

УДК 004.67

5.2.2. Математические, статистические и инструментальные методы экономики (физико-математические науки, экономические науки)

СПОСОБЫ И ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ К ПРОЦЕССУ АНАЛИТИКИ БОЛЬШИХ ДАННЫХ В СЕЛЬСКОМ ХОЗЯЙСТВЕ

Параскевов Александр Владимирович
старший преподаватель
SPIN-код: 2792-3483
e-mail paraskevov.a@kubsau.ru
ФГБОУ ВО Кубанский государственный аграрный университет имени И. Т. Трубилина, г. Краснодар, РФ

Применение больших данных в сельском хозяйстве является актуальной темой исследований в настоящее время. С использованием различных технологий и методов анализа данных, большие объемы информации, собранные в ходе сельскохозяйственной деятельности, могут быть превращены в ценные знания и рекомендации для улучшения урожайности, сокращения затрат и повышения эффективности производства. Технологии сбора, хранения, обработки и анализа больших данных включают в себя датчики, IoT, облачные вычисления, машинное обучение и другие современные методы. Применение больших данных в сельском хозяйстве может улучшить прогнозирование урожайности, оптимизировать использование ресурсов, помочь в принятии решений по выбору сортов, определению оптимальных доз удобрений и пестицидов, а также повысить качество и безопасность продукции. Результаты исследований показывают, что применение больших данных в сельском хозяйстве имеет большой потенциал для улучшения эффективности производства и роста доходов сельскохозяйственных предприятий. Применение больших данных также может помочь сельским хозяйственным предприятиям улучшить управление поставками, снизить риски и повысить точность прогнозирования спроса на сельскохозяйственную продукцию. Кроме того, большие данные могут быть использованы для мониторинга и прогнозирования климатических условий, что может помочь сельским хозяйственным предприятиям принимать решения о посевах и управлении растениями, учитывая изменения погодных условий. В целом, применение больших данных в сельском хозяйстве может помочь улучшить эффективность и устойчивость сельского хозяйства, уменьшить негативное воздействие на окружающую среду и повысить качество жизни

UDC 004.67

5.2.2. Mathematical, statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

METHODS AND TECHNICAL REQUIREMENTS FOR THE PROCESS OF BIG DATA ANALYTICS IN AGRICULTURE

Paraskevov Alexander Vladimirovich
Senior lecturer
RSCI SPIN-code: 2792-3483
e-mail paraskevov.a@kubsau.ru
FSAU HE Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia

The use of big data in agriculture is an urgent topic of research at the present time. Using various technologies and methods of data analysis, large amounts of information collected during agricultural activities can be turned into valuable knowledge and recommendations to improve yields, reduce costs and increase production efficiency. Technologies for collecting, storing, processing and analyzing big data include sensors, IoT, cloud computing, machine learning and other modern methods. The use of big data in agriculture can improve yield forecasting, optimize resource use, help in making decisions on the selection of varieties, determining optimal doses of fertilizers and pesticides, as well as improve the quality and safety of products. The research results show that the use of big data in agriculture has great potential for improving production efficiency and income growth of agricultural enterprises. The use of big data can also help agricultural enterprises improve supply management, reduce risks and improve the accuracy of forecasting demand for agricultural products. In addition, big data can be used to monitor and predict climatic conditions, which can help agricultural enterprises make decisions about crops and plant management, taking into account changes in weather conditions. In general, the use of big data in agriculture can help to improve the efficiency and sustainability of agriculture, reduce the negative impact on the environment and improve the quality of life of rural communities. However, it is necessary to take into account factors related to data privacy and security, as well as the need to develop appropriate infrastructures and technological solutions for the use of big data in agriculture

сельских сообществ. Однако, необходимо учитывать факторы, связанные с конфиденциальностью и безопасностью данных, а также необходимость разработки соответствующих инфраструктур и технологических решений для использования больших данных в сельском хозяйстве

Ключевые слова: БОЛЬШИЕ ДАННЫЕ, ИНТЕРНЕТ ВЕЩЕЙ, ТЕХНОЛОГИЧЕСКИЕ РЕШЕНИЯ, АНАЛИЗ ДАННЫХ

Keywords: BIG DATA, INTERNET OF THINGS, TECHNOLOGICAL SOLUTIONS, DATA ANALYSIS

<http://dx.doi.org/10.21515/1990-4665-187-018>

Инструментальные методы в экономике сельского хозяйства – это различные методы и технологии, используемые для исследования и анализа производства, распределения, потребления и других экономических процессов в аграрном секторе.

