны как достоинства, так и недостатки различных методов, в том числе метода *k*-средних, анализа латентных профилей, метода пороговых значений [1; 13], иерархической кластеризации [14]. В нашем исследовании применялся метод кластеризации на основе нейросетевого анализа. В отличие от методов иерархической кластеризации, методы на основе нейросетей и машинного обучения могут автоматически извлекать более сложные паттерны в данных. В отличие от метода *k*-средних, они могут учитывать более сложные формы кластеров и не требуют задания числа кластеров. Среди недостатков следует

отметить, что методы на основе нейросетей могут иметь проблемы с выбросами и шумом, если не настроены правильно, они требуют большего объема данных и вычислительных ресурсов для обучения модели [15; 16].

Отметим, что кластеризация для определения типовых траекторий достижений успеха одаренных личностей является более сложной задачей, чем традиционно ставятся исследователями жизненного пути: например, их часто интересует, в каких высших школах работали Нобелевские лауреаты [17], насколько значим фактор мобильности в карьере [18]. Весьма популярными являются вычисления среднего «возраста успеха» [7; 12; 19; 20]. Отдельный интерес представляют биографии спортсменов: их профессиональные достижения легко поддаются оцифровке [21; 22]. Стоит отметить, что исследование социокультурного контекста хотя и декларируется, остается преимущественно частью теоретических размышлений на эту тему [23]. Налицо недостаток поведенческих моделей и новых методов социологического исследования, которые могли бы стать опорными для количественного анализа профессиональных траекторий.

Описание этапа сбора данных и контент-анализа

В выборку исследования вошли биографии 100 выдающихся ученых, чей расцвет профессиональной карьеры приходился на XIX и XX вв. Квотирование выборки осуществлялось по сферам научного знания: 45 человек естественно-научного профиля, 55 человек – гуманитарного. Распределение квот в выборке было сделано на основе двух допущений. Во-первых, оно отражает текущее распределение исследовательских интересов в академическом сообществе, где гуманитарные науки часто включают междисциплинарные направления. Во-вторых, такое распределение позволяет исследовать специфические траектории развития

карьеры в гуманитарных науках, которые исторически меньше исследовались по сравнению с естественными. Это дает возможность оценить различия и сходства в траекториях одаренности. Отметим, что для соблюдения валидности в выборку были включены персоналии, входящие не менее чем в три рейтинга международных и российских издательств, в том числе: американского журнала «Time», «Британской энциклопедии», российской серии научно-популярных книг «Жизнь замечательных людей», «Большой советской энциклопедии», «Большой российской энциклопедии». Для справочной навигации по биографиям использовалась информация из русскоязычной Википедии.

Работа коллектива была разделена на два этапа. На первом этапе была поставлена задача определить время успеха одаренной личности. Годом успеха признавалась публикация (обнародование) первого из значимых (программных) произведений (совершенных открытий), признаваемых обществом ключевым достижением одаренной личности в соответствии с ее доминантой оригинальности. Например, для Д.И. Менделеева, несмотря на его разносторонние увлечения, таким произведением остается его «Таблица периодических элементов» (1869 г.). На втором этапе с помощью контент-анализа проводилась первичная обработка смысловых единиц биографий, которые характеризовали различные модели поведения одаренного относительно выбора профессионального пути. Эти условия (категории) были сгруппированы в соответствии с линейно-стадиальным подходом к описанию жизненного пути (рис. 1).

Из биографий выделялись такие кодовые фразы, которые могли охарактеризовать условия выбора. Так как выборка состояла из биографий ученых, значимыми условиями признавались те, в которых одаренный имел возможность получить образование и занять сильную профессиональную позицию. Образование и сфера карьерной реализации должны были соответствовать доминанте оригинальности ученого и профилю его будущих достижений.

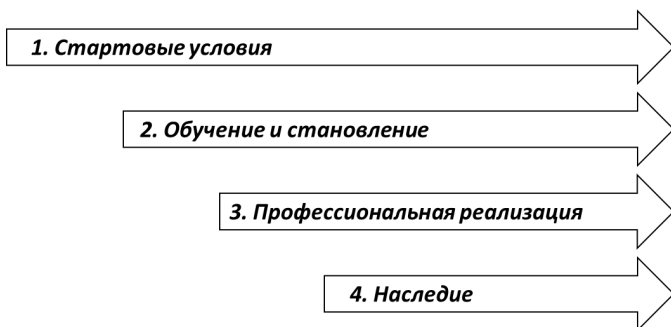


Рис. 1. *Линейно-стадиальная модель жизненного пути*

На первом этапе исследования мы определили, что основными факторами, влияющими на успех одаренной личности, являются условия, создаваемые семьей. Это заключение было сделано на основе анализа биографических данных, где особое внимание уделялось двум аспектам: уровню финансовой обеспеченности семьи и уровню образования и мотивации к обучению и развитию ближайшего круга, в первую очередь родителей. Мы провели систематический контент-анализ биографий, в ходе которого выявили и категоризировали эти факторы как ключевые. Было установлено, что сочетание этих условий создает благоприятную среду для одаренных личностей, позволяя им более свободно выбирать образовательные учреждения и направление профессиональной самореализации. Это, в свою очередь, способствует развитию их потенциала (рис. 2А). Например, из биографии Нильса Бора мы узнаем, что он *«родился в семье Христиана Бора, дважды кандидата на Нобелевскую премию, и Эллен Адлер, дочери весьма влиятельного банкира»*. Такое содержание позволяет отнести условия его жизни к ситуации 1.3 (рис. 2А).

На втором этапе «Обучение и становление» к значимым условиям были отнесены наличие и профильность образования (см. рис. 2Б). Здесь учитывалось соответствие сферы достижений ученого (доминанты одаренности) направлению его образова-



Рис. 2А. Матрица сочетания условий на стартовом этапе



Рис. 2Б. Матрица сочетания условий на этапе обучения

ния – как первого (ранняя профессиональная ориентация), так и последующих (осознанный выбор для карьеры). Так, из биографии В. Даля мы узнаем, что он «*после нескольких лет службы на флоте поступил в Дерптский университет на медицинский факультет*». Однако доминанта оригинальности его достижений лежит в области лексикографии и лингвистики. Следовательно, условия на втором этапе могут быть отнесены к ситуации 2.0 (рис. 2Б). Иная ситуация была у Н. Бора. Он «*успешно изучал физику, химию и математику в том же вузе, где преподавал его отец*», поэтому его ситуация кодируется как 2.2 (рис. 2Б).

На третьем этапе – «Профессиональная самореализация» – условиями, определяющими модель поведения, стали степень инновационности сферы приложения усилий одаренного, а также баланс между ролью коллектива (в том числе сильного влияния наставника и научной школы) и личной оригинальной позицией одаренного (рис. 3).

