World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(REVIEW ARTICLE)

Check for updates

# Trustworthy AI for high-stakes decision support across critical sectors

Mustafizur Rahman Shakil [1], Mehedi Hasan [2], Mohammed Imam Hossain Tarek [2], Fakhru Islam Polash [2] and Erin Jahan Meem [3,*]

[1] *Department of Engineering Management, Westcliff University, Irvine, CA 92614, USA.*
[2] *Department of Business Administration, International American University, Los Angeles, CA 90010, USA.*
[3] *Department of Computer Science, Pacific States University, Los Angeles, CA 90010, USA.*

World Journal of Advanced Engineering Technology and Sciences, 2026, 18(03), 241-253

Publication history: Received on 05 February 2026; revised on 12 March 2026; accepted on 14 March 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.18.3.0152

## Abstract

Artificial intelligence (AI) is increasingly embedded in high-stakes decision support across domains ranging from cybersecurity and energy management to healthcare, finance and public governance. Ensuring trustworthiness is essential, yet evidence remains fragmented across sectors. This title-driven scoping review synthesises recent publications that investigate AI-enabled decision support under conditions of criticality and risk. Following a portfolio-bounded methodology, the review infers from titles a cross-sector taxonomy of trust dimensions—including interpretability, robustness, privacy, fairness and governance—and maps them onto sector-specific tasks such as threat detection in critical infrastructure, optimisation of renewable energy systems, diagnosis and prognosis in healthcare, fraud detection in digital finance, and policy compliance in welfare management. The extraction schema codes sectors, tasks, methodological families and deployment settings based solely on title information, and a conceptual rigor rubric is proposed to appraise future full-text evidence. Narrative synthesis highlights common design principles such as interpretable-by-design models and privacy-preserving federated learning, recurring failure modes such as dataset bias and adversarial vulnerabilities, and deployment considerations spanning edge, cloud, 6G and management information systems. A research agenda is outlined to guide systematic evidence gathering, with priorities for unified trust metrics, cross-sector federated collaboration, human-centric explainability, adversarial resilience and governance integration. Limitations due to title-only inference are acknowledged, and the review serves as a structured foundation for future evidence-based synthesis.

**Keywords:** Trustworthy AI; High-Stakes Decision Support; Cross-Sector Analysis; Interpretability; Robustness; Privacy; Fairness; Governance

## 1 Introduction

High-stakes decision support increasingly relies on AI to forecast, classify and optimise critical processes across diverse sectors such as cybersecurity [1], business analytics [2], energy management [10], healthcare diagnostics [14] and digital finance [57]. These applications promise improved efficiency and accuracy but raise concerns about trust when predictions inform national security, patient outcomes, or economic stability.

Recent studies have emphasized the expanding role of AI and MIS in critical national domains. Goffer et al. [1] focused on AI-enhanced cyber threat detection and response for protecting critical infrastructure, while Haldar et al. [2] examined the contribution of AI-driven business analytics and MIS to data-driven economic growth. In the energy sector, Hassan et al. [3] addressed the implementation of MIS solutions to support national energy dominance strategies. Complementing these perspectives, Mahmud et al. [4] investigated AI-powered workforce analytics for forecasting labor market trends and identifying skill gaps linked to economic competitiveness.

* Corresponding author: Erin Jahan Meem

Trustworthy AI encompasses dimensions of interpretability, robustness, privacy, fairness and governance. Interpretability implies that model outputs can be understood by stakeholders [14], while robustness refers to resilience under distributional shifts, perturbations and attacks [1]. Privacy and security address data confidentiality through techniques such as federated learning and secure transactions [13], [18], fairness seeks to ensure that decisions do not perpetuate bias or discrimination [5], [27], and governance integrates AI within management information systems (MIS) and policy frameworks [2], [7].

Existing evidence on trustworthy AI is siloed by sector and method. Cross-sector comparison is essential because lessons from cybersecurity, energy, healthcare, finance and public welfare may generalise or reveal unique constraints. This review contributes a portfolio-bounded synthesis to map the landscape, develop a trust taxonomy, compare sector practices and propose a research agenda. The synthesis is intentionally conceptual: it infers themes solely from titles, with detailed evidence to be added in future systematic reviews.

## 2 Review Protocol and Scope

### 2.1 Portfolio-bounded scoping rationale and research questions

A portfolio-bounded scoping approach was adopted to assemble a manageable yet diverse corpus of AI-related studies spanning critical sectors. Titles were selected as provided without exclusions, reflecting developments in cybersecurity [9], energy strategy [3], renewable optimisation [12], industrial IoT [20], healthcare diagnostics [16], [26], digital finance [57], welfare governance [5] and agriculture [51], [66]. The scoping questions were: (i) What sectors and tasks are represented in this portfolio? (ii) Which trustworthiness dimensions emerge from titles? (iii) How do methods and deployment settings vary across sectors? and (iv) What cross-cutting patterns inform future trustworthy AI design?

### 2.2 Title-driven extraction schema

Because only titles were available, we developed an extraction schema to infer attributes: (a) sector (cybersecurity and critical infrastructure, energy, healthcare, digital finance/business, public welfare/governance, agriculture and other domains), (b) task type (classification, prediction, optimisation, governance, review), (c) methodological family (deep learning [22], ensemble learning [49], transfer learning [29], reinforcement learning [24], federated learning [13], [18], digital twin [20], blockchain [11], quantum computing [46]), (d) trust dimension (interpretability [23], robustness [1], privacy and security [13], fairness [27], governance [7]) and (e) deployment setting (cloud, edge [15], IoT [25], 6G [8], MIS [7], web applications [74], clinical environments). These inferences provide a high-level map and will be refined through full-text analysis.

### 2.3 Rigor appraisal rubric

To guide future evidence appraisal, we outline a conceptual rubric with four domains: methodological rigor, trustworthiness, evaluation design and governance integration. Methodological rigor will assess clarity of problem statement, appropriateness of model choice (e.g., neural networks [16], [63], ensembles [60], ensemble transformers [69]) and validation strategies. Trustworthiness appraisal will evaluate interpretability mechanisms (saliency maps, surrogate models), robustness measures (stress tests, adversarial training), privacy safeguards (differential privacy, federated learning [18]), fairness audits (demographic parity, equality of opportunity [27]) and security features (threat models [1]). Evaluation design will examine dataset diversity, cross-validation, external validation and human-in-the-loop assessments. Governance integration will consider whether studies link AI outputs to decision workflows, MIS audit trails [9], compliance frameworks [5] and stakeholder engagement. Scores will not be assigned in this scoping review; instead the rubric sets expectations for subsequent systematic appraisal once full texts are obtained.

