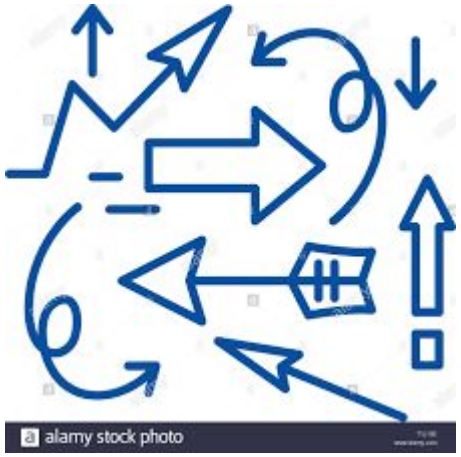
LEHIGH UNIVERSITY

When There is no Clear Answer; A Case Study of Classification Comparison

Amin Hosseiny Marani
CSE 447; Spring 2020

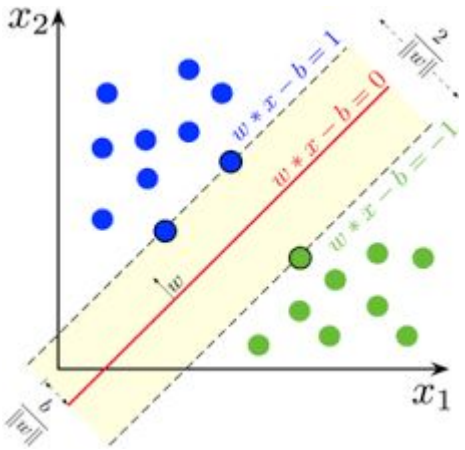
Motivation

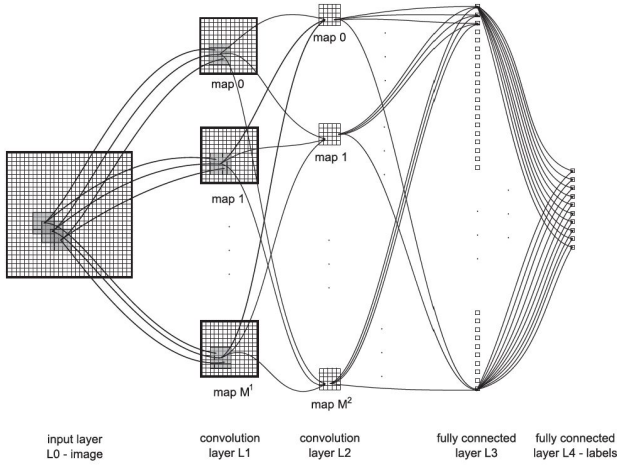
- Classification differs from model to model
  - Parameter selection
  - Different dataset
  - Random initialization
  - Train and test split

