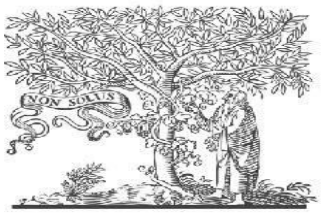
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

Link : https://ijiemr.org/downloads/Volume-12/Issue-05

10.48047/IJIEMR/V12/ISSUE05/16

Title **Fast Privacy-Preserving Text Classification Based on Secure Multiparty Computation**

Volume 12, Issue 05, Pages: 146-156

Paper Authors

Dr. N Swapna, P. Kruthika, . G. Manasa, G. Nithin, G. Deeraj reddy

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Fast Privacy-Preserving Text Classification Based on Secure Multiparty Computation

1. **Dr. N Swapna**, Associate professor & Head of the Department, MTech(Phd),Department of CSE, (Vijay Rural Engineering College(VREC)) swapnanaralas@gmail.com

2. **P. Kruthika**, BTech, Department of CSE, (Vijay Rural Engineering College(VREC)) heykruthika@gmail.com

3. **G. Manasa,** BTech, Department of CSE, (Vijay Rural Engineering College(VREC)) gangonimanasa@gmail.com

4. **G. Nithin**, BTech, Department of CSE, (Vijay Rural Engineering College(VREC)) nithin.gundela55@gmail.com

5. **G. Deeraj reddy**, BTech, Department of CSE, (Vijay Rural Engineering College(VREC)) deerajreddy082@gmail.com

ABSTRACT: We propose and use a privacy-preserving Naive Bayes classifier to the issue of private text classification. In this scenario, one side (Alice) is holding a text message, while another (Bob) is holding a classifier. Alice will only learn the outcome of the classifier applied to her text input at the conclusion of the protocol, whereas Bob will learn nothing. Secure Multiparty Computation is the foundation of our solution (SMC). Our Rust implementation offers a quick and safe solution for unstructured text categorization. In the event when Bob's model's dictionary size covers all words (n = 5200) and Alice's SMS comprises at most m = 160 unigrams, we can identify an SMS as spam or ham in less than 340 ms (the solution is general and may be applied in any other scenario in which the Naive Bayes classifier can be utilised). Our method takes just 21 ms for n = 369 and m = 8 (the average of a spam SMS in the database).

Keywords *– Privacy-Preserving Classification, Secure Multiparty Computation, Naive Bayes, Spam.*

1. INTRODUCTION

Classification is a supervised learning strategy in Machine Learning (ML) that aims to build a classifier using a collection of training data that includes class labels. Classification techniques include Decision Tree, Naive Bayes, Random Forest, Logistic Regression, and Support Vector Machines (SVM). Many issues may be solved using these methods, including: identifying an email/Short Message Service (SMS) as spam or

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

ham (not spam); diagnosing a medical condition (illness vs no sickness); hate speech detection; face classification; fingerprinting identification; and picture categorization. The classification in the first three instances above is binary, with just two class labels (yes or no), but the final three are multiclass, with more than two classes. Consider the following scenario: one party has the private data to be categorised, while the other party has a private model used to classify such data. In such a case, the party holding the data (Alice) is interested in obtaining the classification result of such data against a model held by a third party (Bob), so that at the end of the classification protocol, Alice only knows the input data and the classification result, and Bob knows nothing beyond the model itself. This is a highly significant situation. There are several instances when a data owner is unwilling to provide a piece of data that requires categorization (think of psychological or health related data). Furthermore, a machine learning model owner may not want to/cannot publish the model in the open due to intellectual property concerns or because the model provides information about the data set used to train it. As a result, both parties have enough motivation to engage in a protocol that provides shared functionality of secret categorization.

