

9. Zadorozhny A. et al. first flight-testing of LORA modulation in satellite radio communications in Low-Earth orbit. *IEEE Access*, 2022, vol. 10, pp. 100006–100023. DOI: 10.1109/ACCESS.2022.3207762
10. Prokopyev V.Y. et al. NORBY cubesat nanosatellite: design challenges and the first flight data. *Journal of Physics: Conference Series*, 2021, pp. 012–038. DOI: 10.1088/1742-6596/1867/1/012038

Received 11.12.2023

УПРАВЛЕНИЕ И ПОДГОТОВКА КАДРОВ ДЛЯ ОТРАСЛИ ТЕЛЕКОММУНИКАЦИЙ

УДК 004.855.5

ИСПОЛЬЗОВАНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ТРУДОУСТРОЙСТВА ВЫПУСКНИКОВ

Селиванова Е.С.¹, Газизов Т.Т.^{1,3}, Долганова Н.Ф.¹, Пираков Ф.Д.²

¹Томский государственный педагогический университет, Томск, РФ

²Томский государственный университет систем управления и радиоэлектроники, Томск, РФ

³Сахалинский государственный университет, Южно-Сахалинск, РФ

E-mail: slyvva@tspu.edu.ru

В современном мире технологии продолжают играть все более весомую роль во всех сферах жизни человека. Образование не является исключением и идет в ногу со временем. Одно из более перспективных направлений в этой области является использование методов машинного обучения для анализа данных о выпускниках и прогнозировании их дальнейшего трудоустройства, на основе накопленной информации об обучающихся студентах. Такая информация не только помогает определиться с будущей профессией, но и позволит учебным заведениям более эффективно координировать свою деятельность и повышать качество образования. Кроме того, использование машинного обучения в образовании может привести к созданию новых, более эффективных методов обучения, учитывающих индивидуальные особенности каждого студента. Результатом данной работы является интеллектуальная система обработки больших данных, способная адаптироваться к текущему состоянию рынка труда и помогающая выпускникам в более раннем определении в своей будущей профессии. Анализ данных проводился на выборке обучающихся Томского государственного педагогического университета 2021–2023 учебных годов.

Ключевые слова: анализ данных, трудоустройство, машинное обучение, обработка информации, искусственный интеллект, модель, прогнозирование

Введение и постановка задачи

Ценнейшим товаром на современном рынке услуг является информация о пользователе, анализ которой посредством алгоритмов машинного обучения позволяет обеспечить ее целенаправленное использование в дальнейшем. Информация о человеке, его достижениях и увлечениях считается одним из ценнейших продуктов не только на рынке рекламы, но и в образовательной среде [1; 2]. Система образования является одной из ключевых сфер жизнедеятельности любого человека и активно внедряет и использует тренды в развитии информационных технологий [3; 4]. В образовании целесообразно применять методы машинного обучения. Можно предположить, что в сфере образования возможно прогнозировать будущие места работы студентов педагогического вуза, используя данные выпускников

[5; 6]. Это позволит выявить траекторию профессиональной подготовки молодых специалистов и оценить качество предоставляемых образовательных услуг вузом [7].

В связи с этим целью работы является проведение обработки данных о трудоустройстве и прогнозирование места работы студентов на примере выпускников, используя модели искусственного интеллекта.

Для достижения поставленной цели необходимо решить следующие основные задачи:

1. Определить параметры, влияющие на трудоустройство выпускников.
2. Обработать данные о трудоустройстве выпускников.
3. Разработать интеллектуальную систему для прогнозирования трудоустройства выпускников.

