Report: Assignment 3

Submitted by: Arulkumar S, Shitanshu Kusmakar
CS15S023, ED15F003

30-Oct-2015

1 Introduction

1.1 Problem Definition

In this assignment, we evaluate the performance of probability distribution mixture models to cluster and classify the data. The probability distribution used is Gaussian and gaussian mixture models (GMM) were built for the given data sets. Mixture models are a class of pattern recognition systems. A Gaussian mixture models the probability density function of the given data using a mixture of density functions. GMM can be viewed as a linear superposition of k Gaussian components providing a better estimation of the density model of the observed variables in comparison to a single Gaussian. GMM's find wide application in speech recognition problems. For the given data set, a probability density function was estimated using k-component Gaussian mixture model and the parameters of the mixture model are estimated using EM algorithm.

However, in most cases the data is assumed to be (i.i.d). Whereas, in many applications, the observed data might not be (i.i.d) and describe a sequential data such as those arising from a time series. To cumulate such effects in probabilistic models, we use Markov model. When the latent variables are discrete the model is termed as discrete hidden markov model (DHMM). For a given data set, we used the DHMM library and evaluated the performance of the models using ROC and DET curves.

DTW is an algorithm for measuring similarity between two time series sequences. The technique of dynamic time warping has been employed for digit identification.

2 Approach to the problem

2.1 What is the given information

A Gaussian mixture model has to be build to fit multivariate normal distributions. Full and diagonal covariance matrix has to be tested for making predictions on the test data. A gaussian mixture distribution can be written as a linear superposition of Gaussians in the form as shown in equation 1.

$$p(x) = \Sigma_{k=1}^{K} \pi_k \mathcal{N}(x/\mu_k, \Sigma_k) \tag{1}$$

where, x is the d-component feature vector, μ is the mean vector with d-components, Σ is the $d \times d$ covariance matrix, and π_k is the mixing parameter.

Also, the HMM and DTW algorithm has to be implemented on the given data sets and performance measures for all the pattern recognition systems has to be calculated.

The two cases for which the probability density function (GMM) has to be tested are

- Full covariance matrix (Σ).
- Diagonal covariance matrix (Σ).

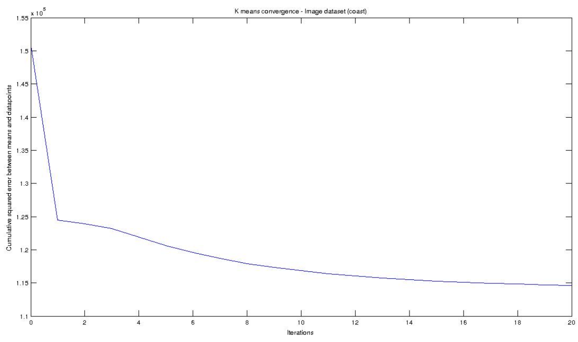
2.2 Approach

The approach to the problem was based on following things

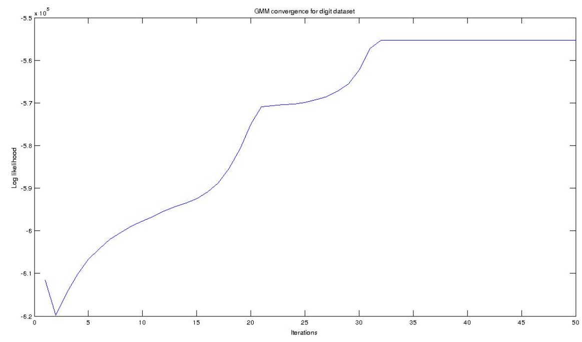
- The data was divided into training and test set. 70% of the data was selected as training set and 30% as test.
- A k-component Gaussian mixture model was trained using the training data. The parameter k was determined empirically, and it was set to 5.
- The initialization of the parameters of the GMM were done using k-means clustering. The figure 1 (a) shows the convergence of the k-means algorithm based on calculating the squared error.
- The parameters are then updated using the EM algorithm and log likelihood of the observed data is calculated in each step. The log likelihood is as shown in equation 2.

$$lnp(X/\mu, \Sigma, \pi) = \Sigma_{n=1}^{N} ln\left\{\Sigma_{k=1}^{K} \pi_k \mathcal{N}(x_n/\mu_k, \Sigma_k)\right\} \tag{2}$$

- The convergence of the algorithm is decided by fixing the number of iterations for the EM step. The figure 1 (b) shows the convergence of the algorithm based on log likelihood calculation.



