

Кукарцев Владислав Викторович

доцент

ФГБОУ ВО «Российский государственный
аграрный университет – МСХА им. К.А. Тимирязева»

г. Москва

Гречаников Никита Дмитриевич

студент

ФГАОУ ВО «Российский университет
дружбы народов имени Патриса Лумумбы»

г. Москва

Глинская Анна Романовна

студентка

ФГБОУ ВО «Сибирский государственный университет науки
и технологий им. академика М.Ф. Решетнева»

г. Красноярск, Красноярский край

МНОГОУРОВНЕВЫЙ АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИЙ МИРОВОЙ ЭКОНОМИКИ С ПРИМЕНЕНИЕМ МЕТОДОВ DATA SCIENCE И МАШИННОГО ОБУЧЕНИЯ

***Аннотация:** в работе анализируются основные тенденции мировой экономики с применением методов Data Science и машинного обучения. На основе открытых макроэкономических данных Всемирного банка (2010–2023 гг.) проведён многоуровневый анализ показателей более чем 170 стран. Цель исследования – выявить взаимосвязи между экономическими индикаторами и факторами роста ВВП. Сравнение линейной регрессии и алгоритма Random Forest показало преимущества ансамблевых методов в выявлении нелинейных зависимостей. Результаты подтверждают эффективность интеллектуального анализа данных для повышения точности экономических прогнозов.*

***Ключевые слова:** макроэкономика, прогнозирование, Data Science, машинное обучение, рост ВВП, нелинейное моделирование.*

I. Введение.

Современная мировая экономическая система характеризуется высокой степенью волатильности и сложной взаимосвязанностью агентов, что ставит новые вызовы в вопросах стратегического планирования [2; 5]. Традиционные эконометрические подходы зачастую сталкиваются с ограничениями при обработке многомерных массивов данных и моделировании нелинейных процессов [1]. В ответ на эти методологические вызовы инструменты Data Science и машинного обучения (МО) становятся необходимым дополнением к классическому инструментарию, позволяя строить более адаптивные прогностические модели [3].

В рамках данного исследования экономический анализ проводится комплексно на трех уровнях: макроуровне, мезоуровне и микроуровне [6]. Исследование направлено на демонстрацию эффективности применения алгоритмов МО для выявления тенденций и построения прогнозов на основе глобального макроэкономического датасета [8].

II. Методы.

Эмпирической базой исследования послужил набор данных, сформированный на основе открытой статистики Всемирного банка. Выборка включает годовые показатели для 175 стран. В качестве целевой переменной для моделирования был выбран годовой рост ВВП (`gdp_growth_annual`), а предикторами выступили: инфляция (ИПЦ), уровень безработицы, ВВП на душу населения и отношение государственного долга к ВВП.

Для решения задачи прогнозирования и оценки взаимосвязей применялся сравнительный подход. В качестве базовой модели использовалась линейная регрессия, позволяющая оценить прямые зависимости. Для улавливания сложных нелинейных взаимодействий между признаками был выбран алгоритм «Случайный лес» [4; 7]. Обучение моделей проводилось на расширенном наборе признаков, включающем также реальную процентную ставку, сальдо текущего счета и показатели государственных финансов. Качество аппроксимации оценивалось метриками среднеквадратичной ошибки и коэффициента детерминации.

