Overview of Differential Privacy in Machine Learning Algorithms

Madhushree M[1], Dr.K.S Jasmine[2]

PG Student, Department of MCA, RV College of Engineering, Bangalore, India[1]

Associate Professor, Department of MCA, RV College of Engineering, Bangalore, India[2]

ABSTRACT: Data comes with the responsibility of ensuring its privacy. With the present usage of internet collection of data has become very easy, but ensuring privacy of the same data still remains difficult. Differential privacy is gaining popularity and attention due to its strong mathematical definition which has been well established. Differential privacy has proved to be useful in preserving data even from the models that try to learn its features. Differential privacy is now being integrated with various machine learning algorithm to preserve the privacy of training data and also the utility of the model. In this paper three machine learning algorithms-Naive Bayesian, Decision Tree, Regression model are discussed with differential privacy.

KEYWORDS: Differential privacy, privacy budget, sensitivity, Naive Bayesian, Regression Model, Decision Tree

I. INTRODUCTION

Differential privacy is a new privacy mechanism that researchers are looking up to as it has been mathematically proved to be strong even against background attack. Organisations have remained hesitant since standard differential privacy is very strict and this could lead to the development of a model with zero utility though it provides complete security to customer data. But with careful selection of parameters the model can remain useful.

II. LITERATURE SURVEY

Differential privacy is now a hot topic for research,[1]gives a brief introduction to all concepts of differential privacy such as sensitivity, privacy budget, and different mechanisms to implement.[2]employs differential privacy in deep learning model layer by layer, i.e. input, hidden and output layer and also suggests that differential privacy helps in overcoming the problem of over fitting due to the added noise. Differential privacy is considered stricter than necessary, [3]suggests an alternative called individual differential privacy where the maximum difference between D and D^1 is exp(ϵ), to reduce the noise added to the model and also proved that any mechanism providing ϵ-Differential privacy can also provide ϵ-Individual differential privacy. Differential privacy is now applicable to every field where data is involved,[4]thoroughly discusses how differential privacy can be integrated with machine learning algorithms both supervised and unsupervised, such as Naive Bayesian, Linear regression, K-means clustering. Privacy budget is an important deciding factor in differential privacy,[5]proposes differential privacy budget allocation to a regression model and also discusses how it helps in avoiding inversion attack. Training data on which a model is developed contains very sensitive data, [6] discusses how Laplace noise can be added to each support vectors of SVM and also proved that it achieves the differential privacy requirements. Differential privacy can be used even in unsupervised algorithm, proposes improved version of K-mean algorithm based on MapReduce and use Laplace noise to achieve the protection. Differential privacy supports number of properties such as sequential, composition. In [7], Bayesian differential privacy is proved to have all the above properties. Bayesian differential privacy was introduced to remove the dependency between attributes in a database which also has proved to be vulnerable to adversary attacks.

The basic definition of differential privacy is, if a query is made to a dataset the result should be the same even if a record is removed or noise is added into the dataset. So this makes an attacker vulnerable as there is no way to know if the information he has belongs to the dataset or not.

The standard definition of differential privacy as per Dwork is -

A randomized mechanism M gives (ϵ,δ)-differential privacy for every set of outputs *S*, and for any neighbouring datasets of *D* and D^1, if *M* satisfies:

$$Pr[M(D)\ \epsilon\ S] <= \exp(\epsilon).\ Pr[M(D^1\ )]\ \epsilon\ S] + \delta$$

The users implementing differential privacy have to set the value for parameters such as sensitivity (δ) and privacy budget (ϵ).

The maximal difference on the result of a query if a record is removed from the data set is called sensitivity. Ex.: If employee database D has 100 records and if one record is removed to form D^1. If a query is executed on D^1 for number of employees then the maximum difference is 1.Hence it is said that D^1 has sensitivity of 1.