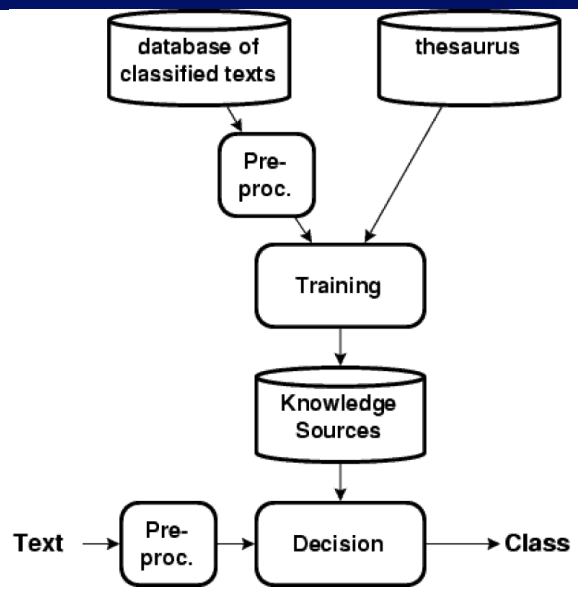


Fig.1: Example figure

Because of these issues, technologies like Secure Multiparty Computation (MPC), Differential Privacy (DP), and Homomorphic Encryption (HE) may be employed to provide privacy-preserving solutions. MPC enables two or more parties to jointly compute a function over their private inputs without giving any information to the other party, while HE is an encryption method that allows calculations to be performed on encrypted material without the need to decode it. Furthermore, DP is a method that adds random noise to queries in order to prevent an adversary from learning information about any specific person in the data collection. Our primary objective is to offer techniques for text categorization that respects privacy. We improve on prior Reich et al. findings by almost one

order of magnitude by carefully picking cryptographic engineering improvements, providing, to the best of our knowledge, the quickest text-classification results in the extant literature (21ms for an average sample of our data set). More specifically, we propose a privacy-preserving Naive Bayes classification (PPNBC) based on MPC in which we classify/predict an example given a trained model without revealing any additional information to the parties other than the classification result, which can be revealed to one specified party or both parties. We next apply our method to a text classification problem: determining whether SMS messages are spam or ham.

2. LITERATURE REVIEW

Privacy-Preserving Training of Tree Ensembles over Continuous Data:

Most known Secure Multi-Party Computation (MPC) algorithms for privacy-preserving decision tree training over dispersed data assume categorical characteristics. In real-world applications, characteristics are often numerical. To construct decision trees on data with continuous values, the traditional "in the clear" technique needs sorting of training instances for each feature in search of an ideal cut-point in the range of feature values in each node. Sorting is a costly process in MPC, thus creating safe methods that bypass this costly phase is an important topic in privacy-preserving machine learning. We propose three more efficient alternatives for secure training of decision tree-based models on data with continuous features in this paper: (1) secure discretization of the data followed by secure training of a decision tree over the discretized data; (2) secure discretization of the data followed by secure training of a random forest over the discretized data; and (3) secure training of extremely randomised trees ("extra-trees") on the original data. Both approaches (2) and (3) entail randomising feature selection. Furthermore, with technique (3), cut-points are generated at random, eliminating the requirement to sort or discretize the data beforehand. We used additive secret sharing-based MPC to execute all offered solutions in a semi-honest context. We empirically analysed and compared all offered techniques in terms of classification accuracy and runtime, in addition to mathematically establishing that they are valid and secure. We train tree ensembles confidentially across data sets with thousands of occurrences or features in a few minutes, with accuracies comparable to those found in the open. As a result, our strategy is more efficient than previous alternatives based on oblivious sorting.

Protecting privacy of users in brain-computer interface applications

Machine learning (ML) is transforming both science and industry. For training and inference, many ML applications depend on significant volumes of personal data. Electroencephalogram (EEG) data is one of the most intimately exploited data sources, a kind of data that is so rich in information that application developers may readily get knowledge beyond the declared scope from unprotected EEG signals, including passwords, ATM PINs, and other personal data. The problem we address is how to do meaningful ML using EEG data while maintaining users' privacy. As a result, we offer cryptographic algorithms based on Secure Multiparty Computation (SMC) to conduct linear regression over EEG data from many users in a completely privacy-preserving (PP) manner, i.e., without revealing each individual's EEG signals to anybody else. To demonstrate the power of our secure system, we demonstrate how it can estimate driver tiredness from EEG data, just as it would in the unencrypted situation, and at a very low computing cost. Our method is the first to use commodity-based SMC to EEG data, as well as the biggest reported experiment of secret sharing-based SMC in general, with 15 participants participating in all calculations.

QUOTIENT: Two-party secure neural network training and ´ prediction

A great deal of work has recently been spent to the development of secure protocols for machine learning activities. Much of this is targeted at making very accurate Deep Neural Network predictions more safe (DNNs). However, since DNNs are trained on data, one important concern is how such models may be learned safely. Prior research on safe DNN training has been on either constructing bespoke protocols for existing training algorithms or inventing customised training algorithms and then using general secure protocols. In this paper, we look at the benefits of creating training algorithms alongside an unique secure protocol, with improvements on both fronts. We offer QUOTIENT, a novel discretized training approach for DNNs, as well as a tailored secure two-party protocol for it. QUOTIENT combines essential components of cutting-edge DNN training, such as layer normalisation and adaptive gradient algorithms, and advances DNN training in two-party computing. In comparison to previous work, we achieve a 50X increase in WAN time and a 6% gain in absolute accuracy.

