

Revision Priority & Action Tables

Bng u Tiên Sa Bài và Hành Đng C Th

Revision Plan for Manuscript 6863b9b0-4db8-4b53-843f-5be5e907cf62

February 5, 2026

Tóm tt tình hung

- **Decision:** Major Revision (c hi accept CAO nu sa tt)
- **S reviewers:** 8 reviewers
- **Thi hn:** 10 ngày làm vic (có th xin thêm 5-7 ngày)
- **Chin lc:** u tiên sa các vn đ CRITICAL trc, sau đó MAJOR

1 BNG 1: Priority Matrix - Ma Trn u Tiêu

1.1 Bng 1A: Tng quan các vn đ cn sa

Table 1: Ma trn u tiênsa bài - Nhóm theo mc đ nghiêm trng

#	Vn đ	Mc đ	Rev.	Thi gian	u tiênsa thc hin
CRITICAL - Chí mng (không sa → b reject)					
1	”Overall across schemas” m h	FATAL	R6, R8	0.5 ngày	1 - LÀM NGAY
2	COCOMO II baseline không fair	FATAL	R1, R7, R8	2-3 ngày	2 - BT BUC
3	Target leakage: Developers = Effort/Time	FATAL	R8	0.5 ngày	3 - NHANH
4	FP n=24 - protocol không hp lý	FATAL	R6, R7, R8	1 ngày	4 - CN LÀM
5	Dataset manifest / provenance thi	FATAL	R7, R8	1 ngày	5 - QUAN TRNG
6	Formatting: no captions, m, thi line #	FATAL	R5, R6, R7	1 ngày	6 - D LÀM
MAJOR - Quan trng (không sa → khó accept)					
7	Equation/label li & lp Time	MAJOR	R6	0.5 ngày	7
8	R ² ct ”–” trong bng	MAJOR	R6	0.5 ngày	8
9	Novelty yu (“pipeline thoi”)	MAJOR	R1, R3, R8	1 ngày	9 - QUAN TRNG
10	Related Work thi so sánh & DOIs	MAJOR	R3, R4, R5	1 ngày	10 - CN LÀM
11	SOTA models thi (XGBoost/LightGBM)	MAJOR	R4, R7	1-2 ngày	11 - nu kp
12	Interpretability claim - thi phân tích	MAJOR	R7	1 ngày	12
13	Ablation study thi	MAJOR	R5, R7	1 ngày	13
14	Generalization: leave-one-out	MAJOR	R7, R8	1-2 ngày	14 - nu kp
MINOR - Ci thin (đ tăng đim)					
15	Language quality / AI-like tone	MINOR	R4, R7	1 ngày	15 - polish
16	Additional metrics (MAPE, MdMRE, RAE)	MINOR	R1	0.5 ngày	16 - nhanh
17	Confidence intervals (bootstrap CI)	MINOR	R1	1 ngày	17
18	Reduce length → Supplementary	MINOR	R1	1 ngày	18

1.2 Bng 1B: Timeline khuyn ngh (10 ngày)

Table 2: Phân b thi gian 10 ngày

Ngày	Công việc	Priority IDs
1-2	CRITICAL block 1: Overall aggregation + COCOMO calibrate + leakage + FP protocol	1, 2, 3, 4
3	CRITICAL block 2: Dataset manifest + Formatting	5, 6
4-5	MAJOR block 1: Novelty + Related Work + DOIs + equations	7, 8, 9, 10
6-7	MAJOR block 2: Interpretability + Ablation + metrics	12, 13, 16, 17
8	OPTIONAL: XGBoost (nu kp) hoc polish language	11, 15
9	Integration & Review: Tích hp tt c, check consistency	All
10	Final check & Submit: Thy review, finalize response letter	All

