

О. В. Зайцева Витебский государственный технологический университет, г. Витебск, Республика Беларусь

Прогнозирование безработицы в Республике Беларусь с использованием данных поисковых запросов Google Trends

Аннотация. Проблема прогнозирования уровня безработицы в Республике Беларусь остается актуальной в условиях динамично изменяющейся экономической среды. Традиционные методы прогнозирования, основанные на официальной статистике, зачастую не учитывают оперативные изменения на рынке труда, что снижает их точность. В то же время данные поисковых запросов доказали свою эффективность в качестве опережающих индикаторов в других странах, однако, их применение в Беларуси остается неизученным. Статья посвящена оценке возможности использования данных поисковых запросов для повышения точности прогнозирования уровня безработицы в Республике Беларусь. Методологическую основу исследования сформировали теоретические положения макроэкономики и модели SARIMA, VAR и SARIMAX. Методы включали декомпозицию временных рядов, тест Дики – Фуллера для оценки стационарности, дифференцирование, стандартизацию данных и тест Грейнджера на причинность. Информационную базу составили данные Национального статистического комитета Республики Беларусь об уровне безработицы за 2015–2024 гг. и данные о поисковых запросах Google, связанные с поиском работы. Выявлено, что модель SARIMAX с включением данных о поисковых запросах превосходит классические модели прогнозирования безработицы, демонстрируя минимальные ошибки. Согласно этой модели, прогнозные значения уровня безработицы показали тенденцию к снижению, что отражает устойчивую динамику улучшения ситуации на рынке труда Республики Беларусь. Полученные результаты подчеркивают значимость комбинирования традиционных данных с цифровыми метриками для повышения точности прогнозов, а также открывают перспективы для дальнейших исследований в области применения интернет-данных для социально-экономического анализа, включая разработку более совершенных моделей прогнозирования безработицы.

Ключевые слова: рынок труда; уровень безработицы; прогнозирование; Google Trends; поисковые запросы.

Для цитирования: Zaitseva O. V. (2025). Forecasting unemployment in the Republic of Belarus using Google Trends data. *Journal of New Economy*, vol. 26, no. 2, pp. 45–63. DOI: 10.29141/2658-5081-2025-26-2-3. EDN: RHBZLH.

Информация о статье: поступила 27 января 2025 г.; доработана 6 марта 2025 г.; одобрена 17 марта 2025 г.

Olga V. Zaitseva

Vitebsk State Technological University, Vitebsk, the Republic of Belarus

Forecasting unemployment in the Republic of Belarus using Google Trends data

Abstract. The problem of unemployment forecasting in the Republic of Belarus keeps its relevance against the background of the dynamically changing economic environment. Conventional forecasting methods based on official statistics often fail to account for real-time changes in the labour market and thus lose their accuracy. At the same time, search query data has proven to be effective as leading indicators in other countries, but its application in Belarus remains unexplored. The paper aims to assess the potential of using search query data to improve the accuracy of unemployment rate forecasting in the Republic of Belarus. Methodologically, the study rests on the theoretical propositions of macroeconomics and SARIMA, VAR, and SARIMAX models. The methods include time series decomposition, the Dickey-Fuller stationarity test, differencing, data standardisation, and the Granger causality test. The evidence is unemployment data of the National Statistical Committee of the Republic of Belarus for 2015–2024 and Google search query data related to job searches. The study found that the SARIMAX model incorporating search query data outperforms classical models of unemployment forecasting by demonstrating a minimum of errors. According to this model, the predicted values of unemployment rate reflect a downward trend indicating the sustained improvement in the labour market of the Republic of Belarus. The findings emphasise the importance of combining traditional data with digital metrics to increase the forecast accuracy as well as open up prospects for further research into the use of internet data for socioeconomic analysis, including the development of more advanced unemployment rate forecasting models.

Keywords: labour market; unemployment rate; forecasting; Google Trends; web query.

For citation: Zaitseva O. V. (2025). Forecasting unemployment in the Republic of Belarus using Google Trends data. *Journal of New Economy*, vol. 26, no. 2, pp. 45–63. DOI: 10.29141/2658-5081-2025-26-2-3. EDN: RHBZLH.

Article info: received January 27, 2025; received in revised form March 6, 2025; accepted March 17, 2025

Введение

С развитием сети Интернет мир столкнулся с преобразующими изменениями, оказывающими значительное влияние на различные аспекты социально-экономической жизни населения. Эти изменения затронули в том числе и рынок труда. Возможности, которые открылись вместе с доступом к онлайн-ресурсам, повлияли на способы поиска работы как среди безработных, так и среди занятых. Постепенно становится доминирующим способ поиска работы через интернет, поскольку он повышает вероятность успешного трудоустройства при минимальных временных затратах. Республика Беларусь не стала исключением: по данным обследования домашних хозяйств по изучению проблем занятости населения, в среднем 70 % безработных предпочитают такой способ поиска работы¹.

¹ Занятость населения в 2022 году (по материалам выборочного обследования). Статистический бюллетень. Минск: Национальный статистический комитет Республики Беларусь, 2023; Занятость населения в 2023 году (по материалам выборочного обследования). Статистический бюллетень. Минск: Национальный статистический комитет Республики Беларусь, 2024.

Анализ международных исследований показывает, что активность экономических субъектов в интернете, связанная с поиском информации о вакансиях, пособиях по безработице, условиях увольнения, может служить индикатором будущих изменений уровня безработицы. Таким образом, данные поисковых запросов могут иметь значительную предсказательную силу, что позволяет органам государственного управления более обоснованно прогнозировать уровень безработицы.

Преимущества использования поисковых запросов в прогнозировании уровня безработицы достигаются благодаря открытому доступу к высокочастотным и своевременно публикуемым данным, которые превосходят традиционные данные опросов. Информация, получаемая из сведений о поисковых запросах – полезный источник экономических данных, в особенности для краткосрочного прогнозирования в периоды экономической нестабильности.

Проблема исследования заключается в том, что существующие методы прогнозирования уровня безработицы в Республике Беларусь, основанные на традиционных источниках данных (опросы домохозяйств, официальная статистика), не всегда позволяют оперативно реагировать на изменения на рынке труда, особенно в условиях экономической нестабильности. Это связано с тем, что традиционные источники имеют временной лаг и не учитывают поведенческие изменения в реальном времени. В то же время в мировой практике для прогнозирования социально-экономических показателей (включая безработицу) активно используются данные поисковых запросов. Однако в Беларуси подобные подходы остаются недостаточно изученными и применяемыми, что создает пробел в исследованиях и ограничивает возможности для более точного и своевременного прогнозирования.

Одним из самых распространенных источников данных о поисковых запросах является инструмент Google Trends, позволяющий пользователям исследовать, как меняется интерес к различным темам, продуктам или явлениям в разных географических регионах и временных периодах, начиная с 2004 г. и до текущего момента.

Для русскоязычного сегмента интернета аналогом Google Trends является Яндекс Вордстат, предоставляющий статистику поисковых запросов в Яндексе. Этот инструмент позволяет получать точные данные о количестве показов по конкретным поисковым запросам, включая их региональную популярность и сезонность. Основным преимуществом этого сервиса является то, что он показывает абсолютные значения количества запросов, а не относительные показатели (в отличие от Google Trends). Однако, основным ограничением для использования сведений о поисковых запросах Яндекс является ограниченность выборки данных (доступны только с января 2018 г.). Кроме того, в Республике Беларусь поисковая система Яндекс является менее распространенной по сравнению с Google: согласно информации Яндекс Радар¹, в декабре 2024 г. доля визитов Google составила 68,84 %, а доля Яндекс – 30,81 %. Следовательно, Google представляет более надежную оценку поисковых запросов белорусов. Эти обстоятельства обусловили выбор в качестве предикторов уровня безработицы в Республике Беларусь данных поисковых запросов Google Trends.

Целью исследования является прогнозирование уровня безработицы в Республике Беларусь с использованием данных поисковых запросов Google Trends и сравнение эффективности полученных моделей с традиционными моделями прогнозирования.

Для решения поставленной цели необходимо решить следующие задачи:

– провести обзор исследований, посвященных прогнозированию безработицы с использованием данных поисковых запросов;

¹ Яндекс.Радар. Поисковые системы в Беларуси. <https://radar.yandex.ru/search?country=149>.

