Journal of Electronic Research and Application, 2026, Volume 10, Issue 8
http://ojs.bbwpublisher.com/index.php/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

Human–AI Collaborative Decision Model for Low-Voltage Diagnosis and Closed-Loop Mitigation in Distribution Transformer Areas

Biwei Li[1], Caiyu Zhang[2]*, Xianwen Zhang[1], Yexuan Liang[1]

[1]Zhaoqing Power Supply Bureau, Guangdong Power Grid Co., Ltd., Zhaoqing 526000, Guangdong, China
[2]Guangdong Technology College, Zhaoqing 526100, Guangdong, China

Corresponding author: Caiyu Zhang, 1509201014@qq.com

Abstract: Low-voltage mitigation in distribution transformer areas is rarely a single-step technical problem. Field diagnosis, plan selection, and post-implementation review are often handled separately, limiting the reuse of operational evidence and professional judgment. We recast mitigation as a six-stage human–AI closed loop linking state sensing, causal diagnosis, collaborative decision-making, implementation, effectiveness evaluation, and knowledge updating. For each candidate measure, the model considers data reliability, AI diagnostic confidence, human judgment reliability, and implementation risk. An expected-loss criterion assigns a machine-led, human–AI collaborative, or human-led mode. Safety and engineering constraints screen the available measures, after which an uncertainty-aware score identifies the preferred plan–mode pair. Evidence collected after implementation updates causal probabilities and the cause–measure knowledge base. Three constructed scenarios illustrate how decision authority moves from machine-led analysis toward professional control as implementation risk rises or diagnostic evidence weakens. These cases establish the internal consistency of the decision logic; they do not demonstrate improved field accuracy. The framework provides a transparent basis for subsequent calibration and validation with operational data.

Keywords: Distribution transformer area; Low-voltage mitigation; Human–AI collaboration; Decision support; Closed-loop feedback; Artificial intelligence

Online publication: Sept 3, 2026

1. Introduction

Low voltage in distribution transformer areas degrades supply quality, increases losses, and complicates refined operation and maintenance. It also varies across time and location because several factors may interact. Long supply radii, undersized conductors, unsuitable tap positions, and phase imbalance are among the recurrent causes, while violation limits should follow the applicable supply-voltage standard [1,2].

Recent work has used customer voltage, load, topology, and environmental data to mine violation patterns, forecast low-voltage events, analyze transformer voltages, and identify low-voltage network topology [1,3–6]. Intelligent fusion terminals and cloud–edge coordination further support multisource processing [7,8]. Most studies, however, concentrate on state identification or forecasting; the link from diagnosis to mitigation decisions and outcome feedback remains underdeveloped.

Mitigation decisions must also satisfy safety rules, cost limits, outage arrangements, equipment ratings, and engineering constraints. Machine learning can help grid operators classify events and prepare response options, yet explanation quality, confidence calibration, and generalization to unusual conditions remain concerns [9–12]. A fixed 'AI recommendation–human confirmation' workflow likewise fails to reflect the different risks and authority boundaries of individual measures [13–15].

This study therefore proposes a six-stage human–AI closed loop. It introduces a dynamic task-allocation rule driven by data reliability, AI diagnostic confidence, human judgment reliability, and implementation risk, together with a plan-evaluation and feedback mechanism that accounts for diagnostic uncertainty. Constructed scenarios are used to test the model's internal logic under different inputs rather than to claim an improvement in field accuracy.

2. Low-voltage mitigation process and problem definition

2.1. Existing mitigation process

Low-voltage cases are usually triggered by monitoring violations, customer complaints, or operational inspections. Utilities then retrieve data, inspect the site, infer causes, formulate measures, obtain approval, implement the work, and verify the result. This sequential workflow relies heavily on manual records, leaving weak structural links among diagnostic evidence, plan selection, and observed outcomes.

2.2. Causes and multisource data requirements

The main causes can be grouped into source-, transformer-, line-, load-, customer-, and operating-mode factors, which may also act jointly [1]. Diagnosis requires voltage measurements, transformer loading and tap positions, line parameters, three-phase currents, power factor, low-voltage topology, and maintenance records [4,7,8]. **Table 1** organizes these elements into a cause–feature–data–measure mapping for probabilistic diagnosis and candidate-plan generation.

