



<http://dx.doi.org/10.35596/1729-7648-2024-30-4-23-32>

Оригинальная статья
Original paper

УДК 331.1+004.89

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ ПРИ АНАЛИЗЕ РЫНКА ТРУДА: ТЕОРЕТИЧЕСКИЕ ПОДХОДЫ И МЕТОДИЧЕСКИЕ ИНСТРУМЕНТЫ

Е. В. ВАНКЕВИЧ, И. Н. КАЛИНОВСКАЯ

Витебский государственный технологический университет (г. Витебск, Республика Беларусь)

Поступила в редакцию 26.07.2024

© Белорусский государственный университет информатики и радиоэлектроники, 2024
Belarusian State University of Informatics and Radioelectronics, 2024

Аннотация. Рассмотрена возможность использования технологий больших данных и больших языковых моделей для анализа рынка труда в Республике Беларусь. Разработаны теоретические подходы с применением больших данных, что подразумевает определение как возможности в проведении аналитики рынка труда по данным онлайн-источников, так и эффективных инструментов для сбора и обработки информации о рынке труда с онлайн-источников. Применение больших данных и больших языковых моделей позволит улучшить качество и точность анализа рынка труда в республике, а использование передовых аналитических инструментов обеспечит более полное и детализированное понимание динамики рынка труда. Исследование основывается на анализе существующих теоретических подходов, практике использования больших данных и больших языковых моделей в зарубежных странах, а также на оценке текущих возможностей и ограничений применения этих технологий в Беларуси. В качестве инструментария использовались методы машинного обучения, анализа больших данных и моделирования. Результаты исследования могут быть применены для улучшения стратегий управления рынком труда, а также для разработки политик и программ занятости, ориентированных на современные вызовы и возможности цифровой экономики.

Ключевые слова: большие данные (Big Data), рынок труда, анализ данных, занятость, большие языковые модели (LLM), машинное обучение, искусственный интеллект, вакансии, резюме, цифровая экономика, инструменты аналитики.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Благодарность. Исследование выполнено в рамках научно-исследовательской работы «Теоретические подходы и методическое обеспечение анализа рынка труда в Республике Беларусь с применением больших данных» при финансовой поддержке Белорусского фонда фундаментальных исследований по договору Г24-013.

Для цитирования. Ванкевич, Е. В. Использование больших данных при анализе рынка труда: теоретические подходы и методические инструменты / Е. В. Ванкевич, И. Н. Калиновская // Цифровая трансформация. 2024. Т. 30, № 4. С. 23–32. <http://dx.doi.org/10.35596/1729-7648-2024-30-4-23-32>.

USING BIG DATA IN LABOR MARKET ANALYSIS: THEORETICAL APPROACHES AND METHODOLOGICAL TOOLS

ALENA V. VANKEVICH, IRYNA N. KALINOUSKAYA

Vitebsk State Technological University (Vitebsk, Republic of Belarus)

Submitted 26.07.2024

Abstract. The article considers the possibility of using big data technologies and large language models to analyze the labor market in the Republic of Belarus. Theoretical approaches using big data have been developed, which implies determining both the possibility of conducting labor market analytics using online sources and effective tools for collecting and processing information on the labor market from online sources. The use of big data and large language models will improve the quality and accuracy of labor market analysis in the republic, and the use of advanced analytical tools will provide a more complete and detailed understanding of the labor market dynamics. The study is based on the analysis of existing theoretical approaches, the practice of using big data and large language models in foreign countries, as well as an assessment of the current capabilities and limitations of using these technologies in Belarus. Machine learning, big data analysis and modeling were used as tools. The results of the study can be used to improve labor market management strategies, as well as to develop employment policies and programs focused on modern challenges and opportunities of the digital economy.

Keywords: Big Data, labor market, data analysis, employment, big language models (LLM), machine learning, artificial intelligence, vacancies, resumes, digital economy, analytics tools.

Conflict of interests. The authors declare no conflict of interests.

Gratitude. The research was carried out within the framework of the research work “Theoretical approaches and methodological support for the analysis of the labor market in the Republic of Belarus using big data” financed by the Belarusian Republican Foundation for Fundamental Research under contract Г24-013.

For citation. Vankevich A. V., Kalinouskaya I. N. (2024) Using Big Data in Labor Market Analysis: Theoretical Approaches and Methodological Tools. *Digital Transformation*. 30 (4), 23–32. <http://dx.doi.org/10.35596/1729-7648-2024-30-4-23-32> (in Russian).

Введение

Актуальность темы исследования обусловлена возрастающими объемами информации о рынке труда, что создает новые возможности для аналитики рынка труда и принятия решений. В современных условиях возрастает количество источников информации, содержащих сведения о вакансиях и соискателях рабочих мест (онлайн-порталы вакансий, сайты предприятий, профессиональные сети, проч.), в которых много дополнительных данных о рынке труда в отличие от статистических и административных источников. Но при работе с большими данными (Big Data) при анализе рынка труда возникают методологические трудности, связанные с неоднородностью самих источников данных, разной степенью их надежности и достоверности, разной структурированностью расположенной на них информации, с недоступностью части информации, что создает проблемы с репрезентативностью. В этих условиях необходимой является разработка технического инструментария для формирования единой базы данных о рынке труда для ее последующего анализа. Использование «сырых» или просто агрегированных данных, даже отслеживаемых за определенный промежуток времени, – это только информация. Аналитикой она становится тогда, когда создан алгоритм (теоретический фундамент и методика), который эти данные может обработать и дать результаты для принятия решений и прогноза.

