Journal of Hunan University

(Natural Sciences)

Vol. 52 No. 10

October 2025

Available online at

https://jonuns.com

Open Access Article

https://doi.org/10.55463/issn.1674-2974.52.10.2

Interpretable Machine Learning for Miscarriage Risk Prediction Using Anchor Rules and Counterfactual Explanations on Large Scale Data

Aswan Supriyadi Sunge[1]*, Harco Leslie Hendric Spits Warnars[2], Maybin K. Muyeba[3], Sri Rahayu[4], Arief Teguh Nugraha[5]

[1]Informatics Engineering Department, Faculty of Engineering, Pelita Bangsa University, Bekasi, Indonesia,

[2]Computer Science Department, BINUS Graduate Program, Doctor of Computer Science (DCS), Bina Nusantara University, Jakarta, Indonesia,

[3]School of Science, Engineering & Environment (SEE), University of Salford, Salford, United Kingdom,

[4]Poltekkes Kemenkes Denpasar, Midwifery Department, Bali Province, Denpasar, Indonesia,

[5]Management Department, Faculty of Economics and Business, Pelita Bangsa University, Bekasi, Indonesia

* Corresponding author: aswan.sunge@pelitabangsa.ac.id

Article History:
Received: September 2, 2025
Revised: October 25, 2025
Accepted: November 9, 2025
Published: November 28, 2025

Abstract: Miscarriage remains a critical concern in maternal health, particularly during early pregnancy when medical intervention options are limited. Although machine learning models have shown promise in identifying high-risk cases, their predictive opacity often undermines clinical trust and adoption. This study proposes an interpretable framework for analyzing miscarriage predictions by combining anchor rules and counterfactual explanations. A Random Forest classification model is applied to a dataset of maternal medical records, and its predictions are interpreted using explainable AI techniques that generate logical rule patterns (anchors) and simulate "what-if" scenarios (counterfactuals). Anchor rules provide human-understandable conditions that strongly influence the model's decisions, while counterfactual explanations highlight the minimal changes required to alter prediction

outcomes. The primary objective of this research is to develop a clinical decision support system that is not only accurate but also transparent and trustworthy for healthcare professionals, particularly for early miscarriage risk detection. The novelty of this approach lies in the integration of two explainability methods rarely combined in maternal health contexts, alongside an evaluation strategy that moves beyond conventional performance metrics such as Accuracy, Precision, Recall, F1-Score, or AUC, focusing instead on interpretability and practical relevance. This work holds significant value for clinicians by delivering actionable insights through clear and consistent risk patterns, thereby facilitating faster, more informed decision-making. Additionally, it contributes to the computer science community by demonstrating how robust machine learning models can be aligned with transparency, ethics, and accountability principles via explainable AI techniques especially in sensitive and high-stakes domains such as maternal and child healthcare.

Keywords: Explainable AI; Miscarriage Prediction; Anchor Rule; Counterfactual Explanation; Pregnancy Health.

使用锚规则和反事实解释的大规模数据流产风险预测的可解释机器学习

摘要： 流产仍然是孕产妇健康领域的一项重要挑战，尤其是在妊娠早期，医学干预手段相对有限。尽管机器学习模型在识别高风险病例方面展现出巨大潜力，但其预测过程的不透明性常常削弱了临床上的信任与应用效果。本研究提出一种可解释的预测框架，结合锚点规则（Anchor Rules）与反事实解释（Counterfactual Explanations）两种方法，用于分析与流产相关的预测结果。我们在一组孕妇的医疗记录数据集上应用了随机森林分类模型，并借助可解释人工智能技术对模型预测进行分析，生成可理解的逻辑规则模式（锚点）以及“假设情境”模拟（反事实）。锚点规则揭示了对模型预测具有显著影响的人类可理解条件，而反事实解释则指出了改变预测结果所需的最小输入变动。本研究的主要目标是开发一个不仅具备准确性，同时也透明且值得信赖的临床决策支持系统，特别适用于流产早期风险的识别。研究的创新点在于整合了两种在孕产健康领域中鲜有联合应用的可解释性方法，并采用一种超越传统性能指标（如准确率、精确率、召回率、F1 分数或 AUC）的评估策略，更加关注模型的可解释性与实际临床相关性。对于临床医生而言，该方法通过清晰一致的风险模式提供可执行的洞见，从而加快并优化决策过程。此外，本研究也为计算机科学领域提供了参考范例，展示了如何通过可解释人工智能技术，使强健的机器学习模型在敏感和高风险的领域中实现透明性、伦理性与可问责性，特别是在母婴健康这样的关键应用场景中。

关键词：可解释人工智能、流产预测、锚点规则、反事实解释、孕期健康

1. Introduction

Miscarriage is one of the most common pregnancy complications and has a significant impact on the physical and especially mental well-being of expectant mothers [1–4]. Studies indicate that approximately 10–15% of pregnancies are at risk of ending in miscarriage, particularly once the pregnancy has been clinically confirmed [5–7]. Early detection of miscarriage risk is crucial to enable appropriate medical intervention and reduce adverse outcomes during pregnancy [8]. In this study, subject selection was based on the availability of well-documented medical pregnancy records, including vital parameters such as blood pressure, body temperature, stress levels, medical history, and early pregnancy laboratory results. The data specifically includes patients in their first trimester, as this stage presents the highest risk of miscarriage and the greatest opportunity for preventive medical action.

In recent years, Machine Learning (ML) approaches have been widely applied to risk prediction across various domains, including industry, natural disasters, and healthcare [9–11]. ML models are capable of identifying complex patterns in data that are often difficult to detect using traditional statistical methods [12]. However, many of these models are recognized as being difficult to interpret often referred to as "black

boxes" which makes it challenging for medical practitioners to understand the reasoning behind the predictions [13–14].

The lack of transparency in ML models also presents challenges in system and application development, as medical decisions require clear and understandable reasoning [15–16]. To address this issue, the field of Explainable Artificial Intelligence (XAI) has emerged, offering techniques that enable the interpretation of ML model outcomes. One such approach is the use of anchor rules, which generate human-understandable if-then statements and provide insights into the factors influencing the predictions [17].

In addition, counterfactual explanations have been proposed as an effective XAI method by identifying the minimal changes in input features that would lead to a different prediction outcome. This approach not only enhances users' understanding of the model but also supports more accurate planning of clinical interventions. By integrating techniques such as anchor rules and counterfactual explanations, miscarriage risk prediction models can become more transparent and reliable in medical practice [18].

