EE 573 – PATTERN RECOGNITION

Final Project Report

Ömer Alperen Katı - 2017401069
Süleyman Batuhan Vatan - 2021717030

Department of Electrical and Electronical Engineering

Instructor: H. Işıl Bozma
T.A.: Serhat İşcan
Due:11.01.2023

Contents

Question 1 – Unsupervised clustering method 1

Question 2 – First order Markov model 2

Question 3 – Supervised decision tree with SVM... 3

Question 4 – One-vs-remaining SVM with decision tree... 5

Comparison... 6

Appendix 8

List of Figures

Figure 1: The confusion matrix for unsupervised clustering method. 1

Figure 2: ROC Curve for unsupervised clustering method. 2

Figure 3: The confusion matrix for the first order Markov model. 3

Figure 4: ROC Curve for the first order Markov model. 3

Figure 5: Resulting supervised tree, SVM to be applied at end of the each node 4

Figure 6: The confusion matrix for supervised decision tree with SVM. 5

Figure 7: ROC Curve for supervised decision tree with SVM. 5

Figure 8: The confusion matrix for SVM with a decision tree 6

Figure 9: ROC Curve each for SVM with a decision tree 6

List of Tables

Table 1: Recall and precision rates for each class, presented for each method applied 7

Important: All MATLAB codes will be shared inside the submission. All codes should be commented and easily followable.

Question 1 – Unsupervised clustering method

- In this part, normalization to the range [0,1] and principle component analysis (PCA) is applied to the data, consecutively. PCA reduced the dimension to 11. Calinski-Harabasz criterion is used to assess the optimal number of clusters for each class. The resulting optimal values are [11 11 11 11 10]. Then, using those k values, mixture of gaussian models are trained using the data of each class separately. Test descriptors are given as input to the mixture models to calculate the likelihoods. After that, the sample is classified to the label of the model that gives the largest likelihood and prior multiplication.
- The results are shown in Figure 1 and 2