## 3 Cross-Sector Taxonomy for Trustworthy AI

### 3.1 Interpretability & explainability (XAI)

Interpretability ensures that AI-driven decision support can be scrutinised by clinicians, engineers, policymakers and citizens. Titles in our corpus highlight the prominence of explainable deep learning for medical diagnosis [14], [22], transformer–Unet frameworks with explainability [23], stacking ensembles with transparent brain tumour diagnosis [60], and hybrid vision transformer models providing post-hoc explanations for mental health detection [69]. Explainable AI is also emphasised in business decision making [44], rare medicinal plant conservation [71], cotton and fabric defect detection [70], mango leaf disease recognition [67] and XAI-enhanced breast cancer diagnosis [63]. These

titles suggest methods ranging from attention mechanisms and saliency maps to model-agnostic explanations. Interpretability may be intrinsic (interpretable-by-design architectures) or extrinsic (post-hoc explainers), and its evaluation requires human-centred studies to assess usability.

### 3.2 Robustness & resilience (shift, perturbations, attacks)

Robustness denotes an AI system's ability to maintain performance under distributional shifts, perturbations, noise and adversarial attacks. Cyber threat detection and response titles highlight national security imperatives [1], while deep learning-based intrusion detection [25] and credit card fraud detection [40] imply the need to withstand adversarial behaviour. Power systems research addresses blackout mitigation [33], resilient grid operation under electric vehicle penetration [34] and fault detection in transmission lines using LSTM models [39]. Industrial robots require real-time analytics for high-speed communication [24], and energy consumption forecasting [32] and energy transaction security [11] depend on resilience to sensor noise and market volatility. The corpus thus suggests a broad spectrum of robustness concerns, including adversarial attacks, sensor noise, network latency and dynamic load.

**Table 1** Trustworthiness taxonomy

| Dimension | Operational definitions | Recommended metrics/tests | Common pitfalls |
|---|---|---|---|
| Interpretability & explainability | Ability of models to provide understandable reasons for their predictions to stakeholders | Explanation fidelity, human satisfaction, qualitative user studies | Trade-offs with accuracy; misleading surrogate explanations |
| Robustness & resilience | Maintaining performance under distributional shifts, noise and adversarial attacks | Out-of-distribution accuracy, adversarial attack success rate, stress test outcomes | Overfitting to training data; ignoring rare events |
| Privacy & security | Protecting data confidentiality and system integrity against threats | Differential privacy budget, encryption overhead, threat model coverage | Decreased utility from privacy mechanisms; hidden vulnerabilities |
| Fairness & bias control | Ensuring equitable treatment across demographic, geographic and socioeconomic groups | Demographic parity, equality of opportunity, subgroup performance gaps | Unrepresentative datasets; complex trade-offs between fairness and accuracy |
| Governance & MIS integration | Embedding AI within accountable decision workflows with audit trails and stakeholder oversight | Audit trail completeness, decision override rates, compliance checklists | Opaque decision processes; insufficient stakeholder involvement |

### 3.3 Privacy & security (federated learning, threat models, governance)

Ensuring privacy and security is crucial when sensitive data are involved. Titles on privacy-preserving federated learning for critical infrastructure [13] and scalable healthcare data processing [18], as well as automated white blood cell diagnostics via federated learning [48], point to decentralised approaches that keep data local while aggregating model updates. Blockchain-enabled secure energy transactions [11] combine distributed ledgers with fraud detection, while AI-driven threat intelligence [9] and deep learning-based intrusion detection [25] target cybersecurity. Zero-touch 6G frameworks for edge-AI applications [8], [79] and low-latency edge intelligence for healthcare [15] suggest next-generation networks with built-in security. Titles on deepfake detection [54], risk management in digital finance [57], and privacy-first sentiment analysis in mental health [59] further underline the diversity of security concerns.

Recent studies have demonstrated the growing role of AI across energy, infrastructure, and healthcare domains. Ahmed et al. [10] showed how AI-driven time-series analysis can support the optimization of solar energy production for smart energy management in the USA. In a related direction, Khan et al. [11] explored the integration of blockchain and AI for secure energy transactions, particularly for fraud detection and market stability. Ahmed et al. [12] further emphasized the contribution of AI to renewable energy generation and advanced storage technologies in smart energy systems. Beyond the energy sector, Ahmed [13] reviewed privacy-preserving federated learning for critical infrastructure, highlighting key security and governance issues. In healthcare, Siam et al. [14] focused on explainable deep learning models for medical diagnosis to help bridge the gap between AI performance and clinical trust.

### 3.4 Fairness, bias control, and compliance

Fairness and bias control address equitable treatment across demographic, geographic and socioeconomic groups. Welfare management titles focus on transparency and bias control [5], while scalable chronic kidney disease screening [27], [82] mortality risk prediction [28], and skin cancer diagnosis [30] suggest potential for bias if models are trained on unbalanced datasets. Mental health indicator identification [50] and suicidal ideation detection [59] imply ethical sensitivities. Inclusive economic development [56] and empowerment of small businesses [55] highlight societal equity, whereas sports analytics [41] raise considerations about fairness in performance assessment. Compliance is emphasised in MIS-driven governance [7], digital security for deepfake detection [54], and blockchain-enabled energy transactions [11]. Collectively, these titles indicate that fairness metrics and regulatory compliance must be explicitly incorporated into AI development and evaluation.

### 3.5 Governance & MIS integration (audit trails, accountability, decision workflows)

Governance concerns the integration of AI into organisational decision workflows with accountability, audit trails and stakeholder involvement. Titles emphasise data-centric governance models for welfare management [5], cloud-based MIS for IT project governance [7], AI-driven cyber threat intelligence as a MIS [9], and secured digital security for MIS with deepfake detection [54]. Business analytics for economic growth [2], workforce analytics [4], and decision making in business management [6] highlight the need for AI within MIS frameworks to support strategic decisions. Quantum computing in supply chains [46], future of AI-powered SaaS [52], and digital innovation for small businesses [55] suggest evolving governance challenges in emerging platforms. Public policy considerations arise in national energy dominance strategies [3] and inclusive economic development [56]. A central theme is establishing auditability and accountability throughout the AI lifecycle, enabling oversight and enabling stakeholders to challenge automated recommendations.

Recent healthcare-oriented AI studies have addressed both predictive performance and data privacy. Khan et al. [16] investigated high-accuracy breast cancer diagnosis using neural networks integrated with dimensionality reduction methods, while Khan et al. [17] focused on optimizing stroke prediction through neural machine learning algorithms. Complementing these prediction-oriented approaches, Ahmed et al. [18] emphasized privacy-first federated learning models for scalable healthcare data processing, highlighting the importance of secure and distributed AI deployment in medical environments. Recent studies have reflected the broadening impact of AI across healthcare, business, and digital markets. In medical applications, Khan et al. [49] proposed generalizable ensemble learning models for early lung cancer detection, while Shakil et al. [31] evaluated the performance of different deep learning architectures for chest disease and lymphoma classification. Beyond clinical prediction, Mosaddeque et al. [46] investigated the integration of AI and quantum computing to improve supply chain efficiency in aerospace and education, and Sufian et al. [47] highlighted the strategic value of machine learning in the healthcare business sector. From a market perspective, Ahmed et al. [52] examined emerging trends and investment opportunities in AI-powered SaaS, illustrating the growing economic relevance of intelligent software ecosystems.

