**Experimental setting (Unbalanced Data):**

- positive class (in training): 800-digit matrix composed of 720 even digits and 80 digits ‘1’;

- positive class (in test): 800-digit matrix composed of 720 even digits and 80 different odd digits (i.e., 3 or 5 or 7 or 9);

- negative class: 800-digit matrix composed of 800 even digits;

- all individual digits in 800-digit matrix are randomly selected from MNIST dataset;

- number of training samples (matrices): 5 samples from positive class and 40 from negative class;

- number of validation matrices: 80 (40 per class);

- number of test matrices: 1000 (500 per class);

- model selection is performed only using validation data with equal cost of two classes;

- every experiment is repeated five times, and the average SS and SP are reported.

**Experimental setting (balanced Data):**

- positive class (in training): 800-digit matrix composed of 720 even digits and 80 digits ‘1’;

- positive class (in test): 800-digit matrix composed of 720 even digits and 80 different odd digits (i.e., 3 or 5 or 7 or 9);

- negative class: 800-digit matrix composed of 800 even digits;

- all individual digits in 800-digit matrix are randomly selected from MNIST dataset;

- number of training samples (matrices): 40 samples from positive class and 40 from negative class;

- number of validation matrices: 80 (40 per class);

- number of test matrices: 1000 (500 per class);

- model selection is performed only using validation data with equal cost of two classes;

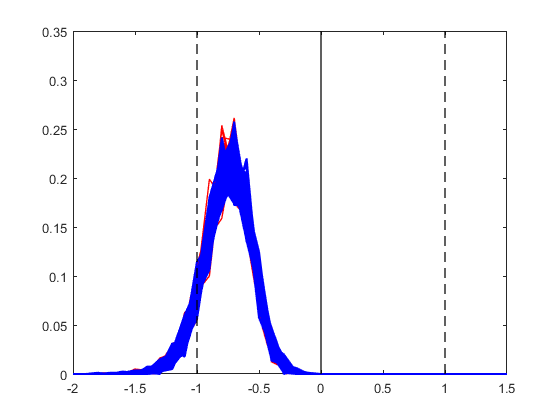
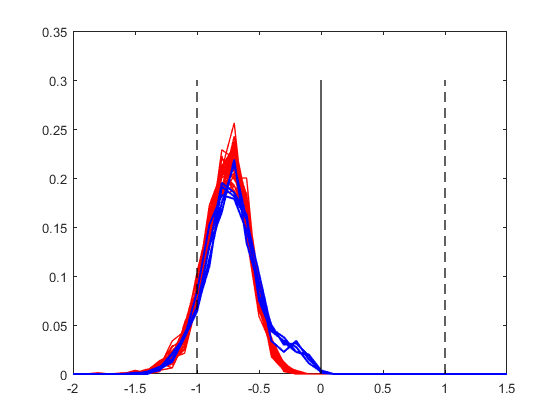
- every experiment is repeated five times, and the average SS and SP are reported.

Results (unbalanced data)

|  |  |  |
| --- | --- | --- |
| Odd digit in test data | SS (%) | SP (%) |
| ‘3’ | 19.56 | 97.16 |
| ‘5’ | 37.28 | 97.20 |
| ‘7’ | 60.32 | 96.68 |
| ‘9’ | 25.96 | 94.18 |

Results (balanced data)

|  |  |  |
| --- | --- | --- |
| Odd digit in test data | SS (%) | SP (%) |
| ‘3’ | 57.36 | 95.24 |
| ‘5’ | 3.56 | 98.80 |
| ‘7’ | 11.82 | 97.68 |
| ‘9’ | 11.24 | 95.60 |



