Handwritten Digit Recognition via Unsupervised Learning
Xinyi Jiang, Huy Pham, Qingyang Xu
xinyij@stanford.edu, huypham@stanford.edu, youngxu@stanford.edu

Introduction
Handwritten digit recognition is a benchmark test of computer vision algorithms. Supervised learning algorithms have error rates lower than 1%, while unsupervised ones in general have error rates ten times larger. We aim to use novel unsupervised learning methods to improve accuracy to the level of supervised learning.

So far we have implemented k-means, k-means++, PCA, sparse autoencoder and spiking neural networks, improved certain features of each, and integrated them to get higher accuracy. For comparison with supervised learning, we also trained our data with the LIBLINEAR SVM library.

The MNIST dataset we used contains 60,000 training images and 10,000 testing images labeled with the digit it represents. Each image has 28x28 pixels on gray scale.

Figure 1: Samples of MNIST training set (left two) and test set (right two)

Improved K-Means
We improve the baseline k-means algorithm which uses 28x28 pixel values as features by modifying (separately)

• the initialization process, by K-Means++, which iterates with distinct test images (probabilistic) rather than randomizing once (deterministic) to find initial centroids

Figure 2: The idealized initialization (which needs labels), with each centroid given an image of a distinct digit, performs better than a random initialization whatever number of features we use.

Figure 2: K-Means++, which repeats initializing before a corresponding objective is achieved, performs better than random initializing in general.

The learning result is the most accurate when we reduce from 784 to around 100 features using PCA.

Figure 3: The learning result is the most accurate when we reduce from 784 to around 100 features using PCA.

Sparse Autoencoder
Sparse autoencoder is a variation of neural network. It tries to reconstruct the input through a smaller number of hidden neurons compared to the input dimension.

Figure 4: Sparse autoencoder selects relevant input features for k-means

Sparsity structure uses fewer hidden units to capture the essential structure of high dimensional inputs.

Figure 4: Sparsity structure uses fewer hidden units to capture the essential structure of high dimensional inputs.

Spiking Neural Networks
The spiking neural network (SNN) is a dynamic system of neurons which evolves in time. We used Brian as the simulator of SNN.

Compared to the sparse autoencoder, the time-dependent neuron activation allows us to update the weight associated with each synapse with more control.

Figure 5: The dynamics of the membrane potential of a single neuron.

Figure 5: Network architecture

Results
We run the algorithms on training sets of different sizes, and report the accuracy we got for distinct algorithms.

• We found in implementing sparse autoencoder that the parameters oscillate without converging.
• We verified that SNN yields the best result for unsupervised approaches.
• K-Means++ with PCA and the original K-Means are much faster than the other two unsupervised learning algorithms.
• SNN requires a very long time to train. We are currently working to improve its efficiency and integrate it with the other unsupervised models to get higher accuracy.

References