Table 1. Mapping of low-voltage causes, observable features, data sources, and candidate mitigation measures

Cause category	Typical cause	Core diagnostic evidence	Main data	Candidate measures and baseline human involvement
Source	Low upstream voltage	Synchronous low voltage across related areas; both near and remote points are depressed	Upstream bus, transformer outlet, and related-area voltages	Recommend upstream voltage regulation, reactive support, or operating-mode changes; human approval for upstream-grid actions
Transformer	Overload, insufficient capacity, or unsuitable tap	Voltage follows loading, or remains area-wide low under light load	Capacity, loading, tap position, and outlet voltage	Capacity expansion, load transfer, or tap adjustment; human approval for investment, outages, and equipment operation

Line	Long radius, undersized conductor, or local impedance anomaly	Voltage declines with distance or drops abruptly between adjacent points	Low-voltage topology, line parameters, end voltages, and currents	Shorten the radius, reinforce the line, or inspect abnormal joints; field verification before engineering work
Load	Phase imbalance, concentrated growth, or insufficient reactive power	Marked phase differences; low voltage tracks load period or reactive demand	Three-phase voltage/ current, P/Q, power factor, and load curves	Rephase customers, shift load, or optimize reactive compensation; human review beyond authorized control limits
Customer	Impact load, service-wire issue, or customer-side equipment anomaly	Low voltage is localised and correlated with load switching or customer-side drop	Customer voltage, current, power curves, and work orders	Revise connection, reinforce service wire, or repair equipment; verify load type and ownership boundary on site
Operating mode	Unsuitable load allocation, topology, or equipment mode	Low voltage follows transfer, switching, or voltage-control device changes	DMS/OMS records, area loading, topology, and equipment status	Optimize service area, load transfer, or device mode; human approval for network operations
Combined state	Coupled factors	Several hypotheses have similar probabilities, and no single cause explains the case	Integrated multisource input	Generate combined or staged plans; raise baseline human involvement, with the final mode set by the allocation model

Note. Observable features are evidence for causal inference, not deterministic rules, and the listed measures are candidates rather than final decisions. Human-intervention requirements are operational priors; the actual mode is determined jointly by data reliability, AI confidence, human judgment reliability, and implementation risk.

2.3. Limitations of conventional manual mitigation

Professional staff provide indispensable field knowledge and safety oversight. Even so, cross-system data are costly to associate, coupled causes are difficult to distinguish, experience is hard to reuse systematically, and candidate measures often lack a structured comparison of effectiveness, economy, and implementation risk [1,7]. The aim is therefore not to replace practitioners, but to improve multisource integration and sustained reuse of professional knowledge.

2.4. Need for a human–AI closed loop

Machines are well suited to data processing, anomaly screening, probabilistic diagnosis, and plan ranking; professionals are better placed to verify field evidence, review safety constraints, judge engineering feasibility, and assume decision responsibility [14,15]. Decision authority should therefore vary with data quality, diagnostic confidence, and plan risk, while observed outcomes should return to the diagnostic and decision process.

3. Human–AI closed-loop architecture for low-voltage mitigation

3.1. Overall architecture

To connect the otherwise fragmented stages of low-voltage mitigation, the proposed architecture comprises state sensing, cause diagnosis, collaborative decision-making, measure implementation, effectiveness evaluation, and knowledge updating (**Figure 1**). Operational flows connect the stages, multisource data support diagnosis and evaluation, and measured outcomes and human feedback update the knowledge state [7,8,13].

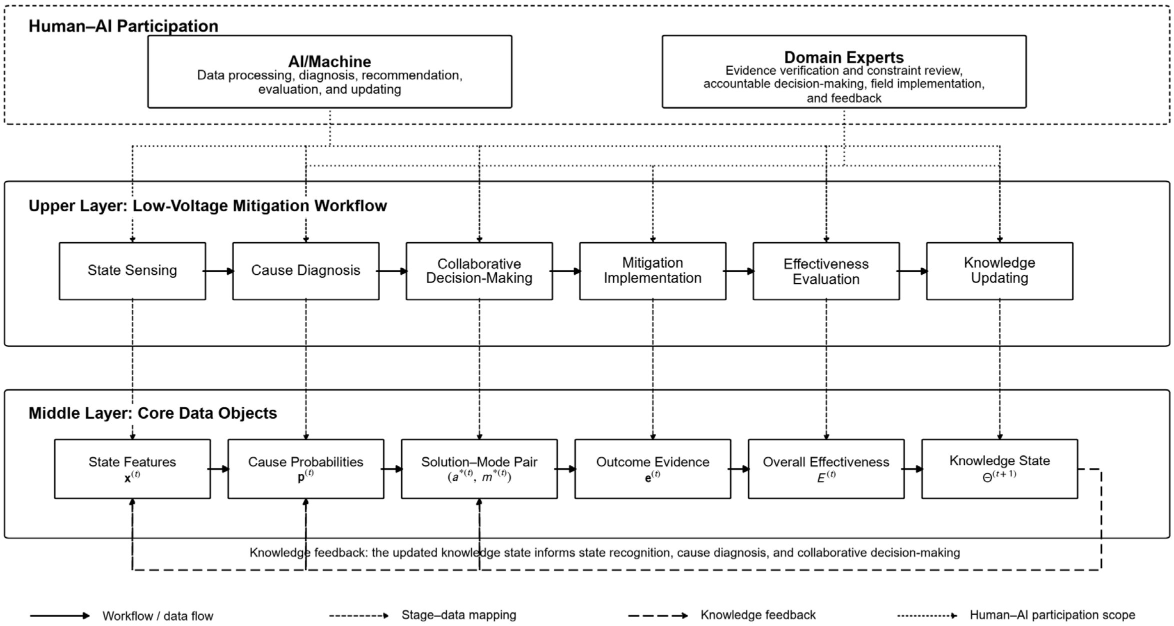

Figure 1. Overall architecture of human–AI collaborative closed-loop low-voltage mitigation in distribution transformer areas.

