International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 3 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

PRE-COMMITMENT PRIVACY RISK INTELLIGENCE: MACHINE-LEARNING-GUIDED ADAPTIVE PROTECTION FOR VERIFIABLE BLOCKCHAIN CREDENTIALS

Omkar Vinayak Bharose[1], Sandip M. Surve[2]

[1]*PG Scholar, M.Sc. Computer Science, Department of Computer Science, Shri Shivaji College, Parbhani- 431401, Maharashtra, India*

[2]*Assistant Professor, Department of Computer Science, Shri Shivaji College, Parbhani- 431401, Maharashtra, India*

Abstract

Blockchain-based credential verification provides durable integrity and decentralized validation, but the persistence of blockchain commitments can make premature disclosure of sensitive attributes difficult to reverse. This study developed a pre-commitment privacy risk intelligence framework that assessed potential information leakage before credential commitment and used the estimated risk to guide adaptive protection. Synthetic credential records and simulated attacker-side information were employed to characterize privacy exposure through uniqueness, re-identification susceptibility, cross-source linkage, attribute inference, and attribute-interaction effects. These indicators were transformed into predictive features and evaluated using machine-learning models for continuous privacy-risk estimation and categorical risk assessment, followed by an optimization stage that considered privacy reduction and retained utility. In the reported training experiment, the gradient-boosted regression model achieved a mean absolute error of 0.03290, root mean square error of 0.04645, and coefficient of determination of 0.84398. Logistic regression provided the strongest classification performance, attaining 85.63% accuracy, 81.33% recall, an F1-score of 82.44%, and a privacy false-negative rate of 18.67%. A subsequent evaluation of 1,100 records reduced the mean privacy-risk score from 0.81826 to 0.67337, corresponding to a 17.71% relative reduction, while maintaining a mean utility score of 0.84227. The findings indicated that pre-commitment privacy assessment could support risk-sensitive protection decisions while retaining substantial utility for verifiable credential processing.

Keywords: ***adaptive privacy protection; blockchain credential verification; privacy leakage assessment; pre-commitment risk analysis; privacy–utility optimization; Machine Learning***

1. INTRODUCTION

The growing adoption of digital credentials has increased the need for verification mechanisms that provide both authenticity and privacy. Blockchain technology offers tamper-resistant records and decentralized verification, while zero-knowledge techniques enable claims to be verified without necessarily revealing the underlying sensitive information. Recent studies have applied these technologies to educational credential verification, demonstrating their potential for privacy-preserving academic record management (Tran et al., 2024; Moya et al., 2025). Cross-chain credential verification has also been investigated using zero-knowledge proofs to reduce unnecessary disclosure during verification (Tran et al., 2024). Despite these developments, existing approaches largely concentrate on secure credential storage, selective disclosure, and privacy-preserving verification. Comparatively less attention has been given to estimating the privacy risk of credential information before it is committed to a persistent blockchain record. Privacy exposure may depend on attribute uniqueness, re-identification potential, linkage with external information, inference, and combinations of attributes. Therefore, a fixed protection strategy may not provide an appropriate balance between privacy and verification utility. This study addressed this gap by developing a pre-commitment privacy risk intelligence framework that quantified privacy-leakage characteristics, predicted disclosure risk using machine-learning models, and adaptively selected protection strategies according to estimated risk and retained utility. Cryptographic commitment and zero-knowledge verification were incorporated to support secure and verifiable credential processing. The work consequently investigated a risk-aware approach in which privacy

exposure was assessed before blockchain commitment, rather than being addressed only after information had been protected or recorded.

2. LITERATURE REVIEW

Recent research on digital credential security has concentrated on three major areas: blockchain-based credential management, privacy-preserving verification, and decentralized access-controlled credential systems. Blockchain-based approaches provide tamper-resistant records, traceability, and decentralized verification, making them suitable for academic and professional credential management (Rustemi et al., 2024; Al Hemairy et al., 2024). However, blockchain primarily addresses integrity and verifiability and does not inherently determine whether the information being committed creates an unacceptable privacy exposure.

A second research direction has focused on zero-knowledge and privacy-preserving credential verification. CrossCert demonstrated the use of zero-knowledge proofs for privacy-preserving educational credential verification across blockchain environments (Tran et al., 2024). Similarly, Moya et al. (2025) developed a blockchain-based academic record verification approach using zero-knowledge proofs, decentralized identifiers, and distributed storage. These studies demonstrate that credential claims can be verified without revealing complete underlying records, thereby reducing unnecessary disclosure.

