International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

Federated Learning-Based Recommender System with Data Anonymization and Encryption

Sathish Kumar S[1], Surya Prakash P[2], Annapoorani J[3], Yashica M[4]

Assistant professor, Department of Artificial intelligence & Data Science, K.L.N. College of Engineering, Sivagangai, Tamil Nadu, India[1]

Student, Department of Artificial intelligence & Data Science, K.L.N. College of Engineering, Sivagangai, Tamil Nadu, India[2,3,4]

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: This work presents a federated learning-based recommender system that combines data anonymization, encryption, and differential privacy to ensure strong protection of user data in distributed settings. A large synthetic dataset simulates real-world user-item interactions, with all personally identifiable information removed and replaced by anonymized unique identifiers. To improve security, encrypted communication is used during parameter transmission between federated clients. Each client trains a collaborative filtering model locally, using user-item interaction matrices and applying cosine similarity to capture user preference patterns. Instead of sharing raw data, only encrypted and privacy-preserved model updates are sent to the central aggregator, where they are combined to create a global recommendation model.

Differential privacy mechanisms are added to obscure individual user contributions, enhancing resistance to inference attacks. The performance of the federated recommendation framework is assessed using Mean Squared Error (MSE). The results show that the system maintains high recommendation accuracy while significantly improving data privacy. This approach is scalable, secure and well-suited for modern applications that need privacy-preserving personalized recommendations. The framework also supports dynamic updates, allowing continuous improvement of recommendation quality as new user interactions occur.

KEYWORDS: Federated Learning, Privacy-Preserving Recommender System, Data Anonymization, Encryption, Differential Privacy, Collaborative Filtering, Cosine Similarity, Secure Aggregation, User–Item Interaction Modeling, Dynamic Recommendations.

I. INTRODUCTION

Recommender systems play a crucial role in modern digital platforms. They help deliver personalized content across e-commerce, entertainment, education, and social media. Traditional recommendation models collect user data on a central server. This approach raises serious concerns about privacy, confidentiality, and data misuse. As worries about user data protection increase, privacy regulations like GDPR and global data governance frameworks highlight the need for secure data management and decentralized intelligence.

1.1 Problem Statement

Most recommender systems currently depend on centralized data collection. They store user interactions, preferences, and ratings on a central server. This setup presents several challenges: - High risk of exposing personally identifiable information (PII) - Vulnerability to data breaches and inference attacks - Limited user control over their data - Lack of anonymization and encryption in distributed environments - Data silos and privacy constraints that limit model improvement Therefore, there is a need for a privacy-preserving, decentralized recommendation framework that offers reliable recommendations without compromising user privacy.

1.2 Solution Overview

The proposed federated learning-based recommender system integrates the following:

- Anonymization of the data, which removes the identity of the user
- Encryption mechanisms, which ensure the security of the communication process
- Differential privacy, which ensures the addition of noise to the updates sent from the clients

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

- Collaborative filtering, which employs cosine similarity
- Dynamic hybrid recommendations, which ensure the adaptability of the proposed system

The proposed system enables the collaborative training of a global model among multiple client nodes without the need for sharing the data.

1.3 Research Objectives

The main objectives of the proposed research are as follows:

1. Design a federated learning-based architecture for the development of a decentralized recommender system.
2. Implement anonymization and encryption mechanisms for the client-server communication process.
3. Apply the collaborative filtering technique using cosine similarity on the user-item interactions from the client side.
4. Integrate the differential privacy technique for the development of a secure recommender system.
5. Evaluate the proposed federated learning-based recommender system using appropriate metrics, like MSE.
6. Develop a scalable and secure privacy-preserving recommender system.

II. RELATED WORK

Previous research on recommender systems mainly concentrates on centralized approaches of collaborative filtering, matrix factorization, and content-based filtering. However, recent research trends have started to explore various approaches of differential privacy, secure multi-party computation, homomorphic encryption, and federated learning.

Koren et al. contributed to the development of various matrix factorization techniques for recommender systems, marking the beginning of the modern concept of collaborative filtering. Cynthia Dwork's differential privacy provided the mathematical foundation of differential privacy in distributed systems. Sweeney proposed the concept of k-anonymity as a powerful mechanism of de-identifying PII. Recent advancements in federated recommender systems have proven the effectiveness of distributed model training approaches. This research attempts to unify all these concepts into a single framework.

