

УДК 338.2

DOI 10.58649/1694-9099-2025-4-403-407

ЛИ ШЭНБО, ХУАН ВЭНЬЯ, ШАМБЕТОВА Э., ЧЖУ МЭЙ, МУКАМБАЕВА И.

Жусуп Баласагын атындагы КУУ

ЛИ ШЭНБО, ХУАН ВЭНЬЯ, ШАМБЕТОВА Э., ЧЖУ МЭЙ, МУКАМБАЕВА И.

КНУ имени Жусупа Баласагына

LI SHENGBO, HUANG WENYA, SHAMBETOVA E., ZHU MEI, MUKAMBAYEVA I.

KNU Jusup Balasagyn

ШААРДЫК ЖАНА АЙЫЛДЫК КАЛКТЫН ОРТОСУНДАГЫ КИРЕШЕЛЕРДИН АЙЫРМАСЫН
ИЗИЛДӨӨДӨГҮ ЧОҢ МААЛЫМАТТАР**БОЛЬШИЕ ДАННЫЕ В ИССЛЕДОВАНИИ РАЗРЫВА В ДОХОДАХ МЕЖДУ ГОРОДСКИМ
И СЕЛЬСКИМ НАСЕЛЕНИЕМ**BIG DATA IN THE STUDY OF THE INCOME GAP BETWEEN URBAN AND RURAL
POPULATIONS

Кыскача мүнөздөмө: Бул макалада шаар менен айылдын кирешелеринин айырмасын изилдөөдө чоң маалымат технологияларынын методологиялык инновациялары каралат. Электрондук коммерция маалыматтары, жумуш менен камсыз кылуу платформасынын маалыматтары жана мамлекеттик статистика сыяктуу көп өзгөрмөлүү маалыматтарды интеграциялоо менен шаар менен айылдын кирешелеринин айырмасын талдоо үчүн кеңири моделди түзүүгө болот. Андан тышкары, макалада чоң маалымат технологияларынын чектөөлөрү, мисалы, айыл жерлеринде чоң маалымат ресурстарынын салыштырмалуу жетишсиздиги жана шаар менен айылдын маалыматтарды чогултуу стандарттарынын шайкеш келбестиги талданат жана максаттуу жакшыртуу стратегиялары сунушталат. Изилдөө чоң маалымат технологиялары кирешелеринин айырмасын изилдөө үчүн жаңы ыкмаларды жана көз караштарды сунуштай тургандыгын көрсөтүп турат, бул айырмачылыкты так мүнөздөөгө мүмкүндүк берет. Бирок, алар маалымат ресурстарына жана техникалык колдонууга байланыштуу кыйынчылыктарга да туш болушат. Бул изилдөө шаар менен айылдын кирешелеринин айырмасы боюнча терең жана ири масштабдуу академиялык изилдөөлөр үчүн методологиялык негиз боло алат.

Аннотация: В данной статье рассматриваются методологические инновации технологий больших данных в области исследований, посвященных разрыву в доходах между населением города и деревни. Интеграция многомерных данных, таких как данные электронной коммерции, данные платформ занятости и государственные статистические данные, позволяет построить более полную модель анализа разрыва в доходах между городскими и сельскими жителями. Кроме того, в статье анализируются ограничения применения технологий больших данных, такие как относительная нехватка ресурсов больших данных в сельской местности и несогласованность стандартов сбора данных о городе и деревне, на основании чего предлагаются целевые стратегии совершенствования. Исследование показывает, что технологии больших данных предоставляют новые методы и перспективы для изучения разрыва в доходах, позволяя более точно охарактеризовать этот разрыв. Однако они также сталкиваются с трудностями, связанными с ресурсами данных и техническим применением. Данное исследование может служить в качестве методологической основы для углубленных и масштабных научных исследований разрыва в доходах между городским и сельским населением.

Abstract: This article explores the methodological innovations of big data technologies in urban-rural income gap research. By integrating multivariate data, such as e-commerce data, employment platform data, and government statistics, a more comprehensive model for analyzing the urban-rural income gap can be constructed. Furthermore, the article analyzes the limitations of big data technologies, such as the relative lack of big data resources in rural areas and the inconsistency of urban-rural data collection standards and proposes targeted improvement strategies. The study demonstrates that big data technologies offer new methods and perspectives for studying the income gap, enabling a more accurate characterization of this gap.

However, they also face challenges related to data resources and technical application. This study can serve as a methodological basis for in-depth and large-scale academic research on the urban-rural income gap.

Ключевые слова: технология больших данных; методы пространственного анализа; модели машинного обучения; разрыв в доходах; интеграция многомерных данных; методологические инновации.

Негизги сөздөр: чоң маалыматтар технологиясы; мейкиндик анализинин ыкмалары; машиналык окутуу моделдери; киреше айырмасы; көп өлчөмдүү маалыматтарды интеграциялоо; методологиялык инновациялар.