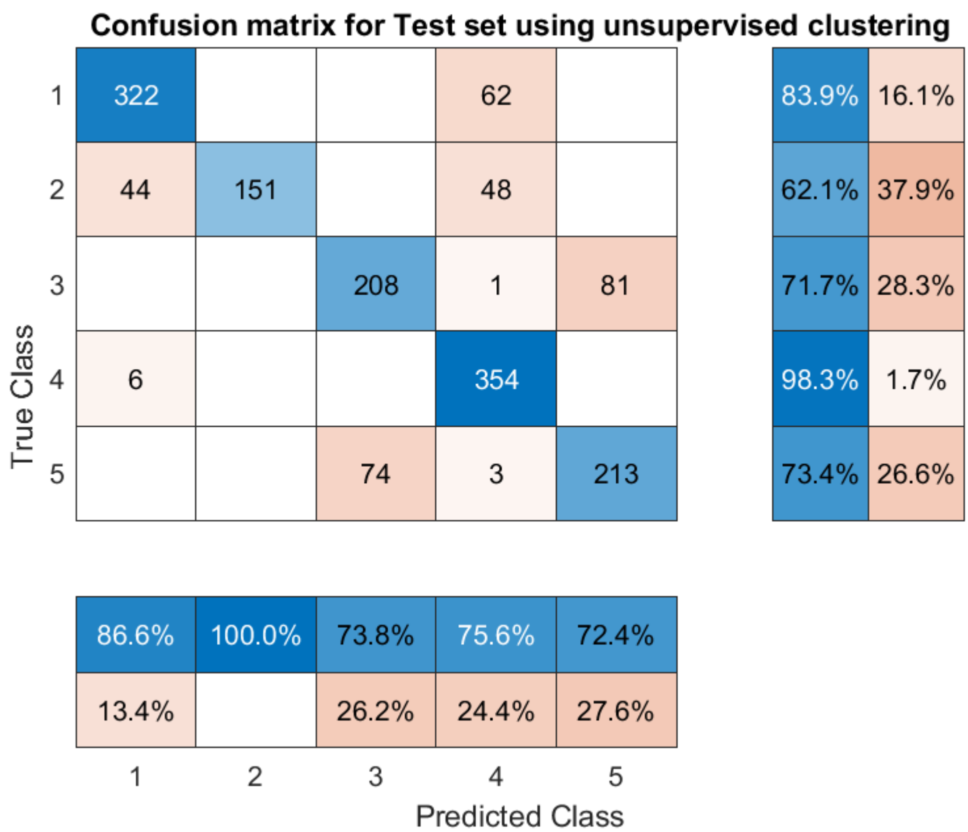

Figure 1: The confusion matrix for unsupervised clustering method.

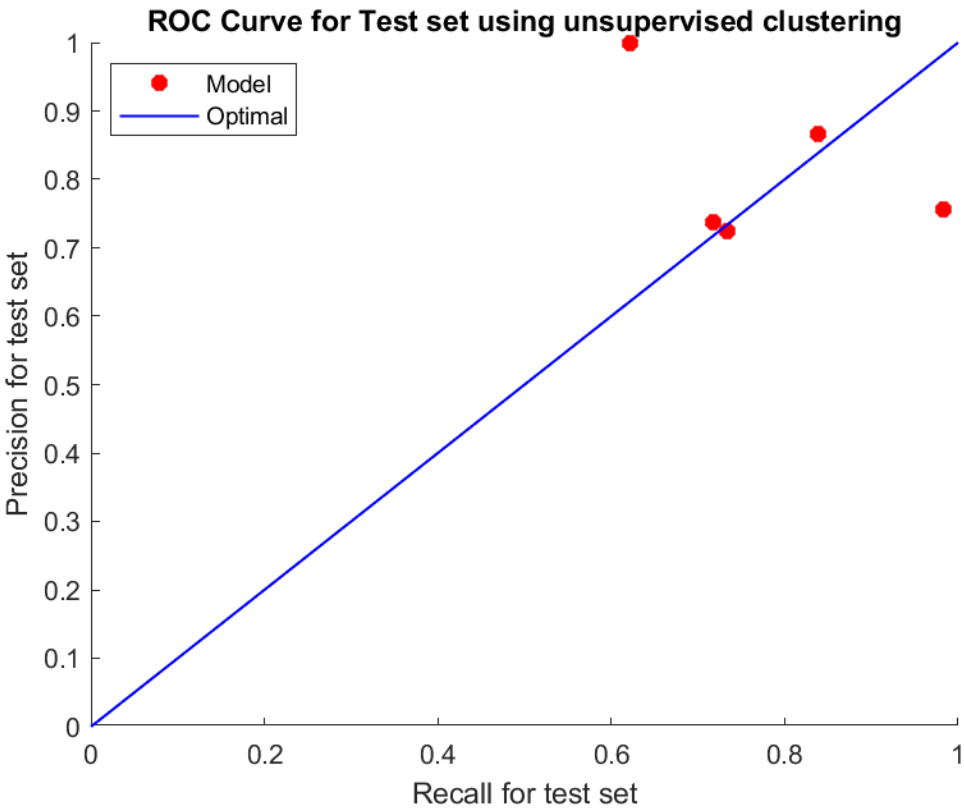

Figure 2: ROC Curve for unsupervised clustering method.

Question 2 – First order Markov model

- Description of the learning algorithm for transition and symbol generation probabilities. This case data is reduced to 15 dimensions using PCA. The whole train set is used for clustering (kmeans) and cluster number is 14. Then we use sequence information given in the dataset to train our First order markov model, train set has 5 sequences while test set has 3 sequences. Estimate of state transition matrix and emission matrix are obtained using hmmestimate function, given the sequences. Later on hmmdecode is used to label the test set sequences.
- The results are shown in Figure 3 and 4

