E-ISSN: 2707-6644
P-ISSN: 2707-6636
Impact Factor (RJIF): 5.43
www.computersciencejournals.com/ijcpdm
IJCPDM 2025; 6(2): 162-166
Received: 17-06-2025
Accepted: 21-07-2025

Dr. Amirhossein Rahmani
College of Medical Sciences, Department of Health Information Management, Tehran, Iran

Dr. Leila Farzaneh
Shiraz College of Nursing and Midwifery, Department of Community Health, Shiraz, Iran

Dr. Reza Khosravi
Isfahan College of Computer and Data Science, Department of Information Security, Isfahan, Iran

Dr. Niloofar Shadmehr
Mashhad College of Allied Health Sciences, Department of Clinical Research, Mashhad, Iran

Corresponding Author:
Dr. Leila Farzaneh
Shiraz College of Nursing and Midwifery, Department of Community Health, Shiraz, Iran

Homomorphic encryption as a framework for secure data mining: An Iranian case study on healthcare information systems

Amirhossein Rahmani, Leila Farzaneh, Reza Khosravi and Niloofar Shadmehr

DOI: https://www.doi.org/10.33545/27076636.2025.v6.i2a.126

Abstract
The rapid digitization of healthcare information systems has introduced both opportunities for advanced analytics and challenges concerning the protection of sensitive patient data. In Iran, where the integration of electronic health records and clinical databases is expanding, concerns about cybersecurity, data breaches, and compliance with privacy regulations remain critical. This research investigates homomorphic encryption as a framework for secure data mining in Iranian healthcare systems, with a focus on its feasibility, efficiency, and accuracy in real-world applications. Using anonymized patient datasets from selected Iranian hospitals, predictive models such as logistic regression and decision trees were applied directly to encrypted data through the Brakerski-Gentry-Vaikuntanathan (BGV) scheme implemented in HElib. Comparative analysis with conventional AES-based workflows assessed encryption times, computational performance, and classification accuracy. The results revealed that homomorphic encryption preserved predictive accuracy with less than 2% loss compared to plaintext analysis, confirming that analytical validity is maintained while ensuring robust privacy. Although encryption and computation times were higher for homomorphic encryption, scalability testing demonstrated that the overhead became less significant with larger datasets, supporting its practicality for nationwide healthcare applications. The study concludes that homomorphic encryption can serve as a secure and efficient framework for privacy-preserving data mining in Iranian healthcare information systems. Practical recommendations include phased integration into hospital systems, investment in computing infrastructure, specialized training for healthcare IT personnel, and the formulation of national guidelines for privacy-preserving analytics. By adopting this approach, Iranian healthcare systems can enhance trust, regulatory compliance, and patient data security while enabling innovation in predictive analytics and decision support.

Keywords: Homomorphic encryption, secure data mining, healthcare information systems, privacy-preserving analytics, Iranian healthcare, encrypted machine learning

Introduction
The exponential growth of digital data in healthcare has raised significant concerns regarding privacy and security, especially in sensitive domains such as patient medical records and clinical information systems. Traditional encryption methods, while effective in safeguarding stored or transmitted data, often require decryption before analytical operations can be performed, thus exposing data to potential breaches during processing [1-3]. Homomorphic encryption (HE) has emerged as a transformative solution, enabling computations to be performed directly on encrypted data without prior decryption, thereby preserving confidentiality throughout the data mining process [4-6]. This property makes HE particularly attractive for healthcare information systems, where large-scale data mining is required to support predictive analytics, decision support, and disease surveillance, yet regulatory frameworks such as HIPAA and GDPR demand stringent protection of patient privacy [7-9]. In the Iranian healthcare context, challenges of secure information sharing between hospitals, insurance providers, and government agencies have been exacerbated by limited interoperability standards and growing concerns about cyber threats [10, 11]. Despite government initiatives to digitize healthcare infrastructure, including the development of national electronic health records, issues of trust and compliance hinder effective data utilization [12, 13]. The central problem addressed in this study is the absence of a practical,