Note. Solid arrows denote the operational flow, dotted arrows map stages to core data objects, and the feedback loop carries updated knowledge. Knowledge supports the next round of identification, diagnosis, and decision-making but does not alter the physical measurements directly.

3.2. Six-stage closed-loop process

At mitigation round t, sensing produces a feature vector; diagnosis returns a probability distribution; collaborative decision-making selects the best plan–mode pair; implementation and evaluation produce outcome evidence; and knowledge updating revises the basis for subsequent diagnosis and decisions:

$$\boldsymbol{x}^{(t)} \to \boldsymbol{p}^{(t)} \to \left(a^{*(t)}, m^{*(t)}\right) \to \boldsymbol{e}^{(t)} \to E^{(t)} \to \Theta^{(t+1)} \tag{1}$$

The next diagnosis is driven jointly by new measurements and the updated knowledge state:

Here, x(t+1) is obtained from the next set of measurements and state sensing, whereas Θ(t+1) contains revised causal knowledge, mitigation rules, plan outcomes, and records of human–AI collaboration. Together they form the input to the next diagnostic round.

$$\left\{\boldsymbol{x}^{(t+1)}, \Theta^{(t+1)}\right\} \to \boldsymbol{p}^{(t+1)} \tag{2}$$

3.3. Functional boundaries between humans and AI

Human and machine roles are not reduced to a fixed ‘recommend–confirm’ pattern. They are assigned dynamically according to information reliability, implementation risk, and operational authority, as summarised in **Table 2** [11,12,14,15]. Machine-led operation is limited to authorized, low-risk analytical or operational adjustments. Outages, investment, field electrical work, and network reconfiguration remain subject to established dispatch, maintenance, and safety procedures.

Table 2. Human and AI responsibilities in the main stages of low-voltage mitigation

Stage	AI responsibilities	Professional responsibilities	Collaborative output
Cause diagnosis	Analyze multi-source features, detect anomalies, estimate causal probabilities and confidence, rank evidence	Verify field evidence, supply missing information, identify exceptional factors, correct abnormal diagnoses	Causal probabilities, key evidence, and correction records
Mitigation decision	Generate candidates, flag constraint conflicts, score plans and modes; handle authorized low-risk tasks	Set authority boundaries and review safety and engineering constraints when human involvement is triggered	Plan–mode pair (a*, m*)
Outcome feedback	Calculate effectiveness and prediction error, identify re-diagnosis conditions, propose updates	Confirm actual outcomes, explain deviations, and record revisions or rejections	Outcome evidence, knowledge updates, and re-diagnosis information

3.4. Information exchange and feedback

An AI module should report more than a single diagnosis. Its output should include causal probabilities, key evidence, candidate measures, constraint conflicts, and uncertainty. Human confirmations, revisions, rejections, and reasons should be recorded in structured form alongside implementation records and measured outcomes. Feeding these records back to the knowledge base makes the interaction explainable, correctable, and capable of continued refinement [10,11,13].

4. Low-voltage state identification and cause diagnosis

Cause diagnosis links state sensing with mitigation decisions. Rather than training a particular learning algorithm, this study defines a common interface from multisource features to causal probabilities and diagnostic confidence for downstream plan generation and task allocation.

4.1. Low-voltage state representation

Voltage violations are identified against the relevant supply-voltage standard, while their magnitude, duration, and affected range describe the operating state [1,2]. The feature vector includes minimum voltage, violation duration and frequency, share of affected users, voltage and current imbalance, transformer loading, power factor, and time-of-occurrence:

$$\boldsymbol{x}=\left[U_{\min},T_{lv},N_{lv},\rho_u,\eta_U,\eta_I,K_T,\cos\varphi,\tau_t\right] \tag{3}$$

where τt denotes the occurrence period or its time category. All indicators are standardized before diagnosis; no single violation value is treated as sufficient evidence of cause.

4.2. Cause classification

The taxonomy distinguishes six physical categories: source, transformer, line, load, customer, and operating mode. Examples include low upstream voltage, transformer overloading, unsuitable tap settings, long supply radii, undersized conductors, local impedance anomalies, phase imbalance, insufficient reactive power, and time-concentrated load growth [1]. Voltage regulation, reactive compensation, phase balancing, storage dispatch, and network reinforcement address different mechanisms [16–20]. The six categories may occur alone or in combination; a coupled case is not treated as a seventh physical category. See **Figure 2**.