Под степенью инновационности понимается не столько новизна результатов, сколько осознанный или случайный выбор сферы приложения творческих усилий одаренного. Например, из биографии В. Даля узнаем, что на его замыслы и творчество

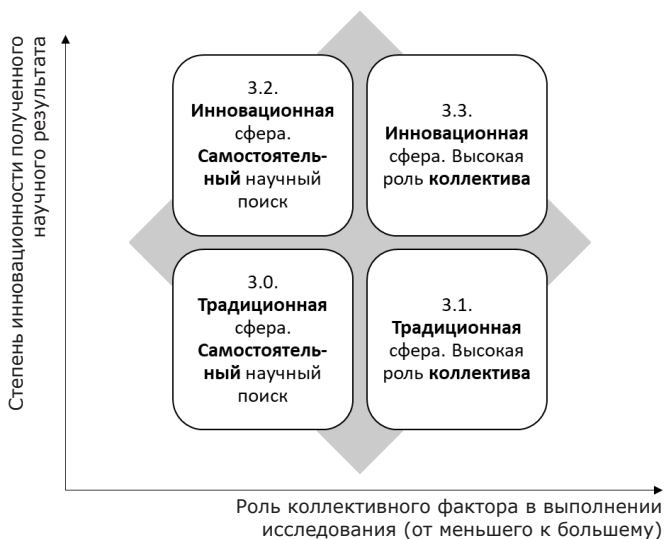


Рис. 3. Матрица сочетания условий на этапе «Профессиональная реализация»

серьезное влияние оказал А.С. Пушкин. Такие условия могут быть закодированы как ситуация 3.1 (рис. 3): работа в традиционной сфере при существенном влиянии наставника и его окружения.

Процедура кодировки биографических данных проводилась в два этапа. На первом этапе был осуществлен контент-анализ биографий, выделены данные по ключевым стадиям жизненного пути ученых. Эти данные включали как категориальные переменные (например, ФИО, сектор науки), так и количественные (скорость успеха в годах). На втором этапе была проведена кластеризация с использованием карт Кохонена, где каждый ученый был представлен вектором из шести координат. Эти координаты включали как номинальные, так и преобразованные в интервальные вещественные переменные (например, оценки стартовых условий, условий обучения и профессиональной реализации). Оценки этих факторов были определены экспертным путем на

основе данных первого этапа анализа. Результаты кластеризации использовались для анализа распределения ученых по различным категориям. Вопросы о неоднозначности данных и их репрезентативности решались привлечением нескольких экспертов для усреднения оценок и получения более объективных результатов. Отметим, что в исследовании на данном этапе не рассматривался заключительный этап жизненного цикла («Наследие», рис. 1). Он располагается за пределами жизни одаренной личности и, по сути, не имеет ограничений во времени.

Метод нейросетевого кластерного (типологического) анализа

В рамках исследования стояла задача выявления типологии профессиональных траекторий одаренных личностей, основываясь на данных, собранных через контент-анализ биографий 100 выдающихся ученых. Целью было определить общие группы с аналогичными профессиональными путями и исследовать их уникальные черты. Для достижения этой цели мы выбрали метод, который уже успешно применялся в смежных областях социально-экономического исследования [24; 25; 26; 27], где он демонстрировал эффективность в типологизации сложных социальных феноменов. Такая задача формулируется следующим образом: как разбить на типовые группы массив объектов из сотни векторов, имеющих несколько вещественных интервальных координат. Отметим, что этот метод подходит для работы с большими объемами многомерных данных и не требует предварительных предположений о структуре данных. Это делает его удобным для анализа биографических данных, структура которых заранее неизвестна.

Для количественного описания траектории отдельно взятого ученого использована система из трех показателей, обоснованная выше поведенческой моделью и представленная вектором (x_{j1}, x_{j2}, x_{j3}) . Координаты этого вектора соотносятся с этапами

жизненного пути одаренной личности, которые были представлены на рис. 1. Первая координата x_{j1} отражает уровень стартовых условий ученого и может принимать числовые значения на вещественном отрезке (1; 4), причем единица соответствует трудным стартовым условиям, когда родители имеют невысокий уровень образования и финансового достатка, а 4 – самым благоприятным условиям, когда семья образована и имеет высокий уровень финансовой состоятельности.

Вторая координата x_{j2} связана с этапом обучения и становления ученого, который проиллюстрирован на рис. 3. Эта координата в общем случае может принимать порядковые значения 1, 2 и 3 на вещественном отрезке (1; 3), причем единица соответствует случаю отсутствия профильного высшего образования в сфере успеха одаренной личности, а 3 – случаю, когда ученый добился успеха в сфере своего первого высшего профессионального образования, то есть, другими словами, рост числового значения от 1 до 3 отражает все более быструю и эффективную профессиональную ориентацию одаренной личности.

Третья координата x_{j3} отражает условия третьего этапа профессиональной реализации ученого, который проиллюстрирован на рис. 4. Она может принимать порядковые значения на отрезке (1; 4), причем единица соответствует случаю самостоятельного поиска ученого по одному из традиционных направлений науки, а 4 – случаю его работы в рамках научной школы, разрабатывающей инновационные научные направления. Таким образом, рост численного значения этой координаты отражает уровень инновационности исследований одаренной личности и роль коллектива, командной работы в достижении успеха (в том числе наличие научной школы, наставника).

Кроме этих трех основных координат каждый вектор имеет три информационных показателя (эти показатели не участвуют в процессе дальнейшей кластеризации): ФИО – фамилия (имя, отчество) ученого; SR – скорость его успеха, которая является вещественным

показателем и исчисляется в годах до первого значимого достижения; сектор науки, в котором ученый добился успеха. Для иллюстрации векторизации профессиональных траекторий рассмотрим примеры, представленные в табл. 1. Аналогичным образом были «оцифрованы» биографии других ученых. Все векторы были сведены в единый массив данных, состоящий из 100 объектов.

Таблица 1

ПРИМЕРЫ ВЕКТОРОВ ПРОФЕССИОНАЛЬНЫХ
ТРАЕКТОРИЙ УЧЕНЫХ

ФИО	x_{j_1}	x_{j_2}	x_{j_3}	SR	Сектор науки
Бор Н.	4,0	3,0	4,0	31	ЕН
Даль В.И.	2,0	1,0	2,0	29	ГУМ

Выявление групп однотипных профессиональных траекторий проводилось путем кластеризации полученного набора данных, причем число кластеров (типов) и их характеристики заранее не были известны.

Задачу кластеризации можно сформулировать следующим образом. Дано множество профессиональных траекторий ученых $S = \{s_1, s_2, \dots, s_p, \dots, s_n\}$, $n = 100$, каждая из которых представлена вектором $x_j = \{x_{j_1}, x_{j_2}, x_{j_3}\}$, $j = 1, 2, \dots, n$. Требуется построить множество кластеров C и отображение F такое, что $F : S \rightarrow C$, где

$$C = \{c_1, \dots, c_k, \dots, c_m\}, \rightarrow m = 2, 3, 4, 5, 6, 7, 8;$$

c_k – кластер, содержащий однотипные траектории из множества S :

$$c_k = \{s_i, s_j | s_i \in S, s_j \in S \text{ и } d(s_i, s_j) < \sigma\}.$$

Здесь σ – критерий близости траекторий, $d(s_i, s_j)$ – мера близости между траекториями. В случае $d(s_i, s_j) < \sigma$ траектории помещаются в один кластер, а если $d(s_i, s_j) \geq \sigma$, то они попадают в разные кластеры (типы). При этом характер распределения объектов в трехмерном пространстве вещественных координат заранее не известен.