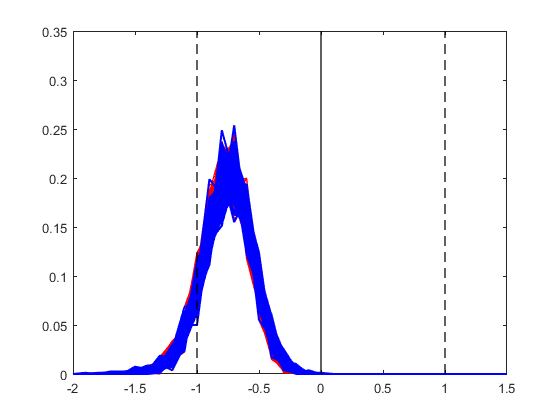
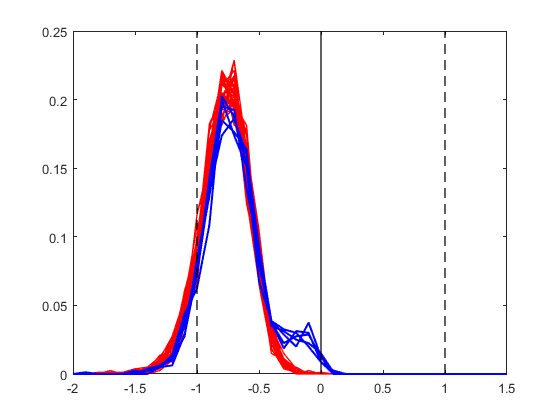
**Fig. 1.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘3’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 5 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 2.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘3’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 5 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).



**Fig. 3.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘5’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 5 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

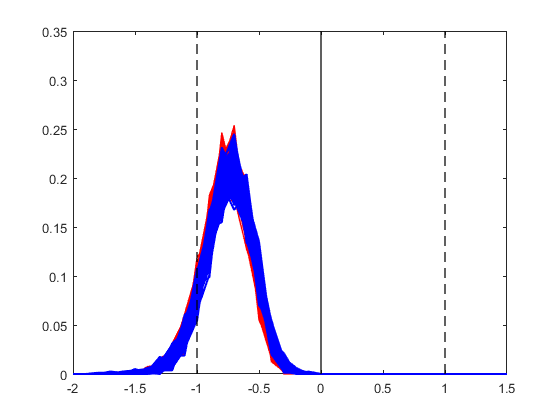
一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 4.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘5’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 5 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).

一張含有 地圖 的圖片

自動產生的描述

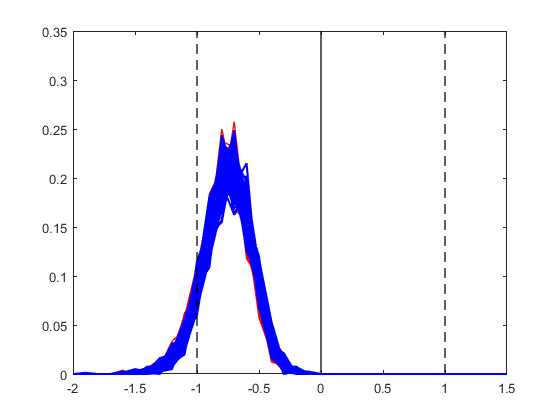
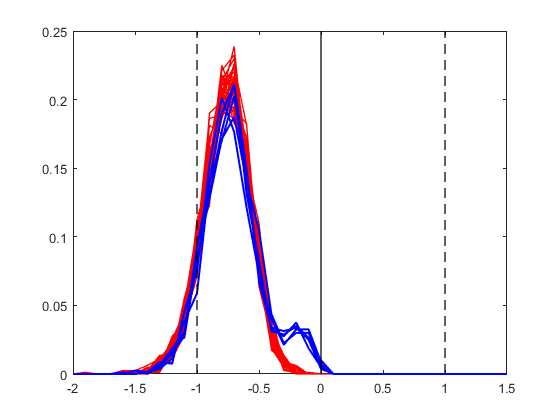
**Fig. 5.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘7’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 5 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

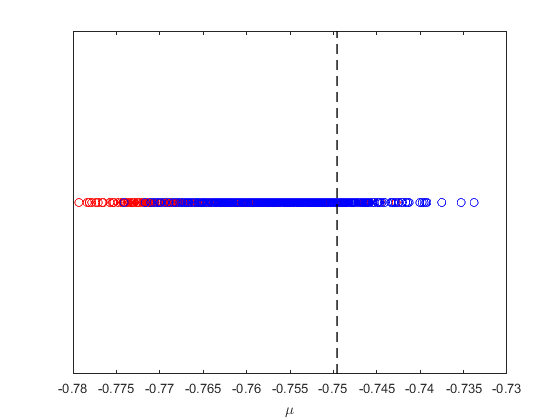
自動產生的描述

**Fig. 6.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘7’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 5 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).