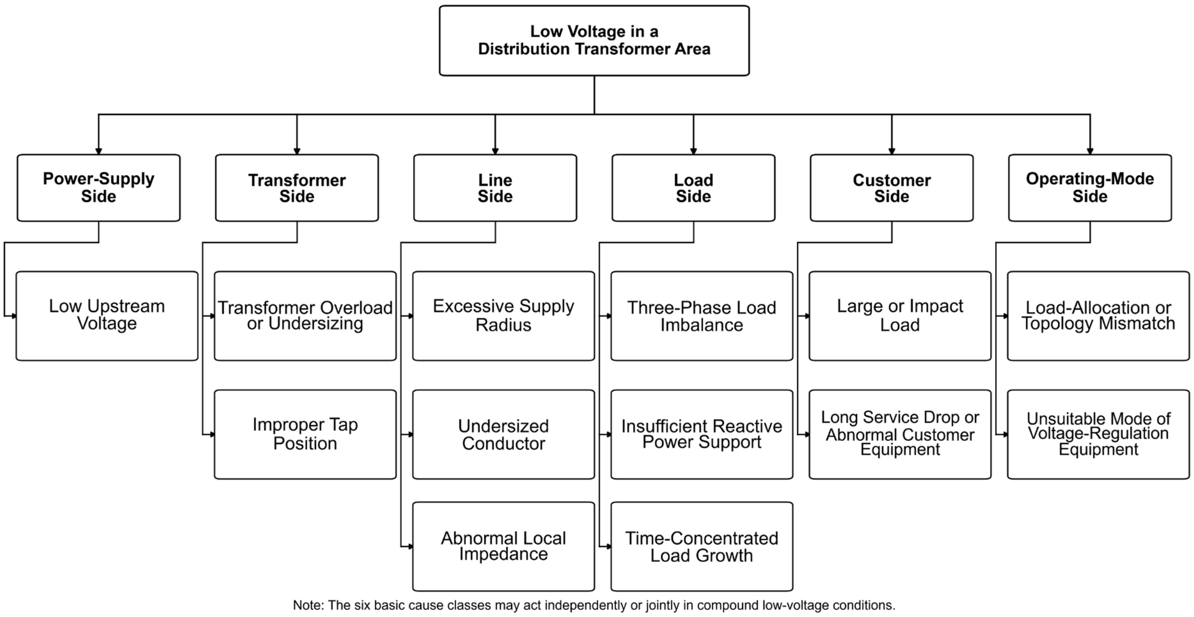

Figure 2. Classification of low-voltage causes in distribution transformer areas.

4.3. Mapping multisource features to causes

No single measurement uniquely identifies a low-voltage cause. Electrical quantities, asset parameters, low-voltage topology, and maintenance records must be considered together [3–8]. A normal sending-end voltage accompanied by distance-dependent decline points towards a line factor; coincident peak loading and area-wide voltage decline support transformer overloading; and phase-specific violations warrant checking phase imbalance. **Table 1** supplies diagnostic evidence, while supervised learning, probabilistic inference, or knowledge-based reasoning may implement the mapping. No particular algorithm is prescribed here.

4.4. Causal probabilities and AI diagnostic confidence

Let C = {C1, C2, …, Cn} denote mutually exclusive diagnostic-state hypotheses. A hypothesis may represent one basic cause or a predefined combination of causes. Given x, the diagnostic module returns the posterior probability of each hypothesis:

$$p_j = P\left(C_j \mid \boldsymbol{x}\right), \sum_{j=1}^{n} p_j = 1 \tag{4}$$

The leading diagnostic state is C* = arg maxj pj.

To reflect both the advantage of the leading hypothesis and the concentration of the full distribution, this study defines AI diagnostic confidence as:

$$q = \alpha \max_j p_j + \left(1-\alpha\right)\left(1 + \frac{\sum_{j=1}^{n} p_j \ln p_j}{\ln n}\right), \quad 0 \le \alpha \le 1 \tag{5}$$

The convention 0 ln 0 = 0 is used, and $n \geq 2$. The second term equals one minus normalized entropy; it approaches 1 as the distribution becomes more concentrated. Thus q measures diagnostic certainty for the current input rather than empirical accuracy [10–12,21]. This confidence measure is a theoretical interface whose parameter α and functional form require calibration with historical diagnostic data. Section 3 therefore provides $x \rightarrow (p, q)$; Section 4 combines p and q with data reliability d, human judgment reliability h, and plan risk ri to obtain (a*, m*).

5. Human–AI collaborative mitigation decision model

For each candidate measure, the model combines data reliability, AI diagnostic confidence, human judgment reliability, and implementation risk to assign decision authority and select the best feasible plan–mode pair.

5.1. Data, machine, and human reliability

Data reliability d, AI confidence q, and human judgment reliability h are normalized to [0, 1]. Completeness, timeliness, consistency, and measurement credibility are aggregated as:

$$d = \sum_{k=1}^{K} \alpha_k d_k, \qquad \sum_{k=1}^{K} \alpha_k = 1 \tag{6}$$

Here dk and αk are the kth data-quality indicator and its weight. The product dq is the effective machine reliability after accounting for data quality. Human reliability h is not an administrative score; it represents agreement between prior professional judgments and subsequent outcomes in comparable cases, the match between expertise and the current task, and the adequacy of field verification [11,12,14]. To avoid double counting, h excludes completeness, timeliness, and consistency already represented by d.

