Applying Homomorphic Encryption as a Solution to Privacy Concerns in Artificial Intelligence-Based Medical Diagnostic Algorithms

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Zachary Holden

Spring 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Kent Wayland, Assistant Professor, Department of Engineering and Society

The growth of the Internet of Things (IoT) has propelled the world into an "information Big Bang" in which 2.5 quintillion bytes of data are generated daily (Marr, 2018). Three of the core laws of computing outline the "network effect" phenomenon, conjecturing en masse that the value of a telecommunications network generally scales with size (Delony, 2013). Accordingly, this means that the near ubiquity of smartphones and other Internet-enabled devices—in conjunction with the overwhelming popularity of social media and widening deployment of quantum computers and 5G networks—will continue to exponentially compound network effects and in turn propel the rate of data expansion (Kerry, 2020). Per the three "Vs" of "volume", "variety", and "velocity", more data will open the door for increasingly robust (volume), diversified (variety), and timely (velocity) statistical analysis (Kerry, 2020, para. 2).

As such, artificial intelligence (AI), a sweeping moniker for machine learning, algorithmic decision-making, and other modern statistical methods, is experiencing a boom in today's big data-driven world. The proliferation of AI-based technologies (e.g., speaker identification, image classification, sentiment analysis, and facial recognition) presents an opportunity to achieve unparalleled levels of productivity and efficiency (Stahl, 2021). AI's unique abilities to "link data, find patterns, and yield outcomes across domains and geographic boundaries" allow it to "be more consistent than humans, quickly adapt to changing inputs, and free humans from tedious or repeated tasks" (Stahl, 2021, n.p.). While the technology has been employed in everything from baseball analytics to search engine optimization (SEO) for digital marketing, one of the most promising applications to date lies in healthcare. Using large sets of clinical data, retrieved directly from wearable devices or more indirectly from disease registries or health surveys, AI has found use in diagnosis, clinical decision-making, and personalized medicine, with impressive performance (Rigby, 2019). In one instance, a trained deep neural network was actually able to more accurately identify skin cancer than a board-certified dermatologist (Rigby, 2019).

However, this promising technology does not come without serious ethical, legal, and social implications rooted in issues of privacy and data protection, bias and transferability, and moral and professional responsibility (Carter et al., 2020). Currently, the privacy-related concern of patient confidentiality stands as the most frequently cited and hotly contested obstacle to the widespread implementation of AI-based diagnostic algorithms in the medical industry (Stahl, 2021). The artificial neural networks at the core of these algorithms rely on large training datasets, and accessing those datasets poses clear privacy issues from potential unwanted exposure to sensitive information (Stahl, 2021). Thus, as an increasingly data-reliant society that values privacy, we find ourselves caught in an impasse produced by a chain reaction of expansion: more data produces more AI-based analysis, which produces more privacy-centered ethical headaches.

If we continue to rush the deployment of AI-based diagnostic algorithms in the healthcare industry, then we will overlook critical effects of socio-political forces that threaten to not only compromise the effectiveness of this innovative technology but also severely infringe on the long-respected intrinsic right to patient confidentiality. Thus, this paper seeks to advance the working knowledge of possible AI-based solutions in medical diagnostics by evaluating whether state-of-the-art cryptographic methods like homomorphic encryption can be employed to develop privacy-preserving machine learning (PPML) models that maintain the confidentiality of sensitive clinical data. In order to conduct a properly comprehensive analysis, the paper will employ a bottom-up approach. First, AI and homomorphic encryption will be independently reviewed, exploring their respective technical backgrounds and associated practical challenges.

Then, these elements will be black-boxed, or opaquely viewed purely in terms of inputs and outputs, and subsequently combined in order to holistically determine the privacy-preserving and functional utilities of the resulting sociotechnical system, a homomorphic encryption-equipped PPML model. All necessary supporting evidence will be collected from a combination of scholarly journal articles, industry documentation, and research papers.