The findings of this study carry significant practical implications in the context of maternal healthcare. By uncovering consistent predictive patterns in early pregnancy data, healthcare professionals are better equipped to perform targeted screenings and initiate timely interventions potentially even before clinical symptoms become evident. This proactive approach may reduce delays in medical response and help prevent adverse outcomes for both the mother and the fetus. From a theoretical perspective, this research contributes to the advancement of XAI in the healthcare domain, particularly through the integration of anchor rules and counterfactual explanations. This dual-method approach highlights the potential for ML models to move beyond “black-box” predictions by offering clear, interpretable reasoning that can be understood by clinical practitioners. In doing so, the study supports the development of AI systems that are not only technically robust but also ethically aligned, transparent, and context-sensitive especially in high-stakes and sensitive applications such as maternal health risk assessment.

This study offers several significant contributions, namely:

1. Developing an interpretable framework that integrates ML models with anchor rules and counterfactual explanation techniques.
2. Applying XAI techniques that have been rarely used in the context of maternal health, particularly on structured datasets.
3. Providing clinical intervention scenario simulations based on counterfactual explanations.
4. Proposing an evaluation approach based on interpretability and clinical relevance, rather than relying solely on conventional metrics such as Accuracy, Precision, Recall, F1-Score, or AUC.

2. Similar Previous Research

Research on miscarriage prediction using ML approaches continues to evolve alongside the increasing availability of maternal medical record data. Several recent studies have applied ML to identify miscarriage risk factors, such as idiopathic spontaneous miscarriage cases, using patient data [19]. Automated systems capable of early prediction based on maternal health parameters have also been developed [20]. In addition, ML-based approaches for identifying threatened miscarriage cases have demonstrated higher accuracy compared to traditional methods [21]. Findings from these studies highlight the potential of ML as a powerful tool for more efficient and precise early detection of miscarriage.

Among various ML algorithms, Random Forest stands out due to its ability to handle complex and nonlinear data as well as its resistance to overfitting. In the context of miscarriage prediction, this model has been used to evaluate the importance of medical features in patient records, such as age, body mass index (BMI), and history of previous miscarriages based on early trimester data [22]. Beyond accuracy, the model also provides feature importance values that assist medical professionals in understanding which factors contribute most significantly to the predictions [23–24].

However, accuracy alone is insufficient in medical contexts that require transparency and trust from both patients and healthcare professionals. Therefore, interpretability approaches such as Anchor Rules have begun to be integrated into miscarriage prediction models. This method generates conditional logical rules (if-then statements) that explain model predictions in a way that is easily understandable by humans [25–26].

In addition to Anchor Rules, Counterfactual Explanation models offer an intuitive and informative approach by demonstrating how small changes in input features can lead to different prediction outcomes. By providing “what-if” scenarios, this method helps non-technical users, including healthcare professionals and patients, understand what needs to be altered to achieve more favorable prediction results. This approach not only enhances model transparency but also supports clinical decision-making by offering actionable insights [27–28].

Although interpretability models such as SHAP and LIME have been widely used and are relatively easy to understand [29–30], the application of methods like Anchor Rules and Counterfactual Explanations in the context of miscarriage risk prediction remains scarce in the literature. This gap indicates an opportunity to explore approaches that are more transparent and interpretable for non-technical users, such as healthcare professionals. This study aims to address this gap by integrating Anchor Rules and Counterfactual

Explanations into a ML–based miscarriage risk prediction framework.

3. Proposed Model

This study is an experimental research based on a quantitative approach, grounded in a systematic literature review and relevant theoretical frameworks within the domain of reproductive health. The primary objective of this research is to apply ML methods to identify the most significant attributes contributing to increased miscarriage risk. By emphasizing model interpretability and transparency, this approach aims to provide deeper computational insights into key risk factors and support more informative clinical decision-making processes. Through analyzing feature contributions to prediction outcomes, this study seeks to make a meaningful scientific contribution in the context of data-driven miscarriage prevention and management. Furthermore, the feature identification strategy employed enables the exploration of hypothetical conditions that could alter model outputs, thereby offering richer contextual insights into individual risk, as illustrated in Figure 1.

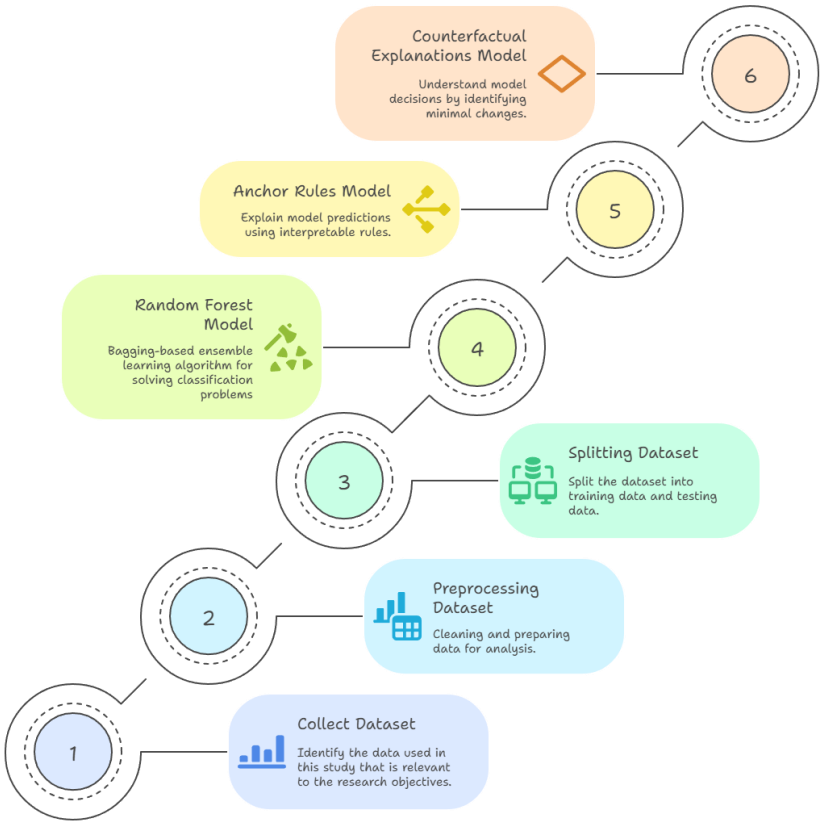

Figure 1. Framework Model

(Source: Developed by the author)

3.1. Collect Dataset

This section aims to identify the data used in this study and ensure that it can be effectively tested and is relevant to the research objectives. Table 1 presents the dataset utilized in this study, which was obtained from a GitHub repository and is licensed for open use according to the source license terms. The dataset is publicly available via the link https://github.com and consists of data collected from various reproductive health sources for analysis. It includes a total of 1,000,000 samples, each comprising 13 predictor features and 1 target class (0 = No Miscarriage, 1 = Miscarriage). The purpose of this dataset is to identify the key factors contributing to miscarriage risk, thereby supporting the development of accurate and interpretable predictive models in a clinical context.

