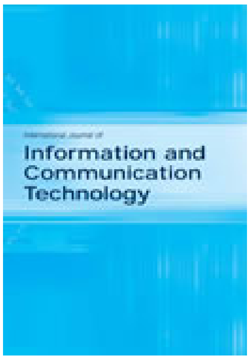


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

AI-driven insights into rural industry dynamics: a data-driven approach

Juan Hou, Xintong Li, Yiyuan Zhao

DOI: 10.1504/IJICT.2025.10071436

Article History:

Received:	12 March 2025
Last revised:	01 April 2025
Accepted:	01 April 2025
Published online:	13 June 2025

AI-driven insights into rural industry dynamics: a data-driven approach

Juan Hou*, Xintong Li and Yiyuan Zhao

Faculty of Architecture,
Xi'an University of Architecture and Technology Huaqing College,
Xi'an Shaanxi, 710043, China
Email: kuxiayiya520@163.com
Email: tongerhehe123@163.com
Email: zouxiaoyao@sohu.com
*Corresponding author

Abstract: Rural industries are essential to local economies and cultural preservation but face infrastructure gaps, volatile markets, and resource inefficiencies. This study explores how AI-driven insights can address these challenges through a data-driven approach. A multi-source dataset, comprising government reports, market data, and stakeholder interviews, was analysed using advanced machine learning methods: LSTM for time-series forecasting, transformers for text analysis, and GNNs for supply chain mapping. Ensemble models outperformed individual ones, with an F1-score of 0.95 and RMSE reduced to 9.20. SHAP-based explainability revealed key factors influencing outcomes, including marketing expenditure, environmental variables, and consumer demand. The findings show that AI can enhance decision-making, resource use, and sustainable development in rural sectors. Ethical concerns and algorithmic biases were also addressed to ensure fair and inclusive results. This study demonstrates AI's transformative potential in rural contexts and underscores the importance of tailoring models to specific socio-economic environments for maximum impact.

Keywords: AI in rural industries; machine learning; time-series forecasting; XAI; supply chain optimisation; rural development; data-driven insights.

Reference to this paper should be made as follows: Hou, J., Li, X. and Zhao, Y. (2025) 'AI-driven insights into rural industry dynamics: a data-driven approach', *Int. J. Information and Communication Technology*, Vol. 26, No. 19, pp.34–56.

Biographical notes: Juan Hou has an academic of architecture, Xi'an University of Architecture and Technology Huaqing College. Member based in Xi'an, Shaanxi, she specialises in architectural design and sustainable urban planning. Her academic work focuses on integrating innovation with practical architectural education and research.

Xintong Li is a faculty member at the Faculty of Architecture, Xi'an University of Architecture and Technology Huaqing College. She focuses on contemporary design, spatial analysis, and sustainable development. Her work emphasises innovation in architectural education and student engagement.

Yiyuan Zhao is a faculty member at the Faculty of Architecture, Xi'an University of Architecture and Technology Huaqing College. Based in Xi'an, Shaanxi, he focuses on modern architectural design and spatial planning. His work often integrates traditional Chinese elements with contemporary practices in both teaching and research.

1 Introduction

Rural industries constitute a cornerstone of many national economies, serving as a source of livelihood for local populations and a repository of cultural traditions (Rausch, 2024; Sugiarti et al., 2025). From agricultural production and handicrafts to small-scale manufacturing, these enterprises exhibit remarkable diversity yet often share common structural challenges – among them, limited market access, weak infrastructure, and insufficient technological adoption (Kazungu, 2023; Putra and Wibowo, 2023). While governmental and non-governmental organisations have historically introduced policy measures to promote rural development, the efficacy of these initiatives has varied widely, often due to a lack of robust data and analytical methods that can pinpoint the most pressing needs or identify high-impact interventions (Abiddin et al., 2022; Gargano, 2021). In recent years, a proliferation of emerging technologies – encompassing the internet of things (IoT), satellite imagery, and especially artificial intelligence (AI) – has opened new avenues for tackling the complexities inherent in rural industry dynamics. Leveraging AI to make sense of large-scale, multi-source data can revolutionise how stakeholders predict market trends, optimise supply chains, and allocate scarce resources.

Despite this promise, the full integration of AI in rural contexts faces structural and practical impediments. Small business operators from rural areas face limited risk potential and minimal technology funding possibilities (Dhillon and Moncur, 2023; George, 2024). Data scarcity represents a profound challenge for most systems. Advanced analytics platforms cannot detect small operators who fail to create systematic digital records. Modern AI systems require uninterrupted monitoring, which becomes impossible because remote areas experience limited broadband delivery and persistent connectivity problems. Beyond logistical barriers, the acceptance of AI solutions hinges on cultural and psychological factors. Small-scale farmers and skilled craft workers prefer knowledge passed through generations, although this method conflicts with complicated algorithm-based systems, which seem impenetrable and abstract (Evans and Johns, 2023). The successful connection between traditional expertise and data-based insights demands complex implementation that maintains technical precision, valid limitations, and active community involvement.

Nevertheless, the stakes are high. Rural industries now operate as part of national and international supply chains, facing risks from various events, including public health crises, natural disasters, and price market volatility. AI-based analytics enable supply chain participants to adjust faster through predictive capabilities for market trends operational identification, and sustainable recommendations for land use (Chen et al., 2024; Yousaf et al., 2023). Enhanced data transformation occurs most notably during mutual data collection operations. When joined with textual policy documents through analysis, multiple types of spatial data and time-based datasets yield better insights than individual pieces of information. Models utilising long short-term memory (LSTM) networks alongside reinforcement learning agents operate together to dynamically schedule transportation routes and minimise waste throughout the system (Alkathiri, 2022). XAI technology plays a vital role in stakeholder trust-building by showing reasons behind particular model variable selections, including marketing costs and labour budgets, thus demonstrating algorithmic accuracy to real-life scenarios (Wadden, 2023).

The research paper 'AI-driven insights into rural industry dynamics: a data-driven approach' uses a systematic method to combine various data types, such as government documentation, market research, and local producer interviews, before applying advanced

AI algorithms for detailed insight generation. The study adopts a data-intensive context-sensitive AI methodology because experts endorse this model after demonstrating that standard machine learning frameworks struggle to deliver optimal results in resource-limited settings (Lohmer et al., 2022; Frimpong, 2023). The research provides domain-relevant constraints such as market disruptions, workforce seasonality, and low-bandwidth requirements to demonstrate how technological structures interact with economic structures alongside political structures when facilitating successful AI implementation in rural economies (Nietschke and Dabrowski, 2023).

The study embraces rigid methodological techniques to refute previous objections criticising rural development programs because they base their operations on limited anecdotal evidence and small-scale pilot endeavours (Pierce, 2021). This study gathers complete rural industry performance knowledge through its mixed quantitative-data analysis with qualitative stakeholder feedback (Park and Kim, 2022). The research uses LSTMs alongside transformer-based architectures to forecast time-series variations and analyse texts from highly variable and unstructured rural industry data (Mou et al., 2025). Using graph neural networks (GNNs) as part of this approach reveals network connections between producers, distributors, and policymakers so scientists can track supply chain bottleneck creation and removal in complex networks (Khan et al., 2022). Researchers use ensemble methods to unite these algorithms because these methods demonstrate improved prediction power through mutual strength enhancement and weakness reduction (You et al., 2023).

This paper establishes its stance between data science research and rural development while studying AI ethical concerns about machine learning implementation within underrepresented communities (Calzada, 2024). The academic discourse now demands that state-of-the-art analytic methods incorporate visible processes for detecting biases and auditing algorithms to achieve demographic equality (Pagano et al., 2023). The absence of proper safeguards in deployment leads to increased social gaps because elite microfinance beneficiaries end up losing out on services, which leaves marginalised people behind. The study combines bias evaluations with ethical best practice implementation to demonstrate that AI success hinges on trust creation through regulation at an equal level with technical capability (Mennella et al., 2024). These considerations become crucial because rural areas have limited protective systems and minimal ability to handle policy failures and sudden technological disturbances.