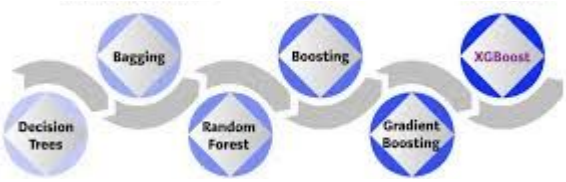


Methods

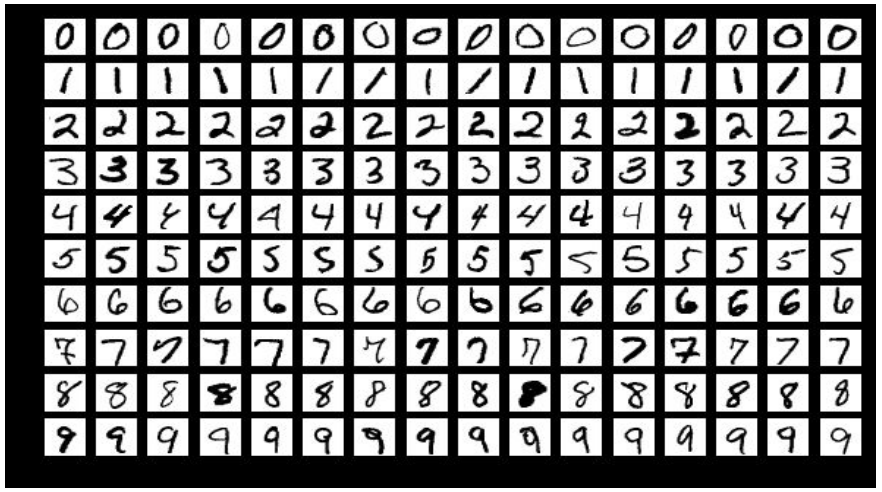
- Support Vector Machine
  - Easy to use
  - Fast to Train
- Convolutional Neural Networks
  - Flexible to new samples
  - Powerful for sparse learning
- XGBoost
  - Robust to noise
  - High Generalizability



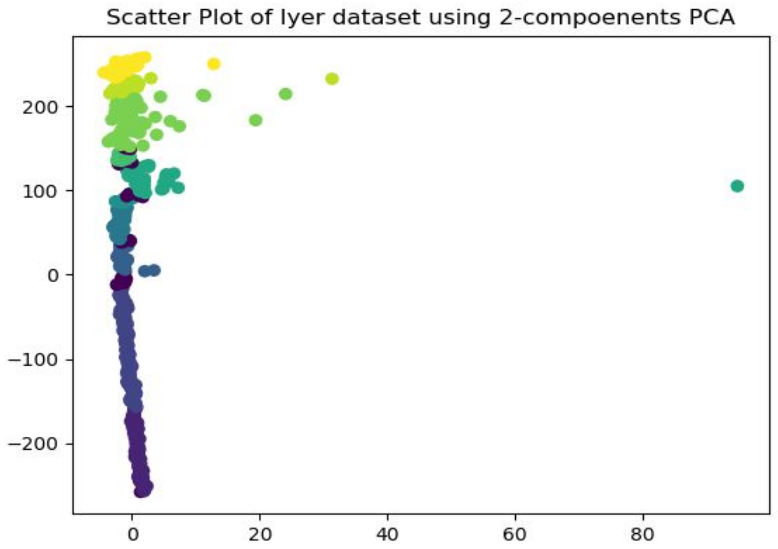




Data



- MNIST
  - Hand-written digits
  - 6000 train & 2000 test
- Iyer
  - Continues valued features
  - Highly imbalanced classes



Number of classes	samples	Samples per class	Lowest number of sample	Highest number of samples
10 (11 with noise)	517	47.00 (SD=39.51)	145 (class 2)	7 (class 5)

Table 1. Details of Iyer dataset.

