MULTI-KERNEL, DEEP NEURAL NETWORK AND HYBRID MODELS FOR PRIVACY PRESERVING MACHINE LEARNING

Mert Al[1], Thee Chanyaswad[2], and Sun-Yuan Kung[3]
Princeton University, Princeton, NJ, USA

Introduction and Motivation

- The advent of IOT and Big Data creates privacy concerns.
- The threat to privacy motivates the *Principle of Least Privilege* to be applied to Big Data.
- We consider the application of privacy preserving classification.
- We look for data representations that are helpful with the utility, but nothing else.
- We perform lossy compression in the private sphere, before the data is released.

Method

Our methodology combines two regimes; *Kernel Based Learning* and *Deep Learning.*

Step 1: Kernel Based Compression

We apply the utility maximizing lossy compression method called KDCA [1]. A KDCA projection can be derived via the optimization:

$$\mathbf{A}_{KDCA} = \underset{\mathbf{A}:\, \mathbf{A}^{\mathrm{T}}(\bar{K}^2+\rho\bar{K})\mathbf{A}}{\operatorname{argmax}} \operatorname{trace}\left(\mathbf{A}^{\mathrm{T}}\mathbf{K}_B\mathbf{A}\right)$$

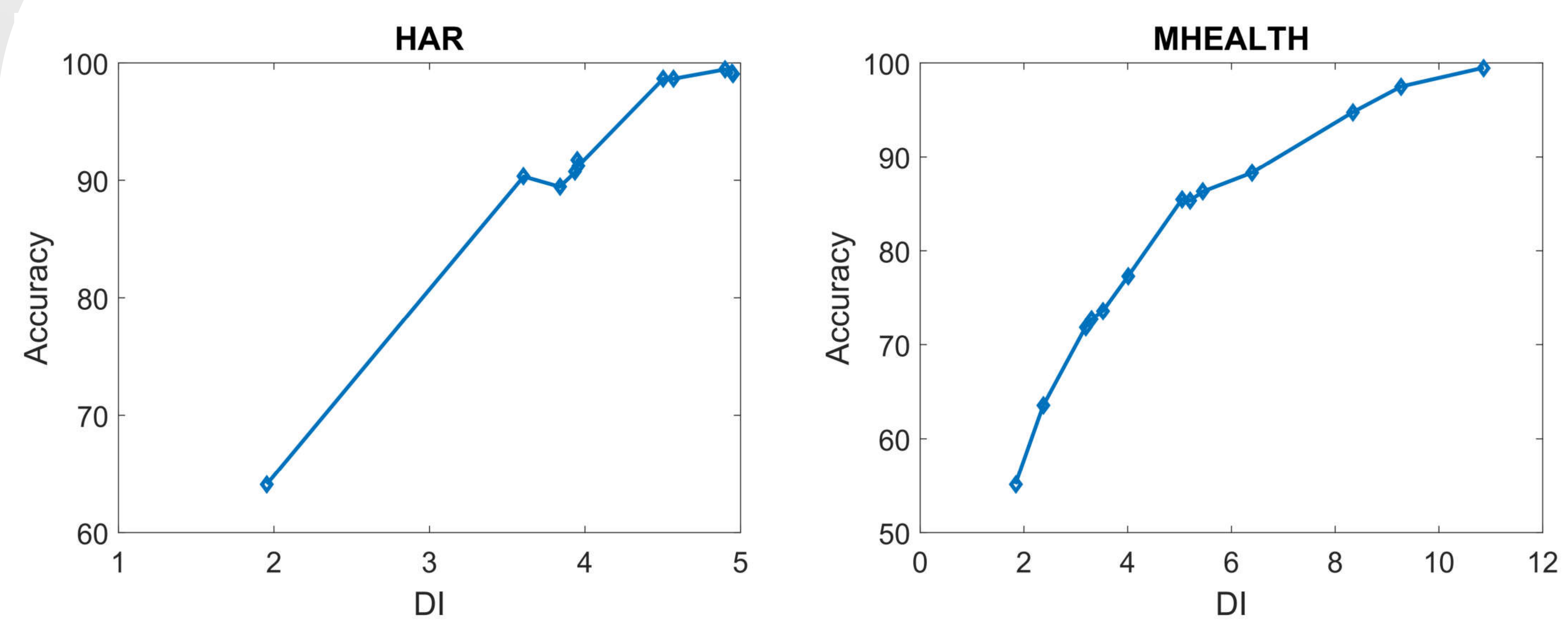
where $\bar{\mathbf{K}}$ is the centered kernel matrix and $\mathbf{K}_B$ is the kernel between-class scatter matrix.