This information reflects the complexity of understanding and predicting miscarriage risk due to the numerous biological, medical, and demographic factors that can influence pregnancy outcomes. With the advancement of data-driven approaches in healthcare, the use of new techniques and tools to analyze miscarriage risk has also increased. However, there are several limitations, such as the lack of up-to-date medical data that comprehensively reflects changes in patient conditions or recent features.

Additionally, reliance on certain variables that may not be fully relevant to specific contexts or populations in different countries can affect the model's generalizability. The data used in this study originates from a single open-source repository, and therefore may not fully represent practices and conditions across various healthcare systems or geographic regions. Nonetheless, the conducted analysis provides meaningful insights into identifying key factors influencing miscarriage risk and supports more accurate, data-driven clinical decision-making.

Table 1. Dataset Features and Description (Source: From Github)

No	Features	Description	Values
1	Age	Age at pregnancy	Numerical
2	BMI	Weight during pregnancy	Numerical
3	Nmisc	History of previous misccarriage	Categorical
4	Activity	The level of the activity of the woman during the day.	Numerical
5	Location	Location where the woman spends her time	Categorical
6	Temp	Body Temperature	Numerical
7	BPM	Heart Rate Variability	Numerical
8	Stress	Stress motion	Categorical
9	Bp	Blood Pressure	Categorical
10	Miscarriage	Class	Categorical

3.2. Preprocessing Dataset

The process begins with a preprocessing stage aimed at preparing the data before testing, thereby minimizing errors and supporting optimal ML data analysis in the context of miscarriage risk prediction. The steps in this stage may vary depending on the analysis objectives and the models used; however, data quality must be thoroughly checked prior to the training process. Common techniques employed include data inspection and cleaning, such as identifying unrecognized symbols, duplicates, typographical errors, or missing data. The primary goal is to ensure that all data fields are complete, contain no gaps, and are error-free. To check for missing or invalid data, the syntax *data.isnull().sum()* is often used to display the number of errors or missing values, as presented in Table 2.

Table 2. Checking and Cleaning Data (Source: Developed by author with Google Colab)

No	Features	Results
1	Age	0
2	BMI	0
3	Nmisc	0
4	Activity	0
5	Location	0
6	Temp	0
7	BPM	0
8	Stress	0
9	Bp	0
10	Miscarriage	0

3.3. Splitting Dataset

At this stage, the dataset is partitioned into two main subsets, with 80% allocated for training data and 20% for testing data. This split aims to provide sufficient data for the ML model to learn important patterns during training while also offering a representative testing set for independent model performance evaluation. This approach ensures the development of a reliable and valid model for miscarriage risk prediction and facilitates an objective assessment of the model's generalizability through testing on data that was not used during training.

3.4. Random Forest Model

Random Forest is an ensemble learning algorithm based on bagging, widely used to address classification and regression problems. In classification tasks, it builds a collection of decision trees, each trained using bootstrap sampling and a random subset of features at each split node. In the context of this study, the dataset is a binary classification dataset with labels 0 or 1, where the model aims to map input features to one of the two target classes. Although this model exhibits high predictive performance and robustness against overfitting, its complexity limits interpretability. To overcome this, two model interpretability approaches are integrated: Anchor Rules and Counterfactual Explanations. Anchor Rules are employed to extract stable and easily understandable conditional rules underlying the model's predictions, while Counterfactual Explanations aim to identify minimal changes in input features that could lead to different predictions. Combining Random Forest with these two approaches enables a more transparent and informative interpretation of classification results.

Mathematically, the final prediction $\hat{y}$ for an input x given by:

$$\hat{y} = \arg\max_{c \in \mathcal{C}} \sum_{i=1}^{B} I\left(h_i(x) = c\right) \quad (1)$$

Where:

- C is the set of all classes (in this case {0, 1})
- B is the total number of trees in the Random Forest
- $h_i(x)$ is the prediction from the tree to-i
- $\mathbb{I}(\cdot)$ is the indicator function (1 if true, 0 if false)
- arg max selects the class c with the highest number of votes.

The combination of training with random data samples and random feature selection at each node makes Random Forest highly effective in reducing model variance and improving generalization. Additionally, this model supports feature importance analysis, providing insights into the contribution of each feature to the classification outcome.

3.5. Anchor Rules Model

It is a model-agnostic interpretability method aimed at explaining ML model predictions through a series of simple and understandable conditional rules. Each anchor consists of a set of logical conditions which, when satisfied by an instance, ensure that the model's prediction remains consistent despite variations in other features. The quality of an anchor is evaluated using the precision metric, which indicates how often the prediction remains stable on data that meet the anchor's conditions.

The precision formula for an anchor A is defined as follows:

$$\text{Precision}(A) = E_{z\sim \mathcal{D}(z|A)}\left[I\left(f(z) = f(x)\right)\right] \quad (2)$$

Where:

- A is the anchor, a set of rules (conditions) that constrain specific feature values.
- x is the instance being explained.
- $f(x)$ is the model's prediction for the instance x.
- $z\sim D(z|A)$ is a sample from the conditional distribution of A, i.e., data that satisfy the anchor rules A
- $I(f(z)=f(x))$ is the indicator function: it returns 1 if the prediction for sample z is the same as the prediction for x and 0 otherwise.
- The precision of anchor A indicates how consistently the rule preserves the same prediction.

The model aims to find an anchor A with high precision—typically above a certain threshold (e.g., 95%) that explains why the model makes a specific prediction for an instance x, while also providing a guarantee that the prediction will remain the same for the majority of samples satisfying the anchor conditions.

