

Разработка модели формирования индивидуальных образовательных траекторий с использованием методов машинного обучения*

Е.С. Морозевич^a 

E-mail: ktyamorozevich@mail.ru

В.С. Коротких^b 

E-mail: vskor@bk.ru

Е.А. Кузнецова^a 

E-mail: zhenya.kuz-1997@yandex.ru

^a Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева
Адрес: Россия, 660037, г. Красноярск, проспект им. газеты Красноярский рабочий, д. 31

^b Муниципальное бюджетное общеобразовательное учреждение средняя общеобразовательная школа №4
Адрес: Россия, 663091, г. Дивногорск, ул. Набережная, д. 9

Аннотация

Сегодня экономика переживает цифровую трансформацию, ключевыми барьерами которой являются нехватка квалифицированных кадров, компетенций и знаний, а также внутреннее сопротивление в организациях. Преодолеть их можно через качественное развитие и обучение персонала. Актуальной задачей является построение индивидуальных образовательных траекторий. Современные исследования направлены на имплементацию рекомендательных систем в целях подбора релевантного материала. Однако основой таких рекомендаций являются цифровые следы; полный личностный профиль обучающегося, а также организационные ценности не учитываются. Целью исследования является создание интеллектуального проводника, который бы сопровождал сотрудника на протяжении всей жизни в организации, вовлекая его в процесс обучения согласно персонифицированной траектории на основе данных о комплексном личностном профиле и реакциях на образовательный материал, тренируя soft и hard skills в соответствии с ценностями организации и сотрудника. В исследовании использовались методы системного анализа, системной инженерии, психодиагностического исследования (модель DISC, методика Алана Роу «Стиль принятия решений», методика определения стилей деятельности Хони и Мамфорда, тест на психотип), проектирования программного обеспечения и искусственного интеллекта (матричная факторизация и нейронные сети). Исследование проведено на уникальной базе данных, собранной в рамках его проведения и состоящей из образовательных задач для развития soft skills, данных по их выполнению пользователями с разными soft skills профилями. Разработана модель интеллектуального проводника и произведена ее реализация в виде программного компонента для системы управления предприятием, основу которого составили модули психодиагностики, организационного управления, обучения и рекомендаций. Интеллектуальность разработанной системы позволяет качественно сформировать индивидуальную образовательную траекторию, которая вовлечет

* Статья опубликована при поддержке Программы НИУ ВШЭ «Университетское партнерство»

сотрудника не только в процесс обучения и развития, но и в достижение организационных целей. Вкладывая в развитие сотрудников, организация получает T-shaped специалистов, обладающих проактивной позицией и способных к самоорганизации. Результаты исследования могут использоваться предприятиями не только на организационном уровне, но и через трансляцию в системе образования для формирования экосистемы образования в соответствии с требованиями инновационного развития экономики региона.

Ключевые слова: обучение и развитие, надпрофессиональные компетенции, адаптивность, рекомендательная система, машинное обучение, матричная факторизация, нейронная сеть

Цитирование: Морозевич Е.С., Коротких В.С., Кузнецова Е.А. Разработка модели формирования индивидуальных образовательных траекторий с использованием методов машинного обучения // Бизнес-информатика. 2022. Т. 16. № 2. С. 21–35. DOI: 10.17323/2587-814X.2022.2.21.35

Введение

Сегодня экономика переживает цифровую трансформацию, которая представляет собой революционный процесс преобразования бизнес-модели организации не только с применением цифровых технологий, но и внесением коренных организационных изменений в технологии, культуру, операции и принципы создания новых продуктов.

Связующим звеном интеграции глубинных преобразований организаций и новых технологий выступают люди — сотрудники с их знаниями, навыками и ценностями. По данным ООО «Делойт Консалтинг» уже сегодня 53% организаций понимают, что в текущих условиях от 50 до 100% их сотрудников должны приобрести новые навыки и способности [1].

Организации создаются людьми с вложением своей системы ценностей в создаваемую структуру. Сотрудник, приходящий в организацию на определенном этапе жизненного цикла, не осознает, что происходит, потому что: он не знает — что было, он не догадывается — что будет. Ввиду этого актуальной задачей является развитие персонала с учетом точек соприкосновения ценностей сотрудника и организации.

Цифровизация деятельности порождает рост объема данных, которые требуют последующей обработки. Именно данные лежат в основе цифровой трансформации, являясь источником жизненной силы процесса перехода к цифровой экономике.

Уже в 2020 году по данным Data Age Report человечеством было сформировано около 51 зеттабайта информации, а к 2025 году объем этих данных увеличится почти в 3,5 раза и составит 175 зеттабайтов [2], что показывает тенденцию экспоненциального роста цифровых данных.

Активность пользователей, фиксируемая различными регистрирующими устройствами, представляет собой их цифровые следы. Сегодня цифровые следы занимают существенную часть облака больших данных и ключевое направление их использования — это извлечение информации о предпочтениях потенциальных клиентов и предложение продукции, которая им будут интересна, то есть повышение объема продаж [3].

Однако извлекаемую информацию из цифровых следов можно использовать не только во внешних взаимодействиях, но и во внутренней деятельности компании. Современные компании все чаще используют данные для повышения вовлеченности сотрудников в достижение и разделение целей, то есть для повышения продуктивности как сотрудников, так и процессов организации, тем самым создавая новые источники для своего конкурентного преимущества [4].

Таким образом, неотъемлемой частью крупных игроков рынка сегодня являются рекомендательные системы. Рекомендательная система представляет собой комплекс алгоритмов, программ и сервисов, предназначенных для формирования релевантных рекомендаций пользователям в отношении объекта информационного поиска [5].

Во многих исследованиях отмечается [6–8], что преодолеть ключевые барьеры цифровой трансформации в виде нехватки квалифицированных кадров, компетенций и знаний, а также внутреннего сопротивления в организациях [9–10] можно путем создания системы обучения и развития персонала, в основе которой лежит формирование индивидуальной образовательной траектории [11] на базе ценностей организации и сотрудника и релевантном образовательном материале, то есть на тех объ-

ектах, действиях, задачах, которые сотрудник будет выполнять с повышенной заинтересованностью и как следствие с максимальной эффективностью.

В сфере образования использование информационных систем, основанных на рекомендациях, призвано решать одну из главных проблем обучающегося – проблему выбора образовательного материала. Зачастую таким материалом являются программы обучения и образовательные курсы. При подборе образовательного материала в таких системах обычно не учитываются личностные характеристики обучающегося, а используются его предпочтения [12]. В последнее время исследователи начали активно изучать вопрос использования личностных характеристик в образовательных рекомендациях. Например, использование информации о стиле обучения, эмоциональных реакциях [13]. Кроме того, в исследованиях все чаще поднимается вопрос об использовании рекомендательных систем в онлайн обучении [14].

Исходя из вышесказанного, актуальной целью исследования является создание интеллектуального проводника, который бы сопровождал сотрудника на протяжении всей жизни в организации, вовлекая его в процесс обучения согласно персонализированной траектории на основе данных о личностном профиле и реакциях на образовательный материал, тренируя soft и hard skills в соответствии с ценностями организации и сотрудника.

1. Концепция интеллектуальной системы обучения и развития персонала

В основу разработки интеллектуального проводника для развития и обучения сотрудников легли принципы системной инженерии. Отрывной точкой в создании интеллектуального проводника

явился сбор требований заинтересованных сторон к разрабатываемой системе.

