

МОДЕЛИРОВАНИЕ СИСТЕМЫ ПРОФОРИЕНТАЦИИ С ИСПОЛЬЗОВАНИЕМ АНАЛИЗА ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ*

Иващенко А. О.¹, научный сотрудник, ✉ aok@dscs.pro

Вяткин А. А.¹, младший научный сотрудник, aav@dscs.pro

Бушмелёв Ф. В.¹, научный сотрудник, fvb@dscs.pro

Абрамов М. В.^{1,2}, канд. техн. наук, ведущий научный сотрудник, доцент, mva@dscs.pro

¹ Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН),
14 линия, 39, 199178, Санкт-Петербург, Россия

² Санкт-Петербургский государственный университет,
Университетская наб., д. 7–9, 199034, Санкт-Петербург, Россия

Аннотация

В статье рассматривается проблема автоматизации профориентации на основе анализа цифрового следа пользователей социальной сети «ВКонтакте». Цель исследования заключается в повышении доступности и точности диагностики профессиональных интересов посредством мини-приложения «AI профориентатор» на платформе VK Mini Apps. В основе методологии лежит задача классификации с несколькими метками: для каждого из шести профессиональных типов модели RIASEC формируется отдельный бинарный выход с сигмоидальной активацией, а обучение осуществляется с использованием функции потерь Binary Cross-Entropy. Модель обучена на данных пользователей, прошедших тест Голланда, и их подписках на сообщества, собранных через приложение «Психологические тесты». Техническая реализация построена по принципу клиент-серверного взаимодействия с архитектурой REST API, включающей обработку данных, взаимодействие с ML-сервером и формирование результатов. Разработанная система обеспечивает полный цикл работы: извлечение и валидация подписок, предсказание профессионального типа личности по шести категориям Голланда и отображение результатов пользователю. Модель продемонстрировала Top-1 точность 47,1 % и Top-2 — 72,3 %, что подтверждает её применимость в профориентационной диагностике. Созданная система автоматизирует процесс первичной профориентации и может использоваться как индивидуальными пользователями, так и организациями для оптимизации подбора и развития персонала. Новизна работы заключается в комплексной интеграции нейросетевых методов и архитектуры клиент-серверного приложения в социальную сеть, обеспечивающей персонализированный и масштабируемый подход к профориентации.

Ключевые слова: профориентация, цифровой след, машинное обучение, социальные сети, клиент-серверное приложение, предсказательная модель, тест Голланда.

* Работа выполнена в рамках проекта по государственному заказу СПб ФИЦ РАН Мол_лаб №FFZF-2024-0003, поддержана Санкт-Петербургским государственным университетом, проект №75254082.

Цитирование: Иващенко А. О., Вяткин А. А., Бушмелёв Ф. В., Абрамов М. В. Моделирование системы профориентации с использованием анализа данных социальных сетей // Компьютерные инструменты в образовании. 2025. № 2. С. 36–47. doi:10.32603/2071-2340-2025-2-36-47

1. ВВЕДЕНИЕ

Проблема профессиональной ориентации сохраняет актуальность как для обучающихся, так и для взрослого населения и работодателей. Традиционные подходы, основанные на анкетировании и тестировании, обладают ограниченной масштабируемостью и требуют значительных временных затрат. В условиях цифровизации актуальным становится поиск инструментов, которые позволяют использовать данные о поведении пользователей в сети и методы искусственного интеллекта для ускорения и персонализации профориентационной диагностики [1–3]. Современные алгоритмы машинного обучения демонстрируют высокую эффективность при прогнозировании карьерных предпочтений на основе интеграции разнородных источников данных — социальных сетей, образовательных траекторий и психологических характеристик [2]. Однако доступ к подобным системам остаётся ограниченным, что делает востребованным создание решений, доступных широкой аудитории и ориентированных на практическое применение в учебном процессе и кадровой политике.

Настоящая статья посвящена созданию интеллектуальной системы профориентации, интегрированной в социальную сеть «ВКонтакте». Центральным элементом системы выступает модуль анализа цифрового следа, в котором данные о подписках пользователя на сообщества обрабатываются методами машинного обучения с целью прогнозирования профессионального типа. Тем самым устраняется необходимость прохождения длительных тестов и анкетирования. Научная задача исследования заключается в разработке полноценной программной платформы, использующей многоэтапный сбор данных из социальной сети, их предобработку и применение алгоритмов машинного обучения для автоматизированного определения профессиональных склонностей.

2. РЕЛЕВАНТНЫЕ РАБОТЫ

Разработка интеллектуальных профориентационных систем активно ведётся в контексте поддержки абитуриентов и студентов. Так, в работе [4] описана университетская веб-система на базе анализа профилей «ВКонтакте», сопоставляющая цифровые признаки с академическими данными вуза и формирующая рекомендации по направлениям подготовки. Другой класс решений базируется на опросниках и психометрических тестах, результаты которых обрабатываются с помощью моделей XGBoost и Random Forest [5–7]. В ряде исследований применяются алгоритмы SVM, XGBoost и методы нечёткой логики для анализа академических и личностных характеристик обучающихся [8–10], наряду с этим разрабатываются ансамблевые подходы [11, 12]. Однако большинство указанных систем нацелено прежде всего на выбор направления обучения и сопровождение образовательных траекторий, а не на выявление и прогнозирование профессиональной идентичности в более широком смысле.