- подобрать наиболее релевантные поисковые запросы, связанные с поиском работы;
- проанализировать взаимосвязь поисковых запросов с уровнем безработицы и выбрать запросы, обладающие наибольшей предиктивной силой;
- построить и сравнить традиционные модели прогнозирования уровня безработицы с моделями, включающими данные поисковых запросов Google в качестве экзогенных переменных;
- оценить эффективность моделей и определить, какие из них обладают наибольшей прогностической силой.

Объектом исследования является динамика уровня безработицы в Республике Беларусь. Предмет исследования – прогностическая сила традиционных моделей прогнозирования безработицы, а также моделей временных рядов, включающих данные поисковых запросов Google Trends.

Гипотезой H_0 исследования является то, что модели, включающие данные о поисковых запросах Google Trends, обладают значительно большей прогностической силой по сравнению с традиционными моделями прогнозирования безработицы.

Впервые для Республики Беларусь предложено использование данных поисковых запросов Google Trends в качестве экзогенных переменных для прогнозирования уровня безработицы. Это позволяет не только повысить точность прогнозов, но и учесть оперативные изменения в поведении экономических субъектов, что необходимо в условиях быстро меняющейся экономической среды.

Теоретический обзор

Статистика поисковых запросов открывает новые возможности для прогнозирования и анализа различных экономических и социальных показателей. С каждым годом увеличивается количество исследований, посвященных использованию поисковых запросов в различных областях, таких как наукастинг и прогнозирование цен [Choi, Varian, 2009], продаж [Зубарев, Голованова, 2021], показателей демографии [Калабихина и др., 2020] и миграции [Броницкий, Вакуленко, 2022], а также уровня безработицы. Как показывает практика, связь между данными поисковых запросов и официальной статистикой становится все более очевидной, давая исследователям и экспертам ценные инструменты для анализа рыночных тенденций и социальных изменений.

Одно из первых значимых исследований в сфере безработицы, проведенное рядом ученых в 2005 г., продемонстрировало значительное влияние данных поисковых запросов на прогнозирование уровня безработицы в США, что стало катализатором для дальнейших исследований в этом направлении [Ettredge, Gerdes, Karuga, 2005].

Некоторые исследователи выделяют высокий уровень корреляции между уровнем безработицы и конкретными поисковыми запросами в системе Google по ключевым словам [Askitas, Zimmermann, 2009].

В другой работе авторы используют ежемесячные статистические данные и также подтверждают, что сведения о поисковых запросах Google улучшают прогноз уровня безработицы, а модели, дополненные Google Trends Index (GTI), демонстрируют более высокую точность прогнозирования по сравнению с традиционными моделями [D'Amuri, Marcucci, 2010].

В одной из работ точность прогноза уровня безработицы при использовании инструмента Google Trends совместно с официальными статистическими данными превосходит точность прогноза, основанного на использовании лишь официальных статистических данных [Simionescu, 2020]. Дальнейшие исследования автора также это подтверждают [Simionescu, Cifuentes-Faura, 2022; Simionescu, 2024].

В другой работе ученые анализируют возможности прогнозирования уровня безработицы в Гане с использованием данных Google Trends и приходят к выводу, что комбинация моделей VAR, ARIMAX и ARIMA в сочетании с поисковыми запросами в реальном времени позволяет эффективно предсказывать динамику показателей безработицы в развивающихся странах [Adu, Appiahene, Afrifa, 2023].

Эффективность применения данных поисковых запросов подтверждается и множеством других зарубежных и русскоязычных исследований [Fondeur, Karamé, 2013; D'Amuri, Marcucci, 2017; Naccarato et al., 2018; González-Fernández, González-Velasco, 2018; Nagao, Takeda, Tanaka, 2019; Claveria, 2019; Куровский, 2019; Borup, Christian, Schütte, 2020; Петрова, Трунин, 2020; Юревич, Ахмадеев, 2021; Mulero, Garcia-Hiernaux, 2023; Costa, Silva, Galvão, 2024].

Таким образом, эмпирические результаты доказывают, что данные поисковых запросов Google Trends добавляют релевантную информацию для прогнозирования безработицы по сравнению с показателями бизнес-цикла или традиционными моделями временных рядов. Это подтверждается множеством исследований, проведенных в различных странах. В частности, такие исследования демонстрируют, что поисковые запросы, связанные с поиском работы, социальными пособиями и другими аспектами занятости, могут служить надежными индикаторами изменения уровня безработицы.

Важно отметить, что исследований, в которых для прогнозирования безработицы используются данные Google Trends, в Республике Беларусь еще не было. Большинство работ опирается на традиционные методы прогнозирования и статистического анализа [Слаут, Цехан, 2019; Королецкая, 2021; Vankevich, Kalinouskaya, 2021; Ванкевич, Зайцева, 2023; Vankevich et al., 2023].

Несмотря на отсутствие таких исследований, потенциал использования Google Trends для прогнозирования безработицы в Республике Беларусь остается высоким. Учитывая опыт других стран, можно предположить, что данные поисковых запросов могут быть полезны для анализа текущих тенденций на рынке труда и прогнозирования уровня безработицы.

Однако для успешного применения этого подхода в Республике Беларусь необходимо учитывать специфику ее рынка труда, включая особенности поведения пользователей в поисковых системах. Проведение подобных исследований в Беларуси может не только расширить методологическую базу для прогнозирования безработицы, но и предоставить новые инструменты для разработки более эффективной социально-экономической политики. Внедрение данных Google Trends в прогнозирование безработицы может стать значимым шагом в развитии аналитических методов, что позволит оперативно реагировать на изменения в экономической ситуации и улучшать качество принимаемых решений.

Материалы и методы

Теоретической и методологической базой исследования стали ключевые положения макроэкономики, связанные с анализом рынка труда и прогнозированием уровня безработицы. В статье также будем учитывать современные подходы к анализу рынка труда, которые подчеркивают важность поведенческих факторов и оперативных данных для прогнозирования уровня безработицы. Использование сведений о поисковых запросах Google Trends позволяет учитывать изменения в поведении экономических агентов в реальном времени, что необходимо в условиях быстро меняющейся экономической среды. Таким образом, методологическая база исследования сочетает в себе как традиционные

макроэкономические теории, так и современные подходы к анализу данных, что позволяет повысить точность прогнозов уровня безработицы.

В качестве источников информации для прогнозирования безработицы были использованы официальные статистические данные Республики Беларусь (Национального статистического комитета, Министерства труда и социальной защиты, а также выборочного обследования домашних хозяйств по изучению проблем занятости населения) и поисковые запросы Google Trends.

Данные по безработице. В качестве показателя безработицы был использован уровень безработицы, рассчитанный по методологии Международной организации труда. Так как информация о фактическом уровне безработицы стала формироваться с момента начала проведения выборочных обследований домашних хозяйств по изучению проблем занятости населения, для построения прогноза будут использоваться квартальные данные с декабря 2015 г. по сентябрь 2024 г.

Для сопоставимости сведений была проведена интерполяция квартальных данных по уровню безработицы с целью получения оценочных месячных значений с декабря 2015 г. по сентябрь 2024 г. (рис. 1). Для решения задачи был применен метод кубической сплайн-интерполяции (CubicSpline) из библиотеки SciPy языка Python, который обеспечивает непрерывность функции и ее первой и второй производных в узловых точках, что позволяет получить гладкую аппроксимацию исходных данных с сохранением их динамических характеристик.