Машинное обучение в прикладных задачах

Одним из наиболее распространенных направлений использования машинного обучения является оптимизация подборов туров в турагентствах [8]. Именно методы автоматического поиска закономерностей в больших объемах информации (данных) являются фундаментом задач прогнозирования и одним из наиболее перспективных направлений использования современных информационных технологий в различных отраслях экономики, в том числе и в туризме. В туристическом бизнесе важно понимать и оперативно предлагать клиенту определенные направления проведения досуга, в зависимости от его предпочтений, образа жизни, социального статуса и прочего. Минимизировать трудовые затраты на удовлетворение потребностей конкретного клиента, предоставив лучший вариант отдыха для него, зачастую помогают инструменты аналитики, механизмы моделирования и разработка алгоритмов. Анализ пользовательских предпочтений осуществляется на фиксации данных и их изучении. Отзывы посетителей о выбранных туристических объектах и их сегментирование по каким-либо схожим характеристикам (как со стороны объектов, так и клиентов) лежат в основе «рекомендательных систем, предлагающих пользователю наиболее подходящие для него точки притяжения, отели, рестораны и прочее». Система спрашивает у гостя при посещении сайта о его предпочтениях и возможностях и на основании отзывов пользователей предоставляет наиболее подходящие варианты отдыха данному клиенту. Из чего следует, что уже на этапе посещения сайта клиенту сразу «предсказывается» его «идеальный» тур, и на услугах менеджера по данному фронту работы можно экономить.

Data Science и Machine Learning

Чтобы научиться извлекать полезную информацию из полученных данных, необходимо их правильно обрабатывать. Для этого используется Data Science – наука о данных, которая включает в себя все инструменты, методы и технологии и позволяет обрабатывать данные и использовать их для собственной выгоды, выверяя закономерности в больших массивах данных и на их основе прогнозировать вероятность получения того или иного исхода [9].

Три основных составляющих Data Science:

1. Организация информации – это процесс ее хранения и форматирования.

2. Агрегация данных – это объединение начальных сведений в новый вид и (или) представление.

3. Доставка данных – это действие, обеспечивающее доступ к коллекциям агрегированных данных.

Существует большое количество областей для работы с данными, использующих искусственный интеллект, и одна из них – это машинное обучение (machine learning, ML) [10; 11], позволяющее создавать программы, улучшающиеся в процессе обучения.

Основные задачи ML:

1. Регрессия (предсказание числовых значений на основе заданных признаков).

2. Классификация (на основе набора признаков происходит разделение объектов).

3. Кластеризация (распределение данных на группы).

Большинство задач, которое решается посредством ML, использует определенные методы, то есть алгоритмы, позволяющие компьютеру обучаться на основе опыта и предоставленных данных, и сделать предсказание или принять решение на основе этого обучения.

Способы машинного обучения [12; 13]:

1. *Машинное обучение с учителем.*

В этом способе человек обучает машину с помощью огромной выборки данных, параметры которой варьируются до тех пор, пока не получится нужного результата.

2. *Машинное обучение без учителя.*

Машинное обучение без учителя предполагает исследование компьютером определенного набора данных, где выявляются скрытые закономерности корреляции между переменными.

3. *Машинное обучение с частичным привлечением учителя.*

Гибридный способ, в основе которого лежит обучение с учителем и без. Учитель, отметив небольшую часть данных, дает возможность понять, каким образом сгруппировать остальные.

4. *Обучение с подкреплением.*

При данном способе обучения машине позволяет взаимодействовать с окружением и в зависимости от настроек получать вознаграждение при правильном выполнении задания.

Для анализа данных выпускников педагогического вуза и реализации прогноза трудоустройства студентов выбран способ машинного обучения с учителем.