В результате анализа требований было выявлено, что модульная архитектура системы позволит не только учесть все требования заинтересованных сторон, но и разграничит доступ к различным функциям в зависимости от занимаемых позиций. Предлагаемая к реализации модульная архитектура в соответствии с выявленными требованиями представлена на *рис. 1*.

Для реализации интеллектуального проводника и поддержки работы предложенных модулей была создана база данных с помощью СУБД MySQL.

Основной задачей модуля диагностики является формирование цифрового профиля сотрудника, который позволит учесть все особенности личности и станет основой для ее эффективного обучения и развития.

Современные тенденции таковы, что организации все чаще делают упор на надпрофессиональные навыки или так называемые soft skills, которые гораздо труднее воспроизвести с помощью инструментов цифровой трансформации и которые лежат в основе происходящих изменений [15]. Soft skills являются основой для эффективного развития профессиональных навыков или так называемых hard skills.

Ввиду этого было предложено сформировать цифровой профиль сотрудника, состоящий не только из персональных данных (пол, возраст, образование, должность, опыт работы), но и раскрывающий уровни развития soft skills сотрудника, востребованных на рынке труда в условиях цифровой экономики.

Для определения уровня владения тем или иным soft skills была разработана комплексная автоматизированная

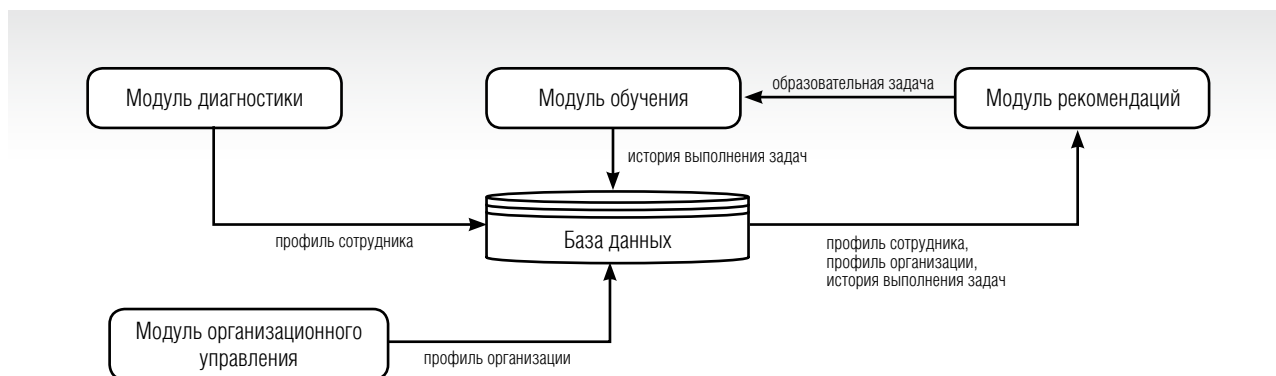


Рис. 1. Модель интеллектуального проводника.

зированной диагностика (рис. 2). В качестве диагностического материала использовались модель DISC, методика Алана Роу «Стиль принятия решений», методика определения стилей деятельности Хони и Мамфорда, тест на психотип. Данные методики раскрывают психоэмоциональные особенности личности в 16 проекциях с соответствующими им компетенциями.

Модуль организационного управления предназначен для фиксации организационных ценностей через формирование профилей должностей. В первую очередь через данный модуль производится создание организационной структуры. После этого для каждой рассматриваемой должности руководителем структурного подразделения определяются точки критического и желаемого профиля сотрудника в соответствии с трудовым функционалом и трудовыми действиями, а также обозначенными на основе организационных ценностей, надпрофессиональными навыками.

Таким образом, составив профиль должности для каждого элемента организационной структуры, формируется комплексный профиль организации, в котором отражены все особенности ее деятельности.

Основой модуля обучения является организация образовательного процесса. Пользователи в рамках

развития soft skills выполняют уникальные образовательные задачи (рис. 3). Для каждой задачи пользователь оценивает насколько понравилась ему задача, ее эффективность и сложность. Таким образом формируется набор данных, характеризующий выполнение пользователями образовательных задач и применяемый в дальнейшем для формирования релевантных рекомендаций.

Модуль рекомендаций призван формировать индивидуальную образовательную траекторию в соответствии с ценностями сотрудника и организации и подбирать максимально релевантный образовательный материал для сотрудника.

Имея профиль сотрудника и профиль должности можно определить направления развития soft skills сотрудника в рамках интересов организации. Однако, как было обозначено ранее, необходимо синхронизировать ценности сотрудника и организации, поэтому требуется также определить желаемые направления развития для сотрудника. Для этого в рамках модуля рекомендаций разработан опросник потребностей развития сотрудника в зависимости от его ценностных ориентиров.

После прохождения диагностики сотруднику предлагается ответить на перечень вопросов для формирования желаемого профиля сотрудника и

Вопрос: 5/64.

Когда я сталкиваюсь с проблемой, я:

Опираюсь на проверенные подходы	0	1	2	3	4
Провожу тщательный анализ	0	1	2	3	4
Ищу творческое решение	0	1	2	3	4
Опираюсь на мои чувства	0	1	2	3	4

выставьте каждому ответу оценку от 0 до 4, характеризующую на сколько утверждение близко вам. 4 - наиболее близко

[Вернуться в личный кабинет](#)

Рис.2. Модуль диагностики.

ДИАГНОСТИКА

ОБУЧЕНИЕ

ОРГАНИЗАЦИОННОЕ УПРАВЛЕНИЕ

ЛИЧНЫЕ ДАННЫЕ

Задачи

Активная задача

[компетенция: Доброжелательность]

Прочитайте стихотворение о дружелюбии «Путь Дружелюбия»

завершить задачу

Доступные задачи (подобраны случайно, переопределено администратором)

[компетенция: Объективность]

Читайте научно-популярную литературу: Кратчайшая история времени, Стивен Хокинг, Леонард Млодинов

взять задачу

[компетенция: Энергичность]

Во время занятия спортом включайте любимую музыку

взять задачу

[компетенция: Экспериментальность]

Почаще подвергайте себя новым испытаниям. Сходите в церковь не вашей веры.

взять задачу

[компетенция: Логическое мышление]

Старайтесь думать и делать выбор сам, иначе за вас будут делать другие.

взять задачу

[компетенция: Исполнительность]

Найдите способ тактично сказать своему сотруднику о увольнении.

взять задачу

[компетенция: Креативное мышление]

Прислушайтесь к звукам за окном, вдохните запах, доносящийся с кухни, осмотритесь вокруг, почувствуйте прикосновение теплого одеяла к ногам. Попытайтесь сохранить это состояние хотя бы на несколько минут. С каждым разом увеличивайте время.

взять задачу

Последние завершённые задачи (всего: 18)

[компетенция: Нетерпимость]

Как вы относитесь к взяточничеству в стране?

36.0 сек. 1 2

[компетенция: Абстрактность]

Сравните мопед и мотоцикл найдите общее и различное

55.0 сек. 1 3

Рис.3. Модуль обучения.

определения тех компетенций, которые он хотел бы развить.