Для решения такой задачи в общем случае можно применять классические алгоритмы машинного обучения (k -средних, DBSCAN или иерархическую кластеризацию), а также нейросетевые подходы (самоорганизующиеся карты или сети Кохонена). Классические алгоритмы не являются универсальными, эффективность их применения зависит от характера распределения объектов (векторов) массива данных в многомерном пространстве состояний. Сравнительный анализ алгоритмов и их недостатки неоднократно освещались в обзорных публикациях как российских [28], так и зарубежных [29; 30] исследователей. Как следует из предыдущих исследований, метод k -средних весьма критичен к выбору координат центроидов и их количеству. Он хорошо работает только в случае группировки объектов в пространстве в плотные сгустки. DBSCAN зависим от выбора корневых объектов, игнорирует выбросы (шумы) и подходит для выявления кластеров на базе компактных группировок объектов специфической (кольцевой или ленточной) конфигурации. Алгоритм иерархической кластеризации, объединяющий на каждом шаге в один кластер два ближайших объекта (кластера), критичен к выбору функции близости, а самое главное, не дает ответа на основной вопрос об оптимальном количестве кластеров (типов) объектов в изучаемом массиве данных. Таким образом, для применения того или иного классического алгоритма необходимо предварительное изучение распределения объектов в пространстве, что легко сделать для случая двухмерного пространства. Для более сложных случаев требуется применение методов понижения размерности пространств, что снижает точность анализа и повышает вероятность ошибок при выборе метода кластеризации.

В свою очередь, при использовании метода нейросетевого кластерного анализа происходит несколько разбиений, которые анализируются с точки зрения критериев качества кластеризации. В настоящем исследовании применялись два критерия. Первый определяется как минимум сумм ошибок квантования (подробнее об этом критерии можно найти в предыдущих работах членов

творческого коллектива [25; 31]). Критерий был рассчитан в программной среде Deductor (версия Deductor Academic 5.3). В ней же проводились все вычислительные эксперименты с искусственными нейронными сетями (ИНС). Первый критерий, таким образом, позволил настроить начальные условия процесса моделирования сети и вычислить средние и максимальные ошибки квантования, которые нужны для выбора лучшей карты. Вторым критерий предполагает расчет коэффициентов силуэта [28, с. 46]. Выбирая разбиение с максимальным коэффициентом силуэта, можно определить оптимальное количество кластеров. Наилучшее разбиение, таким образом, подтверждается сразу двумя критериями.

Метод нейросетевого кластерного анализа не требует предварительной подготовки данных и изучения распределения объектов в пространстве. Применение самоорганизующихся карт Кохонена [32; 33], которые представляют собой специальные искусственные нейронные сети, позволяет в автоматическом режиме группировать объекты исходного датасета и формировать кластеры, состоящие из однотипных векторов. Для получения достоверного результата процесс кластеризации повторяется многократно с разными начальными условиями. После завершения вычислительных экспериментов выбирается лучшая карта с использованием критерия минимума сумм ошибок квантования.

Таким образом, в рассматриваемом случае кластеризации профессиональных траекторий ученых, представленных векторами в трехмерном пространстве вещественных координат, целесообразно использовать именно нейросетевой кластерный анализ, поскольку предварительно не известно количество типовых групп и не ясен характер распределения объектов в пространстве. Кроме того, в этом случае не требуется никакой предварительной подготовки данных.

Ход исследования

Предварительный анализ позволил сделать несколько общих замечаний относительно всей выборки исследования. Средним

возрастом успеха по всей выборке стали 33,5 года, при этом персоналии гуманитарного профиля имели чуть более длинный путь, чем естественно-научного направления (рис. 4).

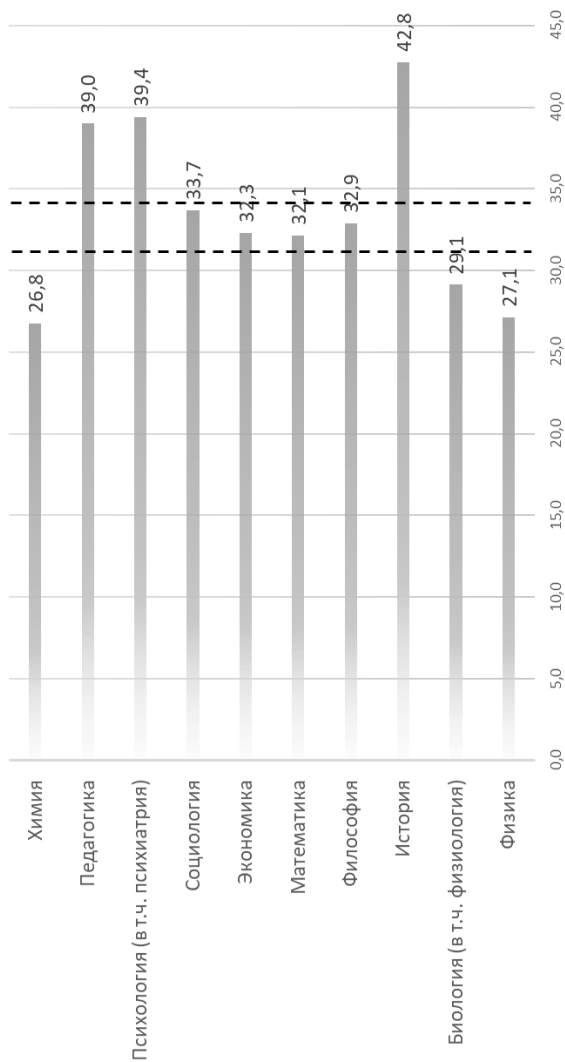
Рассмотрим распределение количества биографий в соответствии с заданной ранее матрицей классификации жизненных условий, в которой можно отразить либо низкие, либо высокие значения факторов (рис. 2А, 2Б, 3). Например, на рис. 5 сегмент 1.0 соответствует низкому уровню образования и мотивации к действию близкого круга одаренного и низкому уровню финансовой обеспеченности семьи. К сегменту 1.1 отнесем тех одаренных, у которых низкий уровень финансовой обеспеченности родителей сочетался с их высокой заинтересованностью в развитии доминанты оригинальности ребенка, либо в обратном сочетании (сегмент 1.1). В сегмент 1.3 вошли биографии ученых, для которых были созданы наиболее благоприятные стартовые условия.

Анализ данных, представленных в виде предложенной матрицы условий (рис. 5), позволяет предположить, что большая часть одаренных на старте жизненного пути либо имела удачное сочетание финансовой обеспеченности и образованности своего окружения (сегмент 1.3, рис. 5), либо не имела никакой финансовой и мотивационной поддержки от близкого окружения (сегмент 1.0, рис. 5).

По мере усложнения социального института науки добиться значительных успехов без качественного образования стало сложнее (рис. 2Б – рис. 6).

Если в XIX в. одаренные личности достигали успеха и без образования, и без ранней профориентации (сегменты 2.0 и 2.1, рис. 6), то в XX в. преимущество за теми из ученых, кто сразу определился с профильным образованием и сумел уклониться от «метаний».

Что касается коллективного фактора (рис. 3 – рис. 7), то ожидаемо его роль имела решающее значение для научной сферы (сегмент 3.2, рис. 7). Лишь немногим удался полностью самостоятельный профессиональный путь.