## 4 Synthesis by Sector

### 4.1 Cybersecurity & critical infrastructure

Cybersecurity and critical infrastructure titles span threat detection [1], [78], cyber threat intelligence as MIS [9], fortified healthcare infrastructure [21], deep learning-based intrusion detection [25], credit card fraud detection [40], deepfake detection [54], zero-touch 6G edge frameworks [8], machine learning for blackout mitigation [33] and resilient grid operation [34]. These suggest tasks such as anomaly detection, intrusion prevention, fraud monitoring and resilient control. Methods include deep learning, reinforcement learning [24], ensemble approaches and secure communication protocols. Trust dimensions emphasised are robustness against adversaries, privacy in threat intelligence sharing, and governance through MIS integration. Deployment settings range from edge and IoT devices [8], [24], [25] to cloud-based MIS [7].

### 4.2 Energy systems & smart grids/renewables

Energy sector titles reflect optimisation of solar energy production [10], secure energy transactions via blockchain and fraud detection [11], renewable energy generation and storage [12], energy consumption prediction with support vector regression [32], ML-based relays to mitigate blackouts [33], resilient grid operations under electric vehicle penetration [34], hybrid renewable energy system optimisation [42] and fault detection in transmission lines [39]. Digital twin technology for predictive maintenance in industrial IoT [20] complements these by improving asset longevity. These studies imply the use of time-series forecasting, genetic algorithms, reinforcement learning, federated

learning and blockchain. Trust themes include robustness to grid fluctuations, privacy in energy data, fairness in energy distribution, and governance through standardised MIS to support national energy strategies [3].

**Table 2** Sector-specific assurance and evaluation practices

| Sector | Assurance & evaluation practices | Specific trust considerations |
|---|---|---|
| Cybersecurity & critical infrastructure | Adversarial testing, intrusion detection benchmarks, penetration testing, incident response simulations | Robustness to attacks, privacy in threat intelligence sharing, auditability of automated responses |
| Energy systems & smart grids | Time-series forecasting validation, blackout mitigation simulations, hardware-in-the-loop testing | Resilience to grid fluctuations, fairness in energy allocation, security of transactions |
| Healthcare | Clinical trials, retrospective studies, cross-institutional validation, human-in-the-loop evaluation | Interpretability of medical decisions, fairness across demographics, privacy of patient data |
| Digital finance & business analytics | Fraud detection benchmarks, stress testing for market volatility, A/B testing in decision support systems | Security against financial fraud, transparency for regulators, fairness in lending and hiring |
| Public sector & welfare | Policy compliance audits, stakeholder consultations, pilot deployments within government MIS | Transparency and bias control, accountability, legal compliance |

### 4.3 Healthcare (diagnosis/decision support, imaging, risk prediction)

Healthcare constitutes the largest share of titles, covering explainable deep learning for medical diagnosis [14], breast cancer detection with neural networks and dimensionality reduction [16], stroke prediction [17], federated learning for scalable data processing [18], optimisation of neural architectures for cancer diagnosis [19], brain tumour classification [22], [80] transformer–Unet segmentation [23], multimodal breast cancer detection [26], scalable chronic kidney disease screening [27], explainable mortality risk prediction [28], transfer learning for eye disease classification [29], skin cancer diagnosis challenges and translation [30], [81], performance analysis of chest disease classification [31], EMG signal denoising [35], fault detection and classification models [39], ensemble models for lung cancer detection [49], mental health indicator identification [50], and numerous transformer and ensemble models for brain tumours [60], leukemia [73], cervical cancer [65], [83], lung cancer [64], breast cancer [63], medicinal plant recognition [71], cotton leaf disease identification [72], soybean and mango leaf disease recognition [66], [67], and web-based diagnostics in low-resource settings [74]. Titles also address telehealth and edge intelligence for critical disease monitoring [15], integrating regenerative therapies and smart technologies for wound healing [75], polymer–drug nanoconjugates for glioblastoma therapy [76], and molecular erasers for cancer immunity reprogramming [77]. These tasks emphasise classification, segmentation, prediction and therapy, with methodological diversity encompassing deep learning, ensemble learning, transfer learning, federated learning, reinforcement learning, transformer architectures and digital twins. Trust concerns focus on interpretability (saliency maps and attention mechanisms), privacy (federated approaches [18], [48]), robustness to data heterogeneity (EMG noise [35], [38]) and fairness across demographic groups [27], [30].

### 4.4 Digital finance/fraud and business analytics

Digital finance and business analytics titles include AI-driven business analytics for economic growth [2], ensemble and deep learning models for credit card fraud detection [40], AI-driven risk management and fraud prevention [57], machine learning for strategic business gains in healthcare [47], impact of IT product innovation [53], empowerment of small businesses through digital innovation [55], and supply chain optimisation using AI and quantum computing [46]. Additional titles explore the future of AI-powered SaaS [52], digital security in MIS for deepfake detection [54], and market trends with investment opportunities [52]. Financial and business applications often adopt machine learning, ensemble and reinforcement learning methods, and emphasise interpretability for decision makers [44], robustness against fraudulent attacks [40], and governance through MIS integration [7]. Fairness is relevant where AI influences credit scoring, workforce analytics [4] and socioeconomic development [56].

Recent studies have demonstrated the breadth of AI applications across energy, language processing, business intelligence, and industrial inspection. Ahamed et al. [42] reviewed machine learning strategies for optimizing hybrid

renewable energy systems in decentralized smart grids to enhance energy efficiency and grid stability. In a different application area, Ahamed et al. [43] introduced a sentiment recognition approach integrating bidirectional deep learning with an extended fuzzy Markov model. From the perspective of organizational analytics, Hossain et al. [44] investigated how explainable AI can strengthen business decision-making by helping bridge the trust gap. In industrial quality assurance, Haque et al. [45] proposed a data-centric approach for leather quality control using advanced vision transformer models.

### 4.5 Welfare/public-sector governance and policy compliance

Public sector and welfare governance titles focus on data-centric governance models to enhance transparency, bias control and policy compliance in welfare management [5], the role of cloud-based MIS in IT project governance [7], AI-driven cyber threat intelligence within MIS [9], national energy dominance strategies supported by MIS solutions [3], inclusive economic development and AI's role in bridging socioeconomic divides [56], and empowerment of small businesses through digital innovation [55]. Titles on policy compliance intersect with healthcare security via AI-driven cybersecurity [21], deepfake detection [54] and risk management in digital finance [57]. Methods range from machine learning to blockchain [11] and federated learning [13], while trust themes emphasise fairness, transparency, accountability and compliance. Deployments often involve MIS and cloud infrastructures.

