Team 51
Dimensionality Reduction

Govind Shukla, Kartavya Gupta, S Abhilash, Vrushank Vyas

Comparing different dimensionality reduction techniques

PCA, MDS on SVM

Work

- Built SVM with 7 kernels and 80-20 random shuffling
- Ran SVM on original data - stored time and error rate
- Built PCA - reduced from 28X28 to 14X14 - stored time and error rate
- Built MDS
- Built LDA

Dataset

MNIST

- Handwritten digits
- Training set of 60,000 examples
- Test set of 10,000 examples.
- The digits have been size-normalized and centered in a fixed-size image.

SVM Classifier

- Poly
  - 2
  - 3
  - 4
- RBF
- Sigmoid
- Linear
  - Hinge
  - Squared Hinge

SVM

- Gamma variable in all 3 non-linear kernels
- One vs all
  - Train one classifier per class in total
- One vs one
  - Train a separate classifier for each different pair of labels.
- 3 parameters - Gamma, Kernel, 1v1 or 1vRest
- Loss is varied in Linear

PCA

- 28x28 image was the input
- Computed co-variance of the data matrix
- And took maximum 14 eigen vectors corresponding to 14 maximum eigen values as principal components
- Used the new 14 principal components to form the projection matrix
- Computed the reduced data by projecting the components on to the original dataset
- Reducing it to a 14X14 image

PCA Results

- On the original dataset - the max accuracy was 92%
- While the time was X
- On the new dimensionally reduced dataset - the accuracy was 85%
- While the time was Y

Experiments - Tables and Graphs

Comparison

Compare the accuracies of different dimensionality reduction techniques.

PCA, MDS