Наложение полученных профилей друг на друга позволяет наглядно представить уровень выраженности компетенций и обозначить направления личностного развития сотрудника в соответствии с потребностями организации, что приводит к формированию их общей траектории развития.

Следующим шагом является непосредственный подбор релевантного образовательного материала с помощью рекомендательной системы. Для выбора алгоритма ее работы было проведено исследование таких технологий, как машинное обучение на основе матричной факторизации и нейронная сеть.

2. Вычислительный эксперимент

2.1. Постановка задачи

Есть множество пользователей $U = \{u_1, u_2, \dots, u_n\}$ и множество образовательных задач $T = \{t_1, t_2, \dots, t_m\}$.

Матрица $V_{n \times m}$ содержит оценки, выставленные пользователями образовательным задачам. На месте v_{ij} ($i \in 1, \dots, n; j \in 1, \dots, m$) будет стоять некоторое число, если пользователь u_i оценил задачу t_j и пусто в противном случае.

Требуется найти вектор $\hat{v}_i$, содержащий уже известные оценки пользователя v_{ij} , а также предполагаемые $\hat{v}_{ij}$. На основе полученного вектора произвести ранжирование списка образовательных задач T для пользователя u_i .

2.2. Машинное обучение на основе матричной факторизации

Матричная факторизация подразумевает декомпозицию исходной матрицы в произведение двух матриц малого ранга. Взаимодействие пользователя с объектом моделируется как скалярное произведение векторов представления пользователя и объекта в факторном пространстве. Факторизационные модели эффективно работают с сильно разреженными матрицами [16–18].

Представим матрицу оценок V в виде произведения двух матриц:

- ♦ матрицы $W_{n \times k}$, которая содержит в себе как латентные характеристики пользователей, так и явные (пол, образование, трудовой стаж и значения профилей пользователей, состоящих из 16 параметров);
- ♦ матрицы $H_{k \times m}$, характеризующей образовательные задачи.

Заполним латентные характеристики матриц W и H случайными величинами на основе закона равномерного распределения на интервале

$$[0; \sqrt{\max\{v_{ij}\} / k}].$$

Далее решим задачу минимизации:

$$\operatorname{argmim} \|V - \hat{V}\| + \alpha \|H\| + \beta \|W\|, \quad (1)$$

где $\hat{V}$ – матрица, полученная аппроксимацией из W и H ;

α, β – параметры алгоритма.

Каждая итерация алгоритма поиска решения минимизации ошибки состоит из следующих шагов:

1. Фиксируем матрицу H .
2. Находим ошибку $\delta = |v_{ij} - \hat{v}_{ij}|, j \in 1, \dots, n$.
3. Находим новые значения

$$W_{ip} = W_{ip} - \epsilon (\delta H_{pj}^T + \lambda W_{ip}),$$

где $p \in 1, \dots, k$;

λ – регуляризационный параметр;

ϵ – скорость обучения.

4. Фиксируем матрицу W .
5. Находим ошибку $\delta = |v_{ij} - \hat{v}_{ij}|, j \in 1, \dots, m$.
6. Находим новые значения

$$H_{pj} = H_{pj} - \epsilon (\delta W_{ip}^T + \lambda H_{pj}),$$

где $p \in 1, \dots, k$;

λ – регуляризационный параметр;

ϵ – скорость обучения.

При решении поставленной задачи опытным путем был найден ранг матриц W и H $k = 40$. Данное значение является граничным для заданного случая, при котором среднеквадратичная и абсолютная ошибки решения задачи факторизации являются оптимальными. При значении $k > 40$ преимущество в результате не прослеживается, время расчетов увеличивается. И наоборот, при значении $k < 40$ получается слишком грубая аппроксимация.

На основе результатов тестирования (таблица 1) при увеличении скорости обучения ϵ со значения 0,005 до 0,03 наблюдается улучшение результатов (уменьшение погрешности, увеличение точности предсказания). При значении $\epsilon = 0,04$ результаты как правило эквивалентны результатам при $\epsilon = 0,03$ либо хуже. При количестве итераций 20 ($\epsilon = 0,03$) уже наблюдается высокая точность факторизации на заданных выборках. На основе данных результатов используем следующие параметры для алгоритма: $\epsilon = 0,03$, количество итераций 20.

При увеличении объемов полученных данных, следовательно, уменьшении разреженности данных, может потребоваться увеличение количества итераций. Контроль количества итераций был автоматизирован с помощью выполнения тестирования после окончания факторизации и сравнения точности предсказания с прошлым тестированием. При уменьшении точности предсказания, производится увеличение количества итераций, иначе количество итераций остается прежним. Еще одна альтернатива – задание допустимой точности факторизации, на которой необходимо останавливать алгоритм.

Одной из основных задач при построении рекомендательных систем является решение проблемы «холодного» старта [19]. Она возникает, когда в системе появляются новые элементы: будь то пользователь или объекты предпочтений.

В рассматриваемой задаче можно выделить два варианта холодного старта пользователей: пользователь не выполнял образовательные задачи; пользователь выполнял образовательные задачи, но задача факторизации решалась до его регистрации, то есть в системе отсутствуют факторы пользователя.

Рассмотрим первый случай. О предпочтениях пользователя ничего неизвестно. Только его профиль, являющийся некоторым вектором, который сформирован из значений пола, стажа, должности. Если пользователь прошел тестирование, то в векторе будет также заложен soft skills профиль из шестнадцати компонент.

Таблица 1.

Скорость и ошибки факторизации ($n = 14000$)

Количество итераций	ϵ	0,005	0,01	0,02	0,03	0,04
20	ср. абс.:	0,197	0,119	0,064	0,05	0,048
	ср. кв.:	0,24	0,09	0,013	0,007	0,0064
	время, с.	6	6	5	5	5
40	ср. абс.:	0,113	0,057	0,028	0,02	0,022
	ср. кв.:	0,082	0,013	0,002	0,001	0,001
	время, с.	11	10	10	11	11
60	ср. абс.:	0,075	0,03	0,017	0,014	0,014
	ср. кв.:	0,03	0,003	0,0009	0,0006	0,0006
	время, с.	16	15	16	15	15

Проблему холодного старта в данном случае предлагается решать с помощью косинусного сходства. Для этого находится косинус между новым пользователем и каждым существующим пользователем, ранжируется по убыванию, выбирается тот пользователь, которому соответствует максимальный косинус. В результате для нового пользователя формируются рекомендации уже на основании найденного через косинус-сходство пользователя.

$$\cos \varphi = \frac{(\overline{W_a}, \overline{W_b})}{\|\overline{W_a}\| \times \|\overline{W_b}\|}. \quad (2)$$

В рамках второго случая проблему предлагается разрешить путем решения задачи факторизации для конкретного пользователя. В этом случае матрицы W и V примут вид векторов-строк. Матрица факторов задач H будет зафиксирована, а поиск значений будет произведен только для вектора-строки W . Алгоритм поиска сводится к частному случаю: $i = 1$, шаги алгоритма матричной факторизации с четвертого по шестой исключаются.

Также стоит выделить проблему «холодного» старта для образовательных задач. На текущем этапе исследования задачи, для которых отсутствуют факторы, случайным образом предлагаются пользователям.

