

PE-Tab2Conv: Clinically structured convolutional representation learning for preeclampsia risk modelling from routine electronic health records

Sindhu Rakasiwi

Universitas Dian Nuswantoro

Abu Salam

`abu.salam@dsn.dinus.ac.id`

Universitas Dian Nuswantoro

Ika Novita Dewi

Universitas Dian Nuswantoro

Danang Wahyu Utomo

Universitas Dian Nuswantoro

Dhita Aulia Octaviani

Politeknik Kesehatan Kemenkes Semarang

Sasha-Gay Wright

University of the West Indies

Research Article

Keywords: Preeclampsia, electronic health records, tabular-to-matrix representation, convolutional neural network, representation learning, clinical decision support

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10925657/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

PE-Tab2Conv: Clinically structured convolutional representation learning for preeclampsia risk modelling from routine electronic health records

Sindhu Rakasiwi^{1,2,3,4*}, Abu Salam^{1,2,3,4*}, Ika Novita Dewi^{1,2,3,4},
Danang Wahyu Utomo^{1,2,3,4}, Dhita Aulia Octaviani⁵,
Sasha-Gay Wright⁶

^{1*}Faculty of Computer Science, Universitas Dian Nuswantoro, Imam Bonjol Street No. 207, Semarang, Central Java 50131, Indonesia.

²Intelligent Distributed Surveillance and Security (IDSS), Universitas Dian Nuswantoro, Semarang, Central Java 50131, Indonesia.

³Dinus Research Group for AI in Medical Science (DREAMS), Universitas Dian Nuswantoro, Semarang, Central Java 50131, Indonesia.

⁴OncoDoc AI Research Laboratory, Universitas Dian Nuswantoro, Semarang, Central Java 50131, Indonesia.

⁵Department of Midwifery, Poltekkes Kemenkes Semarang, Semarang, Central Java 50268, Indonesia.

⁶Biomedical Engineering, Faculty of Engineering, The University of the West Indies, Mona Campus, 4 St. Johns Close, Kingston 7, Jamaica.

*Corresponding author(s). E-mail(s): sindhu.rakasiwi@dsn.dinus.ac.id;
abu.salam@dsn.dinus.ac.id;

Contributing authors: ikadewi@dsn.dinus.ac.id;
danang.wu@dsn.dinus.ac.id; dhitaaulia@poltekkes-smg.ac.id;
sasha-gay.wright@uwi.edu;

Abstract

Background: Machine-learning decision support from routine electronic health records requires representations that handle class imbalance while retaining clinically interpretable structure. We developed PE-Tab2Conv, a clinically structured tabular-to-matrix representation for preeclampsia risk modelling. **Methods:** Encoded maternal variables were arranged in a fixed 4×6 matrix spanning baseline and socioeconomic characteristics, obstetric history, extended obstetric and

family history, and comorbidities. A lightweight two-dimensional convolutional neural network produced 32-dimensional embeddings for downstream classifiers. Models were evaluated on a locked 2023 temporal cohort from RS Adhyatma ($n = 531$), comparing a portable core feature space (17 predictors) with an extended space (21 predictors). Interpretability used SHapley Additive exPlanations and matrix-cell occlusion. A second hospital contributed 72 preeclampsia-positive records for positive-case transportability assessment. **Results:** In the extended setting, PE-Tab2Conv with logistic regression achieved an area under the precision–recall curve of 0.817 (area under the receiver operating characteristic curve 0.922; Brier score 0.047), numerically above extended tabular XGBoost (0.800). The extended-to-core gain was primarily associated with historical predictors; prior preeclampsia had the largest cell-level occlusion effect ($\Delta\text{AUPRC} = 0.210$). External positive cases retained elevated model scores, but the absence of external controls precluded full external validation. **Conclusions:** PE-Tab2Conv offers a compact, traceable representation-learning method for health-informatics risk modelling and potential clinical decision-support development. Balanced multicentre validation is required before deployment.

Keywords: Preeclampsia; electronic health records; tabular-to-matrix representation; convolutional neural network; representation learning; clinical decision support

1 Introduction

Preeclampsia remains one of the most perilous obstetric complications driving maternal and perinatal morbidity and mortality globally, particularly in resource-constrained developing regions (Riano-Moreno et al. 2025; Henderson et al. 2017). The condition is characterized by new-onset hypertension and multisystem organ dysfunction developing after 20 weeks of gestation (Poon et al. 2021). Early identification of preeclampsia risk during early pregnancy is critical to enable timely clinical interventions, such as intensified antenatal surveillance and early low-dose aspirin prophylaxis administration (Myatt 2022). Nevertheless, most primary and secondary healthcare facilities in developing countries still rely on manual risk scoring derived from standard clinical guidelines, which frequently exhibit limited sensitivity and high predictive uncertainty (Al-Rubaie et al. 2020, 2016; Antwi et al. 2020; De Kat et al. 2019).

Leveraging maternal Electronic Health Record (EHR) data to construct machine learning (ML)-based risk prediction systems offers substantial potential to enhance clinical screening accuracy (Jhee et al. 2019; Escobar et al. 2021; Li et al. 2021; Chen et al. 2024; Kempener et al. 2025). However, processing tabular EHR data presents severe methodological hurdles, notably extreme class imbalance where preeclampsia instances represent a small minority, along with complex non-linear interactions across maternal predictors (Pedersen et al. 2024; Salmi et al. 2024). Conventional decision-tree tabular ML algorithms, such as Random Forest and XGBoost, frequently suffer from false-alarm fatigue or degraded positive-class discrimination performance when evaluated on imbalanced temporal validation cohorts (Ranjbar et al. 2024; Eberhard et al. 2025; Jain et al. 2025; Li et al. 2024).

Converting tabular records into two-dimensional image-like formats (Tabular-to-Image/T2I) has emerged as an innovative strategy to harness the representation extraction capabilities of 2D Convolutional Neural Networks (CNNs) (Damri et al. 2024; Briner et al. 2023; Serra Neto et al. 2022; Gómez-Martínez et al. 2026). Conventional T2I methodologies, such as Image Generator for Tabular Data (Zhu et al. 2021) and the Novel algorithm for Convolving Tabular Data (Alenizy and Berri 2025), predominantly rely on purely data-driven spatial ordering algorithms without incorporating domain-specific clinical structures. These arbitrary pixel mapping approaches yield matrix configurations that lack intuitive medical semantic representation, thereby limiting transparency, physician interpretability, and real-world clinical system adoption (Halladay et al. 2025; Selke et al. 2025).

To address these limitations, this study proposes PE-Tab2Conv (4×6), a hybrid matrix-based representation learning framework that structures tabular EHR variables into a clinically guided semantic layout. The 4×6 matrix is systematically arranged across four primary domain rows: baseline maternal & SES (Row 1), routine obstetric history (Row 2), extended obstetric & family history (Row 3), and maternal comorbidity profiles (Row 4) (Al-Rubaie et al. 2020; Liu et al. 2022). This domain-guided arrangement enables 2D convolutional filters to capture spatial correlations across clinically adjacent features while preserving full cell-level traceability and medical interpretability for healthcare providers (Salinas et al. 2025; Ponce et al. 2024).

The PE-Tab2Conv framework couples a 32-dimensional dense embedding extractor from a lightweight CNN with downstream classifiers (such as logistic regression and support vector machines) to evaluate alternative screening operating profiles (Tripathi et al. 2024; Wang et al. 2026). Evaluation was conducted on a locked temporal validation cohort at RS Adhyatma 2023 ($n = 531$), contrasting two distinct feature spaces: a portable core space (17 routine predictors) and an extended full-information space (21 predictors including historical variables) (Tiruneh et al. 2025; Saito and Rehmsmeier 2015). This experimental design was structured to estimate the incremental value of historical features for preeclampsia risk discrimination (Ballard et al. 2024; Al-Rubaie et al. 2016; Kempener et al. 2025).

The methodological contribution combines a clinically structured matrix, compact convolutional representation learning, direct comparison with tabular classifiers, and multi-level attribution using SHapley Additive exPlanations (SHAP) and matrix occlusion sensitivity (Singh and Rajput 2025; Lara-Abelenda et al. 2025). Matrix occlusion was used to determine whether influential regions corresponded to recognizable clinical variables. The locked portable model was also applied to an external PE-only cohort at RSUD K.R.M.T. Wongsonegoro ($n = 72$) as a positive-case transportability analysis, not as full external validation (Tiruneh et al. 2024; Brunelli et al. 2020). Together, these components address representation modelling, knowledge discovery, and transparent decision-support development from routine health-informatics data (Layton 2025; Zapata et al. 2025).

2 Materials and Methods

2.1 Study Design and Cohort Derivation

This retrospective medical informatics study was designed to develop and evaluate a convolutional representation learning framework for routine tabular clinical data. The overall analytic dataset comprised 1,743 maternal records from two public hospitals in Central Java, Indonesia. The model development cohort used RSUD dr. Adhyatma records from 2021–2022 ($n = 1,140$; 67 PE cases), while 2023 records were held out as a locked temporal validation cohort ($n = 531$; 46 PE cases). An external cohort from RSUD K.R.M.T. Wongsonegoro 2024 containing PE-only records ($n = 72$) was used exclusively for cross-site positive-case transportability assessment.

A locked temporal validation scheme was prioritized over random partitioning to better approximate real-world clinical deployment, where models trained on historical records are applied to future patient cohorts (Eberhard et al. 2025; Tiruneh et al. 2025; Antwi et al. 2020; De Kat et al. 2019). This separation minimizes overoptimistic performance estimates caused by documentation shifts, referral patterns, and disease prevalence changes over time. The summary of cohort composition and temporal partitioning is presented in Table 1, while the overall experimental workflow and scenario are detailed in Figure 1. Ethical approval was granted by the relevant Health Research Ethics Committee (approval No. 373/EA/F.XXIII.38/2026).

2.2 Clinical Feature Spaces

To distinguish full-information temporal performance from cross-site model portability, two feature spaces were defined a priori: a portable core feature space and a full-information extended feature space. The core feature space comprised 17 encoded predictors uniformly available across both hospital sites, including baseline maternal characteristics (age, parity, birth interval), comorbidity profile (hypertension history, obesity, diabetes mellitus), and socioeconomic proxy. Relying on routine clinical variables ensures high deployment feasibility without depending on costly laboratory assays or specialized biomarkers (Al-Rubaie et al. 2020; Escobar et al. 2021; Al-Rubaie et al. 2016; Noël et al. 2021; Mendoza et al. 2022; Ansbacher-Feldman et al. 2022).

The extended feature space contained 21 encoded predictors, comprising all core variables plus two critical historical predictors: prior preeclampsia history and family history of hypertension/preeclampsia. These two additional predictors were available at RS Adhyatma but structurally unavailable at RS Wongsonegoro (see Table 2). Consequently, extended features were reserved for within-site full-information temporal evaluation at RS Adhyatma. All preprocessing and one-hot encoding transformations were anchored strictly to the development cohort to prevent data leakage into validation cohorts.