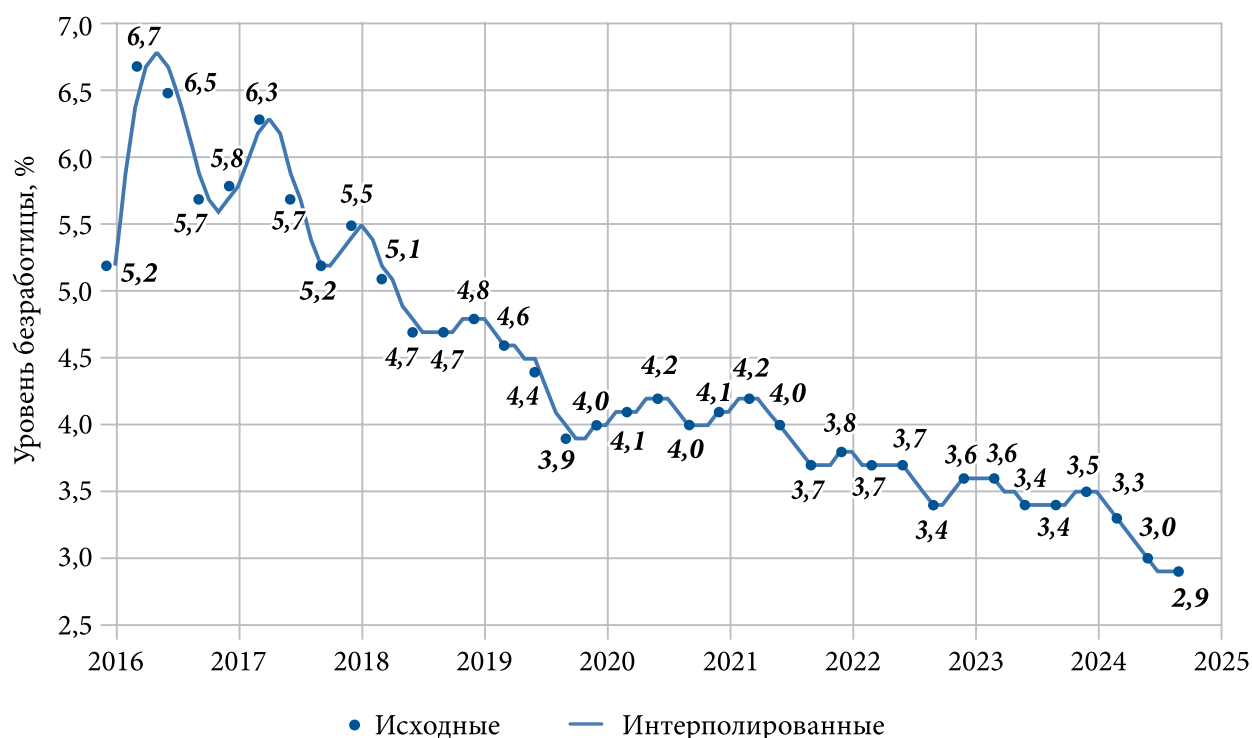


Рис. 1. Исходные и интерполированные данные по уровню безработицы в Республике Беларусь, декабрь 2015 – сентябрь 2024

Fig. 1. Input and interpolated data on the unemployment rate in the Republic of Belarus, December 2015 – September 2024

Базовые характеристики исходного и интерполированного ряда представлены в табл. 1.

Таблица 1. Базовые статистики исходного и интерполированного ряда

Table 1. Basic statistics of the input and interpolated series

Статистический показатель	Исходные данные	Интерполированные данные
Количество наблюдений	36	105
Среднее	4,4	4,41
Стандартное отклонение	1,0	1,0
Минимальное значение	2,9	2,9
25 % (первый квартиль)	3,68	3,6
50 % (медиана)	4,10	4,1
75 % (третий квартиль)	5,12	5,2
Максимальное значение	6,7	6,8
Асимметрия	0,73	0,72
Эксцесс	-0,33	-0,47
Среднее отклонение	0,009	
Максимальное отклонение	0,1	

Анализ, приведенный в табл. 1, показал, что интерполированный ряд хорошо сохраняет статистические характеристики исходного ряда. Это подтверждается минимальными различиями в основных показателях (среднее отклонение всего 0,009, идентичные стандартные отклонения и схожие показатели асимметрии и эксцесса), при этом увеличивая количество наблюдений с 36 до 105.

Данные поисковых запросов. В качестве экзогенных переменных в исследовании были использованы поисковые запросы Google – Google Trends Index (GTI).

Данные Google Trends Index представляют собой временные ряды со значениями относительной популярности запроса по месяцам с 2004 г. по настоящее время. Относительная популярность запроса – это внутренняя метрика сервиса Google Trends, которая представляет собой отношение текущего (сумма всех запросов в течение месяца) количества запросов к количеству всех запросов к Google в этом месяце по региону, нормализованное на максимальное значение временного ряда, то есть значение популярности запроса находится в диапазоне от 0 до 100 %. Популярность, соответствующая 0, говорит о малой доле запросов, посвященных определенной теме. 100 % означает наивысший уровень интереса к запросу в регионе, значение 50 % – уровень популярности запроса вдвое меньше по отношению к самому популярному. В статистику не попадают редкие запросы, одинаковые запросы от одного пользователя за короткий период времени и фразы, содержащие спецсимволы.

Несмотря на то, что Google Trends Index представляет собой индикатор состояния рынка труда, часто предшествующий официальным статистическим данным, и позволяющий оценивать динамику безработицы практически в реальном времени (это является необходимым для принятия быстрых решений в области социальной и экономической политики), он имеет также и очевидные ограничения при использовании в качестве потенциального предиктора.

М. Дилмагани перечисляет четыре проблемы [Dilmaghani, 2019]. Во-первых, источники данных поисковых систем не являются вероятностными выборками населения. Поскольку эти данные отражают часть населения, пользующуюся интернетом, они потенциально могут страдать от предвзятости отбора (например, запросы о поиске работы,

возможно, задаются молодыми людьми из крупных городских центров). Во-вторых, Google Trends не делает различий между теми, кто ищет работу, будучи безработным или только размышляя о смене работы. Это различие может иметь решающее значение, так как запросы о работе для безработных часто демонстрируют антициклическое поведение, в то время как поиск, осуществляемый работниками, обычно предполагается проциклическим. В-третьих, Google Trends не содержит информации о социально-демографических характеристиках пользователей. В-четвертых, индекс Google Trends рассчитывается методом выборочного анализа. Следовательно, ряд Google Trends способен измениться, если добавить несколько новых наблюдений, что может привести к некоторой погрешности в оценках.

Помимо вышеперечисленных ограничений, неточности могут возникать в результате неправильного подбора ключевых слов.

Существуют различные способы подбора ключевых слов: статистика самих поисковых систем, использование искусственного интеллекта и машинного обучения, сезонный и трендовый анализ, экспертные методы и другие.

В нашей работе была использована статистика самих поисковых систем, а также информация, предоставляемая инструментом Google Trends «похожие запросы», которая позволила отследить другие запросы, связанные с выбранными терминами.

Частично используя опыт других исследований [Mulero, Garcia-Hiernaux, 2023], можно сгруппировать поисковые запросы в три группы:

- запросы, связанные с ведущими онлайн-порталами для поиска работы;
- запросы, связанные с государственной службой занятости;
- запросы, связанные со стандартными запросами поиска работы.

Далее был проведен поиск по запросам о работе с декабря 2015 г. по сентябрь 2024 г. и выбраны те, которые имеют наибольший индекс Google Trends. Запросы, имеющие очень низкие или нулевые значения GTI, не были включены в модель.

Результаты анализа показали, что в первой группе наиболее частыми являются запросы, связанные с самыми крупными онлайн-порталами вакансий в Республике Беларусь: *rabota.by*, *prasa.by* и *belmeta.com*. Динамика поисковых запросов по первой группе представлена на рис. 2 и отражает реальную востребованность портала *rabota.by* в Республике Беларусь [Ванкевич, Зайцева, 2020].

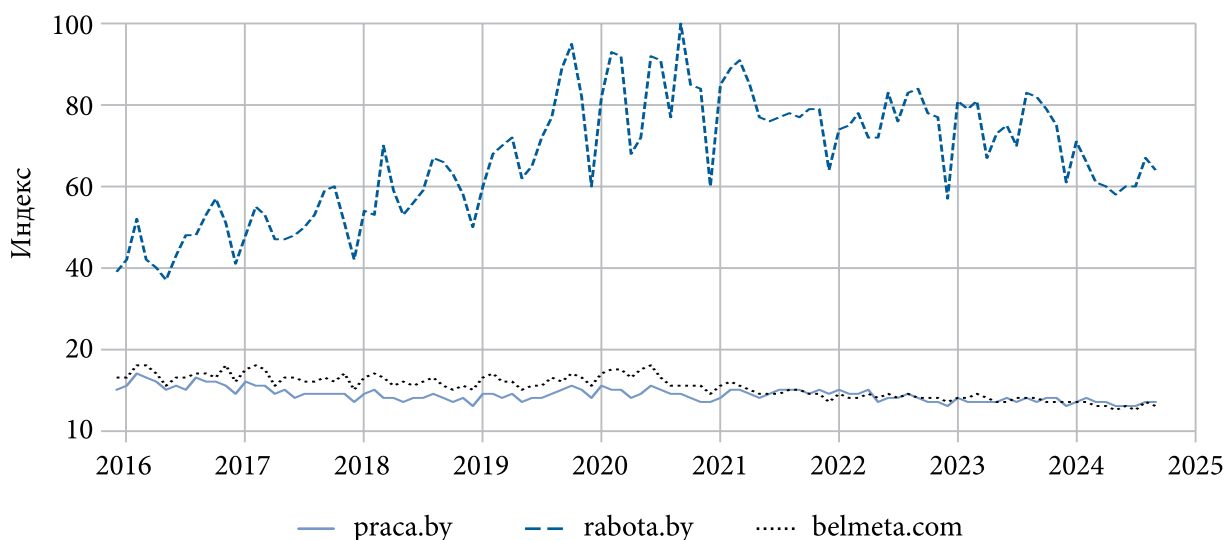


Рис. 2. Динамика GTI по различным запросам в отношении онлайн-порталов вакансий

Fig. 2. GTI dynamics for various queries related to online job portals

Анализ поисковых запросов, относящихся к второй и третьей группе, позволил выделить запросы «биржа труда», «служба занятости» и «работа» (табл. 2).