Для оценки качества предсказаний множество оценок V было разделено на выборки для обучения V_{train} и тестирования V_{test} . Тестирование модели проводилось при следующих количественных характеристиках:

- ◆ количество пользователей: 303;
- ◆ количество задач: 11726;
- ◆ количество выполненных пользователями задач: 17733;
- ◆ обучающая выборка: 75%;
- ◆ тестовая выборка: 25%;
- ◆ ранг матриц: 40;
- ◆ шаги факторизации: 20;
- ◆ диапазон значений исходной матрицы: 0–10 (нормализованные данные).

Матрица V является очень разреженной, большинство ячеек не заполнены (более 95% ячеек). При оценке работы модели проводилось три типа аппроксимации: для эффективности, сложности и предпочтений. Были рассмотрены четыре случая:

- ◆ использование только латентных факторов (на графиках обозначены как *FREE*);
- ◆ использование фиксированных факторов пользователей наряду с латентными (soft skills профиль, пол, должность, образование; обозначены как *UV*);
- ◆ использование фиксированных факторов задач наряду с латентными (средние значения показателей эффективности, сложности и предпочтений, выставленные пользователями, уровень сложности, ориентация задач на компетенции, обозначены как *TV*);
- ◆ использование фиксированных факторов и пользователей, и задач (обозначены как *UV+TV*).

Для каждого случая проводилось 50 факторизаций на случайных данных, далее эти данные усреднялись. При тестировании также учитывался холодный старт пользователей на основе косинус-сходства. Холодный старт задач не учитывался в тестировании.

Данные для ошибки предсказания предпочтений представлены на *рис. 4–7*, для ошибки предсказания эффективности – на *рис. 8–11*.

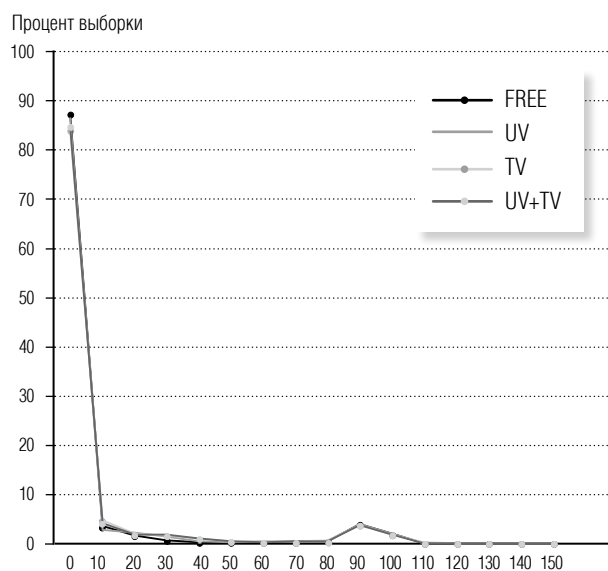


Рис. 4. Ошибка предсказания предпочтения с помощью факторизации на процентных интервалах.

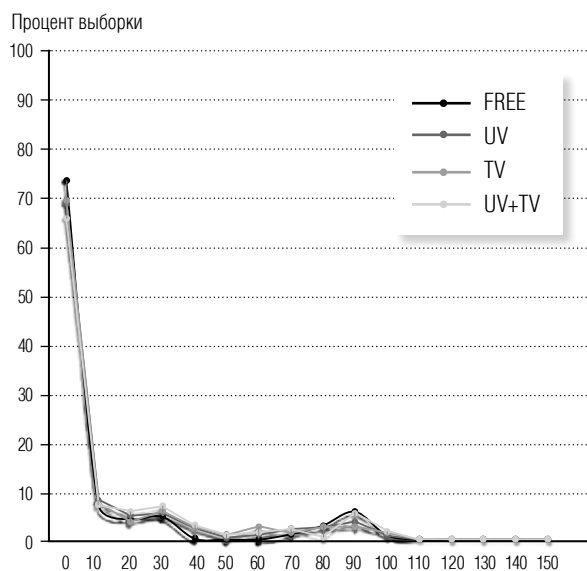


Рис. 6. Ошибка предсказания предпочтения с помощью косинус-сходства на процентных интервалах.

Как видно из графиков (*см. рис. 4–11*), использование фиксированных факторов в основном ухудшает результат. Это может быть следствием малого объема набора данных.

2.3. Нейронная сеть

Нейронная сеть представляет собой вычислительную структуру, собранную в сеть из вычислительных элементов, созданных на подобию биоло-

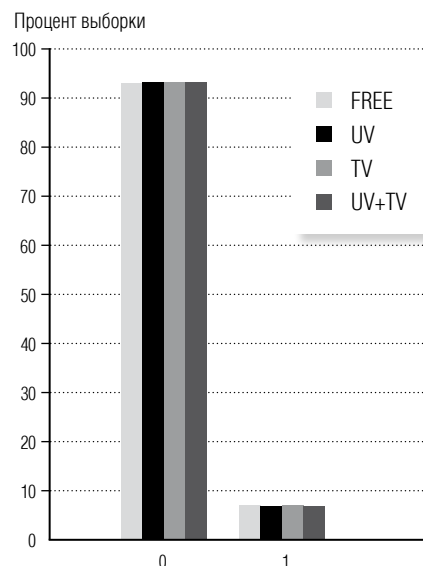


Рис. 5. Ошибка предсказания предпочтения с помощью факторизации на дискретных величинах [0; 1].

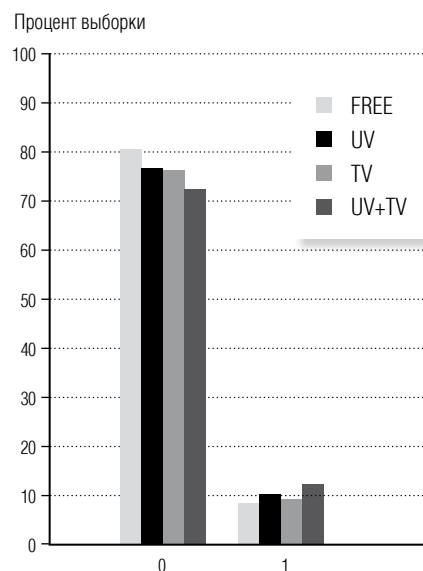


Рис. 7. Ошибка предсказания предпочтения с помощью косинус-сходства на дискретных величинах [0; 1].

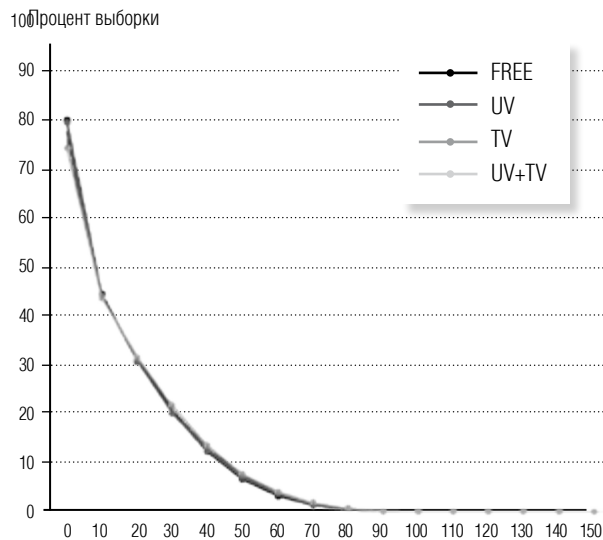


Рис. 8. Ошибка предсказания эффективности с помощью факторизации на процентных интервалах.