2.3 PE-Tab2Conv Matrix Construction

The PE-Tab2Conv framework maps encoded categorical tabular features into a two-dimensional $4 \times 6 \times 1$ matrix representation. Unlike generic tabular-to-image

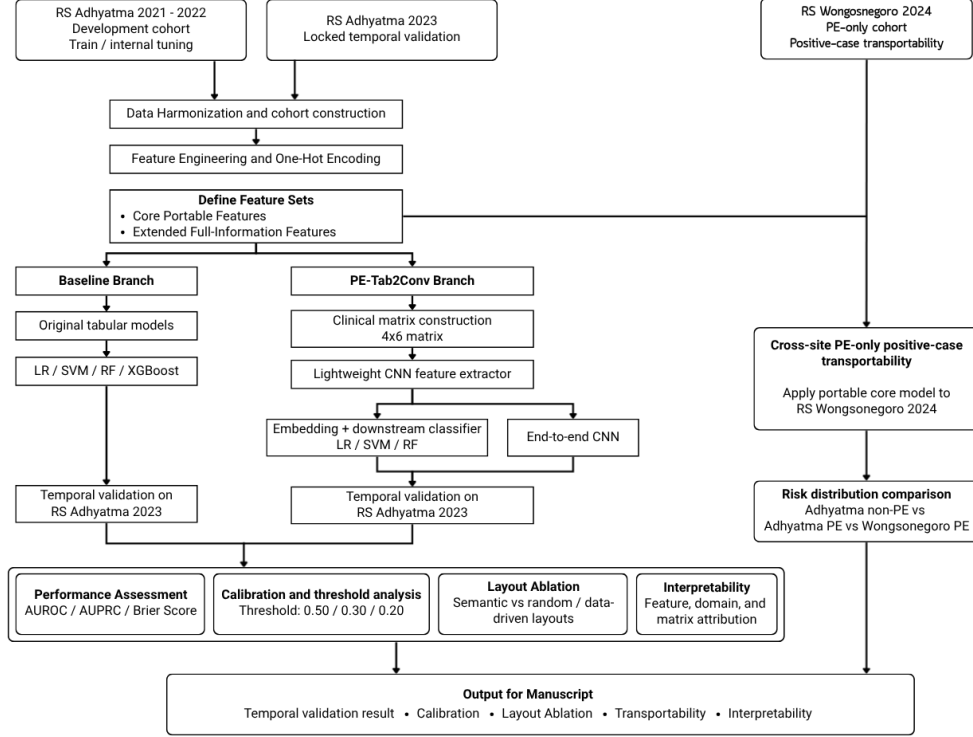


Fig. 1 Overall PE-Tab2Conv workflow. Routine maternal hospital data were harmonized into core and extended feature spaces, transformed into clinically guided matrix representations, processed using a lightweight convolutional feature extractor, and evaluated through original-tabular baselines, PE-Tab2Conv embeddings, layout ablation, temporal validation, cross-site positive-case transportability, calibration analysis, and interpretability assessment

Table 1 Cohort composition and temporal partitioning

Site	Year	Cohort role	Total PE Non-PE PE (%)			
RS Adhyatma	2021	Development cohort	579	26	553	4.49
RS Adhyatma	2022	Development cohort	561	41	520	7.31
RS Adhyatma	2023	Locked temporal validation cohort	531	46	485	8.66
RS Wongsonegoro	2024	PE-only cross-site transportability cohort	72	72	0	100.00

PE, preeclampsia. RS Adhyatma 2021–2022 was used for development, RS Adhyatma 2023 for locked temporal validation, and RS Wongsonegoro 2024 for PE-only cross-site positive-case transportability.

transformations that reorder feature positions based strictly on spatial statistical optimization (Damri et al. 2024; Zhu et al. 2021; Alenizy and Berri 2025; Serra Neto et al. 2022; Gómez-Martínez et al. 2026), the PE-Tab2Conv matrix layout is organized around clinical semantic domains. Each row represents a specific medical domain: (1)

baseline maternal characteristics & SES, (2) Obstetric history, (3) Extended Obstetric & family history, and (4) Comorbidity profile (Figure 2 and Table 2).

The matrix contains 24 spatial cells. Unassigned cells and features that were structurally unavailable in the core setting (such as prior PE and family history) were filled with constant zero-padding. This padding strategy ensures that both core and extended matrices share the same spatial dimensions ($4 \times 6 \times 1$). Domain-based

grouping was designed to impose a meaningful local inductive bias for 2D convolutional filters while enabling matrix-level feature relationship interpretation.

2.4 CNN Feature Extractor and Downstream Classifiers

A lightweight convolutional neural network (CNN) feature extractor was designed to output a 32-dimensional dense embedding from the $4 \times 6 \times 1$ input matrix. The architecture was intentionally kept compact to prevent overfitting risks inherent to moderate-sized medical tabular datasets (Lara-Abelenda et al. 2025). It comprises two sequential convolutional blocks (Block 1: 8 filters of 2×2 , BatchNorm, ReLU, Dropout 0.10; Block 2: 16 filters of 2×2 , BatchNorm, ReLU) followed by Global Average Pooling 2D and a 32-unit Dense Layer with ReLU activation.

Two operational modes were evaluated: end-to-end CNN (connecting the embedding directly to a sigmoid output layer) and a hybrid representation-learning mode. In Hybrid mode, the 32-D embedding was extracted and fed into conventional downstream classifiers: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). Hybrid mode served as the primary configuration because it decoupled feature representation from the final classification rule, enabling direct benchmarking against original tabular models using identical classifier families. Hyperparameter tuning was performed exclusively on the development cohort via stratified cross-validation using average precision as the objective (Liu et al. 2022; Li et al. 2021, 2024; Ma et al. 2025). The complete experimental scenarios and hyperparameter registry are provided in Online Resource 1, Supplementary Tables S1 and S2.

2.5 Layout Ablation Experiments

A layout ablation experiment was specifically designed to evaluate whether the semantic clinical feature arrangement contributed a meaningful performance advantage over arbitrary random or data-driven spatial feature orderings (Tripathi et al. 2024; Wang et al. 2026). The core premise of tabular-to-matrix transformations is that the spatial geometry where features are placed influences pattern extraction by 2D convolutional filters. Consequently, the semantic clinical layout was experimentally benchmarked against four alternative placement strategies: random layouts (repeated 100 times to generate a random sample distribution), correlation-based, mutual-information, and information-gain layouts (Selke et al. 2025; Posada Barrera et al. 2025). All alternative feature layouts were strictly constructed using the RS Adhyatma 2021–2022 development cohort only to prevent data leakage.

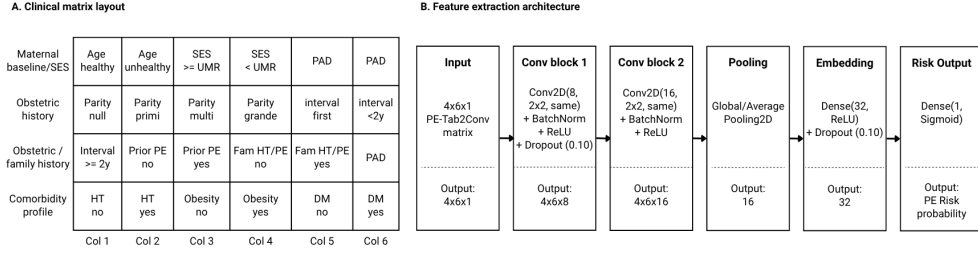


Fig. 2 PE-Tab2Conv matrix architecture. Harmonized encoded maternal predictors were arranged into a fixed $4 \times 6 \times 1$ clinically guided matrix. Rows and regions represent maternal baseline and socioeconomic characteristics, obstetric history, extended obstetric and family-history features, and comorbidity profile. Zero-padding preserved fixed dimensions across core and extended feature spaces

Table 2 PE-Tab2Conv matrix mapping and feature availability

Row	Column	Matrix label	Clinical domain	Encoded feature	Availability
R1	C1	Age healthy	Maternal baseline	maternal_age_group_norm_reproductive_healthy	Core and extended
R1	C2	Age unhealthy	Maternal baseline	maternal_age_group_norm_reproductive_unhealthy	Core and extended
R1	C3	SES $\geq$ UMR	Socioeconomic proxy	low_socioeconomic_status_norm_above_umr	Core and extended
R1	C4	SES $<$ UMR	Socioeconomic proxy	low_socioeconomic_status_norm_below_umr	Core and extended
R1	C5	PAD	Zero-padding	None	Padding cell
R1	C6	PAD	Zero-padding	None	Padding cell
R2	C1	Parity null	Obstetric history	parity_norm_nullipara	Core and extended
R2	C2	Parity primi	Obstetric history	parity_norm_primipara	Core and extended
R2	C3	Parity multi	Obstetric history	parity_norm_multipara	Core and extended
R2	C4	Parity grande	Obstetric history	parity_norm_grandemultipara	Core and extended
R2	C5	Interval first	Obstetric history	birth_interval_norm_first_child	Core and extended
R2	C6	Interval $< 2 y$	Obstetric history	birth_interval_norm_lt_2_years	Core and extended
R3	C1	Interval $> 2 y$	Obstetric history	birth_interval_norm_gt_2_years	Core and extended
R3	C2	Prior PE no	Extended obstetric history	prior_pe_history_norm_no	Extended only
R3	C3	Prior PE yes	Extended obstetric history	prior_pe_history_norm_yes	Extended only
R3	C4	Family HT/PE no	Family history	family_history_ht_pe_norm_no	Extended only
R3	C5	Family HT/PE yes	Family history	family_history_ht_pe_norm_yes	Extended only
R3	C6	PAD	Zero-padding	None	Padding cell
R4	C1	HT no	Comorbidity profile	history_hypertension_norm_no	Core and extended
R4	C2	HT yes	Comorbidity profile	history_hypertension_norm_yes	Core and extended
R4	C3	Obesity no	Comorbidity profile	obesity_norm_no	Core and extended
R4	C4	Obesity yes	Comorbidity profile	obesity_norm_yes	Core and extended
R4	C5	DM no	Comorbidity profile	history_diabetes_mellitus_norm_no	Core and extended
R4	C6	DM yes	Comorbidity profile	history_diabetes_mellitus_norm_yes	Core and extended

DM, diabetes mellitus; HT, hypertension; PE, preeclampsia; SES, socioeconomic status; UMR, local minimum wage proxy; PAD, zero-padding cell. Core and extended matrices used the same $4 \times 6 \times 1$ spatial structure. Extended-only cells were zero-filled in portable core matrices because prior PE history and family history of hypertension or PE were structurally unavailable in RS Wongsonegoro.

The primary layout-ablation configuration used PE-Tab2Conv embeddings paired with Logistic Regression as a standardized downstream classifier to reduce computational overhead while ensuring a consistent baseline. The same locked temporal split was applied across all layouts. For each candidate layout, the CNN feature extractor was trained on the development cohort, embeddings were re-extracted, and the downstream Logistic Regression classifier was refitted before evaluation on the locked RS Adhyatma 2023 temporal validation cohort. Accordingly, the layout-ablation values represent experiment-specific retraining results and are not expected to be numerically

identical to the principal model-comparison results reported in Table 3. To statistically test whether the semantic layout AUPRC outperformed the random layout distribution, the empirical probability (p_{emp}) was calculated as:

$$p_{\text{emp}} = \frac{1 + \# \{ \text{AUPRC}_{\text{random}} \geq \text{AUPRC}_{\text{semantic}} \}}{1 + N_{\text{random}}}$$

where $N_{\text{random}} = 100$ represents the total number of random layout iterations. Complete locked layout parameters and random distributions are documented in Online Resource 1, Supplementary Table S5 and Supplementary Note S2.