Естественно-научный профиль: 31,2 г. Гуманитарный профиль: 34,6 г.
Рис. 4. Средний «возраст успеха» представителей естественно-научного и гуманитарного профиля выборки ученых (100 чел., XIX, XX вв.), годы

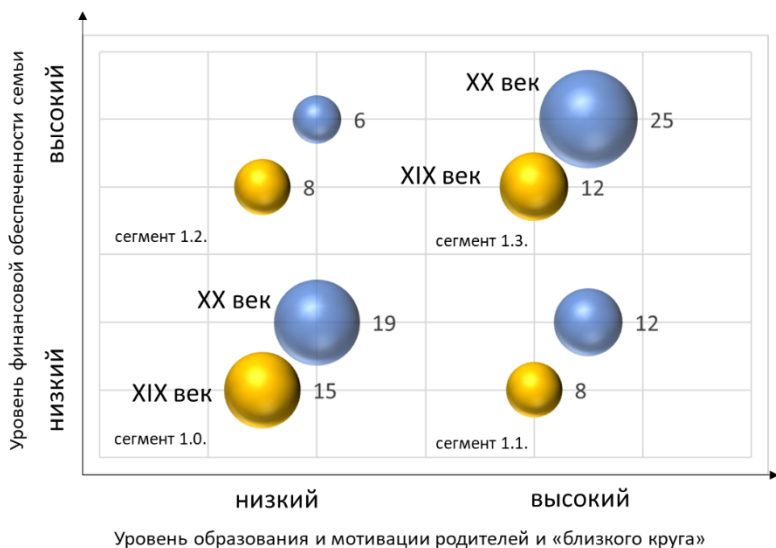


Рис. 5. Матричный анализ по 4 группам стартовых условий

Примечание. Здесь, на рис. 5, и далее на рис. 6–7 размер шара соответствует количеству биографий, которые могут быть отнесены к тем или иным сегментам предложенной классификационной матрицы, а цвет относится к веку: желтый – XIX в., синий – XX в.

Для построения типологии профессиональных траекторий нейросетевой кластерный анализ проводился следующим образом. Весь массив векторов «оцифрованных» биографий ученых разбивался последовательно на 2, 3, 4, 5, 6, 7 и 8 кластеров. На каждом этапе было проведено по 12 вычислительных экспериментов с разными значениями гиперпараметров алгоритма самоорганизации нейронной сети. Заключительный этап продемонстрировал наличие кластера нулевой мощности, что сигнализировало о формировании семи стабильных групп и завершении всей процедуры. Далее, из всех синтезированных карт было выбрано лучшее разбиение на базе критерия минимума сумм ошибок квантования, вычисляемых в программе Deductor. Ошибки квантования характе-

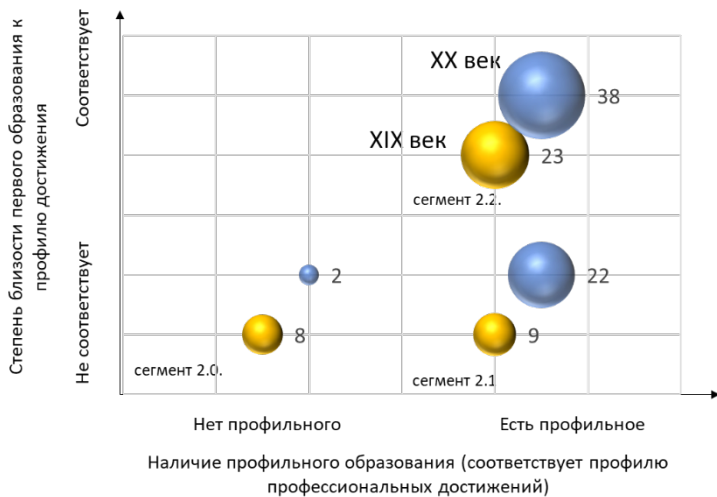


Рис. 6. Матричный анализ по 3 группам этапа «Обучение и становление»

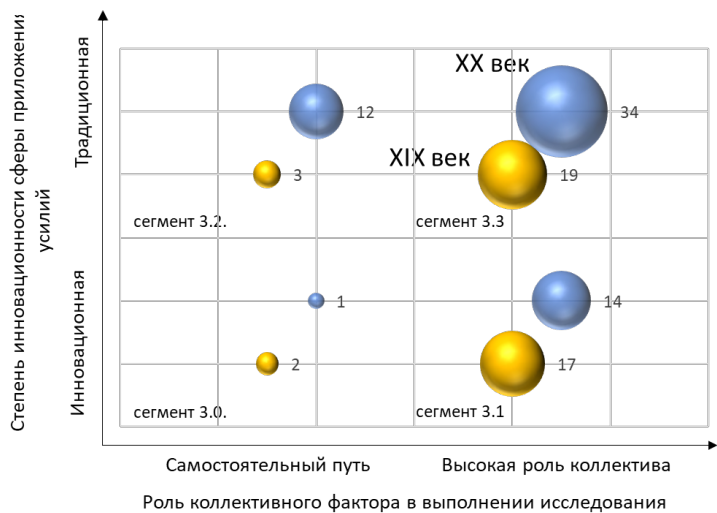


Рис. 7. Матричный анализ по 4 группам этапа «Профессиональная реализация»

ризуют отклонения конкретных векторов кластера от центра соответствующих ячеек карты Кохонена: чем меньше эти ошибки, тем качественнее кластеризация [24]. Сформированные кластеры были проанализированы с точки зрения их состава и характеристик.

Обсуждение полученных результатов

Содержательный анализ составов полученных кластеров и их количественных характеристик, приведенных в табл. 2, позволил выявить характерные типы профессиональных траекторий.

Таблица 2

КОЛИЧЕСТВЕННЫЕ ХАРАКТЕРИСТИКИ КЛАСТЕРОВ

№/цвет	МК	Характеристики кластера	X_1	X_2	X_3	SR	ГУМ	ЕН
1 Голубой	9	Среднее значение	3,11	3	2,44	35,11	1	8
		Стандартное отклонение	0,33	0	0,88			
2 Оранжевый	13	Среднее значение	1,54	2,77	2,15	35,85	9	4
		Стандартное отклонение	0,52	0,44	0,55			
3 Серый	20	Среднее значение	3,75	1,7	3,65	32,2	15	5
		Стандартное отклонение	0,44	0,47	0,49			
4 Желтый	15	Среднее значение	3,67	3	4	31,53	6	9
		Стандартное отклонение	0,49	0	0			
5 Синий	20	Среднее значение	1,15	3	3,7	34,55	10	10
		Стандартное отклонение	0,37	0	0,47			

Окончание табл. 2

№/цвет	МК	Характеристики кластера	X_1	X_2	X_3	SR	ГУМ	ЕН
6 Зеленый	15	Среднее значение	1,27	1,8	3,67	32,67	10	5
		Стандартное отклонение	0,46	0,41	0,62			
7 Темно-синий	8	Среднее значение	4	2,88	2,5	28,75	4	4
		Стандартное отклонение	0	0,35	0,53			

Примечание. В этой таблице приняты следующие обозначения: №/цвет – номер кластера и цвет его линии на рис. 8; МК – мощность кластера (сумма числа ученых, чьи биографии отнесены к кластеру); SR – скорость успеха, лет; X_1 ; X_2 ; X_3 – вещественные интервальные переменные, отражающие уровни трех основных факторов стадиальной модели жизненного пути ученого (рис. 1); ГУМ – число ученых-гуманитариев, попавших в кластер; ЕН – число ученых-естествоиспытателей, попавших в кластер.

