Performance Benchmarks for Detection Problems

Kelsie Larson (Purdue University, Sandia National Laboratories)
Mireille Boutin, PhD (Purdue University)

November 14, 2017

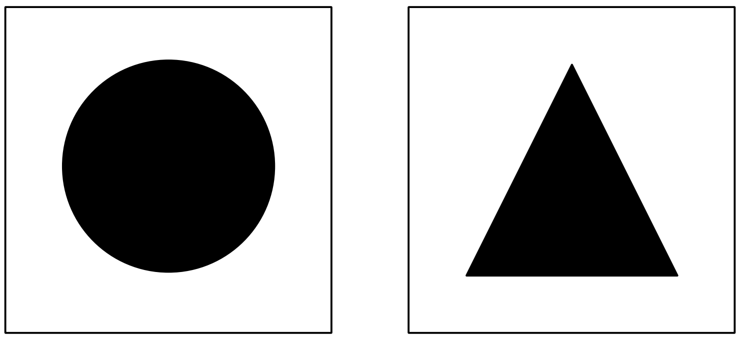
5th IEEE Global Conference on Signal and Information Processing (GlobalSIP 2017)
Montréal, Canada

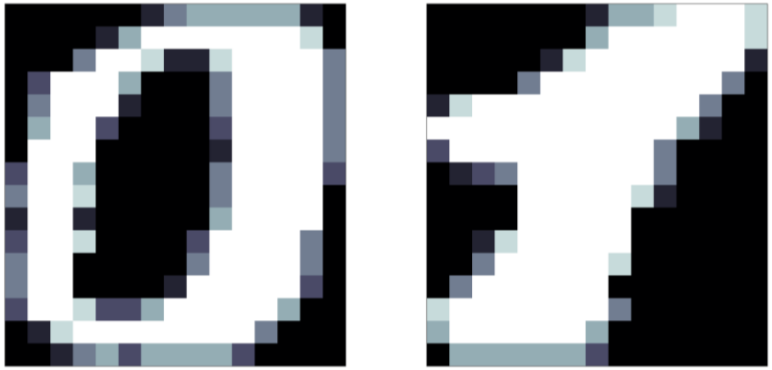
Introduction

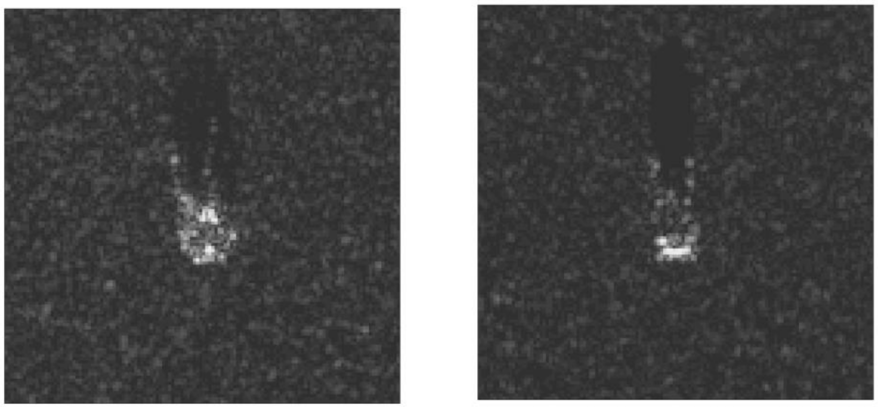
- Let's say you develop an object detection or classification algorithm. How do you determine success?
  - Download a data set
  - Train and test your algorithm
  - Compare your results to others'
  - OR build your own data set, define acceptable results, and test your algorithm
- Is your solution significant?

Problem

- Intuitive feel for difficulty of classification task







- Can we tell if a data set is “inherently separable?”

Previous Work

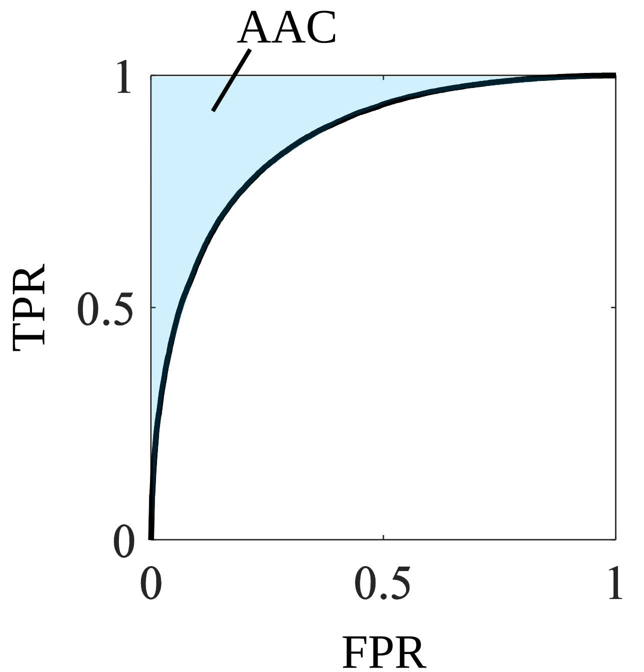
- Previous work suggests:
  - Real data has structure *(Han & Boutin, 2015)*
  - Random projections can reveal structure *(Kaski, 1998; Bingham & Mannila, 2001)*
- TARP (Thresholding After Random Projections) *(Yellamraju et. al, 2015)*
  - Use series of random projections to develop benchmarks