Recent research has further incorporated decentralized access control and cross-chain interoperability into credential ecosystems. DAVE-CC, for example, investigated decentralized and access-controlled academic credential management across blockchain environments (Tran et al., 2025). Such approaches strengthen credential governance and controlled verification; however, they primarily determine how credentials are protected and verified, rather than how privacy risk should be assessed before a credential is committed to a persistent blockchain record. This limitation is significant because privacy exposure can depend on attribute uniqueness, re-identification potential, linkage with external datasets, inference, and interactions among multiple attributes. Consequently, applying a uniform protection mechanism may provide excessive protection for low-risk credentials while remaining insufficient for highly exposed records. Machine-learning-based risk assessment offers a potential solution by learning relationships between measurable leakage characteristics and privacy risk. However, risk prediction alone is insufficient because stronger privacy protection may reduce the information required for legitimate verification. The present study addresses this gap by introducing a pre-commitment privacy risk intelligence approach that connects privacy-leakage measurement with machine-learning-based risk prediction and adaptive protection. The predicted risk is used to guide protection decisions while considering retained verification utility, followed by cryptographic and blockchain-based verification. Thus, rather than treating privacy protection as a fixed property of the credential system, the proposed approach considers it as a risk-aware and utility-conscious decision made before blockchain commitment.

3. METHODOLOGY

The study adopted a quantitative experimental research design to evaluate a machine-learning-guided approach for privacy-risk assessment and adaptive protection of blockchain-based credentials. The proposed methodology consisted of four principal stages: synthetic data generation and privacy-leakage characterization, machine-learning-based risk prediction, adaptive protection, and cryptographic verification. The framework was designed to assess privacy exposure before blockchain commitment, thereby allowing protection decisions to be influenced by the estimated risk associated with individual credential records.

3.1 Data Collection

Real academic credentials were not used because they may contain personally identifiable and sensitive information. Instead, synthetic credential records were generated under controlled experimental conditions. Simulated attacker-side knowledge was introduced to represent potential external information that could be used to identify or infer characteristics of a credential holder. Privacy exposure was characterized using measurable indicators including uniqueness, re-identification susceptibility, linkage potential, inference potential, and attribute interactions. These indicators were transformed into structured numerical features and used to construct the training dataset. The reported experiments included a model-training dataset of approximately 1,000 records and a subsequent evaluation involving 1,100 records. The data-generation and preprocessing procedures were implemented using

Python and the project's reproducible command-line pipeline.

3.2 Data Analysis

Two predictive tasks were considered: continuous privacy-risk estimation and privacy-risk classification. Random Forest and XGBoost regression models were evaluated for continuous risk prediction, while Logistic Regression, Random Forest, and XGBoost classifiers were evaluated for risk classification. Regression models were assessed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Classification performance was evaluated using accuracy, precision, recall, F1-score, Receiver Operating Characteristic Area Under the Curve (ROC-AUC), Precision-Recall Area Under the Curve (PR-AUC), and Privacy False-Negative Rate (PFNR). PFNR was included because failing to identify a genuinely high-risk credential could result in insufficient privacy protection. Following prediction, the estimated risk was provided to the adaptive protection stage. Protection strategies were evaluated according to their ability to reduce privacy exposure while retaining useful information for credential verification. Privacy-risk reduction was calculated as:

$$\Delta R = R_{before} - R_{after}$$

and the relative reduction was calculated as:

$$PRR = \frac{R_{before} - R_{after}}{R_{before}} \times 100$$

where represents the mean risk before protection and represents the mean risk following protection. Mean utility was used as the complementary measure of information retained after protection.

3.3 Experimental Evaluation

The complete experimental workflow was implemented using Python-based machine-learning components together with the project's cryptographic, zero-knowledge, and blockchain modules. Model performance, privacy-risk reduction, retained utility, and execution time were considered as the principal evaluation measures. The study was primarily computational and experimental; therefore, the reported model metrics represented descriptive performance measures rather than statistical significance tests. Where inferential comparisons between repeated experimental runs are subsequently performed, a significance level of $\alpha = 0.05$ will be used. This methodology positioned machine learning as the risk-assessment layer, adaptive optimization as the decision layer, and cryptographic and blockchain mechanisms as the protection and verification layers.

4. ALGORITHMS

The proposed framework employed five supervised learning algorithms to evaluate privacy-risk prediction from complementary perspectives. Random Forest Regression and XGBoost Regression were used to estimate continuous privacy-risk scores, whereas Logistic Regression, Random Forest Classification, and XGBoost Classification were used to categorize records according to their predicted privacy-risk level. The comparative use of linear, bagging-based, and boosting-based approaches enabled the predictive behavior of different model families to be examined under the same experimental setting.