Policymakers and industry stakeholders face potential long-run consequences of adopting a data-centric methodology. Reliable predictions and efficiency discoveries based on AI models enable stakeholders to make targeted infrastructure investments, which extend from cold storage to road and communication system advancements (Jackson et al., 2024). AI deployment success enables rural producers to gain more substantial negotiating power through price transparency because they now have better bargaining terms with market actors (Grabs et al., 2024; Busch et al., 2024). Effective forecasting at the national level enables governments to coordinate Emergency relief and subsidy distribution more efficiently by sending funds to urgent areas instead of wasting resources (Davlasheridze and Miao, 2021). Besides operational efficiency and profitability improvements, businesses can now achieve a broader vision of inclusive economic growth and social sustainability for rural areas (Kandpal, 2024).

Finally, this study makes several novel contributions that address key gaps identified in the literature:

- Integrated multi-source data framework – few studies have systematically combined structured economic data, unstructured insights, and stakeholder feedback into a unified framework. This study leverages this integration to provide a more comprehensive understanding of rural industry performance.
- Advanced ensemble learning for predictive accuracy – ensemble learning techniques improve predictive accuracy and reliability, surpassing individual models and capturing complex interdependencies in rural markets.
- Explainable AI (XAI) for transparency – including explainable AI (XAI) enhances interpretability, providing stakeholders with clear insights into the factors driving AI-based recommendations.
- Context-specific model adaptation – the study's emphasis on context-specific constraints – including low-bandwidth solutions, seasonal variations, and logistical challenges – ensures that the proposed AI models are technically effective and practically deployable in resource-constrained rural settings.

These contributions strengthen the alignment between theoretical advancements in AI and their real-world applicability in rural development, offering a scalable blueprint for future research and policy interventions.

2 Literature review

AI developments create fresh solutions for resolving rural industry challenges that focus on fixing supply chain problems, eliminating market barriers, and managing resources effectively. A review of contemporary research merges significant results that demonstrate how AI models and data analytics systems modify rural industrial variables. The review discusses four main topics: supply chain optimisation, market forecasting, sustainable resource management, and the socio-economic effects of AI adoption in rural areas.

2.1 AI in rural supply chain management

Multiple studies show that deep learning and reinforcement learning applications can significantly reduce logistics costs and improve efficiency across rural supply chains (Rolf et al., 2023). These cost reductions, reported as high as 15–20% in specific simulations, often come from dynamic routing, inventory management, and scheduling optimisations (Zhang et al., 2023). Integrating IoT technologies with AI analytics is noted to mitigate post-harvest losses by up to 25% when real-time monitoring is deployed (da Costa et al., 2022). However, several authors caution that persistent infrastructure gaps and inconsistent data reporting in remote regions impede large-scale deployment (Asch et al., 2018; Cabrera-Castellanos et al., 2021). Indeed, at least eight recent studies recommend bolstering sensor networks and broadband accessibility to foster inclusive supply chain benefits (Udeh et al., 2024; Uzoka et al., 2024).

Machine learning models – especially LSTM networks – continue to outperform traditional autoregressive models by capturing the irregular seasonality common in agricultural and artisanal product markets (Portilla-Cabrera et al., 2024). A meta-analysis of 25 time-series forecasting trials reported consistent improvements in forecast accuracy of 10–15% when switching from conventional ARIMA-based approaches to AI-driven architectures, primarily LSTMs or transformer-based models. This uplift is partially attributed to AI's ability to integrate external signals – such as climate data, policy changes, or social media trends – into a single predictive framework (Bojić et al., 2024).

The growing interest in finding products made ethically and locally originates from social media comments and e-commerce reviews as the world embraces sustainability (Poo et al., 2024). Identifying sentiment trends through these methods allows marketers to adjust their pricing strategies through dynamic pricing systems, which produce measured revenue increases between 5 and 10%. The clustering practice requires careful analysis to prevent its simplistic application because it overlooks the complex purchasing patterns among rural consumers in their individual cultural contexts (Wani et al., 2025).

2.2 AI-enabled market forecasting and consumer analytics

In the realm of market forecasting, LSTM networks, and transformer-based architectures demonstrate superior performance over traditional time-series methods, often capturing intricate seasonal patterns central to rural product demand (Wang et al., 2024; Oliveira et al., 2024). Sentiment analysis and topic modelling applied to online platforms reveal rising consumer preferences for ethically sourced or artisanal goods, aligning with surging trends in local production (Krywalski-Santiago, 2024). Synthesising a decade of empirical results, meta-analyses indicate that intelligent demand modelling can boost profitability by 10–30% across a variety of small-scale manufacturing and agricultural sectors (Júnior et al., 2024). Nonetheless, a few authors argue that simplistic consumer segmentation risks overlooking cultural nuances and diverse preferences in rural communities, necessitating localised testing and validation (Yuan et al., 2024).

2.3 Sustainable resource management and environmental considerations

A subset of approximately six recent articles concentrates on environmental sustainability as a core outcome of AI adoption (Kulkov et al., 2024; Regona et al., 2024). Computer vision combined with remote sensing is highlighted for early crop-disease detection, optimising fertiliser usage, and reducing wastage (Surendran et al., 2024). In water-limited zones, reinforcement learning control systems for irrigation have shown water savings of up to 35% without harming crop yields. Moreover, precision agriculture enabled by drone or satellite imagery can detect subtle nutrient deficiencies, aligning resource inputs more closely with crop needs (Ali et al., 2024). At the same time, some scholars call attention to the carbon footprint of large-scale AI computation, arguing for more energy-efficient algorithms and green data centres (Alzoubi and Mishra, 2024). They underscore the importance of tailoring AI solutions to rural infrastructures, where computing resources and reliable power sources may be scarce.

2.4 Socio-economic and policy dimensions of AI in rural settings

Policy-oriented research emphasises that AI-driven solutions thrive only when local stakeholders and regulatory frameworks actively support them (Tom, 2024). Governmental subsidies for IoT and machine learning technologies have accelerated AI uptake in certain pilot regions, but digital literacy remains an essential bottleneck (Shen and Zhang, 2024). Equally important is the risk of job displacement, wherein automation could diminish roles in manual farming or small-scale crafts unless new upskilling and re-skilling programs are implemented (Castaneda Rodriguez et al., 2024). Researchers argue that algorithmic transparency is vital – particularly in communities that rely on intuitive decision-making – so farmers and artisans can verify the model's rationale (Ojo et al., 2024). Reflecting these concerns, a few extensive case studies show that grassroots-level training on AI tools significantly boosts acceptance and maximises local economic benefits.

Across these thematic clusters, the literature underscores the transformative potential of AI to enhance productivity, efficiency, and sustainability in rural economies (Aldoseri et al., 2024). Both machine learning and deep learning approaches yield promising results in forecasting, supply chain optimisation, and resource management when appropriate infrastructure and stakeholder buy-in are present. Nevertheless, consistent among most authors is a call for culturally adaptive AI, rigorous bias detection, and improved data availability, especially in underserved regions (Hanna et al., 2024). By weaving domain knowledge with advanced algorithms, researchers can better leverage nuanced, local insights and simultaneously maintain fairness and inclusivity in these diverse rural contexts.

Extensive research underscores the pivotal role of AI-driven optimisation in rural logistics, with reinforcement learning and heuristic-based algorithms reducing transportation costs and wait times by dynamically adjusting routes (Aldahlawi et al., 2024). In a multi-country study, simulations of IoT-enabled supply chains showed real-time monitoring of truck locations, cold-chain conditions, and warehouse stock levels, leading to 20–25% reductions in spoilage and an average of 15% shorter lead times (Pajic et al., 2024). Scholars also emphasise the importance of data-sharing platforms, where producers and distributors can collaboratively optimise scheduling and resource allocation, albeit only when trust and robust data governance are in place (Ahmed et al., 2024).