2 BNG 2: Action Tables - Bng Hành Đng Chi Tit

2.1 Bng 2A: Response to Reviewers - Tr li tng câu hi

Table 3: Point-by-point response structure

Rev	Câu hi / Yêu cu	Response Strategy
REVIEWER 1 - Methodological Concerns		
R1.1	Clarify novelty beyond "unified pipeline"	Action: Vit li Abstract + Intro nhn mn 3 novelties: (1) Multi-schema harmonization vi auditable manifest, (2) Fair calibrated baseline, (3) Statistical rigor (Wilcoxon + Cliff's Delta) Changes: Abstract lines 8-12, Intro Section 1.3, NEW Table so sánh vi prior work
R1.2	Recalibrate COCOMO II for fair comparison	Action: Fit A, B parameters trên training data tng schema → báo cáo COCOMO II (original) vs. (calibrated) vs. RF Changes: NEW Methods subsection "Baseline Calibration", NEW Table 3 kt qu, Discussion update
R1.3	Add modern datasets (GitHub/Jira)	Strategy: (A) Nu kp: thu thp mini-validation 30-50 projects t GitHub vi effort trong README → report MMRE. (B) Nu khôn kp: gii thích limitation và da Future Work Changes: Data section + Results hoc Limitations
R1.4	Report MAPE, MdMRE, RAE	Action: Thêm 3 metrics vào code, update tt c tables Changes: Section 2.4 (Metrics), Table 1, 3, 4
R1.5	Confidence intervals for all metrics	Action: Bootstrap 95% CI hoc CI t 10 seeds → format "Mean [95% CI]"

Continued on next page

Rev	Câu hỏi / Yêu cầu	Response Strategy
		Changes: All result tables, add error bars to figures
R1.6	Reduce length → Supplementary	Action: Chuẩn hóa hyperparameter grids, dataset details, extended stats sang Supplementary PDF Changes: Main text giảm t 38 → 28 pages
R1.7	Release data & code	Action: To GitHub repo + Zenodo DOI Changes: Data Availability section

REVIEWER 3 - Structure & Positioning

R3.1	Intro: novelty unclear	Action: Restructure Introduction: What is known → Gap → Our contribution → Why it matters Changes: Section 1 (3 subsections)
R3.2	Related Work thiếu so sánh + cite 4 DOIs	Action: (1) Cite 4 papers h suggest, (2) NEW comparison table (8 studies), (3) Discuss gap Papers: aisly.202300706 (AI in SE), patcog.112890 (ML methods), ACCESS.3480205 (SEE MMRE=1.15), engappai.111655 (ensemble) Changes: NEW Section 2: Related Work (3 pages)
R3.3	Highlight assumptions & limitations	Action: NEW Section 3.6: 5 assumptions + 5 limitations vi mitigation Changes: 2-page section in Methods
R3.4	Figure 1 description unclear	Action: Viết li caption chi tiết 15 lines workflow, redesign figure vi color-code Changes: Figure 1 caption & figure
R3.5	Conclusion: add strengths/weaknesses/implications	Action: Restructure Conclusion thành 4 subsections: Findings, Limitations, Implications, Future Changes: Section 7 expand 2 → 4.5 pages

REVIEWER 6 - Technical Details

R6.1	Abstract: "overall" ambiguous	Action: Clarify "macro-average across LOC/FP/UCP" in abstract
R6.2	Equation labels missing	Action: Fix all <code>\label{eq:...}</code> references
R6.3	FP n=24 too small	Action: Acknowledge explicitly, recommend LOOCV for FP, report wider CIs
R6.4	R ² column "—"	Action: Compute R ² for all models or remove column
R6.5	Time equation duplicated	Action: Delete duplicate (line 124-128)
R6.6	"Enhanced COCOMO II" undefined	Action: Define in Methods or remove term

REVIEWER 7 - Rigorous Critique

R7.1	No figure captions	Action: Add captions to ALL figures/tables
R7.2	Writing "formulaic/AI-generated"	Action: Manual polish: reduce repetition, natural phrasing
R7.3	COCOMO baseline validity	Action: → Same as R1.2 (calibrate)
R7.4	SOTA models missing (XGBoost)	Action: Add XGBoost as 5th model nếu kp, else: explain scope

Continued on next page

Rev	Câu hỏi / Yêu cầu	Response Strategy
R7.5	Interpretability: no feature importance	Action: Add Gini/permutation importance figure + 1 paragraph
R7.6	Ablation study missing	Action: RF with: (no log) → +log → +IQR → full
R7.7	FP n=24 → test set 5 samples	Action: → Same as R6.3
R7.8	Generalization: same pool split	Action: Leave-one-source-out for LOC nu kp
R7.9	Figure anomalies (smooth FP curve)	Action: Check if showing model predictions (smooth) vs. actual data (scatter)