Статистический анализ – это метод, позволяющий определять основные тенденции и закономерности в экономических данных, таких как объемы производства, расходы на производство, цены, спрос и предложение на продукцию сельского хозяйства.

Моделирование – это метод, который используется для разработки математических моделей, которые позволяют прогнозировать экономические процессы, такие как предложение, спрос, цены и т.д. Эконометрика – это метод, который используется для анализа экономических данных и выявления связей между различными переменными, такими как объем производства, цены на продукцию, уровень доходов и другие. Маркетинговые исследования – это метод, который используется для анализа рынка и потребительских предпочтений, чтобы определить оптимальную стратегию маркетинга для продукции сельского хозяйства. Финансовый анализ – это метод, который используется для анализа финансовых показателей предприятия, таких как доходы, расходы, прибыль и убытки, чтобы принимать решения о финансовых стратегиях и инвестициях. SWOT-анализ – это метод,

<http://ej.kubagro.ru/2023/03/pdf/18.pdf>

который используется для анализа сильных и слабых сторон предприятия, а также возможностей и угроз на рынке сельского хозяйства.

Существует множество экономических моделей, которые используются для анализа экономических систем в сельском хозяйстве. Модели производственной функции: эти модели изучают связь между входами (например, землей, трудом, капиталом и т.д.) и выходами (например, урожаем, продукцией и т.д.) в производственном процессе. Модели потребления: эти модели анализируют поведение потребителей и как они принимают решения о покупке товаров и услуг. Модели экономического роста: эти модели изучают факторы, способствующие росту экономики сельского хозяйства, включая инвестиции, технологический прогресс и уровень образования. Модели рыночного равновесия: эти модели анализируют взаимодействие спроса и предложения на рынке сельскохозяйственной продукции и определяют цены на товары и услуги. Модели оптимизации: эти модели используются для оптимизации производства и распределения ресурсов в сельском хозяйстве, чтобы достичь максимальной прибыли или минимальных затрат. Модели макроэкономического анализа: эти модели исследуют экономику сельского хозяйства в целом, включая макроэкономические показатели, такие как ВВП, инфляция и безработица.

Информационные технологии могут играть важную роль в повышении эффективности сельского хозяйства. Ниже приведены несколько способов, которыми информационные технологии могут использоваться в сельском хозяйстве.

1. Управление данными и аналитика. Сельскохозяйственные предприятия могут использовать информационные технологии для сбора и хранения данных о погоде, почве, урожайности и других факторах, которые влияют на производство. Эти данные могут быть анализированы с помощью алгоритмов машинного обучения, чтобы выявить тренды и

помочь принимать решения по управлению сельскохозяйственным производством.

2. Мониторинг растений и животных. Информационные технологии могут быть использованы для мониторинга здоровья и состояния растений и животных. Датчики и устройства мониторинга могут помочь собирать данные о заболеваниях, питании и других факторах, которые влияют на здоровье животных и растений.

3. Управление ресурсами. Информационные технологии могут использоваться для управления ресурсами, такими как вода, удобрения и энергия. С помощью автоматических систем управления ресурсами можно более эффективно использовать ресурсы и уменьшить затраты на их использование.

4. Электронная коммерция. Информационные технологии могут использоваться для создания электронных рынков, где сельскохозяйственные продукты могут быть проданы и куплены. Это позволяет сельскохозяйственным предприятиям получать лучшие цены за свою продукцию и сокращает затраты на продажу и распространение продуктов.

