ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ИССЛЕДОВАНИЯХ МЕДИА И КОММУНИКАЦИИ

Метель Юрий Андреевич, кандидат экономических наук, руководитель направления экономического отдела Отделения по Ставропольскому краю Южного главного управления Центрального банка Российской Федерации, г. Ставрополь, Россия; e-mail: akademik.st.2018@mail.ru, ORCID 0009-0001-5334-0020

Куницына Наталья Николаевна, доктор экономических наук, профессор, заведующая кафедрой финансов и кредита, ФГАОУ ВО «Северо-Кавказский федеральный университет», г. Ставрополь, Россия; e-mail: nkunitcyna@ncfu.ru, ORCID 0000-0001-9336-8100

ПОСТРОЕНИЕ ВЫСОКОЧАСТОТНОГО МЕДИАИНДИКАТОРА РЕГИОНАЛЬНОЙ ЭКОНОМИЧЕСКОЙ АКТИВНОСТИ НА ОСНОВЕ АНАЛИЗА ПУБЛИКАЦИЙ СМИ С ПРИМЕНЕНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ[1]

Наличие временны́х лагов в публикациях официальных статистических материалов обуславливает повышенный интерес к альтернативным источникам информации. Одним из них является массив текстовых данных СМИ, обрабатываемый с помощью больших языковых моделей (LLM). Целью исследования стала разработка показателя экономической активности субъекта Российской Федерации на основе анализа новостных публикаций региональных СМИ с использованием LLM. Материалом для изучения стали тексты интернет-изданий за 2021–2025 гг. Методика исследования включает несколько этапов: формирование базы новостных публикаций, тематическую классификацию текстов по укрупненным секторам экономики с использованием модели zero-shot классификации, определение экономической тональности новостей с помощью большой языковой модели GigaChat, построение агрегированного медиапоказателя. В результате сформирован интегральный медиапоказатель экономической динамики IBVEDMedia, сопоставление которого с официальным индексом выпуска товаров и услуг по базовым видам экономической деятельности продемонстрировало наличие статистически значимой связи между ними. Следовательно, новостной поток региональных СМИ содержит существенную информацию о текущем состоянии региональной экономики. Предложенный медиапоказатель может рассматриваться как композитный индикатор экономической динамики, дополняющий официальную статистику и расширяющий инструментарий анализа экономической конъюнктуры. Валидация его оперативных свойств является предметом дальнейших исследований.

***Ключевые слова**: новостные данные региональных СМИ, большие языковые модели (LLM), обработка естественного языка (NLP), оперативные*

экономические индикаторы, индекс выпуска товаров и услуг по базовым видам экономической деятельности (ИБВЭД).

Как цитировать: *Метель Ю. А., Куницына Н. Н.* Построение высокочастотного медиаиндикатора региональной экономической активности на основе анализа публикаций СМИ с применением больших языковых моделей // Вестник Московского университета. Серия 10. Журналистика. 2026. № 3. С. 3–28. DOI: 10.55959/msu.vestnik.journ.3.2026.328

Yuri A. Metel, PhD in Economics, Head of the Economic Department Division, Stavropol Territory Branch of the Southern Main Directorate of the Central Bank of the Russian Federation, Stavropol, Russia; e-mail: akademik.st.2018@mail.ru, ORCID 0009-0001-5334-0020

Natalia N. Kunitsyna, Doctor of Economics, Professor, Head of the Department of Finance and Credit, North Caucasus Federal University, Stavropol, Russia; e-mail: nkunitcyna@ncfu.ru, ORCID 0000-0001-9336-8100

BUILDING A HIGH-FREQUENCY MEDIA INDICATOR OF REGIONAL ECONOMIC ACTIVITY BASED ON ANALYSIS OF MEDIA PUBLICATIONS USING LARGE LANGUAGE MODELS

The inherent publication lags in official statistical releases have precipitated a growing scholarly and practical interest in alternative data sources for economic assessment. Among these, the vast corpus of textual data from media outlets, processed through Large Language Models (LLMs), presents a significant opportunity. This study aims to develop a media-based indicator of economic activity for a constituent entity of the Russian Federation by analyzing regional news publications with LLMs. The empirical foundation of this research comprises a corpus of texts from online media spanning the period 2021–2025. The methodological framework consists of several sequential stages: (i) the compilation and preprocessing of a regional news publication database; (ii) the thematic classification of texts according to aggregated economic sectors employing a zero-shot classification paradigm; (iii) the extraction of economic sentiment from the news articles using the GigaChat large language model; and (iv) the construction of an aggregated media indicator. The resulting output is an integral media indicator of economic dynamics, designated as the IBVEDMedia index. A comparative analysis with the official Index of Output of Goods and Services for Basic Economic Activities reveals a statistically significant correlation between the two time series. This finding suggests that the news flow from regional media sources encapsulates substantial information pertinent to the current state of the regional economy. The proposed media indicator can be considered a composite economic sentiment indicator, serving as a valuable complement to traditional official statistics. The validation of its real-time operational properties is a subject for future research.

***Key words**: regional media data, large language models (LLMs), natural language processing (NLP), high-frequency economic indicators, index of output of goods and services for basic economic activities.*

To cite this article: Metel Yu. A., Kunitsyna N. N. (2026) Postroenie vysokochastotnogo media-indikatora regional'noy ekonomicheskoy aktivnosti na osnove analiza publikatsiy SMI s primeneniem bol'shikh yazykovykh modeley [Building a high-frequency media indicator of regional economic activity based on analysis of media publications using large language models]. *Vestnik Moskovskogo Universiteta. Seriya 10. Zhurnalistika 3*: 3–28. DOI: 10.55959/msu.vestnik.journ.3.2026.328

DOI: 10.55959/msu.vestnik.journ.3.2026.328

Введение

Современная денежно-кредитная политика во многом зависит от скорости и качества информационного обеспечения. Классические макропоказатели официальной статистики, такие как индекс промышленного производства, динамика ВВП и ВРП, оборот розничной торговли, остаются «золотым стандартом» для ретроспективного анализа. Однако их ключевой недостаток в контексте принятия оперативных решений – значительное запаздывание публикации в статистических сборниках, зачастую с лагом от нескольких месяцев до года (в случае с ВВП/ВРП)[2], что вынуждает Банк России, Министерство экономического развития РФ, Министерство финансов РФ и бизнес принимать решения на основе ретроспективных данных, не отражающих текущие шоки и краткосрочные колебания конъюнктуры рынка.

Эта проблема особенно остро стоит на региональном уровне. Динамика развития субъектов Российской Федерации отличается неоднородностью и зависимостью от локальных факторов: агроклиматических условий, отраслевой структуры, реализации конкретных инвестиционных проектов. Федеральная статистика, обеспечивая необходимую унификацию и сопоставимость, не отражает быстрые точечные изменения в экономической конъюнктуре конкретного региона.