3.6. Counterfactual Explanations Model

This approach to model interpretability aims to provide an understanding of the model's decisions by identifying the minimal changes required to the input in order to produce a different or desired prediction. In this context, counterfactual explanations are formulated as an optimization process to find an alternative instance x' that is as close as possible to the original instance x, but results in a different prediction from the model. Mathematically, this formulation is expressed as follows:

$$\text{CE}(x) = \arg\min_{x'} \mathcal{L}\left(f(x'), y_{\text{target}}\right) + \lambda \cdot d(x, x') \quad (3)$$

Where:

- $f(x')$ an anchor is defined as a set of conditions that fix certain feature values in order to explain a model's prediction.
- y_{target}: is the desired label.
- $L(f(x'))$: a loss function used to minimize the difference between the prediction and the target.
- d(x,x′) : a distance function between the original instance x and counterfactual x′
- λ: a regularization parameter that balances the trade-off between maintaining proximity to the original input and achieving a successful change in the model's prediction.

This approach offers actionable insights by identifying which features should be modified to obtain a different model output. Therefore, it is particularly relevant for ML systems that require transparency, interpretability, and support for accountable decision-making.

4. Discussion and Results

The experimental results present an interpretability analysis of the miscarriage risk classification model using Anchor Rules and Counterfactual Explanations. The primary objective is to identify the most influential features contributing to the model's predictions and to evaluate the model's stability and sensitivity to input changes. These interpretability techniques are employed not only to strengthen understanding of the model's decision-making process but also to promote transparency in data-driven miscarriage prediction systems. This analysis provides clearer insights into model behavior, enabling more focused and contextually relevant discussions of the results, thereby enhancing clinical applicability. Moreover, the findings are compared with those of prior studies to assess the consistency and validity of the approach. The comparison reveals that most previous works have largely focused on evaluating model accuracy, comparing algorithmic performance,

or conducting conventional statistical analyses—often without offering explanations that can be directly interpreted by medical practitioners. In many cases, evaluations have relied solely on standard metrics such as Accuracy, Precision, Recall, F1-Score, and AUC, which, while important, are insufficient to explain why a prediction was made or how it could inform actionable clinical decisions. In this context, the use of interpretability methods in this study offers a significant advancement by providing concrete, actionable, and human-understandable explanations. This contributes to the development of clinical decision-support systems that are not only technically robust but also transparent, ethical, and trustworthy for real-world medical applications.

4.1. Anchor Rules Model

Figure 2 shows instance 0, where the classification model predicts that the individual belongs to class 1, indicating a risk of miscarriage. The explanation for this prediction is generated using the Anchor Explanation approach, which identifies a specific combination of features that consistently lead the model to produce the same prediction.

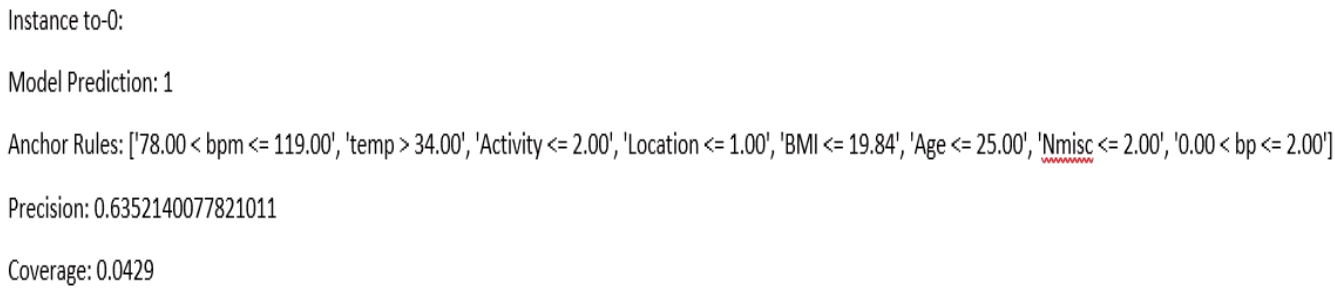

Figure 2. Result Anchor Model
(Source: Developed by author with Google Colab)

The formed rules include a heart rate range of 78–119 bpm, BMI ≤ 19.84, body temperature > 31°C, activity level ≤ 2, location ≤ 1, and age ≤ 25 years. From a computational perspective, these rules achieve a precision of 0.616, indicating that approximately 61% of data instances with similar characteristics are consistently predicted by the model, while a coverage of 0.0718 shows that the rules cover only a small portion of the entire dataset. Medically, these features are associated with several potential risk factors, such as low BMI reflecting poor nutritional status, elevated body temperature indicating physiological stress, and young age and low activity levels which may also affect pregnancy conditions. Therefore, the use of Anchor Rules not only enhances the interpretability of the data-driven classification model but also provides deeper insights into medically relevant factors for predicting miscarriage risk.

4.2. Counterfactual Explanations Model

Based on the obtained results, Figure 3 illustrates a counterfactual analysis, a method used to explain why an artificial intelligence model makes a particular prediction.

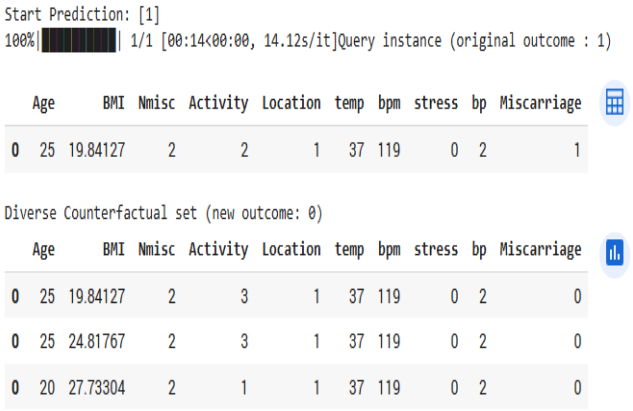

Figure 3. Result Counterfactual Model
(Source: Developed by author with Google Colab)

In the first section, 'Start Prediction,' the data shows the initial prediction result for an individual. With an age of 25 years, a BMI of 19.84, and an activity level of 2, the model predicts a miscarriage risk with a value of 1. This indicates that, based on this combination of characteristics, the model identifies a high risk. Next, in the 'Diverse Counterfactual Set' section, the model presents alternative scenarios. These scenarios demonstrate the minimal changes required in the input data to change the miscarriage prediction from 1 (yes) to 0 (no). The first scenario shows that increasing the activity level from 2 to 3 immediately alters the prediction, implying that increased physical activity is a significant factor in reducing risk. The second and third scenarios provide examples of other effective factor combinations. The second scenario illustrates

that increasing activity combined with a higher BMI can also reverse the prediction. Meanwhile, the third scenario indicates that a younger age and higher BMI, even with a decreased activity level, can result in a low predicted miscarriage risk.