Recent studies have demonstrated the growing role of AI in predictive and screening applications within US healthcare systems. Arafat et al. [26] presented a deep learning framework that integrates mammography and clinical EHR data for early breast cancer detection. Extending AI-based screening to chronic disease management, Rimon et al. [27] proposed a scalable machine learning strategy for chronic kidney disease screening across healthcare systems. In a related direction, Hasan et al. [28] developed an explainable machine learning framework for mortality risk prediction in liver cirrhosis patients, underscoring the need for interpretable and clinically reliable predictive models.

## 5 Cross-Cutting Patterns and Design Principles

Across sectors, several design principles recur. Interpretable-by-design models integrate explainability directly into architectures using attention mechanisms, symbolic reasoning or surrogate models [22], [23], [60], [63], [64], [65], [69], while post-hoc explanation methods attempt to rationalise opaque models [14], [30], [50]. Privacy-by-design approaches leverage federated learning [13], [18], [48], differential privacy and blockchain [11] to minimise data sharing. Assurance-by-design emphasises robustness and security through adversarial training, stress testing and zero-touch network architecture [8], [1], [25], [40]. Fairness-by-design integrates bias mitigation techniques into pipelines [5], [27], [56], and governance-by-design embeds auditability and compliance into MIS [7], [9].

Recent studies have advanced brain tumor analysis through both classification and segmentation frameworks. Oza et al. [22] developed an AI-based ensemble learning approach for brain tumour classification from MRI images, demonstrating the relevance of ensemble strategies in neuroimaging diagnosis. In contrast, Khushubu et al. [23] focused on brain tumor segmentation by proposing TransUNetB, a Transformer–U-Net-based architecture designed to achieve efficient and explainable segmentation. Together, these studies reflect the growing integration of ensemble learning and transformer-based models in brain tumor imaging applications.

Failure modes that recur across domains include dataset bias leading to unfair or inaccurate predictions [5], [27], [30], overfitting and poor generalisation due to limited or homogeneous data [31], vulnerability to adversarial perturbations in cybersecurity and healthcare [1], [25], privacy breaches when models memorise sensitive details [13], lack of transparency hindering user acceptance [14], and governance gaps where AI outputs are not properly audited [7], [54]. Recognising these patterns encourages development of standard stress tests, bias audits and privacy checks.

Recent studies have addressed both predictive intelligence and physical reliability in energy and electrical systems. Hasan et al. [32] applied support vector regression optimized with a genetic algorithm to forecast sector-wise energy consumption. In the context of grid protection, Juel et al. [33] introduced a machine learning-based sophisticated relay approach to reduce blackout risks in power systems. Tonny et al. [34] further explored resilient grid operation through data-driven methods under growing electric vehicle penetration. Beyond grid analytics, Tanbhir et al. [35] compared DWT and EMD techniques for denoising electromyographic signals, while Khan et al. [36] analyzed the effects of inductance and skin effect on transient wave propagation in transformer bushings. Together, these studies reflect the broad application of computational intelligence and system-level analysis in modern engineering problems.

Deployment considerations involve choosing between edge, cloud or hybrid architectures. Low-latency edge intelligence appears in healthcare [15], robotics [24] and 6G frameworks [8], enabling real-time responses but

increasing constraints on model size. Cloud-based deployments offer scalability for business analytics [2], [6] and financial services [57] but raise privacy concerns. Hybrid models combining federated learning with edge computing allow decentralised training [18], [48]. Integration into MIS [7], [9] and compliance with sector regulations demand mechanisms for auditability, version control and user override. Continual monitoring of models in operation is necessary to detect performance drift, bias amplification and emergent security threats.

**Table 3** Governance/MIS patterns

| Pattern | Description | Compliance & stakeholder roles |
|---|---|---|
| Centralised MIS with AI integration | AI models embedded in centralised management information systems, providing dashboards and decision recommendations | Ensures central auditability and compliance; stakeholders include executives, regulators and IT managers |
| Decentralised federated MIS | Distributed MIS that aggregates insights from local nodes via federated learning without sharing raw data | Enhances privacy and local autonomy; stakeholders include local operators and central coordinators |
| Blockchain-based MIS | Uses distributed ledgers to record transactions and model decisions, enabling immutability and transparency | Facilitates auditability and trust; stakeholders include participants and auditors |
| Edge computing & 6G-enabled MIS | Edge devices perform local inference with low latency and synchronise with central MIS through high-speed networks | Supports real-time decisions and resilience; stakeholders include field operators and network providers |

## 6 Research Agenda

### 6.1 Unified trust metrics and evaluation frameworks

Motivation: Existing titles reveal fragmented evaluation across sectors and trust dimensions. A unified set of trust metrics would enable comparable assessment of interpretability, robustness, privacy, fairness and governance.

What to measure: Measures should include human-centred interpretability scores, adversarial robustness indicators, differential privacy budgets, demographic parity indices, auditability compliance checks and computational efficiency metrics.

Suggested evaluation protocol: Develop benchmark datasets and simulation environments representing multiple sectors (cybersecurity, energy, healthcare, finance, public governance). Employ cross-validation, stress tests and fairness audits. Evaluate using stakeholder studies to assess comprehension and trust.

Expected impact: A unified framework will facilitate cross-sector comparison, accelerate regulatory standards and guide practitioners in selecting appropriate models for critical applications.

### 6.2 Privacy-preserving collaborative learning

Motivation: Federated learning and blockchain titles [13], [18], [48], [11] suggest potential for collaborative modelling without centralising data, but open questions remain regarding scalability, fairness and governance.

What to measure: Key measures include model accuracy, communication overhead, privacy leakage estimates, fairness across participating nodes and compliance with sector-specific regulations.

Suggested evaluation protocol: Simulate federated learning across heterogeneous datasets from different sectors; test differential privacy mechanisms; evaluate blockchain integration for auditability. Compare with centralised baselines.

Expected impact: Advancing privacy-preserving collaboration will lower barriers to data sharing, enable richer models and promote equitable participation across organisations and regions.

### 6.3 Human-centric explainability and user trust

Motivation: Multiple titles emphasise explainable AI [14], [22], [23], [60], [69], yet evaluation rarely considers how explanations affect user trust and decision outcomes.

What to measure: Beyond technical fidelity, measure explanation usefulness, cognitive load, user satisfaction and decision quality through user studies with stakeholders (clinicians, engineers, policymakers).

Suggested evaluation protocol: Design controlled experiments comparing different explanation modalities (textual, visual, counterfactual) across tasks such as medical diagnosis, cyber threat analysis and financial fraud detection.