Возникает насущная потребность в разработке и внедрении оперативных индикаторов экономической активности, которые могли бы служить надежным источником сигналов «в режиме, близком к реальному времени», компенсируя лаги официальной статистики. Важным информационным ресурсом становятся потоки неструктурированных текстовых данных, в том числе публи-

кации в СМИ. Новостные ленты, в отличие от статистических сводок, реагируют на события практически мгновенно, аккумулируя информацию о запуске новых производств или модернизации старых, заключении контрактов, колебаниях урожайности, проблемах на транспортных узлах и изменении потребительской активности. Мы выдвинули гипотезу: совокупное информационное поле, формируемое региональными СМИ, несет в себе скрытый, но структурируемый сигнал об изменении экономической динамики отдельного региона. Проверку этой гипотезы мы и осуществили в исследовании.

Обзор литературы

Мониторинг состояния экономики в режиме, близком к реальному времени, на основе оперативного косвенного расчета официальных показателей является сутью методологии наукастинга (*Nowcasting*), которая служит «мостом» между ретроспективной статистикой и потребностями в актуальных данных для принятия решений. В настоящем исследовании мы не стремились к реализации полной процедуры наукастинга. Задача заключалась в разработке высокочастотного медиаиндикатора, который может быть использован в моделях наукастинга. Тем не менее, поскольку данный индикатор ориентирован на применение именно в этой парадигме, а сама методология наукастинга определяет требования к оперативности и качеству входных данных, считаем необходимым рассмотреть существующие подходы к наукастингу с целью позиционирования нашей разработки в рамках сложившейся исследовательской практики.

В экономической науке наукастинг представляет собой оценку текущего состояния экономики, как правило, на основе поступающих высокочастотных данных, позволяющих частично компенсировать информационный разрыв (Ghysels, 2016).

В зарубежной литературе наукастинг получил широкое распространение в 2000-х гг. Изначально в его основе лежали динамические факторные модели, модели со смешанной частотностью (*Mixed-Data Sampling, MIDAS* и др.), векторные авторегрессии (*VAR*), позволяющие использовать данные с разной периодичностью при прогнозировании ВРП, инфляции и уровня безработицы (Clements, Galvão, 2009; Ghysels, 2016; Dahlhaus, Guénette, Vasishtha, 2017; Mikosch, Solanko, 2019).

Впоследствии инструментарий наукастинга пополнился панельными моделями, которые также учитывают и кросс-секционные зависимости исследуемых объектов. В частности, ученые при наукастинге регионального объема выпуска применяли *VAR* (*Vector Autoregression*), *BVAR* (*Bayesian Vector Autoregression*) и *MF-PVAR* (*Mixed-Frequency Panel Vector Autoregression*) с панельной структурой данных. Такой подход обеспечил более качественный прогноз по сравнению с классическими эконометрическими методами, особенно на коротких временных рядах (Koop, Korobilis, Pettenuzzo, 2019; Acedański, 2024; Fosten, Greenaway-McGrevy, 2022).

В последнее время для оперативной оценки макроэкономических показателей экономисты стали использовать и методы классического машинного обучения (*Machine Learning, ML*). Исследователи применили для наукастинга ВВП Новой Зеландии методы градиентного бустинга, *Lasso*- и *Ridge*-регрессии, а также метод опорных векторов (Richardson, van Florenstein Mulder, Vehbi, 2021). В результате авторы выявили, что *ML*-алгоритмы способны значимо повысить точность наукастинга по сравнению с классическими эконометрическими моделями.

Аналогичный вывод на данных о ВВП Китая сделали китайские исследователи (Zhang, Ni, Xu, 2023). Они провели псевдовыборочный эксперимент на основе 89 показателей и установили, что классический *ML*, в частности *Ridge*-регрессия с регуляризацией, превосходят *MIDAS* по точности оперативной оценки, особенно в периоды высокой волатильности.

В российской литературе исследования наукастинга также получили развитие, однако они сосредоточены преимущественно на анализе индикаторов общенационального уровня. Значительная часть работ посвящена оценке текущей динамики ВВП России с использованием динамических факторных моделей, *MIDAS*-подходов и *BVAR*-моделей, *DSGE*-моделей (*Dynamic Stochastic General Equilibrium*) (Поршаков, Пономаренко, Синяков, 2016; Станкевич, 2023; Eliseev, 2025; Зубарев, Рыбак, 2021; Фокин, 2023). В. Бойко с соавторами (Boyko, Kislyak, Nikitin, Oborin, 2020) сравнивают методологию Росстата и методы темпорального дезагрегирования для получения более оперативных оценок регионального выпуска. Авторы показывают, что использование дополнительных индикаторов позволяет повысить точность квартальных оценок ВРП.

В исследованиях ученые используют и альтернативные источники данных, включая поисковые запросы, что позволяет повы-

сить оперативность оценок и чувствительность моделей к изменению ожиданий экономических агентов (Лазарян, Герман, 2018; Shulyak, 2022; Макеева, Станкевич, 2022).

Как можно заметить, большинство существующих исследований – и мировых, и российских – опирается на количественные статистические показатели. При этом вопрос использования текстовых данных и медиаинформации для оперативной оценки региональной экономической активности остается относительно малоизученным. Вместе с тем область обработки текстовых данных и естественного языка (*NLP*) в условиях развития технологий вызывает повышенный интерес и позволяет расширить инструментарий экономического анализа в условиях высокой неопределенности (Гареев, Полбин, 2022).

Так, К. Яковлева (Yakovleva, 2018) для наукастинга динамики экономической активности в России предложила методику расчета высокочастотного индикатора, использовав более 59 000 статей крупных федеральных СМИ. Все тексты прошли предварительную обработку – токенизицию и лемматизацию. Семантическое моделирование осуществлено с использованием метода опорных векторов. Тренировочная выборка размечена автором вручную.

Аналогичный подход при расчете новостного индекса применяет в практической деятельности и Банк России[3]. Методика регулятора также основана на определении частотности вхождения слов в тексты новостей. На первом этапе используется модель, автоматически выделяющая темы, из которых по ключевым словам определяется одна, максимально покрывающая тематику. Затем для каждой новости на основе обученной модели выявляется ее релевантность и тональность путем определения баланса слов с положительной и отрицательной коннотацией. Для этого применяются методы латентного распределения Дирихле и модифицированный стохастический алгоритм, сглаживающий неравномерное распределение количества новостей по источникам.

А. Г. Евстигнеева и Д. А. Карпов[4] исследовали влияние новостей на формирование инфляционных ожиданий населения. После токенизации и лемматизации слов в текстах новостей авторы применили численные метрики: *TF-IDF* (*Term Frequency – Inverse Document Frequency*) и *Count Vectorizer*. Тематическая классификация проведена на основе латентного распределения Дирихле (*Latent Dirichlet allocation, LDA*), метода опорных векторов (*Support Vector Machine, SVM*), случайного леса (*Random Forest*) и градиент-

ного бустинга (*eXtreme Gradient Boosting, XGBoost*). Для обучения авторы сформировали выборку из 100 000 новостей по 25 источникам, представленным в равных долях. При этом для выделения негативных новостей специалисты также применили *LDA*: этот алгоритм требует на вход список экспертно выбранных якорных токенов, описывающих тему.