III. Результаты и обсуждение.

Визуализация глобальных тенденций демонстрирует значительную изменчивость мировой экономики. На рисунке 1 отражена динамика среднегодового роста ВВП: отчётливо виден резкий спад в 2020 году, вызванный пандемией COVID-19, и последующее восстановление в 2021 году. Параллельно изменениям в ВВП, рисунок 2 показывает страновые различия: Китай сохраняет устойчиво высокие темпы роста, тогда как США, Германия и Россия демонстрируют повышенную волатильность, особенно в кризисные периоды. Дополняет анализ рисунок 3, на котором представлено распределение государственного долга. Он указывает на наличие группы стран с критически высокой долговой нагрузкой, создающей дополнительные макроэкономические риски. Для предсказания роста ВВП были обучены линейная регрессия и случайный лес. Тренировочный набор данных включал следующие признаки: `inflation_cpi` (Инфляция, ИПЦ); `unemployment_rate` (Уровень безработицы); `interest_rate_real` (Реальная процентная ставка); `current_account_balance_gdp` (Сальдо счета текущих операций, % от ВВП); `government_expense_gdp` (Государственные расходы, % от ВВП); `government_revenue_gdp` (Государственные доходы, % от ВВП); `tax_revenue_gdp` (Налоговые доходы, % от ВВП); `public_debt_gdp` (Государственный долг, % от ВВП).

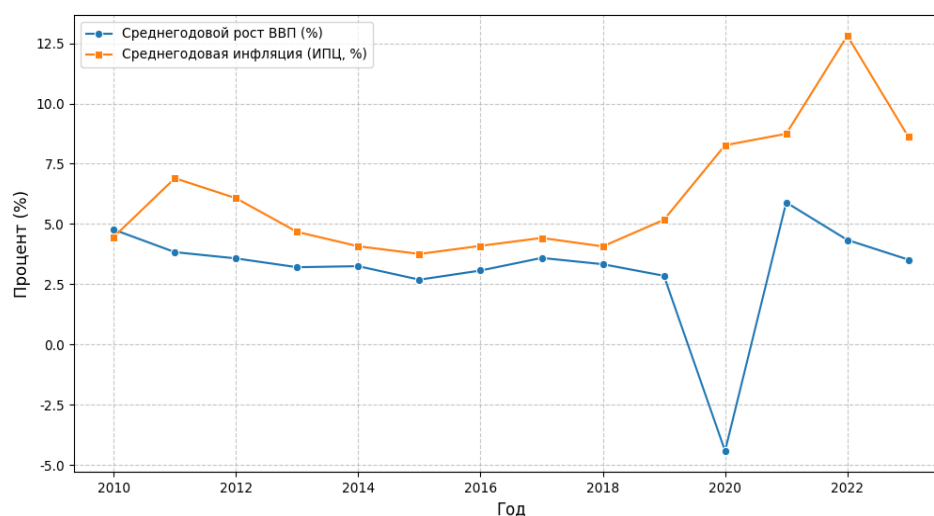


Рис. 1. Глобальные макроэкономические тенденции (2010–2023)

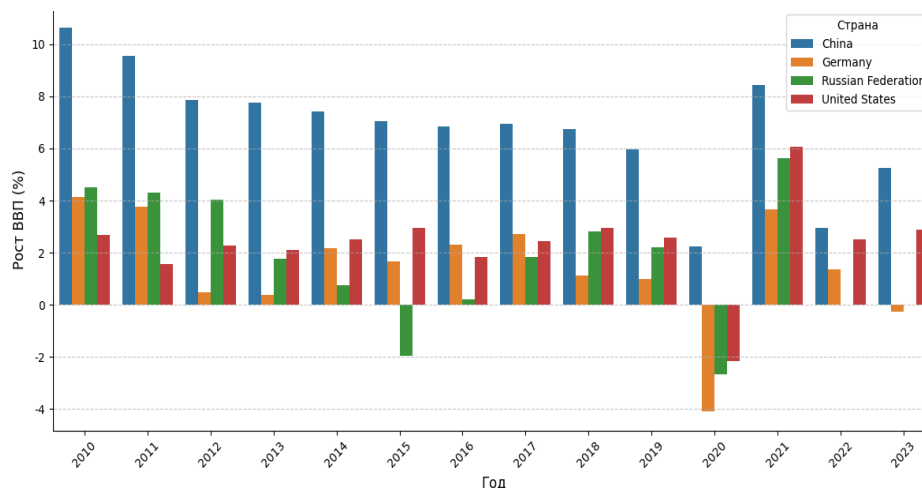


Рис. 2. Сравнение годового роста ВВП по странам (2010–2023)

