A STUDY ON
EYES OFF: ARTIFICIAL INTELLIGENCE

PROJECT REPORT

SUBMITTED BY:

TEAM AI: 30
CIPHER MAGNETS

PARTH GUPTA
VEDITA KAMAT
SOURABHA GANESH

Under the guidance of
MAHESH KUMAR KUSHWAHA

ARTIFICIAL INTELLIGENCE
In
SUMMER OF IDEAS
(Sept 2020)

Table of Contents:

S.No	Title	Page No.
1	Abstract	3
2	Introduction	4
3	Data Privacy	5-6
4	Possible Attacks	7
5	Scenarios	8
6	Current Available Technologies	9
7	Merits and De-Merits	10-12
8	Federated Learning over other techniques	13
9	Use-Cases	14-18
10	References	19

ABSTRACT

As Machine learning has become popular for getting insights from available data, it has potentially given rise to an important and critical issue of data privacy of data owners as well as model owners.

We generate a large amount of data from medical devices, smartphones, census etc. Several corporations collect data for getting deeper insights on issues critical to them. There have been guidelines (GDPR) set up by the EU to protect data and information of private citizens.

Today, there is no consensus on who is responsible for data privacy. Some consumers agree that the responsibility lies with them, but others think governments or businesses are better equipped to deal with this complex issue.

INTRODUCTION

Our main goal is to understand the importance of data privacy and formulate it to create a wrapper to prevent analysts from direct access to data and develop a sustainable and workable framework to train and measure classifier performance without a requirement for direct access to underlying data.

The principal reason to keep user's data protected is to ensure the safety of all the information that may contain the personal details, financial stats, past health conditions, as well as enable the data analyst to analyse the data for gaining insights.

We generate a large amount of data from medical devices, smartphones, census etc. Several corporations collect data for getting deeper insights on issues critical to them. There have been guidelines (GDPR) set up by EU to protect data and information of private citizens.

1. Data Privacy

When data that should be kept private gets in the wrong hands, bad things can happen. A data breach at a government agency can, for example, put top secret information in the hands of an enemy state. A breach at a corporation can put proprietary data in the hands of a competitor. A breach at a school could put students' PII in the hands of criminals who could commit identity theft. A breach at a hospital or doctor's office can put PHI in the hands of those who might misuse it. Since data privacy is such a prevalent issue, many government organizations and corporations spend millions of dollars each year to help protect their data—which could include your PII—from exposure.

For example: If we want to keep the reputation of the business on high stakes, it is necessary to encrypt the data in such a way that it should be accessed by only the lawful peoples.

1.1 Need of Data Privacy

Data Privacy or Information privacy encompasses 3 key elements:

1. Right of an individual to be left alone and have control over their personal data
2. Procedures for proper handling, processing, collecting, and sharing of personal data.
3. Compliance of data protection laws

1.2 Who is responsible for Data privacy?

Issues around privacy are an increasingly pressing concern. But what is less clear is who is responsible for protecting citizens. Is it up to **our governments? Organizations? Manufacturers?**

For their part, governments have been drafting policies to help ensure that personal privacy is being protected. While many of us are most familiar with the European Union's General Data Protection Regulation, it is by no means the only regulation keeping an eye on citizens' privacy.

North and South America, Asia, and the Pacific have also implemented policies aimed at protecting personal privacy. In fact, Malaysia's Personal Data Protection Act came into effect in 2013, Brazil's General Protection Data Law became enforceable in the summer of 2018, and California's Consumer Privacy Act, which recently passed into law, is set to take effect in 2020.

India's forthcoming policy may go even further than its predecessors as a result of a ruling by the country's Supreme Court that found "a right to privacy is part of the fundamental rights to life and liberty enshrined in the constitution." Based on this ruling, the new policy will likely affirm that "it's necessary to protect personal data as an essential facet of informational privacy."