Достоинствами рассмотренных подходов являются простота реализации на практике и «мягкие» требования к вычислительным мощностям компьютера. Главным же недостатком – то, что подобные модели основаны на подсчете лишь частоты упоминаний слов в коллекциях текстов. При этом совершенно не учитывается контекст, который может диаметрально изменить тональность текста.

Ограничения частотных и тематических моделей обусловили интерес к методам, способным учитывать контекст и семантические связи в тексте. Кроме того, применение описанных выше алгоритмов, как правило, требует формирования обучающих выборок и ручной разметки текстов, что связано с существенными трудовыми затратами и снижает масштабируемость исследования. В этой связи внимание исследователей привлекают большие языковые модели (*LLM*), позволяющие осуществлять семантическую интерпретацию текстов без предварительной разметки корпуса и отличающиеся практически полным отсутствием ручной настройки классификаторов.

LLM-модели – это класс статистических моделей, предназначенных для изучения структуры человеческого языка. Такие модели достигают цели, оценивая вероятность токена s_i с учетом всех ранее наблюдаемых токенов в документе $p(s_i \mid s_1, \ldots, s_{i-1})$, используя особую архитектуру нейронной сети, известную как трансформеры (Vaswani, Shazeer, Parmar, Uszkoreit et al., 2017).

Одну из таких моделей в своем исследовании применил Л. Биби (Bybee, 2023). Для воссоздания инфляционных ожиданий опрашиваемых экспертов были использованы новостные тексты и *ChatGPT*. Автор сравнил ожидания модели с существующими экономическими ожиданиями экспертов-аналитиков, такими как *Survey of Professional Forecasters* (*SPF*), *American Association of Individual Investors* (*AAII*), *Duke CFO Survey*. Его основной вывод состоит в том, что ожидания *GPT-3.5* коррелируют с опросными данными и воспроизводят те же отклонения от рациональных ожиданий, которые наблюдаются у людей: недостаточно быструю корректиров-

ку прогнозов при появлении новой информации и склонность формировать ожидания на основе прошлых тенденций в области финансовых рынков. Однако данный подход требует больших финансовых затрат, поэтому автор сильно ограничил выборку текстов.

В последние годы в экономической литературе появились работы, в которых *LLM* используются непосредственно для получения оперативных макроэкономических показателей. Так, в препринте Федеральной резервной системы США *ChatMacro: Evaluating Inflation Forecasts of Generative AI*[5] исследуется возможность применения генеративных языковых моделей для прогнозирования инфляции в режиме, близком к реальному времени. Авторы формируют систему промптов, на основе которых большие языковые модели генерируют количественные прогнозы инфляции и сопоставляют полученные оценки с официальными статистическими данными и результатами профессиональных опросов. Результаты такого подхода следует рассматривать как дополнение в официальной статистике, а не ее замену.

Таким образом, обзор литературы позволяет констатировать, что опыт использования *LLM* в целях наукастинга экономических показателей остается ограниченным. Практически отсутствуют исследования, ориентированные на формирование оперативного индикатора региональной экономической активности на основе публикаций местных СМИ, сопоставимого с официальными статистическими показателями выпуска по базовым видам экономической деятельности. Данный исследовательский пробел определяет методологический подход настоящей работы, направленный на использование больших языковых моделей для классификации и агрегирования информации из новостных текстов различных СМИ с целью построения индикатора региональной экономической активности.

Методика

Исследование ориентировано на решение двух блоков задач – классификацию новостей и построение показателя экономической активности региона. С этой целью нами реализованы следующие этапы.

I. Формирование корпуса и фильтрация данных

Обработка новостных текстов включала в себя 3 этапа:

1) Формирование корпуса текстов

Поскольку формирование коллекции новостных статей и ее преобразование в необходимый вид – трудоемкий процесс, для его автоматизации нами применена процедура *Web Scraping*. В *Python* разработана программа с использованием библиотек *Requests* и *BeautifulSoup*. Источниками данных выступили все тексты региональных СМИ Ставропольского края с 2022 по 2025 гг. включительно: газеты «Ставропольская правда» и «Вечерний Ставрополь», еженедельник «Аргументы и факты: Северный Кавказ», информационный портал АТВ-Медиа, информационное агентство «Победа 26», сетевое издание *NewsTracker*. Данные издания являются наиболее популярными в регионе и охватывают широкий круг социально-экономических тем. Общий объем собранного корпуса составил свыше 100 000 новостных публикаций. Отметим, что из материалов были исключены все статьи до 2022 г. и статьи объемом менее 100 слов.

Во избежание дублирования одного и того же события проведено их исключение: новости, имеющие аналогичные/схожие заголовки и опубликованные в пределах 3 дней определялись как дубликаты, в итоговом списке сохранялась первая по дате новость.

2) Тематическая классификация

Под экономической активностью в настоящем исследовании будем понимать Индекс выпуска товаров и услуг по базовым видам экономической деятельности (ИБВЭД). В соответствии с Приказом Министерства экономического развития Российской Федерации от 30.09.2021 г. № 609 «Об утверждении официальной статистической методологии расчета показателя “Индекс выпуска товаров и услуг по базовым видам экономической деятельности”» в ИБВЭД включены следующие отрасли (см. табл. 1). Поскольку индекс охватывает отраслевые сегменты, которые в регионе представлены слабо, в настоящем исследовании мы их укрупнили.

Таблица 1

Базовые отрасли экономической деятельности

Виды экономической деятельности	Группа / код ОКВЭД	Укрупненная категория видов экономической деятельности
Сельское хозяйство	01 без 01.6 и 01.7	Сельское хозяйство
Добыча полезных ископаемых	Раздел B	Промышленное производства
Обрабатывающие производства	Раздел C	
Обеспечение электроэнергией, газом и паром; кондиционирование воздуха	Раздел D	
Водоснабжение, водоотведение; организация сбора и утилизации отходов, деятельность по ликвидации загрязнений	Раздел E	
Строительство	Раздел F	Строительство
Оптовая торговля	46	Торговля
Розничная торговля	47	
Транспортировка и хранение	Раздел H	Транспортировка и хранение

Из коллекции текстов региональных СМИ отобраны те, которые соответствуют базовым отраслям. Классический подход к анализу текстовых данных для экономических задач традиционно предполагает создание размеченных обучающих выборок с последующим применением методов машинного обучения, таких как алгоритмы на основе *TF-IDF*- векторизации в сочетании с линейными моделями или ансамблевыми методами. Однако данная методология сталкивается с ограничением – необходимостью наличия большого массива текстов для обучения моделей, вручную размеченных экспертами по базовым отраслям экономики.

Создание такой выборки для региональных СМИ представляет собой трудоемкую задачу: требует привлечения специалистов, знакомых с экономической тематикой, и сопряжено со значительными временны́ми затратами.