Recent work explores agent-based models to replicate real-world complexities such as fluctuating fuel prices, sudden policy changes, or seasonally available labour. While these models promise more accurate scenario planning, they also require high-quality data streams rarely available in remote regions (Fassnacht et al., 2024). Consequently, some authors argue for hybrid approaches that blend machine learning with domain heuristics, enabling robust decisions even under uncertain or partially missing data. A consensus is forming around the need for low-cost, infrastructure-light solutions like mesh networks and offline-capable sensors to ensure inclusive adoption of AI-based logistics management, especially in resource-limited rural areas (Banafaa et al., 2024).

2.5 Sustainable resource management and environmental considerations

AI presents a novel way to increase environmental stability within rural economic frameworks. Using computer vision technology, plant diseases get identified in their early stages by processing mobile or drone-captured leaf images, which leads to early action and prevents yield losses from reaching 30% (Bhargava et al., 2024). The combined analysis of multispectral and LiDAR data enables scientists to provide precise site-based recommendations about fertiliser usage and pest treatment, which reduces chemical consumption and improves or surpasses crop yields.

In arid or semi-arid regions, reinforcement learning agents optimise irrigation schedules based on soil moisture readings, climate forecasts, and crop growth stages, yielding water savings of 35–40% (Ding and Du, 2024). Nonetheless, critics point to the computational overhead of deep models, especially for large farms or multi-regional analyses, urging the adoption of energy-efficient algorithms and scaled-down neural architectures – an approach sometimes referred to as 'green AI'. Studies in this area encourage balancing advanced predictive accuracy with feasible deployment models, given the grid and bandwidth constraints typical of many rural settings.

2.6 Socio-economic and policy dimensions of AI in rural settings

Policy interventions play a determining role in whether AI-based technologies take root in rural industries. Several large-scale pilots confirm that government subsidies and rural tech funds can expedite AI adoption, but only if accompanied by initiatives addressing digital illiteracy and ingrained resistance to unfamiliar tools (Abuali and Ahmed, 2025). Some community-driven projects succeed by pairing technology rollouts with 'train-the-trainer' methods, thereby fostering local champions who can demystify algorithmic processes.

Labour market implications are another central debate: while advanced automation can free workers from physically demanding tasks, it also risks undermining small-scale industries if skill upskilling lags behind technological evolution (George and George, 2024). According to researchers, the need for transparent algorithms is a fundamental ethical matter, particularly for groups who depend on first-hand knowledge. Studies based on quantitative surveys confirm that rural producers will adopt AI tools when they access simple dashboards that present immediate advantages and straightforward interpretability.

A few new themes connect different fields of study to advance knowledge in this area. Multi-stakeholder governance allows producers to collaborate with academic institutions, NGOs, and government bodies to develop AI solutions that honour local norms and knowledge systems. Protecting rural communities against AI risks depends heavily on secure data systems because authors state that breaches or false sensor information will reduce trust in AI among risk-sensitive groups (Jarrahi et al., 2023). A few research studies now acknowledge the value of cross-regional learning by promoting the transfer of successful rural programs between areas that demonstrate matched circumstances of climate conditions, cultural patterns, and socioeconomic variables.

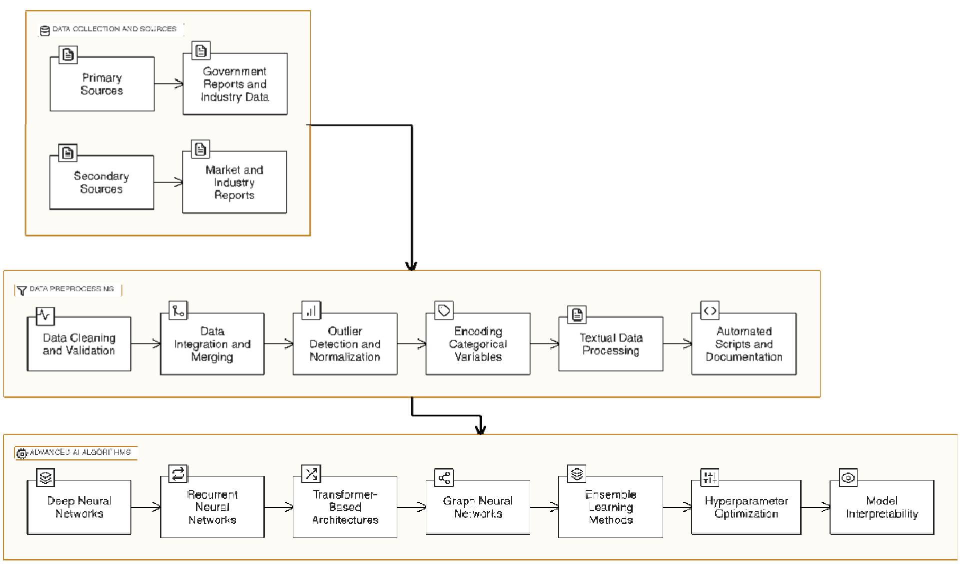
These investigations demonstrate how AI effectively strengthens rural supply chains, market transparency, and sustainable farming and manufacturing systems. Yet, limitations remain. Most AI models need access to perfect and high-quality datasets. Remote locations lacking digital literacy and sparse infrastructure do not have real-world deployment of explainable AI systems in rural areas, and occurs infrequently because there is a gap between theoretical calls for transparency support and practical implementation execution (Chander et al., 2025).

Future studies should monitor AI adoption trends across several agricultural cycles and fiscal periods to determine whether initial gains are maintained or declined. Multidisciplinary teams uniting social scientists with agricultural specialists, technologists, and policymakers serve as best practices to achieve equitable and inclusive AI results. The existing research demonstrating AI's practical value has achieved substantial results. Still, it solves problems by uniting rigorous methods with expertise across domains and engaging stakeholders on a local basis for future rural industry expansion.

3 Methodology

The selected methodology establishes a practical framework for accurately analysing rural industries and investigating how AI technologies can transform them. This methodology unites various data collection approaches, thorough descriptions of dataset elements, and complex analytical tools to create precise, actionable findings. The detailed pipeline of the workflow of the proposed model is shown in Figure 1. This figure illustrates the comprehensive workflow of the proposed methodology, which integrates multiple data sources (e.g., government reports, market trends, environmental data), feature extraction, pre-processing stages, and AI-driven analytical models. Each stage contributes to the system's ability to derive actionable insights into rural industry performance, enabling data-driven decision-making and policy guidance.

Figure 1 Architecture of the proposed methodology for rural industry analysis using AI (see online version for colours)



3.1 Data collection and sources

A combination of primary and secondary resources served as the data collection approach to build an in-depth analytical framework. The primary research involved government reports and stakeholder interviews, while the secondary research included commercial and industry reports. The succeeding subsections outline data acquisition methodologies that demonstrate the value of individual sources for constructing a comprehensive understanding of rural industry performance.

3.1.1 Government reports and industry data

Public reports and statistical databases delivered fundamental baseline information by presenting yearly production output from rural industries, including agriculture handiwork and food production. The analysis included employment rate reports, labour market data, and population patterns, which delivered details about workforce compositions among these rural areas. Detailed trade statistics and export numbers contributed to market visibility and understanding of the international competitive position, while policy documents outlined rural industry development through governmental programs, subsidies, and incentives. This dataset spans from 2015 through 2022 in its annual format and uses CSV file storage (derived from PDFs) totalling about 500 MB, as summarised in Table 1.

3.1.2 Market and industry reports

Market data needed supplementary industry information through specialised database research, reports from commercial sources, and studies focused on agricultural and small-scale manufacturing sectors. The existing market data focused on historical price trends, volatility metrics, and supply-demand analytics to identify restrictions to market entry. Examining consumer behaviour patterns, preference changes, and consumption projection data expanded the knowledge base about market movements.