SVM Performance

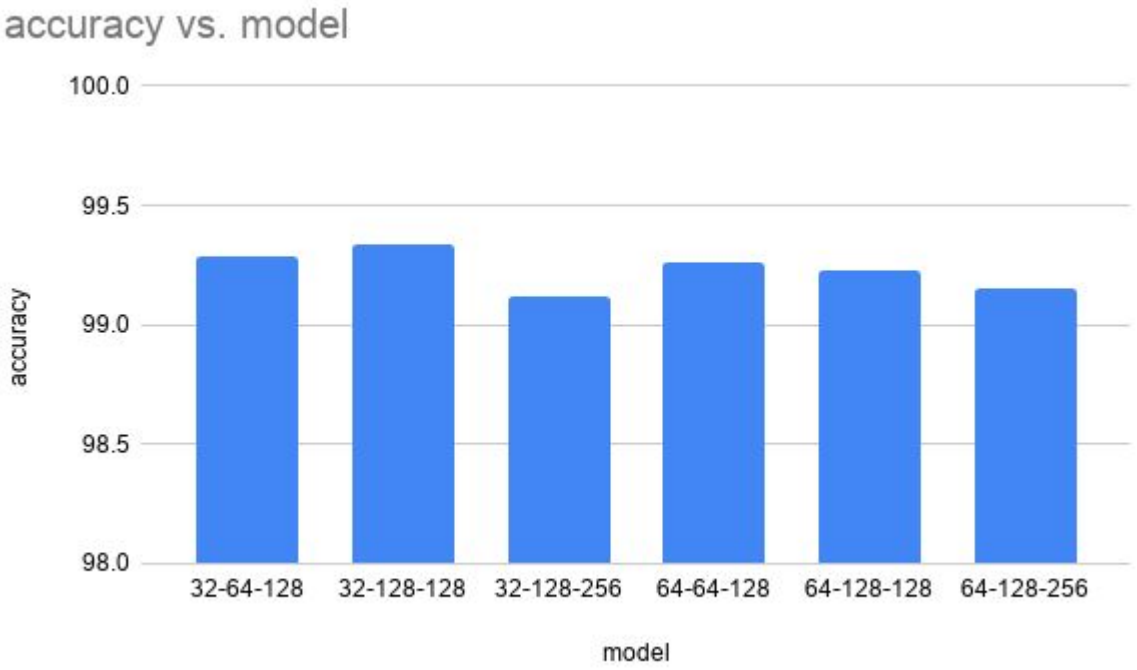
- Different datasets, different kernels, different performance!

Dataset	Sigmoid	RBF	Polynomial (deg. 4)
MNIST	77.59	**97.92**	96.98
Iyer	47.43	76.28	**89.74**

Table 2. SVM classification accuracy over two datasets with different kernels.

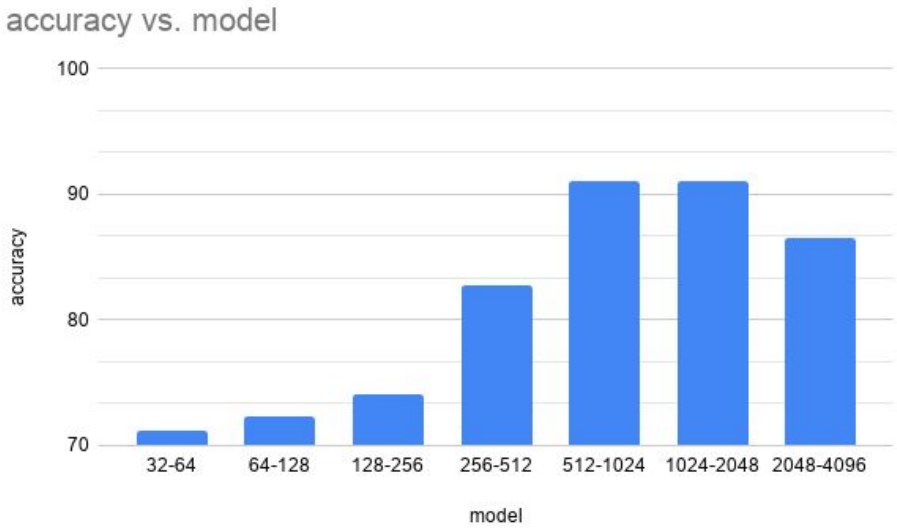
CNN Performance across MNIST

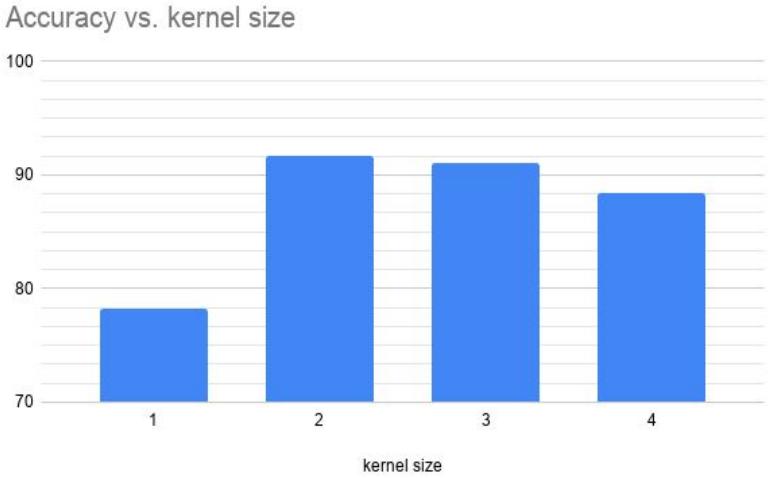
- Changing number of layers, number of filters, and number of units won't improve much!