Таблица 2. Наиболее популярные поисковые запросы, связанные с поиском работы в Республике Беларусь

Table 2. The most popular search queries related to job search in the Republic of Belarus

Группа запросов	Ключевые слова
1) связанные с ведущими онлайн-порталами для поиска работы	праца бай работа бай белмета
2) связанные с государственной службой занятости	биржа труда служба занятости
3) связанные со стандартными запросами поиска работы	работа

Таким образом, анализ поисковых запросов, связанных с поиском работы в Республике Беларусь, показал, что наиболее популярными являются запросы, связанные с ведущими онлайн-порталами вакансий, такими как *rabota.by*, которые отражают реальную востребованность этих платформ. Также значительный интерес вызывают запросы о государственной службе занятости, включая «биржа труда» и «служба занятости», что подчеркивает значимость государственных мер поддержки трудоустройства. Универсальный запрос «работа» из третьей группы указывает на его широкую применимость и отражает общие тенденции в интересе соискателей.

Результаты исследования

В ходе первичного анализа временного ряда уровня безработицы была произведена декомпозиция ряда на тренд, сезонную составляющую и на остатки (рис. 3).

Анализ декомпозиции ряда позволил выявить четкий нисходящий тренд на протяжении всего периода. Снижение происходит неравномерно: более резкое падение в начале периода и более плавное к концу. Тренд показывает общее снижение уровня безработицы примерно с 6 до 3 %. В конце периода наблюдается некоторое замедление темпов снижения.

В ряду присутствует устойчивый сезонный паттерн, амплитуда сезонных колебаний относительно небольшая ($\pm 0,2-0,3$ процентных пункта), пики безработицы приходятся на начало года (1-й квартал), минимумы наблюдаются ближе к концу года (3-й квартал). Сезонность имеет регулярный характер и повторяется из года в год.

Отсутствие четкой структуры в остатках говорит о хорошем качестве декомпозиции.

Таким образом, в данных присутствует явная комбинация сильного тренда и заметной сезонности. Тренд является доминирующей компонентой, что указывает на устойчивые долгосрочные изменения в уровне безработицы. Сезонная составляющая также значима, что говорит о регулярных циклических паттернах внутри года. Эти результаты свидетельствуют о том, что временной ряд имеет четкую структуру с минимальным количеством случайного шума, что делает его подходящим объектом для прогнозирования с потенциально высокой точностью.

Наличие тренда и сезонности свидетельствует о нестационарности временного ряда, что подтверждено также и результатами расширенного теста Дики – Фуллера (ADF). Для приведения ряда к стационарному виду было выполнено две процедуры дифференцирования, что обеспечило значительное улучшение результатов теста ADF.

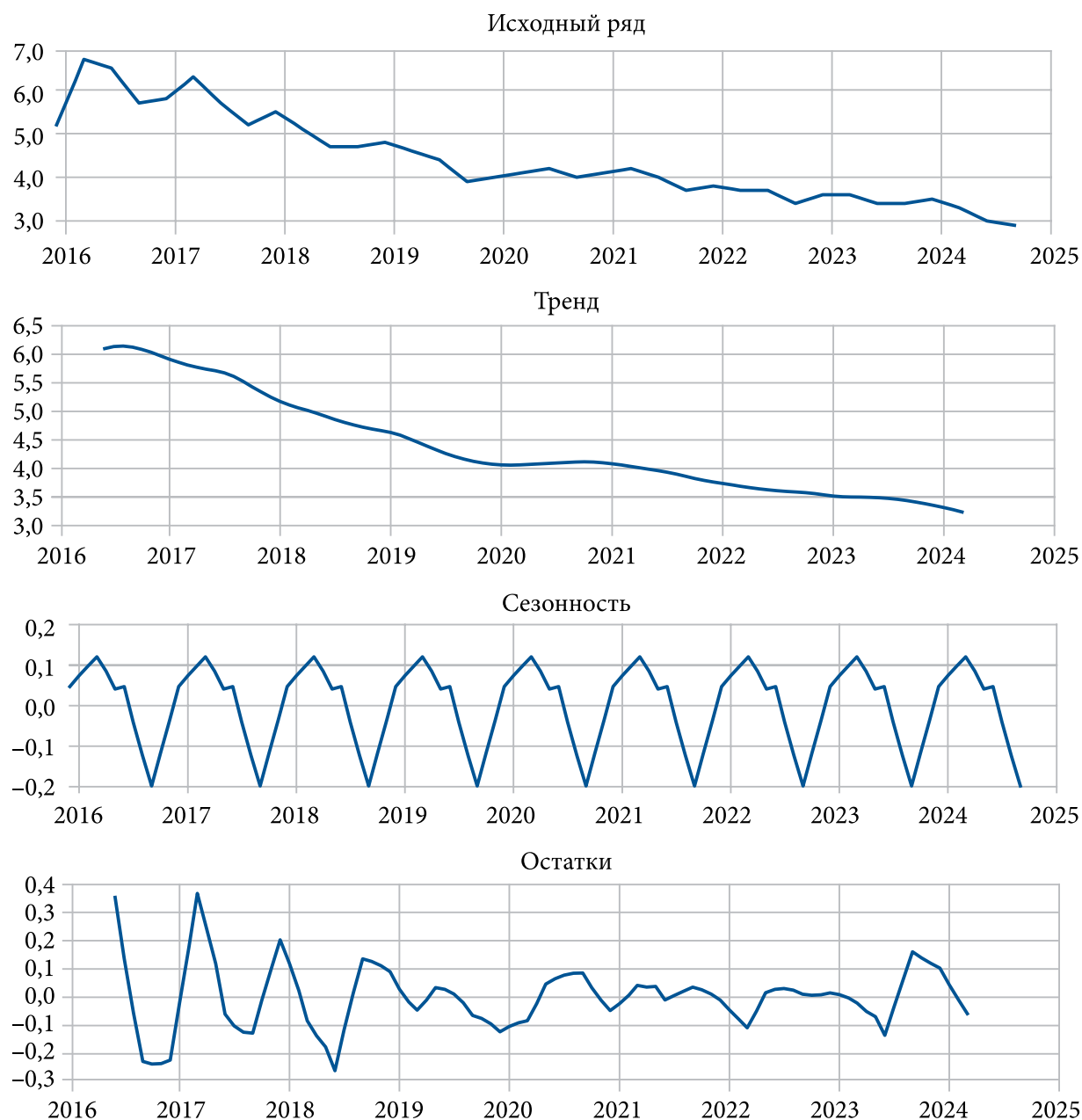


Рис. 3. Декомпозиция временного ряда уровня безработицы

Fig. 3. Decomposition of the unemployment rate time series

В процессе анализа временных рядов с данными о поисковых запросах была также осуществлена последовательная проверка их стационарности и выполнение необходимых преобразований. Ряды, не удовлетворяющие критерию стационарности по тесту ADF, подвергались последовательным преобразованиям, включая первое и второе дифференцирование, а также при необходимости, логарифмированию. В результате такие ряды, как «работа бай» и «работа», потребовали двойного дифференцирования для достижения стационарности, что подтверждалось значением p -value теста ADF ниже критического уровня 0,05. Другие ряды («праца бай», «белмета» и «биржа труда») смогли достичь стационарности через первое дифференцирование. Временной ряд с поисковым запросом «служба занятости» изначально удовлетворял требованиям стационарности.

После приведения данных к стационарности были устранены пропуски, возникшие в результате преобразований.

