Published in Towards Data Science

Alexandre Gonfalonieri
Oct 8, 2020 · 7 min read ★ · Listen

Homomorphic Encryption & Machine Learning: New Business Models

Introduction to Homomorphic Encryption, use cases and impact on Machine Learning projects

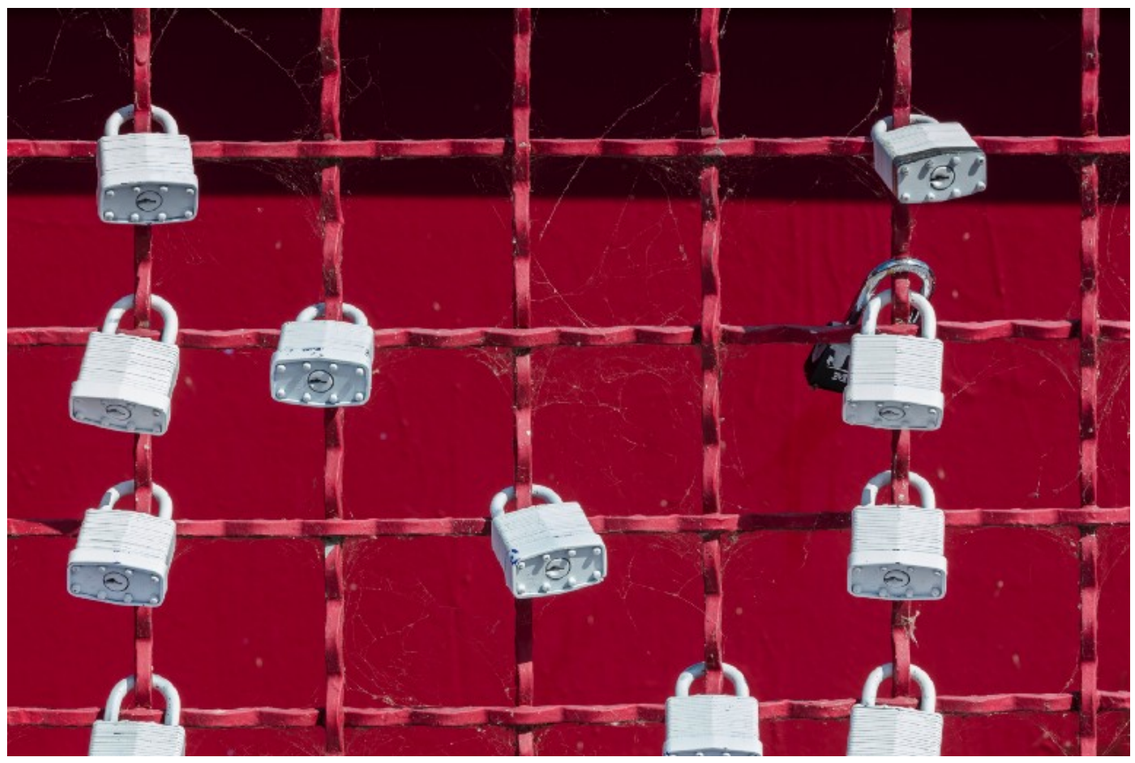


Photo by Jon Moore on Unsplash

One of the main “issues” about AI projects is data privacy. Indeed, you might identify the best use case for your company and then realize that your business project depends on data you are not allowed to use since you can not comply with existing data privacy regulations (for good reasons). This situation hinders our ability to leverage AI in real-life business applications.

Indeed, most Machine Learning systems are fed by data that are very sensitive and personal (customer data, health records, CCTV footage, etc.). After several projects in this industry, I can assure you that concerns over privacy legal issues are a serious barrier to the development of new AI solutions and business models.

Moreover, most companies are unaware that they could create new revenue streams using the data at their disposal while preserving the privacy of it. **A solution called Homomorphic encryption might help to improve the situation.**

In this article, I will explain what is Homomorphic Encryption (HE), the current limits of this technology, and how it will help create new business models while preserving data privacy.

Today’s issues

Let’s start with some common ML issues. Machine learning models are difficult to share across various users. Moreover, sharing modules is not feasible as it requires high computation at every end.

Most companies rely on cloud computing and leverage ML models through cloud APIs. This situation helps fix the issue of computational power but exposes data to the cloud providers.