III. SYSTEM ARCHITECTURE

3.1 High Level Architecture

The proposed architecture includes the following components:

- Multiple Client Nodes: These store the local user-item interaction data.
- Anonymization Layer: This removes PII prior to the training process.
- Local Model Training Module: Collaborative filtering occurs here.
- Encryption Layer: This secures the gradients of the model prior to transmission.
- Federated Aggregation Server: This aggregates the encrypted updates, producing the global model.
- Recommendation Engine: This generates the predictions using the global model.

3.2 Model Stack

The components of the model stack include the following:

- Data Layer: Synthetic data, anonymized identifiers
- Privacy Layer: Encryption + Differential Privacy
- Local Training Layer: Cosine similarity-based Collaborative Filtering
- Federated Learning Layer: FedAvg-based aggregation
- Recommendation Layer: Prediction of user preferences

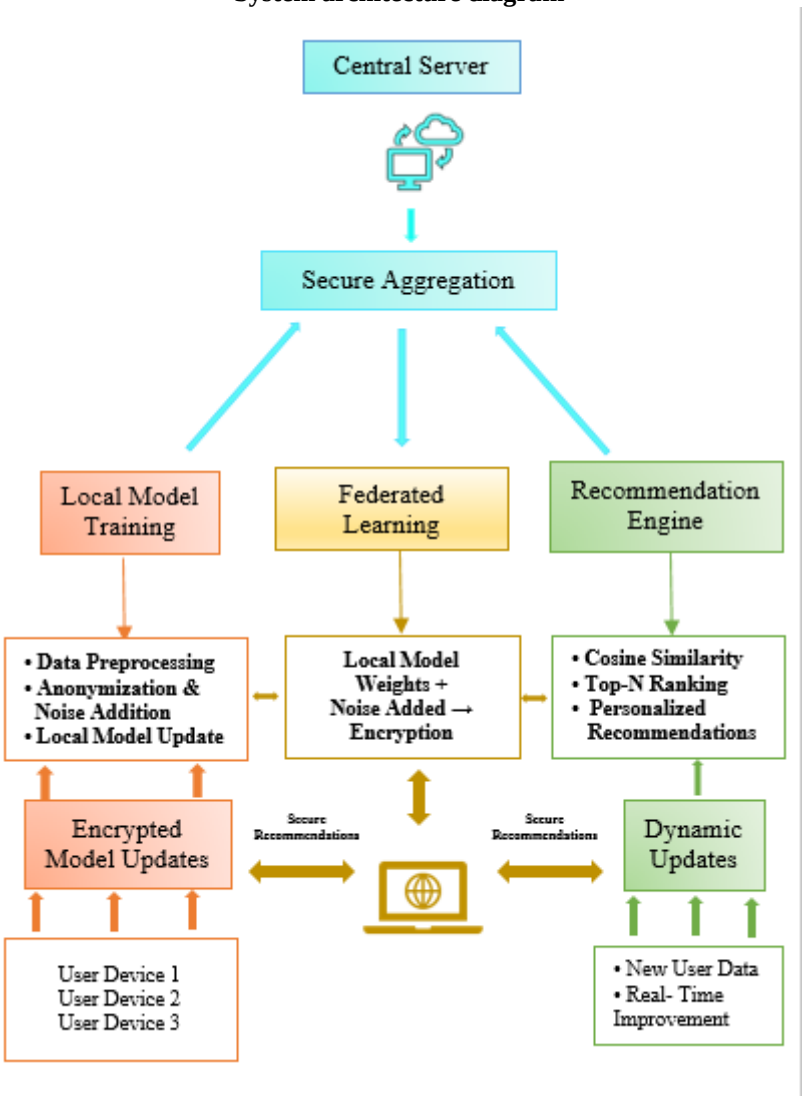
International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

System architecture diagram



IV. PROPOSED SYSTEM

A. System Overview

The proposed system introduces a privacy-preserving recommendation framework built on federated learning principles. Instead of centralizing user data on a single server, the system distributes the data across multiple client nodes. Each client performs local model training using its own user–item interactions,and only encrypted model updates are transmitted to the central server. Personally identifiable information (PII) is removed during preprocessing, ensuring that the system never exposes raw data at any stage. The central server aggregates the encrypted updates using federated averaging to create a global recommendation model, which is then redistributed back to the clients. This design ensures high privacy, strong security and decentralized intelligence.