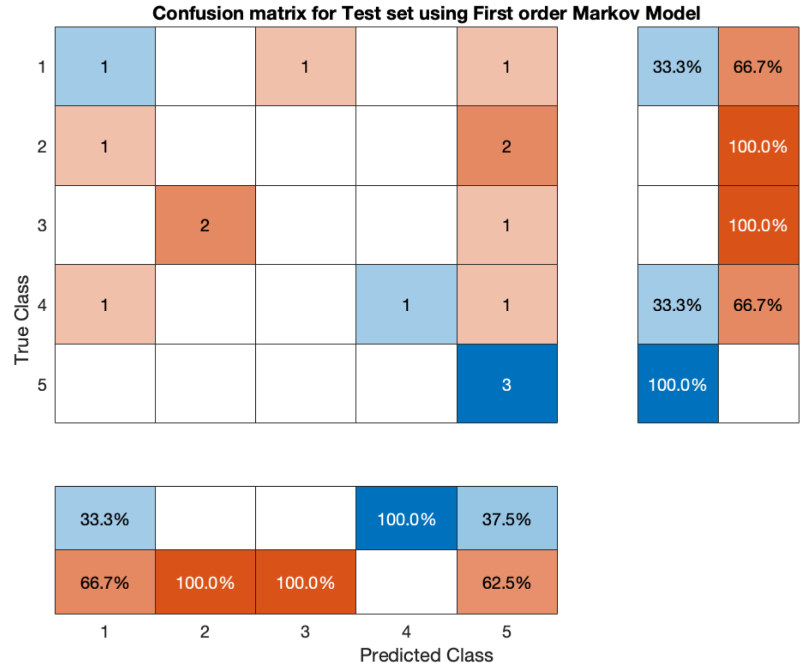

Figure 3: The confusion matrix for the first order Markov model.

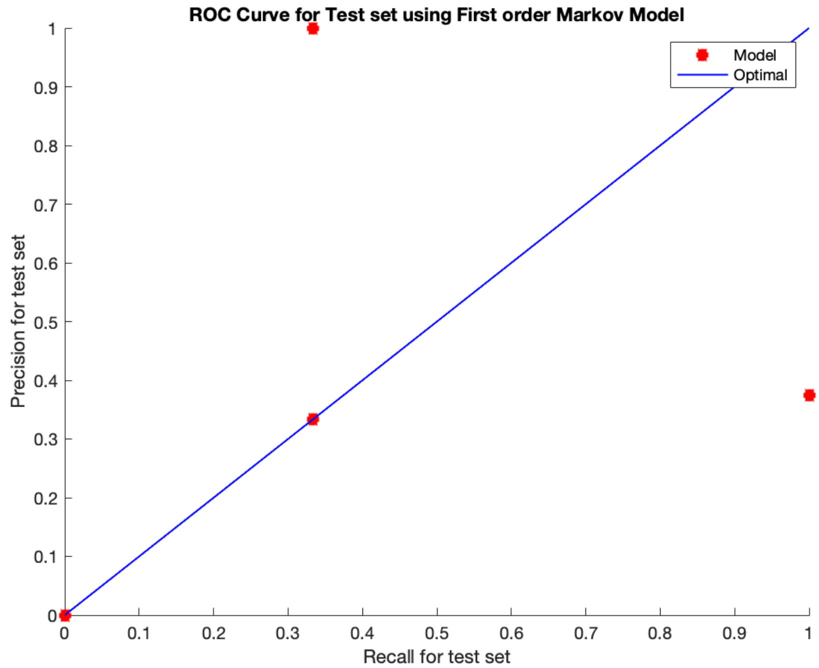

Figure 4: ROC Curve for the first order Markov model.

Question 3 – Supervised decision tree with SVM

- Description of the supervised decision tree with SVM method, supervised tree using classificationtree option along with the fitctree built-in function results following tree, presented in Figure 5.