1.3 Challenges with Data Privacy

- Classifying the data as Non-personal, personal, and personal sensitive. (As per GDPR directives) and apply appropriate algorithms as per requirement.
- Processing the data without having to store it on a single server. (Decentralisation, network and storage theft)
- Protecting the privacy of the model and the training data. (attacker should not be able to modify the hyper parameters or the individual records of interest from the training set)
- Protecting the privacy of model's output.(attacker should not be able to use the model output to test data of his interest)
- Trade-off between maintaining privacy and accuracy.
- Hierarchy among the analysts should decide the access levels.
  For example: Junior analysts do not have the same access to data as that of the senior analysts (update/delete access)

2. Possible Attacks

1. Membership Inference Attack: This attack aims to identify if a sample was present in the training set or not. It can further be extended to obtain individual attributes from the records.

   Soln:
   - Differential Privacy can reduce membership inference attack.

2. Reconstruction Attack: Reconstructing the raw data from the generated feature vectors. It is possible when the feature vectors used for training the model were not deleted after building it. In fact, some ML algorithms such as SVM or KNN store feature vectors in the model itself.
   Eg: Fingerprint raw image reconstruction from features

   Soln:
   - Avoid using SVM, KNN or models that require us to store feature vectors.
   - If usage of these algorithms cannot be avoided, ensure that these results are not available to any result party.

3. Model Inversion Attack: Generating features from model results. These features can further be used to carry out reconstruction attacks. Such attacks use the confidence information to construct the features.

   Soln:
   - Rounding the confidence values associated with the test data.
   - Predict only the class labels without showing the confidence values.

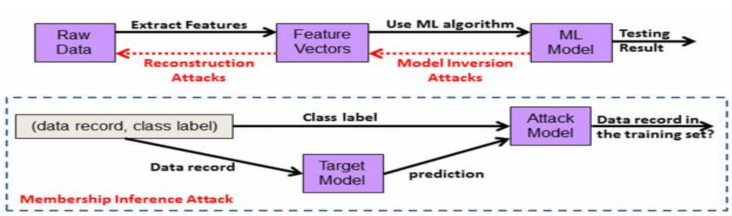


Figure 2: Machine Learning and threats

3. Scenarios

- **Sift Science**, a global fraud detection system uses machine learning technology to predict fraudulent behaviour.

- **Yelp**, **Air-bnb**, and **Jet airways** use data insights to protect themselves from content and promo abuse, payment fraud, fake accounts, and account takeover, etc.

- **E-commerce sites** are providing users with scratch cards and promo cards that enable them to enter their personal information like credit card details or bank details which can be viewed through fraudsters and that leads to data leakage and help them to corrupt the data.

- **Insurance companies** that collect the user's personal and other necessary information should be encrypted and wrapped so that it is not directly accessed by the analysts while the company analyses the user data for various reasons.

4. Currently available technologies

4.1 Anonymisation/Pseudonymisation:

- Includes removing certain sensitive attributes from dataset (anonymisation) and encrypting the sensitive attributes from data (Pseudonymisation).

4.2 Cryptography:

- Homomorphic Encryption: Data owner encrypts the data before sending it to the user. Meaningful calculations can be performed on the encrypted data.
- Paillier Cryptosystem: Partial Homomorphic scheme with different computations on cipher texts.
- Secure Multi-Party Computation: Separate parties can jointly perform the computation on data and receive the outputs without exposing any party's sensitive data.
- Private Set Intersection: If two parties want to test if their datasets contain a matching value but don't want to show their data to each other.

4.3 Privacy-by-design:

Incorporating privacy conserving techniques at every stage of machine learning pipeline.

4.4 Machine learning:

Machines have a tendency to learn from the data on their own. This can disclose some critical individual information. We need to ensure a machine can unlearn the data that it has learnt over time.

4.5 Differential Privacy:

It is used to measure the data leakage associated with machine memorization and reducing the possibility of it happening. It works by adding random noise to data.

4.6 Federated Machine Learning:

Training the ML model on data that is stored on different devices without the need for centralisation. The general model is sent to each of the devices where data is located and updates are sent back to the main server for improvisation of the model.