From a medical perspective, these findings are highly relevant as factors such as BMI, age, and physical activity have been clinically identified as key elements in pregnancy health. Low BMI is often associated with malnutrition and an increased risk of pregnancy complications, including miscarriage, whereas a healthy or slightly higher BMI reflects better nutritional status that can support a more stable pregnancy. Moderate physical activity is also known to contribute to maintaining cardiovascular health and hormonal balance in pregnant women, indirectly reducing miscarriage risk. Young age is generally considered a favorable factor in pregnancy, although overall health conditions and comorbidities must still be taken into account.

Therefore, counterfactual analysis in this context not only provides insights into the most significant features influencing the model's decisions but also aligns with existing medical knowledge regarding miscarriage risk factors. This approach enhances the interpretability of artificial intelligence systems while offering a solid foundation for more targeted and data-driven medical interventions.

5. Conclusions and Future Directions

This study demonstrates that integrating interpretability methods such as Anchor Explanations and Counterfactual Explanations into predictive classification models can offer deep and transparent insights into the model's decision-making process, particularly in the context of miscarriage risk prediction. Anchor Rules effectively identified combinations of key features such as heart rate, low BMI, elevated body temperature, low physical activity, and young maternal age—that consistently contribute to high-risk predictions. Although these rules cover a limited portion of the data, they exhibit stable levels of precision and coverage, making them reliable for specific subgroups within the population. Meanwhile, Counterfactual Explanations provide a more interventional perspective by simulating minimal alternative scenarios that could reverse the model's prediction. These explanations highlight actionable variables, such as physical activity, BMI, and maternal age, that may serve as potential targets for clinical intervention.

The findings of this study are consistent with previous research that links factors such as body mass index, physical activity, and maternal age with miscarriage risk. However, this study offers a novel contribution by introducing an explicitly interpretable approach through explainable AI techniques, particularly via the integration of anchor rules and counterfactual explanations—two methods that are rarely combined in miscarriage prediction studies. Prior research has typically focused on the technical performance of models, such as evaluating classification accuracy, comparing algorithmic performance, or conducting traditional statistical analyses, without providing insights that are readily interpretable by medical practitioners. Often, evaluations are limited to conventional metrics like Accuracy, Precision, Recall, F1-Score, and AUC, which, while important, fall short in explaining why a particular prediction was made or how actionable interventions can be derived. In this context, the present study advances existing approaches by not only identifying risk factors but also offering clinically relevant and concrete explanatory mechanisms. This interpretability enhances trust and promotes the adoption of predictive models within medical environments.

Theoretically, this study reinforces the urgency and relevance of applying interpretable ML approaches within the healthcare domain, particularly when addressing sensitive and complex issues such as miscarriage. The use of anchor rules and counterfactual explanations demonstrates that AI-based predictive models can be developed while maintaining transparency, accountability, and clarity. Practically, the results of this study create opportunities for healthcare professionals to access prediction outputs that are not only accurate but also understandable and actionable, thereby supporting faster and evidence-based clinical decision-making. Nonetheless, limitations related to the subjectivity of certain input variables and the absence of direct validation from medical experts remain critical concerns that must be addressed before these models can be widely adopted in clinical practice.

To improve the validity and utility of the model, future research is encouraged to incorporate larger and more representative datasets that reflect populations from diverse regions or demographic backgrounds. Additionally, integrating more comprehensive clinical variables such as medical history, psychological stress levels, and medication interactions would provide deeper context for predictive modeling. Implementing constraints in the generation of counterfactuals is also necessary to ensure that the resulting scenarios remain medically realistic and clinically meaningful. Moreover, the active involvement of medical professionals in validating the model's interpretability is crucial to ensure that the outputs align with ethical standards and safe clinical practices. This collaborative approach is expected to facilitate the development of AI systems that truly function as reliable, data-driven, and personalized decision-support tools in clinical settings.

Declarations

Author Contributions

Conceptualization, A.S.S; methodology, A.S.S; software, A.S.S., and S.R.; validation, A.S.S, and H.L.H.S.W.; formal analysis, A.S.S., and S.R.; investigation, A.S.S; resources, S.R; data curation, A.S.S.; writing—original draft preparation, H.L.H.S.W., and M.K.M; writing—review and editing, A.S.S.; visualization, H.L.H.S.W., and A.T.N.; supervision, M.K.M; project administration, A.S.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

Data available in a publicly accessible repository that does not issue DOIs: Publicly available datasets were analyzed in this study. This data can be found here: [https://github.com/hibaasri/Miscarriage-Prediction.git]

Funding

This research is supported by Pelita Bangsa University as part of the Research Funding for the Research Information Base and Community Service (BIMA) entitled "Development of a Multi-Parameter Predictive Model for Miscarriage Risk Identification: Integration of Biomarkers and Machine Learning" with the Master Contract Number: 125/C3/DT.05.00/PL/2025, Derivative Contract Date: June 04, 2025 and Derivative Contract Number: 7927/LL4/PG/2025-005/07/KP/.H/UPB/2025.

Institutional Review Board Statement

Not Applicable, this study did not involve direct interaction with human participants or animals. The data used were obtained from publicly available sources where ethical approval had already been addressed by the original data providers.

Informed Consent Statement

Not Applicable, this study did not involve the collection of data from human subjects directly. The dataset used was publicly available and obtained from an open-source repository, where all data were anonymized and published for research purposes.

Conflicts of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

References

[1] AL-ALAMI Z., ABU-HUWAIJ R., HAMADNEH S., and TAYBEH E. Understanding Miscarriage Prevalence and Risk Factors: Insights from Women in Jordan. *Medicina,* 2024, 60(7): 1044. https://doi.org/10.3390/medicina60071044

[2] KIDD C., O'DRISCOLL D., O'BYRNE L.J., O'KEEFFE G.W., KHASHAN A.S., and Maher G.M. The association between threatened miscarriage in early pregnancy and depression or anxiety in offspring in late adolescence. *Journal of Affective Disorders,* 2025, 382: 48-54. https://doi.org/10.1016/j.jad.2025.04.060

[3] REGAN L., and RAI R. Epidemiology and the medical causes of miscarriage. *Best Practice & Research Clinical Obstetrics & Gynaecology.* 2000, 14(5): 839-854. https://doi.org/10.1053/beog.2000.0123

[4] OXFORD UNIVERSITY'S NUFFIELD DEPARTMENT OF WOMEN'S & REPRODUCTIVE HEALTH. *Miscarriage: Research Overview, Risk Factors, and Current Studies*. Nuffield Department of Women's & Reproductive Health, University of Oxford, 2024. https://www.wrh.ox.ac.uk/research/miscarriages

[5] QUENBY, S., GALLOS, I. D., DHILLON-SMITH, R. K., PODESEK, M., STEPHENSON, M. D., FISHER, J., BROSENS, J. J., BREWIN, J., RAMHORST, R., LUCAS, E. S., Mccoy, R. C., ANDERSON, R., DAHER, S., REGAN, L., AL-MEMAR, M., BOURNE, T., Macintyre, D. A., RAI, R., CHRISTIANSEN, O. B., SUGIURA-OGASAWARA, M., and COOMARASAMY, A. Miscarriage Matters: The Epidemiological, Physical, Psychological, and Economic Costs of Early Pregnancy Loss. *The Lancet*, 2021, 397(10285): 1658–1667. https://doi.org/10.1016/S0140-6736(21)00682-6.