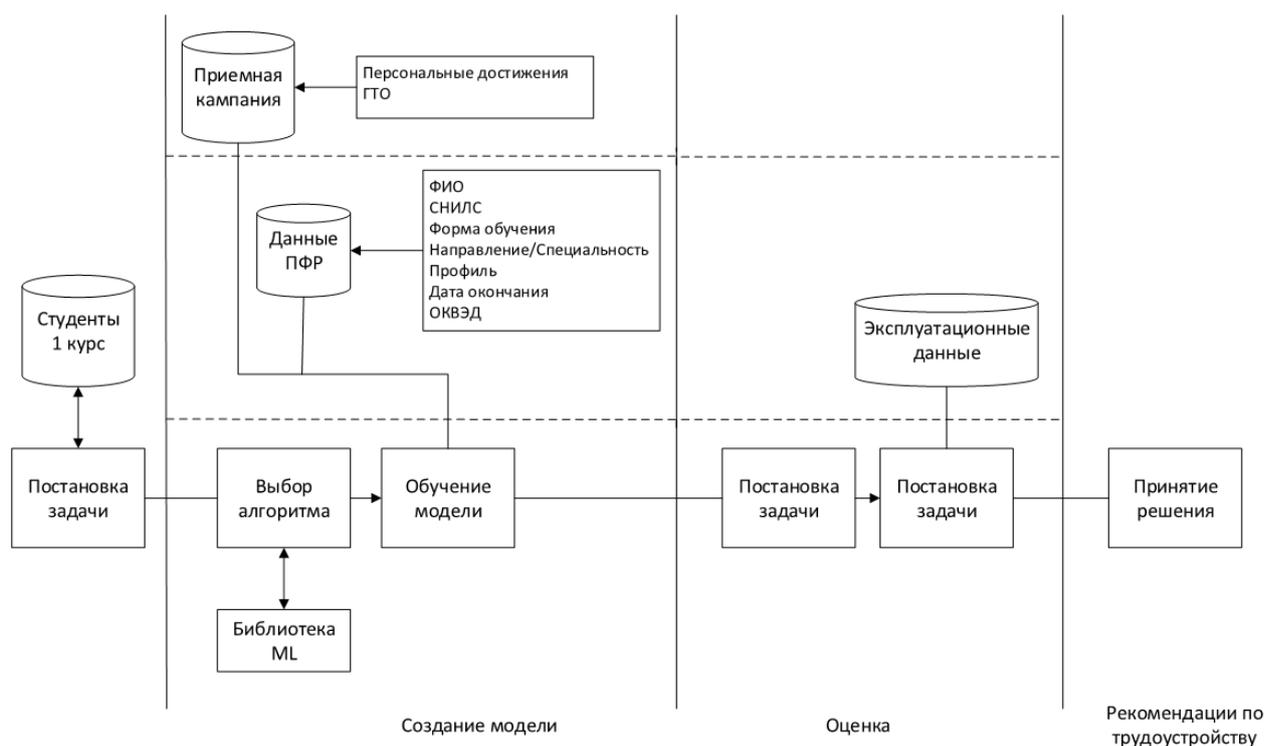


Рисунок 1. Схема работы модели машинного обучения

Таблица 1. Пример данных выпускников из ПФР

Форма	Направление/Специальность	Профиль 1	Дата	Работодатель	Наименование ОКВЭД
заочная	44.03.03: Специальное (дефектологическое) образование	Логопедия	24.10.2022	ОГАУЗ «ОПЦ Им. И.Д. Евтушенко»	Деятельность больничных организаций
очная	44.03.05: Педагогическое образование (с двумя профилями подготовки)	Начальное образование	09.01.2022	Рогожкин Валерий Андреевич	Предоставление услуг по дневному уходу за детьми
очная	44.03.05: Педагогическое образование (с двумя профилями подготовки)	Математика	09.01.2022	МАОУ СОШ № 58 г. Томска	Образование среднее общее
очная	44.03.05: Педагогическое образование (с двумя профилями подготовки)	Математика	09.01.2022	МАОУ Гимназия № 26 г. Томска	Образование среднее общее
заочная	44.03.01: Педагогическое образование	Начальное образование	10.08.2022	МАОУ «ОШ №5 г. Асино»	Образование основное общее
очная	44.03.05: Педагогическое образование (с двумя профилями подготовки)	Технология	09.01.2022	МАОУ Гимназия № 55 им. Е.Г. Версткиной г. Томска	Образование среднее общее
заочная	44.03.01: Педагогическое образование	Дошкольное образование	11.05.2020	МБДОУ Детский сад «Светлячок»	Образование дошкольное
очная	44.03.05: Педагогическое образование (с двумя профилями подготовки)	Русский язык	26.06.2021	МАОУ ДО ДДТ «Созвездие»	Образование дополнительное детей и взрослых

Реализация моделей машинного обучения

Реализовывать алгоритмы машинного обучения было решено на веб-платформе Google Colaboratory (Colab), которая позволяет создать и запускать код на языке Python, не устанавливая на компьютер дополнительных программ.