4.1 Random Forest Regression

Random Forest Regression is an ensemble learning method that combines predictions from multiple decision trees to estimate a continuous target variable. In the proposed framework, it was used to estimate the numerical privacy-risk score associated with a credential record from its extracted privacy-leakage characteristics. The ensemble structure enabled the model to capture nonlinear relationships between privacy indicators and risk while reducing the sensitivity of an individual decision tree to variations in the training data. Random Forest Regression was included as a robust tree-based baseline against which the boosting-based regression model could be compared.

4.2 XGBoost Regression

XGBoost Regression is a gradient-boosting approach in which decision trees are constructed sequentially, with subsequent trees focusing on errors produced by earlier learners. In this study, it was employed to estimate continuous privacy risk values from the engineered leakage features. The algorithm was selected because privacy exposure may arise from complex and nonlinear interactions among uniqueness, linkage, inference, re-identification, and attribute-combination characteristics. Its boosting mechanism provided an alternative to the independently constructed trees used by Random Forest Regression.

4.3 Logistic Regression

Logistic Regression is a statistical classification technique that estimates the probability of an observation belonging to a particular class. In the proposed framework, it was used to classify credential records according to their privacy-risk category. The model served as an interpretable baseline for

determining whether the relationship between the engineered privacy features and risk classification could be represented without a highly complex nonlinear model. Its probabilistic output was also useful for assessing classification performance at different decision thresholds.

4.4 Random Forest Classification

Random Forest Classification applies an ensemble of decision trees to assign observations to discrete classes. In this study, it was used to distinguish between privacy-risk categories based on the extracted leakage characteristics. The algorithm was appropriate for examining nonlinear relationships and interactions among privacy indicators without requiring explicit specification of those interactions. Its ensemble voting mechanism also provided a comparative tree-based approach to the Logistic Regression classifier.

4.5 XGBoost Classification

XGBoost Classification applies gradient-boosted decision trees to classification problems. Within the proposed framework, it was used to determine whether a credential record belonged to the defined privacy-risk categories. The sequential boosting process enabled the model to learn complex relationships among the privacy features and progressively reduce classification errors. Its inclusion provided a comparison with both the linear Logistic Regression model and the bagging-based Random Forest classifier.

5. RESULTS AND DISCUSSION

The proposed framework was evaluated with respect to continuous privacy-risk estimation, risk-state classification, and adaptive privacy protection. Rather than relying on a single performance indicator, the evaluation considered prediction error, discriminative capability, false-negative behavior, privacy-risk attenuation, and retained utility. This provided a more comprehensive assessment of the framework's effectiveness.

5.1 Continuous Privacy Risk Estimation

The regression results are reported in Table 1. XGBoost achieved the lowest MAE (0.0329) and RMSE (0.0464) and the highest R^2 (0.8440), compared with 0.0338, 0.0483, and 0.8312, respectively, for Random Forest. The results indicate that both models captured meaningful relationships between the engineered privacy indicators and the continuous risk score, with XGBoost providing a modest predictive advantage.

Model	MAE ↓	RMSE ↓	R^2 ↑
Random Forest	0.0338	0.0483	0.8312
XGBoost	0.0329	0.0464	0.8440

Table 1. Continuous Privacy-Risk Estimation Performance

The relatively close performance of the two models suggests that the underlying privacy-risk relationships were sufficiently learnable using both ensemble approaches. However, the higher explanatory performance of XGBoost makes it the stronger candidate for continuous risk estimation in the present experimental configuration.

5.2 Privacy Risk Classification

The classification results are illustrated in Figure 1. Logistic Regression achieved the highest accuracy (85.63%), recall (81.33%), and F1-score (82.44%), while XGBoost achieved the highest precision (85.52%).

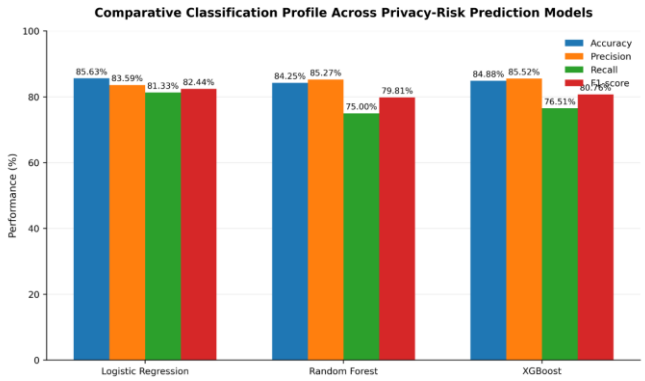

Figure 1. Comparative Classification Profile Across Privacy-Risk Prediction Models

As shown in Figure 1, the differences between the classifiers were relatively small, but Logistic Regression provided the most balanced performance across accuracy, recall, and F1-score. XGBoost offered slightly stronger precision, whereas Random Forest produced the lowest recall. In a privacy-sensitive application, this distinction is important because a model that fails to identify a high-risk case may result in insufficient protection.