Элементы больших данных о рынке труда и их источники

Большие данные – это структурированные и неструктурированные данные значительных объемов и многообразия, обрабатываемые с помощью программных инструментов. Основными характеристиками Big Data являются объем, скорость, разнообразие, достоверность, ценность [1]. Big Data для рынка труда позволяют извлекать новые сведения, поэтому для их использования нужно выполнить шесть этапов: формирование интегрированного массива данных из первоисточников, обработка, классификация, хранение массива данных, анализ, визуализация результатов. На каждом из этапов значима роль эксперта в области рынка труда, поскольку необ-

ходимо генерировать базу данных из релевантных источников, обрабатывать и классифицировать данные по принятым в экономике рынка труда показателям, уметь видеть и оценивать выявленные тенденции и взаимосвязи.

Элементами Big Data о рынке труда являются: информация о вакансиях (число вакансий, наименование организации, название должности, требования к навыкам и образованию, предлагаемая заработная плата); информация о резюме (число резюме, название должности, имеющиеся навыки, образование, требования к условиям труда и заработной плате); профили соискателей (опыт работы, навыки, образование, пол, возраст); метаданные вакансий и резюме (местоположение, отрасль, время открытия вакансии, дата публикации резюме); численность занятых, безработных и выпускников; классификация перечисленных массивов данных по видам экономической деятельности, регионам, занятиям (для поиска зависимостей).

Можно выделить следующие типы источников информации о рынке труда:

– административные источники данных (статистическая и административная отчетность органов государственного, отраслевого, регионального управления о численности и структуре занятых, принятых, уволенных, выпускниках, обучающихся, безработных, количестве свободных рабочих мест и вакансий; данные опросов, др.);

– онлайн-источники (онлайн-порталы вакансий, агрегаторы вакансий, сайты предприятий и кадровых агентств, профессиональные сообщества и сети в интернете).

Каждый из перечисленных типов источников Big Data о рынке труда можно классифицировать по доступности и полноте содержащейся информации, репрезентативности, периоду обновления, используемым классификаторам. Качество больших данных в первую очередь зависит от выбранного источника и способа их обработки.

Зарубежный опыт использования больших данных в аналитике рынка труда

Мировым лидером в аналитике рынка труда, основанной на использовании сформированных из интернет-источников баз данных, является Burning Glass Institute. Данные о вакансиях и резюме собираются методом веб-скрейпинга с более чем 40 тыс. уникальных онлайн-ресурсов, группируются в зависимости от цели исследования с выполнением научно-исследовательских работ (платно). В экономической литературе широко демонстрируются возможности использования больших данных для аналитики различных аспектов рынка труда, обработанных с помощью технологий искусственного интеллекта, по следующим основным направлениям: для классификации навыков и занятий с помощью искусственного интеллекта [2–6], при проведении анализа различных аспектов рынка труда (затрат на персонал [7], структуры занятости [8], динамики и дифференциации спроса на отдельных специалистов [9]), для выявления несоответствий на рынке труда и определения востребованных рынком труда направлений и специальностей подготовки в учреждениях образования [10–13].

В настоящее время на уровне Европейской экономической комиссии обсуждается вопрос консолидации статистических данных и данных, полученных из онлайн-источников¹ [14], что приведет к формированию «статистики граждан». С 2016 г. в ЕС совместно с Евростатом реализуется пилотный проект ESSnet Big Data для интеграции Big Data о рынке труда с официальной статистикой на основе изучения потенциала отобранных онлайн-источников (проект охватывает 22 страны ЕС). Европейским центром развития профессионального образования и обучения (CEDEFOP) создана мультязычная система (32 языка), способная собирать вакансии, извлекать из них навыки и осуществлять мониторинг рынка труда в 28 странах ЕС. Усилиями Европейского фонда образования реализован проект «Большие данные для рынка труда»², в результате которого созданы базы данных и онлайн-платформы их визуализации [1].

В России АНО «Центр научных исследований в сфере профориентации и психологии труда» в 2021–2022 гг. создана цифровая модель рынка труда «Магуча» [15, 16], которая консолидирует информацию о рынке труда с различных онлайн-источников и сравнивает с официальными ста-

¹ Servoz M. (2019) *AI. The Future of work? Work of the Future!* Available: <https://ec.europa.eu/digital-single-market/en/news/future-work-work-future> (Accessed 10 July 2024).

² *Big Data for Labour Market Information – Focus on Data from Online Job Vacancies – Training Workshop (Eastern Partnership Platform 4 Work Programme for Members of the “Make it Match” Experts’ Network)*. Available: <https://www.etf.europa.eu/en/news-and-events/events/big-data-labour-market-information-focus-data-online-job-vacancies-training> (Accessed 10 July 2024).