Для отбора публикаций использована модель для мультиязычной классификации *MoritzLaurer/xlm-v-base-mnli-xnli* (платформа *Hugging Face*, 2025), основанная на архитектуре трансформера и позволяющая осуществлять классификацию текстов без предварительного дообучения на специализированной размеченной выборке. Это позволило отказаться от формирования трудоемкой экспертной разметки корпуса уже на первом этапе исследования (см. табл. 2).

Таблица 2

Пример классификации новостей региональных СМИ (фрагмент)

Дата	Текст	Тематика по видам экономической деятельности	Вероятность[6]
12.01.2022	Первую партию медицинского кислорода выпустят этой весной в Невинномысске. Этот факт подтверждает в том числе Федеральный фонд развития промышленности. Он предоставил займ на строительство завода по производству востребованной продукции...	Промышленность	0,96
13.03.2023	На Ставрополье планируется расширение географии пассажирских перелетов из краевого центра. Сейчас из города можно добраться на самолете в Москву, Санкт-Петербург и Сочи. А с недавнего времени в Махачкалу...	Транспортировка и хранение	0,83
13.02.2024	Эксперты «Открытие Авто» оценили продажи автомобилей с пробегом в январе 2024 г. Количество перерегистраций подержанных машин опустилось ниже уровня 400 тысяч и составило 381,7 тысячи сделок. Это на 8% ниже января 2023 года, но все же на 6% больше января 2022 года...	Торговля	0,82
13.02.2024	В январе 2024 года построено более 175 тысяч квадратных метров жилья. Такие данные привел губернатор. Глава региона сообщил, что в 2023 году в эксплуатацию введено почти 1,8 миллиона квадратных метров жилья. Это на 9,2 процента больше, чем в 2022 г...	Строительство	0,91
18.12.2025	Сельхозпредприятия Ставропольского края активно готовятся к весенним полевым работам. В регионе продолжается осенне-зимняя ремонтная кампания. По данным министерства сельского хозяйства края, всего предстоит привести в порядок около 9,1 тысячи единиц сельскохозяйственной и автомобильной техники...	Сельское хозяйство	0,92

С целью обеспечения качества тематической фильтрации установлен порог вероятности принадлежности текста к категории, равный 0,5. Значение порога выбрано на основе тестирования аль-

тернативных значений (0,2–0,9) в векторе обеспечения баланса между полнотой и точностью отбора.

Для проверки качества классификации случайная подвыборка из 200 текстов была вручную проверена экспертами. Совпадение составило около 80%. После тематической фильтрации итоговая выборка расширена до 8 000 публикаций (7,5% от исходного массива). Отраслевая структура текстов СМИ представлена на рисунке 1.

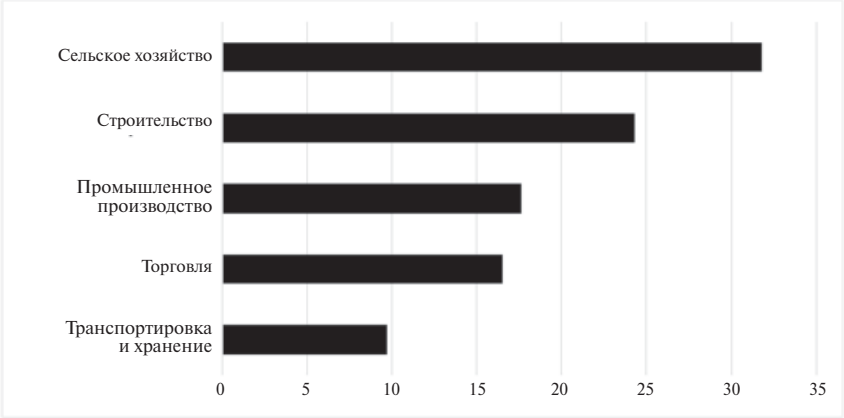


Рисунок 1. Отраслевое распределение текстов, отобранных для анализа, %

II. Оценка экономической тональности публикаций

1) Использование LLM-модели

Для определения тональности отобранных текстов мы использовали *LLM*-модель *GigaChat* с подключением по *API*. Выбор этой модели обусловлен акцентом на работу именно с русским языком, а также возможностью обработки текстов без обучения и разметки.

2) Принципы определения тональности

Экономическая тональность определялась на основе заранее заданного промпта (см. рис. 2), содержащего:

– критерии позитивного, негативного и нейтрального экономического воздействия;

– правила анализа;

– примеры классификации.

Роль: Ты экономический аналитик, специализирующийся на оценке экономической тональности текстов.

Твоя задача - определить, является ли экономическое событие или информация в тексте ПОЗИТИВНОЙ, НЕГАТИВНОЙ или НЕЙТРАЛЬНОЙ с точки зрения экономики.

ЭКОНОМИЧЕСКИЕ КРИТЕРИИ ОЦЕНКИ:
ПОЗИТИВНЫЙ (экономически благоприятный):
- Рост производства, ВВП, экспорта
- Снижение безработицы, инфляции
- Инвестиции, строительство новых объектов
- Улучшение делового климата
- Технологические инновации
- Укрепление курса национальной валюты
- Увеличение доходов населения
НЕГАТИВНЫЙ (экономически неблагоприятный):
- Рост цен, инфляция
- Падение производства, рецессия
- Увеличение безработицы
- Снижение инвестиций
- Санкции, торговые ограничения
- Ослабление курса валюты
- Банкротства компаний
- Снижение доходов населения
НЕЙТРАЛЬНЫЙ:
- Статистические данные без оценки
- Планы и прогнозы без явной окраски
- Техническая информация
- Описательные новости без последствий

ПРАВИЛА АНАЛИЗА:
1. Оценивай ФАКТИЧЕСКОЕ экономическое воздействие
2. Учитывай долгосрочные и краткосрочные эффекты
3. Различай мнение автора и объективные факты
4. Для противоречивых новостей выбирай преобладающую тенденцию

ПРИМЕРЫ:
- "Инфляция выросла с 5 до 10%" → НЕГАТИВНЫЙ
- "Построен новый автомобильный завод" → ПОЗИТИВНЫЙ
- "ЦБ сохранил ключевую ставку" → НЕЙТРАЛЬНЫЙ
- "Экспорт увеличился на 15%" → ПОЗИТИВНЫЙ
- "Безработица достигла 5-летнего максимума" → НЕГАТИВНЫЙ

ТЕКСТ ДЛЯ АНАЛИЗА:
{text}

ФОРМАТ ОТВЕТА:
Только одно слово: ПОЗИТИВНЫЙ, НЕГАТИВНЫЙ или НЕЙТРАЛЬНЫЙ

Рисунок 2. Промпт для определения тональности экономических новостей

Для оценки качества определения тональности моделью *GigaChat* была проведена экспертная проверка. Из итогового корпуса отобранных текстов новостей случайным образом сформирована подвыборка объемом в 100 текстов, вручную размеченная экспертами. Сравнение автоматической классификации с экспертной

показало совпадение на уровне 78%. Анализ матрицы ошибок подтверждает, что основные расхождения приходятся на пары «нейтральный – позитивный» и «нейтральный – негативный». При этом прямые ошибки между позитивными и негативными классами встречаются реже.