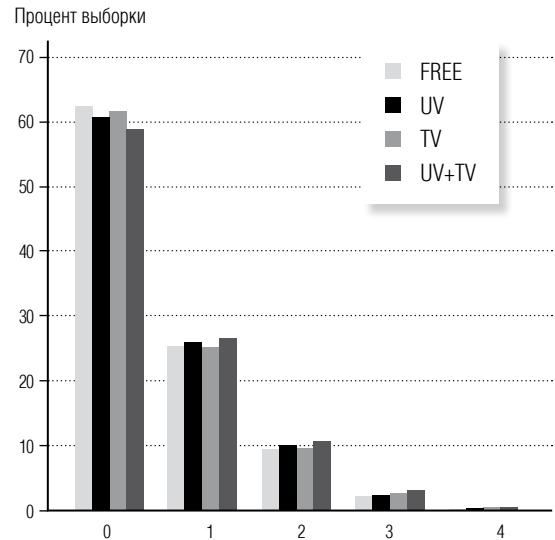


Рис. 9. Ошибка предсказания эффективности с помощью факторизации на дискретных величинах [1; 5].

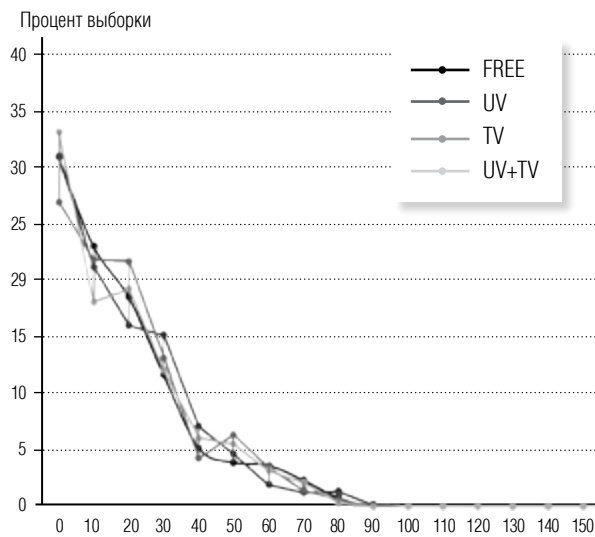


Рис. 10. Ошибка предсказания эффективности с помощью косинус-сходства на процентных интервалах.

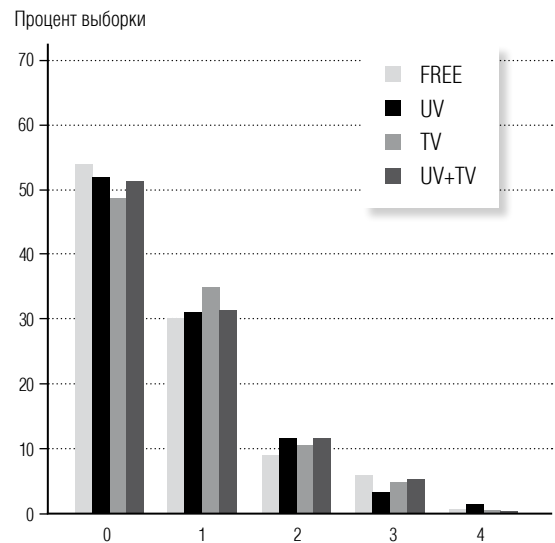


Рис. 11. Ошибка предсказания эффективности с помощью косинус-сходства на дискретных величинах [1; 5].

гического нейрона [20]. Нейронные сети способны решать широкий спектр задач в различных сферах деятельности [21].

Для решения поставленной задачи рассматривались такие архитектуры нейронных сетей как сверточная нейронная сеть [22] и многослойный персептрон по Румельхарту [23]. Наиболее эффективно показала себя последняя. К тому же, сверточ-

ная нейронная сеть требует большего количества данных и времени на обучение.

Для векторизации образовательной задачи было исследовано два способа: кодирование слов уникальным номером и вложение слов (Word2Vec) [24]. Кодирование слов уникальным номером является более быстрым, однако не позволяет фиксировать никакие отношения между словами, а также сходство слов.

Использование модели Word2Vec для вложения слов позволяет фиксировать тонкие взаимосвязи между словами, но требует для обучения большого объема данных и является длительным процессом. Ввиду отсутствия необходимости постоянной векторизации образовательных задач, а лишь при пополнении базы образовательных задач, была использована модель Word2Vec.

Входной слой нейронной сети имеет один нейрон, на который поступает вектор, содержащий данные о пользователе и задаче.

На скрытых слоях задействована функция активации «ReLU», на выходном слое – функция активации «Sigmoid», который позволяет усиливать слабые сигналы и не насыщаться от сильных сигналов [25].

Для обучения и тестирования нейронной сети был использован набор данных, состоящий из 16 862 векторов, содержащих данные о пользователях и задачах. Данные были разделены на обучающую и тестовую выборки в пропорции 80/20.

На исследуемом наборе данных наибольшую точность показала нейронная сеть, имеющая следующую топологию: два скрытых слоя (первый слой – 64 нейрона, второй слой – 128 нейронов), размер батча – 32, количество эпох – 20.

Результаты работы нейронной сети и точность предсказания представлены в *таблице 2*.

Для построения вектора $\hat{v}_i$ запуск сети производится столько раз, сколько значений содержится в базе образовательных задач. Таким образом, после всех итераций формируется вектор $\hat{v}_i$, на основе которого можно произвести выбор наиболее предпочтительных задач для пользователя.

3. Сравнение технологий искусственного интеллекта

В рамках исследования рассмотрены два подхода к формированию рекомендаций: нейронная сеть и машинное обучение на основе матричной факторизации. Оба подхода имеют как достоинства, так и недостатки.

В отличие от векторов, подаваемых на вход нейронной сети, матрица факторов состоит из уникальных векторов пользователей, которые содержат латентные признаки, не учитываемые нейронной сетью. Нейронная сеть, в свою очередь, опирается только на заданный известный вектор пользователя, состоящий из параметров, заложенных исследователем. Это позволяет судить о более персонализированном подходе при использовании матричной факторизации.

Нейронная сеть также требует обязательное наличие не только всех данных о пользователе, но и их корректности. При использовании машинного обучения на основе матричной факторизации можно выдавать рекомендации уже после нескольких оценок, не имея более никакой информации.

Недостатком матричной факторизации является холодный старт, когда пользователь еще не оценивал ни одной задачи, а ему требуется выдать рекомендации. Однако данная проблема решается «на лету» с помощью способа, описанного ранее.

Несмотря на то, что перед исследователями стояла задача ранжирования, необходимо было сравнить работу этих методов по трем критериям: предпочтения (бинарная величина), сложность (дискретная [1; 4]) и эффективность (дискретная [1; 5]). Результаты сравнения представлены в *таблице 3*.

Как видно из *таблицы 3*, на дискретных величинах точность предсказаний нейронной сети падает по сравнению с матричной факторизацией. Исходя из вышесказанного, исследователями было принято решение использовать в качестве ядра рекомендательного модуля машинное обучение на основе матричной факторизации.

Заключение

В статье рассмотрена проблема низкой адаптивности управления в условиях цифровой трансформации. В рамках исследования разработана модель интеллектуального проводника сотрудника, которая позволяет нивелировать данную проблему за счет

Таблица 2.