5.2. Expected loss under alternative decision modes

Building on the principle that automation should vary with task properties, the model considers machine-led (M), human–AI collaborative (HM), and human-led (H) modes [14,15]. Their reliabilities are defined as:

$$\rho_M = dq, \qquad \rho_H = h, \qquad \rho_{HM} = \beta dq + (1-\beta)h, \quad 0 \leq \beta \leq 1 \tag{7}$$

The coefficient β is the machine's relative contribution in collaborative decisions and can be calibrated by task and historical performance. The model does not assume that ρHM must exceed ρM or ρH. Collaboration combines the two information sources and, in higher-risk cases, uses human review to limit risk exposure. The expression is a calculable theoretical interface; its form and parameters require validation with historical mitigation records. For candidate plan ai, the expected loss under mode m is:

$$L_m = C_e (1-\rho_m) + C_s \gamma_m r_i + C_m \tag{8}$$

$$C_M = 0, \qquad C_H = C_t, \qquad C_{HM} = C_t + C_c \tag{9}$$

Ce and Cs denote the losses associated with an incorrect decision and safety risk; ri ∈ [0, 1] is the implementation risk of plan ai; and γm is the mode-specific exposure coefficient. If human review reduces procedural exposure, γM > γHM ≥ γH. Cm is the incremental human and coordination cost above an existing intelligent system. Accordingly, CM = 0 means zero incremental human cost, not cost-free computation; Ct and Cc represent human review and coordination costs. These loss terms are defined here for allocating decision authority in low-voltage mitigation.

5.3. Candidate-plan generation

From the posterior probabilities pj = P(Cj | x), retain the diagnostic hypotheses that satisfy:

$$J = \left\{ j \middle| \ p_j \geq \tau_p \right\} \tag{10}$$

where τp is the screening threshold. The cause–measure mapping in Table 1 then generates:

$$A = \bigcup_{j \in J} A\left(C_j\right) \tag{11}$$

Candidates include tap adjustment, phase-load balancing, reactive compensation, load transfer, feeder reinforcement, and transformer capacity expansion. The AI produces a set rather than a single prescription. Hypotheses below τp are excluded from separate evaluation, and the probabilities retained in J are renormalized but continue to be denoted pj [16–19].

5.4. Dynamic allocation of human and AI tasks

For each candidate, operational attributes, institutional rules, and baseline human-intervention requirements define the admissible mode set Ωi. The selected mode minimizes expected loss:

$$m_i^* = \arg\min_{m \in \Omega_i} L_m \tag{12}$$

A hard constraint applies to high-investment plans, outage work, and measures that affect operational safety:

$$r_i \geq r_{high} \Rightarrow M \notin \Omega_i \tag{13}$$

Thus Ωi and mi* are assigned to each measure rather than to a cause category in advance. Different measures for the same diagnosed cause may carry different authority settings [9,11,12,14].

5.5. Constraints and plan evaluation

Feasible plans Af are first screened against nodal-voltage limits, conductor ampacity, transformer capacity, budget, allowable outage duration, and safe-operation requirements. Plan ai then receives a score S(ai | Cj) under hypothesis Cj based on expected voltage improvement, economy, engineering feasibility, reliability benefit, and implementation risk. Accounting for diagnostic uncertainty, its expected score is:

$$\bar{S}_i = \sum_{j \in J} p_j \, S(a_i \mid C_j) \tag{14}$$

Let i* identify the highest-scoring feasible plan. The selected plan and its associated mode are:

$$i^* = \arg\max_{i:\, a_i \in A_F} \bar{S}_i, \quad (a^*, m^*) = (a_{i^*}, m_{i^*}^*) \tag{15}$$

The final output is therefore the plan–mode pair (a*, m*). Plan evaluation answers which measure should be adopted; dynamic task allocation determines how its decision authority should be shared. Expected loss configures authority, whereas the comprehensive score ranks engineering plans. They are not added, which avoids mixing different units or counting risk twice.