Оценка прогностических моделей показала, что точное предсказание макроэкономических показателей затруднено из-за их высокой вариативности. Сравнение результатов демонстрирует, что линейная регрессия и случайный лес имеют близкие значения качества, однако визуальное сопоставление фактических и прогнозируемых данных (рисунок 4) показывает более устойчивое улавливание общей тенденции моделью Random Forest. Дополнительное преимущество этого алгоритма проявляется в распределении значимости признаков (рисунок 5): наибольшее влияние на рост ВВП оказывают реальная процентная ставка и уровень инфляции, что подчёркивает ключевую роль монетарной политики в динамике экономического развития.

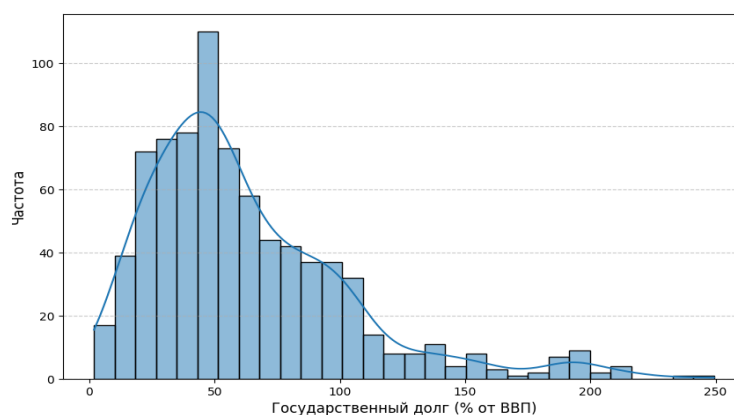


Рис. 3. Распределение государственного долга (% от ВВП) по странам

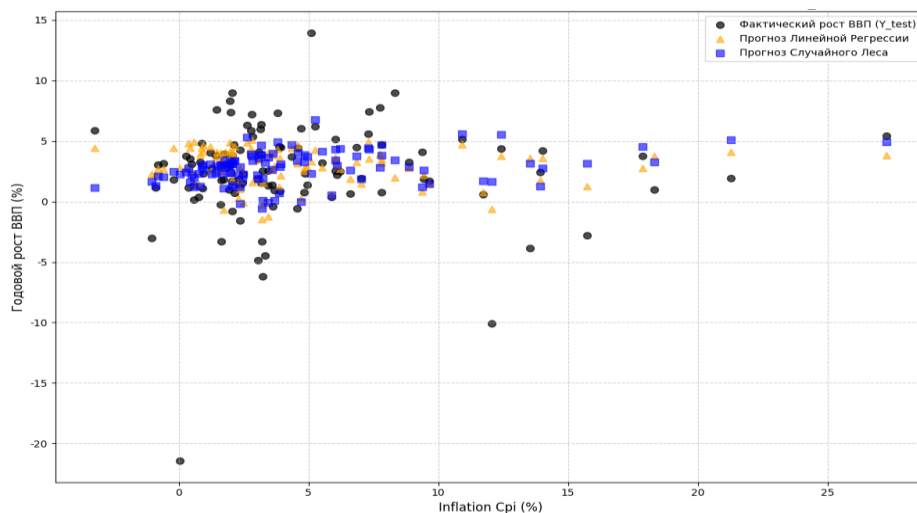


Рис. 4. Сравнение фактического и прогнозируемого роста ВВП в зависимости от inflation_cpi