Вторая линия работ соотносит навыки обучающихся с требованиями рынка труда. Здесь используются анализ вакансий и резюме, KNN, кластеризация и нейросетевые ар-

хитектуры [13–16], а также диалоговые агенты [17, 18] и гибридные модели с использованием RNN и LSTM для повышения точности рекомендаций [19]. Недостатками таких систем являются ограниченность набора профессий и необходимость ручного ввода данных.

Близким направлением исследований является автоматизация оценки личностных особенностей: авторы исследований [20, 21] оценивают данные характеристики на основе анализа резюме соискателей, в [22, 23] предлагается строить описание личности на основе анализа постов в социальных сетях, в работе [24] предсказание строится на основе тематик подписок пользователей. Данные исследования могут служить основой для построения системы, однако сами по себе не являются доступными для конечного пользователя.

Особый интерес представляют приложения платформы VK Mini Apps. Анализ показал наличие восьми релевантных решений, которые можно условно классифицировать на три группы: классические психологические тесты, опросники о предпочтениях пользователя для подбора профессии и приложения информационно-ознакомительного характера. Наиболее близко к задачам профориентационной диагностики стоит приложение, комбинирующее подписки пользователя с кратким опросником и использующее CatBoost [25]. Однако данное решение ограничено небольшой обучающей выборкой (1252 респондента) и не поддерживает анализ сторонних профилей. В целом, существующие решения базируются преимущественно на анкетировании и лишь в ограниченной степени используют современные подходы интеллектуального анализа данных. Нейросетевые архитектуры, методы многометочной классификации и интеграция с внешними ИИ-сервисами пока не находят в них применения. Это определяет сохраняющийся разрыв между возможностями современных технологий и практикой образовательных приложений, что и обуславливает актуальность и новизну представленного исследования.

3. ПОСТАНОВКА ЗАДАЧИ

Разработка инструментов для оперативной и точной профориентационной диагностики на основе цифрового следа пользователей социальных сетей представляет собой актуальную научно-практическую задачу. Целью настоящего исследования является улучшение оперативности профориентационной диагностики путем разработки полноценной программной платформы, использующей сбор данных из социальной сети, их предобработку и применение алгоритмов машинного обучения для автоматизированного определения профессиональных склонностей. Основными этапами для достижения поставленной цели является: разработка модели машинного обучения для предсказания профессионального типа пользователя на основе цифрового следа (тематики сообществ, на которые он подписан), создание архитектуры и пользовательского интерфейса приложения, а также интеграция предсказательной модели в инфраструктуру VK Mini Apps, включая взаимодействие с ML-сервером для обработки и анализа данных.

Новизна работы заключается в комплексном подходе к интеграции multi-label классификации и архитектуры клиент-серверного приложения на платформе VK Mini Apps. В отличие от традиционных решений, ориентированных преимущественно на анкетирование, разработанная система позволяет существенно ускорить процесс выявления профессиональных интересов. Теоретическая значимость исследования связана с расширением методов интеллектуального анализа данных в области профориентации, в том числе с учётом особенностей цифрового следа пользователей в социальной сети. Практиче-

ская значимость выражается в возможности применения разработанного инструмента как в индивидуальной работе с учащимися, так и в рамках учебных курсов и профориентационных программ. Это способствует повышению качества профориентационной поддержки, оптимизации образовательных траекторий и формированию более осознанного выбора профессии, что соответствует задачам устойчивого развития системы образования.

4. МЕТОДЫ ИССЛЕДОВАНИЯ

4.1. Предсказательная модель

Основой разработки модели предсказания профессионального типа пользователя стала гипотеза о взаимосвязи между тематиками сообществ, на которые подписан пользователь в социальной сети «ВКонтакте», и его профессиональным типом [24]. Для построения модели использовался специализированный набор данных, включающий результаты тестирования Голланда и подписки пользователей на сообщества в социальной сети. Данные были получены посредством приложения «Психологические тесты», а информация о тематиках подписок — с помощью разработанного модуля загрузки. Обезличенные данные были предобработаны, после чего сформирована итоговая выборка из 6029 записей, включающая: нормализованный вектор признаков, характеризующий укрупнённые тематики подписок для каждого пользователя; пол пользователя, указанный в профиле социальной сети; оценки выраженности профессиональных типов (кодов Голланда), полученные на основе результатов психологического тестирования пользователей.