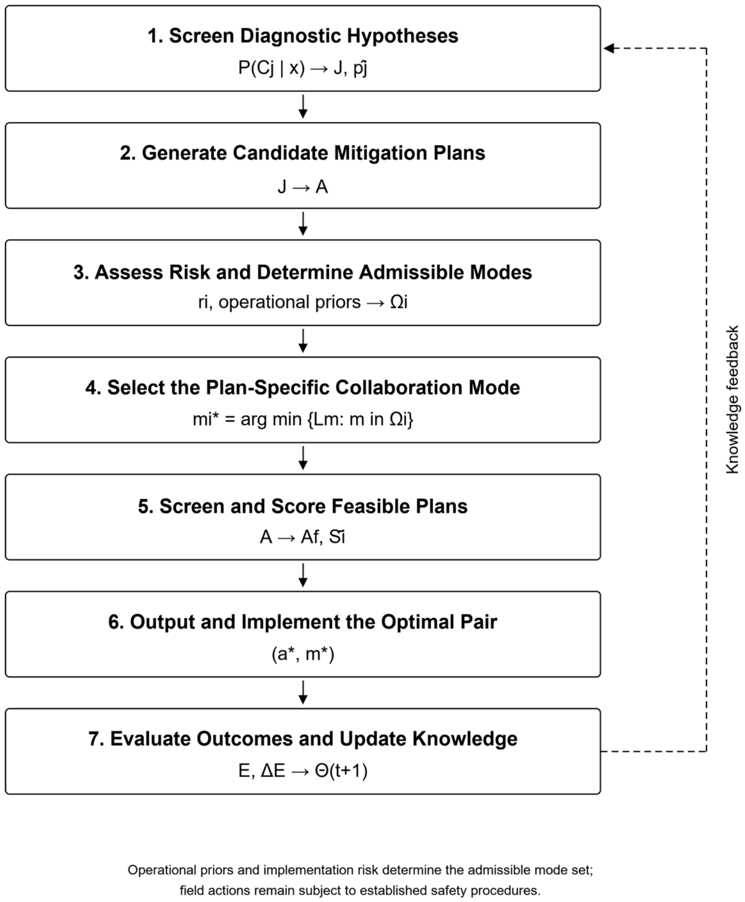

Figure 3. Human–AI collaborative decision-making process for low-voltage mitigation.

Note. 'Decision-making and implementation under the selected mode' refers to plan review and authority routing. Any field electrical work remains governed by existing dispatch, maintenance, and safety rules.

6. Effectiveness evaluation and knowledge updating

To move beyond one-off recommendations, observed mitigation outcomes are returned to diagnosis and decision-making through effectiveness evaluation, knowledge updating, and re-diagnosis.

6.1. Evaluation indicators

Effectiveness is evaluated in seven dimensions: voltage improvement, shorter violation duration, fewer affected users, lower network loss, reduced phase imbalance, mitigation cost, and residual operational risk [1,16,17,19]. All indicators are normalized to [0, 1]. The first five are benefits, whereas cost and residual risk are penalties:

$$E = \mu_1 E_U + \mu_2 E_T + \mu_3 E_N + \mu_4 E_L + \mu_5 E_I - \mu_6 E_C - \mu_7 E_R$$

$$\sum_{k=1}^{7} \mu_k = 1, \qquad \mu_k \geq 0 \tag{16}$$

EU, ET, EN, EL, and EI denote improvements in voltage, violation duration, affected users, network loss, and phase imbalance; EC is mitigation cost; and ER is the operating-safety risk remaining after implementation. Hence, ri in Section 4 is an ex ante implementation risk used for authority allocation, whereas ER is an ex post outcome indicator.

6.2. Judging plan effectiveness

Let Emin < Etar. A plan meets its target when E ≥ Etar and no key indicator deteriorates. It is partly effective when Emin ≤ E < Etar, or when voltage recovers while network loss, imbalance, or another key measure worsens. It fails when E < Emin. The prediction error is:

$$\Delta E = E_{act} - E_{pre} \tag{17}$$

6.3. Updating causal probabilities and knowledge rules

Following principles of calibrated trust, productive delegation, and correctable interaction, both outcomes and human feedback enter the update process [12–15]. Let e(t) be the multi-indicator evidence after round t. For the implemented plan a*, causal probabilities are updated by Bayes' rule:

$$p_j^{(t+1)} = \frac{P\left(e^{(t)} \mid C_j, a^*\right) p_j^{(t)}}{\sum_k P\left(e^{(t)} \mid C_k, a^*\right) p_k^{(t)}} \tag{18}$$

P(e(t) | Cj, a*) is the likelihood of observing the current outcome if Cj holds and a* is implemented. It should be estimated from comparable historical cases; expert elicitation may initialize the model when samples are scarce. Agreement between predicted and observed outcomes raises confidence in the corresponding cause–measure rule. A marked discrepancy lowers that confidence and triggers a check for omitted or coupled causes. Human confirmations, revisions, and rejections are also retained for rule correction and trust calibration.

6.4. Closed-loop re-diagnosis

Re-diagnosis is triggered when E < Emin, |ΔE| > ε, or new low-voltage violations appear after mitigation. Additional information updates d, q, and P(Cj | x) before the process repeats. Successful cases are stored as 'scenario–cause–measure–mode–outcome' records. The loop is not indefinite: it ends when the target is met, or a professional terminates the process; insufficient data suspend automation and prompt field review. Subsequent electrical work remains subject to established operating and safety rules.