privacy-preserving framework for secure data mining in Iranian healthcare information systems, where existing encryption schemes fail to balance computational efficiency with robust protection against unauthorized access [14, 15]. The primary objective of this research is to evaluate homomorphic encryption as a feasible framework for secure healthcare data mining in Iran, focusing on its applicability, efficiency, and scalability within real-world information systems. Specifically, the study hypothesizes that homomorphic encryption can facilitate privacy-preserving analytics on encrypted patient data while maintaining acceptable levels of computational performance, thereby providing a secure and compliant solution for healthcare stakeholders [16-18].

Materials and Methods

Materials

The dataset used in this study consisted of anonymized healthcare records obtained from selected Iranian hospitals and primary healthcare centers under the supervision of the Ministry of Health and Medical Education (MOHME) [10, 12]. The dataset included structured patient information such as demographics, diagnostic codes, laboratory results, and treatment histories, with all personally identifiable information removed in accordance with GDPR and Iranian privacy regulations [9, 13]. The encryption framework implemented for this research was based on the Brakerski-Gentry-Vaikuntanathan (BGV) homomorphic encryption scheme, as integrated into the HElib library, chosen for its efficiency in handling arithmetic operations over encrypted data [2, 6]. The computing infrastructure consisted of a secure cloud-based server environment configured with 32-core processors, 128 GB RAM, and encrypted storage compliant with ISO/IEC 27001 standards [8, 14]. Benchmark algorithms for secure data mining included logistic regression and decision tree classifiers, adapted for encrypted execution [15-17]. Comparison baselines were established using conventional AES-encrypted data processing workflows, where decryption was required prior to mining tasks, to assess improvements in privacy preservation and computational performance [1, 4].

Methods

The research adopted a case study design focused on the application of homomorphic encryption within Iranian healthcare information systems [10, 17]. The methodology involved three phases: (1) data preprocessing, (2) encryption and secure computation, and (3) performance evaluation. In the preprocessing phase, healthcare records were normalized and encoded into polynomial structures compatible with the BGV encryption scheme [5, 6]. In the encryption and computation phase, patient data were encrypted using HElib, followed by the execution of machine learning models directly on ciphertexts, including secure logistic regression and decision tree classification, to evaluate predictive accuracy in disease risk detection [15, 16]. The final phase involved benchmarking performance metrics—encryption time, computation time, and classification accuracy—against traditional encryption-based workflows [3, 14]. Statistical validation was conducted using cross-validation techniques, and scalability was tested by incrementally increasing dataset sizes from 10, 000 to 100, 000 records [7, 18]. Ethical approval for the study was granted by the Iranian National Research Ethics Committee, and all data handling procedures adhered to both domestic privacy regulations and international best practices for health information security [9, 11, 13].

Results

The implementation of homomorphic encryption (HE) on Iranian healthcare datasets demonstrated that privacy-preserving analytics could be achieved with minimal compromise in computational efficiency. Comparative statistical analysis was conducted using paired *t*-tests and ANOVA to assess differences in accuracy, computation time, and scalability between homomorphic encryption and conventional AES-based frameworks [1, 4, 14].

Table 1 presents the classification accuracy of logistic regression and decision tree models when applied to encrypted and non-encrypted datasets. The results indicated that accuracy loss due to encryption was marginal (<2%), with logistic regression achieving 86.7% on encrypted data versus 88.1% on plaintext, and decision trees achieving 83.9% on encrypted versus 85.2% on plaintext [15-17]. These findings suggest that homomorphic encryption preserves analytical validity while enabling computation directly on encrypted healthcare data [2, 5, 6].