(a) Convergence of k-means

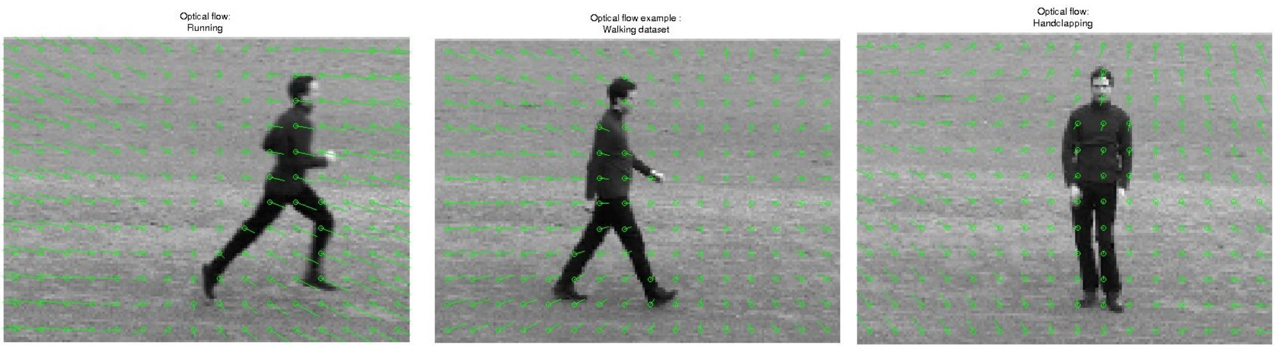


(b) Convergence of EM algorithm for digit data set

Figure 1: Convergence of k-means and EM algorithm

- The testing is then performed using the GMM parameters (π_k, μ_k, Σ_k) for every image. The feature vectors corresponding to a particular image were assigned to individual classes based on maximum likelihood and the class for an image is determined using mode of classes for the feature vectors.

- The ROC and DET curves along with the performance measures for each data set are reported.
- The implementation of the DTW and the HMM is followed by the preparation of the data into sequences represented by the number of classes using k-means clustering.
- The video dataset is first pre-processed to calculate features based on optical flow pattern [1] [2] between consecutive frames. Using these features, a GMM and an HMM is trained for each of the classes (running, walking and hand clapping). The feature extraction were done using the following library [3]. The features include six parameters which are extracted between continuous pair of images.
  - Translation (v_{xo}, v_{y0}) overall shift of the second image w.r.t to first image.
  - Dilation
  - Rotation
  - Shear (S_1, S_2) caused by the depth gradient of the image.
- The figure 2 show the optical flow descriptors for running, walking and hand clapping video data sets.



(a) Optical flow descriptors for running video

(b) Optical flow descriptors for walking video

(c) Optical flow descriptors for hand clapping video

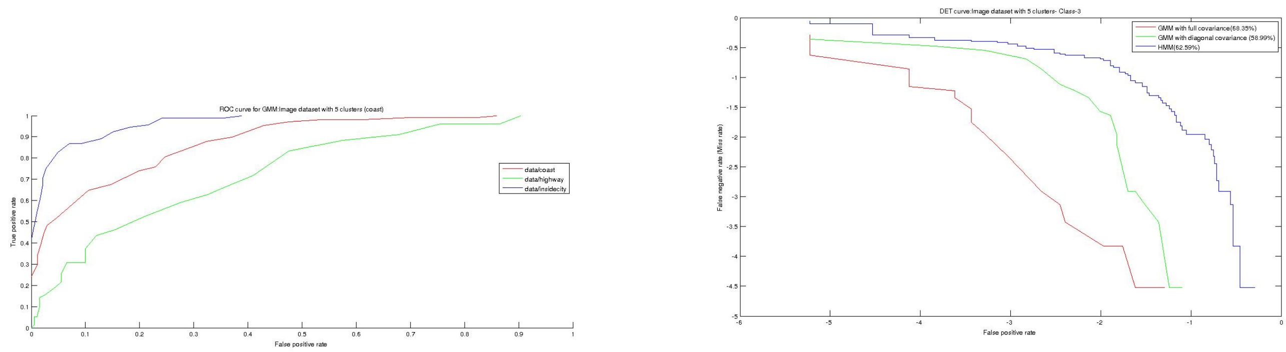
Figure 2: Optical flow descriptors for the video data set

3 Results & Observations

3.1 Image Data Set (Coast, Highway and Inside City)

Gaussian mixture model with k-components and HMM were implemented on the image data set. The figure 3 shows the ROC curve for the GMM and figure 4 shows the comparison of the GMM algorithm with HMM. The figure 4 highlights the performance of the two algorithms on the same data set. The figure 3 (a) also shows the performance of the GMM algorithm for the two cases of full and diagonal covariance and figure 3 (b) shows the DET curve for the GMM algoirthm obtained on the image dataset.