Так как индексы Google Trends являются нормализованной величиной и нормализованы по разным максимумам, сравнивать их между собой и с уровнем безработицы проблематично. В связи с этим была применена процедура стандартизации данных с использованием метода Z-преобразования. Это достигается путем вычисления Z-оценки для каждого значения:

$$Z = (X - \mu) / \sigma, \quad (1)$$

где X – значение ряда; μ – среднее значение ряда; σ – стандартное отклонение ряда.

После стандартизации каждый ряд будет иметь среднее (μ) равное 0 и стандартное отклонение (σ) равное 1. Это делает данные более сравнимыми.

Такие преобразования позволили масштабировать значения всех временных рядов таким образом, чтобы они имели среднее значение 0 и стандартное отклонение 1, что упрощает их последующий анализ и использование в моделировании.

Для выявления причинно-следственных связей между поисковыми запросами и уровнем безработицы был проведен тест Грейнджера на причинность, который позволил сделать следующие выводы:

- наиболее сильную причинно-следственную связь с уровнем безработицы имеют поисковые запросы «биржа труда» и «работа» (для запроса «работа» уровень значимости p -value составил 0,00239, для запроса «биржа труда» – 0,01636, что указывает на статистически значимую связь);

- остальные поисковые запросы, такие как «работа бай» (p -value = 0,11734), «белмета» (p -value = 0,11239) и «служба занятости» (p -value = 0,15598), показали слабую связь с уровнем безработицы.

Таким образом, результаты теста позволили предположить, что поисковые запросы «работа» и «биржа труда» могут быть полезны в качестве экзогенных предикторов для модели временных рядов.

Модель SARIMA. В рамках исследования была построена модель SARIMA для прогнозирования значений уровня безработицы (рис. 4). Для оптимального выбора параметров модели была использована процедура Auto-ARIMA, которая выполняет автоматический подбор гиперпараметров на основе минимизации информационного критерия Акаике (AIC).

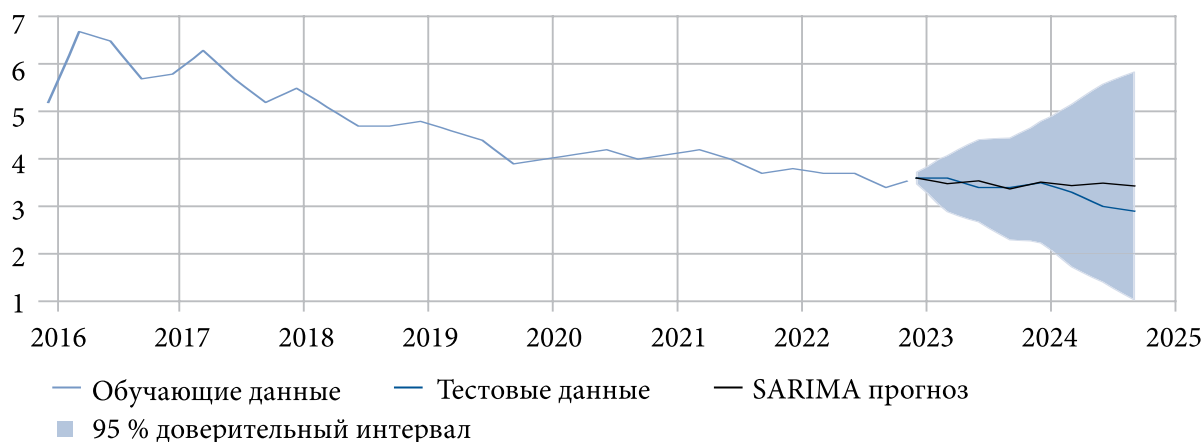


Рис. 4. Прогноз уровня безработицы (модель SARIMA)

Fig. 4. Unemployment rate forecast (SARIMA model)

По результатам алгоритма наилучшая модель была определена как ARIMA(2,1,2) (2,0,0)[12], что указывает на наличие двух авторегрессионных (AR) и двух скользящих средних (MA) компонент, а также на присутствие сезонного компонента с порядком два.

После оценки модели были получены следующие метрики: среднеквадратичная ошибка (MSE) составила 0,0614, а средняя абсолютная ошибка (MAE) равнялась 0,1645. Эти значения свидетельствуют о том, что модель обеспечивает достаточно точный прогноз, с учетом имеющихся относительно низких значений ошибок.

Для оценки соответствия модели предположениям регрессионного анализа была проведена диагностика остатков. Для этого использовались тесты Шапиро – Уилка и Д’Агостино – Пирсона. Оба теста дали статистически значимые результаты (p-value значительно меньше уровня значимости 0,05), так, гипотеза H_0 нормальности распределения остатков была отвергнута.

С целью дальнейшей проверки построены графики автокорреляционной функции (ACF) и частичной автокорреляционной функции (PACF) остатков (рис. 5). Они продемонстрировали отсутствие значимой автокорреляции, что подтверждается результатами теста Льюнга – Бокса, который выявил значение статистики, равное 1,87, с p-значением 0,997. Это указывает на то, что остатки модели представляют собой белый шум, то есть являются случайным и непредсказуемым процессом.

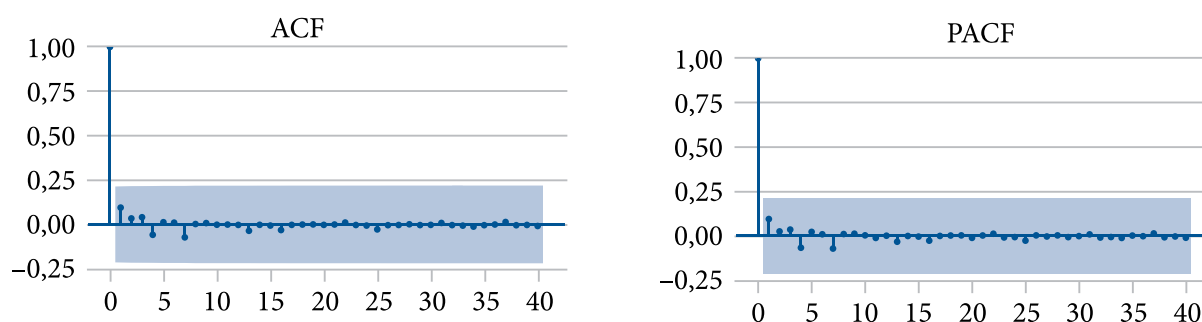


Рис. 5. Автокорреляционные функции остатков модели SARIMA

Fig. 5. Autocorrelation functions of the SARIMA model residuals

Таким образом, несмотря на ненормальность распределения остатков, модель удовлетворяет критерию независимости, что является ключевым требованием для корректной работы модели SARIMA.

Дополнительно стоит отметить, что модель имеет низкое значение AIC, равное 189,704, а также положительные результаты по нескольким другим аспектам диагностики: результаты теста Льюнг – Бокса (p-значение равно 0,69) свидетельствуют об отсутствии автокорреляции остатков, а тест на гетероскедастичность ($\text{Prob}(H) < 0,05$) – о стабильности дисперсии.

На основании проведенного анализа можно заключить, что модель SARIMA обеспечивает адекватное прогнозирование временного ряда уровней безработицы, демонстрируя низкие ошибки прогноза и независимость остатков.

Модель VAR. Для анализа взаимосвязей между уровнем безработицы и поисковыми запросами была построена модель векторной авторегрессии (VAR). В процессе моделирования использовались стационарные и стандартизированные временные ряды «уровень безработицы», «праца бай», «работа бай», «белмета», «служба занятости», «биржа труда» и «работа». На основе минимизации информационных критериев оптимальное количество временных лагов было определено как четыре.

Результаты теста Грейнджера показали, что основным фактором, вносящим значимую долю в динамику уровня безработицы, являются временные лаги переменной «работа». Коэффициенты для ее первого, второго и третьего лагов оказались статистически значимыми с высокой пояснительной способностью. Это подтверждает гипотезу H_0 о сильной причинно-следственной связи между определенными активностями на рынке труда, отраженными в поисковых запросах, и уровнем безработицы. Также переменные «работа бай» (на третьем лаге) и другие запросы внесли вклад в динамику при больших временных интервалах, однако их влияние было менее выраженным.

Модель характеризуется достаточно высоким уровнем объяснительной способности. Значение R^2 для уровня безработицы достигает 0,5823 – это указывает на то, что более половины вариации во временном ряде объясняются переменными и их лагами в системе. Наибольшие значения R^2 среди поясняющих переменных были получены для запросов «работа бай» (0,7654) и «работа» (0,6540), что подтверждает их ключевую роль в прогнозировании динамики.