статистическими данными (без детализации по навыкам)³. Специалистами ФГБУ «ВНИИ труда» Министерства труда Российской Федерации разработана автоматизированная информационная система «Мониторинг рынка труда» – программно-аппаратный комплекс по мониторингу вакансий и резюме, где источниками данных выступили открытые данные HeadHunter, SuperJob, статистические данные Министерства просвещения Российской Федерации⁴, которые доступны для анализа [17].

В целом разработанные страновые модели и их программное обеспечение не в полной мере применимы для Беларуси, так как не рассчитаны на отечественные классификаторы, а используемые ими программное обеспечение и площадка для визуализации не доступны белорусским пользователям. Также остается открытым вопрос экспертной оценки формируемой базы и методологии ее обработки. То есть технически вопрос решен, а с теоретической точки зрения – нет. В республике комплексных исследований рынка труда с применением технологии Big Data не проводится, за исключением некоторых работ в отдельных видах экономической деятельности [18, 19] или для отдельных категорий пользователей (социально уязвимых групп населения) [20]. На некоторых порталах (rabota.by, belmeta.com) есть аналитические разработки, однако в них проводится не анализ, а освещение отдельных аспектов рынка труда, и только по собственным данным (что объясняет нерепрезентативность этих обзоров). Углубленные результаты таких обзоров, как правило, не публикуются, поскольку они выполняются по заказу платно. Но главными ограничениями обзоров является отсутствие аналитики, использование собственных показателей (научная правомерность которых не всегда очевидна) и нерепрезентативность данных. Отсутствие теоретических разработок по оценке инструментов для использования Big Data на рынке труда снижает потребительскую ценность полученной информации. Поэтому нужна экспертная работа для создания инструментов анализа рынка труда с применением Big Data.

Отбор инструментов для анализа рынка труда Республики Беларусь на основе больших данных

Технологии Big Data и большие языковые модели (Large Language Models, LLM) находят все более широкое применение в анализе рынка труда. LLM, такие как GPT-4, Claude и другие, способны обрабатывать и анализировать огромные объемы неструктурированных текстовых данных, открывая новые возможности для исследования тенденций занятости, востребованных навыков и других аспектов рынка труда. Одним из ключевых преимуществ использования LLM является их способность автоматически извлекать ценную информацию из больших массивов данных, таких как описания вакансий, резюме, отзывы сотрудников и дискуссии на профессиональных форумах. Эти модели могут идентифицировать ключевые слова, фразы и концепции, связанные с конкретными профессиями, отраслями и навыками. Это позволяет получать ценные сведения о текущем состоянии и будущих трендах в сфере занятости.

Исследованию возможностей LLM при анализе рынка труда посвящены работы ряда авторов. М. Тамбурри исследовал, как LinkedIn использует LLM BERT (Bidirectional Encoder Representations from Transformers, разработана компанией Google в 2018 г., используется в приложениях, связанных с обработкой естественного языка, интегрирована в поисковую систему Google для улучшения понимания запросов пользователей) для анализа данных о вакансиях и выявления профессий с высоким спросом. LLM применялась для понимания семантики описаний вакансий и улучшения классификации профессий. Авторы [21] использовали GPT-2 для анализа спроса на рабочую силу в США. Они применяли LLM для выявления вакансий, связанных с искусственным интеллектом, и изучали, как спрос на цифровые навыки влияет на различные профессии и демографические группы. Соавторы [22] задействовали BERT для анализа объявлений о вакансиях и резюме с целью изучения связи между автоматизацией и требованиями к навыкам на итальянском рынке труда. LLM извлекала информацию о навыках из неструктурированных текстов для классификации веб-вакансий по профессиям и навыкам. Авторы [23] реализовывали BERT для анализа влияния «Индустрии 4.0» на профили работы и требования к навыкам. LLM собирала информацию о навыках и задачах из описаний вакансий, связанных с «Индустри-

³ Смирнов, А. Ю. Цифровая модель рынка труда: ключевые аспекты работы программного комплекса / А. Ю. Смирнов // Экономика труда. 2023. Т. 10, № 10. С. 1535–1552. DOI: 10.18334/et.10.10.119514.

⁴ Мониторинг рынка труда в сфере общего и среднего профессионального образования (жизненный цикл квалификации). Режим доступа: https://spkobr.ru/upload/docs/Мониторинг_рынка_труда_2023_СПК_в_сфере_образования.pdf. Дата доступа 15.06.2024.

ей 4.0». Эти исследования демонстрируют растущее применение LLM, таких как GPT и BERT, для анализа данных о рынке труда. LLM используются для различных задач, включая классификацию профессий, извлечение информации о навыках, анализ влияния технологических изменений на спрос на рабочую силу и улучшение систем анализа рынка труда.