The results also demonstrate that model complexity alone was not a reliable indicator of classification effectiveness. The comparatively simple Logistic Regression model performed better overall than the two tree-based classifiers under the evaluated configuration.

5.3 Privacy Sensitive Classification Analysis

Because overall accuracy does not fully describe the consequences of privacy-risk misclassification, ROC-AUC, PR-AUC, and Privacy False-Negative Rate (PFNR) were examined separately. The results are given in Table 2.

Model	ROC-AUC ↑	PR-AUC ↑	PFNR ↓
Logistic Regression	93.65%	92.42%	18.67%
Random Forest	92.74%	91.53%	25.00%
XGBoost	93.50%	91.80%	23.49%

Table 2. Privacy-Sensitive Classification and False-Negative Analysis.

Logistic Regression produced the highest ROC-AUC (93.65%) and PR-AUC (92.42%) and the lowest PFNR (18.67%). The PFNR result is particularly relevant to the proposed framework because false negative predictions represent potentially high-risk records that may not receive an appropriate protection level. Consequently, PFNR provides an application-specific perspective that complements conventional classification metrics.

5.4 Effect of Adaptive Privacy Protection

The principal outcome of the adaptive protection stage is illustrated in Figure 2. The mean privacy-risk score declined from 0.8183 before protection to 0.6734 after protection.

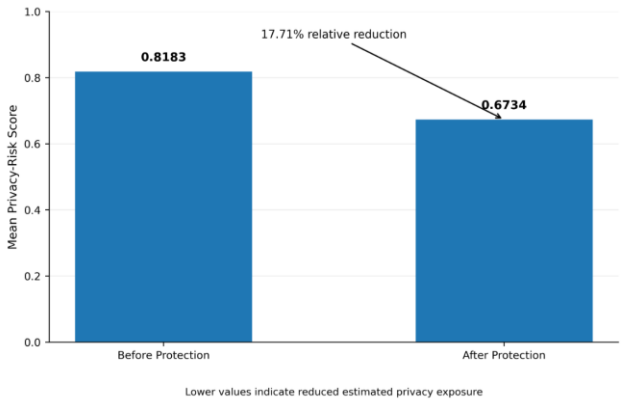

Figure 2. Risk-Aware Attenuation of Privacy Exposure Through Adaptive Protection

The observed decrease of 0.1449 corresponds to a relative privacy-risk reduction of approximately 17.71%. This reduction indicates that the adaptive mechanism was able to translate the estimated risk into a corresponding protection response. In contrast to a uniform protection strategy, the proposed approach was designed to make protection decisions according to the assessed exposure of the credential.

5.5 Privacy Utility Trade-off

The privacy-risk reduction must be considered together with the information retained after protection. Figure 3 presents the two principal outcomes using a common percentage scale.

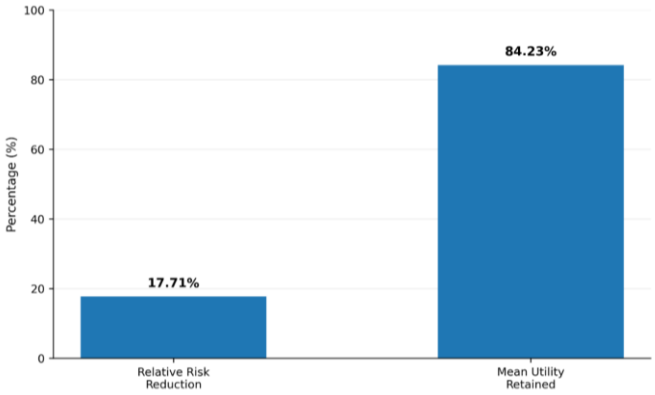

Figure 3. Privacy–Utility Equilibrium Under Risk-Aware Protection

The framework achieved a 17.71% relative reduction in mean privacy risk while retaining 84.23% mean utility. This result is important because privacy protection was not evaluated solely according to the amount of information removed. The retained utility indicates that the protection mechanism preserved a substantial level of information needed for legitimate credential processing. The result therefore supports the underlying privacy–utility balancing principle of the framework: stronger protection should be applied where exposure is greater, while unnecessary degradation of credential utility should be avoided.