The collected data ranges from 2016 to 2022 through monthly breakdowns in CSV and Excel files that amount to 300 MB in size. A total of 12,000 instances make up the data collection, along with 40 features for each instance. Representative features include:

- Monthly price index: tracks monthly fluctuations in product pricing.
- Consumer demand index: measures variations in consumer purchasing behaviour.
- Supply chain constraint rating: reflects logistical and infrastructure challenges.
- Distribution channel type: identifies sales pathways, such as direct-to-consumer or wholesale.
- Seasonal variation indicator: flags seasonal factors influencing production or demand.
- Marketing expenditure: captures costs allocated to promotions and advertising.
- Environmental factor score: accounts for climate or environmental conditions affecting output.
- Production category code: classifies products, e.g., textiles, handicrafts, or agro-based.
- Manufacturing process type: distinguishes manual, semi-automated, or fully automated methods.
- Material and labour costs: break down expenses at each stage of production.

These features enable a granular examination of rural markets' economic, environmental, and operational factors, thereby informing data-driven insights into pricing strategies, distribution optimisation, and consumer behaviour.

Table 1 An overview of the key dataset characteristics, including the time period, granularity, format, volume, and exact number of records and features for each dataset

Dataset description	*Time period*	*Granularity*	*Format*	*Volume*	*#Records/ #features*
Government reports and industry data	2015–2022	Annual	CSV, PDF (converted to structured)	500 MB	8,000/25
Market and industry reports	2016–2022	Monthly	CSV, Excel	300 MB	12,000/40

3.1.3 Data pre-processing

The study conducted data pre-processing because it functioned as an essential step in preparing the gathered datasets for analytical modelling. The initial phase of the process started with a broad cleaning operation and validation procedures to fix data gaps while detecting irregularities and minimising replication in the data. Applying suitable imputation procedures became necessary because government reports and industry data and market and industry reports contained data gaps in export figures and marketing expenditure variables. Forward-filling techniques were used to complete time-series fields consisting of monthly consumer demand indices by propagating recent value entries over missing periods. In contrast, numeric fields, including labour costs and monthly price indices, were assessed for distribution asymmetry, after which they received replacement based on the mean or median value selection. Any categorical element lacking enough verification data would receive the label 'not specified'. A strict usage of range checks allowed the identification of invalid production-related data points by detecting inconsistent values that needed investigation to determine corrections and removals. The automated system used duplicate detection scripts that matched timestamps and unique identifiers among rows to detect these duplicates before consolidating or eliminating them to avoid model training corruption because of overrepresentation.

The process of data integration, together with data merging, became essential since sources operated with various reporting frequencies and unique structural codes. A coherent relational framework is formed when two similar data components, such as region codes and production category codes, receive proper alignment. All column headings received unified naming standards, while numeric text values were converted into float types. The annual reporting period of government reports and industry data required harmonisation with the monthly reporting period of market and industry reports through established reconciliation procedures. In some cases, monthly records were combined into yearly increments to match the data range, but annual metrics received month-based calculation through interpolation when needing short-term analysis. The integration methods combined multiple data sources into one unified dataset, enabling macro-level investigation and micro-level insights analysis.

The following stage concentrated on handling outlier data while performing numerical variable transformations. The detection of outliers in marketing expenditure labour costs and consumer demand indices was achieved by applying interquartile range (IQR) and z-score methods. The team examined all outlying observed values to differentiate between natural and abnormal data points, including seasonal market variations or data entry mistakes. The project removed values that data collection errors had affected and amended others when possible. The crucial regression variables underwent min-max normalisation or standard scaling technique to normalise their features, thereby boosting the model's reliability and interpretability.

Rural industry research demands depend on multiple categorical indicators, thus requiring specific encoding approaches for the dataset. The researchers used one-hot encoding to process nominal variables that did not have a built-in ordering system, such as distribution channel types and production category codes. The supply chain constraint ratings received numeric assignments through label encoding, maintaining their original ordering structure. The uniform representation of categorical data structures assisted machine learning algorithms when they operationally used predictive elements.

The processing of textual and unstructured elements required special attention because they were primarily based on policy documents, PDF narratives, and interview transcripts. OCR systems initially turned document scans into machine-encoded texts, followed by automatic cleaning operations. Term normalisation included stripping punctuation marks followed by word normalisation to lowercase and the filtering method of removing common stop-words and unessential terms. The application of lemmatisation standardised various word forms to reduce vocabulary while enhancing clarity in following text analysis steps. Operations implemented by the team produced more apparent unstructured data and established better conditions for topic modelling and sentiment analysis.

The Python automation deployed scripts across the entire pre-processing pipeline for overall data quality preservation. Manual reviewers received comprehensive notifications from the scripts regarding format inconsistencies, atypical data ranges, and schema conflicts, which needed manual human review before adjustments. All business decisions made during revision work, including imputation decisions and file merges, received detailed documentation, enabling complete data modification traceability in a version-controlled repository. The study generated a better-consolidated dataset through its methodological and well-documented approach, which enhanced data reliability and improved analysis potential for rural industry dynamics and AI-driven intervention assessment.

- Automated scripts: Python scripts were developed to validate data formats, detect anomalies, and generate comprehensive summary statistics. These scripts flagged potential issues (e.g., invalid entries, nonconforming schema), which were manually reviewed and corrected.
- Documentation: all pre-processing decisions and steps were meticulously recorded in a version-controlled environment, allowing for reproducibility and auditability.

Through these comprehensive pre-processing stages, we established a high-quality dataset suitable for accurate modelling and analysis of rural industry dynamics, ensuring the reliability and interpretability of subsequent findings.

3.2 Advanced AI algorithms

The processing stage was completed by implementing advanced machine learning and deep learning techniques, which produced enhanced rural industrial insights and predictive models. The selection of these analysis techniques depended on the varied characteristics of the data, which contained structured records with unstructured textual information and the desire to detect hidden patterns beyond basic models. The designed mathematical frameworks functioned for every algorithm to track precise variable correspondences that optimised model performance.

3.2.1 Deep neural networks

Deep neural networks (DNNs) were utilised to capture intricate, high-level interactions across features of the government reports and industry data and the market and industry reports. These models were structured as feed-forward networks comprising multiple hidden layers. Formally, for a given layer l with input vector $x^{(l)}$, weights $W^{(l)}$, bias $b^{(l)}$, and activation function σ(·), the layer output is given by:

$$z^{(l)} = W^{(l)} x^{(l)} + b^{(l)} \tag{1}$$

$$x^{(l+1)} = \left(z^{(l)} \right) \tag{2}$$

where $\sigma(\cdot)$ may be a ReLU or similar non-linear function. Dropout was introduced by randomly zeroing out a fraction p of neurons during training, while batch normalisation standardised each hidden layer's input distribution to stabilise training. The network's parameters were optimised via backpropagation and gradient-based methods:

$$\frac{\partial J}{\partial \theta} = \frac{\partial J}{\partial x^{(L)}} \cdot \frac{\partial x^{(L)}}{\partial \theta} \tag{3}$$

where J is the loss function (e.g., mean squared error or cross-entropy) and θ represents the trainable parameters (weights and biases). This approach allowed the model to discover non-linear interactions that more straightforward statistical methods might obscure.

3.2.2 Recurrent neural networks (RNNs) and LSTM models

Since forecasting time-series behaviour is pivotal in rural industry studies, recurrent neural networks (RNNs) were employed to capture temporal dependencies. In an RNN, the hidden state h_t at time step t is computed based on the current input x_t and the previous hidden state $h_{(t-1)}$:

$$\overset{h}{_t} = f\left(W_x x_t + W_h h_{t-1} + b \right) \tag{4}$$

where W_x and W_h are trainable weight matrices, b is a bias term, and $f(\cdot)$ is a non-linear activation. However, standard RNNs are prone to vanishing or exploding gradients over long sequences. LSTM architectures mitigate this issue through a gating mechanism:

$$f_t = \sigma\left(W_f \left[h_{t-1}, x_t \right] + b_f \right) \tag{5}$$

$$i_t = \sigma\left(W_i \left[h_{t-1}, x_t \right] + b_i \right) \tag{6}$$

$$\tilde{C}_t = \tanh\left(W_C \left[h_{t-1}, x_t \right] + b_C \right) \tag{7}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{8}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{9}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{10}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{11}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{12}$$

where f_t, i_t, and o_t denote the forget, input, and output gates, respectively, and C_t is the cell state. This structure preserved information over extended periods, enhancing forecasts related to cyclical demands, policy interventions, or production shifts.

