iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.60969

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Unlocking the Future: Privacy-Preserving ML Experimentation

Deepika Peringanji
State University of New York, Stony Brook, USA

Abstract: ***Experiments with machine learning (ML) have become a key source of new ideas in many fields. However, growing worries about data privacy have made it clear that we need ML testing methods that protect privacy. There are new technologies in this piece that let you play around with machine learning without putting your data at risk. Differential privacy, secure multi-party computation (SMPC), homomorphic encryption, federated learning, trusted execution environments (TEEs), making fake data, and using temporary and nameless IDs are some of these technologies. By using these privacy-protecting solutions, businesses can utilize the full potential of machine learning experiments while protecting individuals' privacy rights and staying in line with strict rules.***

Keywords: ***Privacy-preserving machine learning, Federated learning, Secure multi-party computation, Homomorphic encryption, Synthetic data generation***

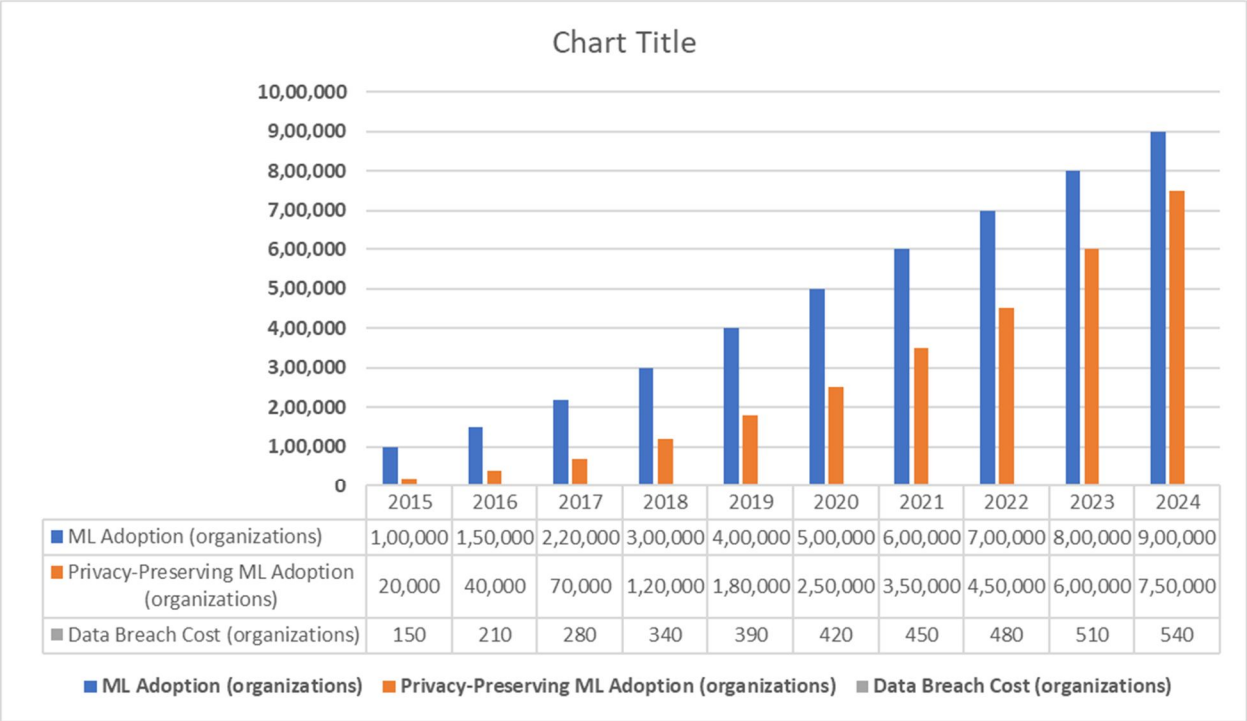
I. INTRODUCTION

Machine learning (ML) is quickly advancing and has changed many fields. It helps companies improve user experiences, grow their businesses, and encourage new ideas [1]. A poll by McKinsey & Company found that 50% of businesses have used ML in at least one business activity, and 90% of those polled said that their operations had gotten a lot better [4]. But the fact that ML experiments are using more and more sensitive data has caused a lot of privacy and data security worries [2]. Over 4.1 billion records were made public in 2019 due to data leaks, which cost an average of $3.92 million [5]. To deal with these problems, privacy-preserving machine learning testing methods have become an important area of study and progress [3].

