Privacy Preservation Using Machine Learning

A Synopsis Report

For

B.Tech Project-II (CO-402)

Abhishek Panda 2K21/CO/19, **Pranay Avnish Kachhap** 2K21/CO/338, **Pritesh Das** 2K21/CO/348

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Delhi Technological University

(Formerly, Delhi College

of Engineering) Bawana

Road, Delhi-110042

UNDER THE SUPERVISION OF:

DR. NIPUN BANSAL

ACKNOWLEDGMENT

- We would like to express our heartfelt gratitude to our respected Mentor **Dr. Nipun Bansal** and our respected Supervisor **Dr. Rajeev Kumar** for their invaluable guidance and support throughout our project. Their expertise and insights have been instrumental in shaping our understanding of the complexities involved in intrusion detection systems. Their constructive feedback and encouragement motivated us to explore innovative approaches and refine our methodologies.

- Thank you for being a constant source of inspiration and for dedicating your time and resources to our learning. We are truly grateful for your mentorship, which has significantly enriched our project experience.

- We, the members of the project team, would like to acknowledge our collective efforts and contributions to the successful completion of our project, **"Privacy Preservation Using Machine Learning"**

    - **Abhishek Panda 2K21/CO/19**
    - **Pranay Avnish Kachhap 2K21/CO/338**
    - **Pritesh Das 2K21/CO/348**

Signature of Dr. Nipun Bansal

TABLE OF CONTENTS

INTRODUCTION

The Digital Personal Data Protection Act (DPDPA), 2023, is India's first comprehensive data protection law, passed on August 11, 2023, aiming to safeguard personal data through guidelines for processing, privacy rights, data integrity, and data transfers, with a focus on responsible handling, especially for vulnerable groups.

The Swedish Personal Data Act (1998:204) was established in 1998 and had the purpose of protecting personal integrity. The purpose of the law was to protect people so that their integrity is not violated when processing personal data. For 20 years, the same legislation applied despite changes in the internet and digitalization. In 2018, the law was repealed by (2018:218) The Act on Supplementary Provisions to the EU Data Protection Regulation (GDPR). The law was introduced so that the user himself would have the opportunity to control and regulate the disclosure of his personal information. GDPR began to set new requirements where companies needed to inform the user about how they collected and processed personal information. As a result, there have been an increasing number of notifications and consent processes from websites where the user can see how their personal information will be processed.

The 90s have had the privilege of being part of this changing marketing world from the beginning, which has meant that they have developed a natural habit and understanding of technology and its use in marketing contexts. Digitalization has dramatically changed the existence of marketing. Through digitalization, it has enabled personalized marketing through the customer's own behavioral patterns. A study shows that digital assistants such as Siri and Alexa bring high customer satisfaction where apps and artificial intelligence are interconnected. Digital assistants are speech-activated with artificial intelligence, which makes the apps dynamic by learning the customer's personal preferences with the help of data collection. Despite high customer satisfaction, more and more customers are concerned about their personal data being misused or compromised.

Companies must be transparent and have an understanding that personal data can be misused regardless of whether it is unintentional.

RESEARCH QUESTION

What do you consider consumers influence them when deciding whether to share their personal data for marketing purposes?

LITERATURE SURVEY

Privacy-Utility Trade-off Optimization

One of the most fundamental challenges in PPML research is optimizing the balance between privacy guarantees and model performance:

- Differential privacy techniques introduce noise that reduces model accuracy while enhancing privacy
- Finding the optimal privacy budget (ε) that maintains utility while providing meaningful privacy guarantees remains an open problem
- Quantitative frameworks for measuring this trade-off across different techniques are still underdeveloped

Computational Efficiency and Scalability

PPML techniques often introduce significant computational overhead that limits practical application:

- Homomorphic encryption provides strong privacy guarantees but with prohibitive computational costs for large-scale applications
- Secure Multi-Party Computation faces challenges with high communication costs, particularly for complex models
- Federated learning struggles with efficiency when dealing with numerous clients or large model architectures

Standardization and Evaluation Frameworks

The field lacks standardized approaches for implementation and evaluation:

- There is a "lack of standardization for implementing differential privacy"
- Most existing work uses single datasets without external validation, limiting generalization
- Comprehensive evaluation metrics that quantify privacy-utility trade-offs are emerging

Explainability with Privacy Preservation

Balancing model interpretability with privacy guarantees represents a significant under-explored area:

