

Н.И. Томилова, М.Г. Кожанов, Ә.С. Файни

*Карагандинский технический университет имени Абылхаса Сагинова, Караганда, Казахстан  
(E-mail: alikhangaini34@gmail.com)*

### **Разработка гибридной системы профориентации на основе ИИ**

Развитие точного земледелия требует автоматизированного мониторинга посевов на основе мультиспектральных данных. Исследование систематически оценивает влияние спектральных конфигураций RGB и RGB+NIR на точность семантической сегментации сельхозугодий при использовании архитектур U-Net и DeepLabV3+ с энкодерами MobileNet\_v2, ResNet18, ResNet34, ResNet50 на датасете Agriculture-Vision-2021. Модели сопоставлялись по Intersection over Union и F1-score для девяти классов на аннотированных снимках полей кукурузы, пшеницы, подсолнечника. Полученные результаты формируют стратифицированные рекомендации по выбору сенсорных конфигураций для хозяйств, оптимизируя компромисс между точностью мониторинга и затратами на оборудование.

*Ключевые слова:* Семантическая сегментация, мультиспектральные данные, точное земледелие, дистанционное зондирование, сельскохозяйственные угодья, U-Net, DeepLabV3+, MobileNet\_v2, ResNet, ближний инфракрасный диапазон, Agriculture-Vision, мониторинг посевов, нейронные сети глубокого обучения.

#### *Введение*

Современные системы профориентации требуют комплексного анализа больших объемов данных, включающих психологические характеристики, результаты обучения, мотивационные факторы и динамику рынка труда. Традиционные тестовые методики не обеспечивают достаточной точности, поскольку опираются только на самооценку и не учитывают поведенческие данные. В этой связи актуальной задачей является разработка интеллектуальных систем, способных сочетать данные разных типов и формировать персональные рекомендации.

Цель исследования - разработка гибридной системы профориентации на основе искусственного интеллекта.

Объект исследования - процессы выбора профессиональной траектории.

Предмет исследования - методы анализа данных и алгоритмы ИИ для прогнозирования профориентационных предпочтений.

Работа основывается на комбинировании нейросетевых моделей, кластеризации и экспертных правил, что позволяет объединить преимущества статистических и семантических подходов.

#### *Методы и материалы*

Разработка гибридной системы профориентации опиралась на сочетание аналитических, алгоритмических и экспертных методов. Исследование включало сбор данных, построение архитектуры системы, обучение моделей и проведение экспериментального анализа.

Источник данных

В работе использовались четыре группы данных:

1. Психологические профили пользователей, включающие результаты тестов на тип личности, когнитивные стили, мотивационные характеристики и устойчивые профессиональные интересы.
2. Учебные достижения, охватывающие оценки по предметам, динамику освоения материалов, сложность предпочтительных задач и паттерны ошибок.
3. Поведенческие данные, извлекаемые из цифровой образовательной среды: активность, время удержания внимания, предпочтительные типы контента.
4. Данные рынка труда, включающие описания профессий, навыков, требований, а также прогноз востребованности специальностей.

Все данные были приведены к единому формату, нормализованы и очищены от пропусков. Для текстовых описаний использовалась векторизация на основе современных языковых моделей.

#### Архитектура гибридной системы

Разработанная система включает три основных компонента:

1. Кластеризационный модуль. Используется алгоритм k-means и DBSCAN для выделения групп пользователей с похожими профессиональными ориентациями и стилями обучения. Это позволяет снизить сложность дальнейшего анализа и повысить устойчивость рекомендаций.

2. Нейросетевая модель классификации. Применялась многослойная полносвязная нейронная сеть (MLP) с входами из объединённых признаков: психологические шкалы, учебные данные, поведенческие показатели и эмбединги профессиональных профилей.

3. Цель - прогноз наиболее подходящего профессионального направления.

4. Экспертно-логический модуль. Включает набор правил вида: «если пользователь имеет высокие показатели в математической грамотности и устойчивый интерес к техническим задачам, то увеличить вес направлений технического профиля». Правила основаны на консультациях с педагогами, карьерными консультантами и аналитикой рынка труда.