Однако следует отметить, что R^2 в модели VAR должен интерпретироваться с осторожностью, поскольку модель описывает взаимозависимости между переменными, а не устанавливает причинно-следственные связи. Следовательно, высокая объяснительная способность не может служить основанием для выводов о фактической причинности.

Помимо объяснительной силы, модель характеризуется удовлетворительными значениями информационных критериев. AIC составил 4,7471, а BIC равен 0,5413, что демонстрирует высокую степень качества модели относительно сложности системы. Однако, эти значения должны рассматриваться в контексте сравнения с другими моделями, а не как абсолютные индикаторы качества модели.

Диагностика остатков модели показала их близость к белому шуму, что свидетельствует об отсутствии существенных автокорреляций, а также подтверждает корректность выбора модели и количества временных лагов. Гистограммы остатков, QQ-диаграммы и функции автокорреляции также подтвердили нормальное распределение и независимость остатков.

Таким образом, модель VAR продемонстрировала свою эффективность для исследования динамики уровня безработицы во взаимосвязи с интернет-поисковыми запросами. Она подтвердила значимость включения поисковых запросов как предикторов, что может быть полезным для мониторинга рынка труда и прогнозирования его будущих тенденций.

SARIMAX модель. В ходе исследования была проведена работа по построению прогнозностической модели временных рядов SARIMAX, основной целью чего являлось предсказание уровня безработицы с учетом экзогенных факторов и последующий анализ точности и производительности модели.

Для подготовки данных была реализована процедура добавления лагов экзогенных переменных, что позволило учесть временную зависимость целевой переменной от ее предшествующих значений. Лагированные предикторы были синхронизированы по индексам с целевой переменной для сохранения строгого соответствия между объясняющими и зависимыми переменными.

Для оптимизации параметров модели был выполнен поиск по сетке (grid search), включающий перебор различных комбинаций значений параметров p , d , q (сезонных и не сезонных). Оптимальной считалась модель с минимальным значением информационного критерия AIC. Согласно результатам оптимизации, модель с параметрами p , d , $q = 0, 1, 2$, и сезонными параметрами $P, D, Q, s = 1, 1, 1, 12$, продемонстрировала наилучший

результат ($AIC = -118,74$), что свидетельствует о низкой сложности модели и ее хорошей способности описывать данные.

На этапе оценки предсказаний модели с использованием тестовых данных были вычислены основные показатели точности. Значение MAE составило 0,1001, а MSE – 0,0172, что указывает на низкий уровень средних абсолютных и квадратичных отклонений прогнозов от реальных значений. Корень из среднеквадратичной ошибки (RMSE) оказался равным 0,1313, что указывает на достаточно низкий уровень ошибки модели и согласуется с ее приемлемой предсказательной способностью. Значение MAPE составило 1,91 %, то есть модель демонстрирует в среднем менее 2 % отклонения от реальных значений в своих прогнозах. Коэффициент детерминации R^2 , равный 0,6364, также указывает на удовлетворительное качество модели, учитывая поведение временного ряда и сложность прогноза.

Модель SARIMAX также была оценена с точки зрения информационных критериев и правдоподобия. Логарифм правдоподобия модели составил 82,37, а низкие значения AIC ($-118,74$) и BIC ($-67,35$) свидетельствуют о предпочтительности этой модели среди рассмотренных вариантов.

Заключительный этап включал визуализацию прогностической способности модели. График сравнения реальных и прогнозных значений, дополненный 95 % доверительным интервалом (рис. 6), продемонстрировал, что прогнозы модели SARIMAX точно следуют динамике реальных данных, что подтверждает ее высокую производительность и потенциал к интеграции в прикладные задачи.

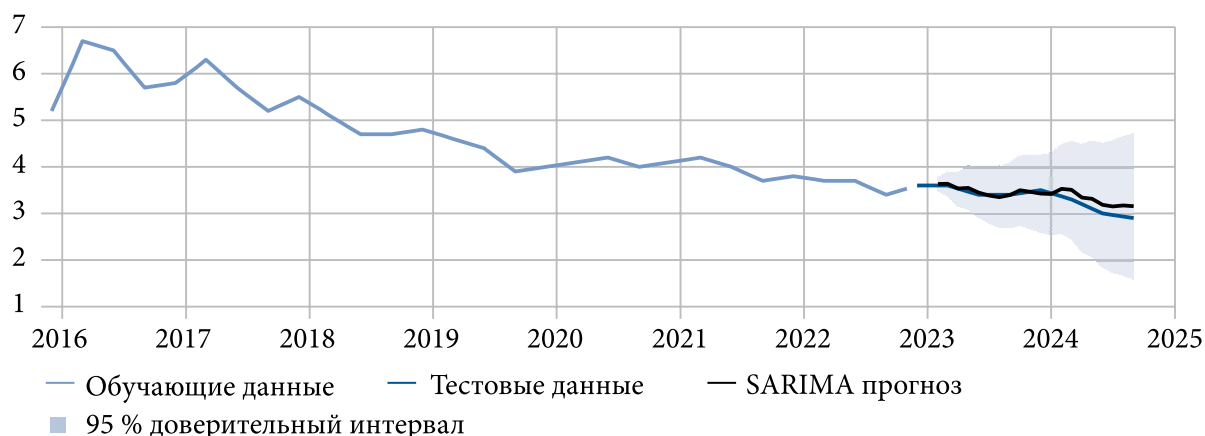


Рис. 6. Прогноз уровня безработицы (модель SARIMAX)

Fig. 6. Unemployment rate forecast (SARIMAX model)

Таким образом, представленная модель удовлетворительно решает поставленную задачу прогнозирования и демонстрирует широкие перспективы применения для анализа временных рядов с экзогенными факторами.

Результаты проведенного сравнительного анализа трех моделей временных рядов (SARIMA, SARIMAX и VAR) представлены в табл. 3.

Результаты анализа позволили сделать вывод, что модель SARIMAX демонстрирует наилучшие характеристики среди представленных методов.

В частности, для модели SARIMAX зафиксированы минимальные значения среднеквадратичной ошибки ($MSE = 0,0172$) и средней абсолютной ошибки ($MAE = 0,1001$), а также корня из среднеквадратичной ошибки ($RMSE = 0,1313$). Эти показатели свидетельствуют об относительно низком уровне прогнозных ошибок по сравнению с другими.

Таблица 3. Критерии эффективности моделей прогнозирования безработицы

Table 3. Performance criteria of the unemployment forecasting models

Модель	MAE	MSE	RMSE	MAPE (%)	AIC	BIC
SARIMA	0,1645	0,0614	0,2477	5,31	-191,92	-174,99
SARIMAX	0,1001	0,0172	0,1313	1,91	-118,74	-67,35
VAR	3,2354	10,6588	3,2648	96,77	-4,7471	0,5413

Кроме того, средняя абсолютная процентная ошибка (MAPE) составила всего 1,91 %, что отражает минимальный уровень отклонения прогнозируемых значений от фактических. Низкие значения информационных критериев AIC (-118,74 и -67,35) дополнительно подтверждают превосходство SARIMAX.

Модель SARIMA также продемонстрировала приемлемые результаты, однако ее значения ошибок были выше по сравнению с SARIMAX. Средняя абсолютная процентная ошибка составила 5,31 %, что почти в три раза превышает аналогичный показатель для модели SARIMAX. Несмотря на хорошие результаты, эта модель уступает SARIMAX из-за отсутствия экзогенных переменных. Ее точность заметно ниже, что ограничивает ее способности в учете дополнительных факторов, влияющих на уровень безработицы.

Модель VAR оказалась менее точной по всем оценочным метрикам. Значения MAE (3,2354), MSE (10,6588) и RMSE (3,2648) указывают на значительное отклонение предсказанных значений от наблюдаемых. Наиболее показательным является крайне высокий уровень MAPE (96,77 %), что свидетельствует о недостаточной способности модели воспроизводить реальные данные. Значения информационных критериев AIC и BIC (-4,7471 и 0,5413 соответственно) для модели указывают на то, что она более сложна и менее предпочтительна по сравнению с SARIMA и SARIMAX. Высокие значения BIC, в частности, подтверждают, что модель VAR неэффективна для данной задачи прогнозирования, так как более простые модели способны лучше объяснять зависимость между переменными.