Таблица 3

Матрица ошибок определения тональности текстов

Факт/Прогноз	Позитивный	Нейтральный	Негативный
Позитивный	28	4	1
Нейтральный	6	24	4
Негативный	2	5	26

Полученная точность в 78% рассматривается нами как приемлемая для построения балансового индекса, поскольку агрегирование большого числа новостей частично компенсирует индивидуальные ошибки классификации, а преобладание ошибок на границе с нейтральным классом снижает риск систематического смещения индекса в ту или иную сторону.

III. Построение индекса $IBVED_{Media}$

1) Расчет отраслевых индексов тональности

Агрегированный показатель тональности для каждой отрасли i в месяц m рассчитывался как баланс скорректированных позитивных сообщений и негативных текстов:

$$Tonality_{i,m} = \frac{Positive_{i,m} - Negative_{i,m}}{All_news_{i,m}} \quad (1),$$

где $Tonality_{i,m}$ – уровень тональности новостей по соответствующей тематике в i-й отрасли в m-й месяц;

$Positive_{i,m}$ – количество текстов, имеющих позитивную тональность;

$Negative_{i,m}$ – количество текстов, имеющих негативную тональность;

$All_news_{i,m}$ – количество всех новостей по соответствующей тематике.

Значение показателя лежит в интервале [-1; 1], где:

– положительные значения отражают преобладание позитивной экономической информации;

– отрицательные – доминирование негативных сообщений;

– значения, близкие к нулю, указывают на нейтральный информационный фон.

Нормализация на общее число публикаций позволяет учитывать различия в интенсивности информационного потока между отраслями и временными периодами.

2) Агрегирование в сводный индекс

Для формирования интегрального показателя экономического информационного фона рассчитывался агрегированный индекс тональности по всем укрупненным отраслям.

Месячное значение сводной тональности определялось как взвешенная сумма отраслевых индексов:

$$Tonality_{union,\ m} = \sum_{i=1}^{5} w_i * Tonality_{i,m} \quad (2),$$

где w_i – вес *i*-й отрасли в структуре валовой добавленной стоимости.

Веса определяются на основе официальных статистических данных о структуре региональной экономики и отражают вклад соответствующих отраслей в формирование валовой добавленной стоимости региона. Использование отраслевых весов позволяет обеспечить сопоставимость медиаиндикатора со структурой экономики.

Именно на данном этапе индикатор обладает потенциальной оперативностью, так как может быть рассчитан сразу по окончании календарного месяца без дополнительных преобразований. Однако для целей сравнения с официальным ИБВЭД, который публикуется в форме годового прироста, нам потребовалось привести индикатор к сопоставимому виду (этапы 3 и 4).

3) Усреднение за 12 месяцев

Поскольку официальный показатель ИБВЭД рассчитывается в годовом выражении, показатель также агрегирован до годового уровня. Подчеркнем, что данная процедура не является необходимым элементом расчета оперативного индикатора, она используется исключительно для сопоставления с официальными данными. Для этого сначала определено среднегодовое значение сводной тональности:

$$Tonality_t^{year} = \frac{1}{12} \sum_{m=1}^{12} Tonality_{union,m,t} \quad (3),$$

где $Tonality_t^{year}$ – среднегодовая тональность;

$Tonality_{union,m,t}$ – значение средней тональности месяца *m* года *t*.

Усреднение позволяет снизить влияние краткосрочных информационных всплесков и повысить устойчивость показателя. Именно по этой причине новости за 2022 г. использованы исключительно с целью формирования среднегодового значения в 2023 г. При этом оперативные свойства построенного индикатора сохраняются на этапе агрегирования тональности отраслевых индексов (этап 2), которые могут использоваться для мониторинга текущего состояния экономической активности без дополнительного усреднения.

4) Расчет индекса $IBVED_{Media}$

Итоговый индикатор экономической динамики рассчитывается как годовой темп прироста тональности:

$$IBVED_{Media,t} = \left(\frac{Tonality_t^{year}}{Tonality_{t-1}^{year}} * 100\right) - 100 \quad (4),$$

где $IBVED_{Media,t}$ – показатель динамики экономической активности в году *t*;

$Tonality_{t-1}^{year}$ – значение показателя тональности в предыдущем году.

Таким образом, структура индекса отражает реальный вклад отраслей в экономику региона.

5) Интерпретация индекса

Индекс $IBVED_{Media}$ интерпретируется следующим образом:

– значения свыше 0 указывают на улучшение экономического информационного фона;

– значение менее 0 свидетельствуют о росте негативной экономической повестки;

– значения, близкие к 0, отражают нейтральный информационный фон.

Результаты и их анализ

1. Динамика отраслевой тональности

Исследование динамики отраслевых индексов тональности новостных публикаций показало, что повестка региональных СМИ характеризуется устойчивым преобладанием положительной информации (см. рис. 3).

Для оценки устойчивости результатов нами проведена ручная проверка подвыборки текстов (n=100), показавшая совпадение классификации на уровне 78%.


Сельское хозяйство
100
80
60
40
20
0
1,0
0,8
0,6
0,4
0,2
0,0
2023
2024
2025
Позитивные тексты
Нейтральные тексты
Негативные тексты
Тональность по группе "Сельское хозяйство" (Правая шкала)


Промышленное производство
40
30
20
10
0
1,0
0,8
0,6
0,4
0,2
0,0
2023
2024
2025
Позитивные тексты
Нейтральные тексты
Негативные тексты
Тональность по группе "Промышленное производство" (Правая шкала)


Строительство
60
50
40
30
20
10
0
1,0
0,8
0,6
0,4
0,2
0,0
2023
2024
2025
Позитивные тексты
Нейтральные тексты
Негативные тексты
Тональность по группе "Строительство" (Правая шкала)