Результаты работы нейронной сети

	Предпочтение	Сложность	Эффективность
Среднеквадратичное отклонение	0,056	0,102	0,098
Точность, %	93,31	68,79	61,0

Таблица 3.

Точность предсказаний технологий искусственного интеллекта

	Предпочтение	Сложность	Эффективность
Нейронная сеть	93,3%	68,8	61%
Матричная факторизация	93,1%	76,4	62,4%

качественного развития компетентностного профиля сотрудника с учетом ценностных ориентиров самого сотрудника и организации, а также произведена ее реализация в виде программного компонента для системы управления предприятием.

Освоение новых компетенций приведет к расширению профиля сотрудника, возможности выполнения новых задач в интересующей области, тем самым получая качественный результат. Организация, в свою очередь, получит новые возможности для развития на рынке и быстрой адаптации к изменяющимся условиям. Таким образом, вкладывая в развитие сотрудников, организация получает T-shaped специалистов, обладающих проактивной позицией и способных к самоорганизации. Что непосредственно приводит к реализации потенциалов как сотрудника, так и организации.

Интеллектуальный проводник имеет модульную архитектуру, что обусловлено междисциплинарностью и новизной исследования, а также простотой внедрения полученных разработок в реальную управленческую практику. Такой подход позволил провести качественные исследования в двух проекциях: разработка комплексной автоматизированной диагностики soft skills профиля сотрудника и релевантного наполнения индивидуальной образовательной траектории образовательным материалом с помощью искусственного интеллекта, а затем объединить полученные результаты в единую систему для обучения и развития персонала.

Искусственный интеллект в предлагаемом подходе выражен с помощью машинного обучения с использованием матричной факторизации. Такая интеллектуальность позволяет качественно подобрать образовательный материал, который будет интересен сотруднику с его личностных позиций и максимально вовлечет его не только в процесс обучения и развития, но и в достижение организационных целей.

Предложенная структура базы данных позволяет собирать цифровые следы пользователей, описыва-

ющие биографические характеристики, навыки и уровень их владения на определенном этапе развития, историю выбора и выполнения образовательных задач, и в дальнейшем использовать собранные данные в качестве основы для работы модуля рекомендаций и формирования персонифицированных предложений.

Междисциплинарность и новизна исследования также обуславливают вариативность его развития. В рамках дальнейшего исследования планируется интеграция программного компонента в систему управления бизнес-процессами предприятия. Предполагается прогнозировать время отсутствия загрузки трудового ресурса и предлагать для его заполнения образовательные задачи. Это станет основной принципиально нового подхода к рассмотрению простоев не как потерь, а как возможности для развития и обучения сотрудников.

Разрабатываемый программный компонент может использоваться предприятиями не только на организационном уровне, но и через трансляцию в систему образования. С одной стороны, предприятия смогут сократить простои персонала, расширить личностно-компетентностный профиль сотрудников, увеличить рост производственного потенциала как самих сотрудников, так и предприятий на рынке. С другой стороны, образовательные учреждения смогут осуществлять подготовку высококвалифицированных кадров в рамках ориентации на ценности рынка и ценности студентов с формированием индивидуальных образовательных траекторий. Такая кооперация позволит сформировать экосистему образования в соответствии с требованиями инновационного развития экономики региона.■

Благодарности

Данное исследование выполнено при поддержке Сибирского государственного университета науки и технологий им. М.Ф. Решетнева в соответствии

с соглашением о выполнении научно-исследовательской работы от 01.12.2020 №04/20-ВГ по проекту «Исследование и разработка интеллектуально-

го проводника для расширения компетентностного профиля сотрудника в рамках индивидуальных образовательных траекторий».

Литература

1. Результаты исследования «Международные тенденции в сфере управления персоналом – 2020» по России // Делойт Инсайт, 2020. [Электронный ресурс]: <https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/human-capital/russian/hc-trends-2020-Russia.pdf> (дата обращения 22.11.2021).
2. Большие данные (Big Data) мировой рынок // TAdviser, 2020. [Электронный ресурс]: <https://www.tadviser.ru/a/129607> (дата обращения 22.11.2021).
3. Фатьянов А.А. Большие данные в цифровой экономике: ценность и правовые вызовы // Экономика. Право. Общество. 2018. № 4. С. 37–40.
4. Reinsel D., Gantz J., Rydning J. The digitization of the world – from edge to core // IDC White Paper, 2018. [Электронный ресурс]: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> (дата обращения 22.11.2021).
5. Ali N.M., Gadallah A.M., Hefny H.A., Novikov B.A. Online web navigation assistant // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2021. Т. 31. Вып. 1. С. 116–131. <https://doi.org/10.35634/vm210109>
6. Что такое цифровая экономика? Тренды, компетенции, измерение / Г.И. Абдрахманова и [др.]; науч. ред. Л. М. Гохберг; Нац. исслед. ун-т «Высшая школа экономики». М.: Издательский дом Высшей школы экономики, 2019 г.
7. Куликова М.Х., Куликова М.Х., Магоматов М.А. ИТ в образовании: сегодня, завтра и всегда // Материалы I студенческой научно-практической конференции «Информационные технологии в бизнесе и образовании», Грозный, 20 февраля 2020 года. Чеченский государственный университет им. А.А.Кадырова, 2020. С. 32–35. <https://doi.org/10.36684/21-2020-1-32-35>
8. Geissbauer R., Lübken E., Schrauf S., Pillsbury S. Digital champions. How industry leaders build integrated operations ecosystems to deliver end-to-end customer solutions // PwC Strategy & Global Digital Operations Study 2018 [Электронный ресурс]: <https://www.strategyand.pwc.com/gx/en/insights/industry4-0/global-digital-operations-study-digital-champions.pdf> (дата обращения 22.11.2021).
9. Harvard Business Review. High-performance sourcing and procurement driving value through collaboration // Harvard Business School Publishing, 2017. [Электронный ресурс]: <https://hbr.org/resources/pdfs/comm/scoutfrp/HighPerformanceSourcing.pdf> (дата обращения 22.11.2021).
10. Долганова О.И., Деева Е.А. Готовность компании к цифровым преобразованиям: проблемы и диагностика // Бизнес-информатика. 2019. Т. 13. № 2. С. 59–72. <http://doi.org/10.17323/1998-0663.2019.2.59.72>
11. Комаров В.А., Сарафанов А.В. Системы интернета вещей в процессе междисциплинарной подготовки кадров для цифровой экономики и их проектирование // Бизнес-информатика. 2021. Т. 15. № 2. С. 47–59. <http://doi.org/10.17323/2587-814X.2021.2.47.59>
12. Rivera A.C., Tapia-Leon M., Lujan-Mora S. Recommendation systems in education: A systematic mapping study // Rocha, Á., Guarda, T. (eds) Proceedings of the International Conference on Information Technology & Systems (ICITS 2018). Advances in Intelligent Systems and Computing. Vol. 721. P. 937–947. https://doi.org/10.1007/978-3-319-73450-7_89
13. Bustos López M., Alor-Hernández G., Sánchez-Cervantes J., Paredes-Valverde M., Salas-Zárate M.P. EduRecomSys: An educational resource recommender system based on collaborative filtering and emotion detection // Interacting with Computers. 2020. Vol. 32. No. 4. P. 407–432. <https://doi.org/10.1093/iwc/iwab001>
14. Urdaneta-Ponte M.C., Mendez-Zorrilla A., Oleagordia-Ruiz I. Recommendation systems for education: Systematic review // Electronics. 2021. Vol. 10. No. 14. Article ID 1611. <https://doi.org/10.3390/electronics10141611>
15. Bughin J., Hazan E., Lund S., Dahlström P., Wiesinger A., Subramaniam A. Skill shift: Automation and the future of the workforce // McKinsey & Company, 2018. [Электронный ресурс]: <https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce> (дата обращения 22.11.2021).
16. Рубцов В.Н. Матричная факторизация на основе глубокого обучения для коллаборативной фильтрации: дис. студента / НИУ ВШЭ, Москва, 2017. [Электронный ресурс]: <https://www.hse.ru/edu/vkr/206744221> (дата обращения 22.11.2021).
17. Strömquist Z. Matrix factorization in recommender systems: How sensitive are matrix factorization models to sparsity? // Uppsala University Publications, 2018. [Электронный ресурс]: <https://uu.diva-portal.org/smash/get/diva2:1214390/FULLTEXT01.pdf> (дата обращения 22.11.2021).
18. Мойсюк-Дранько П.А., Ревотюк М.П. Методы матричной факторизации для систем рекомендации // Информационные технологии и системы 2020 (ИТС 2020): материалы международной научной конференции. Минск: БГУИР, 2020. С. 193–194. [Электронный ресурс]: https://libdoc.bsuir.by/bitstream/123456789/41339/1/Moysyuk_Dranko_Metody.pdf (дата обращения 22.11.2021).
19. Кузнецов И.А. Методы и алгоритмы машинного обучения для предобработки и классификации слабоструктурированных текстовых данных в научных рекомендательных системах: дис. ... канд. техн. наук / ФГАОУ ВО НИЯУ «МИФИ». Москва, 2019. [Электронный ресурс]: https://ds.mephi.ru/documents/90/Кузнецов_И_А_Текст_диссертации.pdf (дата обращения 22.11.2021).
20. Головкин В.А., Краснопрошин В.В. Нейросетевые технологии обработки данных // Минск: БГУ, 2017. [Электронный ресурс]: [https://ds.mephi.ru/documents https://elib.bsu.by/bitstream/123456789/193558/1/Golovko.pdf](https://ds.mephi.ru/documents/https://elib.bsu.by/bitstream/123456789/193558/1/Golovko.pdf) (дата обращения 22.11.2021).