Python один из наиболее популярных языков для машинного обучения. Основными причинами его использования для написания алгоритмов машинного обучения являются простота использования, большой объем доступных библиотек и широкое пользование [14].

Pandas предоставляет набор инструментов для манипулирования данными (чтение, запись, обработка и анализ). Эта библиотека позволяет работать с данными как со структурой, состоящей из столбцов и строк, что облегчает выполнение различных операций над данными [14].

Sklearn написана на языке Python и использует библиотеки NumPy и SciPy для работы с массивами и линейной алгеброй. Библиотека является простым и эффективным инструментом для предиктивного анализа данных [15].

Для задачи регрессии были выбраны два алгоритма машинного обучения [16]:

1. *Логистическая регрессия.* Логистическая регрессия (Logistic Regression) позволяет делать прогнозы для точек в двоичной системе: 0 или 1. Если значение равно или больше 0,5, объект относится к категории «1». Если значение меньше 0,5 – к катего-

рии «0». Каждый признак имеет свою метку, которая равна либо 0, либо 1. Алгоритм решает задачи бинарной классификации, так как алгоритм применяет сигмоидальную функцию. В данном методе выполняется условие, где $0 \leq Y \leq 1$, что достигается применением сигмоидальной (логистической) функции:

$$Y = \frac{1}{1 + e^{-F(x)}},$$

где $F(x)$ – стандартное уравнение регрессии. При этом нужно учесть, что, если значение не равно 0 или 1, значение аппроксимируется.

2. *Случайный лесной регрессор.* Случайный лесной регрессор (Random Forest Regressor) – это модель машинного обучения, которая может решать различные задачи, такие как классификация, регрессия, кластеризация и другие, используя алгоритм случайного леса. Random Forest состоит из множества базовых деревьев решений, которые работают вместе. Каждое дерево предсказывает класс, и класс, за который «проголосует» большинство деревьев, становится ответом классификатора. Одним из преимуществ модели Random Forest Regressor является ее способность обрабатывать большие объемы данных и работать с большим количеством переменных. Она также устойчива к выбросам в данных и способна выявлять сложные нелинейные зависимости между переменными. Теоретическая часть алгоритма,

Таблица 2. Результаты прогноза с помощью метода «Случайный лес»

№	Место работы	Вероятность трудоустройства
1	Образование среднее общее	0,6919
2	Образование основное общее	0,0685
3	Деятельность в области спорта прочая	0,0630
4	Образование дополнительное детей и взрослых	0,0624
5	Образование высшее	0,0263
6	Деятельность по обработке данных, предоставление услуг по размещению информации и связанная с этим деятельность	0,0249
7	Деятельность по предоставлению прочих вспомогательных услуг для бизнеса, не включенная в другие группировки	0,0236
8	Образование профессиональное среднее	0,0219
9	Управление эксплуатацией жилого фонда за вознаграждение или на договорной основе	0,0175

по сравнению с другими методами, достаточно проста, необходима только формула итогового классификатора:

$$a(x) = \frac{1}{N} \sum_{i=1}^N b_i(x),$$

где N – количество деревьев;

i – счетчик для деревьев;

b – решающее дерево;

x – сгенерированная на основе данных выборка.

Конвейер машинного обучения разделяется на несколько основных этапов: определение задачи, данные, оценка, моделирование и применение. Одно не следует за другим, так как от одного пункта можно вернуться к другому. Сбор данных предполагает обработку сырых данных, то есть информацию из ПФР и данные студентов 1 курса. Под «моделированием» понимается использование алгоритма ML для выявления закономерностей в собранных данных. И последним этапом является применение этой модели уже на эксплуатационных данных. На первоначальном этапе прогнозирования рекомендаций по дальнейшему трудоустройству были указаны основные параметры: направление подготовки, направленность (профили), форма обучения. В дальнейшем для более корректного и точного предсказания есть возможность расширения передаваемых в модель параметров из приемной кампании вуза, где указаны персональные достижения поступающего абитуриента (рисунок 1).