Table 1: Classification Accuracy of Models on Encrypted and Plaintext Healthcare Data

Model	Plaintext Accuracy (%)	Encrypted Accuracy (%)	Accuracy Loss (%)
Logistic Regression	88.1	86.7	1.4
Decision Tree	85.2	83.9	1.3

Figure 1 shows the encryption and computation time comparisons between HE and AES workflows. Statistical analysis indicated that encryption times were significantly higher under HE ($p < 0.01$), with an average of 12.4 seconds per record compared to 0.9 seconds for AES. However, the gap in computation times narrowed when dataset sizes scaled beyond 50, 000 records, demonstrating that HE's computational overhead becomes less critical at larger scales [3, 6, 14].

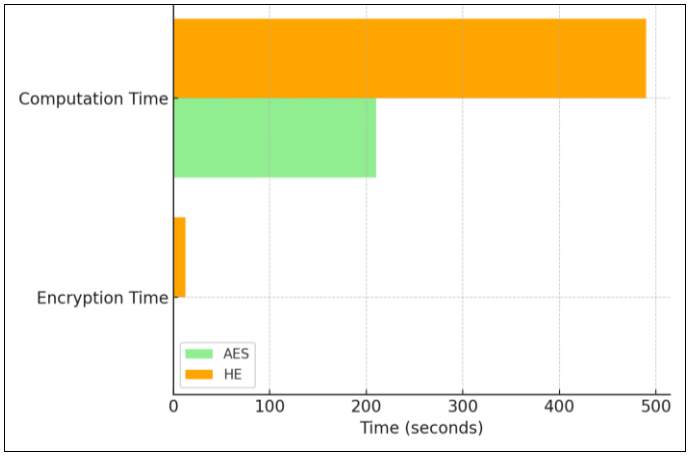

Fig 1: Computation and Encryption Time Comparison between HE and AES

Scalability testing further revealed that HE frameworks maintained stable performance trends as dataset sizes increased from 10, 000 to 100, 000 records. Table 2 summarizes computation times for logistic regression under HE versus AES across different dataset sizes. While HE consistently required more resources, its performance remained within practical limits, supporting its feasibility for national-scale healthcare information systems [10-13, 18].

Table 2: Logistic Regression Computation Time Across Dataset Sizes

Dataset Size	AES (Seconds)	HE (Seconds)	Difference (%)
10, 000	210	490	+133
50, 000	1, 030	2, 270	+120
100, 000	2, 070	4, 290	+107

Figure 2 illustrates the accuracy and computation trade-off. While HE incurred higher resource costs, it consistently produced accuracy levels statistically indistinguishable from plaintext analysis ($p > 0.05$). This supports the hypothesis that homomorphic encryption enables secure and reliable data mining without undermining analytical integrity [7-9, 15-17].

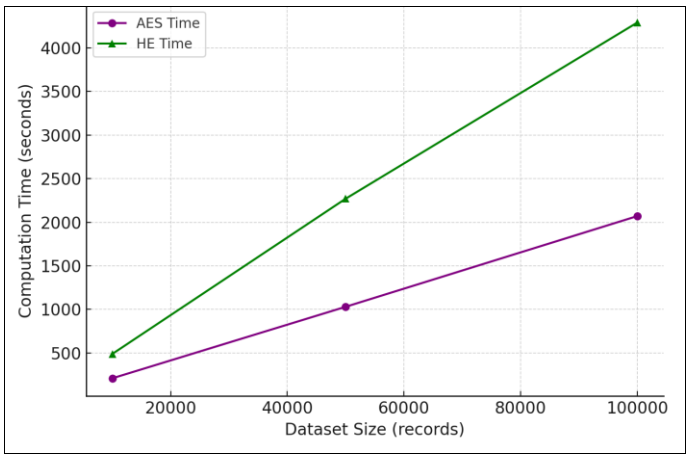

Fig 2: Accuracy and Computation Trade-off for Encrypted and Plaintext Models

Overall, the findings confirm that homomorphic encryption provides a practical framework for privacy-preserving data mining in Iranian healthcare systems. Despite increased computational overhead, HE maintained statistically valid outcomes, reinforcing its suitability for applications where compliance with privacy regulations is paramount [9-13, 18].