> In some European countries, having a US cloud provider will prevent you from doing business with public organizations and state-owned companies (1).

Moreover, current data privacy regulatinons such as the GDPR (2) in Europe limits the use of AI solutions. [illegible]rk in the financial services

industry, your organization handles a lot of personally identifiable information (PII) and financial data that needs to be encrypted both when it is stored and when it is being transmitted.

> ***Personally identifiable information (PII)**: any information relating to an identifiable person (3).*

Finally, it is almost impossible for rival companies to exchange data due to a lack of trust between them (4). This situation limits our abilities to innovate when more data is needed (health research). Having a new encryption method would benefit many companies, create new ecosystems and business models.

What is Homomorphic Encryption?

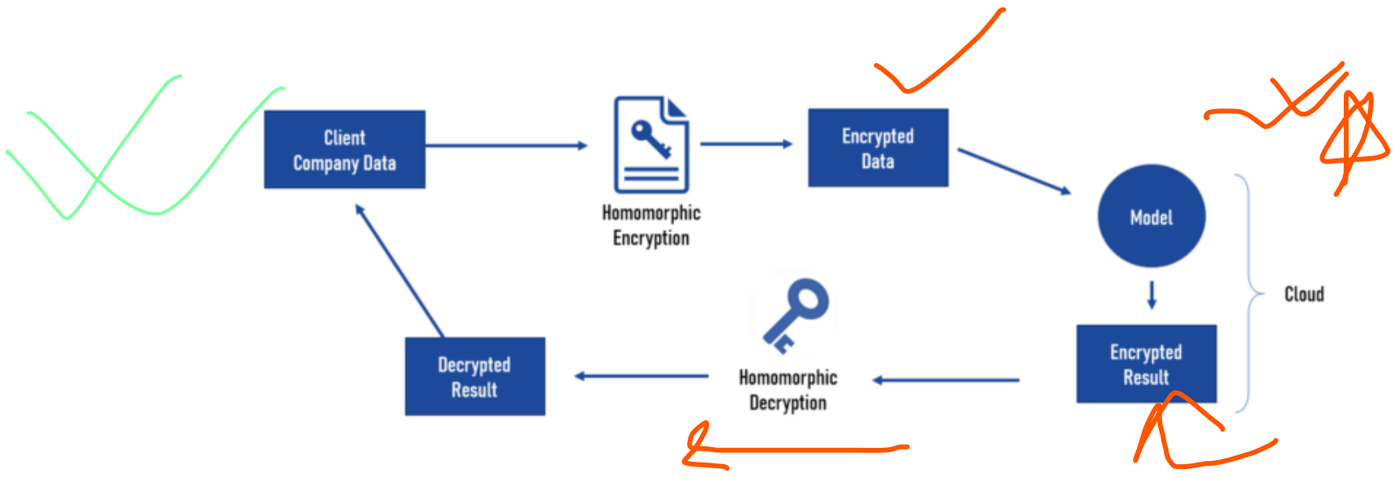
HE allows computations to be performed directly on encrypted data. By using advanced cryptology, it becomes possible to **"run machine learning on anonymized datasets without losing context"** (5).

> ***Computation**: The action of mathematical calculation.*

Whether you are working with data at rest or in transit, the majority of current public-key encryption requires that data must be decrypted before it can be manipulated.

In reality, existing encryption algorithms make it impossible to process the data without first decrypting it — and decrypting your data doesn't make you comply with data privacy laws. If data is encrypted by any other means it must first be decrypted before processing, and this fact makes data vulnerable to unauthorized access.

HE removes the need for decrypting the data before you use it. In other words, data integrity and privacy are protected while you process data. Indeed, homomorphic encryption allows encrypted data to be processed while it is still in an encrypted state.



HE/Cloud Process

Complying with data privacy laws (GDPR, etc.) often makes AI teams spend time trying to go around the regulation or reduce the scope of a project. **HE would help us store all data in the cloud encrypted, and perform computations on encrypted data.**

Despite this potential, the technology is not ready for mass adoption. In the near future, I expect to see more limited forms of HE, such as 'Searchable Encryption' and 'Multi-Party Computation' become mainstream. These solutions can achieve more or less what Homomorphic Encryption can. However, they do not have the main drawback of HE: **Slow computational speed (6).**

Currently, homomorphic encryption is impractically slow. This is, in part, because homomorphic encryption has a larger computational overhead than plaintext operations.