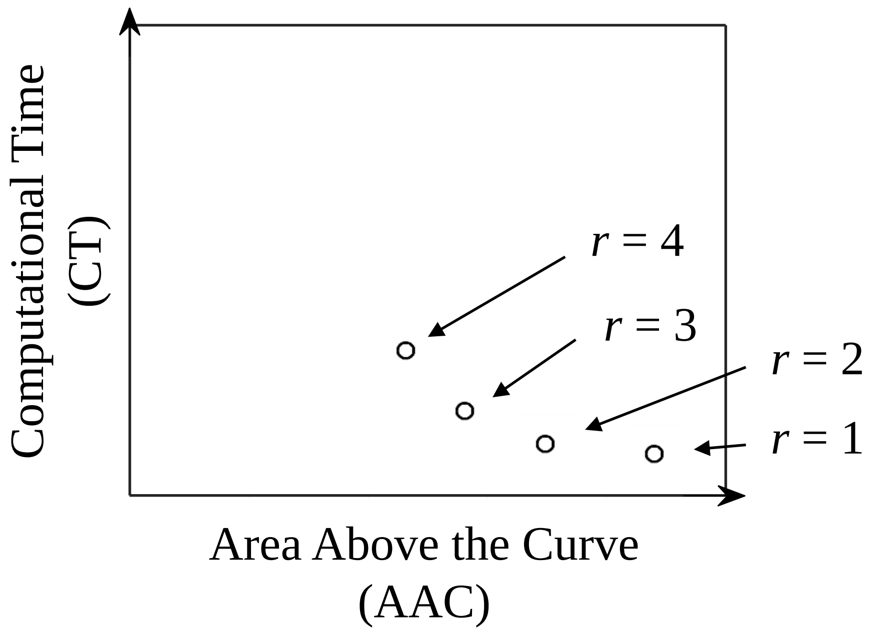
Proposed Solution

- TARP (Thresholding After Random Projections)
  - Randomly project data to 1-D r times
  - Classify with sliding threshold
  - Build ROC curve
  - Find the "best" projection – the one with the lowest AAC (area above the ROC curve)
  - Measure elapsed time (complexity)
  - Repeat for $r = 1, 2, 3, \ldots$

Proposed Solution

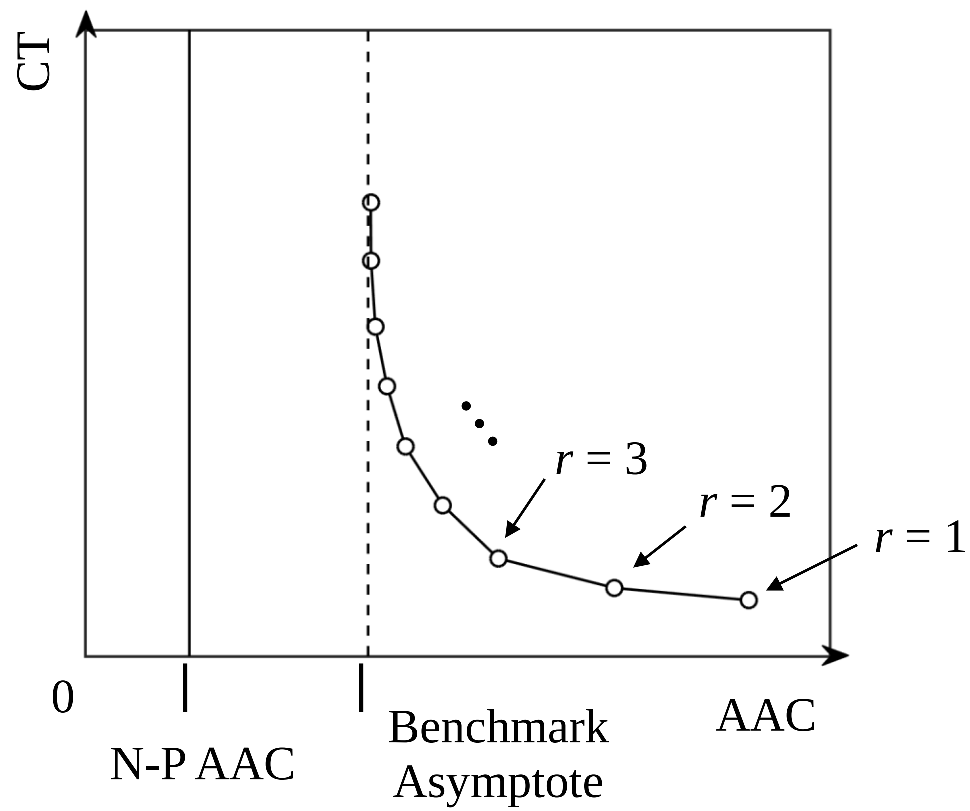
- Use expected results as series of benchmarks
  - Threshold average approximates expected best ROC curve
  - The AAC of the expected best ROC vs. the expected elapsed computational time (CT) for each *r* is considered a benchmark





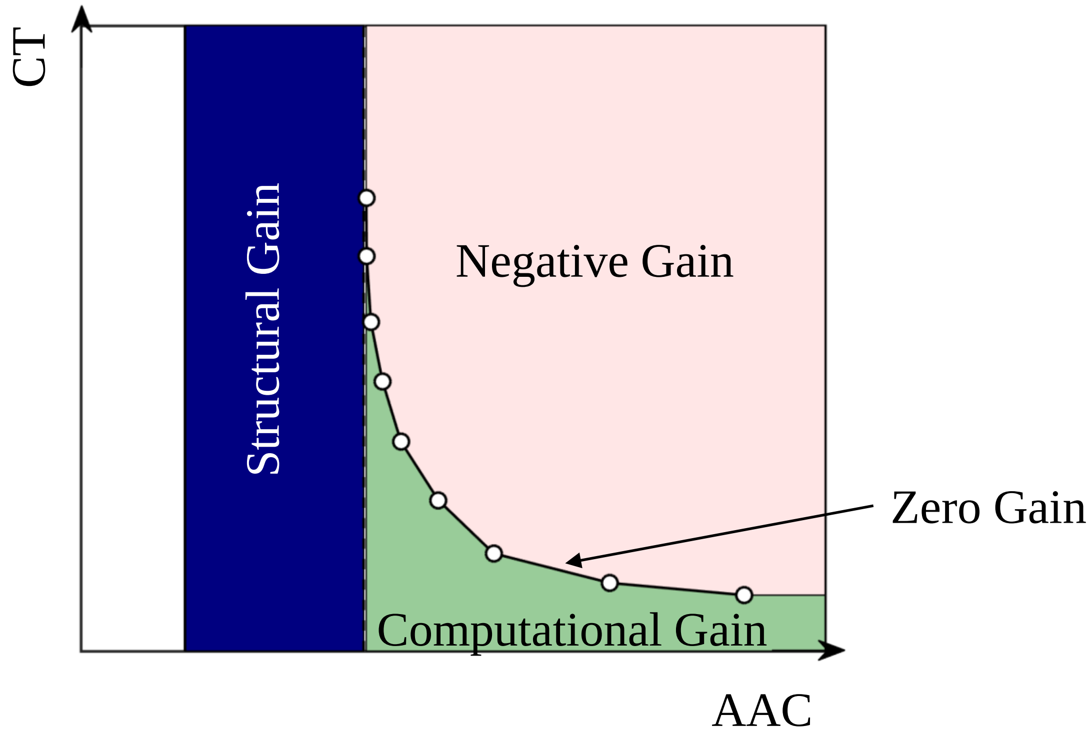
Proposed Solution

- Plotting the benchmarks yields a curve on AAC-CT plane:
  - Neyman-Pearson (N-P) test for theoretical maximum separability