2.6 Cross-site Positive-Case Transportability

Cross-site evaluation was strictly restricted to the portable core feature space because the two extended predictors, prior preeclampsia history and family history, were structurally unavailable in RS Wongsonegoro records. The locked core model trained on the RS Adhyatma 2021–2022 development cohort was applied unchanged to three distinct cohort groups: RS Adhyatma non-PE ($n = 1,558$), RS Adhyatma PE ($n = 113$), and RS Wongsonegoro PE ($n = 72$). The evaluation analyzed predicted risk probability distributions, mean, median, interquartile range (IQR), and the proportion classified as high risk at predefined cutoffs of 0.50, 0.30, and 0.20.

This analysis was explicitly framed as positive-case transportability to assess whether the portable core model preserved an elevated risk signal in external PE cases, rather than a full external validation (Brunelli et al. 2020; Allen et al. 2017; Rezende et al. 2021). This methodological distinction was necessary because the RS Wongsonegoro cohort contained PE-only records without non-PE controls, preventing the estimation of external AUROC, specificity, calibration, or false-positive rates. Online Resource 1, Supplementary Table S7 and Supplementary Note S1 document complete transportability output registries and analytical constraints.

2.7 Performance Metrics, Calibration, and Interpretability

The primary evaluation metric on the locked RS Adhyatma 2023 temporal validation cohort was the Area Under the Precision-Recall Curve (AUPRC). AUPRC was prioritized over the Area Under the Receiver Operating Characteristic (AUROC) due to the low prevalence of preeclampsia (class imbalance), making precision-recall analysis significantly more informative for evaluating positive-class discrimination without being distorted by a large negative class (Salmi et al. 2024; Saito and Rehmsmeier 2015; Ma et al. 2025; Paul et al. 2025; Fardana et al. 2025). Secondary metrics included AUROC for discrimination and the Brier score as an overall measure of probabilistic accuracy.

Threshold-based operating characteristics were evaluated at cutoffs defined a priori at 0.50, 0.30, and 0.20. Reported metrics at each threshold included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1-score, confusion matrix counts (true positives, true negatives, false positives, false negatives), and the proportion of records classified at high risk. This multi-threshold analysis aimed to characterize model behavior under realistic clinical operating conditions, supporting both sensitivity-oriented screening and precision-oriented high-risk alert settings. Calibration quality was evaluated descriptively using calibration curves and

Brier scores without applying recalibration to the validation cohort, preserving the locked evaluation design. The complete threshold output registry is documented in Supplementary Table S6 of Online Resource 1.

Interpretability analysis was conducted at the feature, domain, and matrix levels to assess whether model behavior aligned with clinical knowledge. Feature-level attributions were computed using model coefficients and Shapley Additive exPlanations (SHAP), which were aggregated into five medical domains: baseline maternal characteristics, obstetric history, family history, comorbidity profile, and socioeconomic indicators. For PE-Tab2Conv representations, matrix-level interpretability was evaluated using occlusion sensitivity analysis, where each matrix cell was systematically occluded to measure the resulting change in AUPRC. This analysis determined whether the convolutional network extracted signals from clinically coherent matrix regions (Salinas et al. 2025; Ponce et al. 2024; Singh and Rajput 2025). Summary performance results are reported in Tables 3–7, while calibration and attribution plots are presented in Figures 4 and 7.

2.8 Use of Generative AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist with language editing, readability improvement, and manuscript restructuring. Following the use of this tool, the authors reviewed and edited the resulting text as needed and take full responsibility for the final content of the manuscript.

3 Results

3.1 Cohort Composition and Feature Portability

Following data harmonization and record cleaning, the final analytic dataset comprised 1,743 maternal records from two public hospitals in Central Java, Indonesia. The primary RSUD dr. Adhyatma cohort ($n = 1,671$; 2021–2023) was partitioned into a model development cohort (2021–2022; $n = 1,140$) and a locked temporal validation cohort (2023; $n = 531$). Preeclampsia prevalence at RS Adhyatma exhibited a gradual temporal increase, rising from 4.49% (26/579) in 2021, to 7.31% (41/561) in 2022, and reaching 8.66% (46/531) in the 2023 validation cohort. The external cohort from RSUD K.R.M.T. Wongsonegoro 2024 contributed 72 records, all representing positive preeclampsia cases (100% PE) without non-PE controls.

The definition of core and extended feature spaces established a methodological basis to separate full-information temporal validation within RS Adhyatma from cross-site portability evaluation. The core feature space (17 encoded predictors) represented routine clinical variables consistently documented across both hospitals (maternal age, parity, birth interval, hypertension history, obesity, diabetes mellitus, and socioeconomic proxy). Conversely, the extended feature space (21 encoded predictors) included all core variables plus prior preeclampsia history and family history, which were structurally unavailable in RS Wongsonegoro. Detailed cohort composition and temporal distribution are summarized in Table 1, and the feature availability matrix in Table 2.

3.2 Temporal Validation Performance

Evaluation on the locked RS Adhyatma 2023 temporal validation cohort ($n = 531$; 46 PE cases) demonstrated a marked performance gap between core and extended feature spaces (Table 3 and Figure 3). In the core setting, positive-class discrimination remained limited across all model architectures, with AUPRC values ranging strictly from 0.192 to 0.265. The highest core AUPRC was achieved by original tabular Logistic Regression (0.265), followed by hybrid PE-Tab2Conv + SVM (0.256) and original tabular SVM (0.253). These findings indicate that shared routine predictors alone were insufficient for highly discriminative temporal risk stratification (Tiruneh et al. 2025).

Conversely, integrating the extended feature space substantially improved temporal validation performance across both original tabular and PE-Tab2Conv models. The hybrid PE-Tab2Conv framework paired with Logistic Regression (LR) achieved the overall highest AUPRC of 0.817 (AUROC 0.922; Brier score 0.047). The

end-to-end PE-Tab2Conv CNN mode yielded near-identical performance with an AUPRC of 0.816 and AUROC of 0.924, indicating that the convolutional matrix representation retained substantial extended-feature signal before downstream classifier replacement. Hybrid PE-Tab2Conv with RF achieved an AUPRC of 0.814, whereas hybrid PE-Tab2Conv with SVM attained the highest AUROC (0.929), lowest Brier score (0.026), highest PPV (0.895), and highest F1-score (0.810). The strongest original tabular comparator was extended XGBoost, achieving an AUPRC of 0.800 and AUROC of 0.912.

Bootstrap resampling analysis (2,000 iterations) on the locked validation cohort characterized uncertainty around these estimates. Hybrid PE-Tab2Conv + LR yielded a mean bootstrap AUPRC of 0.820 (95% CI: 0.695–0.917), compared to 0.802 (95% CI: 0.682–0.898) for original tabular XGBoost. Hybrid PE-Tab2Conv + SVM recorded a mean bootstrap AUROC of 0.929 (95% CI: 0.867–0.977) and a Brier score of 0.026 (95% CI: 0.016–0.038). Overall, PE-Tab2Conv performance was highest in the extended setting where prior PE and family history predictors were accessible. Detailed per-seed run histories, full bootstrap confidence intervals, and training convergence curves are documented in Online Resource 1, Supplementary Tables S3 and S4 and Supplementary Figure S2.

3.3 Calibration and Threshold-based Operating Characteristics

Threshold-based evaluation at cutoffs defined a priori (0.50, 0.30, and 0.20) revealed complementary clinical operating profiles among extended feature models (Table 4). At the default 0.50 threshold, hybrid PE-Tab2Conv + LR identified 39 of 46 preeclampsia cases in the locked temporal validation cohort, yielding a sensitivity of 0.848, specificity of 0.951, PPV of 0.619, and NPV of 0.985, while classifying 63 patients (11.9%) as high risk. These operating characteristics remained identical at the 0.30 cutoff, whereas lowering the threshold to 0.20 modestly increased high-risk classification to 69 patients (13.0%) while reducing PPV to 0.565.

Conversely, hybrid PE-Tab2Conv + SVM exhibited a conservative, precision-oriented operating profile. At the 0.50 threshold, it identified 34 preeclampsia cases

Table 3 Temporal validation performance on the locked RS Adhyatma 2023 cohort

Representation	Feature set	Model	AUROC	AUPRC	Brier	Sens.	Spec.	PPV	NPV	F1
Original tabular	Core	LR	0.637	0.265	0.181	0.370	0.934	0.347	0.940	0.358
PE-Tab2Conv embedding	Core	SVM	0.649	0.256	0.074	0.000	1.000	N.A.	0.913	0.000
Original tabular	Core	SVM	0.676	0.253	0.073	0.000	1.000	N.A.	0.913	0.000
PE-Tab2Conv embedding	Core	LR	0.643	0.251	0.184	0.478	0.808	0.191	0.942	0.273
PE-Tab2Conv CNN	Core	CNN	0.640	0.238	0.275	1.000	0.000	0.087	N.A.	0.159
Original tabular	Core	XGBoost	0.678	0.227	0.180	0.348	0.946	0.381	0.939	0.364
PE-Tab2Conv embedding	Core	RF	0.663	0.210	0.163	0.283	0.959	0.394	0.934	0.329
Original tabular	Core	RF	0.612	0.207	0.181	0.326	0.953	0.395	0.937	0.357
PE-Tab2Conv embedding	Core	XGBoost	0.639	0.192	0.169	0.413	0.847	0.204	0.938	0.273
PE-Tab2Conv embedding	ExtendedLR		0.922	0.817	0.047	0.848	0.951	0.619	0.985	0.716
PE-Tab2Conv CNN	ExtendedCNN		0.924	0.816	0.043	0.848	0.951	0.619	0.985	0.716
PE-Tab2Conv embedding	ExtendedRF		0.918	0.814	0.036	0.761	0.977	0.761	0.977	0.761
PE-Tab2Conv embedding	ExtendedSVM		0.929	0.806	0.026	0.739	0.992	0.895	0.976	0.810
Original tabular	ExtendedXGBoost		0.912	0.800	0.047	0.804	0.959	0.649	0.981	0.718
PE-Tab2Conv embedding	ExtendedXGBoost		0.886	0.768	0.040	0.761	0.975	0.745	0.977	0.753
Original tabular	ExtendedLR		0.907	0.740	0.053	0.804	0.965	0.685	0.981	0.740
Original tabular	ExtendedSVM		0.912	0.732	0.038	0.587	0.996	0.931	0.962	0.720
Original tabular	ExtendedRF		0.908	0.723	0.045	0.717	0.975	0.733	0.973	0.725

AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision–recall curve; Sens., sensitivity; Spec., specificity; PPV, positive predictive value; NPV, negative predictive value. Values are reported at the predefined threshold of 0.50 for threshold-dependent metrics. N.A. indicates an undefined value caused by a zero denominator.