21. Лисовский А.Л. Применение нейросетевых технологий для разработки систем управления // Стратегические решения и риск-менеджмент. 2020. Т. 11. № 4. С. 378–389. <https://doi.org/10.17747/2618-947X-923>
22. Bezdan T., Bacanin Džakula N. Convolutional neural network layers and architectures // Sinteza 2019 – International Scientific Conference on Information Technology and Data Related Research. Belgrade, Singidunum University, 2019. P. 445–451. <https://doi.org/10.15308/Sinteza-2019-445-451>
23. Постарнак Д.В. Критический анализ моделей нейронных сетей // Вестник Тюменского государственного университета. 2012. № 4. С. 162–167. [Электронный ресурс]: <https://elibrary.ru/item.asp?id=17758787> (дата обращения 22.11.2021).
24. Young T., Hazarika D., Poria S., Cambria E. Recent trends in deep learning based natural language processing // IEEE Computational Intelligence Magazine. 2018. Vol. 13. No. 3. P. 55–75. <https://doi.org/10.1109/MCI.2018.2840738>
25. Iqbal T., Qureshi S. The survey: Text generation models in deep learning // Journal of King Saud University – Computer and Information Sciences. 2020. <https://doi.org/10.1016/j.jksuci.2020.04.001>

Об авторах

Морозевич Екатерина Сергеевна

процессный архитектор, научно-исследовательское управление, Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, 660037, Сибирский федеральный округ, Красноярский край, г. Красноярск, проспект им. газеты Красноярский рабочий, д. 31;

E-mail: katyamorozevich@mail.ru

ORCID: 0000-0001-8796-3157

Коротких Владимир Сергеевич

системный администратор, Муниципальное бюджетное общеобразовательное учреждение средняя общеобразовательная школа №4, 663091, Сибирский федеральный округ, Красноярский край, г. Дивногорск, ул. Набережная, д. 9;

E-mail: vskor@bk.ru

ORCID: 0000-0001-9270-3376

Кузнецова Евгения Александровна

инженер, научно-исследовательское управление, Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, 660037, Сибирский федеральный округ, Красноярский край, г. Красноярск, проспект им. газеты Красноярский рабочий, д. 31;

E-mail: zhenya.kuz-1997@yandex.ru

ORCID: 0000-0002-3559-8394

The development of a model for a personalized learning path using machine learning methods

Ekaterina S. Morozevich^a

E-mail: katyamorozevich@mail.ru

Vladimir S. Korotkikh^b

E-mail: vskor@bk.ru

Yevgeniya A. Kuznetsova^a

E-mail: zhenya.kuz-1997@yandex.ru

^a Reshetnev Siberian State University of Science and Technology
Address: 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russia

^b Municipal budget general education institution secondary school No. 4
Address: 9, st. Naberezhnaya, Divnogorsk 663091, Russia

Abstract

Today, the economy is undergoing a digital transformation. Its key barriers are a lack of qualified personnel, competencies and knowledge, as well as internal resistance in organizations. It can be overcome through quality staff development and training. An urgent problem is to build a personalized learning path. Modern research is aimed at the implementation of recommendation systems in order to select relevant material. However, these recommendations are based on digital traces; the student's full personal profile, as well as organizational values are not considered. This study aims to create an intelligent guide that would accompany an employee throughout his life in the organization, involving him in the learning process according to a personalized path based on a complex personal profile and reactions to educational material, training soft and hard skills in accordance with the values of the organization and the employee. Methods of system analysis, system engineering, psychodiagnostic research (the DISC model, Rowe's "Decision-making Style" methodology, Honey and Mumford's method of determining activity styles, psychotype test), software design and artificial intelligence (matrix factorization and neural networks) were used in this study. The study was conducted on a unique database collected as part of its implementation and consisting of educational tasks for soft skills development, plus data on their implementation by users with different soft skills profiles. An intelligent guide model has been developed and implemented as a software component for an enterprise management system. The basis consists of psychodiagnostic modules, organizational management, training and recommendations. The intelligence of the system we developed allows you to qualitatively form a personalized learning path that will involve an employee not only in the learning and development process, but also in achieving organizational goals. The organization receives T-shaped specialists who have a proactive position and are capable of self-organization by investing in the development of employees. The results of this study can be used by enterprises not only at the organizational level, but also through broadcasting in the education system to form an education ecosystem in accordance with the requirements of innovative development of a given region's economy.