Proposed Solution

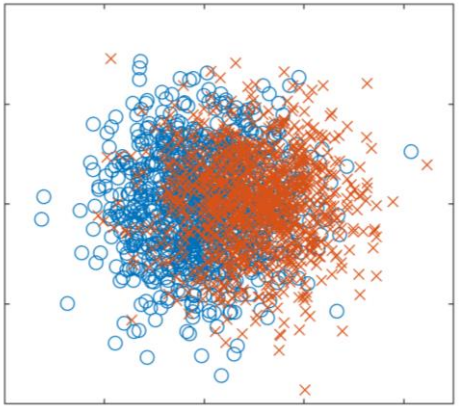
- The AAC-CT space is divided into regions
  - The regions characterize other detection methods



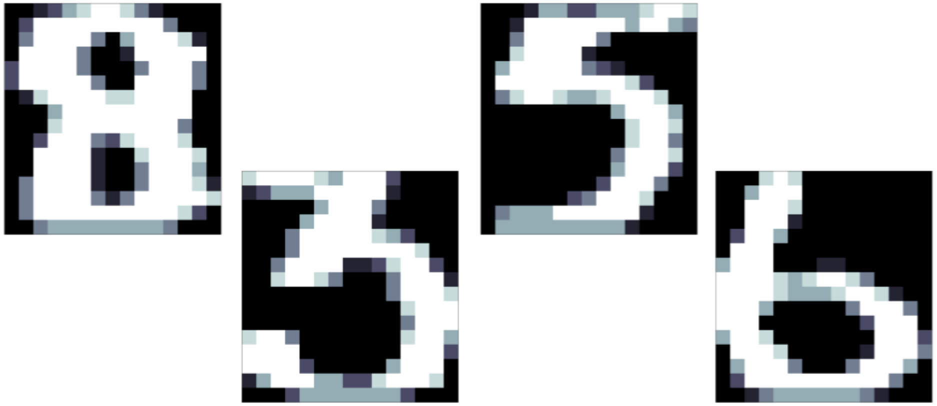
Experimental Results

- Selected results from different data sets:

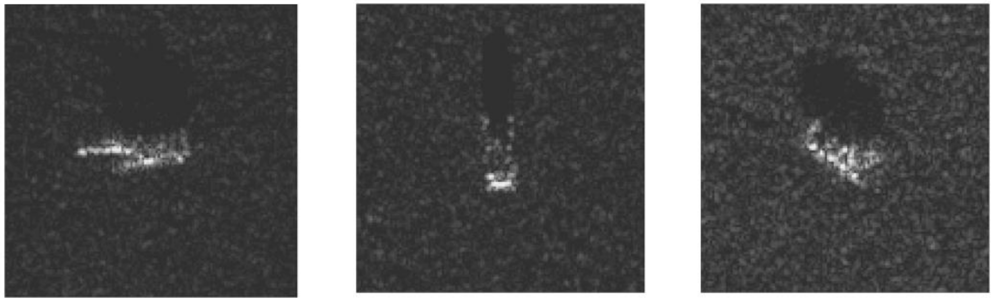
Synthetic 2-D normal



MFEAT handwritten digits *(Duin, 1998)*

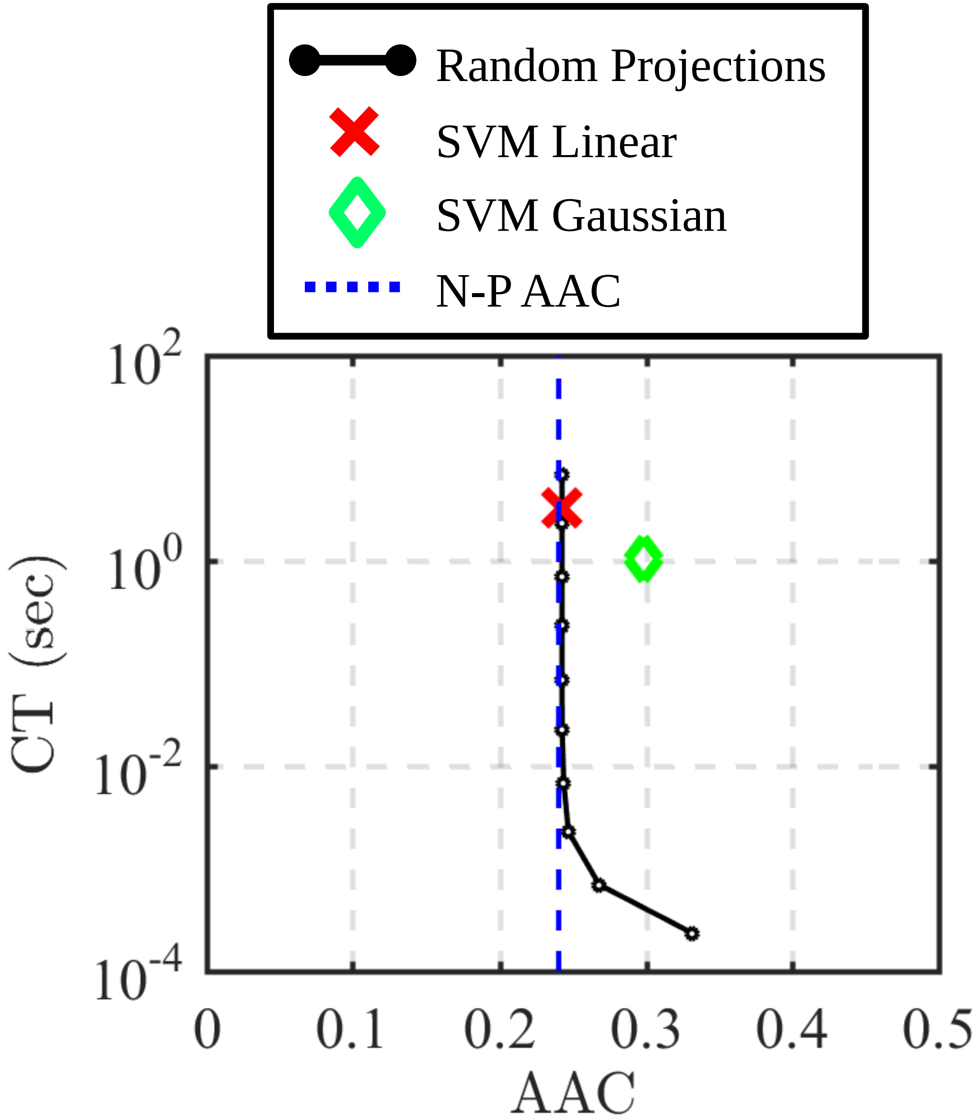


MSTAR SAR (radar) targets *(SDMS, 1995)*



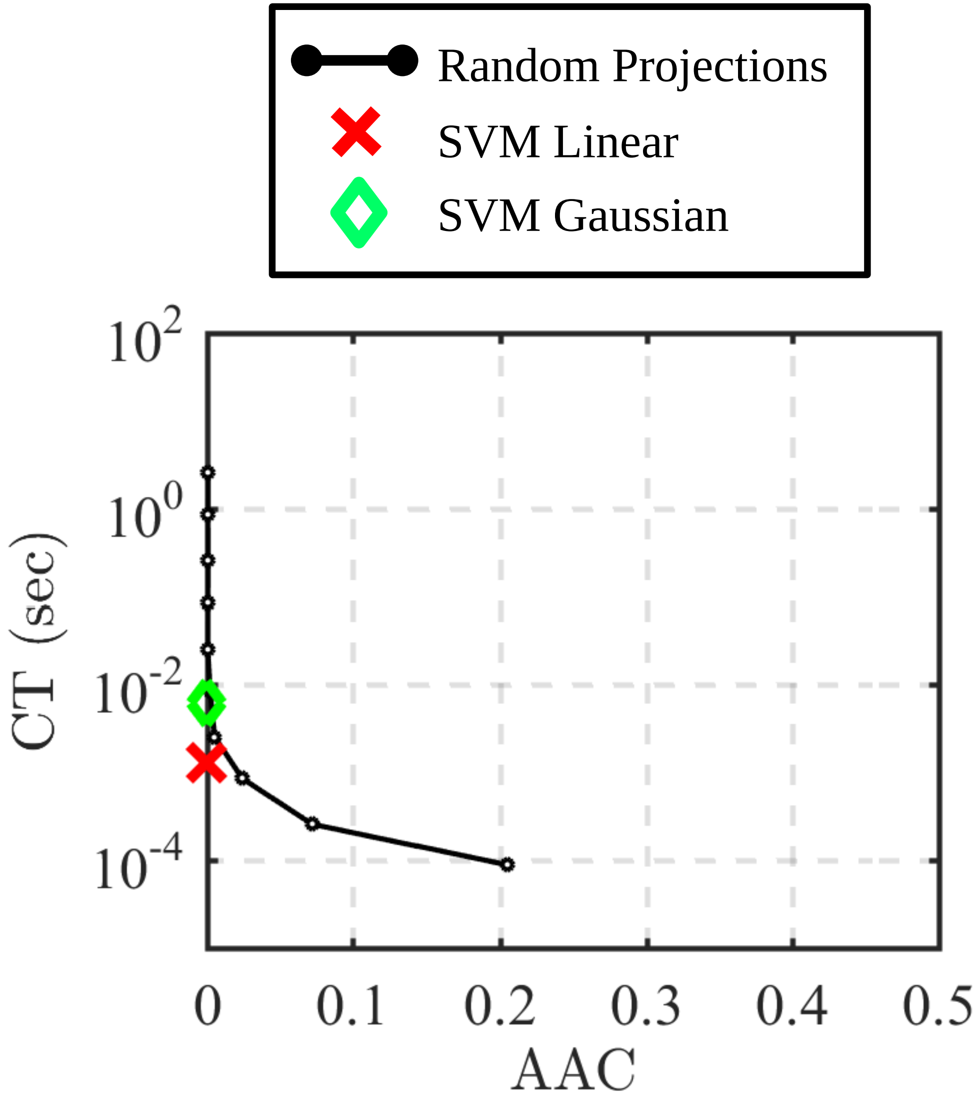
Experimental Results

- Synthetic 2-D normal
  - Covariance I
  - Class means (0, 0) and (0, 1)
  - 4,000 samples