> Although the technology has been around for more than 40 years, computational barriers kept it relegated to the academic/research arena.

New Business Models

From a business standpoint, it will become possible to delegate the execution of a machine learning algorithm to a computing service while retaining the confidentiality of the training and test data. **This can be a major game-changer for the industry and change the cloud industry.**

HE will also make possible cooperation between several rival companies. To understand why, we must first analyze the rise of cloud computing. The success of Cloud service providers (ex: Azure, Google Cloud, etc.) can be explained by large investments made in data centers to offer services that help smaller companies lower their costs.

> *Soon, we could imagine a user that sends its encrypted data to the cloud using API and get the encrypted result from the machine learning models. During the entire process, the data is neither decrypted nor stored in the cloud. As such, the cloud provider could not access the users' data.*

However, one of the barriers to adoption of these services is related to privacy and confidentiality of the data being handled by the cloud provider, and the commercial value of that data/the regulations protecting the handling of sensitive data. By leveraging HE, cloud companies might be able to convince new companies to use their services.

> HE will enable new business models that would previously have been impossible.

HE will help create an entirely new data industry that brings together companies who hold data with companies or individuals who need data. This situation allows existing sensitive or regulated data assets to be used in ways that may previously have been determined as impossible.

I expect to see more data marketplaces enabling companies to sell their data using HE.

With the recent crisis, the need to create new revenue streams has become even more strategic. As such, a growing number of companies are increasingly analyzing how they might leverage existing data assets.

HE can help companies **securely monetized data while preserving the privacy of both customers and the data itself. I expect to see more startups specialized in the creation of data marketplaces enabling companies to sell their data using HE.**

Often, companies can end up having large valuable data sets completely unrelated to their business (data related to the manufacturing process, etc.). As such, they should be looking for a way to monetize these assets and provide them to data scientists who seek additional training data to make their AI systems more accurate, all while keeping the underlying data private and secure.

In Marketing, companies often have to collect data at every step of the customer journey (online and offline) to build a complete picture of the shopping experience. **However, some interesting potential data sources remain out of reach due to data privacy regulations. Homomorphic Encryption might help companies leverage new data sources while complying with privacy regulations.**

> Most AI Marketing projects lack contextual data to be perfectly accurate.

AI vendors might be familiar with the issue of data availability. In several use cases, AI vendors require more data to make a proof of concept successful (lack of accuracy, etc.). However, this is not always possible due to legal reasons. Leveraging HE, AI vendors could have access to additional data and demonstrate their algorithms on real data (banks, health institutions, etc.).

> HE would help rival companies collaborate by sharing data.

When it comes to cooperation, we can imagine that companies in the same industry will start sharing data and create new streams thanks to homomorphic encryption. The possibility to guarantee data privacy as well as financial rewards might help to motivate participants who would normally be reluctant to share their company's data (See: Federated Learning).

> ***Federated Learning:*** *A machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them (7).*

In specific use cases, sharing data with similar companies is the best solution from an AI point of view. Indeed, you could have rival financial institutions that may decide that it's in their interest to cooperate around certain mutual risks, like money laundering. By leveraging HE, companies could pool their data to "jointly build anti-money laundering models while keeping their sensitive customer data private" (8). Similar use cases can be found in the healthcare industry.

It is safe to assume that HE will lead to more breakthrough in the upcoming years. As mentioned by the World Economic Forum (9), data sharing has already helped many industries. For instance:

- Large sets of pooled patient data for medical studies
- Smarter urban transportation thanks to real-time location data

As with many breakthrough technologies, the market adoption of HE will be determined by the number of use cases leveraging HE. I expect to see a growing number of startups becoming specialized in the implementation of homomorphic encryption, while consulting firms will probably help companies accept the idea of sharing more data with their rivals.

For more information on Homomorphic Encryption, I recommend the following links:

- Projects related to Homomorphic Encryption
- Beyond trust: Why we need a paradigm shift in data-sharing
- Fully homomorphic encryption for machine learning
- What is Homomorphic Encryption?
- Using Fully Homomorphic Encryption to Secure Cloud Computing
- Can Homomorphic Encryption be Practical?
- Meet the new twist on data encryption that promises better privacy and security for AI