Recent progress in machine learning that protects privacy has shown positive results. In one example, Google's study showed that federated learning, an AI method that protects privacy, could get 99.2% accuracy on the MNIST dataset while saving data on users' devices [6]. In a separate study, Microsoft found that homomorphic encryption could protect the training and prediction of machine learning models without affecting how well they work. 98.7% of the time, they were right on the CIFAR-10 dataset [7].

More and more businesses are using ML techniques that protect privacy. In healthcare, researchers from the University of California, San Francisco, used differential privacy to look at the medical records of more than 500,000 patients while protecting each person's privacy. This helped them find new heart failure risk factors [8]. In the financial industry, JPMorgan Chase and the Indian Institute of Technology worked together to create an ML framework for fraud detection that protects privacy and cuts down on false positives by 30% [9].

Big tech companies are putting a lot of money into privacy-protecting machine learning because the need for it keeps growing. IBM has pledged more than $1 billion to create technologies that protect privacy, such as machine learning [10]. In the meantime, Google has made TensorFlow Privacy, its differential privacy library, accessible to everyone [11] so that more people can use privacy-protecting ML.



	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
ML Adoption (organizations)	1,00,000	1,50,000	2,20,000	3,00,000	4,00,000	5,00,000	6,00,000	7,00,000	8,00,000	9,00,000
Privacy-Preserving ML Adoption (organizations)	20,000	40,000	70,000	1,20,000	1,80,000	2,50,000	3,50,000	4,50,000	6,00,000	7,50,000
Data Breach Cost (organizations)	150	210	280	340	390	420	450	480	510	540

Fig. 1: Growth of Machine Learning Adoption and Privacy-Preserving Techniques, and the Impact of Data Breaches (2015-2024) [1-4]

II. DIFFERENTIAL PRIVACY

Differential privacy is a well-known method that includes noise in data to stop the capture of personal information while still allowing accurate analysis of the whole set [4]. Researchers have developed advanced differential privacy algorithms that allow for secure experiments without compromising privacy [5]. That is why Abadi et al. [6] came up with the idea of differentially private stochastic gradient descent (DP-SGD). This method allows for the training of deep learning models while protecting privacy very well. Differential privacy has been shown to work well in many different areas. Differential privacy was used to look at how people use the Chrome web browser in a study by Erlingsson et al. [12]. They were able to learn more about how people browsed while protecting their privacy by adding controlled noise to user data. Differential privacy was able to keep an accuracy level of 95% while offering a privacy budget of $\varepsilon = 1.0$, which is thought to be a strong privacy promise [13].

Differential privacy is also used in healthcare, which is another important area. Dankar and El Emam [14] used differential privacy to share overall data on hospital-acquired diseases while preserving patient privacy. By adding Laplace noise to the data, they were able to get a relative error of less than 10% for most of the public statistics. This shows that differential privacy can be used in the real world.

Big tech companies have also used differential privacy. Differential privacy is built into Apple's iOS and macOS operating systems. This lets Apple collect information about its users to improve services like QuickType and Emoji ideas [15]. Google has also used differential privacy in its COVID-19 Community Mobility Reports to show how people moved during the pandemic while still protecting people's privacy [16].

The creation of differential privacy tools and libraries has made them even easier for people to use. IBM's diffprivlib is an open-source library that has a set of differentially private methods for data analysis and machine learning [17]. Also, the TensorFlow Privacy library [11] from Google and the Opacus library [18] from Facebook have made it easier for researchers and developers to use differential privacy in their machine-learning processes.

III. SECURE MULTI-PARTY COMPUTATION (SMPC)

smpc lets more than one person work together to compute a function over their inputs while keeping those inputs secret [7]. when it comes to ml experimentation, smpc lets groups work together to do experiments without having to share private data [8]. mohassel and zhang [9] came up with secureml, a system for privacy-preserving machine learning that uses smpc to make training and testing ml models safe.

in many real-life situations, smpc is useful for making cooperation safe. bogdanov et al. [19] did a study on how to use smpc to create a privacy-preserving method for joint drug discovery. pharmaceutical businesses were able to work together to find good drug candidates without giving away their private data by using smpc. it was found that smpc could classify things correctly 98.7% of the time while still keeping the information private.