На первом этапе задача формулировалась как мультиклассовая классификация с выделением преобладающего профессионального типа. Для этого вектор, описывающий коды Голланда, преобразовывался в один целевой признак — наиболее высокий по величине код для каждого пользователя. Был проведён анализ эффективности различных алгоритмов, включая: Random Forest, Gradient Boosting, LightGBM, Extra Trees, CatBoost, AdaBoost. Наилучшие результаты показала модель Extra Trees ($F1\text{-weighted} = 0,6599$, $n_{\text{estimators}} = 300$). Однако при тестировании на отложенной выборке наблюдалось значительное снижение показателей ($\text{weighted avg } F1 = 0,26$, $\text{macro avg } F1 = 0,28$), что свидетельствовало о слабой обобщающей способности модели.

Дальнейший анализ распределения кодов Голланда показал, что разница между доминирующим и вторым по выраженности типом у многих пользователей составляет менее 2,5 %. Это позволило предположить наличие нескольких преобладающих типов и перейти к постановке задачи классификации с несколькими метками (multi-label classification).

Для реализации данного подхода была применена глубокая нейросетевая архитектура с остаточными связями (Residual Connections). В качестве выходного слоя использовалось шесть нейронов, каждый из которых соответствовал одному из профессиональных типов и имел сигмоидную функцию активации, позволяющую интерпретировать предсказанное значение как вероятность принадлежности к соответствующему классу. Таким образом, каждая предсказанная вероятность отражала степень выраженности того или иного профессионального типа у пользователя. Оптимизация гиперпараметров осуществлялась с использованием алгоритма Optuna. Данные были разделены на обучающую и тестовую выборки для оценки точности предсказаний на обучающих данных применялась k-fold кросс-валидация. На этапе инференса формировался вектор вероят-

ностей принадлежности к каждому из шести профессиональных типов. По сравнению с классическими методами машинного обучения, нейросетевая модель с многовыходной регрессией позволила достичь существенного улучшения качества. В частности, значение Hamming Loss (доля неверных предсказаний меток к общему числу меток) составило 0,348, а вероятность правильного попадания в одну из двух наиболее вероятных категорий для пользователя составила 0,723.

4.2. Проектирование архитектуры программной системы

Для интеграции решения в социальную сеть «ВКонтакте» приложение разрабатывалось согласно правилам платформы VK Mini Apps как одностраничное веб-приложение. Разработка архитектуры приложения велась с использованием модели REST API, что обеспечивает модульность, масштабируемость и возможность дальнейшего расширения функционала.

Клиентская часть приложения отправляет данные пользователей на сервер через HTTP-запросы. Серверная часть принимает запросы, обрабатывает их с помощью контроллеров, выполняет валидацию данных и взаимодействует с внешним сервером машинного обучения для получения предсказаний. Полученные результаты сохраняются в базе данных и передаются обратно клиенту. На рисунке 1 представлена архитектура приложения, демонстрирующая ключевые компоненты системы: клиентская часть, серверная часть, сервер машинного обучения и база данных.

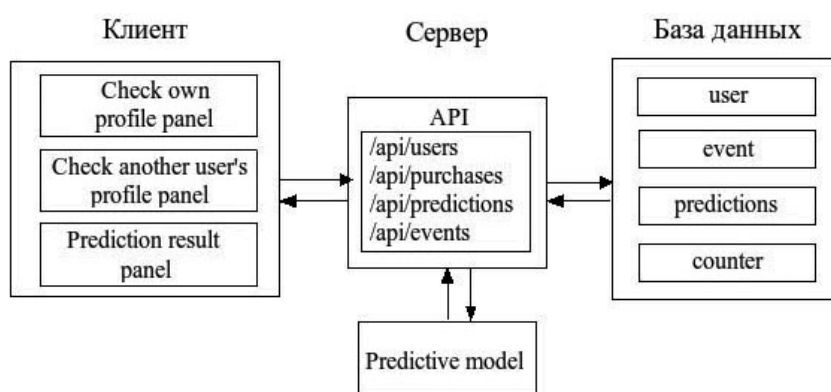


Рис. 1. Архитектура приложения

4.3. Программная реализация и интеграция модели машинного обучения

В приложении предусмотрено два сценария: определение собственной профессиональной направленности и получение предсказаний по профилям других пользователей. На рисунке 2 представлена UML-диаграмма последовательности работы приложения для получения предсказания пользователю по собственному профилю. Опишем её подробнее.

Предсказательная модель использует данные о тематиках подписок пользователей, поэтому в приложении реализовано взаимодействие с VK Bridge. На клиентской стороне выполняется запрос к VK Bridge для получения информации о пользователе и его сообществах, данные приводятся к структурированному виду и отправляются на сервер.