Keywords: big data technology; spatial analysis methods; machine learning models; income gap; multidimensional data integration; methodological innovations.

Введение

Разрыв в доходах между городским и сельским населением является важной проблемой экономического и социального развития Китая. Точное понимание текущего состояния, причин и тенденций изменения разрыва в доходах между жителями городов и сел имеет решающее значение для разработки политики скоординированного развития этих регионов. Традиционные изучения разрыва в рассматриваемых доходах опираются преимущественно на данные выборочных обследований, которые ограничены размером выборки, актуальностью и одним измерением. Появление технологий больших данных открыло новые возможности для исследования

разрыва в городских и сельских доходах. Они позволяют интегрировать многоисточниковые, массивные и актуальные данные о доходах в городе и деревне, предоставляя более полную информацию и более мощные аналитические инструменты [1]. Цель данной статьи – рассмотреть методологические инновации технологий больших данных в изучении разрыва в доходах между городским и сельским населением, проанализировать ограничения их применения и предложить стратегии совершенствования, способствующие научному и уточненному характеру исследований вышеуказанного разрыва в доходах.

Технологии больших данных в изучении разрыва в доходах между городом и деревней

Методологические инновации начинаются с интеграции и анализа многомерных

данных. Традиционные исследования в основном опираются на данные обследований доходов и расходов городского и сельского населения, предоставляемые статистическими управлениями. Однако технологии больших данных позволяют интегрировать многомерные данные, такие как данные о потреблении в городах и сельской местности с платформ электронной коммерции, данные о занятости в городах и сельской местности с платформ трудоустройства, а также данные о социальном обеспечении и налогообложении в городах и сельской местности из государственных систем. При этом данные о потреблении с платформ электронной коммерции могут отражать различия в структуре потребления и уровнях доходов между городскими и сельскими жителями; данные о занятости с платформ трудоустройства могут раскрыть статус занятости и потенциальный доход городских и

сельских жителей; а данные о социальном обеспечении и налогообложении точно отражают масштаб и структуру доходов в городах и сельской местности. Интеграция и анализ этих многомерных данных позволяют более полно и глубоко охарактеризовать особенности разрыва в доходах между городскими и сельскими жителями.

При применении методов пространственного анализа используют пространственную информацию из больших данных, сочетая её с технологиями геоинформационных систем (ГИС). Это позволяет строить модель пространственного анализа разрыва в доходах между городским и сельским населением. Эта модель позволяет анализировать характеристики распределения, пространственные корреляции и пространственные эффекты перетоков разрыва в доходах между городом и деревней в географическом пространстве. Например, анализируя пространственные данные, такие как распределение точек потребления в городах и сельской местности и рабочих мест, мы можем выявить пространственные механизмы, лежащие в основе разрыва в

доходах между городом и деревней, такие как разрыв в доходах между городскими центрами и сельскими районами, а также разрыв в доходах между городскими агломерациями и сельскими окраинами. [2] Применение методов пространственного анализа открывает новые перспективы для изучения разрыва в доходах между городом и деревней, помогая понять его пространственную неоднородность и динамику развития.

Построение моделей машинного обучения, таких как модели регрессии, классификации и кластеризации, мы можем построить модели для прогнозирования и анализа разрыва в доходах между городом и деревней. Например, регрессионные модели чаще всего используются для анализа факторов, влияющих на разрыв в доходах между городом и деревней, таких как уровень образования, сфера занятости и географическое положение; модели классификации используются для стратификации городских и сельских жителей по уровню дохода, выявляя характеристики городских и сельских жителей с разным уровнем дохода; а модели кластеризации – для выявления типичных регионов и типов разрыва в доходах между городом и деревней [3]. Применение моделей машинного обучения повышает точность и эффективность исследований разрыва в доходах между городом и деревней, позволяя определять сложные закономерности и корреляции, которые трудно выявить с помощью традиционных статистических методов.

Ограничения применения технологий больших данных

Относительно ограниченные ресурсы больших данных в сельской местности являются существенным препятствием применения технологий больших данных. По сравнению с городскими районами, сельские районы характеризуются более низким уровнем цифровизации и относительно ограниченными ресурсами больших данных. Платформы электронной коммерции имеют относительно небольшую долю сельских пользователей, платформы трудоустройства не располагают достаточной информацией о занятости в сельской местности, а государственные системы содержат некоторые пробелы в данных о сельской местности. Это приводит к недостаточной репрезентативности и полноте данных о сельских районах при исследовании разрыва в доходах между

городом и деревней, что влияет на точность результатов исследования.

Кроме того, в городских и сельских районах наблюдается несогласованность стандартов сбора данных. Различные стандарты сбора данных в городских и сельских районах приводят к ограниченной сопоставимости данных. Например, данные о доходах городских жителей в основном включают заработную плату, премии и доход от бизнеса, в то время как данные о доходах сельских жителей также включают доход от сельскохозяйственного производства и трансфертный доход. Более того, статистический уровень и методы учета данных о доходах из этих различных источников различаются. Это затрудняет точное сравнение уровней и структуры доходов городских и сельских жителей при анализе разрыва в доходах между городом и деревней.