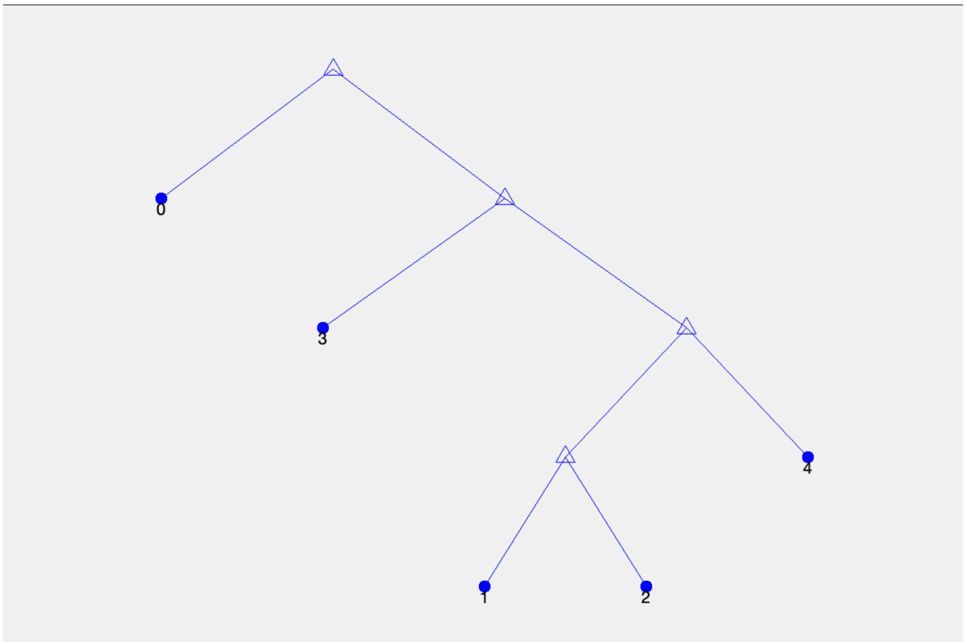

Figure 5: Resulting supervised tree, SVM to be applied at end of the each node

Note that we didn't reduce the dimensionality in this case also, there is no normalization. Then sequentially decision making using SVM is done at the end of each node, training is based on children class information.

- The results are shown in Figure 6 and 7

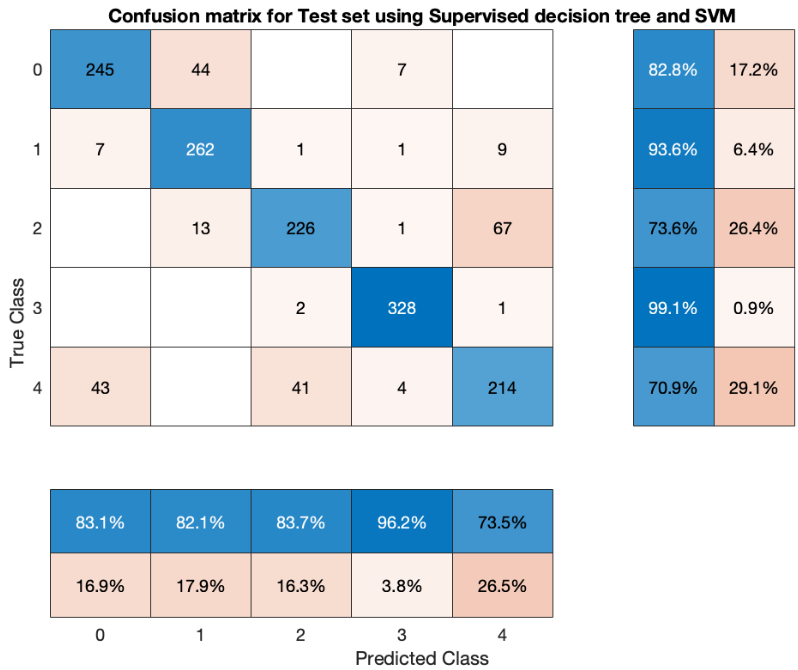

Figure 6: The confusion matrix for supervised decision tree with SVM.