7. Analysis of constructed scenarios

7.1. Scenarios and parameter settings

Three scenarios are constructed to test authority allocation under different data quality, diagnostic confidence, and implementation risk; no claim is made about AI accuracy. The settings are β = 0.60, Ce = 1.00, Cs = 0.40, Ct = 0.10, Cc = 0.02, γM = 1.00, γHM = 0.25, γH = 0.20, h = 0.80, and rhigh = 0.70. Expected plan scores in **Table 3** are normalized illustrative values based on effectiveness, economy, feasibility, reliability benefit, and risk. All parameters serve only to demonstrate model behavior and require calibration from historical cases, project costs, and decision records before field use.

Table 3. Human–AI collaborative decisions in typical low-voltage mitigation scenarios

Scenario and focal plan	Leading raw posterior(s)	d/q/h	Risk ri	Admissible modes Ωi	LM/LHM/LH	Selected mode mi*	Expected score $\bar{S_i}$
A1: Low-risk operational adjustment	Excessive supply radius 0.90	0.95/0.90/0.80	0.10	{M, HM, H}	0.185/0.297/0.308	M	0.70
A2: Low-voltage feeder reinforcement	Excessive supply radius 0.90	0.95/0.90/0.80	0.40	{HM, H}	—/0.327/0.332	HM	0.82
B: Transformer capacity expansion	Transformer overload 0.90	0.95/0.90/0.80	0.85	{HM, H}	—/0.372/0.368	H	0.84
C: Staged combined mitigation	Three-factor 0.38; imbalance–radius 0.34; radius–impedance 0.18	0.55/0.40/0.80	0.65	{HM, H}	—/0.733/0.352	H	0.79

Note. Causal probabilities are raw posteriors. Probabilities exceeding the threshold are renormalized as specified in Section 4.3. Risk, loss, and plan scores are constructed values rather than field statistics or engineering accuracy. A dash indicates a mode excluded by operational priors. Machine-led mode covers only authorized low-risk analysis or operational adjustment, not direct field switching. Scenario C reports the three leading mutually exclusive diagnostic hypotheses; the remaining hypotheses sum to 0.10, and a hypothesis may itself contain several basic causes.

7.2. Single-cause scenario

In Scenario A, the sending-end voltage is normal, but remote users remain undervoltage; the posterior probability of an excessive supply radius is 0.90 [1]. Low-risk operational adjustment admits all three modes and selects M because LM is the smallest. Feeder reinforcement has engineering consequences, so M is excluded and HM is selected. Its higher expected score yields the final pair 'feeder reinforcement–human–AI collaboration'. The example shows that two measures for the same cause can carry different authority settings.

7.3. High-risk mitigation scenario

Scenario B has diagnostic inputs close to Scenario A, but transformer capacity expansion entails investment, outage work, and a change in supply arrangements, raising ri to 0.85. The hard constraint removes M. The remaining losses are LHM = 0.372 and LH = 0.368, so H is selected. Their difference of 0.004 places the case near the HM–H boundary; under the illustrative parameters, human-led decision-making is only slightly preferred. High AI confidence therefore indicates a concentrated diagnosis, not independent authority over a high-risk plan. The boundary must be calibrated from operational records and safety requirements.

7.4. Coupled causes and low confidence

Scenario C combines phase imbalance, an excessive supply radius, and a local impedance anomaly with incomplete measurements. Among predefined mutually exclusive states, the posteriors are 0.38 for the three-factor combination, 0.34 for imbalance plus excessive radius, and 0.18 for excessive radius plus local impedance. The similar values indicate that the available evidence cannot yet distinguish the coupled state. With $d = 0.55$ and $q = 0.40$, effective machine reliability is:

$$\rho_M = dq = 0.55 \times 0.40 = 0.22$$

Collaborative reliability is:

$$\rho_{HM} = 0.60 \times 0.22 + 0.40 \times 0.80 = 0.452$$

The resulting losses, LHM = 0.733 and LH = 0.352, strongly favour H. When data quality is poor and competing compound hypotheses remain unresolved, the model deliberately shifts authority towards professionals. Field checks of conductor size, phase allocation, and connection condition can then update d, q, and P(Cj | x), after which the AI regenerates and reranks the plans. This sequence—human inspection, data completion, AI re-diagnosis, and plan re-evaluation—closes the loop [1,10–12,17].

7.5. Comparison of collaboration modes

Table 3 traces an M → HM → H shift. Machines carry more analytical and decision work when information is reliable, and plan risk is low. Engineering consequences move authority to collaboration; greater implementation risk or diagnostic uncertainty moves it further towards professionals. These results establish logical consistency under the stated inputs, not higher diagnostic accuracy or field effectiveness. Loss parameters, risk thresholds, evaluation weights, and mode boundaries still require calibration with real operating and mitigation data.