Artificial Intelligence (AI)

Technical Background

After years of serving as a leading tech buzzword, the label "artificial intelligence" is barely more than unintelligible jargon, much like "personalization" or "augmented reality". Hence, we need to define the moniker, as well as explain the underlying methodologies, before a truly exhaustive analysis can be performed. Essentially, one can frame AI as an area of study that seeks to perform cognitive problem-solving through a combination of computer science, applied statistics, and data science (IBM Cloud Education, 2020). The entire field can be divided into three subsets: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Super Intelligence (ASI) (IBM Cloud Education, 2020). Today, strong AI—wherein computer systems have intelligence and ability that either rivals that of the human brain (AGI) or surpasses it (ASI)—is completely theoretical (IBM Cloud Education, 2020). All current practical AI-based applications employ weak AI (ANI) to perform targeted functions, with most of these systems falling under the robust sub-field of machine learning (IBM Cloud Education, 2020).

Machine learning models use prepared datasets to uncover patterns and in turn develop predictions without extensive human guidance (Brown, 2021). While software engineers are able to enhance the overall accuracy by choosing the model architecture, tweaking the overarching learning parameters (i.e., hyperparameters), and adjusting the amount of training data set aside for evaluation, ultimately it is the computer itself that performs all of the actual learning (Brown, 2021). In practice, this typically occurs as a form of supervised machine learning, wherein models are trained on human-labeled datasets (Brown, 2021).

Machine learning has seen considerable advancement in the last decade with the advent of artificial neural networks (ANNs) and deep learning. Conceptually based on the human brain, ANNs contain thousands—or even millions—of interconnected processing nodes, or "neurons" (Brown, 2021). Each neuron multiplies inputs with a "synaptic" weight to reflect their relative importance and then outputs the sum of the adjusted values (Ananthaswamy, 2021). During the deep learning process, these weights are adjusted within the scope of the entire network via the two-part backpropagation algorithm (Ananthaswamy, 2021). During the purely inferential "forward" phase, random synaptic weights are used to generate possibly erroneous predictions on the input data (Ananthaswamy, 2021). Model performance is improved in the ensuing "backward" phase, which adjusts the synaptic weights based on their individual contributions to the collective error, working sequentially from the output layer to the input layer (Ananthaswamy, 2021). Simply, the model learns by making mistakes, correcting those mistakes, and repeating the process over and over until it reaches the highest possible accuracy.

Connections Between Dataset Quality & Model Performance

While poor learning performance can be attributed to an overly simplistic neural architecture or an excessive number of features used for generating predictions, more often than not the datasets themselves stand as the root cause of the issue (IBM Cloud Education, 2021). Noisy, fragmented, or otherwise insufficient training datasets—like those containing excessive human- and instrument-based errors that obscure key trends (noisy) or missing observations that hold valuable information (fragmented)—fundamentally hinder the backpropagation algorithm and by virtue the entire machine learning model. Datasets with a meager amount of samples will not supply ample evidence for the algorithm to properly adjust the synaptic weights during the backward phase, so the model may end up "underfitting" the data such that it cannot recognize key patterns. Moreover, datasets containing exceedingly complementary information will likewise create issues during the backward phase, causing the algorithm to adjust the synaptic weights to solely reflect patterns and noise in the training data. In turn, the model will perform poorly when attempting to make inferences with unseen data, typically referred to as "overfitting", or memorizing, the data (Amazon Web Services, 2022).

In this way, the complex trade-off between user privacy and technological efficacy can be fully perceived: stricter privacy legislation inherently limits the depth and variability of the training datasets, which in turn produces overfitting and potential bias in the machine learning models (Kerry, 2020). Conversely, models with access to a greater range of data will be able to generate more accurate, refined, and consistent predictions, albeit at the possible violation of user privacy rights. In fact, properly tuned machine learning models are so analytically powerful that they may be able to formulate patterns and associated conclusions that violate patient confidence and consent, even after applying rigorous data anonymization and randomization procedures (Stahl, 2021). Naturally, continued social and commercial adoption of machine learning-based technologies in medical diagnostics requires a solution that manages to circumvent this significant compromise.