REVIEWER 8 - Deepest Technical Review

R8.1	"Overall across schemas" undefined	Action: → Same as Priority #1: Define macro/micro aggregation
R8.2	COCOMO baseline not reproducible	Action: → Same as R1.2 + R7.3
R8.3	Dataset provenance missing	Action: → Same as Priority #5: Dataset manifest table
R8.4	Developers = Effort/Time → leakage	Action: → Same as Priority #3: Don't use inferred Developers as feature
R8.5	FP protocol inappropriate	Action: → Same as Priority #4
R8.6	Limited novelty (RF & COCOMO known)	Action: → Same as R1.1 + R3.1
R8.7	No cross-schema learning	Action: Clarify: "schema-specific training" (not joint). Future: transfer learning
R8.8	Data imbalance not treated	Action: (A) Nu kp: add quantile regression or weighted loss. (B) Else: discuss in Limitations + Future
R8.9	Suggest focal loss / imbalance-aware	Action: Cite DOI 10.1038/s41598-025-22853-y, explain why not done (scope), add to Future

2.2 Bng 2B: LaTeX File Changes - Sa đâú trong main.tex

Table 4: Mapping: Issue → Location in LaTeX → How to fix

#	Vn đ	V trí trong main.tex	Cách sa
1	Overall aggregation unclear	Abstract (line 10) + Results before Table 1	Add: "Overall metrics computed as macro-average (unweighted mean) across LOC/FP/UCP test sets." + NEW per-schema table
2	COCOMO baseline	Section 2.1 + Section 4.2 (Modeling)	NEW subsection: "Baseline Calibration Protocol" → fit A, B on train → report original vs calibrated
3	Developers leakage	Section 3.2 (Unit Harmonization) + Feature list	REMOVE: Developers = $\text{ceil}(\text{Effort}/\text{Time})$ from features. Only use if in raw data
4	FP n=24 protocol	Section 4.1 (Train-Test)	For FP: "Due to small sample size (n=24), we employ leave-one-out CV and bootstrap 95% CIs. Results considered exploratory."
5	Dataset manifest	Section 3.1 (Sources)	NEW table: Source — Link/DOI — Schema — Raw # — After dedup — Final #
6	Formatting	Preamble + all figures	(1) Add <code>\usepackage{lineno}, \linenumbers;</code> (2) All figures: vector PDF 600dpi + proper captions
7	Equation li	Section 2.1 (lines 120-130)	Delete duplicate Time equation (line 124-128)
8	R ² missing	Table 1 (line 570)	Compute R ² or change caption to "R ² omitted for parametric baseline"
9	Novelty weak	Abstract + Intro Section 1.3	Rewrite: 3 clear novelties (harmonization + fair baseline + statistical rigor)
10	Related Work	NEW Section 2 (before Background)	Add 2-3 pages: cite 4 DOIs + comparison table (8 studies)
11	SOTA models	Section 4.2 (Models)	Add XGBoost: "We also evaluate XGBoost (v2.0) with learning rate 0.01-0.3, max depth 3-6..."
12	Interpretability	Results Section 5.3	NEW figure: Feature importance (Gini or permutation) for RF + 1 paragraph explaining top 3 features
13	Ablation study	Results Section 5.4 (NEW)	NEW table: RF performance with (raw) — (+log) — (+log+IQR) — (full pipeline)

14	Generalization	Experimental Setup 4.1	Add: "To assess cross-source generalization, we perform leave-one-dataset-out evaluation for LOC schema (train on A,B,C → test on D)."
----	----------------	------------------------	--

3 BNG 3: Quyết Định Chín Lớp (Cho Thy Mн)

3.1 Bng 3A: Các quyết định cần Thy cht NGAY

Table 5: Strategic decisions needed immediately

Quyết định	Options	Khuyn nghe
1. COCOMO recalibration	(A) Làm đợt đ (2-3 ngày) (B) Ch gii thích	Option A - Tăng c hi accept 15-20%. Chng minh RF tt hn c optimized baseline → stronger claim
2. GitHub/Jira data	(A) Thu thp mini-set (3-4 ngày) (B) Gii thích limitation	Option B - GitHub không có ground truth tt. Gii thích honest + đà Future Work là đ
3. XGBoost experiment	(A) Thêm XGBoost (1-2 ngày) (B) Không thêm	Option A nu kp - XGBoost là SOTA, d implement, tăng đим vi R4/R7. Nu không kp thì B + explain scope
4. Ablation study	(A) Làm đợt đ (1 ngày) (B) Đn gin hóa	Option A - D làm: ch cn chy RF vi 4 configs (raw/+log/+IQR/full) → 1 table. Reviewer R5/R7 yêu cu rõ
5. Xin thêm thi gian?	(A) Xin thêm 5-7 ngày (B) Không xin	Option A - Gn Tt, 10 ngày hi gp. Xin thêm 5 ngày (tng 15) đ làm tt hn. Journals thng cho