Privft: Private and fast text classification with homomorphic encryption

Because of the severity of privacy concerns and the necessity to comply with new privacy legislation, there is a greater interest than ever in privacy-preserving strategies that attempt to strike a balance between privacy and usefulness. Using Fully Homomorphic Encryption, we provide an effective approach for Text Classification while maintaining the privacy of the material (FHE). Our system (textbfPrivate textbfFast textbfText (PrivFT)) does two things: 1) utilising a plaintext model to infer encrypted user inputs, and 2) using an encrypted dataset to train an effective model. We train a supervised model for inference and propose a system for homomorphic inference on encrypted user inputs with zero loss of prediction accuracy. The second section demonstrates how to train a model using completely encrypted data to build an encrypted model. At different parameter settings, we give a GPU version of the Cheon-Kim-Kim-Song (CKKS) FHE method and compare it to current CPU implementations to achieve 1 to 2 orders of magnitude speedup. We use GPUs to develop PrivFT and obtain a run time per inference of less than 0.66 seconds. Training on a moderately big encrypted dataset requires more computing time, taking 5.04 days.

Contributions to the study of SMS spam filtering: new collection and results

The increase in mobile phone users has resulted in a tremendous increase in SMS spam messages. Fighting mobile phone spam is difficult in reality due to various variables, including the lower SMS rate, which has enabled many users and service providers to disregard the problem, and the restricted availability of mobile phone spam-filtering software. On the other hand, in academic contexts, a fundamental disadvantage is the lack of publicly available SMS spam datasets, which are critical for validating and comparing different classifiers. Furthermore, since SMS messages are very brief, content-based spam filters may suffer performance degradation. In this study, we provide the biggest genuine, public, and non-encoded SMS spam collection we are aware of. Furthermore, we examine the performance of many known machine learning approaches. The findings show that Support Vector Machine beats the other classifiers tested, and so it may be considered as a suitable baseline for future comparison.

3. METHODOLOGY

Because of these issues, technologies like Secure Multiparty Computation (MPC), Differential Privacy (DP), and Homomorphic Encryption (HE) may be employed to provide privacy-preserving solutions. MPC enables two or more parties to jointly compute a function over their

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

private inputs without giving any information to the other party, while HE is an encryption method that allows calculations to be performed on encrypted material without the need to decode it. Furthermore, DP is a method that adds random noise to queries in order to prevent an adversary from learning information about any specific person in the data collection.

Disadvantages:

1. A machine learning model owner may not want/cannot publish the model in the open due to intellectual property concerns, yet the model provides information about the data set used to train it.
2. Insecure

Our primary objective is to offer techniques for text categorization that respects privacy. We improve on prior Reich et al. findings by almost one order of magnitude by carefully picking cryptographic engineering improvements, providing, to the best of our knowledge, the quickest text-classification results in the extant literature (21ms for an average sample of our data set). More specifically, we propose a privacy-preserving Naive Bayes classification (PPNBC) based on MPC in which we classify/predict an example given a trained model without revealing any additional information to the parties other than the classification result, which can be revealed to one specified party or both parties. We next apply our method to a text classification problem: determining whether SMS messages are spam or ham.

Advantages:

1. Our Rust implementation offers a quick and safe solution for unstructured text categorization.
2. We classify/predict an example without providing the parties with any extra information other than the classification result.

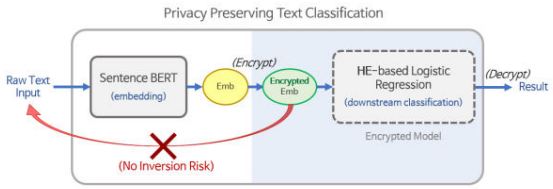


Fig.2: System architecture

MODULES:

To carry out the aforementioned project, we created the modules listed below.

- Data exploration: we will put data into the system using this module.
- Processing: we will read data for processing using this module.

- Splitting data into train and test: Using this module, data will be separated into train and test models.

- Making the model - Logistic Regression - Random Forest Classifier - Decision Tree - Support Vector Classifier - KNN - XGBoost - PPNB - Naive Bayes - Voting Classifier. Calculated algorithm accuracy.

- User registration and login: Using this module will result in registration and login.

- Using this module will provide input for prediction.

- Prediction: final predicted shown

4. IMPLEMENTATION

ALGORITHMS:

Logistic Regression: Logistic regression is a statistical analytic approach that uses past observations of a data set to predict a binary result, such as yes or no. A logistic regression model forecasts a dependent variable by examining the connection between one or more existing independent variables. A logistic regression, for example, might be used to forecast whether a political candidate will win or lose an election, or if a high school student would be accepted or not to a certain institution. These binary outcomes allow for simple choices between two options.

