

Автоматизированный анализ вакансий специалистов психологического и педагогического профиля

Низомутдинов Б.А. (ИТМО), Углова А.Б. (РГПУ им. А.И. Герцена)

Введение. В рамках представленного исследования была создана уникальная база данных вакансий для специалистов в области психологии и педагогики, собранных с ресурса по трудоустройству за последние три года, разработана методика автоматизированного анализа для выявления основных требований к специалистам отдельных направлений. Целью сбора и последующего анализа данных было выявление текущего состояния рынка труда для данных профессий, понимание трендов и изменений в требованиях к квалификации, опыту и профессиональным навыкам специалистов. С использованием алгоритмов тематического моделирования LDA был проведен анализ описаний вакансий, что позволило классифицировать данные предложения работы на группы (кластеры) с учетом схожих характеристик и требований. Этот подход дал возможность выделить ключевые навыки и компетенции, которые являются наиболее востребованными у работодателей, а также проанализировать изменение этих требований. Изучение профиля профессии помогает понять, какие навыки и знания необходимы для успешной работы в данной области, что способствует повышению квалификации, профессиональному развитию и росту в карьере.

Основная часть. Современный рынок труда характеризуется бурными и постоянными изменениями, обусловленными такими факторами, как глобализация, технологический прогресс, социально-экономические сдвиги и изменение культурных нормативов. Особенно это касается области помогающих профессий, включающих психологов, социальных работников, медицинских и образовательных специалистов, где требования к квалификации постоянно эволюционируют. Одним из ключевых аспектов адаптации профессионалов к современным вызовам является способность быстро реагировать на текущие тренды и предвосхищать потребности рынка труда.

В ответ на эту потребность, мы разработали автоматизированный метод анализа рынка труда, который позволяет углубленно исследовать и анализировать ключевые требования к специалистам в помогающих профессиях. Этот подход обеспечивает динамичное слежение за профессиональными трендами и формирование прогностической модели развития профессиональных компетенций, что призвано способствовать повышению эффективности трудовой деятельности и карьерного роста специалистов.

Для анализа вакансий мы использовали массив открытых данных проекта "Инфраструктура научно-исследовательских данных" [1] База содержит архивные и актуальные сведения о резюме, вакансиях, откликах соискателей и приглашениях на собеседование, размещенные на портале Роструда «Работа России».

На первом шаге, с помощью API были выгружены все вакансии по 2 направлениям - педагогического и психологического профиля. На втором шаге описание вакансий было обработано с помощью методов текстовой аналитики - удалены лишние теги, спецсимволы. На третьем шаге, для каждого набора были выделены важные термины, с помощью метода TF-IDF (Term Frequency и Inverse Document Frequency), далее, с помощью метода LDA было проведено тематическое моделирование, выделено 10 основных тем, описанных ключевыми словами для каждого направления.

Метод LDA (латентное размещение Дирихле) эффективно выявляет ключевые темы в больших текстовых корпусах, используя модели вероятностного тематического моделирования [2-3]. Он автоматически обнаруживает и группирует слова в документах в наборы тем, где каждое слово относится к теме с определенной вероятностью.

Применительно к анализу вакансий, метод LDA помог выделить наиболее часто

встречающихся темы, термины и требований, указанных в описаниях рабочих мест. Это помогает понять ключевые навыки и квалификации, ценные для работодателей, и тренды рынка труда, что важно для курсов обучения и профессиональной подготовки специалистов, а также для соискателей, стремящихся улучшить свою конкурентоспособность на рынке труда.

Для агрегации и анализа была задействована ВІ система Yandex.Data Leans. На финальном шаге была проведена визуализация данных, отражающая распределение вакансий по регионам. Данная визуализация позволила наглядно сравнить регионы и выявить районы с наибольшим числом вакансий и спросом на специалистов.

Выводы. Проведен парсинг вакансий, настроен автоматизированный комплекс текстовой аналитики для анализа описаний вакансий, настроена ВІ система Yandex.Data Leans для визуализации результатов.

Предложенный метод был апробирован на массиве вакансий за 2020-2023 год, была собрана база, содержащая более 7500 вакансий педагогов и психологов. Проведенный анализ позволил выявить основные требования, которые предъявляет работодатель к специалистам, а также, определить основные маркеры, описывающие трудовую деятельность помогающих специалистов.

На следующем этапе запланирована выгрузка за 10 лет, для проведения ретроспективного анализа, для выявления основных трендов, а также, анализ резюме специалистов, для сопоставления с полученными ранее результатами.

Полученные результаты могут быть использованы как специалистами, которые находятся в поисках работы, так и образовательных организаций, для корректировки образовательных программ, чтобы отвечать вызовам современной эпохи. Изучение профиля профессии помогает понять, какие навыки и знания необходимы для успешной работы в данной области, что способствует повышению квалификации, профессиональному развитию и росту в карьере.

Статья подготовлена при финансовой поддержке Российского научного фонда (проект № 22-78-10047, <https://rscf.ru/project/22-78-10047/> "Конструктивные и деструктивные коммуникативные практики специалистов помогающих профессий в цифровых медиа". URL: <https://rscf.ru/project/22-78-10047/>)

Список использованных источников:

1. Бабушкина В.О., Тимошенко А.Ш., «Работа в России»: обработанные и объединенные сведения о вакансиях, резюме, откликах и приглашениях портала trudvsem.ru // Роструд; обработка: Инфраструктура научно-исследовательских данных, АНО «ЦПУР», 2021. Доступ: Лицензия CC BY-SA. Размещено: 02.12.2021. (Ссылка на набор данных: <http://data.rcsi.science/data-catalog/datasets/186/>)

2. Yu a, Anran Fang a, Zeshui Xu b. Topic research in fuzzy domain: Based on LDA topic modelling. Information Sciences, Vol. 648, 2023, DOI 10.1016/j.ins.2023.119600

3. Akhtar, N., Sufyan Beg, M.M., Javed, H. (2019). Topic Modelling with Fuzzy Document Representation. In: Singh, M., Gupta, P., Tyagi, V., Flusser, J., Ören, T., Kashyap, R. (eds) Advances in Computing and Data Sciences. ICACDS 2019. Communications in Computer and Information Science, vol 1046. Springer, Singapore. https://doi.org/10.1007/978-981-13-9942-8_54