8. Conclusions and outlook

A six-stage human–AI closed loop was developed for low-voltage mitigation in distribution transformer areas. The model allocates tasks based on data reliability, AI confidence, human judgment reliability, and implementation risk, and it links uncertainty-aware plan evaluation with outcome-based knowledge updating. Constructed scenarios demonstrate coherent shifts in decision authority as information quality and risk change. The work remains a theoretical and logical validation: human reliability, loss parameters, and evaluation weights require real business data. Future work will test the model on operating data, train diagnostic algorithms, examine parameter sensitivity, and develop a decision-support platform.

Author biographies

Biwei Li (born 1996) received the M.S. degree and is an engineer. His research focuses on distribution-network fault management. Email: 498315697@qq.com. Corresponding author: Caiyu Zhang (born 1994) received the M.S. degree. Email: 1509201014@qq.com

Disclosure statement

The authors declare no conflict of interest.

Reference

[1] Shen S, Zhong Q, Xu Z, et al., 2025, Mining User Low-Voltage Violation Patterns Using Hierarchical Affinity-Propagation Clustering. Electric Power Engineering Technology, 44(1): 30–38.

[2] General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of China, 2008, Power Quality—Deviation of Supply Voltage. Standards Press of China.

[3] Guo S, Zhao X, Sun G, et al., 2025, Analysis and Prediction of Over-Limit Characteristics for Regional Distribution-Transformer Voltage Using Bilayer Clustering and Correlation-Feature Screening. Southern Power System Technology, 19(2): 19–27.

[4] Huang Y, Yang S, Cao X, et al., 2023, Topology Identification of Low-Voltage Distribution Networks Using Segmented Current Features and a Random-Forest Algorithm. Power Demand Side Management, 25(2): 63–69.

[5] Wei J, Tian S, Chen W, et al., 2025, Low-Voltage Prediction in Distribution Transformer Areas Using a GA–BP–LSTM Model. Electric Engineering, 46(1): 15–20.

[6] Zhou Y, Zhang B, Huang W, et al., 2023, Low-Voltage Prediction for Distribution Transformer Areas Using an LSTM–BP Combined Model. Journal of Electric Power Science and Technology, 38(5): 177–186.

[7] Chen W, Chen J, Fan Y, et al., 2024, Design and application of smart distribution station areas based on intelligent fusion terminals with cloud-edge coordination. Electric Power, 57(4): 190–199.

[8] Fan Y, Wu H, Li Z, et al., 2024, A flexible orchestration of lightweight AI for edge computing in low-voltage distribution network. Frontiers in Energy Research, 12: 1424663.

[9] Campos V, Klyagina O, Andrade J, et al., 2024, ML-Assistant for Human Operators Using Alarm Data to Solve and Classify Faults in Electrical Grids. Electric Power Systems Research, 236: 110886.

[10] Machlev R, Heistrene L, Perl M, et al., 2022, Explainable artificial intelligence techniques for energy and power systems: review, challenges and opportunities. Energy and AI, 9: 100169.

[11] Naiseh M, Al-Thani D, Jiang N, et al., 2023, How the Different Explanation Classes Impact Trust Calibration: The Case of Clinical Decision Support Systems. International Journal of Human-Computer Studies, 169: 102941.

[12] Yamada S, 2022, Trust Calibration as Rationality for Human–AI Cooperative Decision Making. Cognitive Studies, 29(3): 364–370.

[13] Amershi S, Weld S, Vorvoreanu M, et al., 2019, Guidelines for Human–AI Interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13.

[14] Fügener A, Grahl J, Gupta A, et al., 2022, Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. Information Systems Research, 33(2): 678–696.

[15] Parasuraman R, Sheridan T, Wickens C, 2000, A Model for Types and Levels of Human Interaction with Automation. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 30(3): 286–297.

[16] Guo X, Zheng F, Wang S, 2024, Mobile Energy-Storage Scheduling in Distribution Networks Considering Voltage-Violation Risk. Power Demand Side Management, 26(4): 37–42.

[17] Li B, Li Y, Hai Z, et al., 2025, Online Mitigation of Three-Phase Imbalance in Distribution Networks Using Multi-Agent Deep Reinforcement Learning. Proceedings of the CSEE, 45(5): 1729–1741.

[18] Yang Y, Wen F, Zhou X, et al., 2023, Review of Coordinated Voltage Control in Distribution Systems with High Photovoltaic Penetration. Electric Power Automation Equipment, 43(10): 48–58.

[19] Yuan B, Yang P, Wang T, et al., 2025, Multi-Objective Dispatch of Distributed Energy Storage in High-PV Distribution Transformer Areas Considering Multiple Stakeholders. Modern Electric Power, 42(6): 1268–1277.

[20] Zhang X, Wu Z, Sun Q, et al., 2024, Application and Progress of Artificial Intelligence Technology in the Field of Distribution Network Voltage Control: A Review. Renewable and Sustainable Energy Reviews, 192: 114282.

[21] Schemmer M, Kühl N, Benz C, et al., 2023, Appropriate Reliance on AI Advice: Conceptualization and the Effect of Explanations. Proceedings of the 28th ACM International Conference on Intelligent User Interfaces, 410–422.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.