Однако применение LLM в анализе рынка труда сопряжено и с определенными проблемами. Одна из них – потенциальное смещение, присущее моделям, обученным на исторических данных. Важно обеспечить, чтобы выводы, полученные с помощью LLM, не усугубляли существующие неравенство и дискриминацию на рынке труда. Кроме того, конфиденциальность данных и этические аспекты использования персональной информации в масштабных анализах требуют тщательного рассмотрения.

В рамках исследования проведен сравнительный анализ LLM с другими методами обработки текста в контексте анализа рынка труда, рассмотрены следующие альтернативные методы: методы на основе правил, традиционные методы машинного обучения (ML), классические методы глубинного обучения (DL). Проведено сравнение их по различным критериям (точность, гибкость, вычислительные затраты, простота интерпретации, необходимость в данных, направления использования). Большие языковые модели предоставляют значительные преимущества по сравнению с другими методами обработки текста по вопросам анализа рынка труда благодаря своей высокой точности, гибкости и способности обрабатывать большие объемы неструктурированных данных. Однако их использование связано с очень высокими вычислительными затратами и со сложностью интерпретации результатов. В зависимости от конкретных задач и доступных ресурсов выбор подходящего метода анализа информации о рынке труда может варьироваться от простых правил до сложных моделей глубинного обучения.

В настоящее время существует ряд компаний и платформ, которые используют возможности LLM: LinkedIn, Indeed, Burning Glass Technologies, Textkernel, Gloat. Эти платформы демонстрируют широкий спектр применений LLM в анализе данных о рынке труда – от персонализированного подбора вакансий до прогнозирования будущих потребностей в навыках. С учетом огромной пользовательской базы и растущего внедрения этих технологий они имеют значительный потенциал трансформировать практики рекрутинга и управления талантами во многих отраслях. Каждая платформа имеет свои сильные стороны и ограничения в использовании LLM для анализа данных о рынке труда. LinkedIn и Indeed обладают масштабом и разнообразием данных, Burning Glass Technologies и Textkernel предлагают глубину и специализацию анализа, а Gloat демонстрирует эффективность в оптимизации внутренних рынков труда. Выбор подходящего решения зависит от конкретных потребностей и задач организации в области анализа рынка труда и управления талантами. Дальнейшее развитие и интеграция этих платформ, а также повышение прозрачности в отношении их возможностей и ограничений будут способствовать более эффективному применению LLM в сфере HR-аналитики и принятия управленческих решений о рабочей силе.

В данном исследовании проводился сравнительный анализ как современных, так и более ранних моделей LLM в контексте анализа данных о рынке труда. Рассматривались такие модели, как GPT-4 (OpenAI), BERT (Google), T5 (Google), RoBERTa (Facebook AI) и Claude (Anthropic). Анализ возможностей и ограничений LLM при изучении рынка труда показал, что:

- современные модели, такие как GPT-4 и Claude, демонстрируют высокую гибкость и точность в обработке текста, что позволяет им эффективно анализировать и генерировать текстовые данные, включая анализ вакансий и отзывов. Однако их использование связано с высокими вычислительными затратами и со сложностью интерпретации результатов;

- более ранние модели, BERT, T5 и RoBERTa, также обеспечивают высокую точность, особенно в задачах классификации и анализа текста. Они требуют меньше вычислительных ресурсов по сравнению с современными моделями, но могут быть менее гибкими в задачах генерации текста и обработки длинных контекстов;

- современные модели обладают широкими возможностями для автоматизации и анализа сложных текстовых данных, что делает их полезными для задач, связанных с анализом рынка труда и генерацией аналитических отчетов. Более ранние модели продолжают быть полезными для задач классификации и анализа текста, особенно когда требуется баланс между точностью и вычислительными затратами.

Анализ показал, что как современные, так и более ранние LLM имеют свои сильные и слабые стороны. Выбор модели для конкретного применения должен учитывать требования к точности, гибкости к вычислительным ресурсам.

В задачах машинного обучения и обработки естественного языка для оценки производительности моделей используются показатели точности, полноты и F1-меры. Эти метрики помогают понять, насколько хорошо модель выполняет свои задачи, такие как классификация, извлечение информации или анализ настроений. Эти показатели помогают оценить производительность моделей LLM в контексте задач анализа данных о рынке труда, обеспечивая всестороннее понимание их эффективности и надежности. Анализ показателей LLM при решении задач классификации вакансий и навыков, настроений в отзывах сотрудников и работодателей, точности извлечения ключевой информации из отраслевых отчетов представлены в таблицах 1–3 (составлено по данным⁵).