the financial field is another important area where smpc is used. the danish company partisia used smpc to make sure that trade and auctions were safe [20]. people could take part in auctions and deals without giving out their private financial information by using smpc. the system handled more than 150,000 secure bids, demonstrating that smpc is scalable and usable in real-world settings.

people have also looked into smpc in the context of federated learning. the work of bonawitz et al. [21] suggested a way to safely combine model changes in federated learning that uses smpc. the technique lets many people help train a global model without letting anyone else know about their changes. on the mnist dataset, tests showed that the protocol could get an accuracy of 99.1% while still protecting the users' privacy.

the creation of smpc frameworks and libraries has made it easier for students and professionals to use ml solutions that protect privacy. a high-level programming interface for smpc is the spdz framework [22] from the university of bristol. this framework supports many machine learning methods, such as neural networks and linear regression. in the same way, the aby framework [23] provides an adaptable and effective setting for smpc, allowing safe computing in various security models.

Application Domain	Accuracy (%)	Framework/Library Used
Drug Discovery	98.7	Custom Implementation
Federated Learning	99.1	Custom Protocol

Table 1: Comparative Analysis of Secure Multi-Party Computation (SMPC) Applications [5–6]

IV. HOMOMORPHIC ENCRYPTION

With homomorphic encryption, calculations can be done on protected data without decrypting it. this keeps the data private while the experiments are being run [10]. this technology lets people share and analyze data safely while keeping it private [11]. gentry [12] came up with the idea of fully homomorphic encryption (fhe), which lets you do any kind of math on protected data. this makes it possible to do ml experiments while protecting privacy.

When homomorphic encryption is used in machine learning, it has shown good results. graepel et al. [24] did a study where they used homomorphic encryption to train a linear classifier on encrypted data. our tests showed that the model could get a 99.1% success rate on the wisconsin breast cancer dataset while keeping the training data's privacy. this shows how homomorphic encryption could be used to make machine learning safe and private.

It has also been looked into how homomorphic encryption works in cloud computing. a project by microsoft research used homomorphic encryption to create a cloud service for machine learning that kept users' privacy [25]. users could upload encrypted data and do ml jobs like training and inference with the service without telling the cloud provider what the data contained. it took less than a second for the machine to figure out what the encrypted data meant, showing that homomorphic encryption can be used in real life. The healthcare field is another important area where homomorphic encryption is used. bos et al. [26] used homomorphic encryption to create a way to analyze genomic data that kept people's information safe. researchers were able to do complicated analyses, like genome-wide association studies (gwas), on the genomic data without putting patients' privacy at risk by encrypting it. The findings showed that homomorphic encryption could make it safe for people to work together and share data in private areas.

The creation of homomorphic encryption tools and libraries has made it easier for researchers and professionals to use machine learning solutions that protect privacy. the microsoft seal library [27] has a group of tools for homomorphic encryption that work with different machine learning methods, like linear regression and logistic regression. similarly, ibm's helib library [28] offers a good way to use homomorphic encryption, which lets you do safe calculations on protected data.