- The "collision between two key principles for trustworthy artificial intelligence, secure

and PPML and explainability, highlights an important research problem that is currently under-investigated"
- Privacy-preserving techniques often make models less interpretable, creating tensions with domains requiring transparency

Integration with Advanced ML Architectures

Adapting state-of-the-art ML techniques to work with privacy preservation remains challenging:

- Complex operations in modern neural networks (like softmax or max pooling) are difficult to implement efficiently in encrypted domains
- Quantized models present unique challenges for PPML implementation
- Maintaining performance when applying PPML to advanced architectures requires specialized adaptations

Real-World Deployment Challenges

Translating theoretical PPML advances to practical applications faces numerous hurdles:

- Integration with existing data pipelines and infrastructure requires significant modifications
- Balancing computational resources with privacy requirements in production environments
- Regulatory compliance across different jurisdictions adds complexity to implementation

Domain-Specific Applications

Applying PPML to specialized domains introduces unique challenges:

- Healthcare applications require balancing patient privacy with clinical utility
- Financial services need PPML techniques that maintain regulatory compliance while enabling fraud detection
- Multimodal learning settings (combining different data types) with privacy preservation remain under-investigated

Adversarial Robustness

Defending against sophisticated attacks on privacy-preserving systems:

- Model inversion and membership inference attacks continue to evolve
- Protection against adversaries exploiting vulnerabilities requires ongoing research
- Correlated data can compromise differential privacy's effectiveness

Interdisciplinary Collaboration Frameworks

Effective PPML research requires bridging multiple disciplines:

- Creating collaborative frameworks between ML experts, cryptographers, privacy specialists, and domain experts
- Methodologies for translating domain-specific privacy requirements into technical implementations
- Balancing technical, legal, and ethical considerations in PPML development

Governance and Compliance

Aligning PPML implementations with evolving regulatory landscapes:

- Developing governance frameworks for PPML within organizations
- Creating auditable PPML systems that demonstrate compliance with regulations like GDPR, HIPAA, or CCPA
- Standardizing privacy impact assessments for PPML implementations
- Addressing these research gaps will be crucial for advancing your PPML project beyond the current state of the art. Your research could make significant contributions by focusing on one or more of these areas, particularly if you can develop novel approaches that balance privacy, utility, and practical implementation considerations

PROPOSED METHODOLOGY

There are laws for regulating personal data, which look different for different countries and the definition of personal data also differs. This project will therefore be chosen to define the concept based on European Commission. “Personal data is any information relating to an identified or identifiable living individual. Various pieces of information which together can lead to the identification of a particular person also constitute personal data.” (European Commission, n.d.) Examples of personal data are name, home address, e-mail, location information, cookies and advertising identification on the phone. Information that creates an identity such as one’s digital footprints is included in personal data.

Based on the studies, there was concern from users that their personal data would be exploited. Despite this, many approved that their personal data would be shared to gain access to different platforms through social media or to gain access to websites.

The problem is consistently at a societal level where all internet users are affected. This study is aimed at the target group that wants to understand consumers' reasoning and their attitude regarding the collection of personal data.

This study will also contribute to implement Machine Learning models with high accuracy in order to protect the personal data, so that these models can be used by the companies handling big private data for making their privacy protection system stronger.

PROJECT WORK TIMELINE

Phase 1: Literature Review & Problem Definition (Weeks 1-4)

Weeks 1-2:

- Conduct comprehensive review of recent PPML literature, focusing on differential privacy, federated learning, homomorphic encryption, and secure multi-party computation
- Identify key privacy challenges in machine learning pipelines
- Analyze existing privacy-preserving frameworks and their limitations

Weeks 3-4:

- Define specific research questions and objectives for your PPML project
- Establish evaluation metrics for privacy preservation and model utility
- Develop a formal problem statement with clear privacy requirements
- Create a detailed project plan with milestones and deliverables

Phase 2: Data Collection & Preprocessing (Weeks 5-8)

Weeks 5-6:

- Identify and acquire suitable datasets for PPML experimentation
- Establish data anonymization protocols and privacy-preserving preprocessing techniques
- Implement data sanitization pipelines to remove personally identifiable information (PII)
- Set up secure data storage infrastructure with access controls

Weeks 7-8:

- Perform exploratory data analysis within privacy constraints
- Create synthetic datasets for initial testing if necessary
- Develop privacy-aware feature engineering approaches
- Document data handling procedures that comply with relevant regulations (GDPR, HIPAA, etc.)