Недостаточное применение технологий анализа больших данных также является своеобразным препятствием при применении технологий больших данных. Хотя указанные технологии представляют собой мощный инструмент для изучения разрыва в доходах между жителями города и деревни, их практическое применение по-прежнему сталкивается с трудностями. С одной стороны, нехватка междисциплинарных специалистов, обладающих знаниями как в области экономики, так и в области больших данных, привела к неадекватному применению методов анализа больших данных. С другой стороны, существующие модели анализа больших данных имеют ограничения при решении сложных задач, связанных с разрывом в городских и сельских доходах, такие как трудности с обработкой нелинейных взаимосвязей данных и пространственной неоднородности.

Стратегии совершенствования

Основной стратегией развития технологий больших данных является увеличение ресурсов больших данных в сельской местности. Правительству необходимо увеличить инвестиции в сельскую цифровую инфраструктуру, содействовать развитию платформ электронной коммерции и трудоустройства в сельской местности, а также повысить возможности и качество сбора данных в сельской местности [4]. Кроме того, возможно создать платформу обмена большими данными в сельской местности для интеграции соответствующих данных по

сельскому хозяйству, сельским районам и фермерам, обеспечив богатый источник данных для исследований разрыва в доходах между городом и деревней.

Перспективной стратегией совершенствования является также унификация стандартов сбора данных по городу и деревне. Данной унификации будет способствовать разработка единых стандартов и методов расчета для сбора данных о доходах в городе и деревне, уточняющая статистический охват, источники данных и процедуры обработки. Унификация будет улучшать сравнение и проверку городских и сельских данных для обеспечения сопоставимости и согласованности. [5] Кроме того, следует создать систему оценки качества данных для регулярной оценки и мониторинга качества городских и сельских статистических данных.

Совершенствованию применения технологий анализа больших данных будет способствовать обучение исследователей-экономистов в области технологий больших данных, что позволит формировать группы междисциплинарных специалистов, обладающих опытом как в экономике, так и в области больших данных [6]. Кроме того, это будет стимулировать исследовательские институты и компании к развитию сотрудничества и разработке моделей и инструментов анализа больших данных, подходящих для исследования разрыва в доходах между городом и деревней. В совокупности это будет совершенствовать интерпретацию и валидацию результатов анализа больших данных для повышения надежности и практичности исследовательских выводов.

Заключение

Технологии больших данных привнесли методологические инновации в изучение разрыва в доходах между городским и сельским населением. Применение многомерной интеграции данных, пространственного анализа и моделей машинного обучения позволяет более точно описывать характеристики разрыва в доходах между жителями города и деревни. Однако их применение также сталкивается с ограничениями, такими как нехватка ресурсов для работы с большими данными в сельской местности, несогласованность стандартов сбора данных о городе и деревне и недостаточное применение технологий анализа больших данных.

Укрепление развития ресурсов больших данных в сельской местности, унификация стандартов сбора данных о городских и сельских районах и совершенствование применения технологий анализа больших данных позволят эффективно решать эти проблемы, способствуя научным и углублённым исследованиям разрыва в доходах между городом и деревней и обеспечивая более надёжную поддержку разработки политики скоординированного развития городов и сёл.

В заключение необходимо отметить, что применение технологий больших данных повышает востребованность в инструментариях, обеспечивающих повышение качества моделей и их адекватности существующим данным, а также выравнивание исходных данных, при отсутствии или экстремальном изменении показателей за какой-то период.

Список использованной литературы

1. Лю Фан, Ван Цян. Методологические инновации и проблемы больших данных в изучении разрыва в доходах между городом и деревней // *China Rural Economy*, 2024, № 6, с. 90-105.
2. Чэнь Лян. Пространственный анализ и механизмы формирования разрыва в доходах между городом и деревней. – Шанхай: Shanghai People's Publishing House, 2023, с. 150-165.
3. Чжоу Мин. Развитие применения машинного обучения в изучении разрыва в доходах между городом и деревней // *Quantitative Economic and Technical Economic Research*, 2024, № 4, с. 70-85.
4. Чжан Хун. Цифровая трансформация сельских районов и создание ресурсов больших данных // *Вопросы экономики сельского хозяйства*, 2023, № 11, с. 60-75.
5. Ли Цзяньго. Унификация и оптимизация стандартов сбора данных о доходах городских и сельских жителей // *Статистические исследования*, 2024, № 3, с. 45-58.
6. Ван Фан. Подготовка специалистов с комплексным потенциалом и применение больших данных в экономике // *Образование и экономика*, 2023, №5, с. 80-92.

Рецензент: д-р экономич. наук, проф. Акылбекова Н.И.