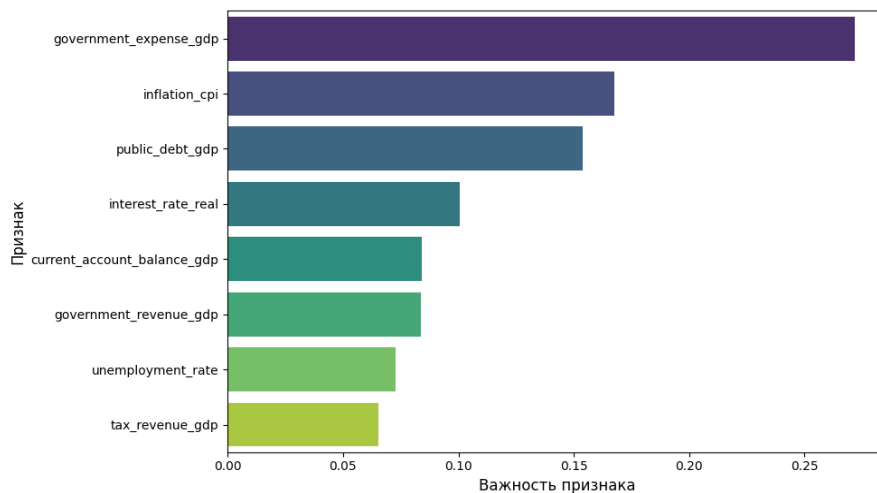


Рис. 5. Важность признаков для прогнозирования роста ВВП (Random Forest)

IV. Заключение.

Проведенное исследование подтверждает гипотезу о том, что методы Data Science являются эффективным инструментом для углубленного экономического анализа, дополняющим классические эконометрические методы. Анализ данных позволил не только визуализировать влияние шоковых событий, таких как пандемия, но и количественно оценить структурные взаимосвязи. Прогнозирование ВВП остается нетривиальной задачей, требующей учета нелинейности

процессов. Перспективы работы связаны с применением более продвинутых моделей временных рядов и использованием неструктурированных данных для повышения точности прогнозов.

Список литературы

1. Вахрушева И.А. К постановке проблемы формирования математической направленности студентов технического вуза в процессе профессиональной подготовки / И.А. Вахрушева // Проблемы современного педагогического образования. – 2019. – №62-2. – С. 46–48. EDN YXUSTB
2. Вахрушева И.А. Сборник индивидуальных заданий по математике / И.А. Вахрушева. – 2012. EDN WDBRTH
3. Карлов А.М. Сравнительная оценка эконометрических моделей по точности прогнозирования / А.М. Карлов, Р.А. Мнацаканян // Балтийский экономический журнал. – 2024. – №2(46). – С. 6–17. DOI 10.46845/2073-3364-2024-0-2-6-16. EDN JAJRAA
4. Лешер О.В. Включение студентов технического вуза в исследовательскую деятельность как педагогическое условие формирования их математической направленности / О.В. Лешер, И.А. Вахрушева, Е.М. Гугина // Перспективы науки и образования. – 2019. – №5(41). – С. 147–157. DOI 10.32744/pse.2019.5.11. EDN HGKUQY
5. Прудкий А.С. Анализ разновременных карт температур на основе дистанционного зондирования / А.С. Прудкий // Методы и программные средства дистанционного зондирования Земли: сборник материалов. – 2022. – С. 111–121. EDN XVBCBJ
6. Прудкий А.С. Применение систем пространственного позиционирования для навигации автотранспортного предприятия / А.С. Прудкий // Методы и программные средства дистанционного зондирования Земли: сборник материалов. – 2022. – С. 85–96. EDN JXFBNQ

7. Шайтура С.В. Использование дистанционного зондирования при мониторинге полей в точном земледелии / С.В. Шайтура // Землеустройство, кадастр и мониторинг земель. – 2023. – Т. 8. DOI 10.33920/sel-04-2308-06. EDN NNGTIB

8. Leshner O.V. Characteristic of personnel readiness for conflict management in the organization / O.V. Leshner // Religación: Revista de Ciencias Sociales y Humanidades. – 2019. – Vol. 4. No. 17. – P. 145–150. EDN XDIJF