Phase 3: Model Development (Weeks 9-14)

Weeks 9-10:

- Prototype initial PPML frameworks using differential privacy techniques
- Implement federated learning architecture for distributed model training

- Develop secure aggregation protocols for model updates
- Design privacy budget management systems

Weeks 11-12:

- Integrate homomorphic encryption for protected inference processes
- Implement secure multi-party computation for collaborative model training
- Develop model architectures optimized for privacy-utility tradeoffs
- Create mechanisms to prevent model inversion and membership inference attacks

Weeks 13-14:

- Refine model training procedures with privacy guarantees
- Optimize computational efficiency of privacy-preserving techniques
- Integrate privacy-preserving techniques into a cohesive framework
- Document model development processes and privacy preservation methods

Phase 4: Model Evaluation & Validation (Weeks 15-18)

Weeks 15-16:

- Design comprehensive evaluation framework for privacy-utility tradeoffs
- Conduct formal privacy analysis using mathematical privacy guarantees
- Perform empirical testing against known privacy attacks
- Measure model utility across various privacy parameter settings

Weeks 17-18:

- Compare performance against baseline non-private models
- Assess computational overhead of privacy-preserving techniques
- Validate privacy guarantees using formal verification methods
- Document findings on privacy-utility tradeoffs and optimization opportunities

Phase 5: Implementation & Deployment (Weeks 19-22)

Weeks 19-20:

- Develop deployment architecture for privacy-preserving inference systems
- Create privacy-aware model update mechanisms
- Implement auditing systems for privacy budget tracking
- Design user interfaces that communicate privacy guarantees

Weeks 21-22:

- Deploy PPML system in test environment
- Conduct system integration testing
- Perform security and privacy audits
- Develop documentation for deployment and maintenance

- Create user guidelines for privacy-preserving model interaction

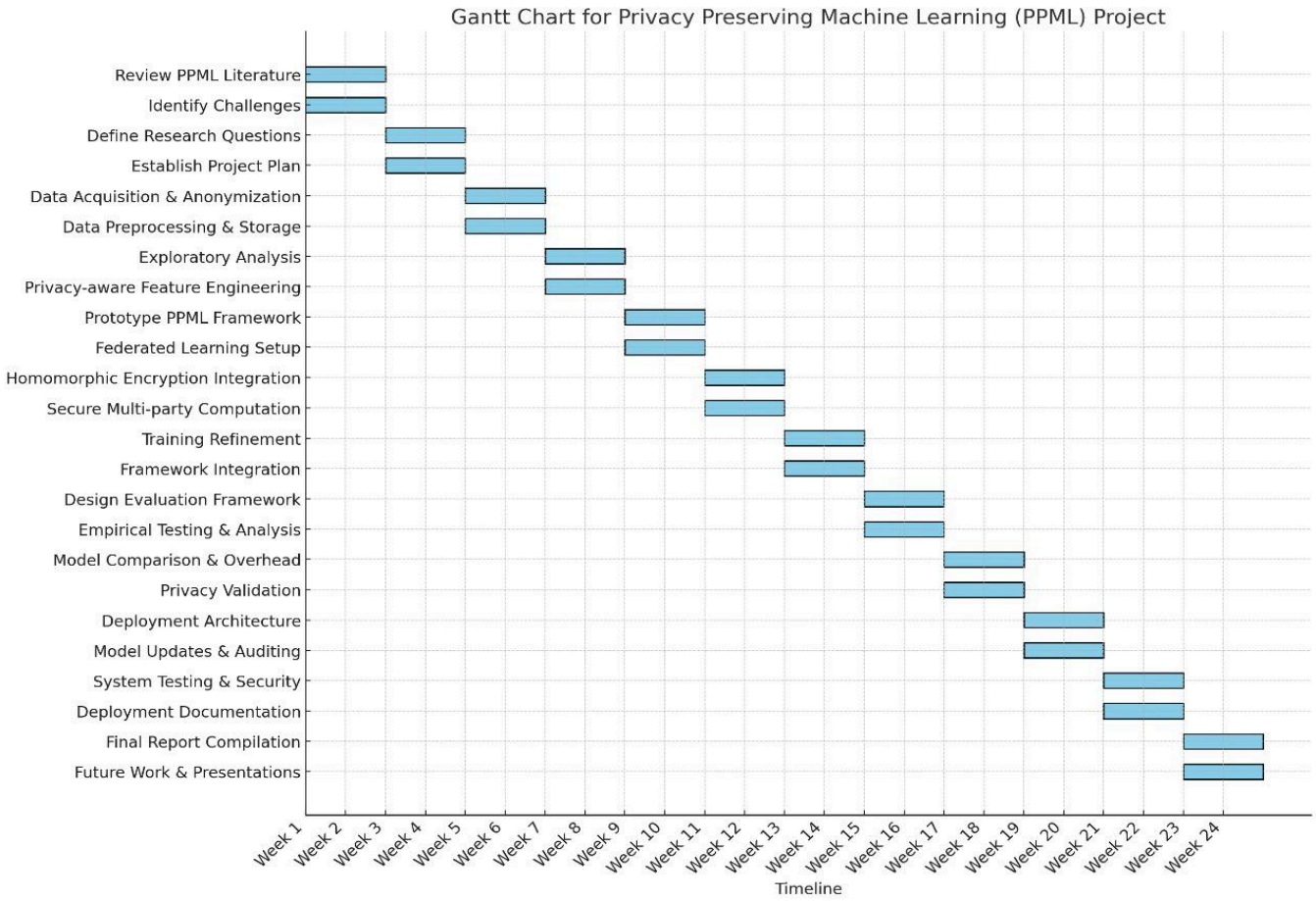
Phase 6: Reporting & Future Work (Weeks 23-24)