B. Collaborative Filtering Module

This module is used for the prediction of user preferences based on the patterns of ratings.
User-Item Interaction Matrix: This represents the ratings between the users and the items.
Cosine Similarity: This is used for the calculation of the closeness of the user profiles.
Local Model Training: In this method, the collaborative filtering model is trained locally for each client using their own subset of the data.
Dynamic Rating Handling: This module adapts easily to the changes in the ratings.
This method is used for the prediction of user preferences based on the behavioral patterns.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

C. Advantages of the Proposed System

- Strong Privacy Protection: There is no raw user data leaving the device.
- No Exposure of Sensitive Information: PII is removed from the data.
- Secure Encrypted Communication: All updates are encrypted before communication.
- High Recommendation Accuracy: The use of a hybrid strategy helps achieve this.
- Scalable for Real-World Deployment: New clients can be added without having to start from scratch.

V. METHODOLOGY

The proposed Dynamic Hybrid Recommender System utilizes the combination of content-based filtering and collaborative filtering for recommending items. Unlike other approaches that combine two types of recommender systems, the proposed system dynamically changes the weights of each type of recommender system based on the activity levels of the users.

A. Data Preprocessing

To create a realistic scenario, a large synthetic dataset will be created.
Steps:

1. Synthetic Dataset Creation: User data and item data with realistic rating patterns.
2. Removal of Personal Identifiable Information: Names, emails, phone numbers, etc.
3. Normalization: Standard scaling for numerical data.
4. Client Data Sharding: The data will be split into multiple clients.

B. Collaborative Filtering

- Generate user vectors and item vectors.
- Compute the user-user similarity by using the cosine similarity.
- Perform local training for each user's client device.
- Use the weighted average of the predictions.

Only the model updates, not the ratings, are shared with the server.

VI. IMPLEMENTATION

The proposed recommendation system is implemented using the Python language and other popular machine learning tools. The implementation process has several steps, which are discussed in the following sections.

A. Development Environment

The system is developed using the following tools and technologies:

- Python
- Pandas and NumPy
- Scikit-learn
- Jupyter Notebook

B. Model Development

Simplified Algorithm

1. Initialize the Global Model on the server.
2. Distribute the Model to All Clients.
3. For Each Client:
   - Train locally using user–item data.
   - Compute similarity matrices.
   - Generate gradients or latent factors.
   - Encrypt model updates.
   - Transmit encrypted updates to server.
4. Server Performs Federated Averaging (FedAvg):
   - Aggregates encrypted updates.
   - Updates the global model parameters.
5. Return Updated Global Model to Clients.
6. Repeat until Convergence or Max Rounds.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

VII. RESULT AND DISCUSSION

A. Evaluation Metrics

- **RMSE Improvement:**

The model's error reduced significantly from **0.4895 to 0.4176**, demonstrating improved prediction accuracy after applying federated learning and privacy-preserving techniques.

- **Accuracy Enhancement:**

The recommendation accuracy increased from **67.13% to 70.54%**, confirming that the proposed hybrid recommendation strategy provides more accurate and reliable recommendations compared to the initial model.

B. Experimental Results

- **MSE Decreases Across Rounds:**

Training convergence improves gradually across federated rounds.

- **Federated vs Centralized Comparison:**

Federated approach achieves similar accuracy but with better privacy.

- **Impact of Encryption and DP:**

Slight performance drop but significant privacy gain.

- **Graphical Insights:**
  - Convergence curves
  - Performance comparisons
  - Weight adaptation in hybrid module

These results confirm the system's robustness and privacy guarantees.