It was observed as seen in figure 3 (a) that higher classification accuracy is observed in the case of full covariance matrix. The reason for such a behaviour can be accounted to the fact, that by making the non-diagonal values to zero we are assuming that the features are i.i.d. However, there might be some inherent dependence between the features which gets neglected on making the non-diagonal elements zero. As a result of which the data might not be represented by standard basis.



(a) ROC curve for all classes: Image data set

(b) DET curve for all classes: Image data set

Figure 3: ROC and DET curves for image data set for GMM and HMM algorithm

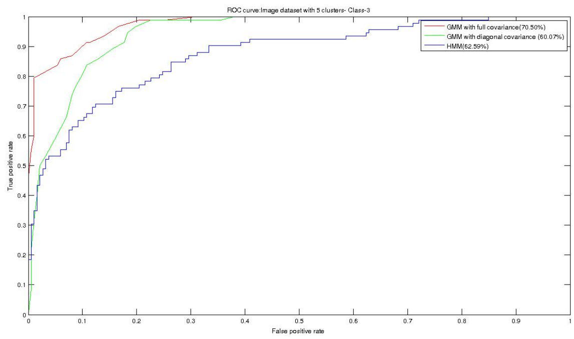


Figure 4: Performance comparison GMM/HMM for image data set (shown for class 3 only).

Table 1 and 2 shows the confusion matrix and the performance measure for the image data set for full covariance.

Table 1: Confusion matrix. GMM image data set. Full covariance case.

Class	Coast-Pred	Highway-Pred	InsideCity-Pred
ClassI-Coast-Trgt	100	3	5
ClassII-Highway-Trgt	46	8	24
ClassIII-Insidecity-Trgt	6	0	86

3.2 Digit Data Set (four, five, seven, eight and nine)

A very high classification accuracy was obtained on the digit data set using GMM and HMM algorithms as shown in figure 5. Also Discrete time warping (DTW) algorithm is also implemented and the results of the DTW implementation are shown in table 3 and the performance measures are shown in table 4.

Table 3 and 4 shows the confusion matrix and the performance measure of the GMM classifier for the digit data set.

Table 2: Table showing the performance measures for the classifier. GMM image data set.

Measure	**Coast**	**Highway**	**InsideCity**
Sensitivity	0.92	0.10	0.93
Specificity	0.69	0.98	0.84
Precision	0.65	0.72	0.74
F1-score	0.76	0.18	0.83
Overall Accuracy	0.70		

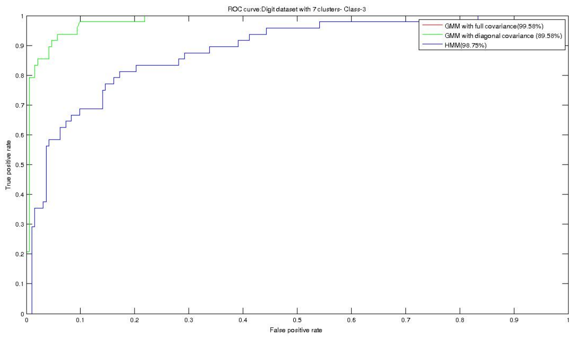


Figure 5: ROC curve for digit data set with 7 components (shown only for class 3).

Table 3: Confusion matrix. DTW algorithm for the digit data set.

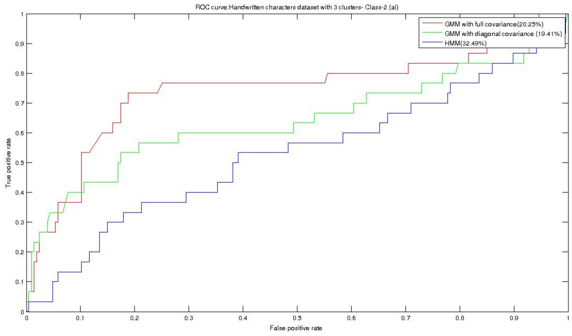
Class	**Five-Pred**	**Eight-Pred**	**Seven-Pred**	**Nine-Pred**	**Four-Pred**
ClassI-Trgt	45	1	2	0	0
ClassII-Trgt	0	48	0	0	0
ClassIII-Trgt	0	2	38	8	0
ClassIV-Trgt	2	2	0	44	0
ClassV-Trgt	2	0	0	0	46

Table 4: Table showing the performance measures for the classifier. DTW algorithm for the digit data set.