Таблица 1. Качество классификации вакансий и навыков

Table 1. The quality of job classification and skills

Модель	Точность, %	Полнота, %	F1-мера, %
GPT-4	91,2	89,7	90,4
BERT	78,5	78,9	78,7
T5	79,9	80,3	80,1
RoBERTa	80,1	80,5	80,3
Claude	91,0	89,4	90,2

Таблица 2. Эффективность анализа настроений в отзывах сотрудников и работодателей

Table 2. The effectiveness of sentiment analysis in employee and employer reviews

Модель	Точность, %	Полнота, %	F1-мера, %
GPT-4	92,3	90,7	91,5
BERT	79,8	80,1	79,9
T5	80,4	80,7	80,5
RoBERTa	81,1	78,5	79,8
Claude	91,8	90,2	91,0

Таблица 3. Точность извлечения ключевой информации из отраслевых отчетов

Table 3. Accuracy of extracting key information from industry reports

Модель	Точность, %	Полнота, %	F1-мера, %
GPT-4	93,0	91,4	92,2
BERT	80,2	78,5	79,3
T5	81,5	79,8	80,6
RoBERTa	82,1	80,5	81,3
Claude	92,8	91,2	92,0

Исследование показателей LLM при решении задач анализа данных рынка труда показало, что использование GPT-4 для анализа базы данных с вакансиями может дать более точные результаты благодаря его способности генерировать тексты и учитывать широкий контекст. Например, GPT-4 может анализировать вакансии, выделяя ключевые навыки и требования, что позволяет работодателям и соискателям лучше понимать текущие тенденции. T5 и GPT-4 должны использоваться для автоматической суммаризации больших отчетов о рынке труда. Это позволяет сократить время, затрачиваемое на анализ и составление отчетов, предоставляя пользователям краткие и информативные резюме. Модели RoBERTa и Claude могут эффективно анализировать отзывы сотрудников и работодателей, помогая компаниям понять настроения своих сотрудников и выявить области для улучшения. BERT и RoBERTa показали высокую точность в извлечении ключевой информации из отраслевых отчетов, что может быть полезно для аналитиков, занимающихся исследованием конкретных отраслей, так как модели позволяют быстро извлекать необходимые данные для анализа.

⁵ *Artificial Analysis LLM Performance Leaderboard*. Available: <https://huggingface.co/spaces/ArtificialAnalysis/LLM-Performance-Leaderboard> (Accessed 10 July 2024).

Результаты исследования

Сравнительный анализ LLM с традиционными методами обработки текста (методы на основе правил, ML, DL) демонстрирует их превосходство в точности, гибкости и способности обрабатывать большие объемы неструктурированных данных. Однако это преимущество сопряжено с повышенными вычислительными затратами и со сложностью интерпретации результатов, что указывает на необходимость дальнейших исследований в области оптимизации и объяснимости моделей.

Количественные показатели эффективности различных моделей (GPT-4, BERT, T5, RoBERTa, Claude) в задачах, связанных с анализом рынка труда, демонстрируют высокий уровень точности и надежности. Это подтверждает потенциал LLM как инструмента для повышения качества принятия решений в сфере управления человеческими ресурсами и формирования политики на рынке труда. Выявленные различия в производительности моделей для разных задач (классификация вакансий, анализ настроений, извлечение информации) подчеркивают важность выбора соответствующей модели для конкретных исследовательских или практических целей. Это указывает на необходимость разработки специализированных подходов к применению LLM в зависимости от специфики анализируемых данных и целей исследования.

Интеграция LLM с экспертными знаниями специалистов по управлению человеческими ресурсами и аналитиков рынка труда открывает перспективы для создания гибридных систем поддержки принятия решений. Такой подход может объединить преимущества автоматизированного анализа Big Data с глубоким пониманием контекста и нюансов рынка труда, доступным человеческим экспертам. В Республике Беларусь LLM модели рекомендуется использовать для решения широкого спектра задач в области рынка труда, управления человеческими ресурсами и обработки больших объемов текстовой информации по следующим направлениям:

- для автоматизированного анализа и сопоставления данных вакансий и резюме – модель GPT-4 определяет ключевые навыки, требуемые в вакансиях, и сопоставляет с данными резюме, что способствует подбору наиболее подходящего кандидата на вакансию;

- для анализа настроений персонала и обратной связи – модели RoBERTa и Claude позволяют оценить настроения сотрудников, их удовлетворенность работой, определить области для улучшения корпоративной культуры;

- для прогнозирования потребностей рынка труда – с помощью моделей T5 и GPT-4 возможно провести анализ тенденций на рынке труда, составить прогноз востребованных профессий и навыков, что поможет учреждениям образования и работодателям лучше адаптироваться к будущим потребностям;

- для обработки и извлечения данных из документов – автоматизация процесса обработки юридических документов, контрактов и других текстовых материалов благодаря моделям BERT и RoBERTa обеспечивает быстрое извлечение нужной информации;

- для создания систем поддержки принятия решений – модель GPT-4 обеспечивает формирование различных вариантов решений, анализируя большие массивы данных.

Однако для эффективного и ответственного применения этих технологий необходимы:

- адаптация моделей под специфику белорусского рынка труда – обучение и донастройка моделей на локальных данных для повышения их релевантности и точности;

- соблюдение правовых ограничений – учет национального законодательства в области защиты персональных данных, конфиденциальности и этики использования искусственного интеллекта;

- учет этических аспектов – обеспечение прозрачности алгоритмов, гарантированность того, что применение искусственного интеллекта не нарушает права и свободы граждан;

- инвестиции в инфраструктуру и обучение персонала – предоставление необходимых ресурсов для внедрения технологий и подготовки специалистов, способных эффективно работать с большими языковыми моделями.