Keywords: training and development, supra-professional competencies, adaptability, recommendation system, machine learning, matrix factorization, neural network

Citation: Morozovich E.S., Korotkikh V.S., Kuznetsova Y.A. (2022) The development of a model for a personalized learning path using machine learning methods. *Business Informatics*, vol. 16, no. 2, pp. 21–35. DOI: 10.17323/2587-814X.2022.2.21.35

References

1. Deloitte Insights (2020) *Results of the study "International trends in human resource management – 2020" in Russia*. Available at: <https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/human-capital/russian/hc-trends-2020-Russia.pdf> (accessed 22 November 2021) (in Russian).
2. TAdviser (2020) *Big Data global market*. Available at: <https://tadviser.com/a/e.php?id=129607> (accessed 22 November 2021).
3. Fatyanov A.A. (2018) Big Data in the digital economy: its value and legal challenges. *Economics. Law. Society*, no. 4, pp. 37–40 (in Russian).
4. Reinsel D., Gantz J., Rydning J. (2018) *The digitization of the world – from edge to core*. IDC White Paper. Available at: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> (accessed 22 November 2021).
5. Ali N.M., Gadallah A.M., Hefny H.A., Novikov B.A. (2021) Online web navigation assistant. *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, vol. 31, no. 1, pp. 116–131. <https://doi.org/10.35634/vm210109>
6. Gokhberg L. (ed.) (2019) *What is the digital economy? Trends, competencies, measurement*. Moscow: HSE (in Russian).
7. Kulikova M.Kh., Kulikova M.Kh., Magomadov M.A. (2020) IT in education: today, tomorrow and always. Proceedings of the I Student Scientific and Practical Conference "Information technologies in business and education", Grozny, 20 February 2020, pp. 85–90 (in Russian). <https://doi.org/10.36684/21-2020-1-32-35>
8. Geissbauer R., Lübken E., Schrauf S., Pillsbury S. (2018) *Digital Champions. How industry leaders build integrated operations ecosystems to deliver end-to-end customer solutions*. PwC Strategy & Global Digital Operations Study. Available at: <https://www.strategyand.pwc.com/gx/en/insights/industry4-0/global-digital-operations-study-digital-champions.pdf> (accessed 22 November 2021).
9. Harvard Business Review (2017) *High-performance sourcing and procurement driving value through collaboration*. Harvard Business School Publishing. Available at: <https://hbr.org/resources/pdfs/comm/scoutfrp/HighPerformanceSourcing.pdf> (accessed 22 November 2021).
10. Dolganova O.I., Deeva E.A. (2019) Company readiness for digital transformations: problems and diagnosis. *Business Informatics*, vol. 13, no 2, pp. 59–72. <http://doi.org/10.17323/1998-0663.2019.2.59.72>
11. Komarov V.A., Sarafanov A.V. (2021) IoT systems in the process of multidisciplinary training of personnel for the digital economy and their design. *Business Informatics*, vol. 15, no 2, pp. 47–59. <http://doi.org/10.17323/2587-814X.2021.2.47.59>

12. Rivera A.C., Tapia-Leon M., Lujan-Mora S. (2018) Recommendation systems in education: A systematic mapping study. Proceedings of the *International Conference on Information Technology & Systems (ICITS 2018)*, Libertad City, Ecuador, 10–12 January 2018 (eds. A. Rocha, T. Guarda). Advances in Intelligent Systems and Computing, vol. 721, pp. 937–947. https://doi.org/10.1007/978-3-319-73450-7_89
13. Bustos López M., Alor-Hernández G., Sánchez-Cervantes J., Paredes-Valverde M., Salas-Zárate M.P. (2020) EduRecomSys: An educational resource recommender system based on collaborative filtering and emotion detection. *Interacting with Computers*, vol. 32, no. 4, pp. 407–432. <https://doi.org/10.1093/iwc/iwab001>
14. Urdaneta-Ponte M.C., Mendez-Zorrilla A., Oleagordia-Ruiz I. (2021) Recommendation systems for education: Systematic review. *Electronics*, vol. 10, no. 14, article ID 1611. <https://doi.org/10.3390/electronics10141611>
15. Bughin J., Hazan E., Lund S., Dahlström P., Wiesinger A., Subramaniam A. (2018) *Skill shift: automation and the future of the workforce*. McKinsey & Company. Available at: <https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce> (accessed 22 November 2021).
16. Rubtsov V.N. (2017) *Matrix factorization based on deep learning for collaborative filtering*. Student Theses. Moscow: HSE. Available at: <https://www.hse.ru/en/edu/vkr/206744221> (accessed 22 November 2021) (in Russian).
17. Strömquist Z. (2018) *Matrix factorization in recommender systems: How sensitive are matrix factorization models to sparsity*. Uppsala University Publications. Available at: <https://uu.diva-portal.org/smash/get/diva2:1214390/FULLTEXT01.pdf> (accessed 22 November 2021).
18. Mojsyuk-Dran'ko P.A., Revotyuk M.P. (2020) Matrix factorization methods for recommendation systems. Proceedings of the *International Conference Information Technologies and Systems 2020 (ITS 2020)*. Minsk: Belarusian State University of Informatics and Radioelectronics, pp. 193–194. Available at: https://libeloc.bsuir.by/bitstream/123456789/41339/1/Mojsyuk_Dranko_Metody.pdf (accessed 22 November 2021) (in Russian).
19. Kuznetsov I.A. (2019) *Machine learning methods and algorithms for preprocessing and classification of semi-structured text data in scientific recommendation systems*. Moscow: NRNU MEPhI. Available at: https://ds.mephi.ru/documents/90/Кузнецов_И_А_Текст_диссертации.pdf (accessed 22 November 2021) (in Russian).
20. Golovko V.A., Krasnoproshech V.V. (2017) *Neural network data processing technologies*. Minsk: Belarusian State University. Available at: <https://elib.bsu.by/bitstream/123456789/193558/1/Golovko.pdf> (accessed 22 November 2021) (in Russian).
21. Lisovsky A.L. (2020) Application of neural network technologies for management development of systems. *Strategic decisions and risk management*, vol. 11, no. 4, pp. 378–389 (in Russian). <https://doi.org/10.17747/2618-947X-923>
22. Bezdan T., Bacanin Džakula N. (2019) Convolutional neural network layers and architectures. Proceedings of the *Sinteza 2019: International Scientific Conference on Information Technology and Data Related Research, Belgrade, 20 April 2019* (ed. Milovan Stanišić), pp. 445–451. <https://doi.org/10.15308/Sinteza-2019-445-451>
23. Postarnak D.V. (2012) The critical analysis of models of neural networks. *Vestnik Tyumenskogo gosudarstvennogo universiteta*, no. 4, pp. 162–167. Available at: <https://elibrary.ru/item.asp?id=17758787&> (accessed 22 November 2021) (in Russian).
24. Young T., Hazarika D., Poria S., Cambria E. (2018) Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75. <https://doi.org/10.1109/MCI.2018.2840738>
25. Iqbal T., Qureshi S. (2020) The survey: Text generation models in deep learning. *Journal of King Saud University — Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2020.04.001>

About the authors

Ekaterina S. Morozovich

Process architect, Research Management Department, Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk 660037, Russia;

E-mail: katyamorozovich@mail.ru

ORCID: 0000-0001-8796-3157

Vladimir S. Korotkikh

System administrator, Municipal budget general education institution secondary school No. 4, 9, st. Naberezhnaya, Divnogorsk 663091, Russia;

E-mail: vskor@bk.ru

ORCID: 0000-0001-9270-3376

Yevgeniya A. Kuznetsova

Engineer, Research Management Department, Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk 660037, Russia;

E-mail: zhenya.kuz-1997@yandex.ru

ORCID: 0000-0002-3559-8394