Homomorphic Encryption

Technical Background

In *Data Hiding Techniques in Windows OS* (2016), Nihad Hassan and Rami Hijazi define cryptography as a discipline concerning mathematics-, computer science-, and electrical engineering-based methods that seek to conceal sensitive information by masking its original value. At its core, this process is executed with a key, which converts unencrypted data (i.e., plaintext) to encrypted data (i.e., ciphertext). Secret key cryptography, or symmetrical encryption, uses the same secret key for both encrypting plaintext and decrypting ciphertext. This presents a glaring security risk, as the recipient of the data must be given the secret key in a secure manner, or else the integrity of the entire cryptographic system could be compromised. Public key cryptography, or asymmetrical encryption, provides a solution with the advent of the keypair. Two different keys are used to encrypt and decrypt data during transmission, but the pair are mathematically linked so that plaintext encrypted with the public key can only be decrypted by the corresponding private key, and vice versa (Hassan & Hijazi, 2016). As a result, system users can exchange their public keys freely without jeopardizing the security of the associated private keys (Hassan & Hijazi, 2016). Moreover, users can encrypt data with a private key—and decrypt the data with the associated public key after transmission—to create a digital signature and verify the sender's identity (Hassan & Hijazi, 2016).

As such, asymmetric encryption has found standard application across an array of information security systems; however, it is far from infallible with regards to data processing. Data encrypted via public key cryptography must be unencrypted for accurate processing, thereby introducing further trust and privacy concerns (Will & Ko, 2015). If data is unencrypted outside of the owner's trusted environment, any intruder with the processing algorithm can readily gain access and exploit or even outright modify sensitive information (IBM, 2021). Just as a deadbolt provides no protection when the door is open, encryption—asymmetric or otherwise—provides no protection when the data is represented as plaintext. Homomorphic encryption is an extension of public key encryption that seeks to definitively resolve these issues, ushering in a new generation of cryptography that allows ciphertext to be processed identically to plaintext (Will & Ko, 2015). While there are multiple forms of homomorphic encryption, this paper will purely focus on modern (i.e., developed after 2017, which are commonly referred to as fourth-generation) fully homomorphic encryption (FHE) schemes that perform multiple operations over ciphertext (Will & Ko, 2015). Fourth-generation FHE schemes are designed to be among the most efficient and widely adaptable cryptosystems to date, yet they still contain serious inherent shortcomings that jeopardize their functional deployment in machine learning applications.

Functional Roadblocks with FHE

FHE, which is often touted as the "holy grail" encryption scheme, has seen minimal practical implementation due to its high computing overhead (Will & Ko, 2015). Elementary computations that may only take a few tenths of a second on plaintext are amplified to anywhere from a few seconds to a few hours with FHE (Will & Ko, 2015). This massive slowdown predominantly occurs due to the unique way in which modern FHE generates ciphertext. All fourth-generation schemes encrypt data using approximations, rather than exact values (van den Nieuwenhoff, 2021). As a result, computations over ciphertext produce a small amount of error that could accumulate and corrupt the data (Kluczniak & Schild, 2021). Accordingly, these cryptosystems require an extra step called bootstrapping to reduce the amount of error-based distortion by periodically re-encrypting the ciphertext (Kluczniak & Schild, 2021). Since it produces no useful computation toward the output, bootstrapping proves highly inefficient, weighing down both run-time and memory usage (Lee et al., 2021). Yet, it is pivotal to accurately evaluate ciphertext over thousands or even millions of function cycles, as would be required in deep learning applications (Lee et al., 2021).