Таким образом, интеграция LLM в анализ рынка труда представляет собой многообещающее направление исследований и практических приложений, способное существенно повысить эффективность принятия решений в сфере управления человеческими ресурсами и формирования политики на рынке труда. Однако для полной реализации потенциала этих технологий необходимы дальнейшие исследования, направленные на решение этических проблем, повышение интерпретируемости моделей и их адаптацию к специфическим задачам анализа рынка труда.

Заключение

1. Рассмотрены возможности применения технологий больших данных (Big Data) и больших языковых моделей (Large Language Models, LLM) для анализа рынка труда в Республике Беларусь. Проанализированы методы, которые могут улучшить анализ и прогнозирование на рынке труда с использованием современных технологий.

2. Изучение источников Big Data и их классификация показали, что доступные онлайн данные могут быть эффективно использованы для анализа тенденций занятости и выявления востребованных навыков. Зарубежный опыт продемонстрировал, что применение Big Data и LLM позволяет значительно повысить точность и качество аналитики рынка труда. В частности, успешные проекты в ЕС, такие как ESSnet Big Data и инициативы CEDEFOP, могут служить примерами для адаптации в Беларуси.

3. Рекомендации, представленные в статье, направлены на внедрение эффективных инструментов для анализа рынка труда в республике. Важным аспектом является адаптация успешных зарубежных практик с учетом местных особенностей. Гипотезы исследования подтвердились: применение Big Data и LLM действительно может улучшить анализ рынка труда при условии правильной интеграции и обработки данных.

4. Внедрение передовых аналитических инструментов на основе Big Data обеспечит более полное и детализированное понимание динамики рынка труда, что позволит принимать обоснованные решения и разрабатывать эффективные стратегии управления занятостью.

5. Дальнейшие исследования будут направлены на углубленный анализ Big Data, интеграцию различных их источников для создания более полной картины рынка труда, разработку методик оценки эффективности использования Big Data и LLM в анализе рынка труда, исследование успешных зарубежных практик и разработку рекомендаций по их адаптации для условий Беларуси, изучение возможностей автоматизации процессов анализа и масштабирования решений на национальном уровне.

6. Результаты исследования могут быть использованы для создания современных политик и программ, направленных на развитие рынка труда и адаптацию к вызовам цифровой экономики.

Список литературы

1. Mezzanzanica, M. Big Data for Labour Market Intelligence – An Introductory Guide / M. Mezzanzanica, F. Mercorio // European Training Foundation. 2019. <https://www.etf.europa.eu/en/publications-and-resources/publications/big-data-labour-market-intelligence-introductory-guide>.
2. Speaking the Same Language: A Machine Learning Approach to Classify Skills in Burning Glass Technologies Data / J. Lassébie [et al.] // OECD Social, Employment and Migration Working Papers. 2021. No 263. <https://doi.org/10.1787/adb03746-en>.
3. Samek, L. The Human Capital Behind AI: Jobs and Skills Demand from Online Job Postings / L. Samek, M. Squicciarini, E. Cammeraat // OECD Science, Technology and Industry Policy Papers. 2021. No 120. <https://doi.org/10.1787/2e278150-en>.
4. Djumalieva, J. An Open and Data-Driven Taxonomy of Skills Extracted from Online Job Adverts / J. Djumalieva, C. Sleeman // Developing Skills in a Changing World of Work. 2018. P. 425–454.
5. Deming, D. Skill Requirements Across Firms and Labor Markets: Evidence from Job Postings for Professionals / D. Deming, L. B. Kahn // Journal of Labor Economics. 2018. Vol. 36, No S1. P. S337–S369.
6. Djumalieva, J. Classifying Occupations According to Their Skill Requirements in Job Advertisements / J. Djumalieva, A. Lima, C. Sleeman // Economic Statistics Centre of Excellence. 2018. P. 1–37.
7. Bayoán, J. Valuing the U. S. Data Economy Using Machine Learning and Online Job Postings U. S. / J. Bayoán, S. Calderón, D. G. Rassier // Bureau of Economic Analysis. 2022. <https://www.bea.gov/research/papers/2022/valuing-us-data-economy-using-machine-learning-and-online-job-postings>.
8. Cammeraat, E. Burning Glass Technologies' Data Use in Policy-Relevant Analysis: An Occupation-Level Assessment / E. Cammeraat, M. Squicciarini // OECD Science, Technology and Industry Working Papers. 2021. <https://doi.org/10.1787/cd75c3e7-en>.
9. Goldfarb, A. Machine Learning Be a General Purpose Technology? A Comparison of Emerging Technologies Using Data from Online Job Postings / A. Goldfarb, B. Taska, F. Teodoridis // SSRN. 2021. <https://ssrn.com/abstract=3468822>.
10. Brüning, N. What Skills do Employers Seek in Graduates? Using Online Job Posting Data to Support Policy and Practice in Higher Education / N. Brüning, P. Mangeol // OECD Education Working Papers. 2020. No 231. P. 1–47. <https://doi.org/10.1787/bf533d35-en>.