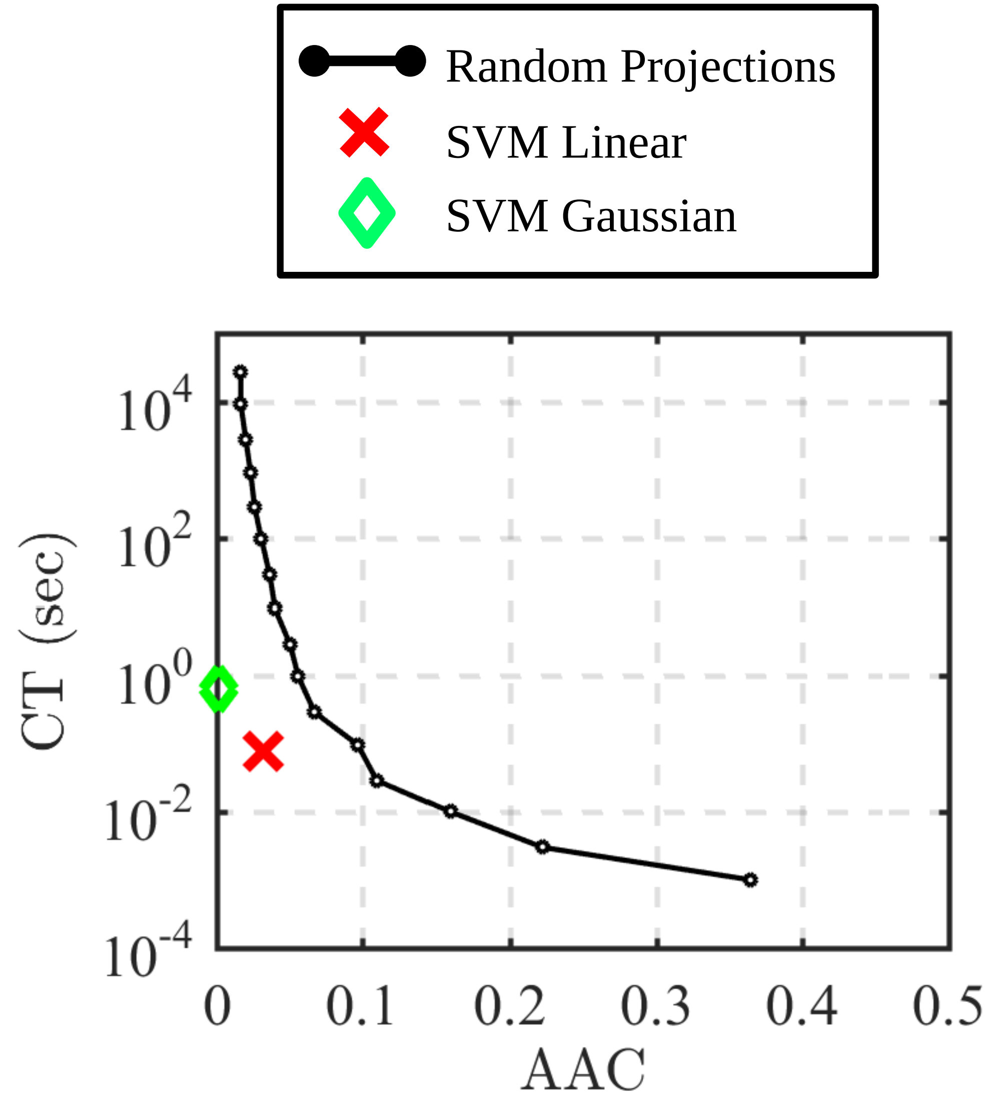
Experimental Results

- MFEAT 0 vs. 1
  - Fourier coefficients
  - 76-D
  - 400 samples



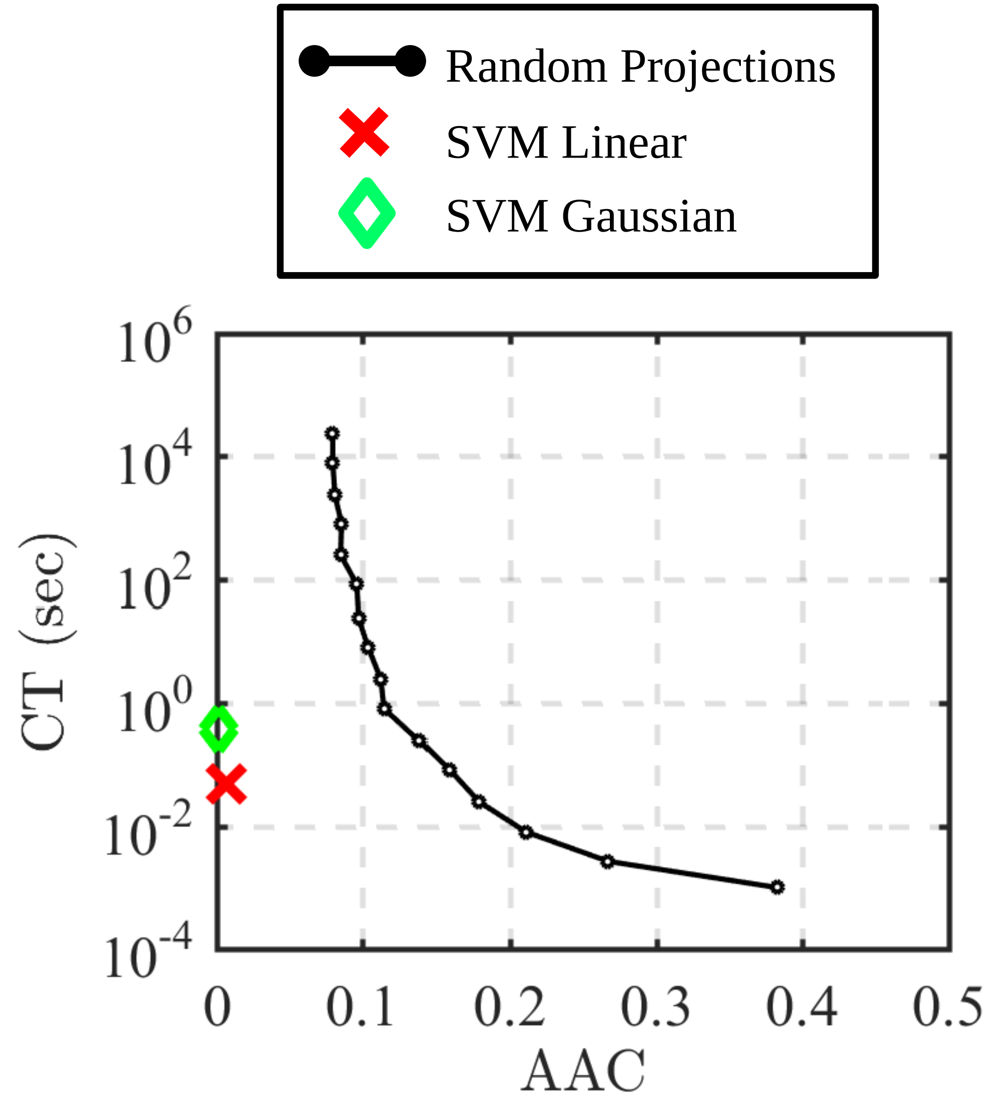
Experimental Results

- MFEAT even vs. odd
  - Profile correlations
  - 216-D
  - 2,000 samples



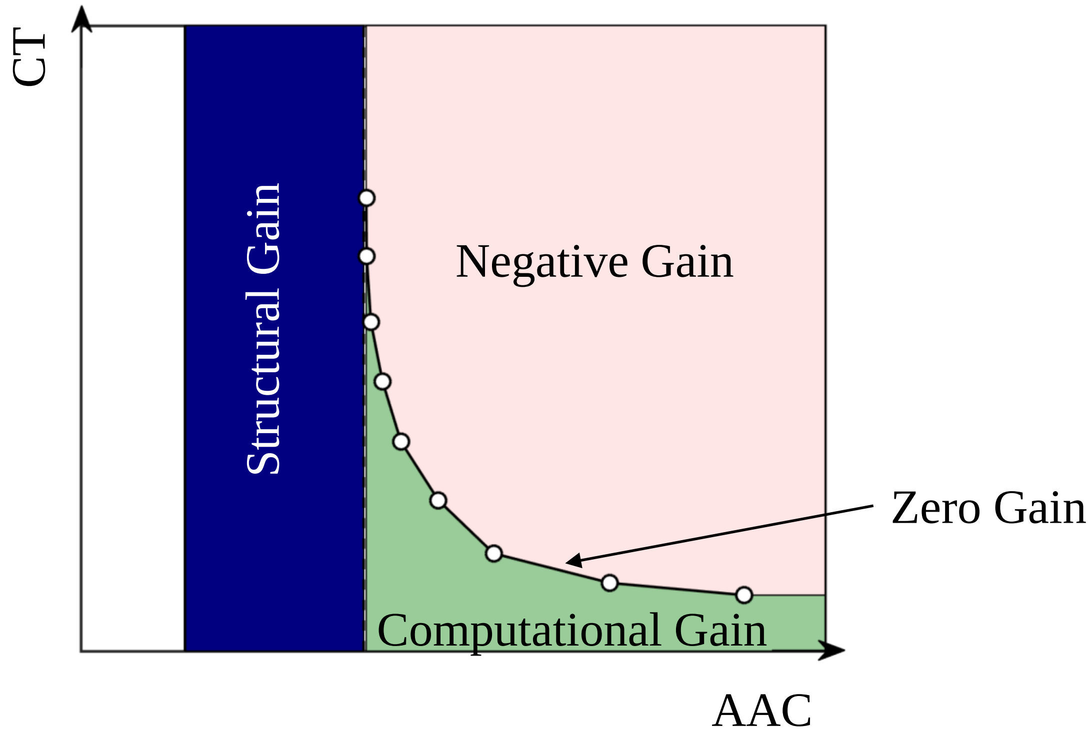
Experimental Results

- MSTAR BTR70 vs. T72
  - PCA coefficients
  - 358-D
  - 1,556 samples
    (392 BTR70/1,164 T72)



Conclusion

- Detection problems have different difficulties
- Investigated benchmark curve



References

E. Bingham and H. Mannila, “Random projection in dimensionality reduction: applications to image and text data,” in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, August 2001, pp. 245-250.

R. Duin, “Multiple Features Data Set,” Delft University of Technology, Department of Applied Physics, 1998. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Multiple+Features

S. Han and M. Boutin, “The Hidden Structure of Image Datasets,” 2015.

S. Kaski, “Dimensionality Reduction by Random Mapping: Fast Similarity Computation for Clustering,” in *Proceedings of ICJNN'98, 1998 IEEE International Joint Conference on Neural Networks*, May 1998.

M. Lichman, “UCI Machine Learning Repository,” University of California, Irvine, School of Information and Computer Sciences, 2013. [Online]. Available: http://archive.ics.uci.edu/ml

“MSTAR Public Targets,” SDMS, 1995. [Online]. Available: www.sdms.afrl.af.mil/index.php?collection=mstar&page=targets

T. Yellamraju, J. Hepp, and M. Boutin, “Benchmarks for Image Classification and Other High-Dimensional Pattern Recognition Problems,” Submitted, 2017.

Acknowledgements

I'd like to thank Sandia National Laboratories for their support.

For more information:

K. Larson and M. Boutin, "Performance Benchmarks for Detection Problems," in *Proceedings of the 2017 IEEE Global Conference on Signal and Information Processing*, November 2017.

Contact:

Kelsie Larson
k.mich.lar@gmail.com

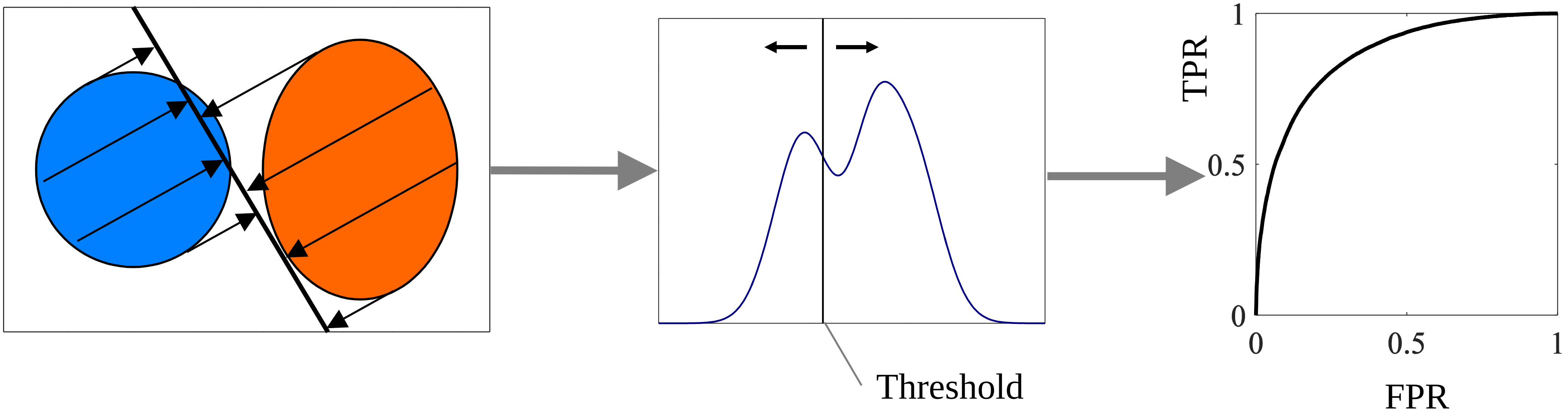
Extra Slides

Demonstration for $r = 2$

Extra Slides

r = 2 demo

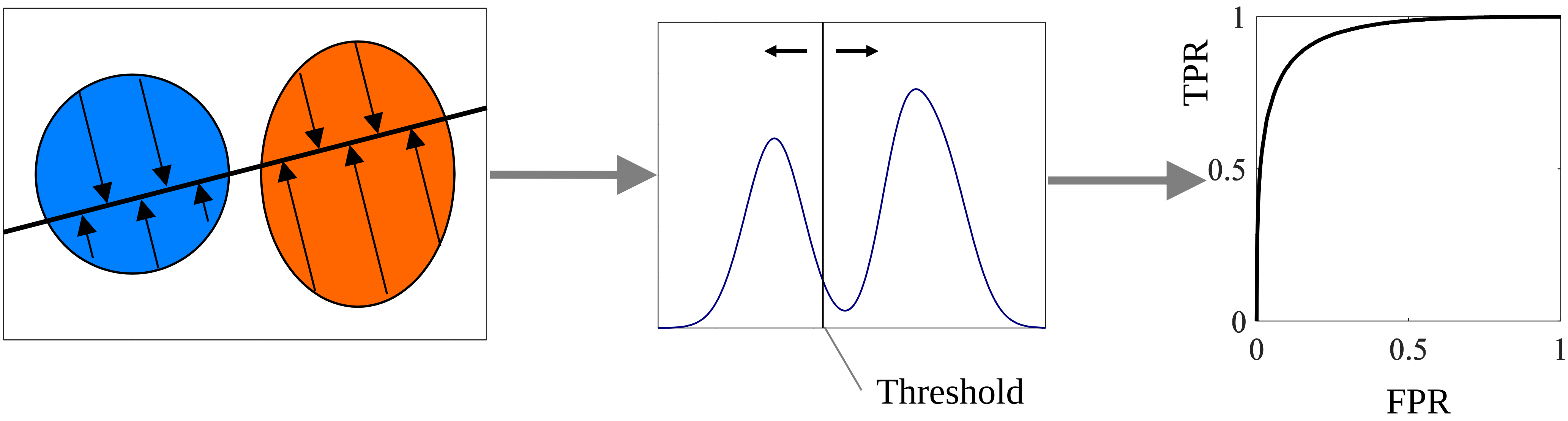
First random vector projection



Extra Slides

r = 2 demo

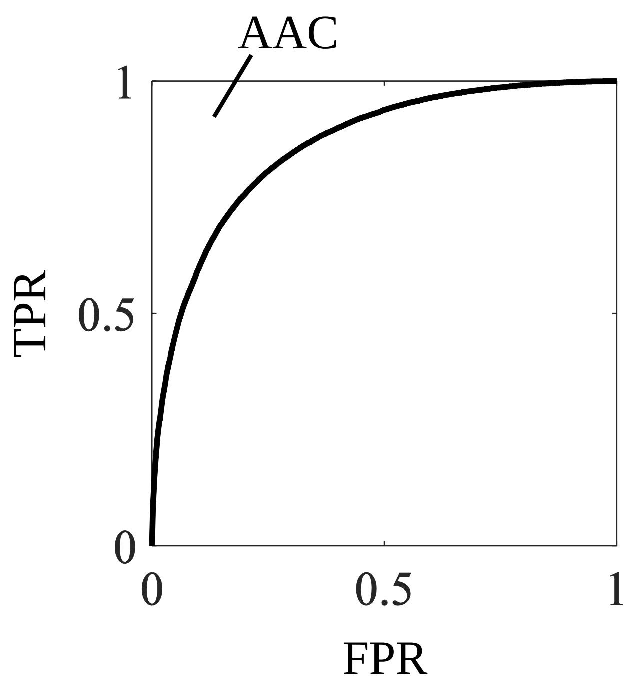
Second random vector projection