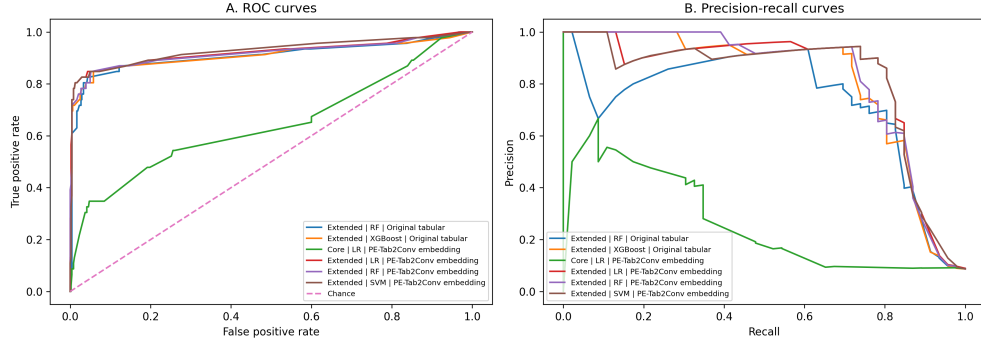


Fig. 3 Receiver operating characteristic and precision–recall curves on the locked RS Adhyatma 2023 temporal validation cohort for representative original tabular and PE-Tab2Conv models

with only 4 false positives, yielding a sensitivity of 0.739, specificity of 0.992, PPV of 0.895, NPV of 0.976, and F1-score of 0.810, while generating a low high-risk classification burden (38 patients, 7.2%). Lowering the SVM threshold to 0.30 increased sensitivity to 0.804 (37 cases identified) while maintaining high specificity (0.988), PPV of 0.860, and a peak F1-score of 0.831. Descriptive calibration curves (Figure 4) and Brier scores showed that PE-Tab2Conv + SVM had the lowest Brier score (0.026), followed by PE-Tab2Conv + RF (0.036), end-to-end CNN (0.043), and original tabular XGBoost (0.047). The complete threshold output registry is documented in Supplementary Table S6.

3.4 Layout Ablation Results

Layout ablation experiments were evaluated to determine whether the semantic clinical spatial feature arrangement provided a statistically meaningful performance advantage over arbitrary random or data-driven spatial orderings (Table 5 and Figure 5). In the extended feature setting, the semantic clinical layout achieved an AUPRC of 0.765 on the locked temporal validation cohort. This value corresponds specifically to the standardized layout-ablation experiment and is distinct from the principal PE-Tab2Conv embedding + LR result (AUPRC 0.817) reported in Table 3.

The semantic-layout AUPRC was higher than the mean AUPRC across 100 random layouts (0.744), yielding a semantic-minus-random mean difference of 0.021. However, the semantic layout performance did not exceed the 95th percentile upper tail of the random distribution (0.810), resulting in an empirical probability of $p_{\text{emp}} = 0.297$. Among non-random data-driven alternatives, the correlation-based layout achieved an AUPRC of 0.680 with a Brier score of 0.052, whereas the mutual-information and information-gain layouts each achieved an AUPRC of 0.754.

In the portable core feature setting, the semantic clinical layout achieved an AUPRC of 0.195 (AUROC 0.624; Brier score 0.183). This performance fell below the mean random layout AUPRC (0.221) and the 95th percentile random cutoff (0.257), yielding $p_{\text{emp}} = 0.931$. Overall, these ablation findings demonstrate that while the semantic clinical layout is highly competitive and provides a clinically interpretable

Table 4 Threshold-based operating characteristics for representative temporal validation models

Model	Feature set	Threshold	Sens.	Spec.	PPV	NPV	F1	TP/TN/ FP/FN	High-risk, n (%)
PE-Tab2Conv embedding + LR	Extended	0.50	0.848	0.951	0.619	0.985	0.716	39/461/24/7	63 (11.9)
PE-Tab2Conv embedding + LR	Extended	0.30	0.848	0.951	0.619	0.985	0.716	39/461/24/7	63 (11.9)
PE-Tab2Conv embedding + LR	Extended	0.20	0.848	0.938	0.565	0.985	0.678	39/455/30/7	69 (13.0)
PE-Tab2Conv embedding + SVM	Extended	0.50	0.739	0.992	0.895	0.976	0.810	34/481/4/12	38 (7.2)
PE-Tab2Conv embedding + SVM	Extended	0.30	0.804	0.988	0.860	0.982	0.831	37/479/6/9	43 (8.1)
PE-Tab2Conv embedding + SVM	Extended	0.20	0.826	0.963	0.679	0.983	0.745	38/467/18/8	56 (10.5)
PE-Tab2Conv embedding + RF	Extended	0.50	0.761	0.977	0.761	0.977	0.761	35/474/11/11	46 (8.7)
PE-Tab2Conv embedding + RF	Extended	0.30	0.783	0.967	0.692	0.979	0.735	36/469/16/10	52 (9.8)
PE-Tab2Conv embedding + RF	Extended	0.20	0.848	0.948	0.609	0.985	0.709	39/460/25/7	64 (12.1)
Original tabular XGBoost	+ Extended	0.50	0.804	0.959	0.649	0.981	0.718	37/465/20/9	57 (10.7)
Original tabular XGBoost	+ Extended	0.30	0.804	0.944	0.578	0.981	0.673	37/458/27/9	64 (12.1)
Original tabular XGBoost	+ Extended	0.20	0.870	0.860	0.370	0.986	0.519	40/417/68/6	108 (20.3)
PE-Tab2Conv embedding + LR	Core	0.50	0.478	0.808	0.191	0.942	0.273	22/392/93/24	115 (21.7)
PE-Tab2Conv embedding + LR	Core	0.30	0.978	0.074	0.091	0.973	0.167	45/36/449/1	494 (93.0)
PE-Tab2Conv embedding + LR	Core	0.20	1.000	0.004	0.087	1.000	0.160	46/2/483/0	529 (99.6)

Thresholds were prespecified at 0.50, 0.30, and 0.20. TP/TN/FP/FN denotes true positives, true negatives, false positives, and false negatives. The core PE-Tab2Conv logistic regression model is shown as a portability-oriented comparator; its low-threshold operating characteristics demonstrate the effect of limited predictor information on high-risk classification burden.

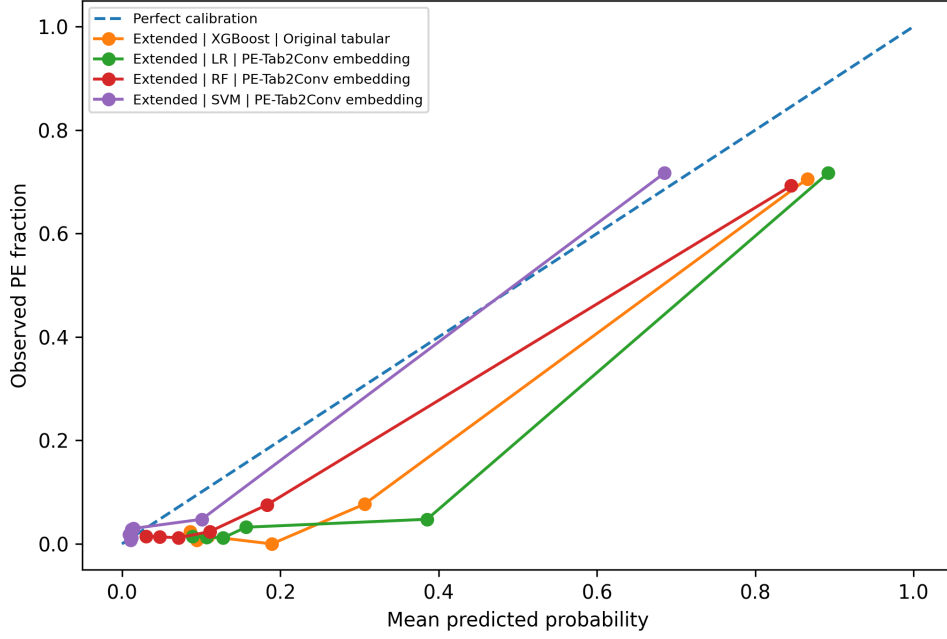


Fig. 4 Calibration plot on the locked RS Adhyatma 2023 temporal validation cohort for representative extended-feature models

matrix geometry in the extended setting, the semantic ordering itself is not statistically superior to random spatial placements. The primary performance gains of PE-Tab2Conv are driven by the synergy between richer clinical predictor availability and convolutional embedding extraction. Complete random parameter distributions are documented in Supplementary Table S5 and Supplementary Note S2 of Online Resource 1.

3.5 Cross-site Positive-Case Transportability

Cross-site evaluation was specifically conducted using the portable core feature space to assess whether the locked core model trained at RS Adhyatma preserved elevated risk signals in external positive preeclampsia

cases (RS Wongsonegoro 2024; $n = 72$ PE-only) (Table 6 and Figure 6). For the core hybrid PE-Tab2Conv + logistic regression (LR) model, mean predicted risk increased from 0.408 (median 0.376; IQR 0.3522–0.3949) in the internal RS Adhyatma non-PE cohort to 0.572 (median 0.503; IQR 0.3522–0.8516) in RS Adhyatma PE and 0.675 (median 0.796; IQR 0.4437–0.8666) in the external RS Wongsonegoro PE cohort. For the core hybrid PE-Tab2Conv + RF model, mean predicted risk was 0.378 in RS Adhyatma non-PE, 0.540 in RS Adhyatma PE, and 0.546 in RS Wongsonegoro PE.

High-risk classification burden across predefined thresholds further supported positive-case signal transportability. At the default 0.50 threshold, the core PE-Tab2Conv + LR model classified 72.2% of RS Wongsonegoro cases as high risk, compared with 51.3% in RS Adhyatma PE cases and only 19.1% in RS Adhyatma

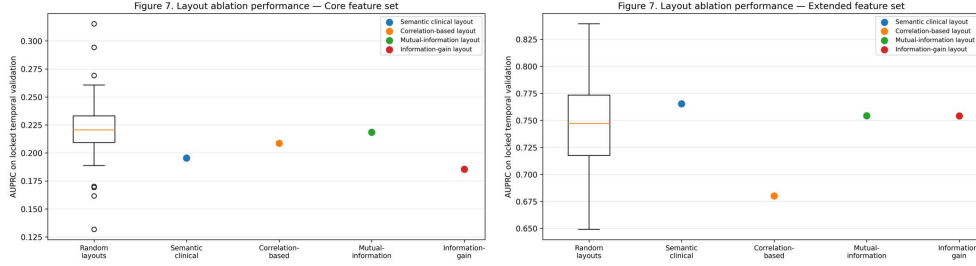


Fig. 5 Layout ablation performance comparing semantic clinical layout against random, correlation-based, mutual-information, and information-gain layouts

Table 5 Layout ablation analysis for PE-Tab2Conv embeddings with logistic regression