Expected impact: Understanding human factors will guide the design of explanations that genuinely improve trust and enable responsible adoption of AI.

### 6.4 Resilience to adversarial and distributional shifts

Motivation: Robustness titles [1], [25], [33], [34], [39] signal concerns about attacks and shifts, but systematic evaluation frameworks are lacking.

What to measure: Define metrics for adversarial robustness (attack success rate, certified robustness), distributional robustness (performance under covariate shift), and detection latency for anomalies.

Suggested evaluation protocol: Create adversarial and distributional shift benchmarks for each sector, including synthetic perturbations and realistic attack scenarios. Evaluate defence mechanisms and stress test models.

Expected impact: Building resilient AI will reduce catastrophic failures in national security, power grids, healthcare diagnostics and financial systems.

### 6.5 Governance frameworks and auditability in MIS

Motivation: Titles on MIS governance [5], [7], [9], [54], [55] underline the need for transparent decision workflows, yet operational frameworks remain underdeveloped.

What to measure: Metrics should capture audit trail completeness, decision override frequencies, compliance with regulatory requirements, and stakeholder satisfaction.

Suggested evaluation protocol: Develop prototype MIS modules integrating AI models with logging, provenance tracking and user feedback. Conduct case studies in sectors such as welfare management, energy strategy and digital finance.

Expected impact: Enhanced governance will foster accountability, enable regulatory approval and build public trust in AI-driven decisions.

*Limitations*

This portfolio-bounded scoping review is restricted to information contained in article titles. No full texts or abstracts were examined, and therefore the classification of sectors, tasks, methods and trust dimensions may be imprecise. Important details such as sample sizes, datasets, methodological nuances, numerical results and context-specific findings remain unknown. The narrative synthesis avoids attributing specific outcomes or performance metrics to individual studies and instead infers conceptual themes. As such, the review should be viewed as an organising framework for future systematic reviews rather than a definitive evidence synthesis.

## 7 Conclusion

AI is transforming high-stakes decision support across sectors including cybersecurity, energy, healthcare, finance and public governance. Trustworthiness is a multidimensional requirement encompassing interpretability, robustness, privacy, fairness and governance. This title-driven scoping review synthesised recent publications, producing a cross-sector taxonomy, sector-specific narratives, design principles and a research agenda. The findings underscore common concerns such as adversarial resilience, bias mitigation, privacy preservation and auditability, and highlight the need for unified evaluation metrics and governance frameworks. By structuring the evidence landscape and

identifying gaps, this review provides a foundation for comprehensive future work that will incorporate full-text analysis and empirical rigor.

## Compliance with ethical standards

*Disclosure of conflict of interest*

There is no conflict of interest.

## References

[1] Goffer MA, Uddin MS, Hasan SN, Barikdar CR, Hassan J, Das N, Chakraborty P, Hasan R. Ai-enhanced cyber threat detection and response advancing national security in critical infrastructure. Journal of Posthumanism. 2025 Apr 17;5(3):1667-89.

[2] Haldar U, Alam GT, Rahman H, Miah MA, Chakraborty P, Saimon AS, Chy MA, Siddiqa KB, Manik MM. AI-driven business analytics for economic growth leveraging machine learning and MIS for data-driven decision-making in the US economy. Journal of Posthumanism. 2025 Apr 24;5(4):932-57.

[3] Hassan J, Rahman H, Haldar U, Sultana S, Rahman MM, Chakraborty P, Das N, Barikdar CR. Implementing MIS solutions to support the national energy dominance strategy. Journal of Posthumanism. 2025 May 20;5(5):4343-63.

[4] Mahmud F, Goffer MA, Chakraborty P, Sultana S, Rozario E, Miah MA, Chy MA, Haldar U. AI-powered workforce analytics forecasting labor market trends and skill gaps for US economic competitiveness. Journal of Computer Science and Technology Studies. 2024 Dec 31;6(5):265-77.

[5] Uddandarao DP, Valiveti SS, Varanasi SR, Rahman H, Chakraborty P. Data-Centric Governance Models Using Trustworthy AI: Strengthening Transparency, Bias Control, and Policy Compliance in Welfare Management. International Journal on Engineering Artificial Intelligence Management, Decision Support, and Policies. 2025;2(4):29-44.

[6] Chakraborty P, Siddiqa KB, Rahman H, Miah MA, Das N, Goffer MA, Das S. Leveraging artificial intelligence and machine learning for decision-making in business management: A comprehensive analysis. Journal of Management World. 2025:46-56.

[7] Mahmud F, Goffer MA, Rahman H, Chakraborty P, Das N, Orthi SM, Alam GT. The Role of Cloud-Based Management Information Systems in Enhancing IT Project Governance and Stakeholder Collaboration. InInternational Conference on AI and Robotics 2025 May 9 (pp. 1-14).

[8] Mahin MR, Chakraborty P, Das N, Kaur H, Himel HU, Kaur J, Mohapatra AG. Secured and Standardized Intelligent Zero-Touch 6G Framework for Edge-AI Applications. IEEE Communications Standards Magazine. 2026 Feb 25.

[9] Orthi SM, Chakraborty P, Siam MA, Shan-A-Alahi A, Al Zaiem A, Hasan SN, Kaur J, Mahmud F, Goffer MA. AI-Driven Cyber Threat Intelligence as a Management Information System: Integrating Cybersecurity Governance and IT Project Management for Organizational Resilience. The Eastasouth Journal of Information System and Computer Science. 2023 Dec 30;1(02):194-213.

[10] Ahmed I, Khan MA, Islam MD, Hasan MS, Jakir T, Hossain A, Abed J, Hasanuzzaman M, Shatyi SS, Hasnain KN. Optimizing Solar Energy Production in the USA: Time-Series Analysis Using AI for Smart Energy Management. arXiv preprint arXiv:2506.23368. 2025 Jun 29.

[11] Khan MA, Islam MD, Ahmed I, Rabbi MM, Anonna FR, Zeeshan MD, Ridoy MH, Chowdhury BR, Rabbi MN, Sadnan GM. Secure Energy Transactions Using Blockchain Leveraging AI for Fraud Detection and Energy Market Stability. arXiv preprint arXiv:2506.19870. 2025 Jun 21.

[12] Ahmed I, Khan MA, Islam MA, Ahamed A, Siam MA, Islam MZ. Utilizing ai to enhance renewable energy generation and advanced storage technologies for smart energy solutions. In2025 International Conference on Metaverse and Current Trends in Computing (ICMCTC) 2025 Apr 10 (pp. 1-10).

[13] Ahmed I. Privacy-Preserving Federated Learning for Critical Infrastructure: A Systematic Review of Security Threat Models, Deployment Patterns, And Governance. International Journal of Scientific Interdisciplinary Research. 2026 Feb 13;7(1):204-35.