[6] MAGNUS M.C., HOCKEY R.L., HABERG S.E., and MISHRA G.T. Pre-pregnancy lifestyle characteristics and risk of miscarriage: the Australian Longitudinal Study on Women's Health. *The Australian Longitudinal Study on Women's Health*. BMC Pregnancy Childbirth, 2022, 22: 169. https://doi.org/10.1186/s12884-022-04482-9

[7] SOFIA N., FADLYANA E., IRIANTI S., KRISNADI S., and SUSIARNO H. Maternal Risk Factors among Pregnant Women with Miscarriage. *Althea Medical Journal,* 2024, 11(1):57-62. https://doi.org/10.15850/amj.v11n1.3037

[8] ALVES C., JENKINS S.M., and RAPP A. *Early Pregnancy Loss (Spontaneous Abortion)*. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing, 2023 https://www.ncbi.nlm.nih.gov/books/NBK560521/

[9] HEGDE J., and ROKSETH B. Applications of machine learning methods for engineering risk assessment – A review. *Safety Science*, 2020, 122:104492.

https://doi.org/10.1016/j.ssci.2019.09.015
[10] YANG Z., GOKON H., and Yu Q. Machine learning-based identification and assessment of snow disaster risks using multi-source data: Insights from Fukui prefecture, Japan. *Progress in Disaster Science*, 2025, 26:100426. https://doi.org/10.1016/j.pdisas.2025.100426
[11] KARUNANIDHI D., RAJ M.R.H., PRAPANCHAN V.N., and SUBRAMANI T. Predicting groundwater fluoride levels for drinking suitability using machine learning approaches with traditional and fuzzy logic models-based health risk assessment. *Geoscience Frontiers*. 16(4):102087. https://doi.org/10.1016/j.gsf.2025.102087
[12] GUO J., FEI T., YANG X., LILBURNE L., MOOT D., MARTIN B., TEIXEIRA., and YANG M. Can we use APSIM's embedded expert knowledge to 'train' a machine learning model to identify, at high resolution, land suitable to grow lucerne (Medicago sativa L.) using simplified climate data?. *European Journal of Agronomy,* 2025, 171:127815. https://doi.org/10.1016/j.eja.2025.127815
[13] PETCH J., DI S., and NELSON W. Opening the Black Box: The Promise and Limitations of Explainable Machine Learning in Cardiology. *Canadian Journal of Cardiology*, 2022, 38(2):204-213. https://doi.org/10.1016/j.cjca.2021.09.004
[14] HASSIJA V., CHAMOLA V., MAHAPATRA A., SINGAL A., GOEL D., HUANG K., SCARDAPANE S., SPINELLI I., MAHMUD M., and HUSSAIN A. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation Journal*, 2024, 16:45-74. https://doi.org/10.1007/s12559-023-10179-8
[15] ALANAZI A. Using machine learning for healthcare challenges and opportunities. *Informatics in Medicine Unlocked*, 2022, 30:100924. https://doi.org/10.1016/j.imu.2022.100924
[16] MANIKANDA M., VENKATESH., ILAKKIYA T., KRISHNAMOORTHI M., RAMU M., and Senthilnathan C.R. Advancements and Difficulties of Machine Learning Applications in Customised Medicine. International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2024, pp. 1-4. https://doi.org/10.1109/ICPECTS62210.2024.10780344
[17] UKWATHTHA J., HERATH S., and MEDDAGE D.P.P. A review of machine learning (ML) and explainable artificial intelligence (XAI) methods in additive manufacturing (3D Printing). *Materials Today Communications*, 2024, 41:110294. https://doi.org/10.1016/j.mtcomm.2024.110294
[18] AL-NAJJAR H.A.H., PRADHAN B., BEYDOUN G., SARKAR R., PARK H., and ALAMRI A. A novel method using explainable artificial intelligence (XAI)-based Shapley Additive Explanations for spatial landslide prediction using Time-Series SAR dataset. *Gondwana Research,* 2023, 123:107-124. https://doi.org/10.1016/j.gr.2022.08.004
[19] SHERPA D., ABHIJIT R.D., MITRA I., DHAR D., SHARMA S., CHAKRABORTY P., and CHAUDHURY K. Prediction of Idiopathic Recurrent Spontaneous Miscarriage using Machine Learning. International Conference on Computer, Electrical & Communication Engineering (ICCECE), Kolkata, India, 2023, pp. 1-8. https://doi.org/10.1109/ICCECE51049.2023.10085363
[20] ALJAMEEL S.S., ALJABRI M., ASLAM N., ALOMARI D.M., ALYAHYA A., ALFARIS S., BLHARITH M., ABAHUSSAIN H., BOUJLEA D., and ALSULMI E.S. An Automated System for Early Prediction of Miscarriage in the First Trimester Using Machine Learning. *Computers, Materials and Continua*, 2023, 75(1):1291-1304. https://doi.org/10.32604/cmc.2023.035710
[21] SINGH S., TIWARI S., GOEL P., and TIWARI D. A Retrospective: Sightseeing Excursion of Threatened Miscarriage Pertaining Ensemble Machine Learning Algorithms. 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-7. https://doi.org/10.1109/ISCON57294.2023.10111961
[22] AMITAI, T., KAN-TOR, Y., OR, Y., SHOHAM, Z., SHOFARO, Y., RICHTER, D., HAR-VARDI, I., BEN-MEIR, A., SREBNIK, N., and BUXBOIM, A. Embryo classification beyond pregnancy: early prediction of first trimester miscarriage using machine learning. *Journal of assisted reproduction and genetics*, 2023, 40(2): 309–322. https://doi.org/10.1007/s10815-022-02619-5
[23] BROWN M.G.L., PETERSON M.G., TEZAUR I.K., PETERSON K.J., and BULL D.L. Random Forest regression feature importance for climate impact pathway detection. *Journal of Computational and Applied Mathematics,* 2025, 464:116479. https://doi.org/10.1016/j.cam.2024.116479
[24] HEIDARI M., MAATAR M.H., GHAFFARI H. Forward propagation dropout in deep neural networks using Jensen–Shannon and random forest feature importance ranking. *Neural Networks*, 2023, 165:238-247. https://doi.org/10.1016/j.neunet.2023.05.044
[25] THANATHAMATHEE P., SAWANGARREERAK S., and NIZAM D.N.M. Enhancing Going Concern Prediction with Anchor Explainable AI and Attention-Weighted XGBoost. 2024, *IEEE Access*, 12, pp. 68345-68363. https://doi.org/10.1109/ACCESS.2024.3401007
[26] DELAUNAY J., GALÁRRAGA L., and LARGOUËT C. Improving Anchor-based Explanations. International Conference on Information & Knowledge Management (CIKM), New York, USA, pp. 3269-3272.