Lamentably, this is not the only complication that has precluded the widespread implementation of FHE-based privacy protection in data processing with ANNs. Leading fourth-generation FHE schemes only support multiplication and addition over encrypted data, but ANNs typically rely on activation functions (i.e., algorithms that determine an individual neuron's output, or "activate" the neuron) employing more complex non-arithmetic operations (Lee et al., 2021). For instance, the default activation function for a majority of neural networks, rectified linear activation (ReLU), employs a piecewise linear function that outputs the input value directly if it is positive or zero if it is less than or equal to zero (Brownlee, 2019). As a workaround, several proposed machine learning models incorporating FHE have reduced the number of neural layers, eliminated bootstrapping, and developed arithmetic activation functions; however, the resulting distilled models are too simple to grasp the complexities of real-world datasets, severely underperforming against standard deep learning architectures (Lee et al., 2019).

The Phase, Guarantee, & Technical Utility (PGU) Triad

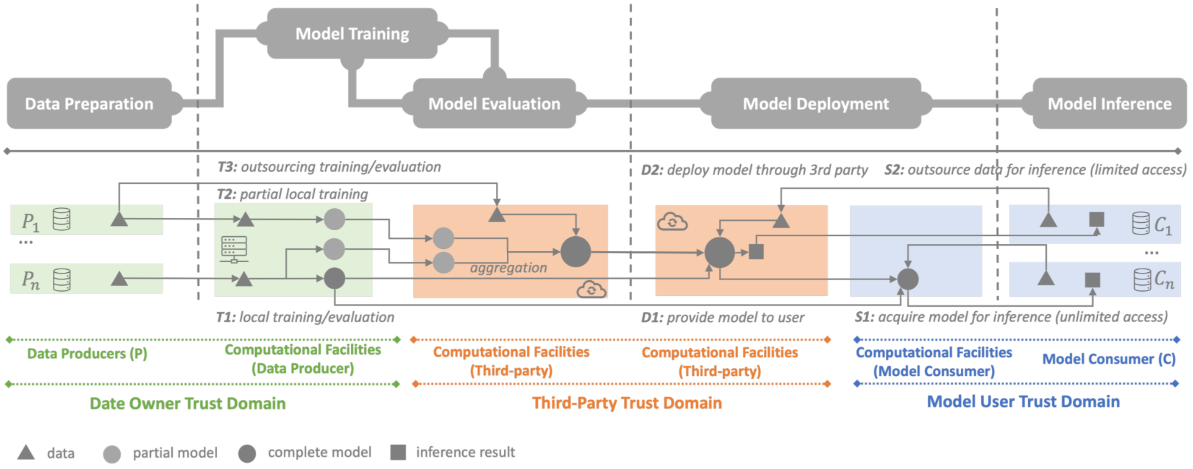
With a growing number of AI-based algorithms finding application in situations that require the use of sensitive information, vociferous concerns over data privacy and security have prompted the development of a new form of machine learning, termed privacy-preserving machine learning (PPML). Fundamentally, PPML attempts to quell these tensions by fortifying machine learning models with stringent data concealment procedures, sitting at the nexus of AI and information security. Duly, a suitable analysis of this intricate, multidisciplinary system requires a methodical interpretive framework. This paper will engage the Phase, Guarantee, and Technical Utility (PGU) triad, a targeted PPML diagnostic strategy proposed by Runhua Xu and Nathalie Baracaldo of IBM Research and James Joshi of the University of Pittsburgh's School of

Computing and Information in "Privacy-Preserving Machine Learning: Methods, Challenges, and Directions".

As defined by Xu et al., the "phase" pillar of the triad divides the machine learning pipeline into four phases, data preparation, model training and evaluation, model deployment, and model inference, in order to examine how a given approach offers privacy protection from a narrower, more categorical lens (2021). In the context of PPML, the former half can be collectively classified as privacy-preserving model generation and the latter half as privacy-preserving model serving. Recognizing that any part of this sequence is susceptible to privacy leakage, three associated trust domains identify potential sources of external risk: the data owner's trust domain, the third-party trust domain, and the model user's trust domain (Xu et al., 2021). Figure 1 draws explicit connections between these trust domains and the aforementioned pipeline phases.

Figure 1

The machine learning pipeline and associated trust domains, as outlined by Xu et al.