[14] Siam MA, Ahmed I, Khan MA, Islam MA, Milon MH, Ahamed A, Islam MZ. Explainable Deep Learning Models for Medical Diagnosis: Bridging the Gap between AI and Healthcare. In2025 3rd International Conference on Business Analytics for Technology and Security (ICBATS) 2025 May 1 (pp. 1-7).

[15] Islam MZ, Ahmed I, Parveen R, Rimon ST, Janjua JI. Revolutionizing Healthcare for Critical Diseases with AI and Enhancing Security via Low-Latency Edge Intelligence. In2025 International Conference on AI-Driven STEM Education and Learning Technologies (AISTEMEDU) 2025 Dec 1 (pp. 1-6).

[16] Khan MA, Parveen R, Ahmed I, Milon MH, Khan TA. High-Accuracy Breast Cancer Diagnosis Using Neural Networks and Dimensionality Reduction Techniques. In2025 IEEE 19th International Conference on Open Source Systems and Technologies (ICOSST) 2025 Dec 1 (pp. 1-6).

[17] Khan MA, Papel MS, Milon MH, Ahmed I, Islam MZ, Raja MR. Optimizing Stroke Prediction in Healthcare with Neural Machine Learning Algorithms. In2025 3rd International Conference on Cyber Resilience (ICCR) 2025 Jul 3 (pp. 1-7).

[18] Ahmed I, Papel MS, Raja MR, Islam MZ, Khan MA, Milon MH. Privacy-First Federated Learning Models for Scalable Healthcare Data Processing. In2025 3rd International Conference on Cyber Resilience (ICCR) 2025 Jul 3 (pp. 1-6).

[19] Raja MR, Milon MH, Ahmed I, Papel MS, Khan MA, Islam MZ. Optimizing Neural Architectures for Accurate Diagnosis of Breast Cancer from Morphological Features. In2025 3rd International Conference on Cyber Resilience (ICCR) 2025 Jul 3 (pp. 1-6).

[20] Milon MH, Rahman MM, Papel MS, Raja MR, Semi MM, Tarafder MT. Digital Twin Technology for Predictive Maintenance in Industrial Iot Environments: Enhancing Operational Efficiency and Asset Longevity. In2025 3rd International Conference on Business Analytics for Technology and Security (ICBATS) 2025 May 1 (pp. 1-8).

[21] Islam MZ, Siam MA, Ahmed I, Khan MA, Islam MA, Milon MH. Fortifying Healthcare and Essential Infrastructure with AI-Driven Cybersecurity Technologies. In2025 International Conference on Metaverse and Current Trends in Computing (ICMCTC) 2025 Apr 10 (pp. 1-9).

[22] Oza BA, Khan IR, Kanulla NS, Hasan SM, Davuluri SK, Pothakamuri LN. Artificial Intelligence based Ensemble Learning Approach for Brain Tumour Classification Using MRI Images. In2024 7th International Conference on Contemporary Computing and Informatics (IC3I) 2024 Sep 18 (Vol. 7, pp. 1583-1588).

[23] Khushubu KG, Al Masum A, Rahman MH, Hasan SM, Bhuiyan MI, Mahmud MR, Swapno SM, Appaji A. Transunetb: An advanced transformer–unet framework for efficient and explainable brain tumor segmentation. Informatics in Medicine Unlocked. 2025 Nov 1:101706.

[24] Hasan SM, Gudala L, Soni M, Quraishi A, Keshta I, Reegu FA. Deep Reinforcement Learning-Based Real-time Data Analytics of Automatic Industrial Robot for High Speed Communication. In2024 IEEE 2nd International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP) 2024 Dec 6 (pp. 1-8).

[25] Hasan SM, Keshta I, Thakkar D, Gudala L, Soni M, Maaliw RR. Deep Learning-Based Intrusion Detection Model for Internet of Things Environment. In2024 IEEE 2nd International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP) 2024 Dec 6 (pp. 1-8).

[26] Arafat Y, Asha NB, Ahmed S, Sakib SH, Shakil MR, RISHTA AZ, Rahat SR. A Deep Learning Framework for Early Breast Cancer Detection Among US Women: Integrating Mammography and Clinical EHR Data. British Journal of Nursing Studies. 2025 Dec 17;5(3):44-59.

[27] Rimon ST, Hasan E, Asha NB, Sultana R, Sharmin F, Rahaman M, Shakil MR. A Scalable Machine Learning Strategy for Chronic Kidney Disease Screening Across US Healthcare Systems. Frontiers in Computer Science and Artificial Intelligence. 2026 Jan 28;5(1):40-54.

[28] Hasan E, Asha NB, Rahaman M, Shakil MR, Haque MM, Paul D, Al Amin M. An Explainable Machine Learning Framework for Mortality Risk Prediction of Liver Cirrhosis Patients in the US Healthcare System. Frontiers in Computer Science and Artificial Intelligence. 2026 Feb 22;5(4):15-26.

[29] Asha NB, Shakil MR, Rahat SR, Sakib SH, Rahaman M, Hasan E, Ahmed S. A Transfer Learning–Based Deep Convolutional Neural Network Framework for Automated Multi-Class Eye Disease Classification in the USA Using Retinal Fundus Image. Journal of Medical and Health Studies. 2023 Jul 23;4(4):209-17.

[30] Rahat SR, Rahaman M, Sakib SH, Ahmed S, Asha NB, Shakil MR, Hasan E. Deep Learning–Based Skin Cancer Diagnosis in the United States: Advances, Challenges, and Clinical Translation. Journal of Medical and Health Studies. 2023 Dec 25;4(6):150-60.

[31] Shakil MR, Hasan M, Tarek MIH, Polash FI, Meem EJ. Performance analysis of deep learning architectures for chest disease and lymphoma classification. World Journal of Advanced Engineering Technology and Sciences. 2025 Oct 31;17(1):233–47.

[32] Hasan MS, Tarequzzaman M, Moznuzzaman M, Juel MA. Prediction of energy consumption in four sectors using support vector regression optimized with genetic algorithm. Heliyon. 2025 Jan 30;11(2).

[33] Juel MA, Uddin N, Abid MA, Islam MZ, Nasim AA, Sourov S, Jaman MS. Machine Learning Approached for Sophisticated Relay to Mitigate Blackouts in Power System. In2025 15th International Conference on Power, Energy, and Electrical Engineering (CPEEE) 2025 Feb 15 (pp. 231-235).

[34] Tonny FE, Rahman M, Kajal MK, Juel MA, Hassan M. Data-driven Resilient Grid Operation Under Growing EV Penetration. In2025 International Conference on Electrical, Computer and Communication Engineering (ECCE) 2025 Feb 13 (pp. 1-6).

[35] Tanbhir R, Juel MA, Abdullah A, Moznuzzaman M. Performance study of dwt and emd techniques for denoising electromyographic signal. In2025 International Conference on Electrical, Computer and Communication Engineering (ECCE) 2025 Feb 13 (pp. 1-6).