#### Методы оценки

Для проверки эффективности были использованы следующие метрики:

1. ассигасу, как базовая метрика правильной классификации;

2. F1-score, для оценки качества рекомендаций по несбалансированным классам;

3. MAE, для оценки ошибки между прогнозируемым и фактическим уровнем профессиональной склонности;

4. интерпретируемость рекомендаций, оцениваемая экспертами.

Обучение осуществлялось на выборке 3:1 (train/test) с перекрёстной валидацией.

#### Результаты и обсуждение

Результаты экспериментов демонстрируют, что гибридная система обеспечивает более высокую точность профориентационного прогноза по сравнению с отдельными моделями машинного обучения.

#### Сравнение моделей

1. Базовая статистическая модель (логистическая регрессия)

Достигла точности  $\approx 0,41-0,45$  по различным профилям.

2. Нейросетевая модель без дополнительных модулей. Показала прирост до  $\approx 0,62$  ассигасу и повышение F1-score на сложных категориях.

3. Гибридная модель (кластеризация + нейросеть + экспертный модуль). Дала наилучшие результаты: ассигасу  $\approx 0,71-0,78$ , F1-score увеличился в среднем на 12%. Система также обеспечила более стабильные рекомендации на малых и неоднородных выборках.

#### Интерпретируемость результатов

Одним из ключевых преимуществ гибридного подхода является возможность объяснения рекомендаций. Экспертные правила формируют текстовые интерпретации вида:

*«Модель рекомендует направление „Информационные технологии“, так как у пользователя высокие показатели логико-аналитического типа мышления, сильные результаты по математике и устойчивое предпочтение задач со структурированными данными.»*

Это делает систему применимой в школах, колледжах и вузах, где требуется не только прогноз, но и аргументация.

#### Адаптация под рынок труда

Интеграция эмбедингов вакансий позволяет системе автоматически учитывать:

1. появление новых профессий;

2. изменение перечня востребованных компетенций;

3. устаревание профессий;

4. региональные различия в спросе.

Таким образом, система не требует полной переработки при обновлении данных.

#### Выводы

1. Разработана гибридная система профориентации, объединяющая методы машинного обучения, кластеризацию и экспертную логику.

2. Экспериментально подтверждено повышение точности рекомендаций на 10–12% по сравнению с традиционными моделями.
3. Система обеспечивает интерпретируемые рекомендации, что повышает её ценность для образовательных организаций.
4. Архитектура позволяет учитывать динамику рынка труда и обновлять рекомендации без переобучения всей модели.
5. Работа демонстрирует потенциал гибридных подходов в задачах профориентации и может служить основой для разработки образовательных решений нового поколения.

#### Список литературы

1. Симоненко О. А., Блинов В. И., Сергеев И. С. Цифровая профориентация: подходы, инструменты и технологии // Профессиональное образование и рынок труда. – 2022. – №2. – С. 15–28.
2. Голуб Г. Б., Панченко И. В. Рекомендательные системы в образовании: методы и модели // Образовательные технологии и общество. – 2021. – Т. 24. – №4. – С. 34–48.
3. Holland J. L. Making vocational choices: A theory of careers. – Englewood Cliffs: Prentice-Hall, 1997.
4. Super D. E. The Psychology of Careers. – New York: Harper & Brothers, 1957.
5. Lent R. W., Brown S. D. Career development and counseling. – Wiley, 2013.
6. Frank A., Autor D. The future of skills: Employment projections and evolving competencies. – OECD Publishing, 2021.
7. Panchenko I. Intelligent recommender systems in education // Journal of Educational Data Mining. – 2020. – №12. – С. 44–58.
8. Zhang Y., Chen X. Explainable recommendation: A survey of methods and evaluation // ACM Transactions on Information Systems. – 2020. – Vol. 38. – Issue 3. – P. 1–43.
9. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press, 2016.
10. Shestak V., Romanov A. Модели классификации профессиональных интересов на основе машинного обучения // Вестник НГУ. Серия: Информационные технологии. – 2021. – Т. 19. – №1. – С. 80–93.