Model Training
Data Preparation
Model Evaluation
Model Deployment
Model Inference
T3: outsourcing training/evaluation
T2: partial local training
D2: deploy model through 3rd party
S2: outsource data for inference (limited access)
P_1
...
P_n
aggregation
C_1
...
C_n
T1: local training/evaluation
D1: provide model to user
S1: acquire model for inference (unlimited access)
Data Producers (P)
Computational Facilities (Data Producer)
Computational Facilities (Third-party)
Computational Facilities (Third-party)
Computational Facilities (Model Consumer)
Model Consumer (C)
Date Owner Trust Domain
Third-Party Trust Domain
Model User Trust Domain
▲ data
● partial model
● complete model
■ inference result

In turn, trust domain-derived potential sources of risk shape the security assumptions that underscore the "guarantee" pillar. Xu et al. frame PPML-based privacy guarantees as either pipeline-oriented (i.e., offering definitive security across the entire PPML pipeline) or object-oriented (i.e., offering definitive security at a specific element in the PPML architecture, such as the data samples or trained model weights) (2021). Moreover, object-oriented privacy guarantees are data-oriented if “an adversary cannot learn private information directly from input training/inference data samples or associate private information with a specific person's identification” or model-oriented if “an adversary cannot derive any private information from a given model by querying it a number of times" (p. 10). Finally, the triad's "technical utility” pillar importantly gauges the functional costs of applying these solutions to standard machine learning models via several key performance metrics, including computation utility, communication utility, model utility, and scalability utility (Xu et al., 2021).

Essentially, the PGU triad's three strategic pillars offer a complete lifecycle perspective, with each centered around the following driving questions:

1. Phase: *How does the approach preserve privacy?*
2. Guarantee: *How effective is the approach at preserving privacy?*
3. Technical Utility: *How does incorporating the approach impact model functionality?*

Along these lines, the framework examines the technical, social, and political factors surrounding the implementation of PPML models. Developers can apply the PGU triad to determine whether a certain privacy-preserving approach provides ample performance—in terms of both information security and prediction accuracy—before the model is implemented and adjust accordingly.

A PGU-Based Analysis of FHE-Equipped PPML

Phase

FHE-based privacy-preserving solutions are particularly interesting from a phase-oriented perspective, as they manage to offer high-level security without actually attaining full-chain privacy preservation. Essentially, they introduce a two-fold approach to data protection that combines encryption and secure computation to offer privacy-preserving functionalities across three of the four phases of the machine learning pipeline. Functionally, we find that FHE provides obvious safeguards during both data preparation and model inference by encrypting the training dataset and model prediction class with the owner's public key. At face value, neither the input data nor the associated model output will be intelligible to an intruder without access to the appropriate private key. Likewise, secure computation produces privacy-preserving model training and evaluation by allowing the supervised learning process to occur over encrypted data. On the contrary, this data-centric approach does little, if anything, to address privacy concerns with the model architecture itself, so it is impossible to claim that FHE attains any level of privacy preservation in the model deployment phase.

Guarantee

FHE-based privacy-preserving approaches only target security at specific phases of the pipeline by encrypting sensitive data and output values, leaving the inner workings of the neural network itself unprotected. Hence, they only offer an object-oriented, specifically data-oriented, privacy guarantee. This introduces the possibility to circumvent encryption through a membership inference attack, wherein an intruder uses the machine learning model's output to determine whether a particular instance was included in the training data (Shokri et al., 2017). Every time a model is queried, it generates a vector of decimal probabilities, with each value corresponding to the confidence that the given input represents a certain classification type (Shokri et al., 2017). Due to overfitting, machine learning models tend to predict with higher confidence on data that was included in the training set, and membership inference attacks take advantage of this Achilles heel to circumvent anonymization and encryption tactics (Shokri et al., 2017). Accordingly, such attacks are particularly damaging to data-oriented solutions, as they allow intruders to gain knowledge of material contained in the training dataset purely by detecting patterns in the output. Minimal prior insight into the PPML model's architecture or encryption scheme is necessary.