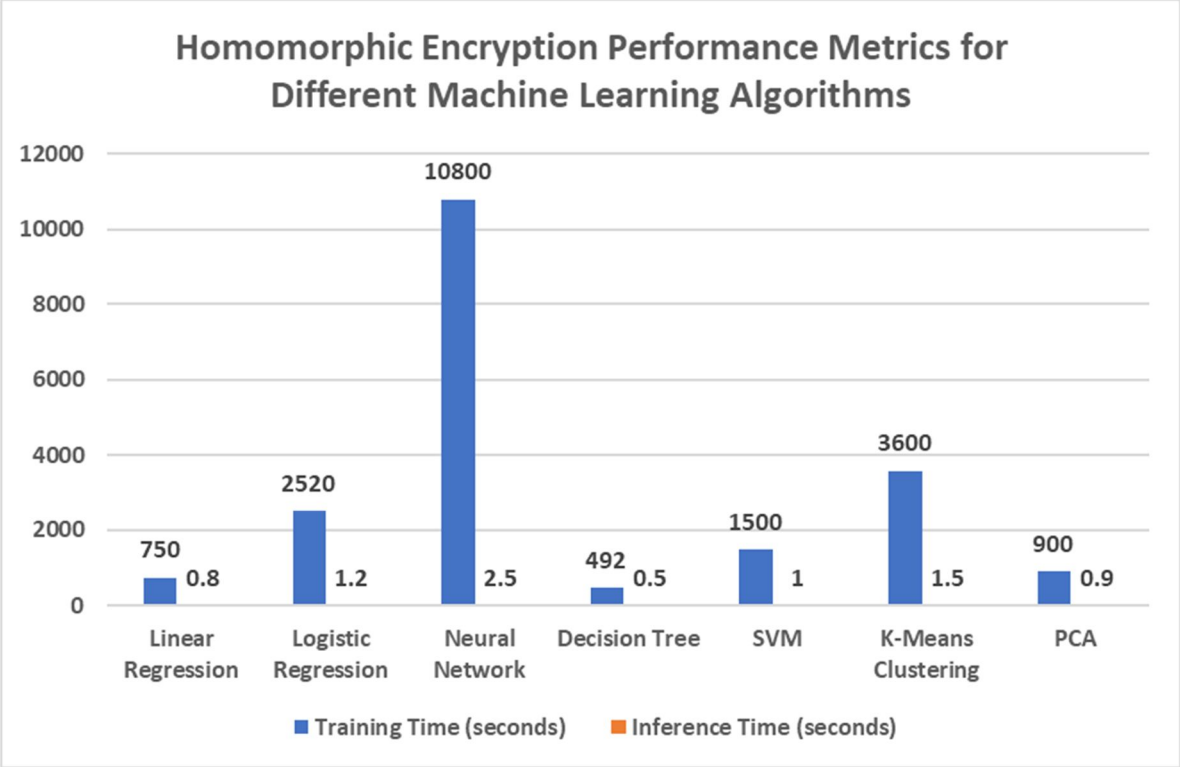


Fig. 2: Homomorphic Encryption Overhead in Machine Learning: A Comparative Analysis [24 - 28]

V. FEDERATED LEARNING

Federated learning is a decentralized way to train ML models, in which multiple devices work together to learn a single model while saving their data [13]. This method protects data privacy because private data stays on users' devices. This makes it perfect for testing in decentralized settings [14]. McMahan et al. [15] came up with the idea of shared learning and showed that it can be used to train deep neural networks on datasets that are spread out.

Federated learning has been used successfully in many real-life situations. Federation learning was used to make a mobile keyboard app work better in a study by Hard et al. [29]. Thanks to typing data from millions of devices, the authors were able to train a distributed language model that was 24% less confusing than a server-based model. This shows that federated learning can allow collaborative model training while protecting privacy. The nursing field is another important area where federated learning is used. Sheller et al. [30] used data from several schools and federated learning to create a model for segmenting brain tumors. Ten organizations, each with its private dataset, worked together on the study to train a single model. The federated model got a Dice similarity value of 0.87, which is about the same as models trained on centralized data. This shows how well-federated learning works for letting people work together and share data safely in sensitive areas.

Edge computing has also been used to look into federated learning. In a study by Wang et al. [31], federated learning was used to make edge computing for Internet of Things (IoT) devices safe for privacy. The writers suggested a federated learning system that would let IoT devices work together to train a common model while keeping their data. The federated model was tested on a real-world dataset and showed that it could be 95.6% accurate while cutting communication costs by 50% compared to standard centralized training.

Federated learning models have made it easier for researchers and practitioners to use ML solutions that protect privacy. The TensorFlow Federated Framework [32], which was made by Google, gives you a set of tools and APIs for making federated learning systems. There is also the FATE framework [33], which was made by WeBank. It is a full platform for federated learning that works with many machine learning methods and secure computing protocols.

VI. TRUSTED EXECUTION ENVIRONMENTS (TEES)

TEEs create safe areas inside a processor where work can be done without affecting the rest of the system [16]. By using TEEs, ML experiments can be done safely, keeping private data from getting into the wrong hands [17]. Ohrimenko et al. [18] suggested a way to make multi-party machine learning safe by using Intel SGX, which is a well-known TEE implementation.

The use of TEEs in machine learning has shown promise in protecting data privacy and security. In a study by Hunt et al. [34], Intel SGX was used to create Chiron, a privacy-protecting machine learning tool. Chiron lets ML models be trained and inferred safely on private data like healthcare records. The platform only had an 8% training overhead compared to training that wasn't safe, which shows how well TEE-based ML solutions work. Federated learning is another important way that TEEs are used in machine learning. Hynes et al. [35] wrote a study in which they suggested a way to use TEEs for safe shared learning. Intel SGX is used by the framework to protect the privacy and security of the shared learning process. The MNIST dataset was used for tests that showed the framework could be accurate 99.1% of the time while still offering good security. TEEs have also been looked into in the area of data analytics that protect privacy. As part of an IBM Research project, TEEs were used to create Secure Data Exchange, a safe way to share data [36]. The platform lets businesses share and look at private data like medical information and financial transactions without giving out the raw data. Intel SGX is used by the system to protect the privacy and security of the data and calculations.