The projection obtained from N training samples can then be applied to the data via the kernel trick:

$$\widehat{\boldsymbol{\Phi}} = \mathbf{A}^T\left(\mathbf{I} - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right)\mathbf{K}.$$

For classification with L classes, $L-1$ dimensional projections can capture all the discriminant power, allowing for a high compression rate.

Step 2: Kernel Selection



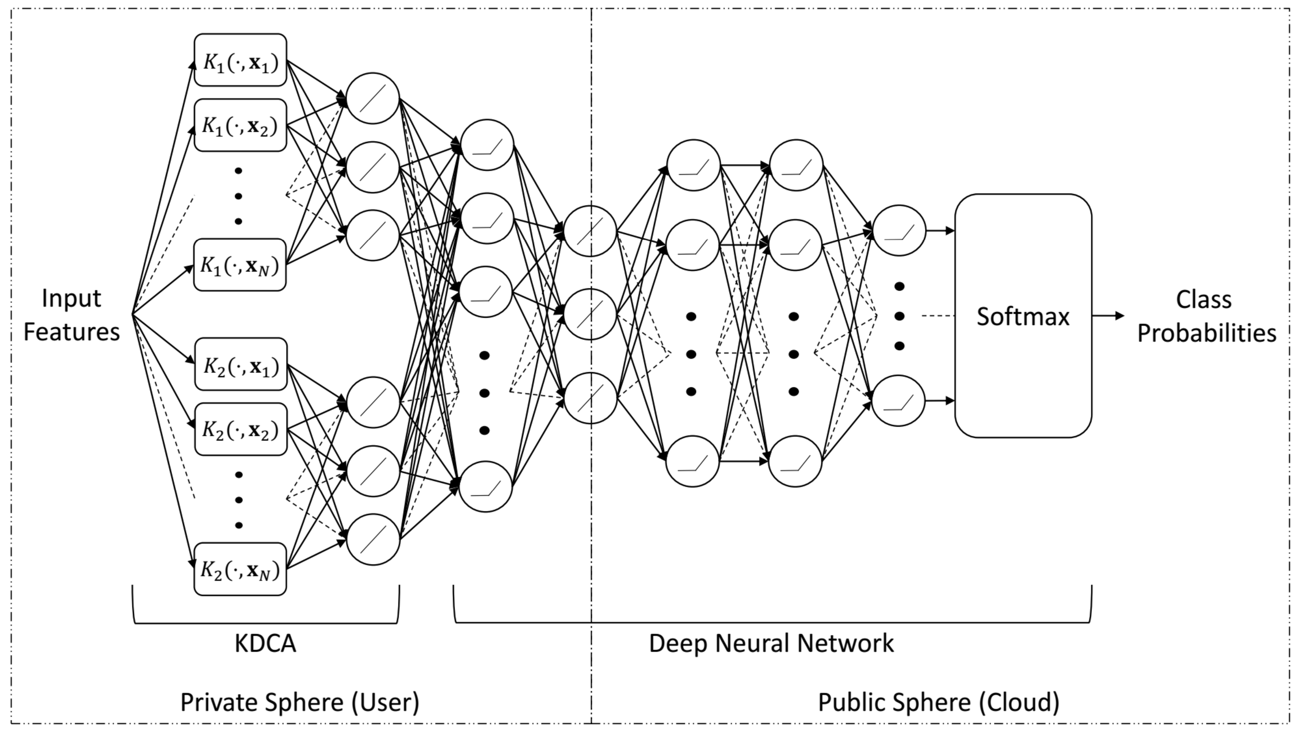
To select the best kernels for the utility, we perform a filtering procedure based on the ***Discriminant Information* (DI)** metric [2,3]:

$$DI = \operatorname{trace}\left((\bar{\mathbf{S}} + \rho\mathbf{I})^{-1}\mathbf{S}_B\right)$$

where $\bar{\mathbf{S}}$ and $\mathbf{S}_B$ are the centered and the between class scatter matrices.