6. CONCLUSIONS

This research established the viability of integrating machine learning, dynamic privacy safeguards, and blockchain-powered credential validation into a unified architecture. The findings indicate that privacy threats can be quantified effectively as both continuous numerical metrics and discrete risk categories. Among the evaluated algorithms, XGBoost demonstrated superior performance for continuous risk prediction, whereas Logistic Regression delivered the strongest overall classification results alongside the lowest false-negative rate for privacy detection. Through the adaptive protection experiments, the average privacy risk score dropped from 0.8183 to 0.6734—representing a 17.71% reduction—while sustaining a high average utility of 84.23%. These outcomes confirm that the proposed methodology successfully lowers estimated privacy exposure without compromising the practical usefulness of the credentials. The primary contribution of this work is the implementation of a risk-sensitive

decision-making layer prior to cryptographic encryption and blockchain ledger commitment. Instead of enforcing a rigid, uniform defense mechanism across all credentials, this strategy dynamically scales privacy protections based on calculated exposure levels. Nevertheless, because these findings were derived from controlled simulations and modeled threat scenarios, further validation is required. Subsequent work should incorporate real-world deployment data, replication trials, rigorous statistical significance testing, and expanded threat scenarios. Additionally, future research could explore supplementary privacy metrics and more advanced machine learning architectures to enhance the framework's overall resilience and scalability.

FUTURE SCOPE

Subsequent phases of this research will prioritize enhancing the precision, adaptability, and scalability of the proposed framework. Future investigations will integrate advanced machine learning and deep-learning architectures to refine predictive accuracy, alongside real-time risk evaluation and granular, dynamic protection scaling. Furthermore, optimizing the delicate equilibrium between privacy preservation and utility remains a key objective against increasingly sophisticated threat vectors. Additional development efforts will explore robust zero-knowledge verification protocols, enterprise-grade distributed ledger integration, and empirical validation across diverse operational environments to ensure the framework's long-term resilience and practical viability in next-generation digital credential ecosystems.

REFERENCES

[1] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel, "A survey on essential components of a self-sovereign identity," *Computer Science Review*, vol. 30, pp. 80–86, 2018, doi: 10.1016/j.cosrev.2018.10.002.

[2] Y. Liu, D. He, M. S. Obaidat, N. Kumar, M. K. Khan, and K.-K. R. Choo, "Blockchain-based identity management systems: A review," *Journal of Network and Computer Applications*, vol. 166, p. 102731, 2020, doi: 10.1016/j.jnca.2020.102731.

[3] Š. Čučko and M. Turkanović, "Decentralized and self-sovereign identity: Systematic mapping study," *IEEE Access*, vol. 9, pp. 139009–139027, 2021, doi: 10.1109/ACCESS.2021.3117588.

[4] L. Xu, T. Li, and Z. Erkin, "Verifiable credentials with privacy-preserving tamper-evident revocation mechanism," in *Proc. 2023 Fifth International Conference on Blockchain Computing and Applications (BCCA)*, pp. 266–273, 2023, doi: 10.1109/BCCA58897.2023.10338923.

[5] J. Zhang, J. Xu, *et al.*, "Privacy protection during the issuance and revocation of verifiable credentials in self-sovereign identity," *Concurrency and Computation: Practice and Experience*, 2025, doi: 10.1002/cpe.70084.

[6] R. Roio, R. Selvaggini, G. Bellini, and A. D'Intino, "SD-BLS: Privacy preserving selective disclosure of verifiable credentials with unlinkable threshold revocation," in *Proc. 2024 IEEE International Conference on Blockchain*, 2024, doi: 10.1109/Blockchain62396.2024.00074.

[7] R. Lavin, X. Liu, H. Mohanty, L. Norman, G. Zaarour, and B. Krishnamachari, "A survey on the applications of zero-knowledge proofs," *arXiv preprint arXiv:2408.00243*, 2024, doi: 10.48550/arXiv.2408.00243.

[8] H. Zheng, J. Chen, X. Du, K. H. Yeh, and Y. Xiang, "Enabling trust and learner agency in lifelong learning: A dual-chain, privacy-preserving credential architecture," *Journal of Information Security and Applications*, vol. 100, p. 104465, 2026, doi: 10.1016/j.jisa.2026.104465.