Distributed one-class learning

Ali Shahin Shamsabadi, Hamed Haddadi, and Andrea Cavallaro

Published in: IEEE International Conference on Image Processing (ICIP) 2018

Centre for Intelligent Sensing
Queen Mary University of London
Outline

• Introduction to privacy in machine learning

• Centralised and distributed learning and their challenges

• Background in Autoencoder and One-Class classifier

• Machine learning solution for revenge pornography

• Proposed Distributed One-Class Learning

• Datasets, results and conclusion
Privacy in machine learning

- Training data, Parameters and Test data
- Training process with users’ collaboration

Centralised

Distributed\cite{1,2}
Challenges

• Users and service provider share
  – Data
  – Parameters

• Training data of each user has different
  – Size
  – Number of classes

• Scalability

• Complex distribution of users’ data (e.g. faces)
Background: Autoencoder\textsuperscript{[3]}

- Encoder-Decoder neural network
Background: One-Class Classifier [4]
Online photo-sharing social media

• Revenge pornography\textsuperscript{[5]}
  – Upload private images without consent

• \textit{How can prevent users from uploading privacy-sensitive images of other users?}

• Cloud-based Filter with users’ collaboration
  – Share permission or block uploading images

• Train blocking filter (\(N\)-class classifier)
  – Private-sensitive training data
    • \(\rightarrow\) Not centralized learning
  – Parameters contain sensitive information, each user one class
    • \(\rightarrow\) Not distributed learning
Distributed One-Class Learning

- $N$ users with $N$ private classes $\rightarrow$ $N$-class classifier
- Decompose $N$-class classifier to $N$ **one-class classifier**
- Distribute $N$ one-class classifiers (= Autoencoders)
- Train $N$ one-class autoencoders locally by users **independently**
- Upload parameters
- New uploaded image
  - Feed to filter
Dataset & accuracy private/non-private images

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$u_0$</th>
<th>$u_1$</th>
<th>$u_2$</th>
<th>$u_3$</th>
<th>$u_4$</th>
<th>$u_5$</th>
<th>$u_6$</th>
<th>$u_7$</th>
<th>$u_8$</th>
<th>$u_9$</th>
</tr>
</thead>
</table>

Per-class accuracy plots for IMDB, CIFAR-10, and MNIST datasets.

Acceptance rate plots for IMDB and MNIST datasets.
Threat & scalability

- **Adversary user:**
  - Access to data of victim user
  - Train one-class classifier with victim & adversary users data

- **Scalability:**
  - Impact of increasing number of user
Conclusion

DISTRIBUTED ONE-CLASS LEARNING

Ali Shahin Shamsabadi*, Hamed Haddadi†, Andrea Cavallaro*

*Queen Mary University of London, †Imperial College London

• Cloud-based Filter with users’ collaboration
  – Each user capture property of their class independently

• Training phase
  – Not uploading users’ data to cloud
  – Not sharing parameters among users
  – Each user data of one class

• Join new user at any time

Thank you!
References