3.2 Bng 3B: Phân công Huy vs. Thy

Table 6: Work division

Công việc	Huy (technical)	Thy (strategic)
Days 1-2: CRITICAL	<ul style="list-style-type: none"> - Code: aggregation definition, COCOMO calibration, remove Developers leakage, FP LOOCV - LaTeX: fix equations, formatting 	Review overall strategy
Days 3-4: Data & Novelty	<ul style="list-style-type: none"> - Create dataset manifest - Draft novelty text (Intro) 	<ul style="list-style-type: none"> - Edit Introduction - Review novelty positioning
Days 5-6: Experiments	<ul style="list-style-type: none"> - XGBoost implementation - Ablation study - Feature importance plots 	Strategy: decide if XGBoost or skip
Days 7-8: Related Work & Writing	<ul style="list-style-type: none"> - Read 4 suggested papers - Create comparison table 	<ul style="list-style-type: none"> - Write Related Work section - Integrate citations
Days 9-10: Integration	<ul style="list-style-type: none"> - Integrate all changes - Consistency check - Generate all figures 	<ul style="list-style-type: none"> - Final review - Approve response letter - Submit

4 BNG 4: Templates - Mu Tr Li

4.1 Template 1: Khi đng ý và làm theo

Template Response:

We thank the reviewer for this excellent suggestion. We have [specific action taken]:

- [Specific change 1 with location in manuscript]
- [Specific change 2]

Changes in manuscript:

- Section X: [description of change]
- Figure/Table Y: [description]

Results: [Outcome of the change, e.g., "Calibrated COCOMO II improved MMRE from 2.79 to 1.85, but RF still outperforms (0.65, p<0.001)."]

4.2 Template 2: Khi gii thíc limitation

Template Response:

We appreciate this valuable comment. While we acknowledge [concern], we respectfully note that [limitation/constraint]:

- [Reason 1]
- [Reason 2]

However, we have [partial mitigation]:

- [Action taken]

We explicitly acknowledge this limitation in Section X.Y and recommend [future work] in our revised Conclusion (Section 7.3).

4.3 Template 3: Khi không dùng CN THN

Template Response:

We respectfully disagree with this assessment because [reason]. Our approach [justification] is appropriate for this study because:

1. *[Evidence 1 with citation]*
2. *[Evidence 2]*

Prior work [citations] supports this methodology. However, to prevent misunderstanding, we have clarified our reasoning in Section X (lines Y-Z).

5 Ph Lc: Key Reminders

5.1 Nhng đm TUYT ĐI phi làm

1. **Overall aggregation:** Define macro-average explicitly (1 câu trong Abstract + 1 đon trong Results)
2. **COCOMO calibration:** Fit A, B on training data → compare original vs calibrated vs RF
3. **Remove leakage:** KHÔNG dùng Developers = Effort/Time làm feature
4. **FP protocol:** Dùng LOOCV cho FP, báo cáo bootstrap CIs, gi kt qu "exploratory"
5. **Dataset manifest:** Table vi 6 ct: Source — Link — Schema — Raw # — Removed — Final
6. **Figure captions:** ALL figures phi có caption đy đ + high quality (vector)
7. **Cite 4 DOIs:** R3 suggest 4 papers → phi cite + discuss trong Related Work
8. **R² values:** Compute hoc remove column + explain

5.2 Likelihood of Acceptance

- Nu làm **CRITICAL (1-6)** đy đ: 60-70% chance
- Nu làm **CRITICAL + MAJOR (1-10):** 75-85% chance
- Nu làm **tt c (1-14):** 85-90% chance

Recommendation: u tiên 1-10, làm 11-14 nu còn thi gian. Có th xin thêm 5 ngày đ làm tt hn.

5.3 Contact Strategy

Email to Editor (nu cn thêm time):

Dear Editor,

We have received the thoughtful and detailed reviews for manuscript 6863b9b0-4db8-4b53-843f-5be5e907cf62. We are committed to addressing all reviewer concerns comprehensively.

Given the extensive revisions requested (8 reviewers with substantial methodological improvements including baseline recalibration, additional experiments, and dataset documentation), we respectfully request a 5-day extension to ensure the highest quality revision. This will allow us to implement all requested changes rigorously rather than superficially.

We expect to submit the revised manuscript within 15 working days from the original deadline.

Thank you for your consideration.