Серверная часть осуществляет валидацию и обработку данных, после чего они передаются на ML-сервер, где формируется вероятностное распределение по шести професси-

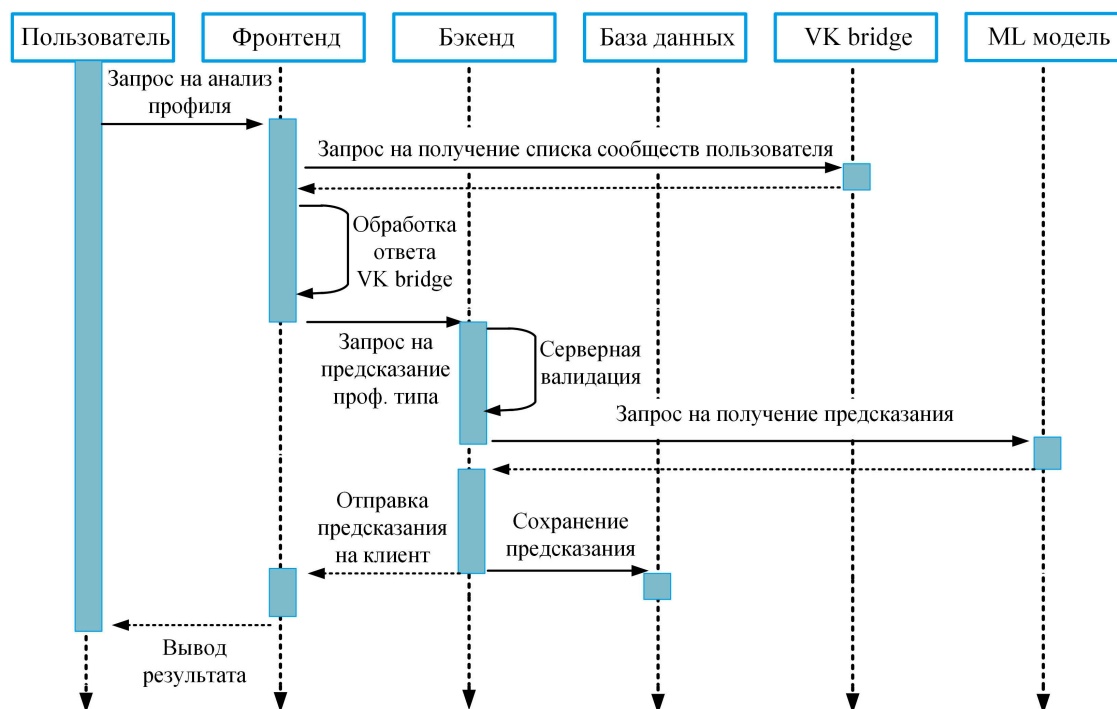


Рис. 2. UML-диаграмма последовательности работы приложения при запросе пользователем собственного предсказания

ональным типам RIASEC. Пользователю отображается как доминирующий тип, так и распределение вероятностей по всем категориям, что обеспечивает более гибкую интерпретацию результата. Полученные данные сохраняются в базе для последующего использования.

Сервер построен по модульной архитектуре, включающей маршруты, контроллеры и модели. Маршруты связывают клиентские запросы с контроллерами, реализующими бизнес-логику, а контроллеры взаимодействуют с моделями для работы с базой данных. Такая структура обеспечивает расширяемость и упрощает интеграцию приложения в образовательные цифровые среды.

На рисунке 3 представлен интерфейс мини-приложения «AI профориентатор» на платформе VK Mini Apps, демонстрирующий результат работы системы.

5. ВЫВОДЫ

Разработанный подход к профориентации, основанный на multi-label классификации, продемонстрировал существенное повышение точности предсказания профессионального типа пользователя по сравнению с традиционной мультиклассовой постановкой и обеспечил Тор-1 точность предсказания 47,1% и Тор-2 — 72,3%, что подтверждает её применимость для практических задач первичной профориентации. Полученные результаты демонстрируют, что мини-приложение «AI профориентатор» (<https://vk.com/app51929417>) для социальной сети «ВКонтакте» эффективно решает задачу предоставления первичных персонализированных рекомендаций по профориентации с использованием алгоритмов машинного обучения.

Результаты готовы!

Ваш тип — Конвенциональный



Люди, у которых выражен конвенциональный тип личности, стремятся к деятельности, предполагающей работу с информацией, ее сбор, хранение, систематизацию, являются превосходными исполнителями.

Для них ценны эффективность, методичность, деловые и экономические достижения. При этом такие люди консервативны, не любят ситуации непредсказуемости и неопределенности, предпочитают знать, что от них ожидают.

Способности

Внимательность к деталям, способность структурировать информацию, навыки работы с числами, канцелярские способности, ручная тонкая моторика, умение организовывать свою работу и расставлять приоритеты, умение следовать инструкциям.

Ценности и личностные качества

Организованность, пунктуальность, педантичность, аккуратность, сознательность, упорство, практичность, самоконтроль, консервативность

Предпочитаемые виды деятельности

Работа, которая требует внимания к деталям и аккуратности, ведение картотек, хранение и систематизация записей, фактов, данных, финансовых книг, написание деловых отчетов, подготовка схем, таблиц, диаграмм, работа в корпорациях, финансовых организациях, архивах, инспекциях, офисное администрирование.

Профессии

Секретарь, патентовед, нотариус, топограф, корректор, страховой агент, офис-менеджер, бухгалтер, аудитор.