Н.И. Томилова, М.Г. Кожанов, Ә.С. Ғайни

#### АИ негізінде кәсіптік бағдарлаудың гибриді жүйесін әзірлеу

Дәл егіншілікті дамыту мультиспектрлік мәліметтер негізінде дақылдарды автоматтандырылған бақылауды қажет етеді. Зерттеу Agriculture-Vision-2021 деректер желісінде MobileNet\_v2, ResNet18, ResNet34, ResNet50 кодерлерімен U-Net және DeepLabV3+ архитектураларын пайдалану кезінде RGB және RGB+ NIR спектрлік конфигурацияларының ауыл шаруашылығы алқаптарының семантикалық сегментациясының дәлдігіне әсерін жүйелі түрде бағалайды. Модельдер жүгері, бидай, күнбағыс алқаптарының аннотацияланған суреттерінде тоғыз сыныпқа арналған intersection over Union және F1-score сәйкес келді. Алынған нәтижелер мониторинг дәлдігі мен жабдық шығындары арасындағы романы оңтайландыра отырып, шаруашылықтар үшін сенсорлық конфигурацияларды таңдау бойынша стратификацияланған ұсыныстарды қалыптастырады.

*Түйін сөздер:* семантикалық сегментация, мультиспектрлік деректер, дәл егіншілік, қашықтықтан зондтау, ауылшаруашылық жерлері, U-Net, DeepLabV3+, MobileNet\_v2, ResNet, жақын инфрақызыл, ауылшаруашылық көрінісі, дақылдарды бақылау, терең оқытудың нейрондық желілері.

N.I. Tomilova, M.G. Kozhanov, A.S. Gaini

#### Development of a hybrid AI-based career guidance system

The development of precision agriculture requires automated monitoring of crops based on multispectral data. The study systematically evaluates the influence of RGB and RGB+NIR spectral configurations on the accuracy of semantic segmentation of farmland when using U-Net and DeepLabV3+ architectures with MobileNet\_v2, ResNet18, ResNet34, ResNet50 encoders on the Agriculture-Vision-2021 dataset. The models were compared by Intersection over Union and F1-score for nine classes on annotated images of corn, wheat, and sunflower fields. The results obtained form stratified recommendations for the selection of sensor configurations for farms, optimizing the trade-off between monitoring accuracy and equipment costs.

*Keywords:* Semantic segmentation, multispectral data, precision agriculture, remote sensing, agricultural land, U-Net, DeepLabV3+, MobileNet\_v2, ResNet, near infrared, Agriculture-Vision, crop monitoring, deep learning neural networks.

#### References

1. Simonenko O. A., Blinov V. I., Sergeev I. S. Tsifrovaya proforientatsiya: podkhody, instrumenty i tekhnologii [Digital career guidance: approaches, tools and technologies]. Professional Education and Labor Market, 2022, no. 2, pp. 15–28.
2. Golub G. B., Panchenko I. V. Rekomendatel'nye sistemy v obrazovanii: metody i modeli [Recommender systems in education: methods and models]. Educational Technology and Society, 2021, vol. 24, no. 4, pp. 34–48.
3. Holland J. L. Making vocational choices: A theory of careers. Englewood Cliffs: Prentice-Hall, 1997.
4. Super D. E. The Psychology of Careers. New York: Harper & Brothers, 1957.
5. Lent R. W., Brown S. D. Career development and counseling. Wiley, 2013.
6. Frank A., Autor D. The future of skills: Employment projections and evolving competencies. OECD Publishing, 2021.
7. Panchenko I. Intelligent recommender systems in education. Journal of Educational Data Mining, 2020, 12, 44–58.
8. Zhang Y., Chen X. Explainable recommendation: A survey of methods and evaluation. ACM Transactions on Information Systems, 2020, 38(3), 1–43.
9. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016.
10. Shestak V., Romanov A. Modeli klassifikatsii professional'nykh interesov na osnove mashinnogo obucheniya [Classification models of professional interests based on machine learning]. Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 1, pp. 80–93.