

CS6690: Pattern Recognition Assignment #3

Group 3: Akshay (ME10B124) & Suthirth (ED10B059)

October 27, 2014

1 Gaussian Mixture Models

Gaussian Mixture Models are a type of classification that sets soft boundaries to data. It estimates data as a probabilistic mixture of its features. Each mixture is represented by a gaussian density function in a gaussian mixture model.

$$p(\bar{x}) = \sum_{i=1}^K \pi_i N(\bar{x} | \mu_i, \Sigma_i)$$

K is the total number of classes, π the weights.

Gaussian Mixture Models are solved using the Expectation maximization Algorithm (EM). This consists of calculating the posterior probability γ which is calculated by using.

$$\gamma_{nk} = \frac{\pi_k N(\bar{x} | \mu_k, \Sigma_k)}{\sum_{i=1}^K \pi_i N(\bar{x} | \mu_i, \Sigma_i)}$$

This posterior for each class gives us the contribution of class k to the mixture, using this, we can maximize the other parameters.

$$\begin{aligned}\pi_k &= \frac{\sum_{i=1}^K \gamma_{nk}}{N} \\ \mu_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} \bar{x}_n \\ \Sigma_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} (\bar{x}_n - \mu_k)(\bar{x}_n - \mu_k)^T\end{aligned}$$

These three update equations are used to maximize and convergence is checked. The initial point for convergence is usually the k-means algorithm.

2 Dynamic Time Warping

Dynamic Time Warping is a template matching algorithm that works by using the principles of dynamic programming. It is a template matching that is capable of matching uneven templates as well. DTW aligns two sequence of feature vectors by warping the time axis iteratively until an optimal match between the two sequences is found. Using this, a distance measure can be computed, and can be used to match to a template.

Since we're dealing with continuous sequences in our dataset, the algorithm can be written as:

$$D[i,j] = |a[i] - b[j]| + \min(D[i-1,j], D[i-1,j-1], D[i,j-1])$$

3 Hidden Markov Models

A Markov Model is one in which the current state of a system depends only on the present and previous states of the system. The order of the system tells us about the dependency of the nth previous value on the conditional probability density of the model.

In a Hidden Markov Model, the state sequence is hidden as is the transition probabilities. We use the HMMs to characterize images and speech data. The performance of the HMM is also benchmarked to that of a GMM.

4 Metrics

The metrics used to evaluate the performance of a classifier are as follows

1. Precision
2. Recall
3. ROC Curves

5 Image dataset

5.1 Data Exploration

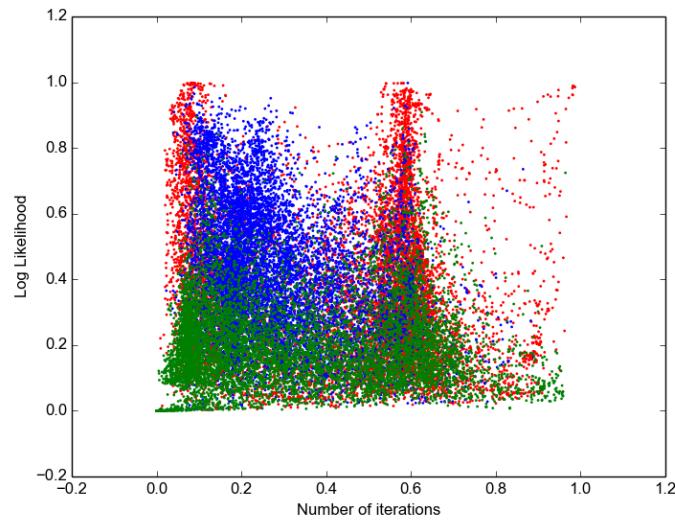
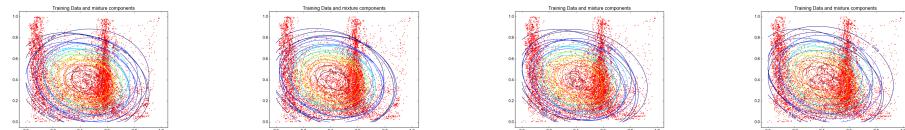