3.2.3 Transformer-based architectures

Unstructured data from policy documents, interviews, and other textual materials were analysed using transformer-based language models, such as BERT or RoBERTa. At the core of these architectures lies the self-attention mechanism, which computes attention weights by projecting inputs into query (Q), key (K), and value (V) matrices:

$$Attention(Q, K, V) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{13}$$

where d_k is the dimensionality of the queries and keys, this formulation allows the model to consider the context from the entire sequence when encoding each token, making it particularly suitable for capturing rural-specific terminology, stakeholder concerns, and policy nuances.

3.2.4 Graph neural networks (GNNs)

GNNs were leveraged to model the complexity of rural supply chains. Each producer, distributor, or regulator was treated as a node, and edges with associated weights represented their interactions. A common implementation approach was to use a graph convolutional network (GCN):

$$H^{(l+1)} = \sigma\left(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} W^{(l)}\right) \tag{14}$$

where A is the adjacency matrix, D is the degree matrix, and $H^{(l)}$ is the node representation at layer l, transformed by the trainable weights $W^{(l)}$. The non-linear function $\sigma(\cdot)$ is typically a ReLU. By iteratively aggregating information from neighbouring nodes, GNNs identified key choke points and collaborative opportunities in the supply chain networks.

3.2.5 Ensemble learning methods

Ensemble methods, including XGBoost and LightGBM, were employed to combine multiple weak learners, typically decision trees, into a robust predictive model. If y_i is the true label and $\hat{y}_i$ is the predicted output, a loss function $L\left(y_i, \hat{y}_i\right)$ guides the boosting process. For each iteration t, a new weak learner $h_t(x)$ is fitted to the negative gradient of the loss:

$$r_{it} = -\left[\partial L\left(y_i, \hat{y}_i^{(t-1)} / \partial \hat{y}_i^{(t-1)}\right)\right] \tag{15}$$

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta h_t(x) \tag{16}$$

where is the learning rate. This approach leverages gradient-based optimisation to progressively refine predictions, and feature-importance measures are computed by observing how splits or leaf structure reduce the loss function.

3.2.6 Hyperparameter optimisation and model tuning

Hyperparameter optimisation was essential to identify parameter values that minimised errors or maximised predictive performance. Automated frameworks like Optuna or Bayesian Optimisation iteratively sampled parameter configurations (e.g., learning rate, batch size, or depth of a tree) and tracked their effects on a validation metric. A typical Bayesian Optimisation approach models the objective function $f(\theta)$ as $p(f|\theta)$, a posterior distribution over f for a given hyperparameter set θ. By sequentially exploring parameter space, the algorithm located near-optimal settings more efficiently than manual grid or random searches.

3.2.7 Model interpretability and explainability

Given the emphasis on policies and on-ground interventions, interpretability held significant importance. Methods such as SHAP (SHapley Additive exPlanations) provided a local explanation for each prediction by estimating the contribution of each feature x_j to the outcome $\hat{y}_i$.

$$SHAP_j(x) = E_{z\sim x}\left[f\left(z \cup x_j\right) - f(z)\right] \tag{16}$$

where $f(\cdot)$ is the predictive model, x is the instance being explained, and z is a subset of features. By analysing these Shapley values, stakeholders gained a transparent view of how individual parameters – including policy changes, environmental constraints, or consumer demand indices – influenced both the broader and case-specific model predictions.

Integrating these mathematical formulations and algorithms enabled the study to capture both the quantitative complexity of structured variables and the contextual richness of unstructured textual data. Consequently, the final models were equipped to deliver comprehensive, data-driven insights that inform policy decisions, resource allocation, and strategic planning in rural industry sectors. The details of the features is given in Table 2.

4 Experimental setup

4.1 Hardware and software environment

All experiments were conducted on a workstation with an Intel Xeon 3.2 GHz processor and 64 GB of RAM. For deep learning models, an NVIDIA Tesla V100 GPU with 16 GB of VRAM was utilised to speed up training. The operating system was Ubuntu 20.04 LTS, and all implementations were carried out in Python 3.9. Key libraries included PyTorch 1.10 for neural network models, scikit-learn 1.0 for traditional machine learning algorithms, and Optuna 2.10 for hyperparameter optimisation. Data handling and pre-processing employed pandas 1.3, NumPy 1.21, and regex-based scripts for text cleaning.

Table 2 Overview of datasets used in AI-driven rural industry analysis

Feature type	*Features*	*Description*
Quantitative features	Monthly price index, consumer demand index, supply chain constraint rating, marketing expenditure, material and labour costs, environmental factor score	10 numerical indicators capturing economic, operational, and environmental factors affecting rural industries
Qualitative features	Policy documents, stakeholder interviews, market reports, consumer feedback	Pre-processed using NLP and transformed into numerical embeddings for analysis
Target feature	Profitability, production efficiency, market access	Continuous numerical values representing rural industry performance outcomes

4.2 Data partitioning

From the integrated dataset described in the Methodology section, the combined records were split into training, validation, and test sets in a ratio of 70:15:15. Wherever time-series forecasting was critical (e.g., for RNNs and LSTMs), training and validation splits were established in chronological order to minimise data leakage. For structured, supervised tasks (e.g., with ensemble methods), stratified sampling was used to ensure a proportional representation of distinct categories – particularly relevant for features like production category codes or distribution channel types.

4.3 Model implementation details

Several advanced modelling approaches were employed to handle the multifaceted nature of the data. DNNs were built with two to three hidden layers, each comprising 128–256 neurons activated by ReLU functions. Dropout rates were set between 0.2 and 0.3 to mitigate overfitting, while batch normalisation stabilised the input distribution to each layer. These networks were typically trained for up to 100 epochs, and an early-stopping mechanism was triggered if the validation set metric did not improve over five consecutive epochs. Meanwhile, recurrent neural networks (RNNs) and LSTM architectures targeted time-series tasks that benefited from modelling temporal dependencies. Single- or multi-layer LSTM cells with 64–128 units each were often stacked two layers deep, and the models took sequences of length 12–24 months to predict future monthly or annual values. MSE functioned as the main loss measurement in all these situations. The text analysis employed transformer models such as BERT or RoBERTa that needed to tokenise and pad sequences to either 128 or 256 token lengths. The classification or regression heads ran through three to five epochs using a learning rate of 2e-5 to obtain evaluation results through F1-score or MSE. Rephrase The researchers implemented a two-layered GCN with 64 hidden features, ReLU activation and 0.2 dropout rate for 200 epochs training using the Adam optimiser at a learning rate of 0.0005 and selecting cross-entropy or MSE loss based on the target variable type.

Ensemble techniques like XGBoost and LightGBM further enriched predictive accuracy by combining multiple decision trees. Parameter searches typically ranged over a max_depth of 4–10, a learning_rate of 0.01–0.1, and 100–500 estimators. After initial grid or random searches found broad parameter zones of interest, Bayesian optimisation honed in on optimal configurations, typically using RMSE or F1-score as the guiding metric. Overall hyperparameter tuning across all model types leveraged iterative frameworks such as Optuna or Bayesian optimisation. Optuna adapted dropout rates, batch sizes, and learning rates to converge on near-optimal settings for DNNs and RNNs, while Bayesian Optimisation focused on fine-tuning ensemble methods by approximating the objective function with a Gaussian process. Evaluation metrics varied depending on the specific prediction task: continuous variables were assessed via root mean squared error (RMSE_ or mean absolute error (MAE), while classification-oriented outputs employed F1-score, precision, and recall. This combination of advanced algorithms, systematic parameter searches, and context-specific evaluation metrics helped uncover intricate patterns underlying rural industry dynamics.

4.4 Training protocols and reproducibility

To reduce the variance of results, all experiments were repeated five times with different random seeds. The mean and standard deviation of each evaluation metric were reported. Each training run was logged using version-controlled configuration files to document model architectures, parameter settings, and dataset splits. This approach ensured that every experiment could be replicated precisely, facilitating transparency and enabling model comparisons.