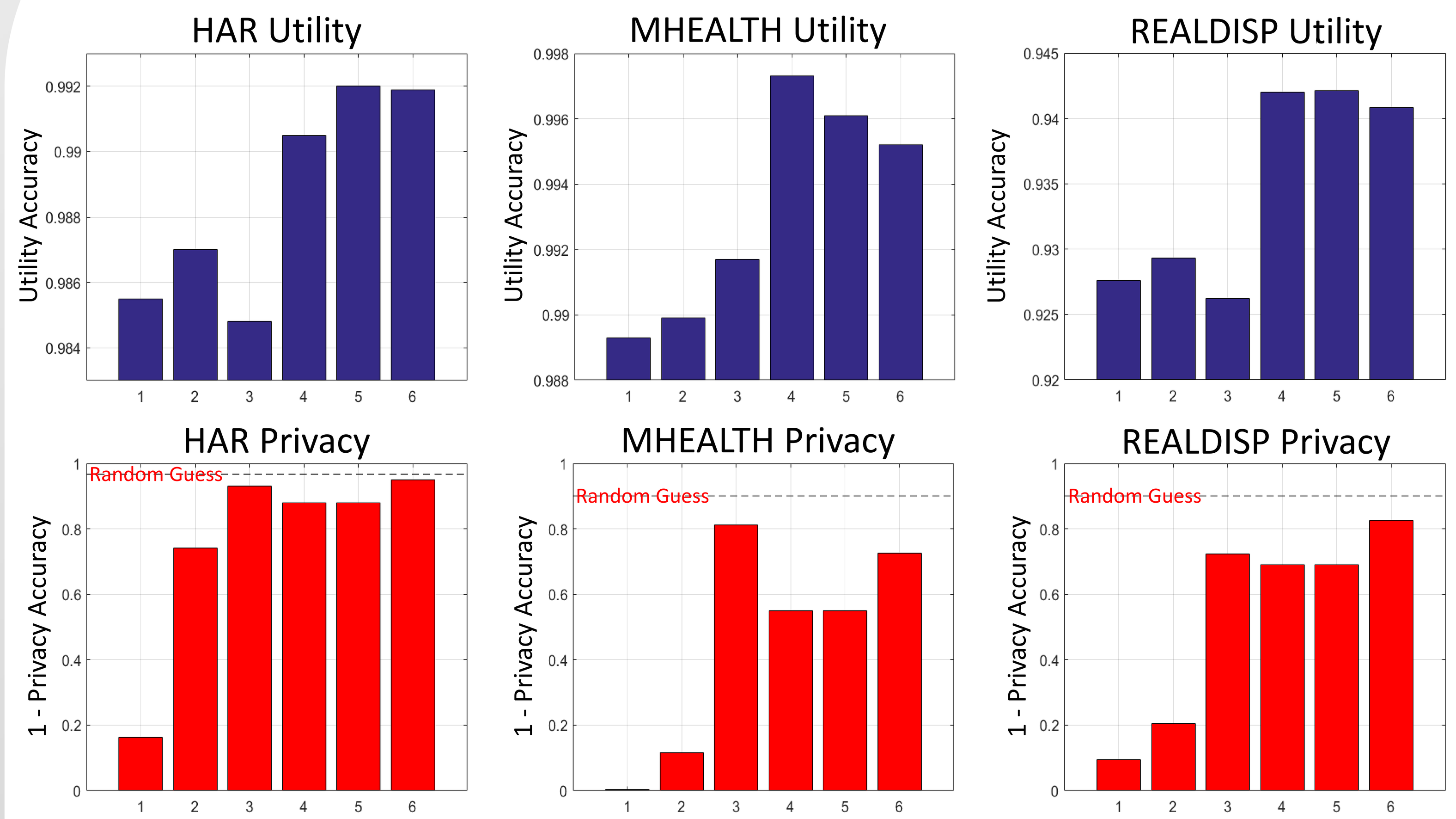
Since KDCA captures the utility information, combining KDCA projections with kernel selection has an effect of utility-maximizing ***space mining***.

Step 3 : Deep Learning Based Compression



To distill the utility information further, we utilize a DNN with a narrow, ***funneling layer***. The DNN processes multiple KDCA projections to form a ***Compressive Hybrid***.

Experimental Results



Methods: 1: DNN, 2: Compressive DNN, 3: Best Single KDCA, **4: Multi-KDCA**, **5: Multi-KDCA+DNN**, **6: Compressive Hybrid**

Conclusion

- Multi-KDCA with kernel selection achieves the best utility performances, demonstrating the importance of the ***space mining*** process.
- Multi-Kernel and Deep Learning based compression can effectively remove private information, while maintaining high utility.

References

1. T. Chanyaswad, J. M. Chang, and S. Y. Kung, "A compressive multi-kernel method for privacy-preserving machine learning," IJCNN 2017.
2. T. Chanyaswad, M. Al, J. M. Chang, and S. Y. Kung, "Differential mutual information forward search for multi-kernel discriminant-component selection with an application to privacy-preserving classification," MLSP 2017.
3. S. Y. Kung, "A compressive privacy approach to generalized information bottleneck and privacy funnel problems," Journal of the Franklin Institute, 2017.

Acknowledgements. This material is based on work supported in part by the Brandeis Program of the Defense Advanced Research Project Agency (DARPA) and Space and Naval Warfare System Center Pacific (SSC Pacific) under Contract No. 66001-15-C-4068. We thank Prof. J. Morris Chang and Prof. Pei-Yuan Wu for making this research possible.

Contact: [1]merta@princeton.edu, [2]tc7@princeton.edu, [3]kung@princeton.edu.