11. Predicting Skill Shortages in Labor Markets: A Machine Learning Approach / N. Dawson [et al.] // 2020 IEEE International Conference on Big Data (Big Data). 2020. P. 3052–3061.
12. Brown, P. The End of the Credential Society? An Analysis of the Relationship Between Education and the Labour Market Using Big Data / P. Brown, M. Souto-Otero // *Journal of Education Policy*. 2018. Vol. 35, No 1. P. 95–118. <https://doi.org/10.1080/02680939.2018.1549752>.
13. Beblavý, M. Demand for Digital Skills in the US Labour Market: The IT Skills Pyramid / M. Beblavý, Br. Fabo, K. Lenaerts // CEPS Special Report. 2016. No 154.
14. Fetsi, A. Changing Skills for a Changing World: Understanding Skills Demand in EU Neighbouring Countries / A. Fetsi, U. Bardak, F. Rosso // European Training Foundation. 2021. https://www.bollettinoadapt.it/wp-content/uploads/2021/02/wcms_771749.pdf.
15. Смирнов, А. Ю. Цифровая модель рынка труда: ключевые аспекты работы программного комплекса / А. Ю. Смирнов // *Экономика труда*. 2023. Т. 10, № 10. С. 1535–1552. DOI: 10.18334/et.10.10.119514.
16. Волгин, Н. А. Спрос на навыки: анализ на основе онлайн-данных о вакансиях* [Электронный ресурс] / А. Д. Волгин, В. Е. Гимпельсон. М.: Изд. дом Высшей школы экономики, 2021.
17. Терников, А. А. Искусственный интеллект и спрос на навыки работников в России / А. А. Терников // *Вопросы экономики*. 2023. № 11. С. 65–80.
18. Ванкевич, Е. В. Изменение подходов к анализу конъюнктуры рынка труда в условиях цифровизации экономики (на примере текстильной промышленности Республики Беларусь) / Е. В. Ванкевич, И. Н. Калиновская // *Известия вузов. Технология текстильной промышленности*. 2022. Т. 401, № 5. С. 27–37. DOI: 10.47367/0021-3497_2022_5_27.
19. Vankevich, A. Better Understanding of the Labour Market Using Big Data / A. Vankevich, I. Kalinouskaya // *Economics and Law*. 2021. Vol. 20, No 3. P. 677–692.
20. Козинец, А. Н. Применение интеллектуального анализа для прогнозирования успешности трудоустройства социально уязвимых групп / А. Н. Козинец // *Цифровая трансформация*. 2024. Т. 30, № 2. С. 33–42. <http://dx.doi.org/10.35596/1729-7648-2024-30-2-33-42>.
21. AI and Jobs: Evidence from Online Vacancies. National Bureau of Economic Research / D. Acemoglu [et al.] // NBER Working Paper. 2020. No 28257. P. 1–55. <https://www.nber.org/papers/w28257>.
22. Colombo, E. AI Meets Labor Market: Exploring the Link Between Automation and Skills / E. Colombo, F. Mercorio, M. Mezzanzanica // *Information Economics and Policy*. 2019. Vol. 47. P. 27–37. <https://doi.org/10.1016/j.infoecopol.2019.05.003>.
23. Estimating Industry 4.0 Impact on Job Profiles and Skills Using Text Mining / S. Fareri [et al.] // *Computers in Industry*. 2020. Vol. 118. <https://doi.org/10.1016/j.compind.2020.103222>.

References

1. Mezzanzanica M., Mercorio F. (2019) Big Data for Labour Market Intelligence – An Introductory Guide. *European Training Foundation*. <https://www.etf.europa.eu/en/publications-and-resources/publications/big-data-labour-market-intelligence-introductory-guide>.
2. Lassébie J., Marcolin L., Vandeweyer M., Vignal B. (2021) Speaking the Same Language: A Machine Learning Approach to Classify Skills in Burning Glass Technologies Data. *OECD Social, Employment and Migration Working Papers*. (263). <https://doi.org/10.1787/adb03746-en>.
3. Samek L., Squicciarini M., Cammeraat E. (2021) The Human Capital Behind AI: Jobs and Skills Demand from Online Job Postings. *OECD Science, Technology and Industry Policy Papers*. (120). <https://doi.org/10.1787/2e278150-en>.
4. Djumalieva J., Sleeman C. (2018) An Open and Data-Driven Taxonomy of Skills Extracted from Online Job Adverts. *Developing Skills in a Changing World of Work*. 425–454.
5. Deming D., Kahn L. B. (2018) Skill Requirements Across Firms and Labor Markets: Evidence from Job Postings for Professionals. *Journal of Labor Economics*. 36 (S1), S337–S369.
6. Djumalieva J., Lima A., Sleeman C. (2018) Classifying Occupations According to Their Skill Requirements in Job Advertisements. *Economic Statistics Centre of Excellence*. 1–37.
7. Bayoán J., Calderón S., Rassier D. G. (2022) Valuing the U. S. Data Economy Using Machine Learning and Online Job Postings U. S. *Bureau of Economic Analysis*. <https://www.bea.gov/research/papers/2022/valuing-us-data-economy-using-machine-learning-and-online-job-postings>.
8. Cammeraat E., Squicciarini M. (2021) Burning Glass Technologies’ Data Use in Policy-Relevant Analysis: An Occupation-Level Assessment. *OECD Science, Technology and Industry Working Papers*. <https://doi.org/10.1787/cd75c3e7-en>.
9. Goldfarb A., Taska B., Teodoridis F. (2021) Machine Learning Be a General Purpose Technology? A Comparison of Emerging Technologies Using Data from Online Job Postings. *SSRN*. <https://ssrn.com/abstract=3468822>.