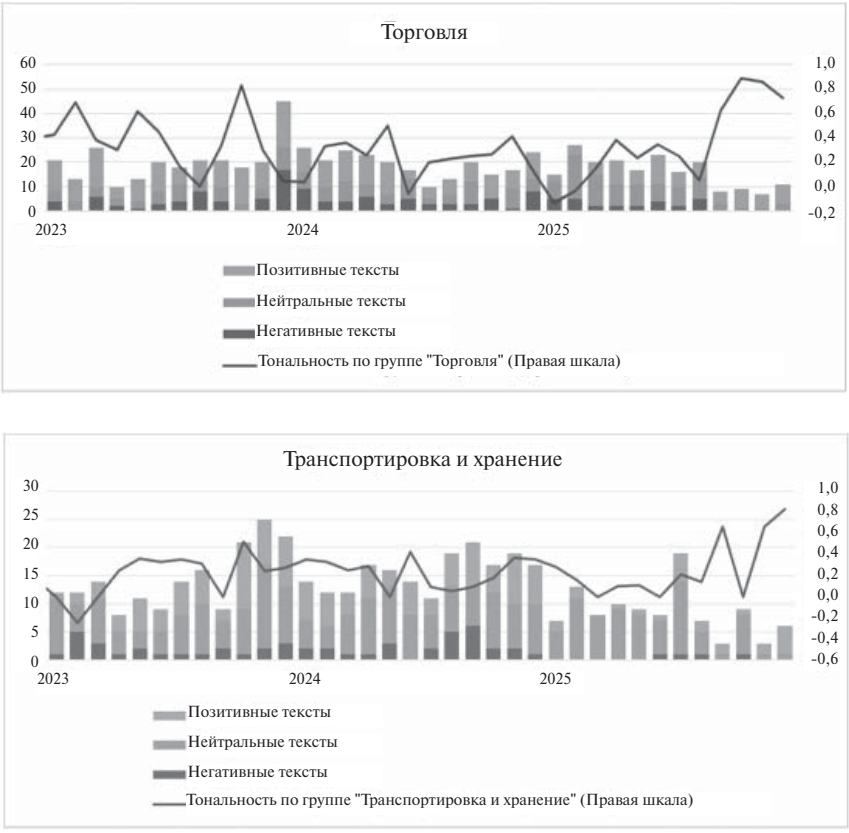


Рисунок 3. Декомпозиция медиатональности в динамике за 2023–2025 гг. по базовым отраслям

Наиболее позитивные значения тональности наблюдаются в сельском хозяйстве и строительстве, где значения индекса на протяжении большей части исследуемого периода находятся в диапазоне 0,6–0,8. Это отражает активное освещение региональными СМИ инвестиционных проектов, развития аграрного производства и инфраструктурного строительства.

В промышленности тональность также сохраняется на относительно высоком уровне (0,4–0,8), однако во второй половине 2024 и в 2025 г. наблюдается заметное снижение позитивного информационного фона.

Наиболее волатильной оказалась тональность в секторе транспорта. Это может быть связано с эпизодическими событиями и меньшим количеством публикаций по данной тематике.

В секторе торговли наблюдается постепенное ослабление позитивной тональности, что может отражать рост обсуждения ценовой динамики и потребительских ожиданий.

Следовательно, повестка местных СМИ во многом соответствует отраслевой структуре экономики региона, где ведущую роль играет аграрный сектор.

2. *Динамика показателя* $IBVED_{Media}$

На основе отраслевых индексов рассчитан интегральный показатель экономической динамики $IBVED_{Media}$, отражающий изменение информационного фона региональной экономики.

Результаты расчета показывают, что индикатор на протяжении всего исследуемого периода находится в отрицательной зоне. При этом он имеет выраженную циклическую динамику (см. рис. 4). В течение 2023 г. наблюдалось постепенное ухудшение тональности, что сопровождалось снижением индекса. Данный период характеризуется массовым обсуждением экономических рисков, инфляционного давления и ограничений внешней торговли.

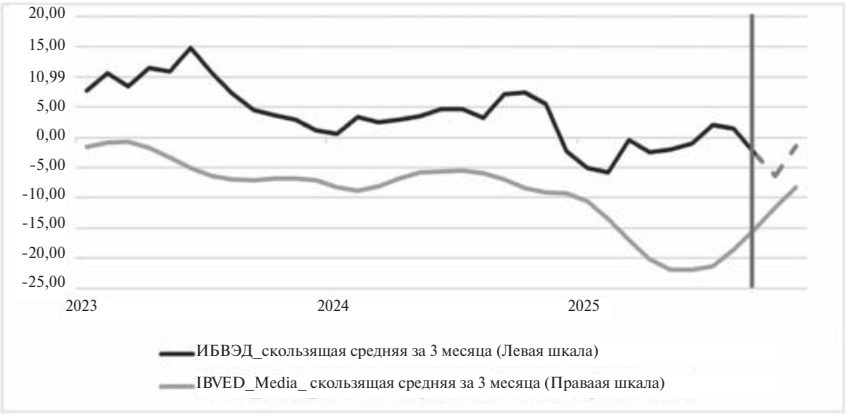


*Рисунок 4. Сопоставление медиаиндекса с официальным показателем ИБВЭД[7] по Ставропольскому краю[8], %**

* Пунктирная линия показывает последнюю дату, на которую были опубликованы данные Росстата на момент исследования.

В 2024 г. показатель $IBVED_{Media}$ становится более стабильным. В этот период экономика как региона, так и страны в целом адаптируется к новой макроэкономической конъюнктуре, что сопровождается большим объемом позитивного контента в СМИ.

Однако уже в 2025 г. динамика показателя демонстрирует повышенную волатильность. Именно в этот период в СМИ наблюдается увеличение числа публикаций, связанных с проблемами почти во всех базовых отраслях.

3. Сравнение с официальным ИБВЭД

Сопоставление $IBVED_{Media}$ с реальным ИБВЭД Ставропольского края свидетельствует, что динамика показателя в значительной степени согласуется с данными официальной статистики. Коэффициент корреляции между предложенным индикатором и значением ИБВЭД составил около 0,4, что указывает на наличие умеренной связи и объясняется высокой волатильностью регионального ИБВЭД, для которого свойственны краткосрочные сезонные колебания. Это подтверждает достаточно высокую степень связи между информационным фоном СМИ и фактической динамикой региональной экономики.

Важно отметить, что $IBVED_{Media}$ измеряет темп прироста тональности, а не ее абсолютный уровень, поэтому отрицательная динамика возникает при снижении тональности относительно соответствующего месяца предыдущего года, что подтверждается данными рисунка 3.

4. Проверка опережающего характера показателя

Дополнительно нами проведена проверка гипотезы о возможном опережающем характере оперативного индикатора. Для этого построены временные лаги $IBVED_{Media}$ относительно официального показателя ИБВЭД.

Результаты расчетов показывают, что учет временных лагов не приводит к устойчивому увеличению коэффициента корреляции между показателями. При анализе различных временны́х интервалов коэффициент корреляции варьируется в диапазоне от 0,40 до 0,80.

Заключение

Проведенное исследование ставило целью разработку высокочастотного медиаиндикатора экономической активности региона на основе анализа новостных публикаций средств массовой

информации с использованием методов обработки естественного языка и больших языковых моделей.

В работе сформирован методический подход к построению показателя экономической динамики, включающий несколько последовательных этапов: сбор массива новостных публикаций региональных СМИ, тематическую классификацию текстов по укрупненным секторам экономики, оценку экономической тональности новостей с использованием большой языковой модели, агрегирование полученных оценок в показатель $IBVED_{Media}$.

Эмпирическая апробация методики на материалах Ставропольского края свидетельствует, что новостной поток региональных СМИ содержит значимую информацию о состоянии экономики региона. Однако СМИ склонны к фиксации позитивных событий.

Расчет показателя $IBVED_{Media}$ позволил оцифровать динамику изменения информационного фона в базовых отраслях экономики региона. Его сравнение с официальным ИБВЭД продемонстрировало наличие статистически значимой взаимосвязи, следовательно, разработанный оперативный индикатор может отражать изменения динамики в отраслях экономики и выступать в роли дополнительного источника информации.