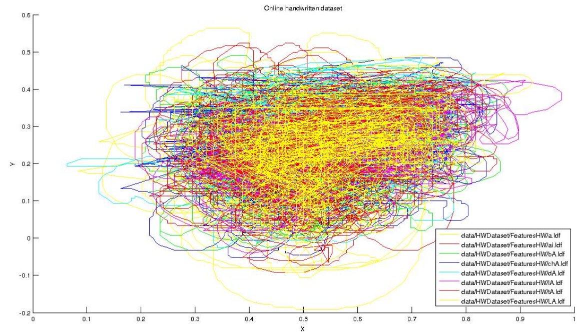
Measure	**Five**	**Eight**	**Seven**	**Nine**	**Four**
Sensitivity	0.93	1.00	0.79	0.91	0.95
Specificity	0.97	0.97	1.00	0.94	1.00
Precision	0.91	0.90	1.00	0.81	1.00
F1-score	0.92	0.95	0.88	0.86	0.97
Overall Accuracy	0.92				

3.3 Hand Written Character Data Set

However, for hand written data set, both GMM and HMM algorithm underperformed as seen in figure 6 (a). The reason for such a behaviour can be attributed to the inherent overlap of the data points in two dimensions as seen in figure 6 (b). Mixture models for data of every class assigns approximately the same likelihood to test data points for every class and hence leading to low classification accuracy.



(a) ROC curve for all classes: hand written characters data set

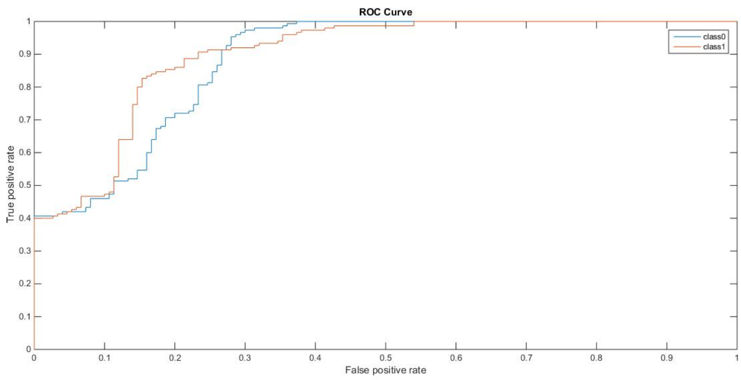


(b) Plot of the data set showing overlapping classes.

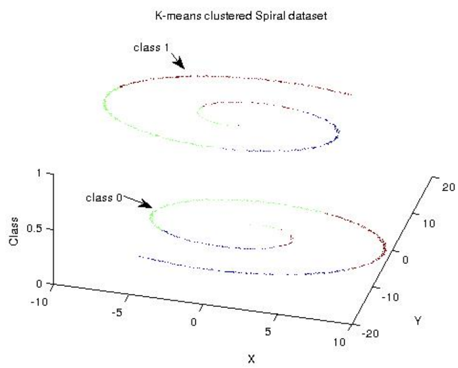
Figure 6: Plots for hand written character data set.

3.4 Spiral Data Set

A fairly good classification accuracy is obtained on spiral data set for GMM algorithm. The figure 7 (a) shows the ROC curve for GMM, for one of the classes. The figure 7 (b) shows the plot of the data after k-means clustering. It can be seen from the figure 7 (b) that the classes are distinctively separated and the data is also also distinctively clustered into k-components. Since, the initialization to the GMM algorithm is fairly good the overall accuracy is increased for such data sets. The table 5 shows the confusion matrix and table 6 shows the performance measures of the GMM algorithm on the spiral data set.



(a) ROC curve for all classes: spiral data set



(b) k-means results for the spiral data set.

Figure 7: Plots for spiral data set.

Table 5 and 6 shows the confusion matrix and the performance measure of the GMM classifier for the spiral data set.

3.5 Video Data Set (Running, Walking and Hand Clapping)

The video data set is first pre-processed for pixel value features using the affine optical flow [3] and features included six parameters which were extracted between continuous pair of images. Based, on

Table 5: Confusion matrix. GMM Spiral data set.

Class	**ClassI-Pred**	**ClassII-Pred**
ClassI-Trgt	143	7
ClassII-Trgt	11	139

Table 6: Table showing the performance measures for the classifier. GMM Spiral data set.