5. Обучение и консультирование. Информационные технологии могут использоваться для обучения и консультирования сельскохозяйственных работников. Веб-сайты, приложения и другие средства обучения могут помочь работникам улучшить свои знания и навыки в области сельского хозяйства, а также получить советы от экспертов.

Большие данные (big data) – это огромные объемы информации, которые могут быть собраны, хранятся, обрабатываются и анализируются для получения ценных знаний и принятия решений. Также под термином «большие данные» понимается комплекс методов обработки и анализа объемов данных с применением алгоритмов и технологий, позволяющих

производить эффективную обработку и выборку. В сельском хозяйстве, большие данные могут помочь улучшить производство и повысить эффективность. Ниже приведены несколько примеров, как большие данные могут использоваться в сельском хозяйстве.

Мониторинг почвы и погодных условий. Данные, полученные от датчиков, установленных на полях, могут помочь фермерам лучше понимать, как погодные условия влияют на качество урожая и здоровье почвы. Например, данные о температуре, влажности и осадках могут быть использованы для прогнозирования погоды, чтобы фермеры могли принимать решения о том, когда и как сеять, поливать и убирать урожай.

Мониторинг растений и животных. Данные, полученные от датчиков, установленных на растениях и животных, могут помочь фермерам мониторить здоровье и состояние своих культур и животных. Это может включать измерение температуры, влажности, освещения, питания и других факторов.

Прогнозирование урожайности. Анализ данных о почве, погодных условиях, использовании удобрений, применении пестицидов и других факторах может помочь фермерам прогнозировать урожайность. Это может помочь фермерам лучше планировать их посевные работы и улучшать производительность.

Управление ресурсами. Анализ данных о расходе воды, энергии и других ресурсов может помочь фермерам оптимизировать их использование и уменьшить затраты.

Рыночный анализ. Анализ больших данных о ценах на сельскохозяйственную продукцию, спросе и предложении может помочь фермерам определить наилучшее время для продажи своих товаров и лучшее место для их распространения.

В целом, большие данные могут помочь фермерам принимать более осознанные решения, улучшать эффективность производственных процессов.

Мониторинг растений и животных с помощью анализа больших данных может помочь фермерам определить, какие культуры или животные находятся в здоровом состоянии, а какие нуждаются в дополнительном уходе или лечении.

Для мониторинга здоровья растений, датчики могут собирать данные о температуре, влажности, освещении и других факторах, которые влияют на рост и здоровье растений. Анализ этих данных может помочь фермерам определить оптимальные условия для выращивания различных культур и определить момент, когда растения нуждаются в поливе, подкормке или защите от вредителей и болезней.

Для мониторинга здоровья животных, фермеры могут использовать датчики для сбора данных о температуре тела, пульсе, дыхании, питании и других физиологических показателях. Эти данные могут помочь фермерам определить, когда животные находятся в здоровом состоянии и когда им нужно дополнительное лечение. Например, фермеры могут использовать данные о температуре тела, чтобы определить, когда животные болеют или когда у них возникают другие проблемы со здоровьем.

Анализ больших данных также может помочь фермерам определить, какие виды культур и животных наиболее успешны в определенных условиях, что может помочь им лучше планировать свои посевные работы и улучшать производительность. В целом, мониторинг растений и животных с помощью анализа больших данных может помочь фермерам улучшить качество своей продукции и повысить эффективность своей деятельности.

Мониторинг здоровья растений с помощью больших данных может быть осуществлен с помощью различных методов, таких как сбор и анализ данных о погоде, заболеваниях растений и условиях окружающей среды.

Один из способов мониторинга здоровья растений с помощью больших данных - это использование датчиков и IoT-устройств (устройства интернета вещей) для сбора данных о различных параметрах, таких как температура, влажность, уровень освещенности, содержание удобрений и т.д. Собранные данные могут быть обработаны с помощью алгоритмов машинного обучения и анализа данных, чтобы выявить тенденции и паттерны в развитии растений.