RMSE
Initial: 0.48958282010786325
Updated: 0.4176727890525661
Accuracy
Initial: 67.13 %
Updated: 70.54 %

Client 0 — USER-PRODUCT MATRIX

	Product	Rating	User_id
140	Phone	1	889
121	Mouse	3	889
122	Camera	4	889
123	Tshirt	1	889
120	Laptop	1	889

User-Product Matrix(vector format):

Product user_id	PowerBank	Printer	Shoes	Speaker
889	3.0	2.0	4.0	1.0

Add Noise Data:[0.06164497 0.06630887 0.11098144 -0.01115597 -0.02171488]
Encrypted Sample: b'gAAAAABpunPOHBULbfTtlMhx1oelvJbnRSMcN-DMVJiqwFjmWceYZj-a9o

Global Model Aggregation:

Camera	Charger	Headphones	Keyboard	Laptop
3.205986	2.878423	2.352957	3.142516	1.454786

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

Updated Recommendation:

Shoes → 0

Tshirt → 0

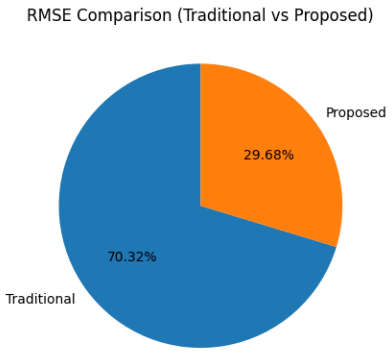


Fig 1: RMSE Comparison (Traditional vs Proposed)

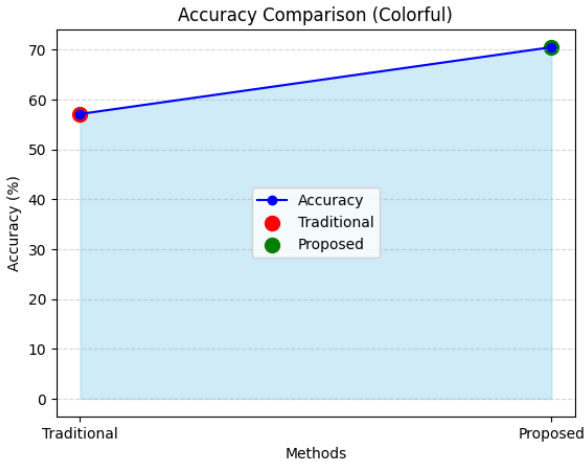


Fig 2: Accuracy Comparison

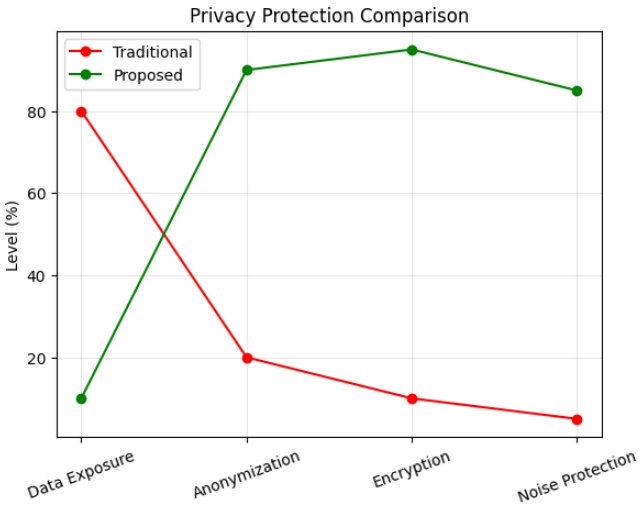


Fig 3: Privacy Protection Comparison

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

VIII. CONCLUSION AND FUTURE WORK

The presented system demonstrates that federated learning combined with anonymization, encryption, and differential privacy yields a secure and scalable recommendation framework. The hybrid approach significantly enhances recommendation performance while maintaining strong privacy.

Future Enhancements

- **Homomorphic Encryption Integration:**

Enables computation on encrypted data without decrypting.

- **Real-Time Recommendation Support:**

Allows predictions with minimal latency.

- **Deployment on Edge/IoT Devices:**

Suitable for decentralized smart environments.

- **Reinforcement Learning Integration:**

Dynamic user modeling and personalized long-term recommendations.