На основе табл. 2 построены лепестковые диаграммы. Они дают возможность визуализировать полученные кластеры в координатах, соответствующих факторам модели (семья, образование, профессиональное становление (реализация)). Цвета линий, представленных на рис. 8, соотнесены с кластерами в табл. 2 в графе №/цвет.

Первый кластер («Инноваторы традиций») характеризуется хорошими стартовыми условиями, точным самоопределением. Большинство представителей этой группы достигли успеха в естественно-научной сфере. Как правило, на жизненном пути одаренные из этого кластера отличались большой самостоятельностью (например, А. Эйнштейн или Э. Шрёдингер).

Во **втором** кластере («Одиссеи науки») – преимущество за гуманитариями. Представители этого кластера имели непростые стартовые условия, но отличались быстрым самоопределением

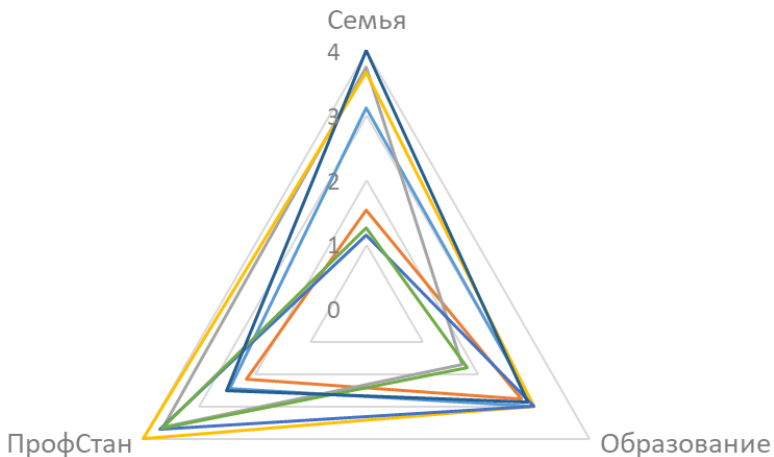


Рис. 8. Лепестковые диаграммы кластеров

на этапе обучения. Однако сфера приложения их усилий не отличается высокой степенью инновационности и высокой научной коллаборацией. Здесь мы найдем много «одиночек», самостоятельно сформировавшихся как профессионалы в науке. Средняя скорость успеха – почти 36 лет, что существенно выше среднего значения по всей выборке. Например, в этот кластер попала биография Ф. Ницше.

Особенностью представителей **третьего** кластера («Пионеры») является длинный этап профессионального самоопределения, связанный с поиском сферы приложения своих талантов. Здесь перевес представителей гуманитарных наук. Однако в отличие от второго кластера, представители данного кластера брались за исследования нового направления. Поэтому, как правило, они сами становились основателями научных школ. Учитывая, что у вошедших в этот кластер ученых была хорошая стартовая семейная поддержка, им удалось добиться успеха, хоть путь самоопределения у многих был непростой. Например, сюда была отнесена биография Ч. Дарвина. Он происходил из обеспеченной

и уважаемой семьи, что дало ему определенную творческую и академическую свободу. Другим примером может быть биография М. Вебера: он происходил из семьи успешного юриста и политика.

Попавших в **четвертый** кластер можно характеризовать как «круглых отличников», «Моцартов науки». У них хороший старт и точное самоопределение, сфера применения их усилий отличается высокой инновационностью. Не случайно средний возраст успеха в этом кластере 31,5 года. Большинство представителей этого кластера работали в сфере точных и естественных наук. Например, в этот кластер попадают биографии Н. Бора и М. Борна.

В **пятый** кластер («Стойкие экспериментаторы») попали ученые, чьи биографии отличались самыми сложными стартовыми условиями, которые им приходилось восполнять «удачей» и трудолюбием на втором и третьем этапах. Здесь много экспериментаторов, например – Д. Менделеев, А. Макаренко. Профессиональный путь таких людей требует времени и большого трудолюбия, но достижения имеют инновационное значение для отрасли.

В **шестом** кластере сосредоточены «Загадочные таланты»: их путь представляется самым неопределенным с точки зрения условий для успеха. Для них характерны сложные стартовые условия и низкая скорость самоопределения в период обучения и становления. Вторая учеба, или второстепенная работа, «мешала» сосредоточиться на доминанте оригинальности. При этом их включение в научную коммуникацию ничем не уступает «экспериментаторскому» пути. Изучение биографий этого кластера дает основание для гипотезы о длинном поиске «себя» в связи с разносторонними талантами (задатками). Например, к этому кластеру относится Л. Пастер, который успел поработать репетитором, администратором в сфере образования, и даже размышлял о карьере художника. Однако страсть к химии в итоге победила другие интересы и занятия. В биографии К.Д. Ушинского встречается история о случайно найденном архиве инспектора Гатчинского сиротского института, Е.О. Гугеля, в котором, как писал сам

Ушинский, он нашел «полное собрание педагогических книг».

Представители **седьмого** кластера («Прорывные мастера») отличаются самыми лучшими стартовыми условиями и самым коротким временем до первого успеха – всего 28,75 года. Это, скорее всего, связано с рано проявленной гениальностью (Галенсон), высокой оригинальностью работ еще на ранних стадиях самоопределения. Здесь поровну представителей обоих научных направлений – и «физиков», и «лириков». Сюда были отнесены биографии М. Планка, Л. Ландау, а из гуманитариев, например, Ж.-П. Сартра.

Сделаем несколько обобщающих замечаний.

1. Высокая скорость достижения успеха в науке в первую очередь зависит от стартовых условий (семейного благосостояния вкпе с образованностью семьи). 3-й, 4-й и особенно 7-й кластеры показывают сильную связь между факторами первой группы и «средним возрастом» успеха.

2. Правильный выбор научной школы (мирового научного центра) может компенсировать «низкий старт». В этом результаты исследования согласуются с выводом о важной роли мобильности в профессиональной карьере [17;18].

3. Представители естественно-научной сферы, как правило, быстрее достигали успеха, чем гуманитарии. Их отличают более высокие показатели самоопределения на этапе «Образование и становление». Это может быть обусловлено недостаточной развитостью мировых центров гуманитарной науки, сильной фрагментацией гуманитарного научного знания, слабой преемственностью школ и наставничества.

Одним из основных ограничений исследования является его ориентированность на конкретную выборку выдающихся ученых XIX и XX вв. Выбранное временное окно может не учитывать факторы и изменения, которые могли бы повлиять на профессиональные траектории в современных условиях. Ограничением является также допущение о роли формального образования в успехе

одаренной личности: в течение XIX–XX вв. институт образования претерпевал существенные изменения. Другим ограничением является качество данных. Результаты контент-анализа зависят от доступности и точности биографических материалов, а также выбранных источников информации. Метод нейросетевого анализа с самоорганизующимися картами Кохонена представляет собой новый для этой сферы исследовательский подход, однако следует помнить, что результаты могут быть зависимы от параметров и начальных условий.