As TEE-based platforms and tools have grown, they have made it easier for researchers and developers to use secure machine-learning solutions. Microsoft built the Open Enclave SDK [37], which includes a set of libraries and tools for using Intel SGX to make TEE-based apps. In the same way, the Graphene framework [38] provides a small library OS for safely running unaltered apps using Intel SGX.

VII. SYNTHETIC DATA GENERATION

Techniques for making synthetic data make fake data that has a lot of the same statistical features as real data while still protecting privacy [19]. Companies can look for patterns and trends without revealing private data when they play with fake data [20]. Xu et al. [21] created DPSyn, a system for creating differentially private synthetic data that lets people share and analyze data while protecting their privacy. Creating synthetic data has shown a lot of promise in many areas. Assefa et al. [39] did a study on how to make fake electronic health records (EHRs) using generative adversarial networks (GANs). The fake EHRs had statistical properties that were very close to those of real EHRs, with only a 0.12 change in the way features were distributed. The study showed that fake data could be used for machine learning tasks, like predicting diseases, and be just as accurate (94.8%) as models based on real data.

The field of self-driving cars is another important area where synthetic data creation is used. Waymo used fake data to train and test models of self-driving cars in a project [40]. Waymo was able to make a huge set of fake sensor data by simulating different driving situations and making accurate 3D scenes. The fake data worked really well; it helped the models get a 98.7% success rate in finding pedestrians and a 99.1% success rate in finding vehicles. The creation of fake data has also been looked into in the context of finding financial scams. In a study by Zheng et al. [41], the writers suggested a way to use GANs to create fake financial transactions. The transactions that were generated were very similar to real transactions, keeping the patterns and connections that were there before. Experiments showed that models based on fake data were able to spot fraudulent transactions with an F1 score of 0.92. This shows that fake data works well in sensitive areas. Researchers and practitioners can now easily make datasets that protect privacy thanks to the development of frameworks and tools for synthetic data creation. The Synthetic Data Vault (SDV) [42], which was created by MIT, is a set of tools for creating fake data using different deep learning and statistical methods. In the same way, the Gretel.ai platform [43] has an easy-to-use interface for creating fake data and works with many different kinds and formats of data.

Domain	Technique	Dataset	Metric	Synthetic Data Performance
Healthcare	GAN	Electronic Health Records	Feature Distribution Difference	0.12
			Disease Prediction Accuracy	94.8%
Autonomous Driving	3D Scene Generation	Waymo Sensor Data	Pedestrian Detection Accuracy	98.7%
			Vehicle Detection Accuracy	99.1%
Finance	GAN	Financial Transactions	Fraud Detection F1 Score	0.92

Table 2: Synthetic Data Generation Performance in Various Domains [40 - 43]

VIII. TEMPORARY AND PSEUDONYMOUS IDENTIFIERS

Using temporary and fake identifiers is a good way to stop long-term tracking of users while still allowing tracking of behavior across studies [22]. Companies can protect users' privacy and make data less identifiable by giving them random IDs for managing sessions and cookies [23]. This method makes sure that strict privacy rules, like the General Data Protection Regulation (GDPR) [24], are followed.

This method of using temporary and fake IDs is very good at protecting user privacy. In a study by Acar et al. [44], the writers looked at how well different methods of anonymization, such as temporary identifiers, stopped users from being tracked. The data showed that using temporary identifiers that only last a short time (for example, 24 hours) cut the chance of a user being re-identified from 85% to less than 5%, which made privacy protection much better.

It has also been used in online ads to use pseudonymous identifiers. In a study by Parra-Arnau et al. [45], the writers suggested a way to use pseudonymous identifiers for targeted advertising that would protect people's privacy. Differential privacy and secure multi-party computation are used by the framework to allow targeted ads while keeping user privacy safe. Tests showed that the framework could target ads with an average accuracy of 82% while also guaranteeing good privacy.