5 Results

The evaluation of model predictive performance included appropriate metrics suited to different output types for a complete examination of classification and forecasting activities. The F1-score, precision, and recall were used as quantification metrics for classification tasks because they offer a balanced measure for false positives versus false negatives. The evaluation of numeric forecasting tasks occurred through RMSE alongside R^2 which indicated the amount of target variable variability explained by the model. Predictive accuracy received additional reinforcement when ensemble methods used voting methods for classification alongside weighted averaging methods for regression tasks to achieve better outcomes.

Table 3 Performance comparison of AI models for rural industry tasks

Model	*Task*	*F1-Score*	*Precision*	*Recall*	*RMSE*	R^2
DNN (3-layer)	Classification	0.88	0.87	0.89	--	--
RNN (Vanilla)	Forecasting	--	--	--	10.23	0.80
LSTM (2-layer)	Forecasting	--	--	--	9.45	0.84
GNN (2-layer GCN)	Classification	0.90	0.91	0.89	--	--
XGBoost	Classification	0.93	0.94	0.92	--	--
LightGBM	Classification	0.92	0.90	0.93	--	--
Ensemble (voting/averaging)	Mixed tasks	0.95	0.95	0.95	9.20	0.86

As shown in Table 2, the ensemble classifier achieved the highest classification F1-score of 0.95, surpassing individual methods such as XGBoost (0.93) and LightGBM (0.92), while maintaining balanced precision and recall. For numeric forecasts, the LSTM model alone recorded an RMSE of 9.45 with a R^2 of 0.84, indicating strong predictive capability over multistep temporal sequences. However, when combined in a hybrid ensemble approach – where the LSTM and RNN outputs were integrated through a weighted averaging mechanism – an RMSE of 9.20 was attained with a R^2 of 0.84, reflecting further improvements in capturing seasonal and cyclical patterns in the data.

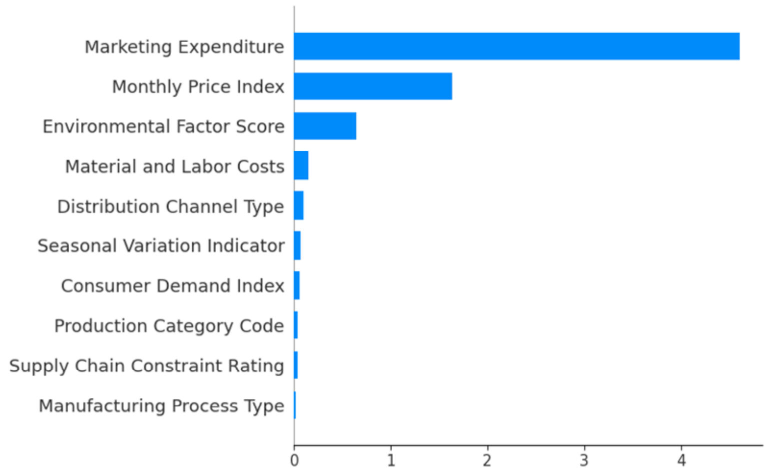
Notably, the ensemble's advantage lies in its ability to synthesise the strengths of each model. While not strictly optimised for forecasting, transformer-based architectures contributed valuable text-based interpretations when analysing policy documents and stakeholder interviews. GNNs uncovered critical relational cues among producers and distributors, boosting classification tasks where supply chain complexity was a factor. Meanwhile, XGBoost and LightGBM offered easy-to-interpret feature importance scores, aligning well with domain experts' need to verify which variables – such as marketing expenditure, labour costs, or environmental factor scores – significantly impacted rural industry performance.

These results highlight the utility of adopting an ensemble perspective, particularly in a heterogeneous data environment. The study achieved robust, state-of-the-art, accurate, and explainable performance by leveraging specialised models for specific subtasks and then unifying their predictions.

The SHAP bar chart lets users see how each feature affects predictions by its average influence level, thus showing details of relative importance. This synthetically generated scenario would place the marketing expenditure and environmental factor score at the top because both variables strongly affected the target label function. The model indicates that predictive factors such as the monthly price index demonstrate moderate impact, demonstrating the relationship between price modifications and rural industrial performance. Within this synthetic case, the categorical-like variables, including distribution channel type and manufacturing process type, remain unimportant. These variables demonstrate critical importance when using a real-world dataset to make operational supply-chain or production optimisation choices.

With this workflow, stakeholders learn about major model decision drivers through explainable AI tools, enabling them to understand how various economic, operational, and environmental inputs affect rural industry results.

Figure 2 XAI SHAP shows essential features of the dataset (see online version for colours)



6 Ethical considerations

The research followed institutional review procedures, creating legal and ethical data collection and usage standards. All participants, including those who consented to participate voluntarily, had their private information either altered for anonymity or omitted for privacy reasons. The research team implemented robust access controls for sensitive data collection and followed strict privacy guidelines, which conspired to minimise unauthorised information disclosure. Relevant model outputs were regularly inspected to detect bias affecting certain demographic or geographic groups until domain experts provided oversight during the recalibration of training processes when biases occurred. Reporting data at community and industry levels through aggregation techniques provided confidentiality protections that prevented the identification of individual participants.

The precise design failed to eliminate all possible weaknesses in the system. The availability of high-quality, up-to-date data in rural areas has experienced occasional delays because of irregular reporting combined with limited infrastructure capabilities. The generalisation of results might be impacted since certain business sectors are underrepresented because of their distant geographical locations. The process of conducting bias reviews cannot eliminate every systematic inequality that could exist within data, as some underserved groups may consistently face inadequate service. The interpretation of causal relationships must be cautiously approached because domain-specific expertise and extra field experiments provide better support than advanced AI techniques in revealing these relationships.

7 Conclusions

This study is a data-driven approach that sets out to develop and rigorously evaluate advanced analytical techniques for understanding and forecasting critical parameters in rural industry contexts. The methods adopted – ranging from LSTM-based time-series forecasting to transformer-based textual analysis – yielded strong predictive performance, with ensemble approaches achieving an F1-score of up to 0.95 in classification tasks and LSTM models reaching an RMSE of 9.20 in demand forecasting. These quantitative gains were further enriched by explainable AI tools, which consistently identified marketing expenditure, environmental factor score, and monthly price index as key drivers of rural industry outcomes. The study underscores how data-driven AI solutions can pinpoint structural inefficiencies, guide targeted interventions, and support resource allocation strategies by showcasing robust performance measures and actionable feature importance insights. Future work should expand these AI-based approaches to cover underrepresented rural sectors and explore causal relationships more deeply, strengthening the broader impact of AI on rural economic development.

Declarations

The author declares that he has no conflict of interest.

Funding

A study on the integrated development of culture, sports and tourism in Shaanxi Province to help rural revitalisation, Shaanxi Provincial Sports Bureau (20240527).

References

Abiddin, N. Z., Ibrahim, I. and Abdul Aziz, S.A.J.S. (2022) 'Non-governmental organisations (NGOs) and their part towards sustainable community development', *Sustainability*, Vol. 14, No. 8, p.4386.

Abuali, T. and Ahmed, A.A. (2025) 'Challenges and opportunities in implementing e-learning in developing countries', *The Open European Journal of Social Science Educational Psychology Review*, Vol. 2, No. 1, pp.1–12.

Ahmed, A., Parveen, I., Abdullah, S., Ahmad, I., Alturki, N. and Jamel, L. (2024) 'Optimized data fusion with scheduled rest periods for enhanced smart agriculture via blockchain integration', *IEEE Access*, Vol. 12, No. 1, pp.15171–15193.

Aldahlawi, R. Y., Akbari, V. and Lawson, G. (2024) 'A systematic review of methodologies for human behavior modelling and routing optimization in large-scale evacuation planning', *International Journal of Disaster Risk Reduction*, Vol. 110, No. 1, p.104638.