Знаменитые представители



Скидвард
«Спанч Боб»



Мальвина
«Буратино»



Жираф Мелман
«Мадагаскар»



Морти
«Рик и Морти»



Ёжик
«Смешарики»



Валли
«ВАЛЛ-И»

Согласны ли Вы с результатом?

☐ Да

☐ Нет

Поделиться результатом

Рис. 3. Интерфейс мини-приложения «AI профорientатор»

Перспективы дальнейших исследований включают интеграцию дополнительных источников данных (например, анкетных данных со страницы пользователя), применение современных архитектур глубокого обучения (в частности, Transformers), расширение формата результатов за счёт отображения нескольких преобладающих типов и использование моделей LLM для интерпретации и объяснения рекомендаций. Важным направлением развития также является адаптация системы для образовательных учреждений, где она может служить инструментом сопровождения индивидуальных учебных траекторий и повышения эффективности профориентационной поддержки.

6. ЗАКЛЮЧЕНИЕ

Таким образом, в работе описана совокупность архитектурных решений и методов машинного обучения, помогающих решить проблему профориентации на основе анализа цифрового следа. Разработанная система объединяет алгоритмы multi-label на основе нейросетевой архитектуры с остаточными связями и клиент–серверное приложение VK Mini Apps. Такой подход обеспечивает персонализированную и масштабируемую диагностику профессиональных интересов, снижая зависимость от традиционных анкетных методов.

Полученные результаты подтверждают эффективность интеграции алгоритмов машинного обучения с данными социальных сетей и демонстрируют перспективность применения разработанной системы в практике профориентации. Работа вносит вклад в развитие интеллектуальных систем поддержки образовательных и кадровых решений и может служить основой для дальнейших исследований, направленных на расширение источников данных, повышение интерпретируемости предсказаний и адаптацию технологий к задачам учебного процесса.

Список литературы

1. *Basnet S.* Artificial Intelligence and machine learning in human resource management: Prospect and future trends // *International Journal of Research Publication and Reviews*. 2024. Vol. 5, № 1. P. 281–287. doi:10.55248/gengpi.5.0124.0107
2. *Song L.* Application of Association Rule Analysis in Vocational Education Student Career Path Planning // 2023 International Conference on Intelligent Computing, Communication & Convergence (ICI3C), Bhubaneswar, India, 2023. P. 92–98. doi:10.1109/ICI3C60830.2023.00028
3. *Al-Dhari S., Al-Alawi A.I.* The Application of Data Analytics to Career Choice Prediction: A Literature Review // 2023 International Conference on Cyber Management and Engineering (CyMaEn). IEEE, 2023. P. 260–265. doi:10.1109/CyMaEn57228.2023.10051101.
4. *Забокрицкая Л. Д., Орешкина Т. А., Обабков И. Н., Чепуров Е. Г.* Применение алгоритма машинного обучения для профориентации абитуриентов высшего учебного заведения // *Вестник Томского государственного университета*. 2022. № 485. С. 217–225. doi:10.17223/15617793/485/24
5. *Kamal A., Naushad B., Rafiq H., Tahzeeb S.* Smart career guidance system // 2021 4th International Conference on Computing & Information Sciences (ICCIS). IEEE, 2021. P. 1–7. doi:10.1109/ICCIS54243.2021.9676408
6. *Безруких А. Д., Черепанов М. Д., Мельников В. А., Мельникова Е. В.* Разработка программного проекта информационного сервиса профориентации абитуриентов СФУ // *Современные наукоемкие технологии*. 2023. № 4. С. 19–27. doi:10.17513/snt.39575
7. *Sorokin V., Tovbis E., Kazakovtsev L.* Browser Game as a New Way of Career Guidance // *European Proceedings of Computers and Technology*. European Publisher, 2022. Vol. 1. P. 235–240. doi:10.15405/epct.23021.28