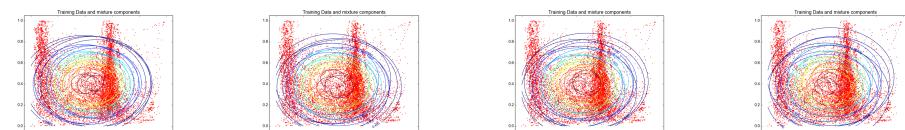
Figure 1: Scatter plot of the image dataset

5.2 Coast

Full Scatterplot:

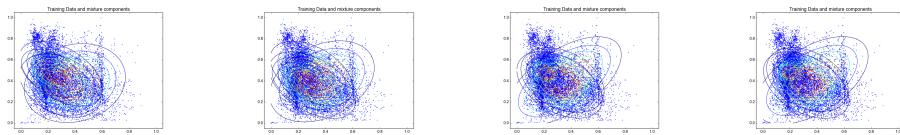


Diagonal Scatterplot:

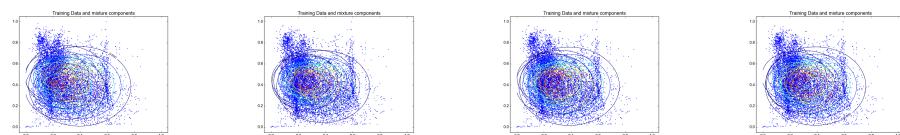


5.3 Forest

Full Scatterplot:

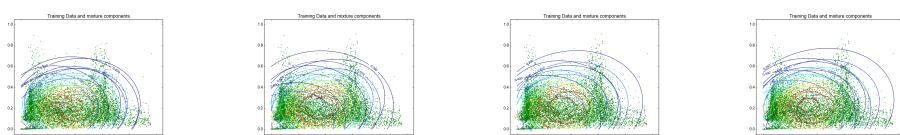


Diagonal Scatterplot:

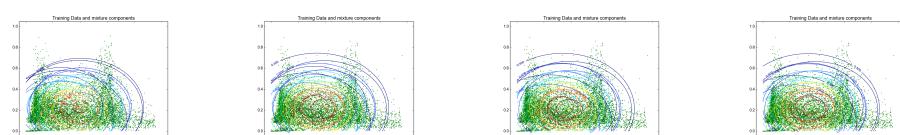


5.4 Street

Full Scatterplot:



Diagonal Scatterplot:



5.5 Results

5.6 Confusion Matrices

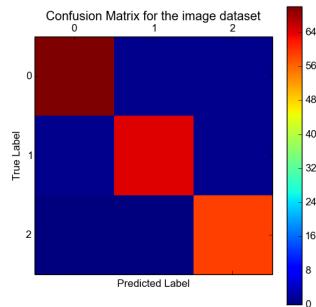


Figure 2: Confusion Matrix for full co-variance

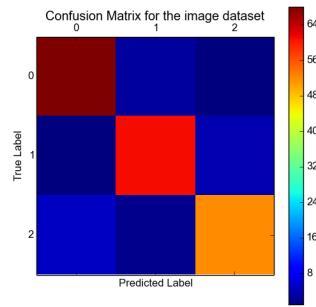


Figure 3: Confusion Matrix for Diagonal Covariance

5.7 Precision and Recall

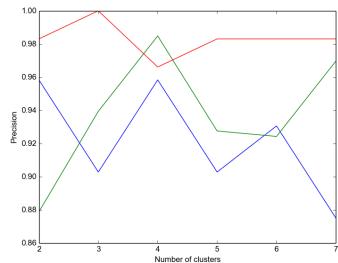


Figure 4: Precision

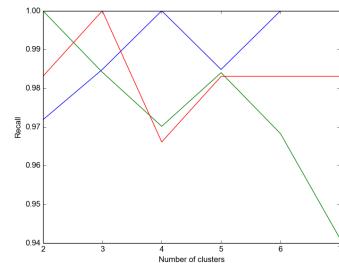


Figure 5: Recall