Кроме того, полученные результаты подтверждают возможность использования *LLM*-моделей для извлечения экономически значимой информации из текстовых данных. Вместе с тем мы признаем, что текущая валидация ограничена внутривыборочным корреляционным анализом на коротком временном промежутке и не может быть рассмотрена как полноценная проверка предсказательных свойств построенного индикатора. Позиционирование данного показателя как инструмента наукастинга требует проведения псевдовыборочного эксперимента в реальном времени. В этой связи перспективные направления дальнейших исследований включают: формирование базы данных реального времени для ИБВЭД, проведение псевдовыборочного эксперимента для строгого выявления оперативных свойств индикатора, апробацию и внедрение новых языковых моделей.

Примечания

[1] Настоящая статья отражает личную позицию авторов. Содержание и результаты данного исследования не следует рассматривать, в том числе цитировать

в каких-либо изданиях, как официальную позицию Банка России или указание на официальную политику или решение регулятора.

[2] На момент подготовки итоговой версии статьи (март 2026 г.) имеются данные по ВРП и ВВП только за 2023 г.

[3] *Селезнев С., Шибитов Д., Яковлева К., Власов С. и др.* Новая методология новостного индекса // Доклады об экономических исследованиях. 2021. Режим доступа: https://cbr.ru/Collection/Collection/File/35503/index_2107.pdf (дата обращения: 26.11.2025).

[4] *Евстигнеева А. Г., Карпов Д. А.* Влияние негативных новостей на восприятие инфляции населением // Доклады об экономических исследованиях. 2023. № 111. Режим доступа: https://cbr.ru/Content/Document/File/144918/wp_111.pdf (дата обращения: 20.11.2025).

[5] Alam M. J., Boyle S., Li H., Sekhposyan T. (2026) ChatMacro: Evaluating Inflation Forecasts of Generative AI. *Federal Reserve Bank of San Francisco, Working Paper Series* 2026-04. Available at: https://ideas.repec.org/p/fip/fedfwp/102407.html (accessed: 24.02.2026).

[6] Показана вероятность отнесения текста к соответствующей теме, т. е. «уверенность». Рассчитывается моделью *MoritzLaurer/xlm-v-base-mnli-xnli* автоматически при проведении текстовой классификации.

[7] Для отражения сопоставимых данных на графике ИБВЭД уменьшен на 100 п. п.

[8] Для дополнительной оценки устойчивости связи было применено скользящее трехмесячное усреднение. Мы признаем, что данная процедура дополнительного сглаживания снижает оперативность индикатора. Она используется нами исключительно как аналитический прием для демонстрации того, что выявленная связь не является элементом высокочастотного шума. В практике оперативного мониторинга трехмесячное сглаживание не является обязательным.

Библиография

Гареев М. Ю., Полбин А. В. Наукастинг: оценка изменения ключевых макроэкономических показателей с использованием методов машинного обучения // Вопросы экономики. 2022. № 8. С. 133–157. DOI: 10.32609/0042-8736-2022-8-133-157

Зубарев А. В., Рыбак К. С. Наукастинг ВВП: динамическая факторная модель и официальные прогнозы // Экономическое развитие России. 2021. № 12. С. 34–40.

Лазарян С. С., Герман Н. Е. Прогнозирование текущей динамики ВВП на основе данных поисковых запросов // Финансовый журнал. 2018. № 6. С. 83–94. DOI: 10.31107/2075-1990-2018-6-83-94

Макеева Н. М., Станкевич И. П. Наукастинг элементов использования ВВП России // Экономический журнал Высшей школы экономики. 2022. № 4. С. 598–622. DOI: 10.17323/1813-8691-2022-26-4-598-622

Поршаков А. С., Пономаренко А. А., Синяков А. А. Оценка и прогнозирование ВВП России с помощью динамической факторной модели // Журнал Новой экономической ассоциации. 2016. № 2 (30). С. 60–76. DOI: 10.31737/2221-2264-2016-30-2-3

Станкевич И. П. Применение MIDAS-моделей с марковским переключением для наукастинга ВВП и его компонентов // Прикладная эконометрика. 2023. № 2. С. 122–143. DOI: 10.22394/1993-7601-2023-70-122-143

Фокин Н. Д. Наукастинг и прогнозирование основных российских макроэкономических показателей с помощью MFBVAR-модели // Экономическая политика. 2023. Т. 18. № 3. С. 110–135. DOI: 10.18288/1994-5124-2023-3-110-135

Acedański J. (2024) Disaggregation and Nowcasting of Regional GDP Series with a Simple Smoothing Algorithm. *Journal of Official Statistics* 40 (4): 508–529. DOI: 10.1177/0282423X241277716.

Boyko V., Kislyak N., Nikitin M., Oborin O. (2020) Methods for Estimating the Gross Regional Product Leading Indicator. *Russian Journal of Money and Finance* 79 (3): 3–29. DOI: 10.31477/rjmf.202003.03.

Bybee L. (2023) Surveying Generative AI's Economic Expectations. *arXiv*. DOI: 10.48550/arXiv.2305.02823. Available at: https://arxiv.org/abs/2305.02823 (accessed: 17.01.2026).

Clements M. P., Galvão A. B. (2009) Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models. *Journal of Applied Econometrics* 24 (7): 1187–1206. DOI: 10.1002/jae.1075.

Dahlhaus T., Guénette J.-D., Vasishtha G. (2017) Nowcasting BRIC+M in Real Time. *International Journal of Forecasting* 33 (4): 915–935. DOI: 10.1016/j.ijforecast.2017.05.002.

Eliseev A. (2025) Nowcasting Russian GDP in a Mixed-Frequency DSGE Model with a Panel of Non-Modelled Variables. *Russian Journal of Money and Finance* 84 (3): 63–93.

Fosten J., Greenaway-McGrevy R. (2022) Panel data nowcasting. *Econometric Reviews* 41 (7): 675–696. DOI: 10.1080/07474938.2021.2017670.

Ghysels E. (2016) Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193 (2): 294–314. DOI: 10.1016/j.jeconom.2016.04.008.

Koop G., Korobilis D., Pettenuzzo D. (2019) Bayesian Compressed Vector Autoregressions. *Journal of Econometrics* 210 (1): 135–154. DOI: 10.1016/j.jeconom.2018.11.009.

Mikosch H., Solanko L. (2019) Forecasting Quarterly Russian GDP Growth with Mixed-Frequency Data. *Russian Journal of Money and Finance* 78 (1): 19–35. DOI: 10.31477/rjmf.201901.19.

Richardson A., van Florenstein Mulder T., Vehbi T. (2021) Nowcasting GDP using Machine-Learning Algorithms: A Real-Time Assessment. *International Journal of Forecasting* 37 (2): 941–948. DOI: 10.1016/j.ijforecast.2020.10.005.

Shulyak E. (2022) Macroeconomic Forecasting Using Data from Social Media. *Russian Journal of Money and Finance* 81 (4): 86–112.

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. (2017) Attention Is All You Need. *arXiv*. DOI: 10.48550/arXiv.1706.03762. Available at: https://arxiv.org/abs/1706.03762 (accessed: 24.12.2025).