8. Padma E., Sowdharshini P., Shanmugapriya P., Reshma K.M., Srimathi C.N. Career guidance system for students using machine learning // *Challenges in Information, Communication and Computing Technology*. CRC Press, 2025. P. 666–671. doi:10.1201/9781003559092-115
9. Panthee S., Rajkarnikar S., Begum R. Career Guidance System Using Machine Learning // *Journal of Advanced College of Engineering and Management*. 2023. Vol. 8, № 2. P. 113–119. doi:10.3126/jacem.v8i2.55947
10. Cui C. Career interest assessment: College students career planning based on machine learning // *Journal of Electrical Systems*. 2024. Vol. 20, № 6s. P. 1633–1644. doi:10.52783/jes.3083
11. Reddy K., Reddy M.A., Kaur V., Kaur G. Career guidance system using ensemble learning // *Proceedings of the Advancement in Electronics & Communication Engineering*. 2022. P. 33–39. doi:10.2139/ssrn.4157249
12. Wakde A., Maywade R., Pandey A., Kumar J., Singh A.K. An Ensemble Learning Based Career Prediction Model // *The Future of Artificial Intelligence and Robotics*. Cham: Springer Nature Switzerland, 2023. P. 503–512. doi:10.1007/978-3-031-60935-0_45
13. Диков М. Е., Широбокова С. Н. О варианте формализации задачи определения востребованности направлений подготовки и возможных сфер трудоустройства выпускников на основе семантического анализа описаний вакансий // *Инженерный вестник Дона*. 2022. № 5 (89). С. 214–222.
14. Wulandari R.S., Setianingsih C., Kusuma P.D. Analysis of Big Five Personality Factors to Determine the Appropriate Type of Career Using the C4.5 Algorithm // *Data Science and Emerging Technologies. DaSET 2022*. Singapore: Springer Nature Singapore, 2022. Vol. 165. P. 18–36. doi:10.1007/978-981-99-0741-0_2
15. José-García A., Sneyd A., Melro A., Ollagnier A., Tarling G., Zhang H., Stevenson M., Everson R., Arthur R. C3-IoC: A career guidance system for assessing student skills using machine learning and network visualisation // *International Journal of Artificial Intelligence in Education*. 2023. Vol. 33, № 4. P. 1092–1119. doi:10.1007/s40593-022-00317-y
16. Vignesh S., Priyanka C.S., Manju H.S., Mythili K. An intelligent career guidance system using machine learning // *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2021. Vol. 1. P. 987–990. doi:10.1109/ICACCS51430.2021.9441978
17. Goyal R., Chaudhary N., Singh M. Machine Learning based Intelligent Career Counselling Chatbot (ICCC) // *2023 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 2023. P. 1–8. doi:10.1109/ICCCI56745.2023.10128305
18. Kumbhar V. R., Maddel M. M., Raut Y. Smart model for career guidance using hybrid deep learning technique // *2023 1st International Conference on Innovations in High-Speed Communication and Signal Processing (IHCSP)*. IEEE, 2023. P. 327–331. doi:10.1109/IHCSP56702.2023.10127152
19. Yuan C., Hong Y., Wu J. Who Are You Meant to Be? Predicting Psychological Indicators and Occupations based on Personality Traits // *Journal of Systems Science and Systems Engineering*. 2023. Vol. 32, № 5. P. 571–602. doi:10.1007/s11518-023-5576-6
20. Grunenber E., Peters H., Francis M.J., Back M.D., Matz S.C. Machine learning in recruiting: predicting personality from CVs and short text responses // *Frontiers in Social Psychology*. 2024. Vol. 1. P. 1290295. doi:10.3389/frsps.2023.1290295
21. Nirmala M., Rajalakshmi B., Sarvepally S.R.M., Mandava S., Megana M. Personality Detection for Recruitment Using Machine Learning // *Proceedings of the 6th International Conference on Communications and Cyber Physical Engineering, ICCCE 2024*. Singapore: Springer, 2024. Vol. 1096. P. 399–406. doi:10.1007/978-981-99-7137-4_38
22. Iwasaki T., Seki Y., Kashino W., Keyaki A., Kando N. Estimating Citizen Personality Traits Using Social Media Posts // *International Conference on Asian Digital Libraries*. Springer, Singapore, 2024. P. 119–135. doi:10.1007/978-981-96-0868-3_10
23. Oliseen V. D., Khlobystova A. O., Korepanova A. A., Tulupyeva T. V. Automating the temperament assessment of online social network users // *Doklady Mathematics*. Moscow: Pleiades Publishing, 2023. Vol. 108, № S2. P. S368–S373, doi:10.1134/S1064562423701041
24. Хлобыстова А. О., Абрамов М. В., Столярова В. Ф. Исследование тенденций взаимосвязи между профориентационными предпочтениями пользователей и его цифровыми следами в соци-

альной сети // Научно-технический вестник информационных технологий, механики и опти-
ки. 2023. Т. 23, № 3. С. 564–574. doi:10.17586/2226-1494-2023-23-3-564-574

25. Kiselev P., Kiselev B., Matsuta V., Feshchenko A., Bogdanovskaya I., Kosheleva A. Career guidance based on machine learning: social networks in professional identity construction // Procedia Computer Science. 2020. Vol. 169. P. 158–163. doi:10.1016/j.procs.2020.02.128

Поступила в редакцию 20.06.2025, окончательный вариант — 15.07.2025.

Иващенко Анастасия Олеговна, научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), ✉ aok@dscs.pro

Вяткин Артём Андреевич, младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), aav@dscs.pro

Бушмелёв Фёдор Витальевич, научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), fvb@dscs.pro

Абрамов Максим Викторович, канд. техн. наук, ведущий научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), доцент, Санкт-Петербургский государственный университет, mva@dscs.pro

Computer tools in education, 2025

№ 2: 36–47

<http://cte.eltech.ru>

doi:10.32603/2071-2340-2025-2-36-47

Modelling a Career Guidance System Using Social Network Data Analysis

Ivashchenko A. O.¹, Researcher, ✉ aok@dscs.pro

Vyatkin A. A.¹, Junior Researcher, aav@dscs.pro

Bushmelev F. V.¹, Researcher, fvb@dscs.pro

Abramov M. V.^{1,2}, Cand. sc., Senior Researcher, Associate Professor, mva@dscs.pro