[36] Khan MN, Juel MA, Nasim AA, Islam MZ, Jamil MS, Jaman MS, Robbani MS. Influence of Inductance and Skin Effect in Multi-layer Coaxial Cables on Transient Wave Propagation in Transformer Bushings. In2025 Interdisciplinary Conference on Electrics and Computer (INTCEC) 2025 Sep 15 (pp. 1-6).

[37] Mondal R, Bose D, Juel MA, Tarequzzaman M, Moznuzzaman DM. Efficient EMG Signal Classification Using Domain-Optimized Features. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[38] Datta TC, Tanbhir R, Juel MA, Islam M, Moznuzzaman M. Advancing Electromyography Signal Analysis: Wavelet-Based Filtering for Low Back Pain Patients Assessment. In2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM) 2025 Jun 27 (pp. 1-6).

[39] Pasha SM, Ye C, Sumon MS, Jaman MS, Sohag SR, Ali MM, Juel MA. Generalized and Robust LSTM Model for Fault Detection and Classification in Power Transmission Line. In2025 7th Asia Energy and Electrical Engineering Symposium (AEEES) 2025 Mar 28 (pp. 422-426).

[40] Juel MA, Islam M, Jahan SN, Tarek A, Al Hasib A, Jaman MS. Evaluating Machine Learning and Deep Learning Models for Credit Card Fraud Detection in Cybersecurity. In2025 10th International Conference on Information and Network Technologies (ICINT) 2025 Mar 12 (pp. 142-146).

[41] Ahmed H, Tanbhir R, Shohag JA, Islam MM, Juel AA, Moznuzzaman M. Evaluation of the Importance of Shoulder Strength, Hand strength and Heart Rate Recovery for Badminton Players. In2025 International Conference on Electrical, Computer and Communication Engineering (ECCE) 2025 Feb 13 (pp. 1-5).

[42] Ahamed A, Fardin H, Hasan E, Rimon ST, Haque MM. A Review of Machine Learning Approaches for Optimizing Hybrid Renewable Energy Systems (HRES) in Decentralized Smart Grids: Enhancing Energy Efficiency and Grid Stability. Iranian Journal of Science and Technology, Transactions of Electrical Engineering. 2025 Oct 23:1-4.

[43] Ahamed A, Tarafder MT, Rimon ST, Ahmed N. Bidirectional deep learning and extended fuzzy Markov model for sentiments recognition. ICCK Transactions on Neural Computing. 2025 Mar 30;1(1):11-29.

[44] Hossain Z, Hossain IS, Kabir MF, Rahman A, Gharami AK, Sumi SS, Ahamed A. Implementing Explainable AI to Enhance Business Decision Making & Bridging the Trust Gap. In2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 2025 Feb 18 (pp. 0072-0077).

[45] Haque R, Khan M, Pranto MN, Islam MA, Islam MB, Rawat U, Bhattacharya P, Jangid M. Data-Centric Approach for Leather Quality Control Using Advanced Vision Transformer Models. In2025 International Conference on Next Generation Communication & Information Processing (INCIP) 2025 Jan 23 (pp. 200-205).

[46] Mosaddeque AI, Guria ZM, Morshed N, Sufian MA, Ahamed A, Rimon ST. Transforming AI and Quantum Computing to Streamline Business Supply Chains in Aerospace and Education. In2024 International Conference on TVET Excellence & Development (ICTeD) 2024 Dec 16 (pp. 231-236).

[47] Sufian MA, Rimon ST, Mosaddeque AI, Guria ZM, Morshed N, Ahamed A. Leveraging Machine Learning for Strategic Business Gains in the Healthcare Sector. In2024 International Conference on TVET Excellence & Development (ICTeD) 2024 Dec 16 (pp. 225-230).

[48] Fakrul Islam Polash M, Khan S, Imam M, Tarek H, Hasan M, Shakil R, et al. Automated white blood cell diagnostics using federated learning and distributed deep learning. World Journal of Advanced Engineering Technology and Sciences. 2025 Oct 31;17(1):218–32.

[49] Khan S, Hasan M, Tarek MIH, Shakil MR, Polash MF islam, Kabir I. Generalizable Ensemble Learning Models for Early Lung Cancer Detection. World Journal of Advanced Engineering Technology and Sciences. 2025 Oct 31;17(1):200–17.

[50] Meem EJ, Fakrul Islam Polash M, Imam M, Tarek H, Hasan M, Shakil MR. Identifying Critical Mental Health Indicators Using Ensemble and Explainable AI Techniques. World Journal of Advanced Engineering Technology and Sciences. 2025 Oct 31;17(1):186–99.

[51] ZakirHossain M, Khan MM, Thapa S, Uddin R, Meem EJ, Niloy SK, Rafi MA, Shazib MS, Kabir MF, Bhavani GD. Advanced deep learning techniques for precision diagnosis of tea leaf diseases. In2025 IEEE International Conference on Emerging Technologies and Applications (MPSec ICETA) 2025 Feb 21 (pp. 1-6).

[52] Ahmed KR, Ahmmed MR, Kowser S, Faruq O, Shammah RS, Sufian MA. The Future of AI-Powered SaaS: Market Trends and Investment Opportunities. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[53] Ahmed KR, Sufian MA, Faruq O, Kowser S, Shammah RS, Ahmmed MR. Impact of IT Product Innovation on Competitive Advantage in the Digital Economy. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[54] Ahmed KR, Shammah RS, Kowser S, Faruq O, Sufian MA, Ahmmed MR. Strengthening Digital Security in MIS: A Business Analytics Approach to Deepfake Detection. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[55] Sufian MA, Morshed N, Rahman I, Guria ZM, Dhar SR. Empowering Small Businesses with Digital Innovation by Leveraging IT Infrastructure, Data Analytics, and AI for Sustained Growth. In2025 International Conference on Artificial Intelligence's Future Implementations (ICAIFI) 2025 Sep 14 (pp. 94-99).

[56] Guria ZM, Morshed N, Rahman I, Dhar SR, Sufian MA. Advancing Global Peace through Inclusive Economic Development and the Role of Artificial Intelligence in Bridging Socioeconomic Divides. In2025 International Conference on Artificial Intelligence's Future Implementations (ICAIFI) 2025 Sep 14 (pp. 100-105).

[57] Guria ZM, Rahman I, Kowser S, Morshed N, Shammah RS, Faruq O. AI-Driven Risk Management and Fraud Prevention in Digital Finance as a Catalyst for US Economic Resilience and Global Financial Leadership. In2025 IEEE 19th International Conference on Open Source Systems and Technologies (ICOSST) 2025 Dec 1 (pp. 1-8).