5. Merits and De-Merits for different approaches in Data Privacy

5.1 Anonymisation / Pseudonymisation:

MERITS	DISADVANTAGES
Stronger information security and counterpart for cyber security measures	Collecting anonymous data and deleting identifiers from the database limit your ability to derive value and insight from your data
Risk minimization regarding information transfers.	Even when everything is anonymised, malicious actors can figure out identities.
Possible information reuse.	Anonymised data could be the biggest threat to identity in the post-GDPR world.
Application of automated Big Data techniques.	Data suppression

5.2 Cryptography

MERITS	DE-MERITS
Classical Cryptography is independent of the medium of transmission, and hence can be used to transmit highly sensitive data over long distances.	difficult to access even for a legitimate user at a crucial time of decision-making
The one-time pad algorithm in classical cryptography is practically uncrackable.	High availability
When used properly, classical cryptography protects your plain text from all sorts of casual	selective access control also cannot be realized through the use of cryptography

snooping.	
Classical Cryptography is highly flexible and can be implemented in hardware, software, or a combination of both.	Cryptography does not guard against the vulnerabilities and threats that emerge from the poor design of systems, protocols, and procedures

5.3 Privacy-by-design

MERITS	DE-MERITS
It means potential problems with data protection can be more easily identified.	Privacy is a fuzzy concept and is Thus difficult to protect. We need to come to terms on what it is we want to protect.
It promotes greater awareness of privacy and data protection issues across an organisation.	System development life cycles rarely leave room for privacy considerations
Any action taken by a company is less likely to be privacy intrusive –	Little knowledge exists about the tangible and intangible benefits and risks associated.

5.4 Machine learning:

MERITS	DE-MERITS
Handling multidimensional and multi variety data	Data Acquisition
Easily identifies trends and patterns	Interpretations of Results
No human intervention needed	high error susceptibility
Continuous improvement	Time and resource

5.5 Differential Privacy:

MERITS	DE-MERITS
No need of attack modelling	It only works for interactive scenarios.
Privacy loss can be quantified	It can't provide good results for complex queries.
Multiple mechanisms can be composed	When there is a diversity in data, it includes too much noise which ultimately reduces the data utility.

5.6 Federated Machine Learning

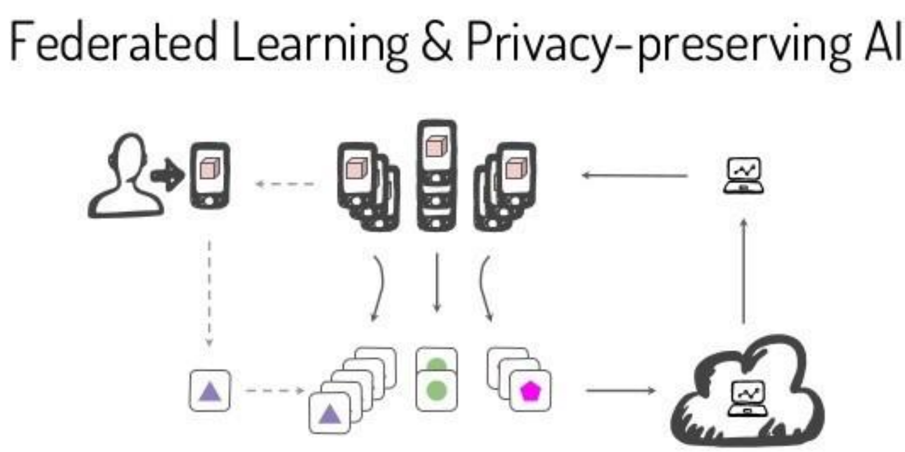
MERITS	DE-MERITS
This protocol uses cryptography to prevent the server from accessing the individual information summaries.	Expensive Communication
differential privacy can be used to add random data noise	Systems Heterogeneity
the learning process can be conducted when a device is charging, connected to Wi-Fi and not in use	Statistical Heterogeneity
This model of the algorithm then learns from the private data on the phones of a select group of users.	