Еще один способ мониторинга здоровья растений – это использование дистанционного зондирования, которое позволяет получать информацию о растительных культурах с помощью спутников или беспилотных летательных аппаратов (дронов). Эти данные могут быть обработаны для определения различных показателей, таких как плотность посева, общее количество зеленой массы, уровень урожайности и другие.

Для анализа больших данных о здоровье растений также могут быть использованы методы глубокого обучения, которые могут помочь выявить скрытые паттерны в данных. Например, алгоритмы нейронных сетей могут быть обучены распознавать симптомы болезней растений по изображениям листьев, что позволяет выявлять их на ранней стадии и принимать меры по лечению.

В целом, мониторинг здоровья растений с помощью больших данных позволяет фермерам быстро выявлять проблемы и принимать меры по предотвращению их распространения, что ведет к улучшению качества урожая и повышению производительности сельского хозяйства.

Предсказание урожайности на основе анализа больших данных может быть выполнено с использованием различных методов и технологий.

Анализ данных погоды: погода играет важную роль в росте и развитии растений. С помощью анализа данных о погоде за прошедшие годы можно выделить паттерны и тренды, которые могут быть использованы для прогнозирования урожайности в будущем.

Сбор и анализ данных о почве: качество почвы также оказывает значительное влияние на урожайность. С помощью анализа больших данных о почвенных параметрах, таких как pH-уровень, содержание азота и других элементов, можно прогнозировать, как будет происходить рост и развитие растений в данном участке.

Использование дистанционного зондирования: дистанционное зондирование позволяет получать информацию о растительных культурах с помощью спутников или беспилотных летательных аппаратов (дронов). Собранные данные могут быть использованы для определения различных показателей, таких как плотность посева, общее количество зеленой массы и другие, что может быть использовано для прогнозирования урожайности.

Использование алгоритмов машинного обучения: алгоритмы машинного обучения могут быть использованы для анализа больших данных и построения моделей прогнозирования урожайности. Например, алгоритмы регрессии могут использоваться для определения связи между урожайностью и погодными условиями и другими факторами, в то время как алгоритмы классификации могут использоваться для выделения групп почвенных параметров, которые могут влиять на урожайность.

В целом, прогнозирование урожайности на основе анализа больших данных может помочь сельскохозяйственным предприятиям повышать эффективность своей деятельности, улучшать качество урожая и оптимизировать расходы на производство.

Точность анализа погодных данных нейросетями может варьироваться в зависимости от нескольких факторов.

Качество и объем данных: точность нейросетевой модели напрямую зависит от качества и объема данных, на которых она обучается. Если данные не соответствуют реальным условиям погоды, то модель не сможет дать точный прогноз.

Архитектура модели: выбор подходящей архитектуры модели также может влиять на ее точность. Некоторые модели могут быть более подходящими для анализа погодных данных, чем другие.

Параметры обучения: правильная настройка параметров обучения может существенно повлиять на точность модели. Например, выбор оптимального шага обучения может ускорить процесс обучения и улучшить точность прогноза.

Объем данных для тестирования: для оценки точности модели необходимо использовать данные, которые не использовались для обучения. Также важно, чтобы объем тестовых данных был достаточным для того, чтобы достоверно оценить точность модели.

В целом, современные нейросетевые модели способны обеспечивать высокую точность прогнозирования погоды. Но чтобы достичь наилучших результатов, необходимо учитывать все перечисленные выше факторы и настраивать модель в соответствии с конкретными условиями и требованиями.

Существует несколько способов, которые могут помочь увеличить точность прогноза нейросетевой модели.

Увеличить объем данных для обучения: чем больше данных используется для обучения модели, тем лучше она сможет обобщать и предсказывать результаты на новых данных. Поэтому, если возможно, стоит попытаться собрать и использовать больше данных.

Настроить параметры модели: настройка параметров модели может существенно повлиять на ее точность. При этом, следует помнить о переобучении модели на тренировочных данных, поэтому необходимо

использовать методы регуляризации для улучшения обобщающей способности модели.