Machine learning as a service (MLaaS), which provides cloud-based machine learning tools via a third-party like Google or Amazon, is becoming increasingly popular since it can greatly simplify the development process and offset high recurring maintenance and computing costs, not to mention offer unparalleled portability and accessibility (Dickson, 2021). For these reasons, medical diagnostic PPML models will most likely employ MLaaS; however, this also means that they will require a model-oriented privacy guarantee transcending the data-oriented guarantee offered by FHE alone. MLaaS PPML models would be hosted on a cloud service, so little assumed trust can be placed in the third-party developer's domain. Under the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, third-party developers constructing the diagnostic models for a healthcare provider or health plan would be legally required to "use appropriate safeguards to prevent a use or disclosure of the protected health information" (OCR, 2019, n.p.). Consequently, MLaaS services must formulate exhaustive whole-pipeline PPML solutions that protect sensitive health information not only from direct identification (through FHE) but also from circumstantial association with a membership inference attack (through another privacy-preserving solution).

Technical Utility

Recent academic work has developed promising new bootstrapping algorithms that strive to resolve the aforementioned utility concerns regarding PPML models equipped with fourth-generation FHE schemes by allowing for deeper neural architectures and faster processing times. Addressing concerns regarding neural network depth, Lee et al. propose a high-fidelity bootstrapping algorithm that works with the Residue Number System (RNS) variant of the Cheon-Kim-Kim-Song (CKKS) homomorphic encryption scheme—the most efficient and accurate method to execute arithmetic operations over both real and complex numbers, commonly referred to as RNS-CKKS—and cutting-edge, low-error approximations of non-arithmetic activation functions like ReLU in "Privacy-Preserving Machine Learning with Fully Homomorphic Encryption for Deep Neural Network" (2019). To demonstrate the cryptosystem's practicality, the authors applied RNS-CKKS FHE with bootstrapping to ResNet-20, an industry-standard deep learning architecture for computer vision tasks, and performed object recognition with the widely used CIFAR-10 image dataset (Lee et al., 2019).

The proposed RNS-CKKS FHE-equipped model achieved a classification accuracy of 90.67%, which is 98.7% identical to the original ResNet-20 model working on plaintext (Lee et al., 2019). Importantly, this demonstrates FHE-based solutions' rare ability to fortify existing models with privacy-preserving functionalities, all while inducing negligible losses in model accuracy. Substantial previous research and development has generated powerful deep learning frameworks that apply image classification to diagnose everything from cardiovascular irregularities to blood-borne bacterial infections and even certain types of cancer, and FHE's practical flexibility allows for their continued use and advancement. Rather than having to go through the arduous and often largely ineffectual process of model retraining and restructuring, RNS-CKKS FHE can be used with modified activation functions and bootstrapping to adapt these highly successful machine learning models for computation over ciphertext.

Secondly, in "FDFB: Full Domain Functional Bootstrapping Towards Practical Fully Homomorphic Encryption" Kamil Kluczniak and Leonard Schild of the CISPA Helmholtz Center for Information Security introduce a flexible bootstrapping algorithm that reduces noise while simultaneously performing useful functions with minimal processing costs (2021). Accordingly, the proposed algorithm reduces run-times by a factor of 3000 compared to fourth-generation FHE schemes, bringing computations that may have taken weeks or days down to hours or even minutes (Kluczniak & Schild, 2021). This is particularly promising in the medical diagnostic space, as we see that state-of-the-art FHE-based privacy-preserving solutions achieve computational run-times that would not contribute to significant slowdowns in medical diagnoses, and may actually expedite the process in certain cases.