Заключение

В рамках исследования произведен анализ моделей прогнозирования уровня безработицы на основании показателей прогностической точности и информационных критериев. Так, можно сделать вывод, что модель SARIMAX обладает наивысшей прогностической точностью в рассматриваемой задаче. Это подтверждает гипотезу о том, что учет экзогенных переменных, таких как динамика поисковых запросов, позволяет значительно улучшить качество прогноза уровня безработицы, что делает SARIMAX наиболее подходящим методом для этого типа анализа.

Таким образом, полученные результаты имеют существенное практическое значение для прогнозирования динамики рынка труда в Республике Беларусь. Установлено, что поисковые запросы способны служить надежными опережающими индикаторами изменений уровня безработицы. При этом наиболее точные прогнозы могут быть получены при использовании комбинации различных поисковых терминов с учетом их сезонных колебаний.

Исследование открывает перспективы для разработки более совершенных моделей прогнозирования безработицы с использованием данных поисковых запросов. Дальнейшие исследования могут быть направлены на построение комплексных прогностических моделей, учитывающих более широкий спектр факторов и их взаимное влияние.

Источники

- Броницкий Г. Т., Вакуленко Е. С. (2022). Прогнозирование миграции из России в Германию с использованием Google-трендов // Демографическое обозрение. Т. 9, № 3. С. 75–92. <https://doi.org/10.17323/demreview.v9i3.16471>.
- Ванкевич Е. В., Зайцева О. В. (2020). Онлайн-порталы вакансий: перспективы использования в практике регулирования рынка труда Республики Беларусь // Цифровая трансформация. № 2 (11). С. 33–42. <https://doi.org/10.38086/2522-9613-2020-2-33-42>.
- Ванкевич Е. В., Зайцева О. В. (2023). Развитие рынка труда Республики Беларусь на современном этапе: тенденции и приоритеты // Общество и экономика. № 5. С. 75–100. <https://doi.org/10.31857/S020736760025548-1>.
- Зубарев А. В., Голованова Е. А. (2021). Использование Google Trends для прогнозирования: обзор и применение для прогнозирования розничных продаж: препринт. Москва: РАНХиГС. 30 с.
- Калабихина И. Е., Банин Е. П., Архангельский В. Н., Абдуселимова И. А., Клименко Г. А., Колотуша А. В., ... Шамсутдинова В. Ш. (2020). Краткосрочное прогнозирование демографических тенденций на основе данных Google Trends // Прикладная информатика. Т. 15, № 6. С. 91–118. <https://doi.org/10.37791/2687-0649-2020-15-6-91-118>.
- Королецкая Е. С. (2021). Прогнозирование молодежной безработицы и меры по ее регулированию // Бюллетень Министерства труда и социальной защиты Республики Беларусь. № 8. С. 48–53.
- Куровский Г. С. (2019). Использование текстовой информации для прогнозирования в макроэкономике // Вестник Московского университета. Серия 6. Экономика. № 6. С. 39–58. <https://doi.org/10.38050/01300105201965>.
- Петрова Д. А., Трунин П. В. (2020). Прогнозирование основных макроэкономических показателей с использованием поисковых интернет запросов. Москва: РАНХиГС. 71 с.
- Слаут А. И., Цехан О. Б. (2019). Построение ARIMA-моделей для анализа и прогнозирования уровня безработицы по Республике Беларусь // Тенденции экономического развития в XXI веке: материалы Международной научной конференции (г. Минск, 28 февраля 2019 г.). Минск: Белорусский государственный университет. С. 283–286.
- Юревич М. А., Ахмадеев Д. Р. (2021). Возможности прогнозирования уровня безработицы на основе анализа статистики запросов (в поисковых системах) // Terra Economicus. Т. 19, № 3. С. 53–64. <https://doi.org/10.18522/2073-6606-2021-19-3-53-64>.
- Adu W. K., Appiahene P., Afrifa S. (2023). VAR, ARIMAX and ARIMA models for nowcasting unemployment rate in Ghana using Google Trends. *Journal of Electrical Systems and Information Technology*, vol. 10, 12. <https://doi.org/10.1186/s43067-023-00078-1>.
- Askatas N., Zimmermann K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, vol. 55, issue 2, pp. 107–120. <https://doi.org/10.3790/aeq.55.2.107>.
- Borup D., Christian E., Schütte M. (2020). In search of a job: Forecasting employment growth using Google Trends. *Journal of Business & Economic Statistics*, vol. 40, issue 1, pp. 186–200. <https://doi.org/10.1080/07350015.2020.1791133>.
- Choi H., Varian H. (2009). *Predicting the present with Google Trends. (Technical report)*. Mountain View: Google. 23 p. <https://doi.org/10.2139/ssrn.1659302>.
- Claveria O. (2019). Forecasting the unemployment rate using the degree of agreement in consumer unemployment expectations. *Journal for Labour Market Research*, vol. 53, 3. <https://doi.org/10.1186/s12651-019-0253-4>.
- Costa E. A., Silva M. E., Galvão A. B. (2024). Real-time nowcasting the monthly unemployment rates with daily Google Trends data. *Socio-Economic Planning Sciences*, vol. 95, 101963.
- D'Amuri F., Marcucci J. (2010). 'Google It!' Forecasting the US unemployment rate with a Google job search index (FEEM Working Paper no. 31.2010). Milan: Fondazione Eni Enrico Mattei. 58 p. <https://doi.org/10.2139/ssrn.1594132>.
- D'Amuri F., Marcucci J. (2017). The predictive power of Google searches in forecasting unemployment. *International Journal of Forecasting*, vol. 33, issue 4, pp. 801–816. <https://doi.org/10.1016/j.ijforecast.2017.03.004>.

Dilmaghani M. (2019). The racial 'digital divide' in the predictive power of Google trends data for forecasting the unemployment rate. *Journal of Economic and Social Measurement*, vol. 43, no. 1. <https://doi.org/10.3233/JEM-180458>.

Ettredge M., Gerdes J., Karuga G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, vol. 48, issue 11, pp. 87–92. <https://doi.org/10.1145/1096000.1096010>.

Fondeur Y., Karamé F. (2013). Can Google data help predict French youth unemployment? *Economic Modelling*, vol. 30, pp. 117–125. <https://doi.org/10.1016/j.econmod.2012.07.017>.

González-Fernández M., González-Velasco C. (2018). Can Google econometrics predict unemployment? Evidence from Spain. *Economics Letters*, vol. 170, pp. 42–45. <https://doi.org/10.1016/j.econlet.2018.05.031>.

Mulero R., Garcia-Hiernaux A. (2023). Forecasting unemployment with Google Trends: Age, gender and digital divide. *Empirical Economics*, vol. 65, pp. 587–605. <https://doi.org/10.1007/s00181-022-02347-w>.

Naccarato A., Falorsi S., Loriga S., Pierini A. (2018). Combining official and Google Trends data to forecast the Italian youth unemployment rate. *Technological Forecasting and Social Change*, vol. 130, pp. 114–122. <https://doi.org/10.1016/j.techfore.2017.11.022>.

Nagao S., Takeda F., Tanaka R. (2019). Nowcasting of the U.S. unemployment rate using Google Trends. *Finance Research Letters*, vol. 30, pp. 103–109. <https://doi.org/10.1016/j.frl.2019.04.005>.

Simionescu M. (2020). Improving unemployment rate forecasts at regional level in Romania using Google Trends. *Technological Forecasting & Social Change*, vol. 155, 120026. <https://doi.org/10.1016/j.techfore.2020.120026>.

Simionescu M. (2024). Is Google Trends useful in nowcasting unemployment rate during the pandemic at regional and national level in Romania? *Romanian Statistical Review*, no. 1, pp. 65–84.

Simionescu M., Cifuentes-Faura J. (2022). Forecasting national and regional youth unemployment in Spain using Google Trends. *Social Indicators Research*, vol. 164, pp. 1187–1216. <https://doi.org/10.1007/s11205-022-02984-9>.

Vankevich A., Kalinouskaya I. (2021). Better understanding of the labour market using Big Data. *Ekonomia i Prawo. Economics and Law*, vol. 20, issue 3, pp. 677–692. <https://doi.org/10.12775/EiP.2021.040>.

Vankevich A., Kalinouskaya I., Zaitsava V., Alyakseeva A. (2023). Theoretical and methodological approaches to analysis and forecasting of the labour market within the framework of sustainable development. *Proc. 2nd Int. Conf. on Environmental Sustainability Management and Green Technologies (E3S Web of Conf., vol. 451, 05019)*. Novosibirsk: Novosibirsk State Technical University. 11 p. <https://doi.org/10.1051/e3sconf/202345105019>.