[58] Haque R, Islam N, Tasneem M, Das AK. Multi-class sentiment classification on Bengali social media comments using machine learning. International journal of cognitive computing in engineering. 2023 Jun 1;4:21-35.

[59] Haque R, Islam N, Islam M, Ahsan MM. A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning. Technologies. 2022 Apr 29;10(3):57.

[60] Haque R, Khan MA, Rahman H, Khan S, Siddiqui MI, Limon ZH, Swapno SM, Appaji A. Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. Computers in Biology and Medicine. 2025 Jun 1;191:110166.

[61] Haque R, Laskar SH, Khushbu KG, Hasan MJ, Uddin J. Data-driven solution to identify sentiments from online drug reviews. Computers. 2023 Apr 21;12(4):87.

[62] Ahmed MR, Rahman H, Limon ZH, Siddiqui MI, Khan MA, Pranta AS, Haque R, Swapno SM, Cho YI, Abdallah MS. Hierarchical swin transformer ensemble with explainable ai for robust and decentralized breast cancer diagnosis. Bioengineering. 2025 Jun 13;12(6):651.

[63] Debnath J, Hossain A, Sakib A, Rahman H, Haque R, Ahmed MR, Reza AW, Swapno SM, Appaji A. LMVT: A hybrid vision transformer with attention mechanisms for efficient and explainable lung cancer diagnosis. Informatics in Medicine Unlocked. 2025 Jul 9:101669.

[64] Siddiqui MI, Khan S, Limon ZH, Rahman H, Khan MA, Al Sakib A, Swapno SM, Haque R, Reza AW, Appaji A. Accelerated and accurate cervical cancer diagnosis using a novel stacking ensemble method with explainable AI. Informatics in Medicine Unlocked. 2025 Jan 1;56:101657.

[65] Pranta AS, Fardin H, Debnath J, Hossain A, Sakib AH, Ahmed MR, Haque R, Reza AW, Dewan MA. A novel maxvit model for accelerated and precise soybean leaf and seed disease identification. Computers. 2025 May 18;14(5):197.

[66] Noman AA, Hossain A, Sakib A, Debnath J, Fardin H, Sakib AA, Haque R, Ahmed MR, Reza AW, Dewan MA. ViX-MangoEFormer: An Enhanced Vision Transformer–EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with Explainable Artificial Intelligence. Computers. 2025 May 2;14(5):171.

[67] Ahmed MR, Haque R, Rahman SMA, Reza AW, Siddique N, Wang H. Vision-audio multimodal object recognition using hybrid and tensor fusion techniques. Information Fusion. 2026 Feb 1;126:103667.

[68] Islam S, Haque R, Khan MA, Mohiuddin AB, Siddiqui MI, Limon ZH, Khushbu KG, Swapno SM, Ahmed MR, Appaji A. Ensemble transformer with post-hoc explanations for depression emotion and severity detection. iScience. 2026 Feb 20;29(2).

[69] Swapno SM, Sakib A, Hossain A, Debnath J, Al Noman A, Al Sakib A, Ahmed MR, Haque R, Appaji A. Explainable transformer framework for fast cotton leaf diagnostics and fabric defect detection. IScience. 2026 Feb 20;29(2).

[70] Khan S, Rahman H, Siddiqui MI, Limon ZH, Khan MA, Haque R, Ahmed MR, Reza AW, Ripon SH. Ensemble-Based Explainable Approach for Rare Medicinal Plant Recognition and Conservation. In2025 10th International Conference on Information and Network Technologies (ICINT) 2025 Mar 12 (pp. 88-93).

[71] Rahman MM, Hossain MS, Dhakal K, Poudel R, Islam MM, Ahmed MR, Rahman S. A Novel Transformer Model for Accelerated and Efficient Cotton Leaf Disease Identification. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[72] Hasan J, Hasan K, Al Noman A, Hasan S, Sultana S, Arafat MA, Islam SA, Sarker A, Rahman S, Ahmed MR. Transforming leukemia classification: A comprehensive study on deep learning models for enhanced diagnostic accuracy. In2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON) 2024 Sep 12 (pp. 1-6).

[73] Haque R, Sultana S, Rafy A, Islam MB, Arafat MA, Bhattacharya P, et al. A transfer learning-based computer-aided lung cancer detection system in smart healthcare. IET Conference Proceedings. 2025 Mar;2024(37):594–601.

[74] Mohiuddin AB, Rahman MM, Gony MN, Shuvra SM, Rafy A, Ahmed MR, Rahman S. Accelerated and Accurate Cervical Cancer Diagnosis Using a Novel Stacking Ensemble Method with Explainable AI. In2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) 2025 Jul 31 (pp. 1-6).

[75] Malik AH, Rahman S. Toward precision wound healing: Integrating regenerative therapies and smart technologies. International Journal of Science and Research Archive. 2025 Sep 30;16(3):244–57.

[76] Malik AH, Rahman S. Hybrid Temozolomide Nanoconjugates: A polymer–drug strategy for enhanced stability and glioblastoma therapy. International Journal of Science and Research Archive. 2025 Sep 30;16(3):258–68.

[77] Malik AH, Rahman S. Molecular erasers: Reprogramming cancer immunity through protein degradation. World Journal of Advanced Engineering Technology and Sciences. 2025 Sep 30;16(3):277–91.

[78] Sultana S, Rahman MM, Hossain MS, Gony MdN, Rafy A. AI-powered threat detection in modern cybersecurity systems: Enhancing real-time response in enterprise environments. World Journal of Advanced Engineering Technology and Sciences. 2022 Aug 30;6(2):136–46.

[79] Abid SM, Xiaoping Q, Islam MM, Islam MA, Rahman MM, Alam ARM. Edge-Conditioned GAT for Journal Ranking in Citation Networks. 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology, AINIT 2025. 2025;1612–9.

[80] Chakraborty P. CatBoost-Stacked Heterogeneous Deep Ensembles for Explainable Multi-Class Brain Tumor MRI Classification. Journal of Medical and Health Studies. 2026 Mar 8;7(5):18–24.

[81] Linkon AA, Shakil MR, Ahmed S, Miah MR, Malik AH. Explainable Transformer-Based Skin Lesion Classification from Clinical Images. Journal of Medical and Health Studies. 2026 Mar 8;7(5):46–55.

[82] Ahmed S, Miah MR, Shakil MR, Linkon AA, Siddiqui MIH, Malik AH. Global–Local Attention Modeling for Reliable Multiclass Kidney Disease Classification from CT Images. Journal of Medical and Health Studies. 2026 Mar 8;7(5):36–45.

[83] Shakil MR, Malik AH, Siddiqui MIH, Ahmed S, Miah MR, Linkon AA. Swin Transformer–Driven Cervical Cell Classification with Explainable AI and Web-Based Screening. Journal of Medical and Health Studies. 2026 Mar 8;7(5):25–35.