Feature set	Layout strategy	n layouts	AUR	AUP	Brier	Ran. mean AUP	Ran. p95 AUP	Sem.-ran. mean	p_{emp}
Core	Semantic clinical layout	1	0.624	0.195	0.183	0.221	0.257	-0.026	0.931
Core	Correlation-based layout	1	0.646	0.209	0.184	0.221	0.257	-0.026	0.931
Core	Mutual-information layout	1	0.612	0.219	0.201	0.221	0.257	-0.026	0.931
Core	Information-gain layout	1	0.614	0.185	0.192	0.221	0.257	-0.026	0.931
Core	Random layout	100	0.656	0.221	0.207	0.221	0.257	-0.026	0.931
Extended	Semantic clinical layout	1	0.910	0.765	0.131	0.744	0.810	0.021	0.297
Extended	Correlation-based layout	1	0.912	0.680	0.052	0.744	0.810	0.021	0.297
Extended	Mutual-information layout	1	0.920	0.754	0.048	0.744	0.810	0.021	0.297
Extended	Information-gain layout	1	0.919	0.754	0.065	0.744	0.810	0.021	0.297
Extended	Random layout	100	0.909	0.744	0.066	0.744	0.810	0.021	0.297

AUR denotes AUROC, area under the receiver operating characteristic curve; AUP denotes AUPRC, area under the precision-recall curve; Brier denotes the Brier score; Ran. denotes random layout; Sem. denotes semantic clinical layout; p95 denotes the 95th percentile; Sem.-ran. mean denotes the difference between the semantic-layout AUPRC and the mean AUPRC of random layouts. For non-random layouts, values correspond to the single evaluated layout. For random layout, values summarize 100 random arrangements. p_{emp} is the empirical probability that a random layout achieved an AUPRC greater than or equal to the semantic layout.

non-PE controls. Lowering the threshold to 0.30 and 0.20 increased high-risk classification in RS Wongsonegoro to 94.4% and 100.0%, respectively. Due to the absence of

non-PE controls at RS Wongsonegoro, this analysis is strictly interpreted as positive-case transportability (the ability to preserve risk elevation in external positive cases) rather than a full external validation.

3.6 Interpretability and Domain-level Attribution

Multi-level attribution analysis (feature, domain, and matrix levels) was performed to evaluate whether representation behavior learned by PE-Tab2Conv models aligned with clinical logic (Table 7 and Figure 7). In the portable core feature setting, SHAP attribution and matrix occlusion analyses demonstrated that the comorbidity profile (hypertension history, obesity, and diabetes mellitus) dominated predictive contributions, yielding a mean $|SHAP| = 0.485$ (37.2% attribution share) and a performance drop of $\Delta AUPRC = 0.188$ when row 4 matrix cells were occluded. Conversely, in the extended feature setting, a marked attribution shift occurred toward the extended obstetric and family history domain (row 3 of the PE-Tab2Conv matrix). The prior

Table 6 Cross-site PE-only positive-case transportability for portable core PE-Tab2Conv embedding models

Model	Cohort group	<i>n</i>	Mean risk	Median risk	IQR	High-risk, <i>n</i> (%)		
						0.50	0.30	0.20
LR	RS Adhyatma Non-PE	1558	0.408	0.376	0.3522–0.3949	297 (19.1)	1458 (93.6)	1556 (99.9)
LR	RS Adhyatma PE	113	0.572	0.503	0.3522–0.8516	58 (51.3)	110 (97.3)	113 (100.0)
LR	RS Wongsonegoro PE	72	0.675	0.796	0.4437–0.8666	52 (72.2)	68 (94.4)	72 (100.0)
RF	RS Adhyatma Non-PE	1558	0.378	0.338	0.3250–0.3748	80 (5.1)	1508 (96.8)	1552 (99.6)
RF	RS Adhyatma PE	113	0.540	0.486	0.3381–0.7595	43 (38.1)	111 (98.2)	112 (99.1)
RF	RS Wongsonegoro PE	72	0.546	0.560	0.3747–0.6255	42 (58.3)	72 (100.0)	72 (100.0)
XGBoost	RS Adhyatma Non-PE	1558	0.379	0.343	0.3348–0.3632	248 (15.9)	1473 (94.5)	1487 (95.4)
XGBoost	RS Adhyatma PE	113	0.565	0.519	0.3433–0.8772	58 (51.3)	109 (96.5)	110 (97.3)
XGBoost	RS Wongsonegoro PE	72	0.505	0.408	0.3433–0.7316	26 (36.1)	58 (80.6)	69 (95.8)
SVM	RS Adhyatma Non-PE	1558	0.050	0.038	0.0382–0.0393	0 (0.0)	1 (0.1)	82 (5.3)
SVM	RS Adhyatma PE	113	0.106	0.047	0.0382–0.2076	0 (0.0)	1 (0.9)	42 (37.2)
SVM	RS Wongsonegoro PE	72	0.138	0.149	0.0468–0.2110	0 (0.0)	2 (2.8)	27 (37.5)

PE, preeclampsia; IQR, interquartile range; LR, logistic regression; RF, random forest; SVM, support vector machine. High-risk columns report the number and percentage of cases classified as high risk at thresholds of 0.50, 0.30, and 0.20. Only core portable models were applied to RS Wongsonegoro because extended predictors were unavailable at that site. The analysis evaluates positive-case transportability only and does not constitute balanced external validation.

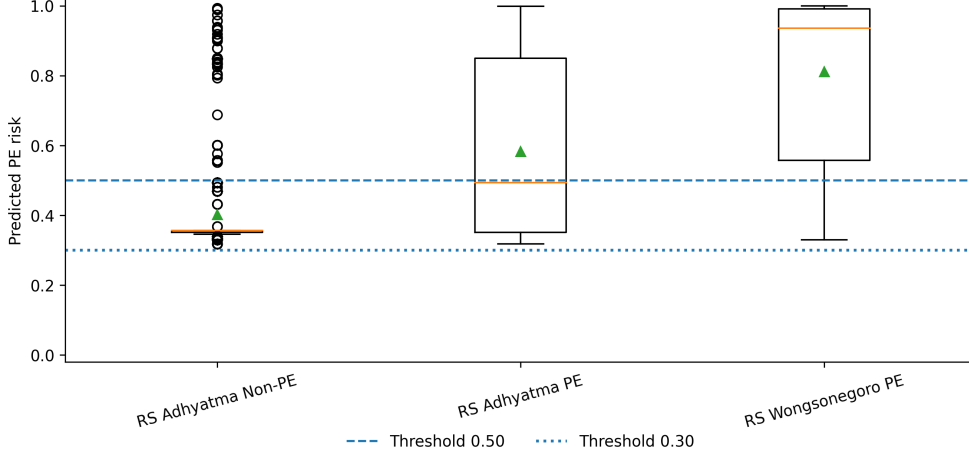


Fig. 6 Predicted-risk distributions for cross-site PE-only positive-case transportability analysis using portable core models

preeclampsia history cell (Prior PE: yes) recorded the single highest individual cell occlusion impact ($\Delta AUPRC = 0.210$).

Domain-level attribution aggregation indicated that the performance difference between the extended and core models (AUPRC 0.817 vs 0.251) coincided with greater reliance on historical predictors. Cumulative occlusion of the extended obstetric history domain (Online Resource 1, Supplementary Figure S4B) induced the largest total performance drop ($\Delta AUPRC = 0.249$), whereas occluding the baseline maternal domain (age and SES) exerted minimal influence ($\Delta AUPRC = 0.040$). The most influential observed regions were consistent with established clinical preeclampsia risk factors, although this analysis cannot exclude unmeasured or residual spurious associations. Visual checks of sample matrices and complete matrix occlusion heatmaps are provided in Online Resource 1, Supplementary Figures S1 and S4.

4 Discussion

4.1 Principal Findings and Methodological Contribution

This study evaluates PE-Tab2Conv, a 4×6 hybrid matrix-based representation-learning framework for imbalanced maternal electronic health record (EHR) data. Evaluation on the locked RS Adhyatma 2023 temporal validation cohort ($n = 531$) confirmed that coupling a convolutional embedding extractor with a downstream LR classifier achieved peak positive-class discrimination performance, yielding an AUPRC of 0.817 (AUROC 0.922; Brier score 0.047). This performance was numerically higher than the strongest traditional tabular machine-learning comparator by AUPRC, extended XGBoost (AUPRC 0.800; AUROC 0.912), although the bootstrap intervals overlapped. The result suggests that compact 2D convolutional representations can retain useful cross-feature information, but it does not establish superiority

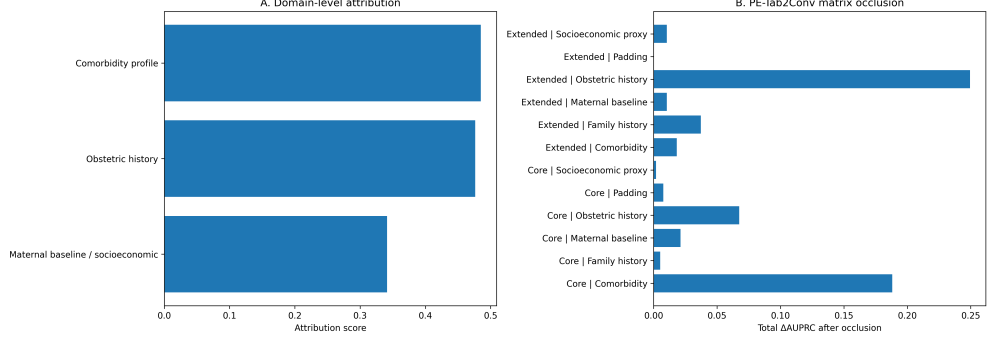


Fig. 7 Domain-level interpretability and matrix-level attribution for original tabular and PE-Tab2Conv models

Table 7 Interpretability and domain-level attribution summary

Source	Feature set	Attribution method	Top domain	Score	Detail
Original tabular model	Core	Mean absolute SHAP	Comorbidity profile	0.485	Note 1
PE-Tab2Conv matrix occlusion	Core	Cell occlusion Δ AUPRC	Comorbidity	0.188	Note 2
PE-Tab2Conv matrix occlusion	Extended	Cell occlusion Δ AUPRC	Obstetric history	0.249	Note 3

DM, diabetes mellitus; HT, hypertension; PE, preeclampsia; SHAP, SHapley Additive exPlanations; AUPRC, area under the precision–recall curve. Domain attribution was used to assess clinical plausibility rather than causality. Matrix cells are denoted by row (R) and column (C) in the PE-Tab2Conv matrix.

Note 1: Top variables corresponded to hypertension history, obesity, and diabetes mellitus indicators.

Note 2: Top core matrix cells corresponded to hypertension, obesity, and diabetes mellitus positions in the comorbidity row.

Note 3: Top extended matrix cells corresponded to prior PE history, parity, and birth-interval positions in the obstetric-history region.

over all tabular algorithms (Zhu et al. 2021; Serra Neto et al. 2022; Briner et al. 2023; Damri et al. 2024).

Another important finding was the pronounced performance gap observed between the portable core (17 predictors) and extended full-information (21 predictors) feature spaces. In the core setting, positive-class discrimination remained limited across all architectures (AUPRC ranging from 0.192 to 0.265), whereas adding four encoded indicators representing two historical predictors—prior preeclampsia and family history of hypertension/preeclampsia—was associated with a more than threefold higher AUPRC (0.817 vs 0.251; absolute Δ AUPRC = 0.566) (Al-Rubaie et al. 2016, 2020; Kempener et al. 2025). Remarkably, the end-to-end PE-Tab2Conv CNN achieved an AUPRC of 0.816 and AUROC of 0.924 in the extended setting, nearly identical to the hybrid PE-Tab2Conv + LR configuration (AUPRC 0.817). This indicates that the 32-dimensional dense embedding extracted via Global Average Pooling retained substantial risk information before downstream classification and supports the representation efficiency of the lightweight CNN architecture in this dataset.