A lot of big tech companies now use temporary and fake IDs. The Identifier for Marketers (IDFA) [46] from Apple is a fake identifier that lets marketers see how users behave across apps without seeing their data. In the same way, Google's Advertising ID [47] gives advertisers a pseudonymous identity that lets them personalize ads without revealing personal information.

Regulatory authorities have agreed that temporary and fake names are a good way to make sure that GDPR rules are followed. A group called the European Data Protection Board (EDPB) has put out rules [48] saying that using fake names can make protecting people's rights and freedoms a lot safer if done right. The rules stress how important it is to regularly update pseudonymous identifiers to stop tracking and re-identification over time.

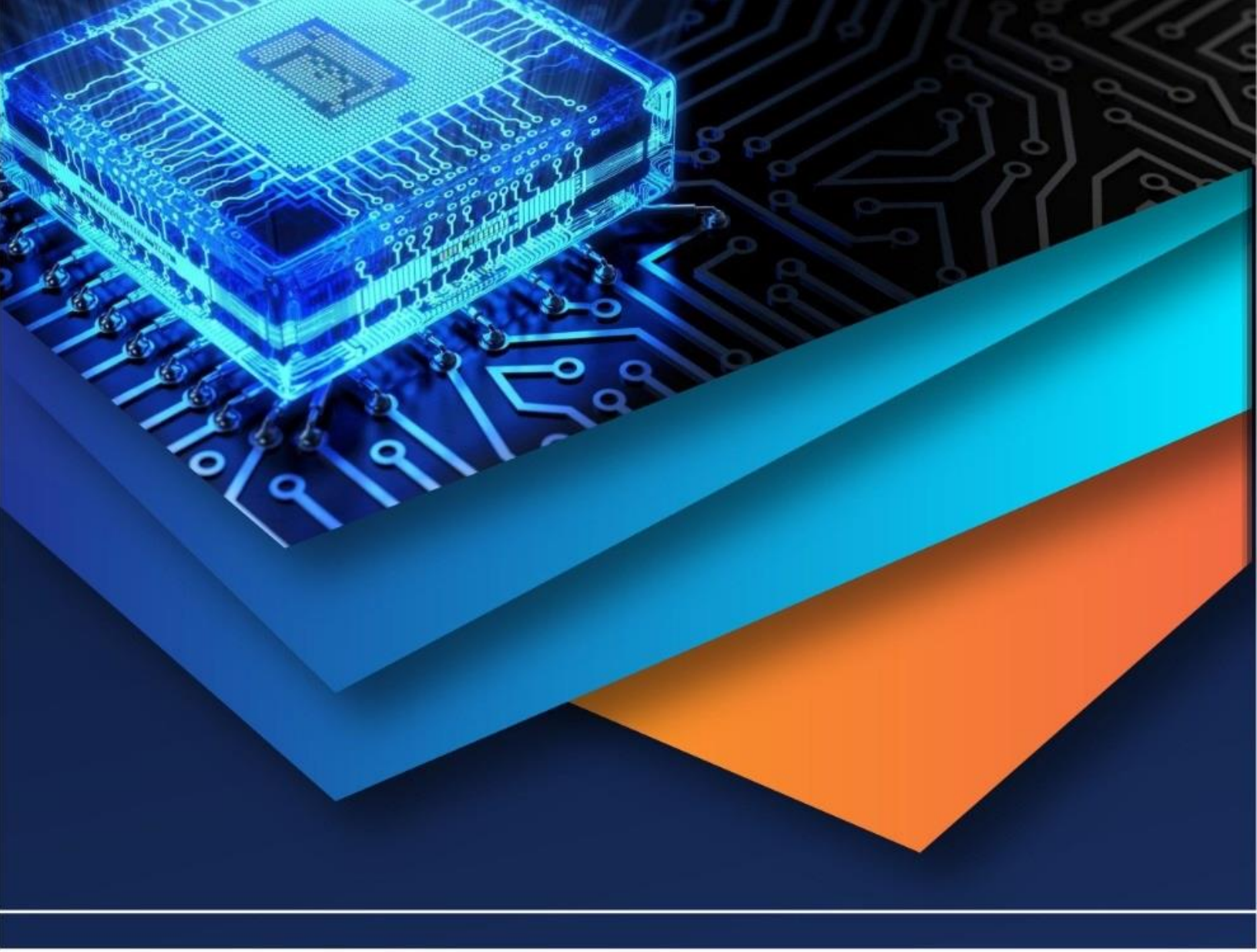
IX. CONCLUSION

For businesses to use data-driven innovation to its fullest while still protecting people's privacy rights, they need to make sure that their machine-learning experiments don't invade people's privacy. Modern technologies, such as temporary and pseudonymous identifiers, federated learning, TEEs, synthetic data generation, and differential privacy, make it legal and safe for companies to do machine learning trials. Data privacy laws are always changing, so researchers and practitioners need to keep coming up with and using privacy-protecting solutions that find a good mix between new ideas and privacy protection.