Для обучения модели прогнозирования будущего трудоустройства студентов использовались данные о выпускниках, полученные из ПФР, пример которых представлен в таблице 1. Информация, полученная о выпускниках 2021–2022 года, представляет собой Excel-таблицу с данными бывших студентов, которая содержит: ФИО, СНИЛС, форму обучения, направление/специальность, профиль, дату окончания, работодателя и ОКВЭД (общероссийский классификатор видов экономической деятельности) (таблица 1).

Таким образом, выбрав язык программирования Python, файл с данными обработан программной библиотекой pandas и применен в работе с библиотекой, предназначенной для машинного обучения, scikit-learn (Алгоритм 1 и 2).

Алгоритм 1. Реализация модели машинного обучения LogisticRegression на определенном наборе данных.

```
job = pd.read_excel("job_2021.xlsx")
df = pd.get_dummies(job,
columns=["form", "spec_napravl", "profile"])
y = df.position
x = df.drop('position', axis=1)
model = LogisticRegression(solver='lbfgs',
max_iter=1000)
model.fit(x,y)
val_of = pd.DataFrame(val)
model.predict(val_of)
```

Алгоритм 2. Реализация модели машинного обучения RandomForestRegressor на определенном наборе данных.

```
job = pd.read_excel("job_2021.xlsx")
df = pd.get_dummies(job,
columns=["form", "spec_napravl", "profile"])
y = pd.get_dummies(job.position,
columns=["position"])
x = df.drop(['position'], axis=1)
model = RandomForestRegressor(random_
state=0)
model.fit(x,y)
val_of = pandas.DataFrame(val)
a = model.predict(val_of)
```

Результаты прогнозирования

Модели, обученные на конкретном примере, могут спрогнозировать, кем студент педагогического вуза будет работать в будущем. Для проверки «предсказания» выбраны такие метки для параметров.

1. Направление подготовки: 44.03.05 Педагогическое образование (с двумя профилями подготовки).

2. Направленность (профили): русский язык и литература.

3. Форма обучения: очная.

В результате проделанной работы была разработана интеллектуальная система для прогнозирования трудоустройства выпускников.

После проведенного анализа и вывода прогноза был получен результат: «Студент с такими параметрами, вероятнее всего, станет учителем (оказание услуг по предоставлению общего среднего образования)» (таблица 2).

В результате использования данного алгоритма с учетом входных параметров, описывающих студента педагогического вуза, определяется наиболее вероятный результат его трудоустройства. Путем применения логистической регрессии был получен прогноз, согласно которому студент станет учителем в будущем (рисунок 2).

```
array(['Образование среднее общее'], dtype=object)
```

Рисунок 2. Результат прогноза с помощью метода логистической регрессии

Рассмотренный алгоритм «Случайный лес» представляет собой метаоценщик, который состоит из ряда классификационных деревьев принятия решений, построенных на различных подвыборках набора данных, и использует технику усреднения для повышения точности прогнозов и предотвращения переобучения. Этот алгоритм позволяет получить вероятностное распределение потенциального места работы студента в будущем, которое представлено в виде таблицы 2.

Заключение

В заключение следует отметить, что в ходе работы была обработана информация о трудоустройстве выпускников педагогического вуза, а также разработана интеллектуальная система анализа больших данных, которая позволяет прогнозировать возможное место работы студента в будущем и оценивать качество образовательных услуг. Использование машинного обучения значительно улучшает точность таких прогнозов, а также помогает определить востребованные навыки и знания для работодателей. Это, в свою очередь, позволяет адаптировать учебные программы и повысить конкурентоспособность образовательного учреждения. Кроме того, такая система может быть использована для определения наиболее перспективных направлений для будущих специалистов и оценки качества образовательных услуг.

Исследование выполнено за счет гранта Российского научного фонда № 23-28-00577.