6. Federated Learning over Other Techniques

With the advent of big data and computational power, stakeholders are becoming more and more interested in leverage the power of Artificial Intelligence (AI) and data-driven methods. However, one specific setback is the privacy restrictions regarding sensitive and private data. One example of confidential such data is the patients' healthcare records that should not be shared with any unauthorized individuals. To remedy these restrictions, researchers proposed and employed **federated learning** as a method to train Machine Learning methods. Federated learning eliminates the need to collect data from data holders which can drastically augment the data privacy.

In specific domains such as healthcare**, it is impractical to assume we can have a great centralized data to work with for Machine Learning purposes. Due to** privacy restrictions **enforced *by laws and regulations*, in healthcare, the majority of the data holders may not or cannot share their data nor willing to share any trained model on the sensitive data.**

Federated learning has the potential to solve some of the hurdles faced by methods that need the centralization of sensitive health data regarding privacy. By using federated learning, **healthcare stakeholders are not obliged to share confidential health data**. *The individual data providers or even the patients themselves can keep the full control of their data without the need to sharing it*. Such a system can *revolutionize the application of Artificial Intelligence* and Machine Learning in healthcare



7. Use Cases of Federated Learning

7.1.1 Smart Farming

Problem description: Smart farming requires us to produce more in an efficient and cost effective manner with the help of technology and innovation. Precision agriculture has been defined as a management system that is information and technology based, is site-specific and uses one or more of the following sources of data: soils, crops, nutrients, pests, moisture, or yield, for optimum profitability, sustainability, and protection of the environment.

This data has been classified into three categories: agronomic data, which refers to information regarding the yields of crops and the amount of input products applied; machine data, which refers to information about farm equipment; and weather data.

Key stakeholders:

- Farmers
- Data Analysts
- Researchers

Desirability:

1. GPS receivers, yield monitors, and variable rate application systems are now combined with other tools such as cell phones, personal computer systems and tablets to permit agricultural producers to collect and store a sizable and comprehensive amount of information about their farming operations.
2. The data about farming operations collected using the tools of precision agriculture has many characteristics that make it sensitive. The data collected may contain the personal information of individual producers. This information may include names and addresses, property locations, as well as crop yield information, which may lead to inferences about a producer's income and the value of their farmland. Producers are understandably unenthusiastic about the risk of this information getting into the hands of unauthorized third parties.
3. Producers are also concerned about competitors gaining access to their private data and using it against them.

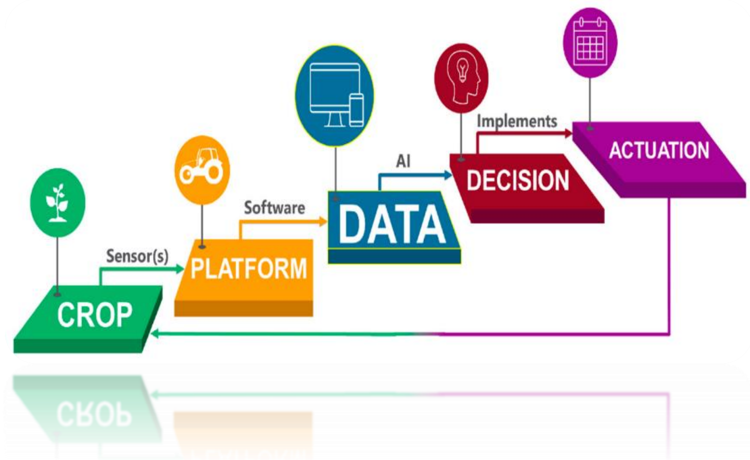
Feasibility:

1. Smart farming is already in use with edge computing, but a large number of farmers are reluctant in sharing the necessary information due to privacy issues pointed out earlier.
2. The use of federated Learning can motivate farmers to participate and practice smart farming.
3. Each plant can be equipped with a sensor box (e.g., with humidity and temperature sensors) and light-weight pre-trained models can be deployed locally to detect anomaly (Entity view) efficiently but with some acceptable error range. According to the idea of a Federated Learning framework, all the plants within a specific region (e.g., inside 2km range) are equipped with an Edge Node with pre-trained regional models along with pre-trained local models to be downloaded for each plant's end node within its service range. The regional model is more complex than the local models (and slower in giving prediction results) but is able to provide a more comprehensive view (Edge view) over the plants within its service range (e.g., anomaly detection). The global model is deployed at the Cloud Server, which is as complex as required to handle the vast amount of data generated by all plants covered by the system. The global model is able to give an overarching view (Global view) of the entire plant farm.