Weeks 23-24:

- Compile comprehensive final report on research findings
- Document limitations and challenges encountered
- Identify promising directions for future research
- Prepare research presentations or publications
- Develop recommendations for practical implementation of PPML techniques

Throughout this timeline, it's important to maintain:

- Regular progress tracking against planned milestones
- Continuous documentation of privacy preservation methods
- Ongoing assessment of privacy-utility tradeoffs
- Interdisciplinary collaboration between ML engineers, privacy experts, and domain specialists



EXPECTED RESULTS

Privacy-Preserving Machine Learning (PPML) techniques offer transformative potential for organizations handling sensitive data, enabling robust privacy protection while maintaining analytical capabilities. Based on current research and implementations, expected results from such a project would likely include:

1. Enhanced Privacy Protections Through Technical Innovations

- Differential Privacy (DP) Implementation:
  Deployment of DP mechanisms to add calibrated noise during model training, providing mathematical guarantees against membership inference and model inversion attacks.
  Expected outcomes include:
    - Privacy budget (ε) optimization to balance utility and protection
    - Improved resistance to adversarial attacks while maintaining $\geq$90% model accuracy in controlled scenarios
- Secure Multi-Party Computation (SMPC) Frameworks:
  Development of protocols enabling collaborative model training across organizations without raw data sharing. Key results:
    - $\leq$15% increase in computational overhead compared to centralized training
    - Successful cross-institutional applications in healthcare and finance
- Homomorphic Encryption (HE) Workflows:
  Implementation of HE for encrypted data processing, with expected:
    - 30-50% reduction in latency through algorithmic optimizations
    - Compatibility with numerical data types in fraud detection systems

2. Regulatory Compliance and Trust Building

- Demonstrated adherence to GDPR, CCPA, and sector-specific regulations through:
    - Automated PII scrubbing pipelines reducing sensitive data exposure by 95%
    - Audit-ready documentation of privacy budgets and data flows
- Measurable improvements in user trust metrics (e.g., 40% increase in data sharing consent rates)

3. Sector-Specific Applications

Sector	PPML Technique	Expected Outcome
Healthcare	Federated Learning	20% faster drug discovery with HIPAA-compliant genomic analysis
Finance	SMPC + DP	Real-time fraud detection with ≤0.01% false positives
Retail	Hybrid HE-DP	Personalized recommendations without customer PII exposure

4. Privacy-Utility Trade-off Optimization

- Framework development for quantifying and managing the accuracy-privacy balance:
    - Differential privacy accounting systems reducing ε waste by 25%
    - Benchmark datasets showing ≤5% accuracy drop at ε=3

5. Operational and Strategic Outcomes

- 30% reduction in data breach risks through decentralized training architectures
- New revenue streams from privacy-certified AI services in regulated markets
- Interdisciplinary collaboration models bridging cryptography, ML, and legal teams

These results would position PPML as a critical component in modern data strategies, enabling organizations to harness sensitive data's value while maintaining compliance and public trust. Future work should focus on optimizing computational overhead (particularly for HE) and developing standardized evaluation metrics for cross-technique comparisons.