If δ=0 then a mechanism M gives ϵ-differential privacy which is the strictest form of differential privacy. In this form the model provides complete privacy but the utility of the model is 0 as it does not give any correct answer. Hence it is important to strike a balance between utility and privacy which is achieved with privacy budget.

Every result of a query in differential privacy consists of original result with some added noise. Hence if one is allowed to query a model any number of times, then the adversary can figure out the original value of the data because differential privacy has parallel composition property. Hence the number of queries that can be executed on a model is set, which is the privacy budget, i.e. The model is paying the cost of giving partial original result for every query it answers.

A smaller value of ϵ represents a strong privacy which means less number of queries. But if ϵ=0 then again it yields no valuable model. In practice the ϵ is usually set as less as 1, 0.1 or ln2. The mechanism M can be implemented using Laplace mechanism and Exponential mechanism. But this choice depends on the machine learning algorithm used.

III. DIFFERENTIAL PRIVACY IN MACHINE LEARNING ALGORITHM

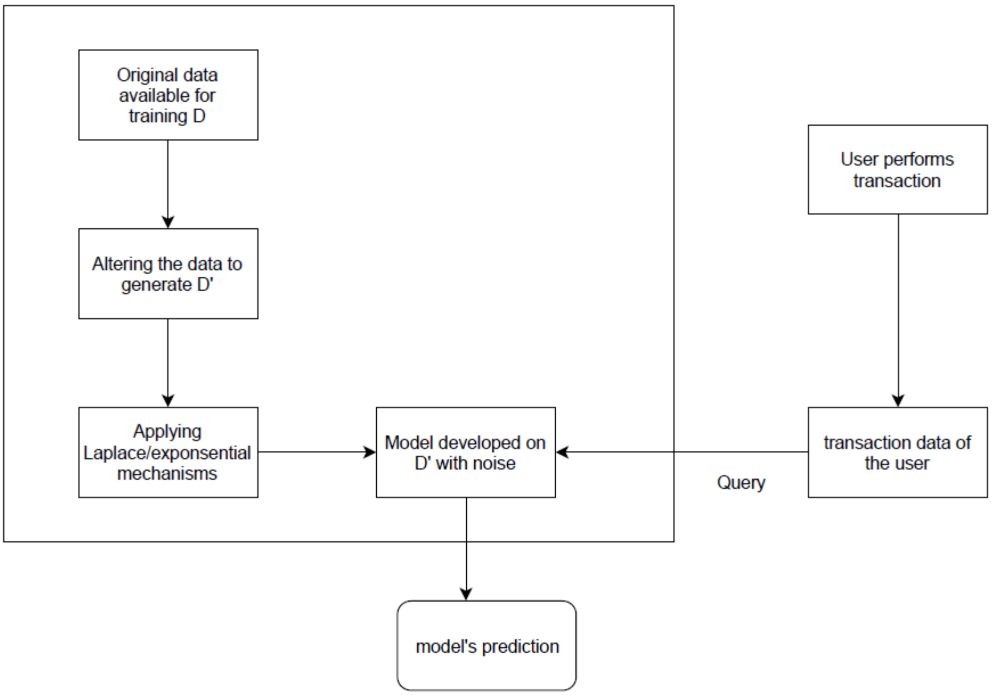
Original data available for training D
Altering the data to generate D'
Applying Laplace/exponential mechanisms
Model developed on D' with noise
User performs transaction
transaction data of the user
Query
model's prediction

Fig.1: Block diagram of differential privacy integrated with machine learning models.

Fig.1 is a block diagram depicting how differential privacy can be integrated with any system using training data to build their model. At first the organisation removes certain records from customer database before giving it for training ,next based on the machine learning algorithm chosen the differential privacy mechanism is chosen to add noise i.e. Laplace or exponential mechanisms. The same data is then finally used for training and testing the developed model.

Following section explains how differential privacy can be implemented for Naive Bayesian, regression model and decision tree

Table1: Description of the three algorithms that can integrate differential privacy.