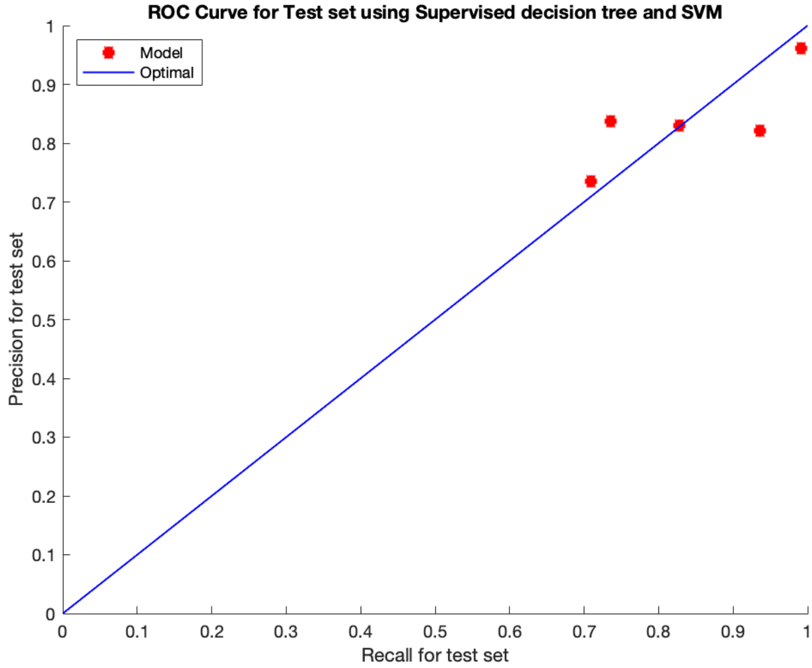

Figure 7: ROC Curve for supervised decision tree with SVM.

Question 4 – One-vs-remaining SVM with decision tree

- In this part, the dimension of both learning and test set is reduced to 73 using PCA. Then, by using one-vs-remaining approach, SVM with a linear kernel function is applied to standardized learning descriptors. Using each model, all samples are classified, and scores are saved. Before constructing the three, to have an intuition of how the learning data of each class is separated from the others, means of each class are calculated. Then, starting from the model of the best separated class, the tree is constructed, and test cases are labeled. Basically, considering the priority of each class, positive samples are labeled. After all, for the samples that cannot be labeled positively, the decision of the the model that gives the maximum score is considered.
- The results are shown in Figure 8 and 9