Выводы

В исследовании с помощью нейросетевого анализа на выборке в 100 биографий выдающихся личностей из науки XIX и XX вв. были выявлены и интерпретированы семь траекторий. Было обосновано, что для поставленной задачи целесообразно использовать именно нейросетевой кластерный анализ, поскольку предварительно не известно количество типовых групп и не ясен характер распределения объектов в пространстве. Кроме того, в этом случае не требуется никакой предварительной подготовки данных.

Анализ траекторий проводился с точки зрения скорости достижения успеха (среднего возраста успеха) и тех факторов и условий жизненного пути, которые могли повлиять на более быстрое или медленное достижение профессиональных целей и самореализацию одаренной личности. На выборке исследования были сформулированы гипотезы, открывающие возможности для дальнейших исследований с использованием разработанной методологии контент-анализа, а также развития опыта применения нейросетевого анализа. Ограничения метода заключаются в повышенных требованиях, которые предъявляются к вычислительным мощностям при проведении вычислительных экспериментов, и в возможных проблемах с шумом и выбросами. Однако в контексте поставленной задачи поиска конечного числа типовых про-

фессиональных траекторий одаренных личностей нейросетевой анализ показал свои лучшие стороны, так как позволил работать со сложными формами кластеров и сосредоточиться на поиске их оптимального числа.

ЛИТЕРАТУРА

1. Mayer K., Pfeiffer J. Computational Social Science // Schlüsselwerke der Netzwerkforschung / Ed. by B. Holzer, C. Stegbauer. Wiesbaden: Springer VS, 2019. P. 721–723. DOI: 10.1007/978-3-658-21742-6_77.

2. Beytia P., Schobin J. Networked pantheon: a relational database of globally famous people: Social and behavioural sciences // Research Data Journal for the Humanities and Social Sciences. 2020. Vol. 5, № 1. P. 50–65. DOI: 10.17632/twvsjygw3m.1.

3. Chisholm A., Radford W., Hachey B. Learning to generate one-sentence biographies from Wikidata // Cornwall University [site]. 21.02.2017. URL: arXiv preprint arXiv:1702.06235. (дата обращения: 01.02.2023).

4. Reznik I., Shatalov V. Hidden revolution of human priorities: An analysis of biographical data from Wikipedia // Journal of Informetrics. 2016. Vol. 10, № 1. P. 124–131. DOI: 10.1016/j.joi.2015.12.002.

5. Samoilenko A., Yasseri T. The distorted mirror of Wikipedia: a quantitative analysis of Wikipedia coverage of academics // EPJ Data Science. 2014. Vol. 3, № 1. P. 1–11. DOI: 10.1140/epjds20.

6. Beyond One-Dimensional Portraits: A Synoptic Approach to the Visual Analysis of Biography Data / F. Windhager, M. Schlögl, M. Kaiser [et al.] // Conference: Biographical Data in a Digital World 2017 (BD 2017). Volume: CEUR Vol-2119. Linz, Austria, 2017. P. 67–75.

7. Collison P., Nielsen M. Science is getting less bang for its buck // The Atlantic [site]. 16.11.2018. URL: <https://www.theatlantic.com/science/archive/2018/11/diminishing-returns-science/575665/> (дата обращения: 19.01.2023).

8. Sternberg R.J. Identification for utilization, not merely possession, of gifts: What matters is not gifts but rather deployment of gifts // Gifted Education International. 2022. Vol. 38, № 3. P. 354–361. DOI: 10.1177/02614294211013345.

9. Леонтьев Д.А., Лебедева А.А., Костенко В.Ю. Траектории личностного развития: реконструкция взглядов Л.С. Выготского // Вопросы образования. 2017. № 2. С. 98–112. DOI: 10.17323/1814-9545-2017-2-98-112. EDN: YUPYIX.

10. Шадриков В.Д. Отношение понятий «жизнь», «поведение», «деятельность» // Мир психологии. 2020. № 2(102). С. 57–65. EDN: BCDJMW.

11. Galenson D.W. Old Masters and Young Geniuses: The Two Life Cycles of Human Creativity // Journal of Applied Economics. 2009. Vol. 12, № 1. P. 1–9. DOI:

10.1016/S1514-0326(09)60002-7.

12. *Weinberg B.A., Galenson D.W.* Correction to: Creative Careers: The Life Cycles of Nobel Laureates in Economics // *De Economist*. 2019. Vol. 167, № 3. P. 241–241. DOI: 10.1007/s10645-019-09342-0.

13. *Кольцова О.Ю., Маслинский К.А.* Выявление тематической структуры российской блогосферы: автоматические методы анализа текстов // *Социология: методология, методы, математическое моделирование (Социология: 4М)*. 2013. № 36. С. 113–139. EDN: RCFOWJ.

14. *Ким А.В., Мальцева Д.В., Щеглова Т.Е.* Блокмоделинг для анализа социальных структур: пример изучения структуры сообщества петербургских социологов // *Социология: методология, методы, математическое моделирование (Социология: 4М)*. 2021. № 53. С. 7–38. DOI: 10.19181/4m.2021.53.1. EDN: HYNUSK.

15. The effects of outliers' data on neural network performance / A. Khamis, Z. Ismail, K. Haron, A. Tarmizi // *Journal of Applied Sciences*. 2005. Vol. 5, № 8. P. 1394–1398. DOI: 10.3923/jas.2005.1394.1398.

16. *Mijwel M.M.* Artificial neural networks advantages and disadvantages // *Mesopotamian Journal of Big Data*. 2021. Vol. 2021. P. 29–31. DOI: 10.58496/mjbd/2021/006.

17. *Schlagberger E.M., Bornmann L., Bauer J.* At what institutions did Nobel laureates do their prize-winning work? An analysis of biographical information on Nobel laureates from 1994 to 2014 // *Scientometrics*. 2016. Vol. 109, № 2. P. 723–767. DOI: 10.1007/s11192-016-2059-2.

18. *Lucchini L., Tonelli S., Lepri B.* Following the footsteps of giants: modeling the mobility of historically notable individuals using Wikipedia // *EPJ Data Science*. 2019. № 8. P. 36. DOI: 10.1140/epjds/s13688-019-0215-7.

19. *Cui H., Wu L., Evans J.A.* Aging Scientists and Slowed Advance // *Cornwall University [site]*. 08.02.2022. URL: arXiv preprint arXiv:2202.04044. 2022 (дата обращения: 20.02.2023).

20. *Jones B.F., Reedy E.J., Weinberg B.A.* Age and scientific genius // *The Wiley handbook of genius / Ed. by D. K. Simonton*. Hoboken: John Wiley & Sons, Ltd, 2014. P. 422–450. DOI: 10.1002/9781118367377.ch20.

21. Developmental biographies of Olympic super-elite and elite athletes: A multidisciplinary pattern recognition analysis / A. Güllich, L. Hardy, L. Kuncheva [et al.] // *Journal of Expertise*. 2019. Vol. 2 (1). P. 23–46.