¹Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS),

14 Linia V. O., 39, 199178, Saint Petersburg, Russia

²Saint Petersburg State University, 7–9 Universitetskaya emb., 199034, Saint Petersburg, Russia

Abstract

The paper addresses the problem of automating career guidance through the analysis of digital footprints of VKontakte users. The aim of the study is to enhance the accessibility and accuracy of diagnosing professional interests by means of the "AI Proforientator" mini-application on the VK Mini Apps platform. The methodology is based on a multi-label classification approach: each of the six RIASEC professional types is represented by a separate binary output with sigmoid activation, and the model is trained using the Binary Cross-Entropy loss function. Training data included users who had completed the Holland test, along with their community subscriptions collected through the Psychological Tests app. The technical implementation follows a client-server architecture with REST API, covering data preprocessing, interaction with the ML server, and result delivery. The developed system provides a full processing cycle: extraction and validation of subscriptions, probabilistic prediction of professional personality types according to Holland's six categories,

and presentation of results to the user. The model achieved a Top-1 accuracy of 47.1% and a Top-2 accuracy of 72.3%, confirming its applicability to career guidance diagnostics. The system automates the initial career guidance process and can be employed both by individual users and organisations to support educational trajectories, optimise recruitment, and improve staff development. The novelty of the work lies in the comprehensive integration of neural network methods for multi-label classification with a client-server application embedded in a social network, ensuring a personalised and scalable approach to career guidance.

Keywords: *career guidance, digital footprint, machine learning, online social media, client-server application, predictive model, RIASEC.*

Citation: A. O. Ivashchenko, A. A. Vyatkin, F. V. Bushmelev, and M. V. Abramov, "Modelling a Career Guidance System Using Social Network Data Analysis," *Computer tools in education*, no. 2, pp. 36–47, 2025 (in Russian); doi:10.32603/2071-2340-2025-2-36-47

References

1. S. Basnet, "Artificial intelligence and machine learning in human resource management: Prospect and future trends," *International Journal of Research Publication and Reviews*, vol. 5, no. 1, pp. 281–287, 2024; doi:10.55248/gengpi.5.0124.0107
2. L. Song, "Application of Association Rule Analysis in Vocational Education Student Career Path Planning," in *2023 Int. Conf. on Intelligent Computing, Communication & Convergence (ICI3C)*, Bhubaneswar, India, 2023, pp. 92–98; doi:10.1109/ICI3C60830.2023.00028
3. S. Al-Dhari and A. I. Al-Alawi, "The Application of Data Analytics to Career Choice Prediction: A Literature Review," in *2023 Int. Conf. on Cyber Management and Engineering (CyMaEn)*, 2023, pp. 260–265; doi:10.1109/CyMaEn57228.2023.10051101
4. L. D. Zabokritskaya, T. A. Oreshkina, I. N. Obabkov, and E. G. Chepurov, "Application of a machine learning algorithm for career guidance of university applicants," *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*, no. 485, pp. 217–225, 2022; doi:10.17223/15617793/485/24
5. A. Kamal, B. Naushad, H. Rafiq, and S. Tahzeeb, "Smart career guidance system," in *2021 4th Int. Conf. on Computing & Information Sciences (ICCIS)*, 2021, pp. 1–7; doi:10.1109/ICCIS54243.2021.9676408
6. A. D. Bezrukikh, M. D. Cherepanov, V. A. Melnikov, and E. V. Melnikova, "Development of a software project for an information service for career guidance of applicants at Siberian Federal University," *Modern High Technologies*, no. 4, pp. 19–27, 2023; doi:10.17513/snt.39575
7. V. Sorokin, E. Tovbis, and L. Kazakovtsev, "Browser Game as a New Way of Career Guidance," in *Proc. of Int. Workshop "Hybrid methods of modeling and optimization in complex systems" (in the framework of The Eleventh Int. Conf. on Mathematical Models and their Applications)*, November 22–24, 2022, Krasnoyarsk, the Russian Federation, vol. 1, 2023, pp. 235–240; doi:10.15405/epct.23021.28
8. E. Padma, P. Soudharshini, P. Shanmugapriya, K. M. Reshmaa, and C. N. Srimathi, "Career guidance system for students using machine learning," in *Challenges in Information, Communication and Computing Technology*. CRC Press, 2025, pp. 666–671; doi:10.1201/9781003559092-115
9. S. Panthee, S. Rajkarnikar, and R. Begum, "Career Guidance System Using Machine Learning," *Journal of Advanced College of Engineering and Management*, vol. 8, no. 2, pp. 113–119, 2023; doi:10.3126/jacem.v8i2.55947
10. C. Cui, "Career interest assessment: College students career planning based on machine learning," *Journal of Electrical Systems*, vol. 20, no. 6s, pp. 1633–1644, 2024; doi:10.52783/jes.3083
11. K. Reddy, M. A. Reddy, V. Kaur, and G. Kaur, "Career guidance system using ensemble learning," in *Proceedings of the Advancement in Electronics & Communication Engineering*, 2022, pp. 33–39; doi:10.2139/ssrn.4157249
12. A. Wakde, R. Maywade, A. Pandey, J. Kumar, and A. K. Singh, "An Ensemble Learning Based Career Prediction Model," in *The Future of Artificial Intelligence and Robotics*. Cham: Springer Nature Switzerland, 2023, pp. 503–512; doi:10.1007/978-3-031-60935-0_45
13. M. E. Dikhtov and S. N. Shirobokova, "On a variant of formalizing the task of determining the demand for training areas and possible spheres of employment of graduates based on the semantic analysis of vacancy descriptions," *Engineering Journal of Don*, no. 5, pp. 214–222, 2022.
14. R. S. Wulandari, C. Setianingsih, and P. D. Kusuma, "Analysis of Big Five Personality Factors to Determine the Appropriate Type of Career Using the C4.5 Algorithm," in *Data Science and Emerging Technologies (DaSET 2022)*, vol. 165, 2022, pp. 18–36; doi:10.1007/978-981-99-0741-0_2