Naive Bayesian	Regression Model	Decision Tree
Naïve Bayesian is selected when the end result is a conditional probability that a set of features $a_1,a_2…a_n$ belong to a class X.	In regression model, given a database $(x_1,y_1),(x_2,y_2)..(x_n, y_n)$ The model should be able to predict the value of y for any given x.	Decision Tree is supervised algorithm with no assumptions of underlying distribution in the training data. Recursively at each level of the tree the best attributed is determined to divide the dataset to finally identify the different labels.
In naive Bayesian algorithm the probabilities are calculated with mean and standard deviation of the attributes. Hence sensitivity is calculated for mean and standard deviation based on which noise is added[10]	In order to realize the privacy preservation the input features are divided into features that are most important for determining output and less important. In the objective function, to the coefficients of features which are more important less noise are added, while to coefficients of features with less importance more noise are added [11].	Decision Tree is more vulnerable to privacy leakage because the dataset is divided based on the description of attributes at each level. In DT if label generation of leaf node, splitting attributes and the number of instances in each node is protected with Laplace or exponential mechanism then the decision tree is considered DP protected [12].
Naive Bayesian model builds a model by learning the conditional probability of all the attributes and its class. In order to implement differential privacy noise are added to the attributes and not to the class. The naive Bayesian classifier is observed to provide approximate accuracy when noise is below 0.6. But above 0.6 the accuracy of the model drops to 0.Hence we have add to noise to the training data set in controlled manner and set sensitivity above 0.6 [9].	Regression analysis is often used to find relationship and pattern between attributes. For example, date and time of the transaction is dependent certain extent to the amount withdrawn. The place of transaction is also an important attribute. New place of transactions which is far away from the card holder's current location are variables for suspicion. But this relationship between attributes can also help adversary to gain information about the customers if they have access to the model using inversion attack. Example: the adversary can get details of the places a customer has made his purchases and the places where he has made the most transactions. To overcome this, noise should be added to the features in controlled measures and also to the polynomials of the function.	Decision tree classifiers can be used in systems such as fraud detection system to classify a transaction as safe or unsafe. The training data can be divided first if the value of transaction is less than average transaction, next if the place of transaction is new, next date and time and so on. To implement differential privacy mechanism noise must be added to the training data, but decision tree has the disadvantage of over fitting when there is noise in the training data. To overcome this problem the maximum of height of the tree and also privacy budget for each level of the tree must be set. Some authors also suggest using all privacy budget to find one best attribute that has the maximum information gain that can determine if the transaction is fraudulent or not.

Advantages

- Adding noise in controlled measures ensures the result is not far from actual result
- Adding noise also is helpful to avoid over fitting model when learning the training data
- Since the attributes considered sensitive are not removed or masked an efficient machine learning model is developed

Limitations

- With every result a part of original result is provided by the system
- In Decision Tree due to added noise the nodes continues to split indefinitely
- There is always danger of the developed model becoming useless due to incorrect selection of parameters

IV. ALGORITHM

Based on the above points, a general algorithm to use differential privacy is formulated.
Input: Data of customer's transaction from many years for training with labels
Processing: Generate synthetic data (D^1), the most common method is by adding noise
Based on the model used, add noise to the original data.-Noise can be sampled from Laplace Distribution centred at 0
Set the parameter values such as privacy budget and sensitivity- both should be greater than 0.

1. For SVM, add Laplace noise
   Lap (Δ/ϵ)
   Where Δ is sensitivity and ϵ is privacy budget.
   to the support vectors from Laplace distribution.

2. For Decision Tree, select the best attribute to stop splitting of the nodes in the tree using privacy budget(Ex, average withdrawal of a customer for a year can be an important attribute to check if the current transaction should be verified or not).

Develop a model with the synthetic data and the chosen algorithm
Output: A model with differential privacy mechanism. The model prediction even if includes noise should have the same trends as the original results would have.