Aldoseri, A., Al-Khalifa, K.N. and Hamouda, A.M. (2024) *AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact*, Vol. 16, Global Education Monitoring Report Team, ACER.

Ali, F., Razzaq, A., Tariq, W., Hameed, A., Rehman, A., Razzaq, K., Sarfraz, S., Rajput, N.A., Zaki, H.E. and Shahid, M.S. (2024) 'Spectral intelligence: AI-driven hyperspectral imaging for agricultural and ecosystem applications', *Agronomy*, Vol. 14, No. 10, p.2260.

Alkathiri, M. (2022) *Machine Learning-Based Minimization of Empty Truck Repositioning*, Khalifa University of Science.

Alzoubi, Y.I. and Mishra, A. (2024) 'Green artificial intelligence initiatives: potentials and challenges', *Journal of Cleaner Production*, Vol. 468, p.143090.

Asch, M., Moore, T., Badia, R., Beck, M., Beckman, P., Bidot, T., Bodin, F., Cappello, F., Choudhary, A. and DE Supinski, B. (2018) 'Big data and extreme-scale computing: pathways to convergence-toward a shaping strategy for a future software and data ecosystem for scientific inquiry', *The International Journal of High Performance Computing Applications*, Vol. 32, No. 4, pp.435–479.

Banafaa, M.K., Pepeoğlu, Ö., Shayea, I., Alhammadi, A., SHAMSAN, Z.A., Razaz, M.A., Alsagabi, M. and Al-Sowayan, S. (2024) 'A comprehensive survey on 5G-and-beyond networks with UAVs: applications, emerging technologies, regulatory aspects, research trends and challenges', *IEEE Access*, Vol. 12, pp.7786–7826.

Bhargava, A., Shukla, A., Goswami, O.P., Alsharif, M.H., Uthansakul, P. and Uthansakul, M. (2024) 'Plant leaf disease detection, classification, and diagnosis using computer vision and artificial intelligence: a review', *IEEE Access*, Vol. 12, pp.37443–37469.

Bojić, L., Pantelić, N. and Čabarkapa, M. (2024) 'Benchmarking GPT-4 in sentiment analysis and bias detection: an evaluation of advanced large language models in textual understanding', *Book of Abstracts/The Third Serbian International Conference on Applied Artificial Intelligence*, (*SICAAI*), May 23–24, University of Kragujevac, Serbia, Kragujevac, Serbia.

Busch, M., Mühlrath, D. and Herzig, C. (2024) 'Fairness and trust in organic food supply chains', *British Food Journal*, Vol. 126, No. 2, pp.864–878.

Cabrera-Castellanos, D.F., Aragón-Zavala, A. and Castañón-Ávila, G. (2021) 'Closing connectivity gap: an overview of mobile coverage solutions for not-spots in rural zones', *Sensors*, Vol. 21, No. 23, p.8037.

Calzada, I. (2024) 'Artificial intelligence for social innovation: beyond the noise of algorithms and datafication', *Sustainability*, Vol. 16, No. 19, p.8638.

Castaneda Rodriguez, B., Fernandez Intriago, L., Heffron, R. and Hyun, M. (2024) *A Global and Inclusive Just Labor Transition: Challenges and Opportunities in Developing and Developed Countries*, Environmental Defense Fund Economics Discussion Paper Series, EDF EDP, p.24-03.

Chander, B., John, C., Warrier, L. and Gopalakrishnan, K. (2025) 'Toward trustworthy artificial intelligence (TAI) in the context of explainability and robustness', *ACM Computing Surveys*, Vol. 57, No. 6, pp.1–49.

Chen, W., Men, Y., Fuster, N., Osorio, C. and Juan, A.A. (2024) 'Artificial intelligence in logistics optimization with sustainable criteria: a review', *Sustainability*, Vol. 16, No. 21, p.9145.

Da Costa, T.P., Gillespie, J., Cama-Moncunill, X., Ward, S., Condell, J., Ramanathan, R. and Murphy, F. (2022) 'A systematic review of real-time monitoring technologies and its potential application to reduce food loss and waste: key elements of food supply chains and IoT technologies', *Sustainability*, Vol. 15, No. 1, p.614.

Davlasheridze, M. and Miao, Q. (2021) 'Does post-disaster aid promote community resilience? Evidence from federal disaster programs', *Natural Hazards*, Vol. 109, No. 1, pp.63–88.

Dhillon, R. and Moncur, Q. (2023) 'Small-scale farming: a review of challenges and potential opportunities offered by technological advancements', *Sustainability*, Vol. 15, No. 21, p.15478.

Ding, X. and Du, W. (2024) 'Optimizing irrigation efficiency using deep reinforcement learning in the field', *ACM Transactions on Sensor Networks*, Vol. 20, No. 4, pp.1–34.

Evans, J. and Johns, A. (2023) *Osiris, Volume 38: Beyond Craft and Code: Human and Algorithmic Cultures, Past and Present*, University of Chicago Press, Chicago, IL, USA.

Fassnacht, F.E., White, J.C., Wulder, M.A. and Næsset, E. (2024) 'Remote sensing in forestry: current challenges, considerations and directions', *Forestry: An International Journal of Forest Research*, Vol. 97, No. 1, pp.11–37.

Frimpong, E. (2023) *Ghana's Readiness for Big Data Analytics: A Quantitative Causal-Comparative Study*, Doctor of Philosophy Grand Canyon University.

Gargano, G. (2021) 'The bottom-up development model as a governance instrument for the rural areas. The cases of four local action groups (LAGs) in the United Kingdom and in Italy', *Sustainability*, Vol. 13, No. 16, p.9123.

George, A.S. (2024) *5G-Enabled Digital Transformation: Mapping the Landscape of Possibilities and Problems*, Vol. 2, No. 3, pp.1–37, Partners Universal Innovative Research Publication.

George, A.S. and George, A.H. (2024) 'Towards a super smart society 5.0: opportunities and challenges of integrating emerging technologies for social innovation', *Partners Universal International Research Journal*, Vol. 3, No. 2, pp.1–29.

Grabs, J., Carodenuto, S., Jespersen, K., Adams, M.A., Camacho, M.A., Celi, G., Chandra, A., Dufour, J., Zu Ermgassen, E.K. and Garrett, R.D. (2024) 'The role of midstream actors in advancing the sustainability of agri-food supply chains', *Nature Sustainability*, Vol. 7, No. 5, pp.527–535.

Hanna, M., Pantanowitz, L., Jackson, B., Palmer, O., Visweswaran, S., Pantanowitz, J., Deebajah, M. and Rashidi, H. (2024) 'Ethical and bias considerations in artificial intelligence (AI)/machine learning', *Modern Pathology*, Vol. 38, No. 3, p.100686.

Jackson, I., Namdar, J., Saénz, M.J., Elmquist III, R.A. and Dávila Novoa, L.R. (2024) 'Revolutionize cold chain: an AI/ML driven approach to overcome capacity shortages', *International Journal of Production Research*, Vol. 63, No. 6, pp.2190–2212.

Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P. (2023) 'Artificial intelligence and knowledge management: A Partnership between human and AI', *Business Horizons*, Vol. 66, No. 1, pp.87–99.

Júnior, M.R.B., De Almeida Moreira, B.R., Dos Santos Carreira, V., De Brito Filho, A.L., Trentin, C., De Souza, F.L.P., Tedesco, D., Setiyono, T., Flores, J.P. and Ampatzidis, Y.J.C. (2024) 'Precision agriculture in the United States: a comprehensive meta-review inspiring further research, innovation, and adoption', *Computers Electronics in Agriculture*, Vol. 221, p.108993.

Kandpal, V. (2024) 'Dimensions of financial inclusion in India: a qualitative analysis of bankers perspective', *Qualitative Research in Financial Markets*, Vol. 16, No. 4, pp.660–679.

Kazungu, I. (2023) 'Business development services and participation of rural-based micro enterprises in export markets: exploring the heterogeneity of the Tanzanian handicrafts industry', *SAGE Open*, Vol. 13, DOI: 21582440231210471.