10. Brüning N., Mangeol P. (2020) What Skills Do Employers Seek in Graduates? Using Online Job Posting Data to Support Policy and Practice in Higher Education. *OECD Education Working Papers*. (231). <https://doi.org/10.1787/bf533d35-en>.
11. Dawson N., Rizoiu M.-A., Johnston B., Williams M.-A. (2020) Predicting Skill Shortages in Labor Markets: A Machine Learning Approach. *2020 IEEE International Conference on Big Data (Big Data)*. 3052–3061.
12. Brown P., Souto-Otero M. (2018) The End of the Credential Society? An analysis of the Relationship Between Education and the Labour Market Using Big Data. *Journal of Education Policy*. 35 (1), 95–118. <https://doi.org/10.1080/02680939.2018.1549752>.
13. Beblavý M., Fabo Br., Lenaerts K. (2016) Demand for Digital Skills in the US Labour Market: The IT Skills Pyramid. *CEPS Special Report*. (154).
14. Fetsi A., Bardak U., Rosso F. (2021) Changing Skills for a Changing World: Understanding Skills Demand in EU Neighbouring Countries. *European Training Foundation*. https://www.bollettinoadapt.it/wp-content/uploads/2021/02/wcms_771749.pdf.
15. Smirnov A. Y. (2023) Digital Model of the Labor Market: Key Aspects of the Software Package. *Labor Economics*. 10 (10), 1535–1552. DOI: 10.18334/et.10.10.119514 (in Russian).
16. Volgin N. A., Gimpelson V. E. (2021) *Demand for Skills: Analysis Based on Online Job Data*. Moscow, Publishing House of the Higher School of Economics (in Russian).
17. Chernikov A. A. (2023) Artificial Intelligence and the Demand for Workers' Skills in Russia. *Questions of Economics*. (11), 65–80 (in Russian).
18. Vankevich A. V., Kalinovskaya I. N. (2022) Changing Approaches to the Analysis of Labor Market Conditions in the Context of Digitalization of the Economy (on the Example of the Textile Industry of the Republic of Belarus). *News of Universities. Technology of the Textile Industry*. 401 (5), 27–37. DOI: 10.47367/0021-3497_2022_5_27 (in Russian).
19. Vankevich A., Kalinovskaya I. (2021) Better Understanding of the Labour Market Using Big Data. *Economics and Law*. 20 (3), 677–692.
20. Kazinets A. N. (2024) Application of Intelligent Data Analysis to Predict the Employment Success of Socially Vulnerable Groups. *Digital Transformation*. 30 (2), 33–42. <http://dx.doi.org/10.35596/1729-7648-2024-30-2-33-42>. (in Russian).
21. Acemoglu D., Autor D., Hazell J., Restrepo P. (2020) AI and Jobs: Evidence from Online Vacancies. *NBER Working Paper*. (28257), 1–55. <https://www.nber.org/papers/w28257>.
22. Colombo E., Mercorio F., Mezzanzanica M. (2019) AI Meets Labor Market: Exploring the Link Between Automation and Skills. *Information Economics and Policy*. 47, 27–37. <https://doi.org/10.1016/j.infoecopol.2019.05.003>.
23. Fareri S., Fantoni G., Chiarello F., Coli E., Binda A. (2020) Estimating Industry 4.0 Impact on Job Profiles and Skills Using Text Mining. *Computers in Industry*. 118. <https://doi.org/10.1016/j.compind.2020.103222>.

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Ванкевич Е. В., д-р экон. наук, проф., проректор по научной работе, Витебский государственный технологический университет

Калиновская И. Н., канд. техн. наук, доц. каф. экономики и электронного бизнеса, Витебский государственный технологический университет

Адрес для корреспонденции

210039, Республика Беларусь,
г. Витебск, просп. Московский, 72
Витебский государственный
технологический университет
Тел.: +375 29 663-05-49
E-mail: vankevich_ev@tut.by
Ванкевич Елена Васильевна

Information about the authors

Vankevich A. V., Dr. of Sci. (Econ.), Professor, Vice-Rector for Research, Vitebsk State Technological University

Kalinovskaya I. N., Cand. of Sci., Associate Professor at the Department of Economics and Electronic Business, Vitebsk State Technological University

Address for correspondence

210039, Republic of Belarus,
Vitebsk, Moskovsky Ave., 72
Vitebsk State
Technological University
Tel.: +375 (29) 663-05-49
E-mail: vankevich_ev@tut.by
Vankevich Alena Vasilievna