Литература

1. Наговицын Р.С. Прогнозирование трудоустройства выпускников педагогического института на основе технологий искусственного интеллекта // Образовательное пространство в информационную эпоху: материалы международной научно-практической конференции. Москва, 2022. С. 244–253.
2. Наговицын Р.С. Искусственный интеллект для реализации прогноза профессиональной занятости будущих учителей физической культуры // Российское государство, право, экономика и общество: проблемы и пути развития: материалы III Национальной научно-практической конференции. Казань, 2021. С. 86–91.
3. Гладкова И.А., Щанина Е.В. Прогнозирование трудоустройства выпускников вузов // Теоретические и практические аспекты раз-

вития современной науки: теория, методология, практика: материалы международной научно-практической конференции. Уфа, 2019. С. 151–155.

4. Наговицын Р.С. Трудоустройство выпускников педагогических профилей на основе технологий искусственного интеллекта и анализа данных // Мир образования – образование в мире. 2023. № 1 (89). С. 165–176.
5. Фадеев А.С., Змеев О.А., Газизов Т.Т. Модель университета 4.0 // Научно-педагогическое обозрение. 2020. № 2 (30). С. 172–178. DOI: 10.23951/2307-6127-2020-2-172-178
6. Основные направления информатизации деятельности томского государственного педагогического университета / А.Н. Клишин [и др.] // Вестник Томского государственного педагогического университета. 2015. № 3 (156). С. 110–118.
7. Наговицын Р.С. Искусственный интеллект по обработке данных абитуриентов для прогнозирования их дальнейшего трудоустройства после обучения // Региональные вузы-драйверы пространственного развития России: материалы Всероссийской с международным участием научно-практической конференции, посвященной 90-летию Удмуртского государственного университета. Ижевск, 2021. С. 92–104.
8. Изучение опыта прогнозирования туристских потоков с применением алгоритмов машинного обучения / С.А. Лочан [и др.] // Известия высших учебных заведений. Серия: Экономика, финансы и управление производством. 2021. № 4 (50). С. 145–155.
9. Наука о данных. URL: [https://www.tadviser.ru/index.php/Статья:Наука_о_данных_\(Data_Science\)](https://www.tadviser.ru/index.php/Статья:Наука_о_данных_(Data_Science)) (дата обращения: 10.09.2023).
10. Лебедев И.С. Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации // Информационно-управляющие системы. URL: <https://cyberleninka.ru/article/n/adaptivnoe-primenenie-modeley-mashinnogo-obucheniya-na-otdelnyh-segmentah-vybornki-v-zadachah-regressii-i-klassifikatsii> (дата обращения: 15.09.2023).
11. Сейдаметова З.С. Задачи и алгоритмы машинного обучения: вероятностные графические модели // Информационно-компьютерные технологии в экономике, образовании и социальной сфере. 2019. № 1 (23). С. 180–187.
12. Згонникова А.О., Прокопенко А.А. Машинное обучение и обучение на протяжении всей

- жизни // Новые научные исследования: материалы VIII Международной научно-практической конференции. Пенза, 2022. С. 22–24.
13. Бородин И.Д. Рефлексия в машинном обучении на примере обучения деревьев решений // Аллея науки. 2017. Т. 4, № 9. С. 857–865.
14. Аналитикам: большая шпаргалка по Pandas. URL: <https://smysl.io/blog/pandas/> (дата обращения: 19.09.2023).
15. Scikit-learn. Машинное обучение в Python. URL: <https://scikit-learn.ru/> (дата обращения: 19.09.2023).
16. Микшина В.С., Павлов С.И. Принятие решений с использованием ансамбля классификаторов // Информационные технологии в науке, образовании и управлении. 2019. № 2 (12). С. 50–54.