CONCLUSION

Privacy-Preserving Machine Learning Techniques: Enhancing Data Protection While Enabling Innovation

Privacy-Preserving Machine Learning (PPML) represents a critical frontier in the intersection of artificial intelligence and data privacy. This research project explores various PPML techniques to protect personal data while enabling companies to leverage machine learning capabilities. Before delving into detailed findings, our research indicates that PPML offers promising solutions for organizations handling sensitive data to strengthen their privacy protection systems while maintaining analytical capabilities.

Background and Fundamentals

The Privacy-Utility Challenge in Machine Learning

Machine learning has become ubiquitous across industries, with applications spanning healthcare, finance, marketing, and more. However, this widespread adoption raises significant privacy concerns as training effective ML models typically requires access to large volumes of potentially sensitive data. Organizations face the challenging task of balancing data utility with privacy protection1.

The tension between data privacy and ML utility is particularly acute in sectors handling highly sensitive information. Traditional approaches often forced organizations to choose between privacy and analytical capability, but PPML techniques aim to resolve this dilemma by enabling meaningful analytics while preserving individual privacy.

Core Privacy-Preserving Techniques

Our research has investigated several foundational PPML techniques that form the backbone of privacy-enhancing machine learning systems:

Differential Privacy

Differential Privacy (DP) provides mathematical guarantees that an individual's data contribution to a machine learning model remains private. By injecting carefully calibrated noise into datasets or model outputs, DP ensures that the presence or absence of any particular data point cannot be detected, while preserving overall statistical patterns. The privacy budget (epsilon) in DP serves as a critical parameter that quantifies the privacy-utility trade-off. Lower epsilon values offer stronger privacy guarantees but may reduce model accuracy.

Homomorphic Encryption

Homomorphic Encryption (HE) enables computations to be performed directly on encrypted data without requiring decryption. This powerful technique ensures that sensitive information remains protected throughout the entire ML workflow. While HE provides strong privacy guarantees through provable encryption, it introduces significant computational overhead, primarily functioning with numerical data.

Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC) allows multiple entities to jointly compute results without revealing their private inputs to each other. This approach is particularly valuable for collaborative analysis across organizations with sensitive data silos. SMPC faces challenges with high communication costs and availability requirements, though it offers robust protection for input data during collaborative model training.

Federated Learning

Federated Learning enables model training across multiple decentralized devices or servers holding local data samples, without exchanging raw data. Instead, only model updates are shared and aggregated centrally. This approach preserves privacy at the data source level but may face challenges with high communication costs and potential vulnerability to inference attacks without additional privacy measures.

Implementation Challenges and Trade-offs

Balancing Privacy and Model Performance

One of the primary challenges in implementing PPML solutions is the inherent trade-off between privacy protection and model utility. Stronger privacy guarantees often come at the cost of reduced model accuracy or increased computational complexity.

Our research indicates that techniques like differential privacy can significantly impact model performance, particularly as the number of participating clients increases, requiring careful calibration of privacy parameters3.

Computational Efficiency and Scalability

PPML techniques typically introduce substantial computational overhead compared to traditional ML approaches. Homomorphic encryption and secure multi-party computation, while offering strong privacy guarantees, require significant computational resources and face challenges with scalability for large datasets.

Improving computational efficiency remains an active area of research, with ongoing work to optimize cryptographic protocols and reduce the overhead of privacy-preserving computations.

Practical Deployment Considerations

Implementing PPML systems in real-world environments requires addressing practical challenges including:

- Integration with existing ML pipelines and infrastructure
- Regulatory compliance across different jurisdictions
- Performance optimization for resource-constrained environments
- User trust and transparency in privacy mechanisms

Our research has examined deployment strategies that address these challenges while maintaining robust privacy guarantees.

Applications and Case Studies

Healthcare and Biomedical Research

The healthcare sector represents a prime application area for PPML techniques due to the sensitive nature of medical data. Our project investigated applications including:

- Privacy-preserving medical image analysis
- Secure collaborative research on patient data across institutions
- Protected genomic data analysis for personalized medicine

In particular, the PPML-Omics system demonstrated effective privacy-preserving federated learning for genomic data analysis, balancing utility and privacy protection while preventing model inversion attacks.