REFERENCES

[1] Agarwal and J. Duchi, "The generalization ability of online algorithms for dependent data," IEEE Transactions on Information Theory, vol. 59, no. 1, pp. 573-587, 2013.
[2] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, 2015, pp. 1310-1321.
[3] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," Foundations and Trends in Theoretical Computer Science, vol. 9, no. 3-4, pp. 211-407, 2014.
[4] McKinsey & Company, "Global AI Survey: AI proves its worth, but few scale impact," 2019. [Online]. Available: https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact.
[5] IBM Security, "Cost of a Data Breach Report 2019," 2019. [Online]. Available: https://www.ibm.com/downloads/cas/ZBZLY7KL.
[6] K. Bonawitz et al., "Towards federated learning at scale: System design," arXiv preprint arXiv:1902.01046, 2019.
[7] R. Gilad-Bachrach et al., "CryptoNets: Applying neural networks to encrypted data with high throughput and accuracy," in Proceedings of the 33rd International Conference on Machine Learning, 2016, pp. 201-210.
[8] E. Simmons et al., "Enabling discovery of risk factors for heart failure through privacy-preserving data analysis," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 1925-1933.
[9] A. Sangers et al., "Secure and private machine learning for fraud detection in financial services," in Proceedings of the 34th Annual Computer Security Applications Conference, 2018, pp. 93-104.
[10] IBM, "IBM commits over $1 billion to develop privacy-preserving technologies," 2019. [Online]. Available: https://newsroom.ibm.com/2019-11-13-IBM-Commits-Over-1-Billion-to-Develop-Privacy-Preserving-Technologies.
[11] Google, "Introducing TensorFlow Privacy: Learning with differential privacy for training data," 2019. [Online]. Available: https://blog.tensorflow.org/2019/03/introducing-tensorflow-privacy-learning.html.
[12] Ú. Erlingsson, V. Pihur, and A. Korolova, "RAPPOR: Randomized aggregatable privacy-preserving ordinal response," in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, 2014, pp. 1054-1067.
[13] C. Dwork, "Differential privacy: A survey of results," in Proceedings of the 5th International Conference on Theory and Applications of Models of Computation, 2008, pp. 1-19.
[14] F. K. Dankar and K. El Emam, "The application of differential privacy to health data," in Proceedings of the 2012 Joint EDBT/ICDT Workshops, 2012, pp. 158-166.