Random Forest Classifier: As the name indicates, a random forest is made up of a huge number of individual decision trees that work together as an ensemble. Each tree in the random forest produces a class prediction, and the class with the most votes becomes the prediction of our model.

Decision tree: A decision tree is a graph that employs a branching mechanism to show every potential result for a given input. Decision trees may be hand-drawn or generated using a graphics application or specialist software. When a group has to make a decision, decision trees may help concentrate the debate.

SVM: A support vector machine (SVM) is a supervised machine learning model that employs classification methods to solve two-group classification problems. They can classify fresh text after providing an SVM model with sets of labelled training data for each category.

KNN: The k-nearest neighbours method, often known as KNN or k-NN, is a non-parametric, supervised learning classifier that employs proximity to create classifications or predictions about an individual data point's grouping.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

XGBoost: XGBoost is a popular and efficient open-source gradient boosted trees solution. Gradient boosting is a supervised learning approach that combines the estimates of a collection of smaller, weaker models to try to correctly predict a target variable.

Naive Bayes: A probabilistic classifier, the Naive Bayes classification technique. It is based on probability models with high independence assumptions. The independence assumptions often have little effect on reality. As a result, they are seen as naïve.

Voting Classifier: Voting Classifier is a machine-learning method that Kagglers often employ to improve the performance of their model and move up the rank ladder. Voting Classifier may also be used to increase performance on real-world datasets, although it has significant limitations.