Использовать более сложные модели: если модель не обладает достаточной точностью, то может помочь использование более сложных моделей, которые способны лучше аппроксимировать данные. Однако, использование более сложных моделей может привести к переобучению модели, поэтому необходимо контролировать параметры модели.

Применить предварительную обработку данных: некоторые методы предварительной обработки данных, такие как нормализация, стандартизация и уменьшение размерности могут улучшить качество модели. Например, нормализация данных может помочь сгладить различия в масштабах значений различных признаков.

Использовать ансамбли моделей: ансамблирование нескольких моделей может помочь улучшить точность предсказаний. Например, можно использовать метод бустинга, где несколько моделей обучаются последовательно, на каждом шаге корректируя ошибки предыдущей модели.

Важно помнить, что для достижения наилучшей точности прогноза необходимо выбирать подходящие методы в зависимости от конкретной задачи и данных.

Существует несколько источников больших данных о погоде, которые могут быть использованы для обучения нейросетевых моделей:

Спутниковые данные: спутниковые данные могут предоставлять информацию о многих аспектах погоды, таких как температура, атмосферное давление, влажность, скорость и направление ветра и т.д. Кроме того, спутниковые данные могут быть доступны в реальном времени, что позволяет использовать их для прогнозирования погоды на более короткие временные интервалы.

Данные с метеорологических станций: метеорологические станции собирают данные о погоде на земле, включая температуру воздуха, осадки, ветер, атмосферное давление и т.д. Эти данные могут быть использованы для обучения нейросетевых моделей и создания точных прогнозов.

Исторические данные: исторические данные о погоде могут быть использованы для обучения моделей и анализа тенденций в погоде на длительных временных интервалах. Эти данные могут быть собраны из различных источников, включая метеорологические станции и базы данных.

Данные от датчиков IoT: многие устройства Интернета вещей (IoT) могут собирать данные о погоде, например, температуру, влажность, давление, ветер и т.д. Эти данные могут быть использованы для обучения нейросетевых моделей и создания более точных прогнозов.

Данные с социальных сетей: некоторые социальные сети могут предоставлять данные о погоде, например, фотографии и сообщения пользователей о погоде в определенном месте и время. Эти данные могут быть использованы для создания более точных и релевантных прогнозов погоды.

Важно отметить, что для использования данных в нейросетевых моделях необходимо убедиться в их качестве и точности, а также произвести необходимую предварительную обработку данных для улучшения точности модели.

Нейросетевые модели могут использоваться для прогнозирования погоды, но точность прогнозов сильно зависит от множества факторов, включая доступность и качество данных, используемых для обучения моделей, а также текущие атмосферные условия.

Тем не менее, существуют специализированные сервисы и приложения, которые предоставляют прогнозы погоды на длительные

временные интервалы. Они могут использовать данные и модели, которые обновляются регулярно, чтобы предоставить более точные прогнозы.

Для анализа больших данных необходимо использовать специализированное программное обеспечение, которое позволяет обрабатывать, хранить, анализировать и визуализировать большие объемы данных.

Существует много различных инструментов и технологий для анализа больших данных, некоторые из которых наиболее распространены.

Hadoop является одним из наиболее популярных фреймворков для обработки больших данных. Он использует распределенную систему хранения и обработки данных, которая позволяет распараллеливать задачи и ускорять обработку данных.

Spark – это быстрый и масштабируемый фреймворк для обработки больших данных, который работает в памяти и позволяет обрабатывать данные на многих уровнях, включая SQL, машинное обучение и потоковую обработку данных.

MongoDB – это база данных NoSQL, которая используется для хранения и обработки больших объемов данных. Она позволяет легко масштабировать базу данных и обеспечивает высокую производительность при работе с большими объемами данных.

Tableau – это инструмент для визуализации данных, который позволяет создавать красивые и информативные диаграммы и графики на основе больших объемов данных.