Viability:

1. Smart farming can be extended to suggest budget farmer friendly resources.

2. If farmers are more comfortable in sharing their private data, each farmer can get a personalised farming plan based on the type of land, soil, income capabilities, region of residence etc.



7.1.2 Healthcare

Problem description: Building AI model for severity and kind of treatment that needs to be provided to a Covid -19 patient.

Key stakeholders

- Hospitals
- Medical centres
- Researchers
- Data scientists

<u>Desirability</u>

1. Predicting the kind of treatment a patient should get based on his symptoms, X-ray /CT scan images, current morbidities, other lab reports.

2. since the data is personal and private to every individual there is need to protect it.

<u>Feasibility</u>

1. Since a large amount of data is generated by the hospitals, this data can be used to build a generalised model.

2. The federated learning technique helps to both preserve patient data privacy, as healthcare facilities keep their data on premises, and expose the AI model to a massive, diverse data set, which broadens the model's predictive ability, making it applicable to a wider range of healthcare facilities.

3. Since it is a global pandemic, a large amount of clinical data is available for research and usage. Federated Learning can be used to mitigate, treat and eradicate the spread of deadly diseases.

Viability

1. Can be extended to prediction of global supplies of vaccines per country /continent.

2. Can aid world health organisations, researchers to carry out medical research without worrying about security and privacy.

3. Can aid national and international bodies, healthcare units to plan and manage infrastructure like oxygen supply requires, kits etc. by using machine learning algorithms.

7.1.3 Suggestion/Personalisation application : Health assistant

Problem description: It takes important health data from each of the individual user, their images, and health conditions and offers them solutions and rewards for their actions. Using the concept of medical selfie for determining age, height, sex giving further insights to disease prediction and suggestions.

Key stakeholders:

- Hospitals
- User(individual)
- Data analysts
- Researchers

Desirability:

1. Medical data is highly sensitive and needs to be protected.
2. Medical selfie works on the concept that your image remains on your cellular phone and the model moves to your data thereby ensuring privacy.

Feasibility:

1. Edge computing combined with federated Learning can revolutionize the aspects of data privacy.
2. The data itself is present on the mobile device of users and used to train the model, thus generalising the model to give better insights.
3. Can further use differential privacy to strengthen the individual records from identification threats.

Viability:

1. We can add several other features to the application using IoT data generated from medical devices.
2. We can incorporate mental health using secure text communication and personalized emotion analysis and suggestions conforming the regulations.

8. References

- Bluemke, E., Trask, A., Lopardo, A., & Kang, N. (n.d.). Privacy Preserving Data Science. OpenMined. https://blog.openmined.org/private-machine-learning-explained/

- Braren, R. (2020, June 08). Secure, privacy-preserving and federated machine learning in medical imaging. Secure, privacy-preserving and federated machine learning in medical imaging. https://www.nature.com/articles/s42256-020-0186-1

- Github Repository for Ethical ML. (n.d.). https://github.com/EthicalML/awesome-production-machine-learning#privacy-preserving-machine-learningb.

- Mohammad Al-Rubaie, & J. Morris Chang. (n.d.). Privacy Preserving Machine Learning: Threats and Solutions. https://arxiv.org/ftp/arxiv/papers/1804/1804.11238.pdf

- *Privacy protection in machine learning: The state-of-the-art for a private decision tree*. (n.d.). https://www.researchgate.net/publication/324314039_Privacy_protection_in_machine_learning_The_state-of-the-art_for_a_private_decision_tree.

- Scientific development of smart farming technologies and their application in Brazil. (2018, March). *5*(1), 21-32. https://www.sciencedirect.com/science/article/pii/S2214317316301184#t0015

- Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming https://www.sciencedirect.com/science/article/pii/S1573521418302616