https://doi.org/10.1145/3340531.3417461
[27] ZHANG H., and ZHANG J. Efficient and effective counterfactual explanations for random forests. *Expert Systems with Applications*, 2025, 293:128661. https://doi.org/10.1016/j.eswa.2025.128661
[28] KURTZ J., BIRBIL Ş. Ĭ., and HERTOG D. D. Counterfactual explanations for linear optimization. *European Journal of Operational Research*, 2025. https://doi.org/10.1016/j.ejor.2025.06.016
[29] ABEKOON, T., SAJINDRA H., RATHNAYAKE N., EKANAYAKE I.U., JAYAKODY A., and RATHNAYAKE U. A novel application with explainable machine learning (SHAP and LIME) to predict soil N, P, and K nutrient content in cabbage cultivation. Smart Agricultural Technology, 2025, 11:100879. https://doi.org/10.1016/j.atech.2025.100879
[30] HERMOSILLA, P., BERRÍOS, S., and ALLENDE-CID, H. Explainable AI for Forensic Analysis: A Comparative Study of SHAP and LIME in Intrusion Detection Models. *Applied Sciences*, 2025, 15(13):7329. https://doi.org/10.3390/app15137329

参考文:

[1] AL-ALAMI Z., ABU-HUWAIJ R., HAMADNEH S., TAYBEH E.
流产的患病率与风险因素解析：来自约旦女性的洞察 [J]. *Medicina*, 2024, 60(7): 1044. https://doi.org/10.3390/medicina60071044
[2] KIDD C.、O'DRISCOLL D.、O'BYRNE L. J.、O'KEEFFE G. W.、KHASHAN A. S. **和** MAHER G. M.
早期妊娠先兆流产与后代青少年晚期抑郁或焦虑之间的关联[J]. *Journal of Affective Disorders*, 2025, 382: 48–54. https://doi.org/10.1016/j.jad.2025.04.060
[3] REGAN L. **和** RAI R.
流产的流行病学与医学原因[J]. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 2000, 14(5): 839–854. https://doi.org/10.1053/beog.2000.0123
[4] **牛津大学**纳菲尔德妇女与生殖健康系.
流产：研究概述、风险因素与当前研究[R]. **牛津大学**纳菲尔德妇女与生殖健康系, 2024. https://www.wrh.ox.ac.uk/research/miscarriages
[5] QUENBY S.、GALLOS I. D.、DHILLON-SMITH R. K.、PODESEK M.、STEPHENSON M. D.、FISHER J.、BROSENS J. J.、BREWIN J.、RAMHORST R.、LUCAS E. S.、MCCOY R. C.、ANDERSON R.、DAHER S.、REGAN L.、AL-MEMAR M.、BOURNE T.、MACINTYRE D. A.、RAI R.、CHRISTIANSEN O. B.、SUGIURA-OGASAWARA M. **和** COOMARASAMY A.
流产的重要性：早期妊娠损失在流行病学、生理、心理及经济方面的代价[J]. *The Lancet*, 2021, 397(10285): 1658–1667. https://doi.org/10.1016/S0140-6736(21)00682-6.
[6] MAGNUS M. C.、HOCKEY R. L.、HABERG S. E. **和** MISHRA G. T.
孕前生活方式特征与流产风险：澳大利亚女性健康纵向研究[J]. *BMC Pregnancy and Childbirth*, 2022, 22: 169. https://doi.org/10.1186/s12884-022-04482-9
[7] SOFIA N.、FADLYANA E.、IRIANTI S.、KRISNADI S. **和** SUSIARNO H.
流产孕妇的母体风险因素[J]. Althea Medical Journal, 2024, 11(1): 57–62. https://doi.org/10.15850/amj.v11n1.3037
[8] ALVES C.、JENKINS S. M. **和** RAPP A.
早期妊娠损失（自然流产）[M/OL]. 见：*StatPearls* [**互**联网]. **特雷**热岛（佛罗里达州）：*StatPearls Publishing*, 2023. https://www.ncbi.nlm.nih.gov/books/NBK560521/
[9] HEGDE HEGDE J. **和** ROKSETH B.
机器学习方法在工程风险评估中的应用综述[J]. *Safety Science*, 2020, 122: 104492. https://doi.org/10.1016/j.ssci.2019.09.015
[10] YANG Z.、GOKON H. **和** YU Q.
基于机器学习的多源数据雪灾风险识别与评估：来自日本福井县的研究启示[J]. *Progress in Disaster Science*, 2025, 26: 100426. https://doi.org/10.1016/j.pdisas.2025.100426
[11] KARUNANIDHI D.、RAJ M. R. H.、PRAPANCHAN V. N. **和** SUBRAMANI T.
利用机器学习方法结合传统和模糊逻辑模型进行地下水氟含量饮用适宜性及健康风险评估的预测[J]. *Geoscience Frontiers*, 16(4): 102087. https://doi.org/10.1016/j.gsf.2025.102087
[12] GUO J.、FEI T.、YANG X.、LILBURNE L.、MOOT D.、MARTIN B.、TEIXEIRA **和** YANG M.
我们能否利用 APSIM **嵌入的**专家知识，通过简化气候数据训练机器学习模型，高分辨率识别适合种植紫花苜蓿（Medicago sativa L）**的土地？**[J]. *European Journal of Agronomy*, 2025, 171: 127815. https://doi.org/10.1016/j.eja.2025.127815
[13] PETCH J.、DI S. **和** NELSON W.
打开黑箱：可解释机器学习在心脏病学中的前景