4.2 Comparative Analysis with Existing Literature

The performance of the PE-Tab2Conv framework provides key insights into both the tabular-to-image (T2I) representation and obstetric risk prediction systems. Conceptual T2I methodologies, such as Image Generator

for Tabular Data (IGTD) (Zhu et al. 2021) and SuperTML (Jain et al. 2022; Serra Neto et al. 2022; Briner et al. 2023; Gómez-Martínez et al. 2026), typically rely on purely data-driven feature ordering algorithms, such as feature correlation or mutual information, without incorporating domain-specific clinical structures. This study extends this paradigm by evaluating a clinically guided semantic spatial arrangement with an explicitly traceable matrix structure. In the standardized layout ablation, the extended semantic layout achieved an AUPRC of 0.765. The layout ablation analysis revealed an empirical probability of $p_{\text{emp}} = 0.297$ (0.021 above the random mean, but below the 95th percentile random threshold of 0.810). These findings indicate that the observed performance cannot be attributed to semantic spatial ordering alone; predictor availability, model fitting, and convolutional representation learning also contributed.

Within the broader context of EHR-based preeclampsia risk prediction systems, PE-Tab2Conv showed complementary discrimination and operating profiles. Conventional risk scoring and machine learning models often rely on complex clinical variables or specialized diagnostic procedures that are inconsistently documented across primary care facilities (Al-Rubaie et al. 2016; Noël et al. 2021; Mendoza et al. 2022; Ansbacher-Feldman et al. 2022). By leveraging 21 routine clinical and historical predictors extracted directly from maternal EHR data, extended PE-Tab2Conv achieved an AUROC of 0.929 (AUPRC 0.806; Brier score 0.026) under the hybrid SVM mode and an AUPRC of 0.817 under the LR mode. These results support continued evaluation of convolutional matrix representations for routine EHR risk modelling. They do not by themselves establish effectiveness or implementation readiness in primary or secondary care.

4.3 Clinical Implications and Operational Deployment

The observed operating profiles illustrate how PE-Tab2Conv might be evaluated within electronic health record-based clinical decision support systems (CDSSs) (Layton 2025; Zapata et al. 2025). They also expose trade-offs between sensitivity, precision, and alert burden. In the locked temporal cohort, the PE-Tab2Conv + LR (extended) configuration at the 0.50 threshold offers a high-sensitivity profile (0.848), successfully identifying 39 out of 46 preeclampsia cases in the locked temporal validation cohort while classifying 11.9% of the population (n=63) as high risk.

The PE-Tab2Conv + SVM (extended) configuration provided a more precision-oriented profile at the 0.50 or 0.30 cutoffs. At the 0.50 threshold, this configuration achieved a PPV of 0.895 and specificity of 0.992, generating only 4 false-positive alerts out of 38 high-risk designations (7.2% classification burden). Adjusting the cutoff to 0.30 enhanced sensitivity to 0.804 (37 cases identified) while preserving high PPV (0.860; 6 false positives) and reaching a peak F1-score of 0.831. Potential portability is further suggested by the positive-case transportability findings, where the portable

core logistic-regression model assigned higher scores to external RS Wongsonegoro cases than to internal non-PE records (median risks 0.796 vs 0.376). Because the external cohort lacked controls, these operating characteristics support prospective evaluation of a triage application but do not establish external false-alarm burden or clinical utility.

4.4 Interpretability and Clinical Plausibility

The clinical plausibility of the representations learned by the PE-Tab2Conv framework was examined through multi-level attribution analyses (SHAP and matrix occlusion sensitivity). Influential observed features aligned with established preeclampsia risk factors, but this analysis cannot rule out spurious or unmeasured associations. In the portable core feature space, model predictive focus was dominated by the maternal comorbidity profile (row 4 of the matrix; mean $|SHAP| = 0.485$; 37.2% attribution share), where occluding all row 4 cells (Online Resource 1, Supplementary Figure S4A) induced a performance drop of $\Delta AUPRC = 0.188$. This aligns with clinical literature establishing chronic hypertension, obesity, and diabetes mellitus as primary comorbidity predictors for preeclampsia risk during pregnancy (Al-Rubaie et al. 2016, 2020; Kempener et al. 2025; Riano-Moreno et al. 2025).

Integrating historical predictors within the extended feature space prompted a substantial attribution shift toward the extended obstetric and family history domain (row 3 of the matrix; 51.8% attribution share; cumulative row drop $\Delta AUPRC = 0.249$). At the individual cell level, occluding the prior preeclampsia history (Prior PE: yes) cell recorded the single highest matrix occlusion impact ($\Delta AUPRC = 0.210$), alongside an overall model AUPRC difference from 0.251 to 0.817. This shift in attribution from general comorbidities to prior preeclampsia indicates that the fitted model prioritized a strong epidemiological risk factor (prior PE increases recurrence risk by 7-8 fold) (Al-Rubaie et al. 2016; Poon et al. 2021; Wainstock and Sheiner 2022). Such mechanistic clarity is vital for establishing clinical trust and adoption of automated risk recommendations generated by CDSS.

4.5 Study Limitations and Analytical Constraints

Although PE-Tab2Conv showed promising temporal discrimination and Brier scores, several methodological constraints must be acknowledged to properly contextualize the findings. The primary limitation stems from the nature of the cross-site evaluation at RSUD K.R.M.T Wongsonegoro ($n = 72$), which comprised exclusively positive preeclampsia cases without a corresponding non-PE control cohort. This structural condition precluded the computation of full external specificity, AUROC, AUPRC, and calibration curves. Consequently, external testing in this study was strictly framed and interpreted as a positive-case transportability assessment, evaluating the locked model’s capacity to preserve elevated risk signals in external positive cases, rather than a fully balanced external validation (Allen et al. 2017; Brunelli et al. 2020; Rezende et al. 2021; Tiruneh et al. 2024).

Another notable constraint involves the structural unavailability of historical predictors in secondary site records. The absence of prior preeclampsia and family

history variables at RS Wongsonegoro restricted cross-hospital testing exclusively to the portable core feature space (17 predictors). Internal validation showed that these historical features were associated with the primary AUPRC difference (from 0.251 to 0.817). Additionally, the moderate analytical sample size ($n = 1,743$) and geographic restriction to two hospitals in Central Java imply that model generalizability to broader maternal populations warrants further verification. Inconsistent documentation practices across primary care facilities also represent an operational hurdle for uniform extended feature collection in routine practice.

4.6 Future Research Directions

Future opportunities to extend the PE-Tab2Conv framework focus on three primary domains: external dataset expansion, temporal data modelling, and record system integration. The immediate priority involves evaluating the model on multi-center balanced datasets comprising broader patient populations. This step is essential to confirm model generalizability and robustness in the extended feature space beyond the current study sites (Antwi et al. 2020; De Kat et al. 2019; Tiruneh et al. 2024). Additionally, adapting the PE-Tab2Conv architecture to process time-series or multi-visit sequential data could enhance predictive sensitivity by tracking dynamic feature trend changes over time (Chen et al. 2024; Tiruneh et al. 2025).

Transitioning from offline model evaluation to real-time integration within EHR system architecture represents another vital focus area. Developing lightweight interface modules connected directly to EHR databases via data interoperability standards would enable automated risk scoring whenever input features are updated (Layton 2025; Zapata et al. 2025). Future research could also explore fusing the 4×6 tabular matrix with multimodal data streams while maintaining the computational efficiency of the lightweight convolutional network. Field implementation trials will provide empirical evidence regarding the computational reliability and user adoption of the PE-Tab2Conv framework.

5 Conclusion

This study presents PE-Tab2Conv, a 4×6 clinically structured representation-learning framework for imbalanced tabular EHR data. On the locked temporal validation cohort, coupling the convolutional embedding extractor with a downstream LR classifier yielded an AUPRC of 0.817 (AUROC 0.922; Brier score 0.047), numerically higher than extended tabular XGBoost (AUPRC 0.800). Historical predictors in the extended feature space were associated with a greater than threefold AUPRC relative to the core PE-Tab2Conv + LR model (AUPRC 0.251).

Multi-level attribution analysis (SHAP and matrix occlusion sensitivity) identified clinically plausible regions, with prior preeclampsia showing the largest single-cell occlusion effect ($\Delta AUPRC = 0.210$). PE-Tab2Conv therefore offers a compact and traceable method for health-informatics risk modelling. Balanced multicentre and prospective evaluation remains necessary before clinical deployment.

Supplementary Information

Online Resource 1

Supplementary methods, Tables S1–S7, Figures S1–S4, and Notes S1–S2 (ESM_1_Supplementary_Material.pdf).

Acknowledgments

The authors gratefully acknowledge the participating hospital teams and clinical staff for facilitating access to maternal clinical records and supporting the clinical interpretation of the data. The authors also thank the UDINUS Digital Health Laboratory for its support in data management, computational analysis, and technical implementation.

Statements and Declarations

Funding

This work was supported by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia through the 2026 research funding program administered via BIMA (Grant No. HK.02.03/F.XV/0695/2026). The funder had no role in the study design, data collection, data analysis, interpretation of the results, manuscript preparation, or the decision to submit the article for publication.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Sindhu Rakasiwi: Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization, Writing – original draft, Project administration.

Abu Salam: Conceptualization, Methodology, Supervision, Validation, Project administration, Writing – review and editing.

Ika Novita Dewi: Methodology, Investigation, Data curation, Validation, Writing – review and editing.

Danang Wahyu Utomo: Software, Data curation, Formal analysis, Validation, Visualization, Writing – review and editing.

Dhita Aulia Octaviani: Conceptualization, Investigation, Resources, Data curation, Validation, Writing – original draft, Writing – review and editing.

Sasha-Gay Wright: Methodology, Formal analysis, Validation, Supervision, Writing – review and editing.

All authors reviewed and approved the final version of the manuscript.

Ethics approval

This study was conducted as a retrospective, non-interventional analysis using de-identified maternal clinical records. Ethical approval for the study, covering both the RS Adhyatma and RS Wongsonegoro cohorts, was obtained from the Health Research Ethics Committee of Poltekkes Kemenkes Semarang (approval No. 373/EA/F.XXIII.38/2026).

Consent to participate

The requirement for individual informed consent was waived by the Health Research Ethics Committee of Poltekkes Kemenkes Semarang because the study used retrospective, de-identified clinical records and involved no direct patient contact, intervention, or patient-level re-identification. All data were handled in de-identified form and used solely for research purposes. Patient privacy and institutional data-governance requirements were maintained throughout the study.

Consent for publication

Not applicable. The manuscript contains no identifiable individual-level information.

Data availability

The data used in this study comprise private retrospective maternal clinical records collected within routine hospital workflows. These data are not publicly available because of patient privacy, ethical restrictions, and institutional data-governance requirements. Access may be considered upon reasonable request to the corresponding authors or an authorized data custodian, subject to institutional approval, applicable regulations, ethics requirements where necessary, and an appropriate data use agreement.