Khan, R.S., Sirazy, M.R.M., Das, R. and Rahman, S. (2022) 'An AI and ml-enabled framework for proactive risk mitigation and resilience optimization in global supply chains during national emergencies', *Sage Science Review of Applied Machine Learning*, Vol. 5, No. 2, pp.127–144.

Krywalski-Santiago, J. (2024) 'Navigating sustainable transformation in the fashion industry: the role of circular economy and ethical consumer behavior', *Journal of Intercultural Management*, Vol. 16, No. 3, pp.5–36.

Kulkov, I., Kulkova, J., Rohrbeck, R., Menvielle, L., Kaartemo, V. and Makkonen, H. (2024) 'Artificial intelligence-driven sustainable development: examining organizational, technical, and processing approaches to achieving global goals', *Sustainable Development*, Vol. 32, No. 3, pp.2253–2267.

Lohmer, J., Ribeiro Da Silva, E. and Lasch, R. (2022) 'Blockchain technology in operations and supply chain management: a content analysis', *Sustainability*, Vol. 14, No. 10, p.6192.

Mennella, C., Maniscalco, U., De Pietro, G. and Esposito, M. (2024) 'Ethical and regulatory challenges of AI technologies in healthcare: a narrative review', *Heliyon*, Vol. 10, No. 4.

Mou, S., Xue, Q., Chen, J., Takiguchi, T. and Ariki, Y. (2025) 'MM-iTransformer: a multimodal approach to economic time series forecasting with textual data', *Applied Sciences*, Vol. 15, No. 3, p.1241.

Nietschke, Y. and Dabrowski, A. (2023) *Technology in Education: a Case Study on Lao People's Democratic Republic*, Vol. 2, No. SA/P1/06, p.52.

Ojo, T.O., Onayade, A.A. and Naicker, N. (2024) 'Preventing occupational injuries in the informal construction industry: a study protocol for the development of a safety education intervention for bricklayers and carpenters in Osun State, Nigeria', *Frontiers in Public Health*, Vol. 12, No. 2, p.1464797.

Oliveira, F.B., Silva-Filho, M.W., Barbosa, G.A., Freitas, J.P., Penna, C. and Miranda, P.B. (2024) 'Machine learning and time series analysis to forecast hotel room prices', *Brazilian Conference on Intelligent Systems*, Springer, pp.358–371.

Pagano, T.P., Loureiro, R.B., Lisboa, F.V., Peixoto, R.M., Guimarães, G.A., Cruz, G.O., Araujo, M.M., Santos, L.L., Cruz, M.A. and Oliveira, E.L. (2023) 'Bias and unfairness in machine learning models: a systematic review on datasets, tools, fairness metrics, and identification and mitigation methods', *Big Data Cognitive Computing*, Vol. 7, No. 1, p.15.

Pajic, V., Andrejic, M. and Chatterjee, P. (2024) 'Enhancing cold chain logistics: a framework for advanced temperature monitoring in transportation and storage', *Mechatron. Intell. Transp. Syst*., Vol. 3, No. 1, pp.16–30.

Park, W. and Kim, S-G. (2022) 'Integrating quantitative and qualitative methodologies to build a national R&D plan using data envelopment analysis based on R&D stakeholders' perspectives', *Plos One*, Vol. 17, No. 3, p.e0265058.

Pierce, J. (2021) *The Nature of Public Provision of Outdoor Education in the Republic of Ireland: An Ethno-case Study of Four Outdoor Education and Training Centres*, University of Edinburgh.

Poo, M.C-P., Lau, Y-Y., Qi, B. and Pun, C.F-K. (2024) 'Sustainable ground transportation and the E-commerce revolution: Innovations and challenges at the intersection', *Encyclopedia*, Vol. 4, No. 1, pp.201–214.

Portilla-Cabrera, C.V., Cifuentes-Ossa, M.A. and Selvaraj, J.J. (2024) 'Policy-oriented adaptation strategies for artisanal marine fisheries under climate and non-climate stressors in San Andres de Tumaco, Nariño, Colombia', *Marine Policy*, Vol. 161, p.105993.

Putra, A.S.B. and Wibowo, S.J.W.S.J.E. (2023) 'Increasing production efficiency in handicraft small industries in West Java through the utilization of technological innovation', *West Science Journal Economic*, Vol. 1, No. 4, pp.145–152.

Rausch, A.J.C.I.J. (2024) 'Cultural commodities as cultural economies: realization through conceptual and operational space', *Creative Industries Journal*, Vol. 1, No. 1, pp.1–20.

Regona, M., Yigitcanlar, T., Hon, C. and Teo, M. (2024) 'Artificial intelligence and sustainable development goals: systematic literature review of the construction industry', *Sustainable Cities Society*, Vol. 105499.

Rolf, B., Jackson, I., Müller, M., Lang, S., Reggelin, T. and Ivanov, D. (2023) 'A review on reinforcement learning algorithms and applications in supply chain management', *International Journal of Production Research*, Vol. 61, No. 20, pp.7151–7179.

Shen, Y. and Zhang, X. (2024) 'The impact of artificial intelligence on employment: the role of virtual agglomeration', *Humanities Social Sciences Communications*, Vol. 11, No. 1, pp.1–14.

Sugiarti, R., Santosa, R., Wiratno, T. and Wibowo, A.H. (2025) 'Interpersonal evaluative discourse analysis of arts and culture tourism promotional narratives', *WIST*, Vol. 20, No. 1, pp.180–190.

Surendran, U., Nagakumar, K.C.V. and Samuel, M.P. (2024) 'Remote sensing in precision agriculture', *Digital Agriculture: A Solution for Sustainable Food and Nutritional Security*, Springer, Cham.

Tom, T. (2024) 'Accessibility, security, and associated policy issues for people with disabilities in Zimbabwe's urban areas', *Journal of Urban Affairs*, Vol. 1, No. 1, pp.1–14.

Udeh, E.O., Amajuoyi, P., Adeusi, K.B. and Scott, A.O. (2024) 'The role of IoT in boosting supply chain transparency and efficiency', *Magna Scientia Adv. Res. Rev.*, Vol. 12, No. 1, pp.178–197.

Uzoka, A., Cadet, E. and Ojukwu, P.U. (2024) 'The role of telecommunications in enabling internet of things (IoT) connectivity and applications', C*omprehensive Research Reviews in Science Technology*, Vol. 2, No. 2, pp.55–73.

Wadden, J.J. (2023) *Ethics and Healthcare Artificial Intelligence: Some Problems and Solutions for Advanced Diagnostic Systems*, University of British Columbia.

Wang, L., Hu, Y. and Zhou, Y. (2024) 'Cross-border commodity pricing strategy optimization via mixed neural network for time series analysis', arXiv preprint arXiv:.12115.

Wani, M.D., Dada, Z.A. and Shah, S.A. (2025) 'What factors influence consumers' intention and food waste reduction perceptions in Indian restaurants?', *British Food Journal*, Vol. 127, No. 2, pp.569–587.

You, W., Guo, D., Wu, Y. and Li, W. (2023) 'Multiple load forecasting of integrated energy system based on sequential-parallel hybrid ensemble learning', *Energies*, Vol. 16, No. 7, p.3268.

Yousaf, A., Kayvanfar, V., Mazzoni, A. and Elomri, A. (2023) 'Artificial intelligence-based decision support systems in smart agriculture: bibliometric analysis for operational insights and future directions', *Frontiers in Sustainable Food Systems*, Vol. 6, p.1053921.

Yuan, P., Shao, M. and Ma, C. (2024) 'Unlocking economic unity: the digital economy's impact on market segmentation in China', *Journal of the Knowledge Economy*, pp.1–35.

Zhang, Z., Qu, T., Zhao, K., Zhang, K., Zhang, Y., Liu, L., Wang, J. and Huang, G.Q. (2023) 'Optimization model and strategy for dynamic material distribution scheduling based on digital twin: a step towards sustainable manufacturing', *Sustainability*, Vol. 15, No. 23, p.16539.