22. *Letiagina E., Perova V., Orlova E.* Neural network analysis of the development of physical education and sports in Russia as an economic factor of country security // *Proceedings of the 4th International Conference on Innovations in Sports, Tourism and Instructional Science (ICISTIS 2019)*. Atlantis Press. 2019. № 11. P. 138–142. DOI: 10.2991/icistis-19.2019.37.

23. *Subotnik R.F., Olszewski-Kubilius P., Worrell F.C.* Environmental factors and personal characteristics interact to yield high performance in domains // *Frontiers in*

Psychology. 2019. № 10:2804. P. 1–8. DOI: 10.3389/fpsyg.2019.02804.

24. Трифонов Ю.В., Сочков А.Л., Соловьев А.Е. Оценка экономического потенциала регионов РФ на основе методологии нейросетевого кластерного анализа // Вестник Нижегородского университета им. Н.И. Лобачевского. Серия: Социальные науки. 2021. № 3(63). С. 38–47. DOI: 10.52452/18115942_2021_3_38. EDN: SIQALB.

25. Трифонов Ю.В., Сочков А.Л., Куликова А.В. Построение и реализация моделей интеллектуальных цифровых коммуникаций в социально-политических сферах // Экономика и предпринимательство. 2021. № 8(133). С. 1087–1095. DOI: 10.34925/EIP.2021.133.8.209. EDN: UGMSUS.

26. Carboni O.A., Russu P. Assessing Regional Wellbeing in Italy: An Application of Malmquist–DEA and Self-organizing Map Neural Clustering // Social Indicators Research. 2015. Vol. 122, № 3. P. 677–700. DOI: 10.1007/s11205-014-0722-7.

27. Regional disaster risk assessment of China based on self-organizing map: Clustering, visualization and ranking / N. Chen, L. Chen, Y. Ma, A. Chen // International Journal of Disaster Risk Reduction. 2019. № 33. P. 196–206. DOI: 10.1016/j.ijdr.2019.101085.

28. Абдурахманова Э.М. Исследование структур обобщенных групп, выделяемых разными методами, на примере результатов исследования СТАРТ // Социология: методология, методы, математическое моделирование (Социология: 4М). 2020. № 50–51. С. 37–63. EDN: MEPMGD.

29. Gülagiz F.K., Sahin S. Comparison of hierarchical and non-hierarchical clustering algorithms // International Journal of Computer Engineering and Information Technology. 2017. Vol. 9, № 1. P. 6–14.

30. Musdholifah A., Hashim S.Z.M., Zaiton S. Cluster analysis on high-dimensional data: A comparison of density-based clustering algorithms // Australian Journal of Basic and Applied Sciences. 2013. Vol. 7, № 2. P. 380–389.

31. Research of the innovative development of the Russian Federation regions and its impact on the eco-friendliness of the economy based on neural network cluster analysis for the purpose of economic security / S. Yashin, Y. Trifonov, A. Sochkov [et al.] // E3S Web of Conferences. 2021. Vol. 291. P. 1–10. DOI: 10.1051/e3sconf/202129103008.

32. Kohonen T. The self-organizing map // Proceedings of the IEEE. Vol. 78, № 9. P. 1464–1480. DOI: 10.1109/5.58325.

33. Engineering applications of the self-organizing map / T. Kohonen, E. Oja, O. Simula [et al.] // Proceedings of the IEEE. Vol. 84, № 10. P. 1358–1384. DOI:10.1109/5.537105.

Chepyuk Olga R.,

Doctor of Philosophy, Professor of the Department of Human Resource Management, National Research Lobachevsky State University, Nizhny Novgorod, Russia, chepyuko@yandex.ru

Angelova Olga Yu.,

Candidate of Economic Sciences, Associate Professor of the Department of Information Technologies and Instrumental Methods in Economics, National Research Lobachevsky State University, Nizhny Novgorod, Russia, oangelova@mail.ru

Sochkov Andrey L.,

Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and Instrumental Methods in Economics, National Research Lobachevsky State University, Nizhny Novgorod, Russia, sochkov@ice.unn.ru

Podolskaya Tatyana O.,

Candidate of Sociological Sciences, Associate Professor of the Department of Human Resource Management, National Research Lobachevsky State University, Nizhny Novgorod, Russia, podolskaya@ice.unn.ru

Typology of professional trajectories of gifted individuals using neural network analysis

Based on a data set (100 biographies) created by the authors through content analysis of biographical material about outstanding scientists of the 19th and 20th centuries in the humanities and natural sciences, the clustering of professional trajectories of gifted individuals was carried out. Neural network analysis based on self-organizing Kohonen maps was used as a clustering method. The professional trajectories were formed within the framework of the behavioral model of the linear-stage approach to studying life cycles. Within this approach, career and professional self-realization are understood as a sequence of evolutionary stages fixed in their order of occurrence. Each stage was encoded, and the biographies were transformed into a vector system. In turn, the task of clustering consisted in dividing a hundred vectors into typical groups with several real-valued coordinates. The criteria for the quality of clustering were the minimum sum of quantization errors and the silhouette coefficient. As a result of the study, seven professional trajectories

of gifted individuals were identified and interpreted. The analysis of trajectories was carried out from the point of view of the speed of success (average age of success) and those factors and conditions of the life path that could affect either rapid or slow achievement of professional goals and self-realization. This example demonstrates the possibilities and limitations of using neural network analysis for solving similar research tasks, especially when working with complex cluster forms and finding their optimal number. *Keywords:* neural network analysis, giftedness, gifted personality, professional trajectory, machine learning, neural network, Kohonen maps

References

1. Mayer K., Pfeffer J. *Computational Social Science*. Wiesbaden: Springer VS, 2019, P. 721–723. DOI: 10.1007/978-3-658-21742-6_77.
2. Beytía P., Schobin J. Networked pantheon: a relational database of globally famous people: Social and behavioural sciences, *Research Data Journal for the Humanities and Social Sciences*, 2020, vol. 5, no. 1, p. 50–65. DOI: 10.17632/twvsjygw3m.1.
3. Chisholm A., Radford W., Hachey B. Learning to generate one-sentence biographies from Wikidata, *Cornwall University* [site]. 21.02.2017, URL: arXiv preprint arXiv:1702.06235. (date of the application: 01.02.2023).
4. Reznik I., Shatalov V. Hidden revolution of human priorities: An analysis of biographical data from Wikipedia, *Journal of informetrics*, 2016, vol. 10, no. 1, p. 124–131. DOI: 10.1016/j.joi.2015.12.002.
5. Samoilenko A., Yasseri T. The distorted mirror of Wikipedia: a quantitative analysis of Wikipedia coverage of academics, *EPJ data science*, 2014, vol. 3, no. 1, p. 1–11. DOI: 10.1140/epjds20.
6. Windhager F., Schlögl M., Kaiser M. et al. Beyond One-Dimensional Portraits: A Synoptic Approach to the Visual Analysis of Biography Data, *Conference: Biographical Data in a Digital World 2017* (BD 2017), Vol.: CEUR Vol-2119, Linz, Austria, 2017. P. 67–75.
7. Collison P., Nielsen M. Science is getting less bang for its buck, *The Atlantic* [site]. 16.11.2018. URL: <https://www.theatlantic.com/science/archive/2018/11/diminishing-returns-science/575665/> (date of access: 19.01.2023).
8. Sternberg R.J. Identification for utilization, not merely possession, of gifts: What matters is not gifts but rather deployment of gifts,