Code availability

The source code used for data harmonization, matrix construction, model development, temporal validation, threshold-based evaluation, layout ablation, interpretability analysis, and cross-site positive-case transportability assessment is not publicly available because of institutional and confidentiality constraints. To support reproducibility, the manuscript and Supplementary Material report the cohort construction logic, feature definitions, preprocessing workflow, model families, hyperparameter settings, validation strategy, evaluation metrics, and experimental procedures in detail. Restricted code components or a runnable analysis package may be made available by the corresponding authors upon reasonable request, subject to institutional approval and, where necessary, an appropriate confidentiality agreement.

References

Al-Rubaie Z, Askie L, Ray J, et al (2016) The performance of risk prediction models for pre-eclampsia using routinely collected maternal characteristics and comparison

- with models that include specialised tests and with clinical guideline decision rules: a systematic review. *BJOG: An International Journal of Obstetrics and Gynaecology* 123(9):1441–1452. <https://doi.org/10.1111/1471-0528.14029>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84979017484&doi=10.1111%2F1471-0528.14029&partnerID=40&md5=8bf8285f7ef8468baf120fee2e2eece>
- Al-Rubaie Z, Hudson H, Jenkins G, et al (2020) Prediction of pre-eclampsia in nulliparous women using routinely collected maternal characteristics: A model development and validation study. *BMC Pregnancy and Childbirth* 20(1). <https://doi.org/10.1186/s12884-019-2712-x>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077586184&doi=10.1186%2Fs12884-019-2712-x&partnerID=40&md5=b94b21673aee6814e0dbf68d1972ca15>
- Alenizy H, Berri J (2025) Transforming tabular data into images via enhanced spatial relationships for CNN processing. *Scientific Reports* 15(1). <https://doi.org/10.1038/s41598-025-01568-0>, URL <https://www.scopus.com/pages/publications/105005444109?origin=resultslist>
- Allen R, Zamora J, Arroyo-Manzano D, et al (2017) External validation of preexisting first trimester preeclampsia prediction models. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 217:119–125. <https://doi.org/10.1016/j.ejogrb.2017.08.031>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85028769259&doi=10.1016%2Fj.ejogrb.2017.08.031&partnerID=40&md5=48d611dc922eb2c1c8917ee664702fc8>
- Ansbacher-Feldman Z, Syngelaki A, Meiri H, et al (2022) Machine-learning-based prediction of pre-eclampsia using first-trimester maternal characteristics and biomarkers. *Ultrasound in Obstetrics and Gynecology* 60(6):739–745. <https://doi.org/10.1002/uog.26105>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143198842&doi=10.1002%2Fuog.26105&partnerID=40&md5=8c779ae82654e4b11149dc0bc65e454a>
- Antwi E, Amoakoh-Coleman M, Vieira D, et al (2020) Systematic review of prediction models for gestational hypertension and preeclampsia. *PLoS ONE* 15(4). <https://doi.org/10.1371/journal.pone.0230955>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083554269&doi=10.1371%2Fjournal.pone.0230955&partnerID=40&md5=8661e3a6362fde5166850cb60fb739db>
- Ballard H, Yang X, Mahadevan A, et al (2024) Five-Feature Models to Predict Preeclampsia Onset Time From Electronic Health Record Data: Development and Validation Study. *Journal of Medical Internet Research* 26. <https://doi.org/10.2196/48997>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85201337845&doi=10.2196%2F48997&partnerID=40&md5=4908660f446c6cd1153b08d450a84fa5>
- Briner N, Cullen D, Halladay J, et al (2023) Tabular-to-Image Transformations for the Classification of Anonymous Network Traffic Using Deep Residual Networks.

- IEEE Access 11:113100–113113. <https://doi.org/10.1109/ACCESS.2023.3323927>, URL <https://www.scopus.com/pages/publications/85174802302?origin=resultslist>
- Brunelli E, Seidenari A, Germano C, et al (2020) External validation of a simple risk score based on the ASPRE trial algorithm for preterm pre-eclampsia considering maternal characteristics in nulliparous pregnant women: a multicentre retrospective cohort study. *BJOG: An International Journal of Obstetrics and Gynaecology* 127(10):1210–1215. <https://doi.org/10.1111/1471-0528.16246>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083963876&doi=10.1111%2F1471-0528.16246&partnerID=40&md5=fe547dbccda6e2d8fa3fbc93ce0fe300>
- Chen G, Chen Y, Shi Y, et al (2024) A dynamic prediction model for preeclampsia using the sFlt-1/PLGF ratio combined with multiple factors. *BMC Pregnancy and Childbirth* 24(1). <https://doi.org/10.1186/s12884-024-06627-4>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85197110449&doi=10.1186%2Fs12884-024-06627-4&partnerID=40&md5=917420af11c628547dee4cec671dd4a4>
- Damri A, Last M, Cohen N (2024) Towards efficient image-based representation of tabular data. *Neural Computing and Applications* 36(2):1023–1043. <https://doi.org/10.1007/s00521-023-09074-y>, URL <https://www.scopus.com/pages/publications/85173916507?origin=resultslist>
- De Kat A, Hirst J, Woodward M, et al (2019) Prediction models for preeclampsia: A systematic review. *Pregnancy Hypertension* 16:48–66. <https://doi.org/10.1016/j.preghy.2019.03.005>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065654913&doi=10.1016%2Fj.preghy.2019.03.005&partnerID=40&md5=b3d0299ca22396415efcecf594dfc9e4>
- Eberhard B, Cohen R, Wheeler N, et al (2025) Development and validation of an interpretable longitudinal preeclampsia risk prediction using machine learning. *PLOS ONE* 20(6 June). <https://doi.org/10.1371/journal.pone.0323873>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105007880816&doi=10.1371%2Fjournal.pone.0323873&partnerID=40&md5=44d1a664ea992028e12a1529291cb64b>
- Escobar G, Soltesz L, Schuler A, et al (2021) Prediction of obstetrical and fetal complications using automated electronic health record data. *American Journal of Obstetrics and Gynecology* 224(2):137–147.e7. <https://doi.org/10.1016/j.ajog.2020.10.030>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85097076666&doi=10.1016%2Fj.ajog.2020.10.030&partnerID=40&md5=561b8ff54645f270486a9cdb9bc054a8>
- Fardana N, Isnanto R, Nurhayati O (2025) Handling Class Imbalance in Health Datasets: A Comparative Study of SMOTE and SMOTEENN with TabNet. In: *Proc. Int. Conf. Informatics Comput. Sci.* Institute of Electrical and Electronics Engineers Inc., pp 305–310, <https://doi.org/10.1109/ICICoS68590.2025.11329876>, URL <https://www.scopus.com/pages/publications/105032346399?origin=resultslist>, journal Abbreviation: *Proc. Int. Conf. Informatics Comput. Sci.*

- Gómez-Martínez V, Chushig-Muzo D, Soguero-Ruiz C (2026) Tabular-to-Image Encoding Methods for Melanoma Detection: A Proof-of-Concept. *Applied Sciences (Switzerland)* 16(5). <https://doi.org/10.3390/app16052459>, URL <https://www.scopus.com/pages/publications/105032643761?origin=resultslist>
- Halladay J, Cullen D, Briner N, et al (2025) BIE: Binary Image Encoding for the Classification of Tabular Data. *Journal of Data Science* 23(1):109–129. <https://doi.org/10.6339/24-JDS1122>, URL <https://www.scopus.com/pages/publications/85218957966?origin=resultslist>
- Henderson J, Thompson J, Burda B, et al (2017) Preeclampsia screening evidence report and systematic review for the US preventive services task force. *JAMA* 317(16):1668–1683. <https://doi.org/10.1001/jama.2016.18315>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85018306277&doi=10.1001%2Fjama.2016.18315&partnerID=40&md5=c39d6c34f8c6549042216051dab655f2>
- Jain P, Saxena J, Joshi A, et al (2025) Advancing machine learning tools for early prediction and clinical diagnosis of pre-eclampsia. *Pregnancy Hypertension* 42. <https://doi.org/10.1016/j.preghy.2025.101269>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105024731544&doi=10.1016%2Fj.preghy.2025.101269&partnerID=40&md5=6fd4d868bb430632af6c6cb360e12a49>
- Jain V, Goel M, Shah K (2022) Deep Learning on Small Tabular Dataset: Using Transfer Learning and Image Classification. In: Dev A., Agrawal S.S., Sharma A. (eds) *Commun. Comput. Info. Sci.*, vol 1546 CCIS. Springer Science and Business Media Deutschland GmbH, pp 555–568, https://doi.org/10.1007/978-3-030-95711-7_46, URL <https://www.scopus.com/pages/publications/85124674885?origin=resultslist>, journal Abbreviation: *Commun. Comput. Info. Sci.*
- Jhee J, Lee S, Park Y, et al (2019) Prediction model development of late-onset preeclampsia using machine learning-based methods. *PLoS ONE* 14(8). <https://doi.org/10.1371/journal.pone.0221202>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85071308561&doi=10.1371%2Fjournal.pone.0221202&partnerID=40&md5=28b8471502ecf079b9c7dc2523e92987>
- Kempener B, Janssen E, Ellerbrock J, et al (2025) Routine mid-gestational prediction of later preeclampsia. *Best Practice and Research: Clinical Obstetrics and Gynaecology* 101. <https://doi.org/10.1016/j.bpobgyn.2025.102627>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105008556273&doi=10.1016%2Fj.bpobgyn.2025.102627&partnerID=40&md5=d89d3268be2ff207017b7849ddc4b69f>
- Lara-Abelenda F, Chushig-Muzo D, Peiro-Corbacho P, et al (2025) Transfer learning for a tabular-to-image approach: A case study for cardiovascular disease prediction. *Journal of Biomedical Informatics* 165. <https://doi.org/10.1016/j.jbi.2025.104821>, URL <https://www.scopus.com/pages/publications/105002031197?origin=resultslist>