Получено 07.11.2023

Селиванова Елизавета Сергеевна, магистрант кафедры информатики Томского государственного педагогического университета (ТГПУ). 634061, Российская Федерация, г. Томск, ул. Герцена, 49. Тел. +7 961 096-90-90. E-mail: slvnva@tspu.edu.ru

Газизов Тимур Тальгатович, д.т.н., профессор кафедры информатики ТГПУ. 634061, Российская Федерация, г. Томск, ул. Герцена, 49; научный сотрудник центра экспертизы и испытаний Сахалинского государственного университета. 693008, Российская Федерация, г. Южно-Сахалинск, пр. Коммунистический, 33. Тел. +7 382-2 31-13-66. E-mail: gtt@tspu.edu.ru

Долганова Надежда Филипповна, старший преподаватель кафедры информатики ТГПУ. 634061, Российская Федерация, г. Томск, ул. Герцена, 49. Тел. +7 923 422-47-45. E-mail: dolganovaNF@tspu.edu.ru

Пираков Фаррух Джамshedович, аспирант кафедры технологий электронного обучения Томского государственного университета систем управления и радиоэлектроники. 634050, Российская Федерация, г. Томск, пр-т Ленина, 40. Тел. +7 952 157-21-24. E-mail: pfd@tspu.edu.ru

USING MACHINE LEARNING METHODS FOR PREDICTING THE EMPLOYMENT OF GRADUATES

Selivanova E.S.¹, Gazizov T.T.^{1,3}, Dolganova N.F.¹, Pirakov F.D.²

¹Tomsk State Pedagogical University, Tomsk, Russian Federation

²Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

³Sakhalin State University, Yuzhno-Sakhalinsk, Russian Federation

E-mail: slvnva@tspu.edu.ru

In today's world, technology continues to play an increasingly important role in all spheres of human life. Education is no exception and keeps pace with the times. One of the more promising directions in this area is the use of machine learning methods to analyze data on graduates and predict their future employment, based on the accumulated information about students. Such information not only helps to determine the future profession, but also will allow educational institutions to co-ordinate their activities more effectively and improve the quality of education. In addition, the use of machine learning in education can lead to the creation of new, more effective teaching methods that take into account the individual characteristics of each student. The result of this work is an intelligent system of big data processing, capable of adapting to the current state of the labor market and helping graduates in an earlier determination of their future profession. Data analysis was conducted on a sample of students of Tomsk State Pedagogical University of 2021-2023 academic years.

Keywords: *Data analysis, employment, machine learning, information processing, artificial intelligence, model, prediction*

DOI: 10.18469/ikt.2023.21.3.13

Selivanova Elizabeth Sergeevna, Tomsk State Pedagogical University, 49, Herzen Street, Tomsk, 634061, Russian Federation; Master's Degree Student of Informatics Department. Tel. +7 961 096-90-90. E-mail: slvnva@tspu.edu.ru

Gazizov Timur Talgatovich, Tomsk State Pedagogical University, 49, Herzen Street, Tomsk, 634061, Russian Federation; Professor of Computer Science Department, Doctor of Technical Science. Sakhalin State University, 33, Kommunisticheskoy Avenue, Yuzhno-Sakhalinsk, 693008, Russian Federation; Researcher at the Center for Expertise and Testing. Tel. +7 382-2 31-13-66. E-mail: gtt@tspu.edu.ru

Dolganova Nadezhda Pilippovna, Tomsk State Pedagogical University, 49, Herzen Street, Tomsk, 634061, Russian Federation; Senior lecturer of Informatics Department. Tel. +7 923 422-47-45. E-mail: dolganovaNF@tspu.edu.ru

Pirakov Farrukh Jamshedovich, Tomsk State University of Control Systems and Radioelectronics, 40, Lenina Prospect, 634050, Tomsk, Russian Federation; Department of E-Learning Technologies, Automation, PhD Student. Tel. +7 952 157-21-24. E-mail: pfd@tspu.edu.ru