15. A. Jose-Garcia et al., “C3-IoC: A career guidance system for assessing student skills using machine learning and network visualisation,” *International Journal of Artificial Intelligence in Education*, vol. 33, no. 4, pp. 1092–1119, 2023; doi:10.1007/s40593-022-00317-y
16. S. Vignesh, C. S. Priyanka, H. S. Manju, and K. Mythili, “An intelligent career guidance system using machine learning,” in *2021 7th Int. Conf. on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, 2021, pp. 987–990; doi:10.1109/ICACCS51430.2021.9441978
17. R. Goyal, N. Chaudhary, and M. Singh, “Machine Learning based Intelligent Career Counselling Chatbot (ICCC),” in *2023 Int. Conf. on Computer Communication and Informatics (ICCCI)*, 2023, pp. 1–8; doi:10.1109/ICCCI56745.2023.10128305
18. V. R. Kumbhar, M. M. Maddel, and Y. Raut, “Smart model for career guidance using hybrid deep learning technique,” in *2023 1st Int. Conf. on Innovations in High-Speed Communication and Signal Processing (IHCSP)*, 2023, pp. 327–331; doi:10.1109/IHCSP56702.2023.10127152
19. C. Yuan, Y. Hong, and J. Wu, “Who Are You Meant to Be? Predicting Psychological Indicators and Occupations based on Personality Traits,” *Journal of Systems Science and Systems Engineering*, vol. 32, no. 5, pp. 571–602, 2023; doi:10.1007/s11518-023-5576-6
20. E. Grunenbergh, H. Peters, M. J. Francis, M. D. Back, and S. C. Matz, “Machine learning in recruiting: predicting personality from CVs and short text responses,” *Frontiers in Social Psychology*, vol. 1, p. 1290295, 2024; doi:10.3389/frsps.2023.1290295
21. M. Nirmala et al., “Personality Detection for Recruitment Using Machine Learning,” in *Proc. of the 6th Int. Conf. on Communications and Cyber Physical Engineering, ICCCE 2024*, Singapore: Springer, 2024, pp. 399–406; doi:10.1007/978-981-99-7137-4_38
22. T. Iwasaki, Y. Seki, W. Kashino, A. Keyaki, and N. Kando, “Estimating Citizen Personality Traits Using Social Media Posts,” in *Int. Conf. on Asian Digital Libraries*. Singapore: Springer, 2024, pp. 119–135; doi:10.1007/978-981-96-0868-3_10
23. V. D. Oliseenko, A. O. Khlobystova, A. A. Korepanova, and T. V. Tulupyeva, “Automating the temperament assessment of online social network users,” *Doklady Mathematics*, vol. 108, no. S2, pp. S368–S373, 2023; doi:10.1134/S1064562423701041
24. A. O. Khlobystova, M. V. Abramov, and V. F. Stolyarova, “Research of trends in the relationship between users’ career guidance preferences and their digital footprints in a social network,” *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 23, no. 3, pp. 564–574, 2023; doi:10.17586/2226-1494-2023-23-3-564-574
25. P. Kiselev, B. Kiselev, V. Matsuta, A. Feshchenko, I. Bogdanovskaya, and A. Kosheleva, “Career guidance based on machine learning: social networks in professional identity construction,” *Procedia Computer Science*, vol. 169, pp. 158–163, 2020; doi:10.1016/j.procs.2020.02.128

Received 20-06-2025, the final version — 15-07-2025.

Anastasiia Ivashchenko, Researcher, Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), ✉ aok@dscs.pro

Arteom Vyatkin, Junior Researcher, Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), aav@dscs.pro

Fedor Bushmelev, Researcher, Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), fvb@dscs.pro

Maxim Abramov, Candidate of Sciences (Tech.), Senior Researcher, Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Associate Professor, Saint Petersburg State University, mva@dscs.pro