- Layton A (2025) Artificial Intelligence and Machine Learning in Preeclampsia. *Arteriosclerosis, Thrombosis, and Vascular Biology* 45(2):165–171. <https://doi.org/10.1161/ATVBAHA.124.321673>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85215966169&doi=10.1161%2FATVBAHA.124.321673&partnerID=40&md5=27c4e38a15a9bfb58e0c7e7726f16a5>
- Li YX, Shen XP, Yang C, et al (2021) Novel electronic health records applied for prediction of pre-eclampsia: Machine-learning algorithms. *Pregnancy Hypertension* 26:102–109. <https://doi.org/10.1016/j.preghy.2021.10.006>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85118505958&doi=10.1016%2Fj.preghy.2021.10.006&partnerID=40&md5=31e22af09243df17a53d54c2cc41b40f>
- Li Z, Dai Y, Yun L, et al (2024) A prediction model for the progression from gestational hypertension to pre-eclampsia complicated with HELLP syndrome. *International Journal of Gynecology and Obstetrics* 165(3):1002–1012. <https://doi.org/10.1002/ijgo.15274>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85178238219&doi=10.1002%2Fijgo.15274&partnerID=40&md5=c8cdb39d4b9e5fe1137dceaae7656533>
- Liu M, Yang X, Chen G, et al (2022) Development of a prediction model on preeclampsia using machine learning-based method: a retrospective cohort study in China. *Frontiers in Physiology* 13. <https://doi.org/10.3389/fphys.2022.896969>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136812794&doi=10.3389%2Ffphys.2022.896969&partnerID=40&md5=a8be1498f12e62172f179974a0d77001>
- Ma Y, Lv H, Ma Y, et al (2025) Advancing preeclampsia prediction: a tailored machine learning pipeline integrating resampling and ensemble models for handling imbalanced medical data. *BioData Mining* 18(1). <https://doi.org/10.1186/s13040-025-00440-1>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105000841419&doi=10.1186%2Fs13040-025-00440-1&partnerID=40&md5=c2bc052c7800edf49742ae75174b6af9>
- Mendoza M, Bonacina E, Serrano B, et al (2022) Implementation of routine first-trimester combined screening for preeclampsia based on the Gaussian algorithm: A clinical effectiveness study. *International Journal of Gynecology and Obstetrics* 159(3):803–809. <https://doi.org/10.1002/ijgo.14192>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85127541591&doi=10.1002%2Fijgo.14192&partnerID=40&md5=6eb7dce4b75f1d6e70bceb9ecfc75447>
- Myatt L (2022) The prediction of preeclampsia: the way forward. *American Journal of Obstetrics and Gynecology* 226(2):S1102–S1107.e8. <https://doi.org/10.1016/j.ajog.2020.10.047>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85104066638&doi=10.1016%2Fj.ajog.2020.10.047&partnerID=40&md5=304435cf854572ac527cf35c7624c382>
- Noël L, Guy G, Jones S, et al (2021) Routine first-trimester combined screening for pre-eclampsia: pregnancy-associated plasma protein-A or placental growth factor?

- Ultrasound in Obstetrics and Gynecology 58(4):540–545. <https://doi.org/10.1002/uog.23669>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114737271&doi=10.1002%2Fuog.23669&partnerID=40&md5=f8ba9d7ceb2e68d0bee9c54e03a4e886>
- Paul P, Ghosh A, Takrim-Ul-Alam M, et al (2025) Handling Imbalanced Datasets with Real-World Positive Samples in Dengue Prediction Using Machine and Deep Learning Models. In: Palaiahnakote S., Saraee M., Palit R., et al (eds) Commun. Comput. Info. Sci., vol 2681 CCIS. Springer Science and Business Media Deutschland GmbH, pp 422–437, https://doi.org/10.1007/978-3-032-11335-1_29, URL <https://www.scopus.com/pages/publications/105027581738?origin=resultslist>, journal Abbreviation: Commun. Comput. Info. Sci.
- Pedersen L, Mazur-Milecka M, Rumiński J, et al (2024) A Review on Machine Learning Deployment Patterns and Key Features in the Prediction of Preeclampsia. Machine Learning and Knowledge Extraction 6(4):2515–2569. <https://doi.org/10.3390/make6040123>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85213463583&doi=10.3390%2Fmake6040123&partnerID=40&md5=946b453f6928fc83755c9b79f7581db8>
- Ponce H, Martínez-Villaseñor L, Martínez-Velasco A (2024) Explainable artificial hydrocarbon networks classifier applied to preeclampsia. Information Sciences 670. <https://doi.org/10.1016/j.ins.2024.120556>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85190447732&doi=10.1016%2Fj.ins.2024.120556&partnerID=40&md5=2b68235e3f35aa4713d188b1fa7b1821>
- Poon L, Magee L, Verlohren S, et al (2021) A literature review and best practice advice for second and third trimester risk stratification, monitoring, and management of pre-eclampsia: Compiled by the Pregnancy and Non-Communicable Diseases Committee of FIGO (the International Federation of Gynecology and Obstetrics). International Journal of Gynecology and Obstetrics 154(S1):3–31. <https://doi.org/10.1002/ijgo.13763>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85111684274&doi=10.1002%2Fijgo.13763&partnerID=40&md5=8ad05c3f95189f47bacecdfa141ac282>
- Posada Barrera A, Rodríguez Peralta L, Rosete Beas E (2025) Tabular Data-to-Image Methods for Improved Indoor Air Quality Analysis and Risk Detection. In: Proc. Int. Conf. Syst. Eng. - Cybersecur. AI: Build. reliable digit. future, CIIS. Association for Computing Machinery, Inc, pp 40–46, <https://doi.org/10.1145/3771678.3771684>, URL <https://www.scopus.com/pages/publications/105025356656?origin=resultslist>, journal Abbreviation: Proc. Int. Conf. Syst. Eng. - Cybersecur. AI: Build. reliable digit. future, CIIS
- Ranjbar A, Montazeri F, Ghamsari S, et al (2024) Machine learning models for predicting preeclampsia: a systematic review. BMC Pregnancy and Childbirth 24(1). <https://doi.org/10.1186/s12884-023-06220-1>, URL <https://www.scopus.com/inw>

- ard/record.uri?eid=2-s2.0-85181232964&doi=10.1186%2Fs12884-023-06220-1&partnerID=40&md5=8bc3197d10c5729e4c3348cf274d63a2
- Rezende K, Bornia R, Rolnik D, et al (2021) External validation of first trimester combined screening for pre-eclampsia in Brazil: An observational study. *Pregnancy Hypertension* 26:110–115. <https://doi.org/10.1016/j.preghy.2021.10.005>, URL <http://www.scopus.com/inward/record.uri?eid=2-s2.0-85118482728&doi=10.1016%2Fj.preghy.2021.10.005&partnerID=40&md5=15699a6a4389f043ee0a19ba7693900a>
- Riano-Moreno J, Vargas-Castellanos E, Pedraza A, et al (2025) Preeclampsia prediction and diagnosis: a comprehensive historical review from clinical insights to omics perspectives. *Frontiers in Medicine* 12. <https://doi.org/10.3389/fmed.2025.1689745>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105021016366&doi=10.3389%2Ffmed.2025.1689745&partnerID=40&md5=a4b566705b073f729053dacd38544990>
- Saito T, Rehmsmeier M (2015) The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE* 10(3):e0118432. <https://doi.org/10.1371/journal.pone.0118432>, URL <https://dx.plos.org/10.1371/journal.pone.0118432>
- Salinas M, Velandia D, Mayeta-Revilla L, et al (2025) An Explainable Fuzzy Framework for Assessing Preeclampsia Classification. *Biomedicines* 13(6). <https://doi.org/10.3390/biomedicines13061483>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105008974841&doi=10.3390%2Fbiomedicines13061483&partnerID=40&md5=40fecd463d7218ddec8b54172601302>
- Salmi M, Atif D, Oliva D, et al (2024) Handling imbalanced medical datasets: review of a decade of research. *Artificial Intelligence Review* 57(10):273. <https://doi.org/10.1007/s10462-024-10884-2>, URL <https://link.springer.com/10.1007/s10462-024-10884-2>
- Selke W, Sung H, Lee C, et al (2025) Exploring Tabular-to-Image Algorithms for Applying CNNs to Tabular Data. In: Liu J., Huang J., Wang X., et al (eds) *Proc. - IEEE Int. Conf. Bioinform. Biomed., BIBM*. Institute of Electrical and Electronics Engineers Inc., pp 5066–5073, <https://doi.org/10.1109/BIBM66473.2025.11356899>, URL <https://www.scopus.com/pages/publications/105033585410?origin=resultslist>, journal Abbreviation: *Proc. - IEEE Int. Conf. Bioinform. Biomed., BIBM*
- Serra Neto M, Dutra I, Mollinetti M (2022) Map-Optimize-Learn: Predicting Cardiac Pathology in Children and Teenagers with a Deep Learning Based Tabular Learning Method. In: *Proc Int Jt Conf Neural Networks*. Institute of Electrical and Electronics Engineers Inc., <https://doi.org/10.1109/IJCNN55064.2022.9889788>, URL <https://www.scopus.com/pages/publications/85140760605?origin=resultslist>, journal Abbreviation: *Proc Int Jt Conf Neural Networks*

- Singh S, Rajput S (2025) Benchmarking Explainable AI Methods for Tabular Healthcare Data: Towards Standardized and Clinically Interpretable Evaluation Frameworks. In: IEEE Pune Sect. Int. Conf., PuneCon. Institute of Electrical and Electronics Engineers Inc., <https://doi.org/10.1109/PuneCon67554.2025.11378728>, URL <https://www.scopus.com/pages/publications/105034700055?origin=resultslist>, journal Abbreviation: IEEE Pune Sect. Int. Conf., PuneCon
- Tiruneh S, Vu T, Moran L, et al (2024) Externally validated prediction models for pre-eclampsia: systematic review and meta-analysis. *Ultrasound in Obstetrics and Gynecology* 63(5):592–604. <https://doi.org/10.1002/uog.27490>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85182447843&doi=10.1002%2Fuog.27490&partnerID=40&md5=e988dc3d633d2cc7117c13b6c15d347a>
- Tiruneh S, Rolnik D, Teede H, et al (2025) Temporal validation of machine learning models for pre-eclampsia prediction using routinely collected maternal characteristics: A validation study. *Computers in Biology and Medicine* 191. <https://doi.org/10.1016/j.compbimed.2025.110183>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105002311422&doi=10.1016%2Fj.compbimed.2025.110183&partnerID=40&md5=217bb1f296b8a72e54d8f7025100a80c>
- Tripathi S, Fritz B, Abdelhack M, et al (2024) Multi-view representation learning for tabular data integration using inter-feature relationships. *Journal of Biomedical Informatics* 151. <https://doi.org/10.1016/j.jbi.2024.104602>, URL <https://www.scopus.com/pages/publications/85184840830?origin=resultslist>
- Wainstock T, Sheiner E (2022) Clinical factors associated with preeclampsia recurrence. *Pregnancy Hypertension* 30:31–35. <https://doi.org/10.1016/j.preghy.2022.08.004>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135700217&doi=10.1016%2Fj.preghy.2022.08.004&partnerID=40&md5=67bddc6de4aad1f21d9103f6ec285a92>
- Wang S, Upadhyay D, Zaman M, et al (2026) Diabetes risk modeling through tabular-to-image transformations and ensemble learning. *Informatics in Medicine Unlocked* 61. <https://doi.org/10.1016/j.imu.2026.101748>, URL <https://www.scopus.com/pages/publications/105031962815?origin=resultslist>
- Zapata R, Tolani T, Reich R, et al (2025) AI in Hypertensive Disorders of Pregnancy: Review. *American Journal of Hypertension* 38(12):1009–1019. <https://doi.org/10.1093/ajh/hpaf052>, URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-105022116932&doi=10.1093%2Fajh%2Fhpaf052&partnerID=40&md5=616ea52eb971586999aed761a46bc5ed>
- Zhu Y, Brettin T, Xia F, et al (2021) Converting tabular data into images for deep learning with convolutional neural networks. *Scientific Reports* 11(1). <https://doi.org/10.1038/s41598-021-90923-y>, URL <https://www.scopus.com/pages/publications/85107364091?origin=resultslist>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1SupplementaryMaterial.pdf.pdf](#)