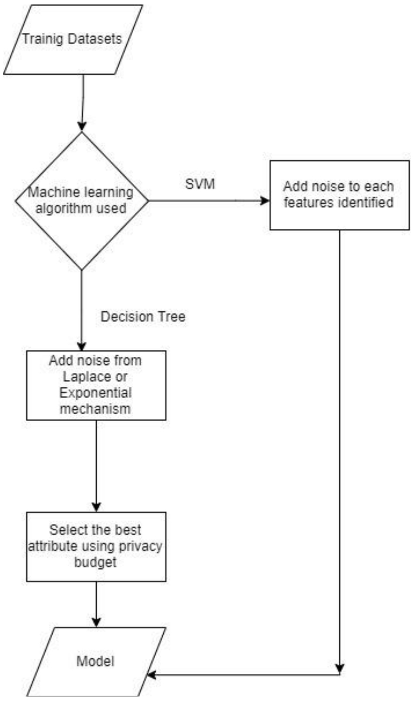
Trainig Datasets
Machine learning algorithm used
SVM
Add noise to each features identified
Decision Tree
Add noise from Laplace or Exponential mechanism
Select the best attribute using privacy budget
Model

Figure 2: Flowchart of the algorithm

Figure 2 is a flowchart of how differential privacy is used based on underlying machine learning algorithm

V. CONCLUSION

Due to the strictness of differential privacy it is often considered to nullify the purpose of a model. But with careful selection of parameters and algorithm it can balance privacy and utility. In this paper the three important algorithms often used are considered and how differential privacy mechanism is integrated with it is studied.

REFERENCES

[1] T. Zhu, G. Li, W. Zhou and P. S. Yu,(Aug. 2017) "Differentially Private Data Publishing and Analysis: A Survey," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 8, pp. 1619-1638, 1.

[2] J. Zhao, Y. Chen and W. Zhang, (2019)"Differential Privacy Preservation in Deep Learning: Challenges, Opportunities and Solutions," in *IEEE Access*, vol. 7, pp. 48901-48911.
[3] J. Soria-Comas, J. Domingo-Ferrer, D. Sánchez and D. Megías, (June 2017)"Individual Differential Privacy: A Utility-Preserving Formulation of Differential Privacy Guarantees," in *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1418-1429.
[4] Ji, Zhanglong & Lipton, Zachary & Elkan, Charles. (2014). Differential Privacy and Machine Learning: a Survey and Review
[5] X. Fang, F. Yu, G. Yang and Y. Qu, (,2019)"Regression Analysis With Differential Privacy Preserving," in IEEE Access, vol. 7, pp. 129353-129361.
[6] Y. Zhang, Z. Hao and S. Wang,(2019) "A Differential Privacy Support Vector Machine Classifier Based on Dual Variable Perturbation," in IEEE Access, vol. 7, pp. 98238-98251.
[7] S. Yao (2018),"An Improved Differential Privacy K-Means Algorithm Based on MapReduce," 2018 11th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China.
[8] J. Zhao,(2017) "Composition properties of Bayesian differential privacy," 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC.
[9] Yang, Yirong & Xia, Yi & Chi, Yun & Muntz, Richard. (2003). Learning naive Bayes classifier from noisy data.
[10] Wenru Tang, Yihui Zhou, Zhenqiang Wu, Laifeng Lu, and Mingshuang Li. (2019). Naive Bayes Classification based on Differential Privacy. In Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM 2019). Association for Computing Machinery, New York, NY, USA, Article 65, 1–6.
[11] Gong, M., Pan, K., & Xie, Y. (2019). Differential privacy preservation in regression analysis based on relevance. Knowledge-Based Systems.
[12] Bai, X., Yao, J., Yuan, M. *et al.(2017)* Embedding differential privacy in decision tree algorithm with different depths. *Sci. China Inf. Sci.* 60, 082104.