CNN Performance across lyer

- Conv1D layer
- Highly dense layer





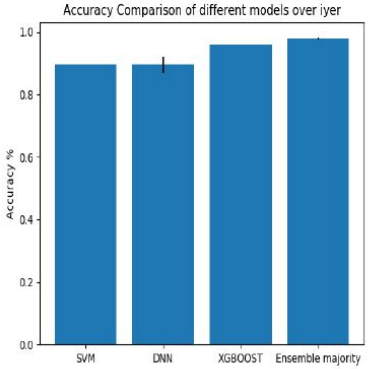
XGBoost

- Changing parameters won't affect much!
- High performance on both dataset

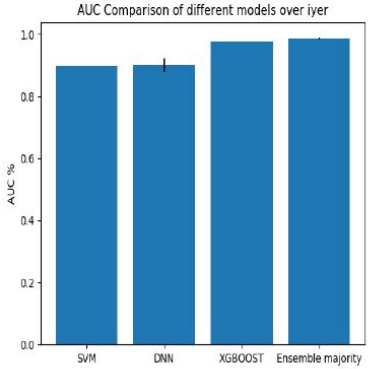
Dataset	Accuracy	AUC	Recall	Precision	F1
MNIST	96.15	97.53	95.45	94.11	94.08
Iyer	97.80	98.76	97.78	97.79	97.78

Table 3. XGBoost classification performance across two datasets

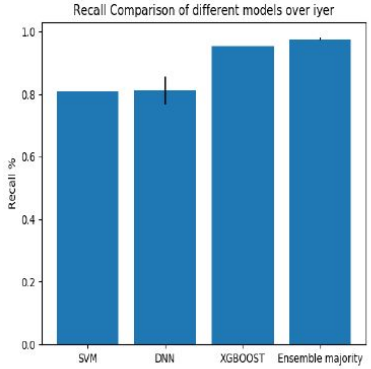
Ensemble Learning



(a) Accuracy comparison

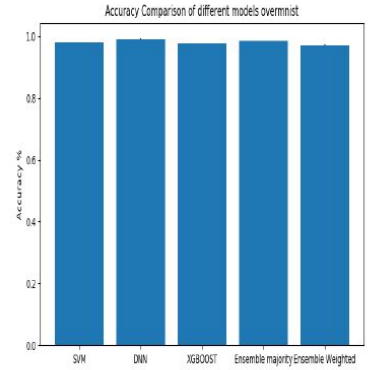


(b) Area Under Curve Comparison

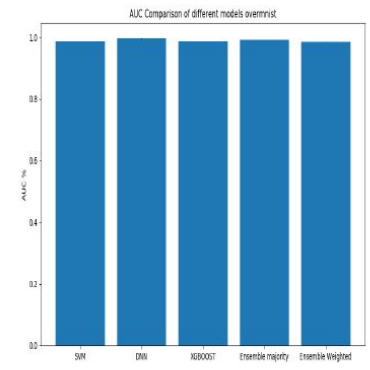


(c) Recall Comparison

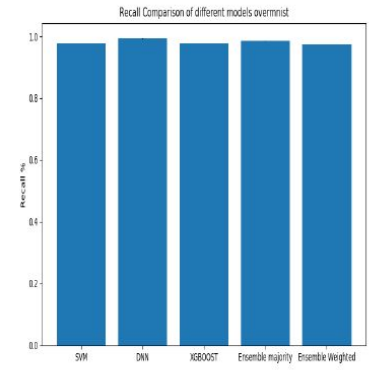
Fig. 6. Comparison of different algorithms on Iyer dataset. Black line over bar charts shows variance of models' performance.