Информация об авторе

Зайцева Ольга Вячеславовна – кандидат экономических наук, доцент, доцент кафедры экономики и электронного бизнеса. Витебский государственный технологический университет, г. Витебск, Республика Беларусь. E-mail: olgazaitseva@gmail.com

...

References

Bronitskiy G. T., Vakulenko E. S. (2022). Using Google Trends for external migration prediction. *Demograficheskoe obozrenie = Demographic Review*, vol. 9, no. 3, pp. 75–92. <https://doi.org/10.17323/demreview.v9i3.16471>. (In Russ.)

Vankevich E. V., Zaitseva V. V. (2020). Online job portals: The future of labor market regulation in the Republic of Belarus. *Tsifrovaya transformatsiya = Digital Transformation*, no. 2 (11), pp. 33–42. <https://doi.org/10.38086/2522-9613-2020-2-33-42>. (In Russ.)

Vankevich A. V., Zaitseva V. V. (2023). Development of the labour market in the Republic of Belarus: Current trends and priorities. *Obshchestvo i ekonomika = Society and Economics*, issue 5, pp. 75–100. <https://doi.org/10.31857/S020736760025548-1>. (In Russ.)

Zubarev A. V., Golovanova E. A. (2021). *Using Google Trends for forecasting: Overview and application for retail sales forecasting*. Moscow: RANEPA. 30 p. (In Russ.)

Kalabikhina I. E., Banin E. P., Arkhangelskiy V. N., Abduselimova I. A., Klimenko G. A., Kolotusha A. V., ... Shamsutdinova V. Sh. (2020). Short-term forecasting of demographic trends based on Google Trends data. *Prikladnaya informatika = Journal of Applied Informatics*, vol. 15, no. 6, pp. 91–118. <https://doi.org/10.37791/2687-0649-2020-15-6-91-118>. (In Russ.)

Koroletskaya E. S. (2021). Forecasting youth unemployment and measures to regulate it. *Byulleten Ministerstva truda i sotsialnoy zashchity Respubliki Belarus = Bulletin of the Ministry of Labor and Social Protection of the Republic of Belarus*, no. 8, pp. 48–53. (In Russ.)

Kurovskiy G. S. (2019). Using textual information to predict in macroeconomics. *Vestnik Moskovskogo universiteta. Seriya 6. Ekonomika = Moscow University Economic Bulletin*, no. 6, pp. 39–58. <https://doi.org/10.38050/01300105201965>. (In Russ.)

Petrova D. A., Trunin P. V. (2020). *Prediction of key macroeconomic indicators using Internet search queries*. Moscow: RANEPA. 71 p. (In Russ.)

Slaut A. I., Tsekhan O. B. (2019). Construction of ARIMA-models for analysing and forecasting the unemployment rate in the Republic of Belarus. *Proc. Int. Sci. Conf. "Trends in economic development in the 21st century"* (pp. 283–286). Minsk: Belarusian State University. (In Russ.)

Yurevich M. A., Akhmadeev D. R. (2021). Predicting the unemployment rate: Analyzing statistics on search engine query. *Terra Economicus*, vol. 19, no. 3, pp. 53–64. <https://doi.org/10.18522/2073-6606-2021-19-3-53-64>. (In Russ.)

Adu W. K., Appiahene P., Afrifa S. (2023). VAR, ARIMAX and ARIMA models for nowcasting unemployment rate in Ghana using Google Trends. *Journal of Electrical Systems and Information Technology*, vol. 10, 12. <https://doi.org/10.1186/s43067-023-00078-1>.

Askitas N., Zimmermann K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, vol. 55, issue 2, pp. 107–120. <https://doi.org/10.3790/aeq.55.2.107>.

Borup D., Christian E., Schütte M. (2020). In search of a job: Forecasting employment growth using Google Trends. *Journal of Business & Economic Statistics*, vol. 40, issue 1, pp. 186–200. <https://doi.org/10.1080/07350015.2020.1791133>.

Choi H., Varian H. (2009). *Predicting the present with Google Trends. (Technical report)*. Mountain View: Google. 23 p. <https://doi.org/10.2139/ssrn.1659302>.

Claveria O. (2019). Forecasting the unemployment rate using the degree of agreement in consumer unemployment expectations. *Journal for Labour Market Research*, vol. 53, 3. <https://doi.org/10.1186/s12651-019-0253-4>.

Costa E. A., Silva M. E., Galvão A. B. (2024). Real-time nowcasting the monthly unemployment rates with daily Google Trends data. *Socio-Economic Planning Sciences*, vol. 95, 101963.

D'Amuri F., Marcucci J. (2010). 'Google It!' Forecasting the US unemployment rate with a Google job search index (FEEM Working Paper no. 31.2010). Milan: Fondazione Eni Enrico Mattei. 58 p. <https://doi.org/10.2139/ssrn.1594132>.

D'Amuri F., Marcucci J. (2017). The predictive power of Google searches in forecasting unemployment. *International Journal of Forecasting*, vol. 33, issue 4, pp. 801–816. <https://doi.org/10.1016/j.ijforecast.2017.03.004>.

Dilmaghani M. (2019). The racial 'digital divide' in the predictive power of Google trends data for forecasting the unemployment rate. *Journal of Economic and Social Measurement*, vol. 43, no. 1. <https://doi.org/10.3233/JEM-180458>.

Ettredge M., Gerdes J., Karuga G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, vol. 48, issue 11, pp. 87–92. <https://doi.org/10.1145/1096000.1096010>.

Fondeur Y., Karamé F. (2013). Can Google data help predict French youth unemployment? *Economic Modelling*, vol. 30, pp. 117–125. <https://doi.org/10.1016/j.econmod.2012.07.017>.

González-Fernández M., González-Velasco C. (2018). Can Google econometrics predict unemployment? Evidence from Spain. *Economics Letters*, vol. 170, pp. 42–45. <https://doi.org/10.1016/j.econlet.2018.05.031>.

Mulero R., Garcia-Hiernaux A. (2023). Forecasting unemployment with Google Trends: Age, gender and digital divide. *Empirical Economics*, vol. 65, pp. 587–605. <https://doi.org/10.1007/s00181-022-02347-w>.

Naccarato A., Falorsi S., Loriga S., Pierini A. (2018). Combining official and Google Trends data to forecast the Italian youth unemployment rate. *Technological Forecasting and Social Change*, vol. 130, pp. 114–122. <https://doi.org/10.1016/j.techfore.2017.11.022>.

Nagao S., Takeda F., Tanaka R. (2019). Nowcasting of the U.S. unemployment rate using Google Trends. *Finance Research Letters*, vol. 30, pp. 103–109. <https://doi.org/10.1016/j.frl.2019.04.005>.

Simionescu M. (2020). Improving unemployment rate forecasts at regional level in Romania using Google Trends. *Technological Forecasting & Social Change*, vol. 155, 120026. <https://doi.org/10.1016/j.techfore.2020.120026>.

Simionescu M. (2024). Is Google Trends useful in nowcasting unemployment rate during the pandemic at regional and national level in Romania? *Romanian Statistical Review*, no. 1, pp. 65–84.

Simionescu M., Cifuentes-Faura J. (2022). Forecasting national and regional youth unemployment in Spain using Google Trends. *Social Indicators Research*, vol. 164, pp. 1187–1216. <https://doi.org/10.1007/s11205-022-02984-9>.

Vankevich A., Kalinouskaya I. (2021). Better understanding of the labour market using Big Data. *Ekonomia i Prawo. Economics and Law*, vol. 20, issue 3, pp. 677–692. <https://doi.org/10.12775/EiP.2021.040>.

Vankevich A., Kalinouskaya I., Zaitsava V., Alyakseeva A. (2023). Theoretical and methodological approaches to analysis and forecasting of the labour market within the framework of sustainable development. *Proc. 2nd Int. Conf. on Environmental Sustainability Management and Green Technologies (E3S Web of Conf., vol. 451, 05019)*. Novosibirsk: Novosibirsk State Technical University. 11 p. <https://doi.org/10.1051/e3sconf/202345105019>.

Information about the author

Olga V. Zaitseva, Cand. Sc. (Econ.), Associate Prof., Associate Prof. of Economics and e-Business Dept. Vitebsk State Technological University, Vitebsk, the Republic of Belarus. E-mail: olgazaitseva@gmail.com

© Зайцева О. В., 2025