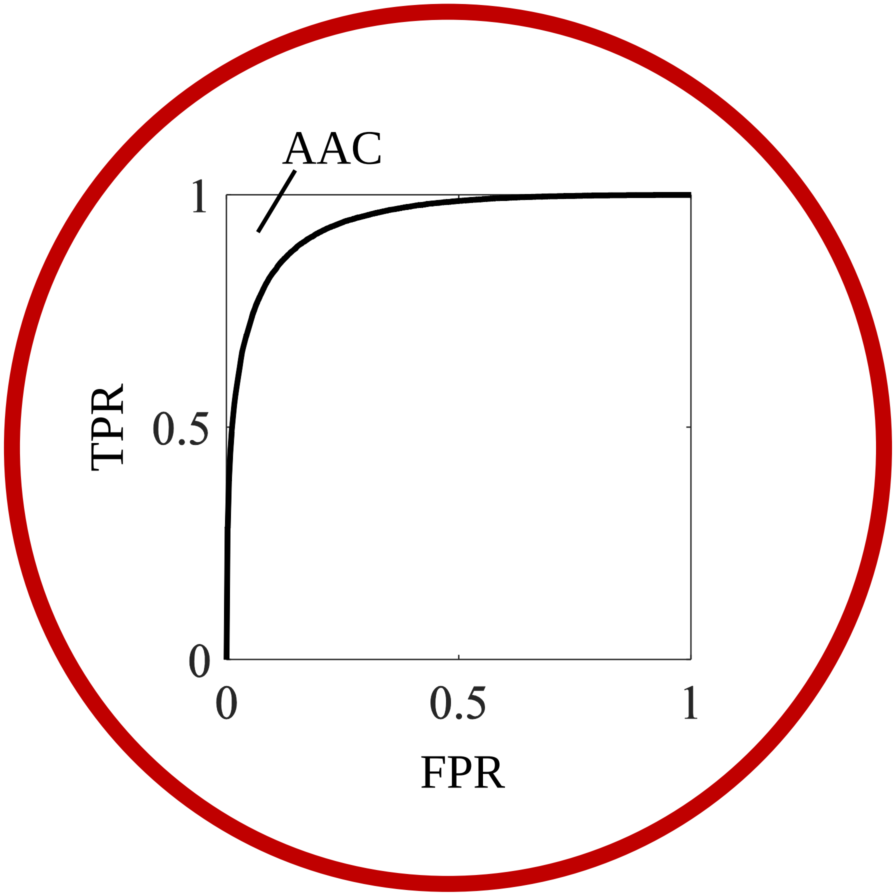


Extra Slides

r = 2 demo

Choose the best ROC curve out of 2

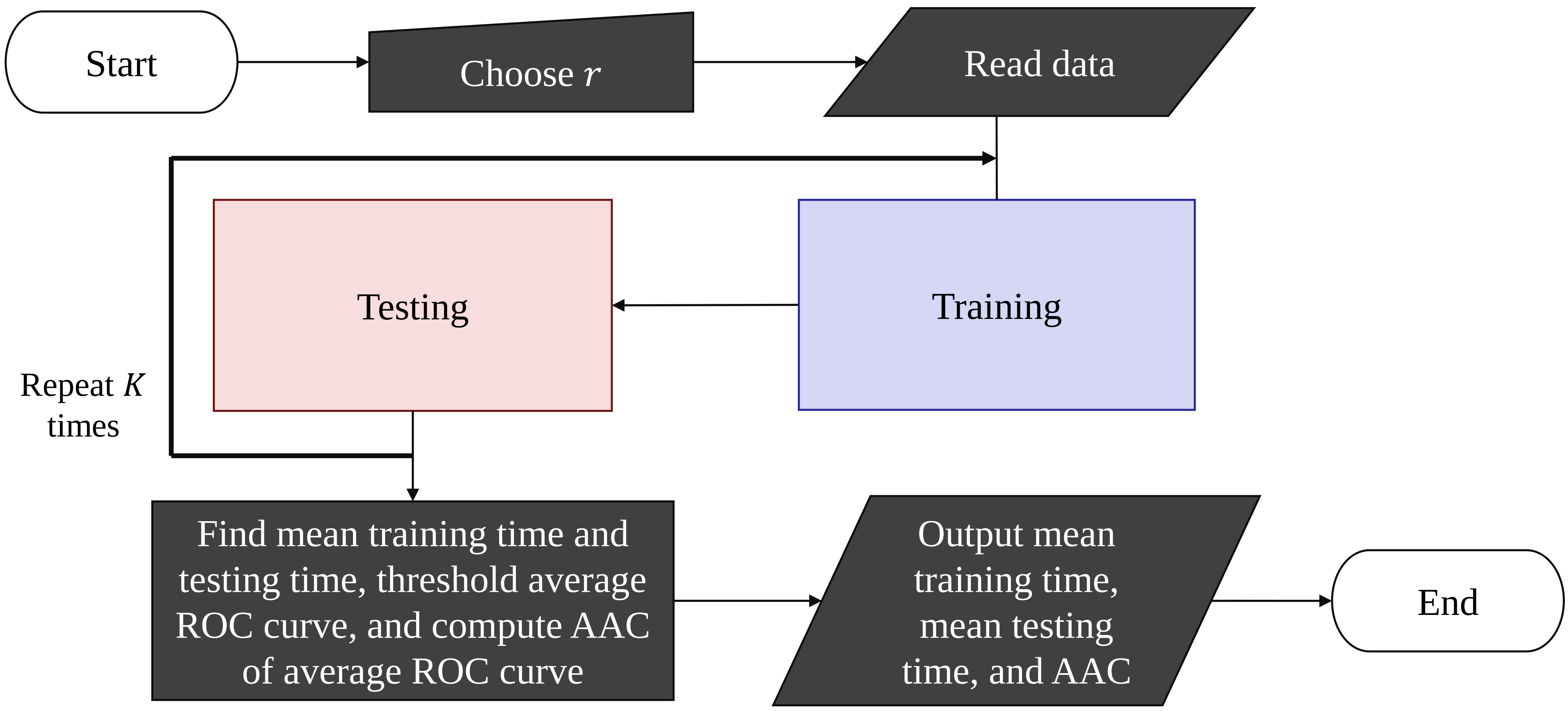




Extra Slides

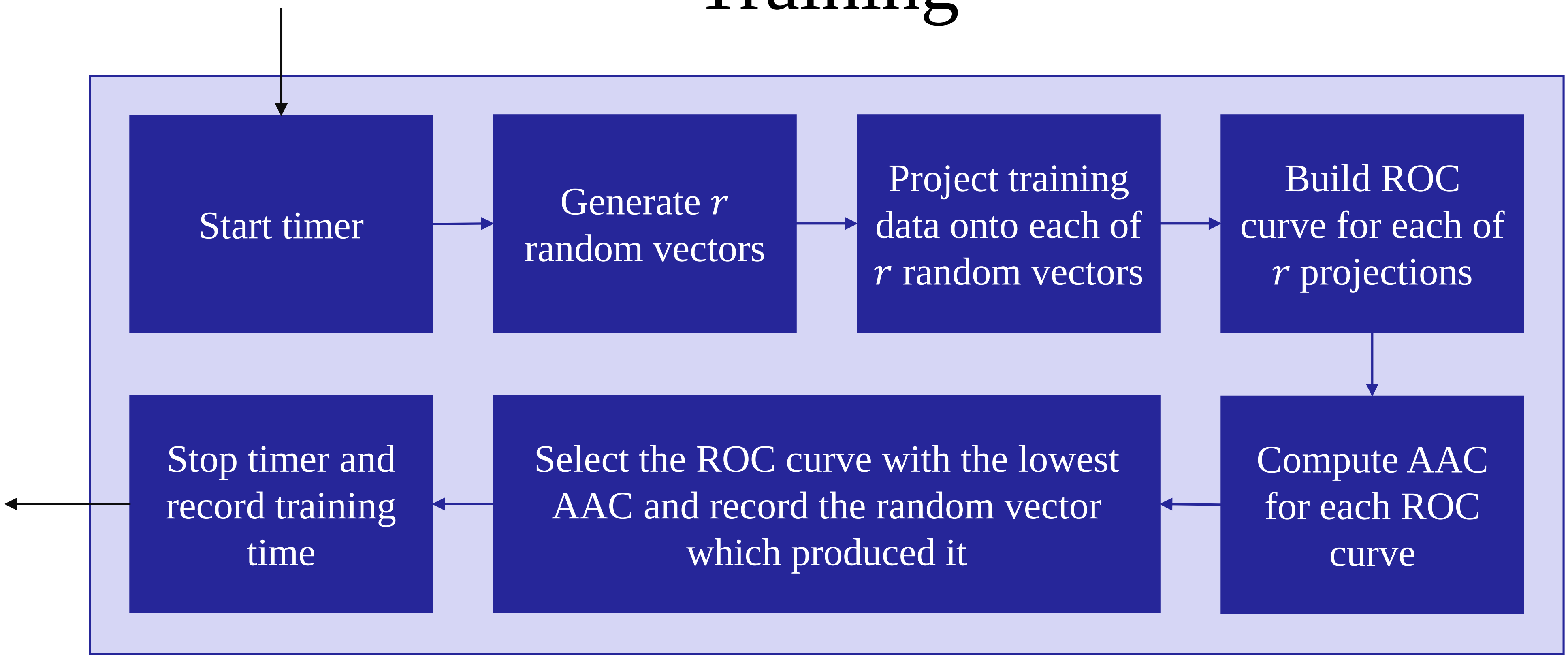
Algorithm Flowchart

Extra Slides



Extra Slides

Training



Extra Slides

Testing