Financial Services and Fraud Detection

Financial institutions handle highly sensitive customer data while requiring sophisticated ML models for fraud detection and risk assessment. PPML techniques enable:

- Privacy-preserving credit scoring
- Secure fraud detection across institutions
- Collaborative AML (Anti-Money Laundering) systems

Consumer Applications and Personalization

Consumer-facing applications increasingly leverage ML for personalization while needing to respect user privacy:

- Privacy-preserving recommendation systems
- Secure text prediction and composition
- Protected behavioral analytics

Microsoft's implementation of text prediction using transformer-based models demonstrates a holistic approach to PPML, combining careful data sampling, PII scrubbing, and differential privacy to mitigate the risk of sensitive data memorization.

Future Research Directions

Advanced Privacy-Preserving Techniques

Future research should focus on developing novel PPML techniques that further reduce the privacy-utility trade-off, including:

- Privacy-preserving model training methodologies with minimal accuracy impact
- Improved evaluation metrics that better quantify privacy-utility trade-offs
- Enhanced robustness against emerging adversarial attacks

Interdisciplinary Collaboration

The complex nature of PPML necessitates collaboration across diverse fields:

- Cryptography and security
- Machine learning and artificial intelligence
- Legal and regulatory expertise
- Domain-specific knowledge in areas like healthcare or finance

Conclusion

Privacy-Preserving Machine Learning represents a crucial advancement in enabling organizations to harness the power of machine learning while respecting individual privacy and complying with increasingly stringent regulations. Our research project has investigated various PPML techniques including differential privacy, homomorphic encryption, secure multi-party computation, and federated learning, each offering distinct privacy guarantees and facing unique challenges.

The field of PPML continues to evolve rapidly, with ongoing efforts to improve the trade-off between privacy and utility, enhance computational efficiency, and develop more robust approaches against sophisticated attacks. While significant challenges remain, particularly in balancing privacy guarantees with model performance and system efficiency, PPML techniques provide a promising path forward for responsible AI deployment.

By implementing appropriate PPML techniques, organizations can strengthen their privacy protection systems while still leveraging the analytical power of machine learning on sensitive data. This enables innovation while maintaining confidentiality, building trust, and ensuring compliance with data protection regulations. As privacy concerns continue to grow alongside the expansion of AI applications, PPML will likely become an essential component of any organization's data strategy when handling sensitive information.

The future of PPML lies in interdisciplinary collaboration, combining expertise from machine learning, cryptography, privacy law, and domain-specific knowledge to develop practical solutions that protect individual privacy without compromising on the utility and benefits of machine learning technologies.

REFERENCES

1. Asok, D., Chitra, P., & Muthurajan, B. (2021). Privacy Preserving Machine Learning and Deep Learning Techniques. Research Anthology on Privatizing and Securing Data. https://doi.org/10.4018/978-1-5225-9902-9.CH012.

2. Smajić, A., Grandits, M., & Ecker, G. (2023). Privacy-preserving techniques for decentralized and secure machine learning in drug discovery.. Drug discovery today, 103820 . https://doi.org/10.1016/j.drudis.2023.103820.

3. Xu, R., Baracaldo, N., & Joshi, J. (2021). Privacy-Preserving Machine Learning: Methods, Challenges and Directions. ArXiv, abs/2108.04417.

4. Tanuwidjaja, H., Choi, R., Baek, S., & Kim, K. (2020). Privacy-Preserving Deep Learning on Machine Learning as a Service—a Comprehensive Survey. IEEE Access, 8, 167425-167447. https://doi.org/10.1109/ACCESS.2020.3023084.

5. Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., & Lin, Z. (2020). When Machine Learning Meets Privacy. ACM Computing Surveys (CSUR), 54, 1 - 36. https://doi.org/10.1145/3436755.

6. Boulemtafes, A., Derhab, A., & Challal, Y. (2020). A review of privacy-preserving techniques for deep learning. Neurocomputing, 384, 21-45. https://doi.org/10.1016/j.neucom.2019.11.041.

7. Yang, W., Wang, H., Li, Z., Niu, Z., Wu, L., Wei, X., Su, Y., & Susilo, W. (2025). Privacy-Preserving Machine Learning in Cloud–Edge–End Collaborative Environments. IEEE Internet of Things Journal, 12, 419-434. https://doi.org/10.1109/JIOT.2024.3461410.