5. EXPERIMENTAL RESULTS

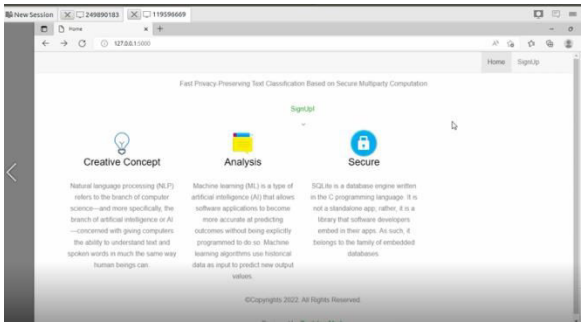


Fig.3: Home screen

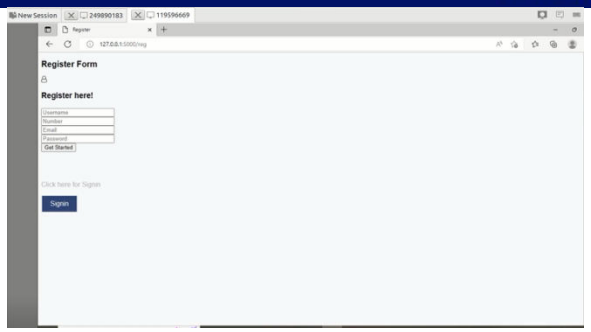


Fig.4: User signup & signin

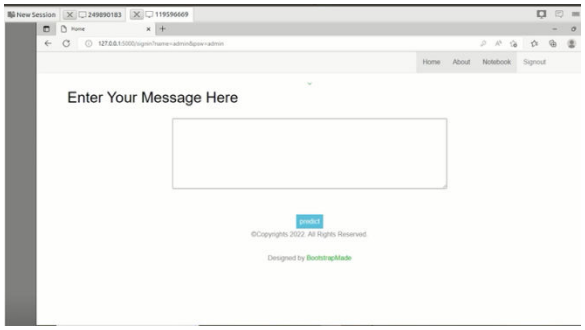


Fig.5: Main screen

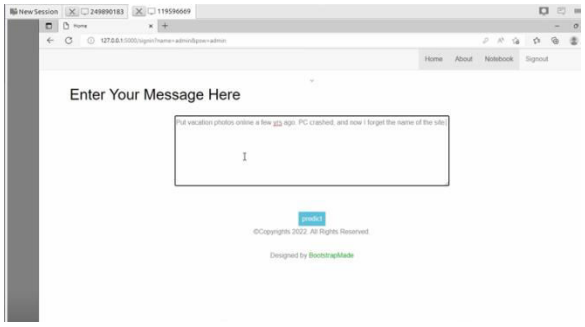


Fig.6: User input

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

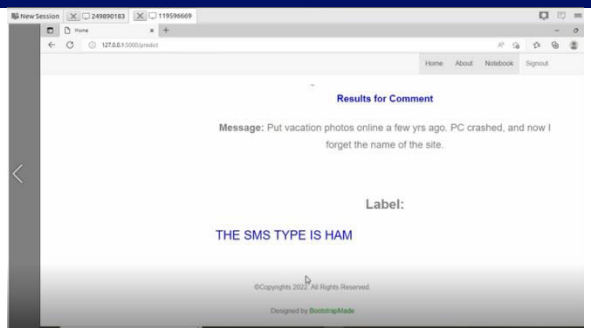


Fig.7: Prediction result

6. CONCLUSION

Privacy-preserving machine learning protocols are strong methods for performing operations on data while ensuring data privacy. We believe this is the first privacy-preserving Naive Bayes classifier with private feature extraction. There is no information provided about Bob's model (including which terms belong to the model) or the words in Alice's SMS. Our Rust implementation offers a quick and safe solution for unstructured text categorization. In the instance of spam detection, we can categorise an SMS as spam or ham in less than 340 ms if Bob's model's dictionary size covers all words (n = 5200) and Alice's SMS has no more than m = 160 unigrams. Our method takes just 21 ms for n = 369 and m = 8 (the average of a spam SMS in the database). Furthermore, the accuracy is almost identical to executing the Naive Bayes classification in the clear. It is vital to highlight that our solution may be utilised in any application that supports Naive Bayes. As a result, we think that our technique is applicable to the categorization of unstructured text while maintaining privacy. Our method is, to the best of our knowledge, the quickest SMC-based solution for private text categorization. Finally, we want to emphasise that anytime Alice is given the categorization result, she will learn something about Bob's model. This is inescapable, yet it does not violate our idea of security. Such a characteristic is, in fact, included in the ideal functionality used to define the security 14 of our proposed categorization protocol. To reduce such information leakage, add differential privacy to the model, such that Alice can never determine with confidence whether a word is in Bob's lexicon or not. This would reduce Alice's knowledge of Bob's lexicon while decreasing the model's accuracy. These are questions for the future.

REFERENCES

[1] Samuel Adams, Chaitali Choudhary, Martine De Cock, Rafael Dowsley, David Melanson, Anderson Nascimento, Davis Railsback, and Jianwei Shen. Privacy-Preserving Training of Tree Ensembles over Continuous Data. IACR ePrint 2021/754, 2021.

[2] Anisha Agarwal, Rafael Dowsley, Nicholas D. McKinney, Dongrui Wu, Chin-Teng Lin, Martine De Cock, and Anderson C. A.

Nascimento. Protecting privacy of users in brain-computer interface applications. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 27(8):1546–1555, Aug 2019.

[3] Nitin Agrawal, Ali Shahin Shamsabadi, Matt J. Kusner, and Adria` Gascon. QUOTIENT: Two-party secure neural network training and ´ prediction. In Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz, editors, ACM CCS 2019: 26th Conference on Computer and Communications Security, pages 1231–1247. ACM Press, November 11–15, 2019.

[4] Ahmad Al Badawi, Louie Hoang, Chan Fook Mun, Kim Laine, and Khin Mi Mi Aung. Privft: Private and fast text classification with homomorphic encryption. IEEE Access, 8:226544–226556, 2020.

[5] Tiago A. Almeida, Jose Mar ´ ´ıa Gomez Hidalgo, and Akebo Yamakami. ´ Contributions to the study of SMS spam filtering: new collection and results. In ACM Symposium on Document Engineering, pages 259–262. ACM, 2011.

[6] Boaz Barak, Ran Canetti, Jesper Buus Nielsen, and Rafael Pass. Universally composable protocols with relaxed set-up assumptions. In 45th Annual Symposium on Foundations of Computer Science, pages 186–195, Rome, Italy, October 17–19, 2004. IEEE Computer Society Press.

[7] Mauro Barni, Pierluigi Failla, Riccardo Lazzeretti, Ahmad-Reza Sadeghi, and Thomas Schneider. Privacy-Preserving ECG Classification With Branching Programs and Neural Networks. IEEE Trans. Information Forensics and Security, 6(2):452–468, 2011.

[8] Paulo S. L. M. Barreto, Bernardo David, Rafael Dowsley, Kirill Morozov, and Anderson C. A. Nascimento. A framework for efficient adaptively secure composable oblivious transfer in the ROM. Cryptology ePrint Archive, Report 2017/993, 2017. http://eprint.iacr.org/2017/993.

[9] Donald Beaver. Commodity-Based Cryptography (Extended Abstract). In STOC, pages 446–455. ACM, 1997.

[10] Raphael Bost, Raluca Ada Popa, Stephen Tu, and Shafi Goldwasser. Machine Learning Classification over Encrypted Data. In NDSS. The Internet Society, 2015.