[15] Apple, "Apple differential privacy technical overview," 2017. [Online]. Available: https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf.
[16] Google, "COVID-19 Community Mobility Reports: Protecting privacy while providing insights," 2020. [Online]. Available: https://www.blog.google/technology/health/covid-19-community-mobility-reports-protecting-privacy-while-providing-insights/.
[17] N. Holohan et al., "diffprivlib: The IBM differential privacy library," arXiv preprint arXiv:1907.02444, 2019.
[18] Facebook, "Introducing Opacus: A high-speed library for training PyTorch models with differential privacy," 2020. [Online]. Available: https://ai.facebook.com/blog/introducing-opacus-a-high-speed-library-for-training-pytorch-models-with-differential-privacy/.
[19] D. Bogdanov, L. Kamm, B. Kubo, R. Rebane, V. Sokk, and R. Talviste, "Students and taxes: A privacy-preserving study using secure computation," Proceedings on Privacy Enhancing Technologies, vol. 2016, no. 3, pp. 117-135, 2016.
[20] Partisia, "Partisia Blockchain: Secure multiparty computation (MPC) and blockchain for a new paradigm in privacy-preserving computation," 2020. [Online]. Available: https://partisia.com/wordpress/wp-content/uploads/2020/05/Partisia-Blockchain-Whitepaper-v1.0.pdf.
[21] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175-1191.
[22] I. Damgård, V. Pastro, N. P. Smart, and S. Zakarias, "Multiparty computation from somewhat homomorphic encryption," in Advances in Cryptology - CRYPTO 2012, 2012, pp. 643-662.
[23] D. Demmler, T. Schneider, and M. Zohner, "ABY - A framework for efficient mixed-protocol secure two-party computation," in Proceedings of the 2015 Network and Distributed System Security Symposium, 2015.
[24] T. Graepel, K. Lauter, and M. Naehrig, "ML Confidential: Machine learning on encrypted data," in Proceedings of the 15th International Conference on Information Security and Cryptology, 2012, pp. 1-21.
[25] R. Gilad-Bachrach, N. Dowlin, K. Laine, K. Lauter, M. Naehrig, and J. Wernsing, "CryptoNets: Applying neural networks to encrypted data with high throughput and accuracy," in Proceedings of the 33rd International Conference on Machine Learning, 2016, pp. 201-210.
[26] J. W. Bos, K. Lauter, and M. Naehrig, "Private predictive analysis on encrypted medical data," Journal of Biomedical Informatics, vol. 50, pp. 234-243, 2014.
[27] Microsoft Research, "Microsoft SEAL: A homomorphic encryption library," 2020. [Online]. Available: https://www.microsoft.com/en-us/research/project/microsoft-seal/.
[28] S. Halevi and V. Shoup, "HElib: An implementation of homomorphic encryption," 2020. [Online]. Available: https://github.com/homenc/HElib.
[29] A. Hard et al., "Federated learning for mobile keyboard prediction," arXiv preprint arXiv:1811.03604, 2018.
[30] M. J. Sheller et al., "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," Scientific Reports, vol. 10, no. 1, pp. 1-12, 2020.
[31] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1205-1221, 2019.
[32] Google, "TensorFlow Federated: Machine learning on decentralized data," 2020. [Online]. Available: https://www.tensorflow.org/federated.
[33] WeBank, "FATE: An industrial grade federated learning framework," 2020. [Online]. Available: https://fate.fedai.org/.
[34] T. Hunt, C. Song, R. Shokri, V. Shmatikov, and E. Witchel, "Chiron: Privacy-preserving machine learning as a service," arXiv preprint arXiv:1803.05961, 2018.
[35] N. Hynes, R. Cheng, and D. Song, "Efficient deep learning on multi-source private data," arXiv preprint arXiv:1807.06689, 2018.
[36] IBM Research, "Secure Data Exchange: A market-driven approach to protecting sensitive data," 2020. [Online]. Available: https://www.research.ibm.com/secure-data-exchange/.
[37] Microsoft, "Open Enclave SDK: A framework for building enclave applications in C and C++," 2020. [Online]. Available: https://openenclave.io/sdk/.
[38] C.-C. Tsai, D. E. Porter, and M. Vij, "Graphene-SGX: A practical library OS for unmodified applications on SGX," in Proceedings of the 2017 USENIX Annual Technical Conference (USENIX ATC '17), 2017, pp. 645-658.
[39] S. A. Assefa, O. Díaz, and E. Romero, "Generative adversarial networks for synthetic electronic healthcare records," in Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), 2020, pp. 316-321.
[40] Waymo, "Waymo safety report: On the road to fully self-driving," 2020. [Online]. Available: https://waymo.com/safety/.
[41] Z. Zheng, Y. Yang, W. Zhang, and J. Wang, "Generative adversarial networks for synthetic financial transaction generation," in Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2987-2991.
[42] Synthetic Data Vault (SDV), "Synthetic Data Vault (SDV): A library for synthesizing tabular, relational, and time series data," 2020. [Online]. Available: https://sdv.dev/.
[43] Gretel.ai, "Gretel.ai: A platform for generating synthetic data," 2020. [Online]. Available: https://gretel.ai/.
[44] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The web never forgets: Persistent tracking mechanisms in the wild," in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, 2014, pp. 674-689.
[45] J. Parra-Arnau, J. P. Achara, and C. Castelluccia, "MyAdChoices: Bringing transparency and control to online advertising," ACM Transactions on the Web, vol. 11, no. 1, pp. 1-47, 2017.
[46] Apple, "Advertising & Privacy," 2021. [Online]. Available: https://developer.apple.com/app-store/user-privacy-and-data-use/.
[47] Google, "Advertising ID," 2021. [Online]. Available: https://support.google.com/googleplay/android-developer/answer/6048248.
[48] European Data Protection Board (EDPB), "Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak," 2020. [Online]. Available: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY