International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN (P): 2319–3972; ISSN (E): 2319–3980
Vol. 11, Issue 2, Jul–Dec 2022; 201–230

International Academy of Science,
Engineering and Technology
Connecting Researchers; Nurturing Innovations

ENHANCING DATA PRIVACY IN MACHINE LEARNING WITH AUTOMATED COMPLIANCE TOOLS

Abhishek Das[1], Archit Joshi[2], Indra Reddy Mallela[3], Dr Satendra Pal Singh[4], Shalu Jain[5] & Om Goel[6]

[1]Researcher, Texas A&M University, North Bend, WA -98045

[2]Scholar, Syracuse University, Syracuse Colma C, 94014, USA

[3]Scholar, Texas Tech University, USA

[4]Ex-Dean, Gurukul Kangri University, Haridwar, Uttarakhand , India

[5]Independent Researcher Maharaja Agrasen Himalayan Garhwal University, Pauri Garhwal, Uttarakhand, India

[6]Independent Researcher, ABES Engineering College Ghaziabad, U.P., India

ABSTRACT

Data privacy has emerged as a critical concern in the era of widespread adoption of machine learning (ML) technologies. As organizations increasingly leverage ML models to extract insights from data, ensuring the protection of sensitive information while adhering to privacy regulations has become paramount. The implementation of privacy-preserving measures is often challenging due to the complexity of regulatory landscapes, including laws such as the General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), and the Health Insurance Portability and Accountability Act (HIPAA). Traditional manual approaches for compliance and privacy enforcement are not only time-consuming but also prone to errors, making them unsuitable for large-scale ML applications. This research paper proposes the adoption of automated compliance tools as a viable solution to address these challenges effectively.

Automated compliance tools, such as data anonymization, differential privacy, and secure multi-party computation, are designed to streamline the enforcement of privacy policies throughout the ML lifecycle. These tools can automatically detect compliance gaps, monitor data usage, and apply privacy-preserving transformations, thereby reducing the risk of data breaches and ensuring adherence to legal and ethical standards. By integrating automated compliance tools into machine learning workflows, organizations can achieve a balance between data utility and privacy protection without compromising the performance of ML models. Furthermore, automated tools facilitate continuous monitoring and real-time alerts, allowing organizations to respond promptly to any potential violations.

The research focuses on evaluating the effectiveness of various automated compliance tools in enhancing data privacy across different ML applications. A framework is proposed that integrates these tools into the standard ML pipeline, covering data preprocessing, model training, and deployment phases. The framework also includes automated mechanisms for consent management, data minimization, and secure data sharing. Using benchmark datasets and real-world case studies, the research demonstrates how automated compliance tools can maintain high levels of privacy while preserving model accuracy. Key findings suggest that the use of these tools can lead to significant reductions in privacy risks compared to traditional methods, particularly in scenarios involving high-dimensional data or large-scale data sharing across multiple stakeholders.

The implications of this research are significant for industries such as healthcare, finance, and e-commerce, where the use of sensitive personal information is prevalent. By adopting automated compliance tools, organizations can not only ensure compliance with data privacy regulations but also build trust with customers by demonstrating their commitment to protecting personal information. Moreover, the research highlights emerging trends in privacy-preserving ML, including the integration of federated learning and the development of new privacy metrics. It also outlines the challenges associated with implementing automated tools, such as scalability and computational overhead, suggesting potential solutions to overcome these limitations.

In conclusion, the adoption of automated compliance tools represents a transformative approach to data privacy in machine learning. It offers a scalable, reliable, and efficient solution to the complex problem of regulatory compliance, paving the way for broader adoption of privacy-preserving techniques in ML applications.

KEYWORDS: *Data Privacy, Machine Learning, Automated Compliance Tools, GDPR, CCPA, HIPAA, Privacy-Preserving Techniques, Differential Privacy, Secure Multi-Party Computation, Data Anonymization, Regulatory Compliance, Federated Learning, Privacy Metrics, Consent Management, Data Minimization, Ethical AI*

Article History

Received: *03 Sep 2022* | ***Revised:*** *12 Sep 2022* | ***Accepted:*** *18 Sep 2022*