Yakovleva K. (2018) Text Mining-based Economic Activity Estimation. *Russian Journal of Money and Finance* 77 (4): 26–41. DOI: 10.31477/rjmf.201804.26.

Zhang Q., Ni H., Xu H. (2023) Nowcasting Chinese GDP in a Data-Rich Environment: Lessons from Machine Learning Algorithms. *Economic Modelling* 122: 106204. DOI: 10.1016/j.econmod.2023.106204.

Notes

Alam M. J., Boyle S., Li H., Sekhposyan T. (2026) ChatMacro: Evaluating Inflation Forecasts of Generative AI. *Federal Reserve Bank of San Francisco, Working Paper Series* 2026-04. Available at: https://ideas.repec.org/p/fip/fedfwp/102407.html (accessed: 24.02.2026).

Evstigneeva A., Karpov D. (2023) Vliyanie negativnykh novostey na vospriyatie inflyatsii naseleniem [Impact of negative news on households' perception of inflation]. *Bank of Russia Working Papers* No 111. Available at: https://cbr.ru/Content/Document/File/144918/wp_111.pdf (accessed: 20.11.2025).

Seleznev S., Shibitov D., Yakovleva K., Vlasov S., Chernyadyev D. (2021) Novaya metodologiya novostnogo indeksa [New news index methodology]. *Bank of Russia Working Papers*. Available at: https://cbr.ru/Collection/Collection/File/35503/index_2107.pdf (accessed: 26.11.2025).

References

Acedański J. (2024) Disaggregation and Nowcasting of Regional GDP Series with a Simple Smoothing Algorithm. *Journal of Official Statistics* 40 (4): 508–529. DOI: 10.1177/0282423X241277716

Boyko V., Kislyak N., Nikitin M., Oborin O. (2020) Methods for Estimating the Gross Regional Product Leading Indicator. *Russian Journal of Money and Finance* 79 (3): 3–29. DOI: 10.31477/rjmf.202003.03

Bybee L. (2023) Surveying Generative AI's Economic Expectations. *arXiv*. DOI: 10.48550/arXiv.2305.02823. Available at: https://arxiv.org/abs/2305.02823 (accessed: 17.01.2026).

Clements M. P., Galvão A. B. (2009) Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models. *Journal of Applied Econometrics* 24 (7): 1187–1206. DOI: 10.1002/jae.1075

Dahlhaus T., Guénette J.-D., Vasishtha G. (2017) Nowcasting BRIC+M in Real Time. *International Journal of Forecasting* 33 (4): 915–935. DOI: 10.1016/j.ijforecast.2017.05.002

Eliseev A. (2025) Nowcasting Russian GDP in a Mixed-Frequency DSGE Model with a Panel of Non-Modelled Variables. *Russian Journal of Money and Finance* 84 (3): 63–93.

Fokin N. D. (2023) Naukasting i prognozirovanie osnovnykh rossiyskikh makroekonomicheskikh pokazateley s pomoshch'yu MFBVAR-modeli [Nowcasting and forecasting key Russian macroeconomic variables with the

MFBVAR model]. *Economic Policy* 18 (3): 110–135. DOI: 10.18288/1994-5124-2023-3-110-135 (In Russian)

Fosten J., Greenaway-McGrevy R. (2022) Panel data nowcasting. *Econometric Reviews* 41 (7): 675–696. DOI: 10.1080/07474938.2021.2017670

Gareev M. Yu., Polbin A. V. (2022) Naukasting: otsenka izmeneniya klyuchevykh makroekonomicheskikh pokazateley s ispol'zovaniem metodov mashinnogo obucheniya [Nowcasting Russia's key macroeconomic variables using machine learning]. *Voprosy Ekonomiki* 8: 133–157. DOI: 10.32609/0042-8736-2022-8-133-157 (In Russian)

Ghysels E. (2016) Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193 (2): 294–314. DOI: 10.1016/j.jeconom.2016.04.008

Koop G., Korobilis D., Pettenuzzo D. (2019) Bayesian Compressed Vector Autoregressions. *Journal of Econometrics* 210 (1): 135–154. DOI: 10.1016/j.jeconom.2018.11.009

Lazaryan S. S., German N. E. (2018) Prognozirovanie tekushchey dinamiki VVP na osnove dannykh poiskovykh zaprosov [Forecasting current GDP dynamics based on search query data]. *Financial Journal* 6: 83–94. DOI: 10.31107/2075-1990-2018-6-83-94 (In Russian)

Makeeva N. M., Stankevich I. P. (2022) Naukasting elementov ispol'zovaniya VVP Rossii [Nowcasting of the components of Russian GDP]. *HSE Economic Journal* 4: 598–622. DOI: 10.17323/1813-8691-2022-26-4-598-622 (In Russian)

Mikosch H., Solanko L. (2019) Forecasting Quarterly Russian GDP Growth with Mixed-Frequency Data. *Russian Journal of Money and Finance* 78 (1): 19–35. DOI: 10.31477/rjmf.201901.19

Porshakov A. S., Ponomarenko A. A., Sinyakov A. A. (2016) Otsenka i prognozirovanie VVP Rossii s pomoshch'yu dinamicheskoy faktornoy modeli [Estimation and forecasting of Russia's GDP using a dynamic factor model]. *Journal of the New Economic Association* 2 (30): 60–76. DOI: 10.31737/2221-2264-2016-30-2-3 (In Russian)

Richardson A., van Florenstein Mulder T., Vehbi T. (2021) Nowcasting GDP using Machine-Learning Algorithms: A Real-Time Assessment. *International Journal of Forecasting* 37 (2): 941–948. DOI: 10.1016/j.ijforecast.2020.10.005

Shulyak E. (2022) Macroeconomic Forecasting Using Data from Social Media. *Russian Journal of Money and Finance* 81 (4): 86–112.

Stankevich I. P. (2023) Primenenie MIDAS-modeley s markovskim pereklyucheniem dlya naukastinga VVP i ego komponentov [Application of MIDAS models with Markov switching for nowcasting GDP and its components]. *Applied Econometrics* 2: 122–143. DOI: 10.22394/1993-7601-2023-70-122-143 (In Russian)

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. (2017) Attention Is All You Need. *arXiv*. DOI: 10.48550/arXiv.1706.03762. Available at: https://arxiv.org/abs/1706.03762 (accessed: 24.12.2025).

Yakovleva K. (2018) Text Mining-based Economic Activity Estimation. *Russian Journal of Money and Finance* 77 (4): 26–41. DOI: 10.31477/rjmf.201804.26

Zhang Q., Ni H., Xu H. (2023) Nowcasting Chinese GDP in a Data-Rich Environment: Lessons from Machine Learning Algorithms. *Economic Modelling* 122: 106204. DOI: 10.1016/j.econmod.2023.106204

Zubarev A. V., Rybak K. S. (2021) Naukasting VVP: dinamicheskaya faktornaya model‘ i ofitsial‘nye prognozy [Nowcasting GDP: dynamic factor model and official forecasts]. *Economic Development of Russia* 12: 34–40 (In Russian)

Поступила в редакцию
25.03.2026