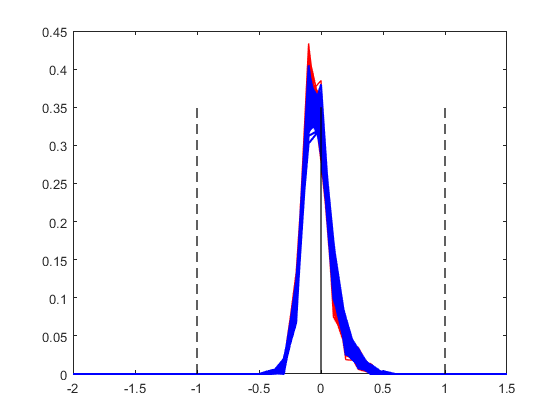
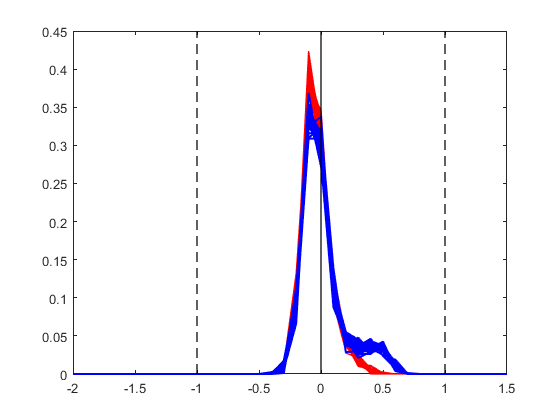


**Fig. 7.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘9’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 5 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

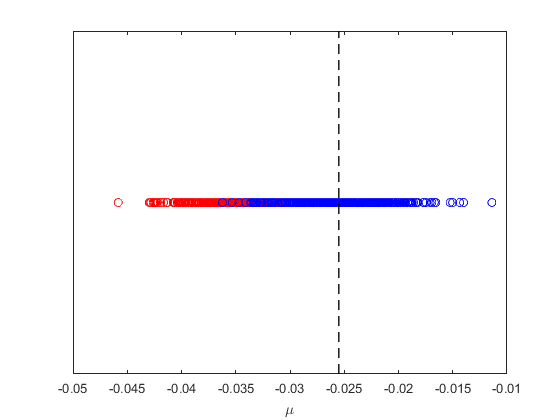
自動產生的描述

**Fig. 8.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘9’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 5 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).

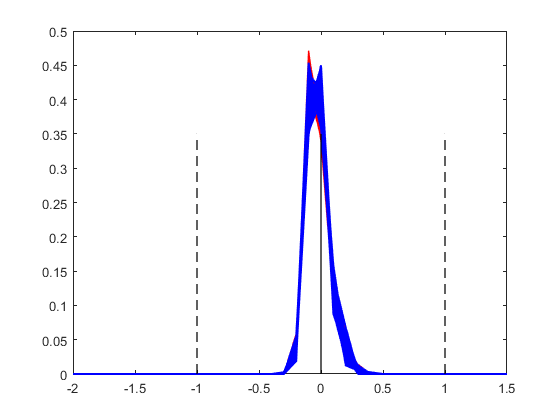
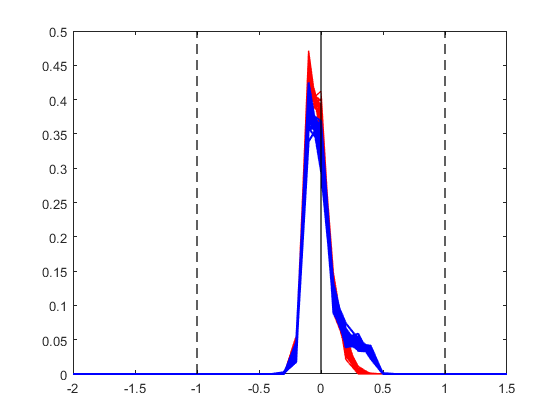


**Fig. 9.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘3’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 40 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 10.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘3’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 40 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).