Measure	**ClassI**	**ClassII**
Sensitivity	0.95	0.92
Specificity	0.92	0.95
Precision	0.92	0.95
F1-score	0.94	0.93
Overall Accuracy	0.94	

these features a GMM and HMM is trained. The figure 8 shows the ROC curve for the three classes (running, walking and clapping) for GMM and HMM. An overall classification accuracy of 80% is achieved using GMM algorithm. Whereas, a classification accuracy of 67% is achieved using HMM algorithm. Table 9 and 7 shows the confusion matrix for GMM and HMM. Table 10 and 8 shows the performance measure of the GMM and HMM algorithm on the video dataset.

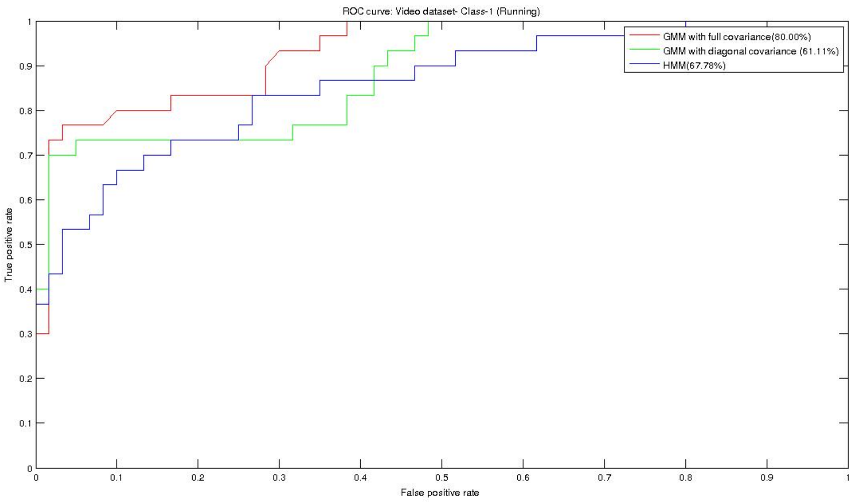


Figure 8: ROC curve for GMM and HMM. Video data set for class (running).

Table 7: Confusion matrix. GMM video data set.

Class	**Running-Pred**	**Hand-Pred**	**Walking-Pred**
ClassI-Running-Trgt	27	0	3
ClassII-Hand Clapping-Trgt	0	30	0
ClassIII-Walking-Trgt	15	0	15

Table 8: Table showing the performance measures for the GMM classifier for video data set.

Measure	**Running**	**Hand Clapping**	**Walking**
Sensitivity	0.90	1.00	0.50
Specificity	0.75	1.00	0.95
Precision	0.64	1.00	0.83
F1-score	0.75	1.00	0.63
Overall Accuracy	0.80		

Table 9: Confusion matrix. HMM video data set.

Class	**Running-Pred**	**Hand-Pred**	**Walking-Pred**
ClassI-Running-Trgt	24	1	5
ClassII-Hand Clapping-Trgt	2	19	9
ClassIII-Walking-Trgt	10	2	18

Table 10: Table showing the performance measures for the HMM classifier for video data set.

Measure	**Running**	**Hand Clapping**	**Walking**
Sensitivity	0.80	0.63	0.60
Specificity	0.80	0.95	0.76
Precision	0.66	0.86	0.56
F1-score	0.72	0.73	0.58
Overall Accuracy	0.67		

4 Conclusion

As seen from the results the classifier based on mixture models and markov model is a good algorithm for performing classification tasks involving multi class data. We performed various experiments based on which we observed that mixture models based on full covariance matrix perform better classification. Similar results were obtained for markov models. Using the mixture model classifier built for full covariance, we achieved a classification accuracy of greater than 70% in classification of image data set and greater than 90% for spiral and digit data set. However, the performance for hand written character data set was low owing to the inherent class overlap in the data. An overall accuracy 20% with GMM and 33% with HMM is achieved. On the video data set based on the optical flow features extracted from the videos, we could achieve an overall classification accuracy of 80% with GMM and 67% with HMM.

References

[1] Berthold K Horn and Brian G Schunck. Determining optical flow. In *1981 Technical symposium east*, pages 319–331. International Society for Optics and Photonics, 1981.

[2] Wikipedia.Inc. Optical flow. In $https://en.wikipedia.org/wiki/Optical_flow$. Wikipedia.Inc.

[3] Mathworks.Inc. Determining optical flow. In $http://in.mathworks.com/matlabcentral/fileexchange/270$ $affine-optic-flow$. Mathworks.Inc.