Python является одним из наиболее популярных языков программирования для анализа данных. Он имеет богатую экосистему библиотек и инструментов, которые позволяют работать с большими объемами данных, включая библиотеки NumPy, Pandas, SciPy и многие другие.

Выбор программного обеспечения зависит от конкретной задачи и требований к обработке и анализу данных.

Анализ больших данных требует мощного оборудования, которое может обрабатывать и хранить большие объемы данных. Ниже перечислены некоторые из наиболее распространенных компонентов оборудования, используемых для анализа больших данных:

Кластеры серверов используются для обработки больших объемов данных и могут состоять из сотен и тысяч серверов, объединенных в единую систему. Кластеры серверов обычно используются вместе с фреймворками, такими как Hadoop и Spark, для распределенной обработки данных.

Системы хранения данных. Для хранения больших объемов данных используются специальные системы хранения, такие как NoSQL базы данных, которые обеспечивают масштабируемость и производительность.

Хранилища данных используются для хранения и обработки больших объемов данных. Они могут включать в себя хранилища на основе технологий RAID и NAS, а также распределенные файловые системы.

Большие сети: Большие сети используются для передачи данных между различными компонентами оборудования и устройствами. Они обычно включают в себя высокоскоростные соединения, такие как 10/40/100 Гбит/с Ethernet.

Графические процессоры (GPU): используются для ускорения обработки данных, особенно для задач машинного обучения и глубокого обучения.

Память большого объема. Обработка больших объемов данных требует большого количества оперативной памяти. Поэтому для анализа больших данных может потребоваться использование специальных модулей памяти, таких как модули RDIMM и LRDIMM.

Многопроцессорные системы используются для параллельной обработки данных. Они могут содержать много ядер процессора, что позволяет быстрее обрабатывать данные.

Оборудование для анализа больших данных должно быть мощным и иметь достаточно ресурсов для обработки больших объемов данных. Точные требования к оборудованию зависят от конкретной задачи и объемов данных.

Список литературы

1. Бабенков И.М. Средства и методы защиты информации в экономической сфере / И. М. Бабенков, А. В. Параскевов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №05(119). С. 653 – 665. – IDA [article ID]: 1191605047. – Режим доступа: <http://ej.kubagro.ru/2016/05/pdf/47.pdf>, 0,812 у.п.л.
2. Пенкина Ю. Н. Предпосылки разработки адаптивной системы поддержки принятия оперативных решений в управлении IT-проектами для организаций IT-аутсорсинга / Ю. Н. Пенкина, А. В. Параскевов // Научное обеспечение агропромышленного комплекса, сборник статей по материалам IX всероссийской конференции молодых ученых. – Краснодар: КубГАУ.
3. Мицкевич И. У. Проблемы и способы продвижения IT-продуктов / И. У. Мицкевич, В. Ю. Кац, А. В. Параскевов // Тенденции развития науки и образования, 2022.

References

1. Babenkov I.M. Sredstva i metody zashchity informacii v ekonomicheskoy sfere / I. M. Babenkov, A. V. Paraskevov // Politematicheskij setevoy elektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Elektronnyj resurs]. – Krasnodar: KubGAU, 2016. – №05(119). S. 653 – 665. – IDA [article ID]: 1191605047. – Rezhim dostupa: <http://ej.kubagro.ru/2016/05/pdf/47.pdf>, 0,812 u.p.l.
2. Penkina YU. N. Predposylki razrabotki adaptivnoj sistemy podderzhki prinyatiya operativnyh reshenij v upravlenii IT-proektami dlya organizacij IT-outsorsinga / YU. N. Penkina, A. V. Paraskevov // Nauchnoe obespechenie agropromyshlennogo kompleksa, sbornik statej po materialam IX vserossijskoj konferencii molodyh uchenyh. – Krasnodar: KubGAU.
3. Mickevich I. U. Problemy i sposoby prodvizheniya IT-produktov / I. U. Mickevich, V. YU. Kac, A. V. Paraskevov // Tendencii razvitiya nauki i obrazovaniya, 2022.