- Gifted Education International*, 2022, vol. 38, no. 3, p. 354–361. DOI: 10.1177/02614294211013345.
9. Leontyev D.A., Lebedeva A.A., Kostenko V.Yu. Trajectories of personal development: reconstruction of the views of L.S. Vygotsky (in Russian), *Issues of education*, 2017, no. 2, p. 98–112. DOI: 10.17323/1814-9545-2017-2-98-112.
 10. Shadrikov V.D. The relationship of the concepts “life”, “behavior”, “activity” (in Russian), *World of Psychology*, 2020, vol. 2, no. 102, p. 57–65.
 11. Galenson D.W. Old Masters and Young Geniuses: The Two Life Cycles of Human Creativity, *Journal of Applied Economics*, 2009, vol. 12, no. 1, p. 1–9. DOI: 10.1016/S1514-0326(09)60002-7.
 12. Weinberg B.A., Galenson D.W. Correction to: Creative Careers: The Life Cycles of Nobel Laureates in Economics, *De Economist*, 2019, vol. 167, no. 3, p. 241–241. DOI: 10.1007/s10645-019-09342-0.
 13. Koltsova O.Yu., Maslinsky K.A. Revealing the thematic structure of the Russian blogosphere: automatic methods of text analysis (in Russian), *Sotsiologiya 4M (Sociology: methodology, methods, mathematical modeling)*, 2013, no. 36, p. 113–139.
 14. Kim A.V., Maltseva D.V., Shcheglova T.E. Block modeling for the analysis of social structures: an example of studying the structure of a community of St. Petersburg sociologists (in Russian), *Sotsiologiya 4M (Sociology: methodology, methods, mathematical modeling)*, 2021, no. 53, p. 7–38. DOI: 10.19181/4m.2021.53.1.
 15. Khamis A., Ismail Z., Haron K., Tarmizi A. The effects of outliers’ data on neural network performance, *Journal of Applied Sciences*, 2005, vol. 5, no. 8, p. 1394–1398. DOI: 10.3923/jas.2005.1394.1398.
 16. Mijwel M.M. Artificial neural networks advantages and disadvantages, *Mesopotamian Journal of Big Data*, 2021, vol. 2021, p. 29–31. DOI: 10.58496/mjbd/2021/006.
 17. Schlagberger E.M., Bornmann L., Bauer J. At what institutions did Nobel laureates do their prize-winning work? An analysis of biographical information on Nobel laureates from 1994 to 2014, *Scientometrics*, 2016, vol. 109, no. 2, p. 723–767. DOI: 10.1007/s11192-016-2059-2.
 18. Lucchini L., Tonelli S., Lepri B. Following the footsteps of giants: modeling the mobility of historically notable individuals using Wikipedia, *EPJ Data Science*, 2019, no. 8, p. 36. DOI: 10.1140/epjds/s13688-019-0215-7.

19. Cui H., Wu L., Evans J.A. Aging Scientists and Slowed Advance, *Cornwall University [site]*. 08.02.2022. URL: arXiv preprint arXiv:2202.04044, 2022 (date of access: 20.02.2023).
20. Jones B.F., Reedy E.J., Weinberg B.A. Age and scientific genius, *The Wiley handbook of genius*, Hoboken: John Wiley & Sons, Ltd, 2014, p. 422–450. DOI: 10.1002/9781118367377.ch20.
21. Gullich A., Hardy L., Kuncheva L. et al. Developmental biographies of Olympic super-elite and elite athletes: A multidisciplinary pattern recognition analysis, *Journal of Expertise*, 2019, vol. 2, no. 1, p. 23–46.
22. Letiagina E., Perova V., Orlova E. Neural network analysis of the development of physical education and sports in Russia as an economic factor of country security, *Proceedings of the 4th International Conference on Innovations in Sports, Tourism and Instructional Science (ICISTIS 2019)*. Atlantis Press, 2019, № 11, P. 138–142. DOI: 10.2991/icistis-19.2019.37.
23. Subotnik R.F., Olszewski-Kubilius P., Worrell F.C. Environmental factors and personal characteristics interact to yield high performance in domains, *Frontiers in Psychology*, 2019, no. 10:2804, P. 1–8. DOI: 10.3389/fpsyg.2019.02804.
24. Trifonov Yu.V., Sochkov A.L., Soloviev A.E. Assessment of the economic potential of the regions of the Russian Federation based on the methodology of neural network cluster analysis (in Russian), *Bulletin of the Lobachevsky University. Series: Social Sciences*, 2021, 3 (63), p. 38–47. DOI: 10.52452/18115942_2021_3_38.
25. Trifonov Yu.V., Sochkov A.L., Kulikova A.V. Construction and implementation of models of intelligent digital communications in socio-political spheres (in Russian), *Economics and Entrepreneurship*, 2021, 8 (133), p. 1087–1095. DOI: 10.34925/EIP.2021.133.8.209.
26. Carboni O.A., Russu P. Assessing Regional Wellbeing in Italy: An Application of Malmquist–DEA and Self-organizing Map Neural Clustering, *Social Indicators Research*, 2015, vol. 122, no. 3, p. 677–700. DOI: 10.1007/s11205-014-0722-7.
27. Chen N., Chen L., Ma Y. , Chen A. Regional disaster risk assessment of China based on self-organizing map: Clustering, visualization and ranking, *International Journal of Disaster Risk Reduction*, 2019, no. 33, p. 196–206. DOI: 10.1016/j.ijdr.2019.101085.

28. Abdurakhmanova E.M. Study of the structures of generalized groups identified by different methods, using the results of the START study as an example (in Russian), *Sotsiologiya 4M (Sociology: methodology, methods, mathematical modeling)*, 2020, no. 50–51, p. 37–63.
29. Gülagiz F.K., Sahin S. Comparison of hierarchical and non-hierarchical clustering algorithms, *International Journal of Computer Engineering and Information Technology*, 2017, vol. 9, no. 1, P. 6–14.
30. Musdholifah A., Hashim S.Z.M., Zaiton S. Cluster analysis on high-dimensional data: A comparison of density-based clustering algorithms, *Australian Journal of Basic and Applied Sciences*, 2013, vol. 7, no. 2, p. 380–389.
31. Yashin S., Trifonov Y. , Sochkov A. et al. Research of the innovative development of the Russian Federation regions and its impact on the eco-friendliness of the economy based on neural network cluster analysis for the purpose of economic security, *E3S Web of Conferences*. 2021, vol. 291, p. 1–10. DOI: 10.1051/e3sconf/202129103008.
32. Kohonen T. The self-organizing map, *Proceedings of the IEEE*, vol. 78, no. 9, p. 1464–1480. DOI: 10.1109/5.58325.
33. Kohonen T., Oja E. , Simula O. et al. Engineering applications of the self-organizing map, *Proceedings of the IEEE*, vol. 84, no. 10, p. 1358–1384. DOI:10.1109/5.537105.