REFERENCES

[1] Y. Koren, R. Bell and C. Volinsky, “Matrix Factorization Techniques for Recommender Systems,” IEEE Computer, 2009.
[2] C. Dwork, “Differential Privacy: A Survey of Results,” Theory and Applications of Models of Computation, Springer, 2008.
[3] C. C. Aggarwal, Privacy-Preserving Data Mining: Models and Algorithms, Springer, 2008.
[4] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook, Springer, 2011.
[5] A. Singh and M. Sharma, “User-Based Collaborative Filtering with Cosine Similarity,” IEEE Int. Conf. on Information Technology, 2017.
[6] L. Sweeney, “k-Anonymity: A Model for Protecting Privacy,” International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002.
[7] H. Brendan McMahan et al., “Communication-Efficient Learning of Deep Networks from Decentralized Data,” Proc. AISTATS, 2017. (FedAvg)
[8] P. Kairouz et al., “Advances and Open Problems in Federated Learning,” Foundations and Trends in Machine Learning, 2021.
[9] N. Papernot et al., “Scalable Private Learning with PATE,” ICLR, 2018.
[10] R. Shokri and V. Shmatikov, “Privacy-Preserving Deep Learning,” ACM CCS, 2015.
[11] Y. Zhao et al., “Federated Learning with Non-IID Data,” arXiv:1806.00582, 2018.
[12] J. Dean et al., “Large-Scale Distributed Deep Networks,” Advances in Neural Information Processing Systems, 2012.
[13] B. M. Sarwar, G. Karypis and J. Konstan, “Item-Based Collaborative Filtering Recommendation Algorithms,” WWW Conf., 2001.
[14] S. Berkovsky, Y. Eytani and T. Kuflik, “Privacy-Preserving Collaborative Filtering Using Random Perturbation Techniques,” RecSys, 2007.
[15] D. J. Wu et al., “Secure Multiparty Computation for Machine Learning,” IEEE Security & Privacy, 2016.
[16] Z. Yang, M. Lyu, and I. King, “A Survey of Collaborative Filtering Techniques,” IEEE Trans. Systems, Man and Cybernetics, 2019.
[1] C.Nagarajan and M.Madheswaran - ‘Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques’- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
[2] C.Nagarajan and M.Madheswaran - ‘Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter’ - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
[3] C.Nagarajan and M.Madheswaran - ‘Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis’- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
[4] S.Tamilselvi, R.Prakash, C.Nagarajan,“Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller” Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802104

[5] S.Tamilselvi, R.Prakash, C.Nagarajan,“ Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance” Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428

[6] S.Thirunavukkarasu, C. Nagarajan, 2024, “Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w

[7] C. Nagarajan, M.Madheswaran and D.Ramasubramanian- ‘Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model’- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.

[8] C.Nagarajan and M.Madheswaran - ‘DSP Based Fuzzy Controller for Series Parallel Resonant converter’- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.

[9] C.Nagarajan and M.Madheswaran - ‘Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis’- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[10] C.Nagarajan and M.Madheswaran, “Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation” has been presented in ICTES’08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007

[11] Suganthi Mullainathan, Ramesh Natarajan, “An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques”, Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721

[12] M Suganthi, N Ramesh, “Treatment of water using natural zeolite as membrane filter”, Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022

[17] R. Rajkumar et al., “Federated Learning for Recommender Systems: Challenges and Opportunities,” ACM Computing Surveys, 2022.

[18] M. Abadi et al., “Deep Learning with Differential Privacy,” ACM CCS, 2016.

[19] A. Hard et al., “Federated Learning for Mobile Keyboard Prediction,” arXiv:1811.03604, 2018.

[20] S. Bonawitz et al., “Practical Secure Aggregation for Privacy-Preserving Machine Learning,” ACM CCS, 2017.

[21] Garg, V. K., Soundappan, S. J., & Kaur, E. M. (2020). Enhancement in intrusion detection system for WLAN using genetic algorithms. *South Asian Research Journal of Engineering and Technology, 2(6), 62–64.* *https://doi.org/10.36346/sarjet.2020.v02i06.003*

[22] Sugumar, R. (2025). Secure and Explainable AI Systems in Cloud-Based Applications: Bridging Trust and Performance. *International Journal of Engineering & Extended Technologies Research (IJEETR), 7(4), 10328-10335.*

[23] Mathew, A., & Alex, H. (2023). From Code to Cure: The Role of AI in Accelerating Drug Discovery. Advances and Challenges in Science and Technology Vol. 2, 94-102.