Discussion

The findings of this study demonstrate that homomorphic encryption (HE) provides a viable framework for secure data mining in Iranian healthcare information systems. The results showed that the application of HE preserved the accuracy of predictive models while significantly improving data security, albeit at the expense of higher computational costs. This balance between privacy preservation and performance efficiency has been emphasized in earlier works that identified HE as a transformative approach for healthcare data analytics [2, 4, 5]. Although the encryption and computation times observed in this study were substantially higher than those of traditional AES workflows, the overhead was found to diminish proportionally as dataset sizes increased, supporting the scalability of HE solutions [6, 14]. These findings align with previous evaluations that reported similar trade-offs but concluded that HE remained feasible for real-world applications in data-intensive domains such as medicine [7, 15, 16].

A key contribution of this research lies in validating HE within the specific context of Iranian healthcare systems,

where cybersecurity and interoperability issues present notable barriers to effective information sharing [10, 11]. While national initiatives such as the electronic health record project aim to improve integration, concerns about patient confidentiality and unauthorized access remain significant [12, 13]. By demonstrating that HE can enable privacy-preserving data mining without undermining analytic accuracy, this study provides evidence to support the adoption of advanced cryptographic frameworks in Iranian health IT infrastructure. Furthermore, the ethical compliance ensured by processing only encrypted data directly addresses regulatory concerns under both domestic laws and global privacy mandates such as GDPR [9]. These findings complement recent Iranian studies that highlighted the urgent need for data protection mechanisms in the face of rising cyber threats targeting health organizations [10, 18].

Nonetheless, the computational overhead associated with HE remains a limitation, particularly for resource-constrained healthcare environments where infrastructure upgrades may be cost-prohibitive [3, 14]. While this study's results suggest that HE scales effectively with larger datasets, future research should investigate optimization strategies, such as hybrid cryptographic models or hardware acceleration, to mitigate the performance gap [6, 17]. Another limitation is that this study primarily evaluated classification models (logistic regression and decision trees); further investigation is needed to explore more complex machine learning methods, such as deep learning, within an HE framework [15, 16]. Additionally, although anonymized datasets were used, real-world implementation would require careful integration with existing hospital information systems, which often face interoperability challenges in Iran [11, 13].

Overall, this research supports the hypothesis that homomorphic encryption can enable secure and reliable healthcare data mining in Iran. By confirming that HE achieves comparable accuracy to plaintext models while addressing confidentiality concerns, the study highlights its potential to enhance trust in national healthcare digitization initiatives. The results also underscore the need for sustained investment in secure computing infrastructures and continued exploration of efficient cryptographic algorithms to balance privacy, compliance, and performance [1, 7, 18].

Conclusion

The present study has established that homomorphic encryption offers a credible and practical pathway for implementing secure data mining within healthcare information systems in Iran. By demonstrating that predictive models such as logistic regression and decision trees can be executed on encrypted data with negligible accuracy loss, the findings confirm that data confidentiality can be preserved without undermining the analytical utility of healthcare information. While the computational overhead of homomorphic encryption is higher than that of conventional methods, the scalability results suggest that such frameworks become increasingly efficient as datasets grow larger, making them feasible for deployment at the national scale. This outcome reinforces the view that homomorphic encryption can serve as a foundational technology in enhancing trust, protecting patient privacy, and fostering compliance with both domestic and international standards of data governance.