与局限[J]. *Canadian Journal of Cardiology*, 2022, 38(2): 204–213. https://doi.org/10.1016/j.cjca.2021.09.004
[14] HASSIJA V.、CHAMOLA V.、MAHAPATRA A.、SINGAL A.、GOEL D.、HUANG K.、SCARDAPANE S.、SPINELLI I.、MAHMUD M. 和 HUSSAIN A.
黑箱模型的解释：可解释人工智能综述[J]. *Cognitive Computation Journal*, 2024, 16: 45–74. https://doi.org/10.1007/s12559-023-10179-8
[15] ALANAZI A.
机器学习在医疗健康中的挑战与机遇[J]. *Informatics in Medicine Unlocked*, 2022, 30: 100924. https://doi.org/10.1016/j.imu.2022.100924
[16] MANIKANDA M.、VENKATESH、ILAKKIYA T.、KRISHNAMOORTHI M.、RAMU M. 和 SENTHILNATHAN C. R.
定制医疗中机器学习应用的进展与挑战[C]// International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 印度钦奈, 2024: 1-4. https://doi.org/10.1109/ICPECTS62210.2024.10780344
[17] UKWATHTHA J.、HERATH S. 和 MEDDAGE D. P. P.
增材制造（3D 打印）中机器学习（ML）和可解释人工智能（XAI）方法综述[J]. *Materials Today Communications*, 2024, 41: 110294. https://doi.org/10.1016/j.mtcomm.2024.110294
[18] AL-NAJJAR H. A. H.、PRADHAN B.、BEYDOUN G.、SARKAR R.、PARK H. 和 ALAMRI A.
一种基于可解释人工智能（XAI）Shapley 加法解释的新方法，用于利用时序 SAR 数据集进行空间滑坡预测[J]. *Gondwana Research*, 2023, 123:107-124. https://doi.org/10.1016/j.gr.2022.08.004
[19] SHERPA D.、ABHIJIT R. D.、MITRA I.、DHAR D.、SHARMA S.、CHAKRABORTY P. 和 CHAUDHURY K.
利用机器学习预测特发性复发性自然流产[C]// International Conference on Computer, Electrical & Communication Engineering (ICCECE), 印度加尔各答, 2023: 1-8. https://doi.org/10.1109/ICCECE51049.2023.10085363
[20] ALJAMEEL S. S.、ALJABRI M.、ASLAM N.、ALOMARI D. M.、ALYAHYA A.、ALFARIS S.、BLHARITH M.、ABAHUSSAIN H.、BOUJLEA D. 和 ALSULMI E. S.
基于机器学习的第 孕期流产早期预测自动化系统[J]. *Computers, Materials and Continua*, 2023, 75(1): 1291-1304. https://doi.org/10.32604/cmc.2023.035710
[21] SINGH S.、TIWARI S.、GOEL P. 和 TIWARI D.
关于威胁性流产的集成机器学习算法回顾性考察[C]// 6th International Conference on Information Systems and Computer Networks (ISCON), 印度马图拉, 2023: 1-7. https://doi.org/10.1109/ISCON57294.2023.10111961
[22] AMITAI T.、KAN-TOR Y.、OR Y.、SHOHAM Z.、SHOFARO Y.、RICHTER D.、HAR-VARDI I.、BEN-MEIR A.、SREBNIK N. 和 BUXBOIM A.
胚胎分类的延伸应用：利用机器学习早期预测第一孕期流产[J]. *Journal of Assisted Reproduction and Genetics*, 2023, 40(2): 309–322. https://doi.org/10.1007/s10815-022-02619-5
[23] BROWN M. G. L.、PETERSON M. G.、TEZAUR I. K.、PETERSON K. J. 和 BULL D. L.
随机森林回归特征重要性在气候影响路径检测中的应用[J]. *Journal of Computational and Applied Mathematics*, 2025, 464: 116479. https://doi.org/10.1016/j.cam.2024.116479
[24] HEIDARI M.、MAATAR M. H. 和 GHAFFARI H.
利用 Jensen–Shannon 和随机森林特征重要性排序在深度神经网络中的前向传播丢弃[J]. *Neural Networks*, 2023, 165: 238-247. https://doi.org/10.1016/j.neunet.2023.05.044
[25] THANATHAMATHEE P.、SAWANGARREERAK S. 和 NIZAM D. N. M.
利用 Anchor 可解释 AI 和注意力加权 XGBoost 提升持续经营预测[J]. *IEEE Access*, 2024, 12: 68345-68363. https://doi.org/10.1109/ACCESS.2024.3401007
[26] DELAUNAY J.、GALÁRRAGA L. 和 LARGOUËT C.
改进基于 Anchor 的解释方法[C]// International Conference on Information & Knowledge Management (CIKM), 美国纽约, 页码 3269-3272. https://doi.org/10.1145/3340531.3417461
[27] ZHANG H. 和 ZHANG J.
随机森林的高效且有效的反事实解释[J]. *Expert Systems with Applications*, 2025, 293: 128661. https://doi.org/10.1016/j.eswa.2025.128661
[28] KURTZ J.、BIRBIL Ş. Ī. 和 HERTOG D. D.

线性优化的反事实解释[J]. *European Journal of Operational Research*, 2025. https://doi.org/10.1016/j.ejor.2025.06.016
[29] ABEKOON T.、SAJINDRA H.、RATHNAYAKE N.、EKANAYAKE I. U.、JAYAKODY A. 和 RATHNAYAKE U. 利用可解释机器学习（SHAP 和 LIME）预测白菜栽培中土壤氮、磷、钾含量的新应用[J]. *Smart Agricultural Technology*, 2025, 11: 100879. https://doi.org/10.1016/j.atech.2025.100879
[30] HERMOSILLA P.、BERRÍOS S. 和 ALLENDE-CID H. 法医分析的可解释人工智能：SHAP 与 LIME 在入侵检测模型中的比较研究[J]. Applied Sciences, 2025, 15(13): 7329. https://doi.org/10.3390/app15137329

Word count: 7100 words, excluding references.

Peer-review record:
Fast-track status: Not fast-tracked
First-round reviews received: 3 reports
Revision cycles completed: 3 rounds
Final version submitted: November 9, 2025