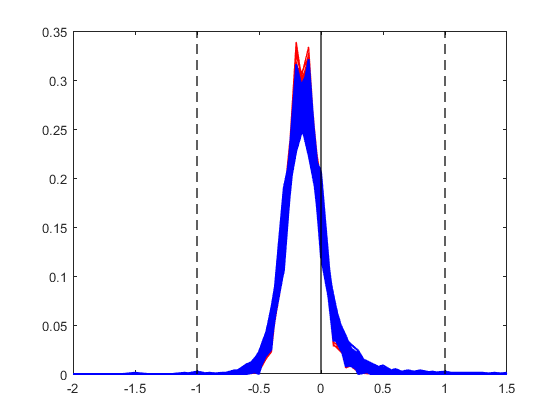
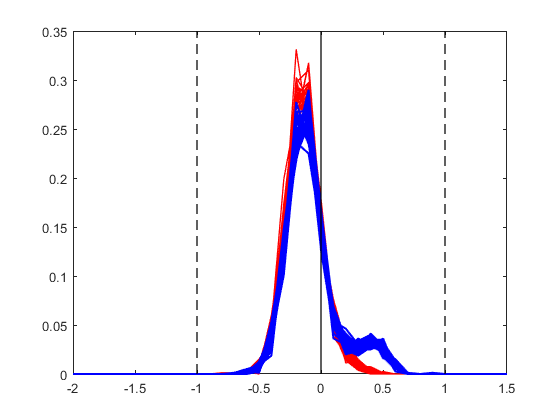
**Fig. 11.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘5’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 40 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 12.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘5’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 40 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).



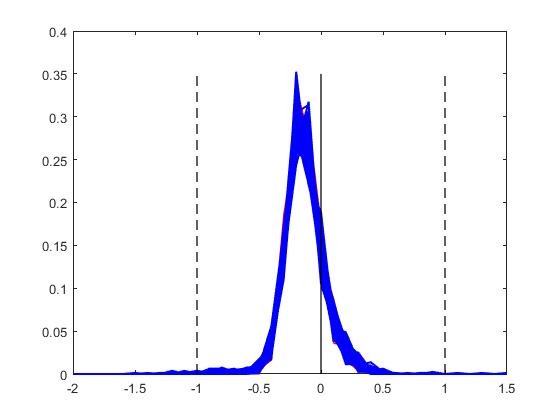
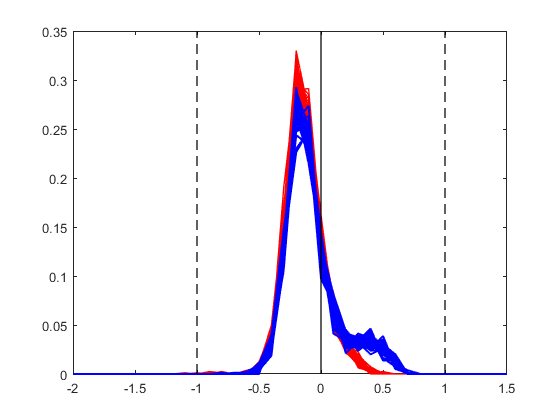
**Fig. 13.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘7’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 40 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 14.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘7’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 40 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).



**Fig. 15.** Histograms of projections for ‘matrix-of-digits’ data formed by 800 handwritten digits. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘9’. *Notation:* red ~ negative class, blue ~ positive class samples; SVM margin borders are labeled -1 and +1. (a) histograms for training data (left, 40 positive and 40 negative matrices); (b) histograms for test data (right, 500 positive and 500 negative matrices).

一張含有 螢幕擷取畫面 的圖片

自動產生的描述一張含有 螢幕擷取畫面 的圖片

自動產生的描述

**Fig. 16.** Clustering of the mean values of 800 outputs of SVM classifier under Group Learning setting. Positive training data contains odd digit ‘1’, and positive test data contains odd digit ‘9’. Dashed line indicates optimal adaptive decision threshold, *Q*\* which separates 100% of negative training samples (in the μ–space). (a) histogram of μ –values for training data (left, 40 positive and 40 negative samples), (b) histogram of μ–values for test data (right, 500 positive and 500 negative samples).