For example, let's assume that a PPML model equipped with flexible bootstrapping FHE requires two hours for inference, a fairly conservative estimate by Kluczniak and Schild's calculations. Diagnostic mammography performed by a radiologist usually takes anywhere from 10-30 minutes (Frazer & Wylie, 2018). As such, the hypothetical PPML model could reasonably be applied in this situation as a second opinion. On the other hand, oncological imaging diagnoses can take anywhere from a few days to a few weeks with delays (American Cancer Society, 2016). In this case, the model could be used to decrease wait times, providing patients with an initial assessment that same day. Thus, it becomes evident that computing overhead for FHE-based privacy-preserving solutions is not so laborious that it hinders the model's functionality entirely. Instead, determining the best means to incorporate these PPML models within existing industry practice and organization merely requires a bit of resourcefulness, which is rather trivial when considering the overall societal benefit acquired from properly maintaining patient confidentiality.

Conclusion

Machine learning models have shown great potential as a core element in the advancement of modern medicine, promising to usher in a new era of diagnostics built on enhanced efficiency and unprecedented accuracy. Notwithstanding, these algorithms rely on large sets of private health information for training and accordingly threaten to jeopardize the basic, long-heralded right of patient confidentiality. From a technological standpoint, homomorphic encryption presents a conceivable solution, yet further analysis is necessary in order to gain a holistic understanding of privacy-related socio-political effects surrounding the practical utilization of FHE-equipped PPML models. Employing the targeted PGU triad as an analytical framework, we can conclude that purely FHE-based solutions are a step in the right direction towards attaining complete privacy preservation but unfortunately not a complete solution: they cannot guarantee privacy across the entire machine learning pipeline, which grants the opportunity for membership inference attacks during model deployment.

That being said, FHE is still a particularly appealing option, as it can add privacy-preserving functionalities to proven, industry-tested deep neural networks with near-lossless accuracy. Proper implementation would require a hybrid approach that takes advantage of other object-oriented privacy-preserving approaches. For instance, private aggregation of teacher ensembles (PATE) creates an ensemble of "teacher" models that are independently trained on disjoint sets from the data (Xu et al., 2021). In turn, their collective predictive ability is distilled to a single "student" model, which induces noise such that it is impossible to glean any information from querying the student model multiple times, like in a membership inference attack (Xu et al., 2021). Adding RNS-CKKS FHE as proposed by Lee et al. produces a PPML model that offers a full-pipeline privacy guarantee with nearly identical performance to the original version.

Ultimately, the analysis presented in this paper answers a knowledge gap in how FHE-based PPML solutions can find practical applications in industry, as little previous work has been done to explicitly connect their use to privacy-preserving functionalities in medical diagnostics. As such, this opens the door for future work with employing PPML to enhance efficiency—while maintaining patient confidentiality—in other essential tasks in medicine, spanning everything from personalized care to natural language processing for report preparation and clinical note transcription (Davenport & Kalakota, 2019). Currently, privacy concerns stand as some of the primary underlying issues preventing the widespread adoption of these progressive technologies, and developing thorough, well-evaluated PPML solutions could help spur necessary social and regulatory acceptance.

References

Amazon Web Services (2022). *Model fit: Underfitting vs. overfitting*. AWS Developer Guide. https://docs.aws.amazon.com/machine-learning/latest/dg/model-fit-underfitting-vs-overfitting.html

American Cancer Society medical and editorial content team (2016, Jun. 6). *Understanding the cancer experience when you're a caregiver*. American Cancer Society. https://www.cancer.org/treatment/caregivers/what-a-caregiver-does/treatment-timeline.html#written_by

Ananthaswamy, A. (2021, Feb. 18). *Artificial neural nets finally yield clues to how brains learn.* Quanta Magazine. https://www.quantamagazine.org/artificial-neural-nets-finally-yield-clues-to-how-brains-learn-20210218/

Brown, S. (2021). *Machine learning, explained.* MIT Sloan School of Management. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Brownlee, J. (2019, Jan. 9). *A gentle introduction to the Rectified Linear Unit (ReLU)*. Machine Learning Mastery. https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/#:~:text=The%20rectified%20linear%20activation%20function,otherwise%2C%20it%20will%20output%20zero.