References

1. Nagovitsyn R.S. Forecasting the employment of graduates of a pedagogical institute based on artificial intelligence technologies. *Obrazovatel'noe prostranstvo v informacionnuyu epokhu: materialy mezhdunarodnoj nauchno-prakticheskoy konferencii*. Moscow, 2022, pp. 244–253. (In Russ.)
2. Nagovitsyn R.S. Artificial intelligence for implementing the forecast of professional employment of future physical education teachers. *Rossiyskoe gosudarstvo, pravo, ekonomika i obshchestvo: problemy i puti razvitiya: materialy III Nacional'noj nauchno-prakticheskoy konferencii*. Kazan, 2021. pp. 86–91. (In Russ.)
3. Gladkova I.A., Shchanina E.V. Forecasting job placement of university graduates. *Teoreticheskie i prakticheskie aspekty razvitiya sovremennoj nauki: teoriya, metodologiya, praktika: materialy Mezhdunarodnoj nauchno-prakticheskoy konferencii*. Ufa, 2019. pp. 151–155. (In Russ.)
4. Nagovitsyn R.S. Employment of graduates of pedagogical profiles based on artificial intelligence technologies and data analysis. *Mir obrazovaniya – obrazovanie v mire*, 2023, no. 1 (89), pp. 165–176. (In Russ.)
5. Fadeev A.S., Zmeev O.A., Gazizov T.T. University Model 4.0. *Nauchno-pedagogicheskoe obozrenie*, 2020, no. 2, pp. 172–178. (In Russ.)
6. Klishin A.N. et al. Main directions for applying information technologies to the automation of tspu activities. *Vestnik Tomskogo gosudarstvennogo pedagogicheskogo universiteta*, 2015, vol. 3 (156), pp. 110–118. (In Russ.)
7. Nagovitsyn R.S. Artificial intelligence for processing the data of applicants to predict their further employment after training. *Regional'nye vuzy-drajvery prostranstvennogo razvitiya Rossii: materialy Vserossiyskoy s mezhdunarodnym uchastiem nauchno-prakticheskoy konferencii, posvyashchennoj 90-letiyu Udmurtskogo gosudarstvennogo universiteta*. Izhevsk, 2021, pp. 92–104. (In Russ.)
8. Lochan S.A. et al. Study of experience in forecasting tourist flows using machine learning algorithms. *Izvestiya vysshih uchebnyh zavedeniy. Seriya: Ekonomika, financy i upravlenie proizvodstvom*, 2021, vol. 4 (50), pp. 145–155. (In Russ.)
9. Data science. URL: [https://www.tadviser.ru/index.php/Статья:Наука_о_данных_\(Data_Science\)](https://www.tadviser.ru/index.php/Статья:Наука_о_данных_(Data_Science)) (accessed: 10.09.2023). (In Russ.)
10. Lebedev I.S. Adaptive application of machine learning models on separate sample segments in regression and classification tasks. *Informacionno-upravlyayushchie sistemy*. URL: <https://cyberleninka.ru/article/n/adaptivnoe-primenenie-modeley-mashinnogo-obucheniya-na-otdelnyh-segmentah-vyboriki-v-zadachah-regressii-i-klassifikatsii> (accessed: 15.09.2023). (In Russ.)
11. Seidametova Z.S. Machine learning problems and algorithms: probabilistic graphical models. *Informacionno-komp'yuternye tekhnologii v ekonomike, obrazovanii i social'noj sfere*, 2019, no. 1 (23), pp. 180–187. (In Russ.)
12. Zgonnikova A.O., Prokopenko A.A. Machine learning and lifelong learning. *New scientific*

research: Novye nauchnye issledovaniya: materialy VIII Mezhdunarodnoj nauchno-prakticheskoy konferencii, Penza, 2022, pp. 22–24. (In Russ.)

13. Borodin I.D. Reflexion in machine learning on the example of decision tree learning. *Alleya nauki*, 2017, vol. 4, no. 9, pp. 857–865. (In Russ.)
14. For Analysts: the big cheat sheet on Pandas. URL: <https://smysl.io/blog/pandas/> (accessed: 19.09.2023). (In Russ.)
15. Scikit-learn. Machine learning in Python. URL: <https://scikit-learn.ru/> (accessed: 19.09.2023). (In Russ.)
16. Mikshina V.S., Pavlov S.I. Decision-making using an ensemble of classifiers. *Informacionnye tekhnologii v nauke, obrazovanii i upravlenii*, 2019, no. 2 (12). pp. 50–54. (In Russ.)

Received 07.11.2023