(a) Accuracy comparison



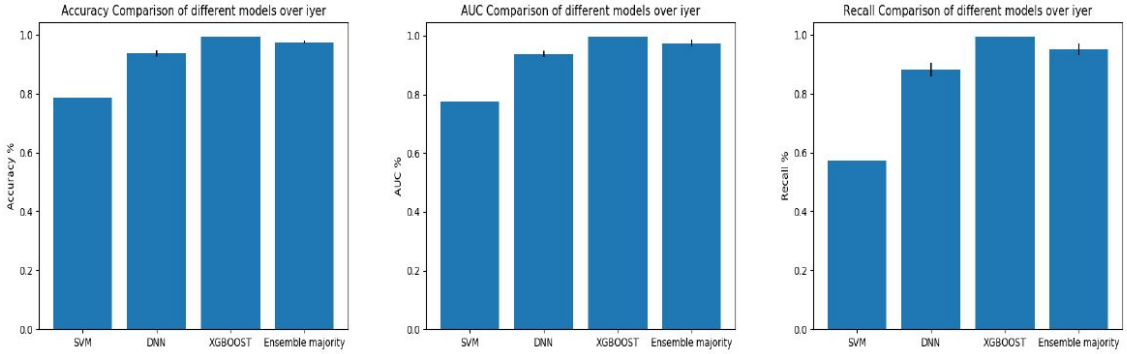
(b) Area Under Curve Comparison



(c) Recall Comparison

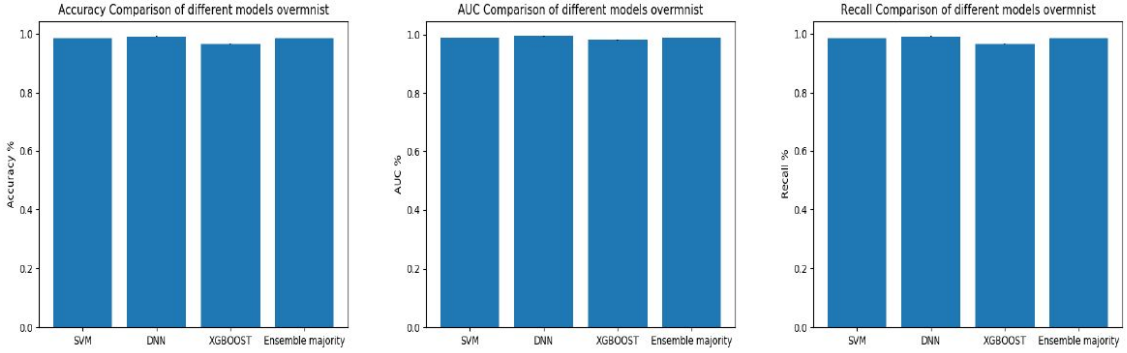
Fig. 7. Comparison of different algorithms on MNIST dataset. Black line over bar charts shows variance of models' performance.

Dimension Reduction



(a) Accuracy comparison (b) Area Under Curve Comparison (c) Recall Comparison

Fig. 8. Comparison of applying PCA on different algorithms across Iyer dataset.



(a) Accuracy comparison (b) Area Under Curve Comparison (c) Recall Comparison

Fig. 9. Comparison of applying PCA on different algorithms across MNIST dataset.

Discussion

- Data features
  - Dimension reduction can be a boost
  - Feature selection
  - PCA is not always a boost!
- Different datasets
  - No model is dominant on every dataset
  - Data distribution and model strength
- Accuracy is not everything
  - Speed
  - Generalizability
  - Ease of access
- Beyond performance!
  - Readability and Explainability
  - Linear relation