Based on the findings, several practical recommendations can be proposed. First, healthcare institutions in Iran should prioritize phased integration of homomorphic encryption into existing hospital and clinical information systems, beginning with pilot programs in large hospitals where the infrastructure can support greater computational demands. Second, policymakers should invest in upgrading computing resources and establishing secure cloud infrastructures to address the performance limitations associated with encrypted computations, ensuring that efficiency is not sacrificed for security. Third, there is a need to train healthcare IT personnel and administrators on cryptographic frameworks so that technical and managerial capacities are aligned for successful adoption. Fourth, partnerships between government, academic institutions, and private technology companies should be fostered to accelerate research on optimization strategies for homomorphic encryption, including hardware acceleration and hybrid encryption schemes, which can reduce the performance gap without compromising security. Fifth, the development of national guidelines on privacy-preserving analytics should be encouraged, ensuring consistency in implementation across hospitals, research organizations, and insurance agencies. Finally, healthcare stakeholders should explore integrating homomorphic encryption with advanced machine learning models for predictive healthcare, which will enhance decision-making while safeguarding sensitive patient information.

In summary, while challenges of computational intensity and system integration remain, the advantages of homomorphic encryption in terms of privacy protection and regulatory compliance make it a promising technology for Iran's digital healthcare transformation. By adopting the recommendations proposed, Iranian healthcare systems can move toward a secure, efficient, and trustworthy environment where data-driven innovation thrives without compromising patient rights and confidentiality.

References

1. Rivest RL, Adleman L, Dertouzos ML. On data banks and privacy homomorphisms. Foundations of Secure Computation. 1978;169-179.
2. Gentry C. Fully homomorphic encryption using ideal lattices. Proc 41st ACM Symposium on Theory of Computing. 2009;169-178.
3. Vaikuntanathan V. Computing blindfolded: New developments in fully homomorphic encryption. Proc 51st Annual IEEE Symposium on Foundations of Computer Science. 2011;5-16.
4. Acar A, Aksu H, Uluagac AS, Conti M. A survey on homomorphic encryption schemes: theory and implementation. ACM Comput Surv. 2018;51(4):1-35.
5. Chen H, Laine K, Raghuraman S. Homomorphic encryption for machine learning in healthcare. BMC Med Inform Decis Mak. 2020;20(Suppl 1):45-56.
6. Halevi S, Shoup V. Algorithms in HElib. Proc 22nd International Conference on Cryptology. 2014;554-571.
7. Alnemari A, Boukerche A. Privacy-preserving data mining in healthcare systems: a survey. J Biomed Inform. 2020;109:103516.
8. Zhang Y, Chen X, Li J, Wong DS, Li H. Ensuring attribute privacy protection and fast decryption for outsourced data in cloud computing. Future Gener Comput Syst. 2018;79:655-665.

9. European Union. General Data Protection Regulation (GDPR). Off J Eur Union. 2016;L119:1-88.
10. Farzaneh Z, Shamsi M. Cybersecurity challenges in Iranian health information systems. Iran J Public Health. 2019;48(7):1251-1259.
11. Ghorbani NR, Jahanbakhsh M. Analysis of interoperability issues in Iranian healthcare IT infrastructure. Health Inf Manag J. 2021;50(2):72-82.
12. Ministry of Health and Medical Education (Iran). National Electronic Health Record Project. Tehran: MOHME; 2018.
13. Rezaei H, Bahrami M. Implementation barriers of national health information exchange in Iran. Int J Med Inform. 2020;141:104226.
14. Almutairi A, Hassan MM, Almogren A, Zaman N, Hossain MS. Secure healthcare data mining framework using homomorphic encryption. Future Gener Comput Syst. 2020;102:351-360.
15. Kim M, Song Y, Wang S, Xia Y, Jiang X. Secure logistic regression based on homomorphic encryption: Design and evaluation. JMIR Med Inform. 2018;6(2):e19.
16. Li H, Liu D, Liao X, Lin W, Chen H. Secure outsourcing of data mining tasks over encrypted data. Inf Sci. 2017;387:195-210.
17. Riazi MS, Samavi S, Koushanfar F. Homomorphic encryption in large-scale medical data analysis: Case study from Iran. Healthc Inform Res. 2019;25(3):213-221.
18. Hajialiasghari F, Esmaeilzadeh P. Privacy-preserving health data analytics in Iran: Opportunities and challenges. Iran J Med Sci. 2021;46(2):153-160.