Carter, S. M., Rogers, W., Than Win, K., Frazer, H., Richards, B., & Houssami, N. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *The Breast*, *49*, 25-32. https://doi.org/10.1016/j.breast.2019.10.001

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6(2)*, 94-98. https://doi.org/10.7861/futurehosp.6-2-94

Delony, D. (2013). *The laws of computing*. Techopedia. https://www.techopedia.com/2/28205/trends/the-laws-of-computing

Dickson, B. (2021, Apr. 23). *Machine learning: What are membership inference attacks?* TechTalks. https://bdtechtalks.com/2021/04/23/machine-learning-membership-inference-attacks/

Frazer, H., & Wylie, L. (2018, Aug. 31). *Diagnostic mammography*. InsideRadiology. https://www.insideradiology.com.au/diagnostic-mammography/#:~:text=How%20long%20does%20diagnostic%20mammography,sure%20clear%20images%20are%20taken.

Hassan, N. A., & Hijazi, R. (2017). Introduction and historical background. In *Data Hiding Techniques in Windows OS: A practical approach to investigation and defense* (pp. 1-22). Syngress. https://doi.org/10.1016/B978-0-12-804449-0.00001-4

IBM (2021, Mar. 5). *Public key cryptography*. IBM Documentation. https://www.ibm.com/docs/en/ztpf/1.1.0.14?topic=concepts-public-key-cryptography

IBM Cloud Education (2020). *Artificial intelligence (AI)*. IBM Cloud Learn Hub. https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

IBM Cloud Education (2021). *Overfitting*. IBM Cloud Learn Hub. https://www.ibm.com/cloud/learn/overfitting#:~:text=When%20the%20model%20memorizes%20the,that%20it%20was%20intended%20for.

Kerry, C. F. (2020, Feb. 10). *Protecting privacy in an AI-driven world*. The Brookings Institution Center for Technology Innovation.

https://www.brookings.edu/research/protecting-privacy-in-an-ai-driven-world/#footnote-1

Kluczniak, K., & Schild, L. (2021). *FDFB: Full Domain Functional Bootstrapping towards practical fully homomorphic encryption*. Computing Research Repository (CoRR). https://doi.org/10.48550/arXiv.2109.02731

Lee, J.L., Kang, H.C., Lee, Y., Choi, W., Eom, W., Deryabin, M., Lee, E., Lee, J., Yoo, D., Kim, Y.S. & No, J.S. (2021). *Privacy-preserving machine learning with fully homomorphic encryption for deep neural network*. Computing Research Repository (CoRR). https://doi.org/10.48550/arXiv.2106.07229

Marr, B. (2018, May 21). *How much data do we create every day? The mind-blowing numbers everyone should read*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=41ec469660ba

Office for Civil Rights (OCR) (2019). *HIPAA for professionals: Business associates*. U.S. Department of Health and Human Services (HHS). https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/business-associates/index.html

Rigby, M. J. (2019). Ethical dimensions of using artificial intelligence in health care. *AMA Journal of Ethics*, *21*(2), E121-124. https://doi.org/10.1001/amajethics.2019.121

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). *Membership inference attacks against machine learning models*. Computing Research Repository (CoRR). https://doi.org/10.48550/arXiv.1610.05820

Stahl, B. C. (2021). Ethical issues of AI. *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies*, 35-53. https://doi.org/10.1007/978-3-030-69978-9_4

van den Nieuwenhoff, T. (2021, Nov. 14). *Fully homomorphic encryption: The history*. TVDN Blog. https://tvdn.me/fhe/2021-05-27-homomorphic-encryption-history/

Will, M. A., & Ko, R. K.L. (2015). A guide to homomorphic encryption. In R. Ko & R. Choo, *The Cloud Security Ecosystem: Technical, legal, business and management issues* (pp. 101-127). Syngress. https://doi.org/10.1016/C2014-0-00456-X

Xu, R., Baracaldo, N., & Joshi, J. (2021). *Privacy-preserving machine learning: Methods, challenges and directions*. Computing Research Repository (CoRR). https://doi.org/10.48550/arXiv.2108.04417