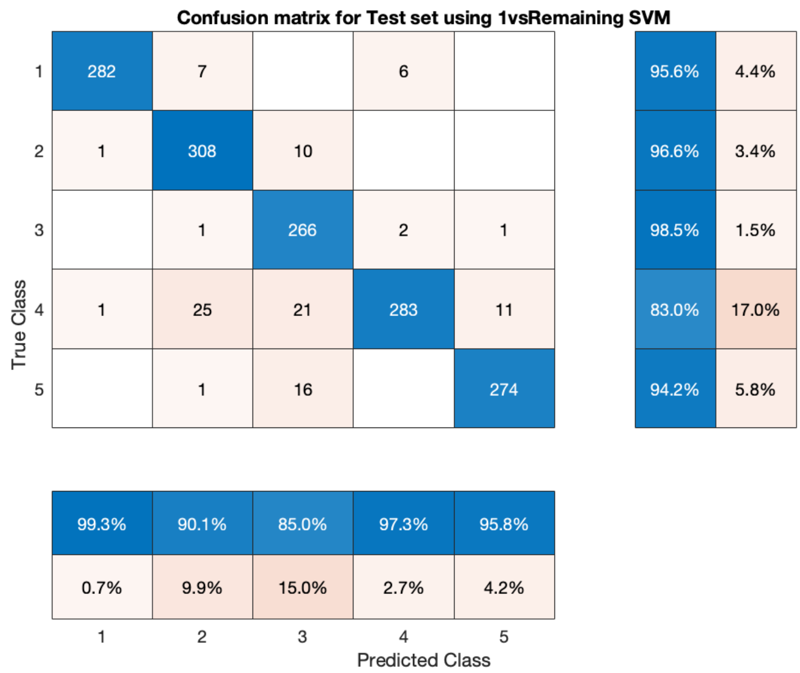

Figure 8: The confusion matrix for SVM with a decision tree

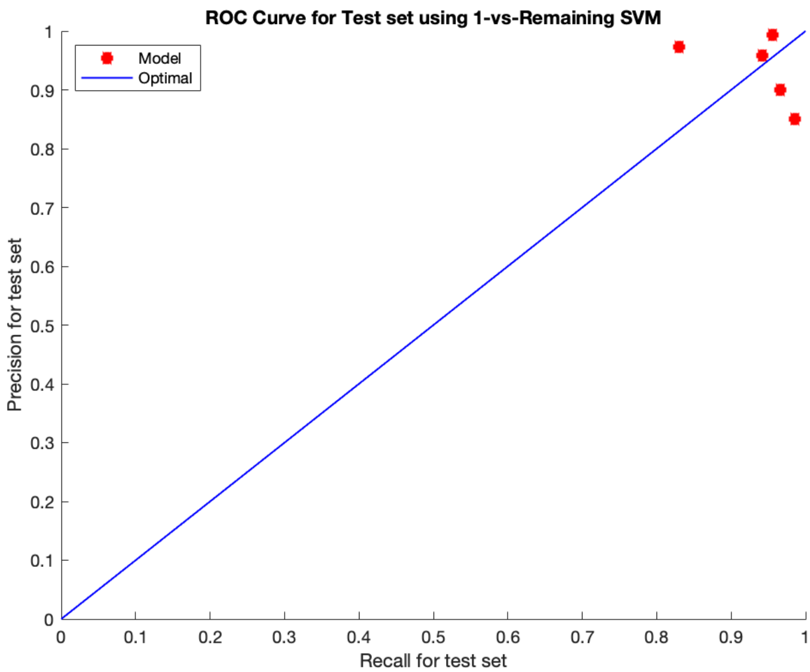

Figure 9: ROC Curve each for SVM with a decision tree

Comparison

- Unsupervised clustering with expectation maximization is relatively a generalized approach. It can be useful to classify classes with simple distributions. On the other hand, it offers relatively less flexibility. Only the number of clusters can be changed. The decision is solely based on the maximum value out of c discriminant functions.

- First Order Markov model is an advantageous approach while classifying temporal processes. If the clusters of each class are located on top of each other, the only solution is to track the transitions between the clusters to detect the true class. But still it is hard to determine the true clusters using kmeans algorithm even if the correct number of clusters is provided.
- For the last two methods, they differ from the first two method in that they are hybrid methods. For instance, in the last part, using decision tree one can prioritize the results of SVM method using the distances between the means of the classes as a measure of the reliability of each SVM model. By this way, one can procastinate the decision for the unreliable data to the last nodes of the tree. With the comfort of having get rid of the decision of other samples, one can set a decision rule for the remaining more easily.
- The overall conclusion is that, before implementing a solution onto a data, it is important to visualize the dataset. After assessing the classification needs of the data, then there should be the search for a convenient method.

Following table demonstrates methods vs their respective recall and precision rates for each class.

Table 1: Recall and precision rates for each class, presented for each method applied

Method number		Chair	Chair Set	Sofa	Human	Desk
Method1 (Results may vary)	Precision	0.7237	0.6305	0.7536	0.9943	0.5961
	Recall	0.9136	0.8533	0.7317	0.7317	0.7320
Method2	Precision	0.3333	0.0000	0.5000	1.0000	0.42
	Recall	0.3333	0.0000	0.3333	1.0000	1.0000
Method3	Precision	0.8305	0.8213	0.8370	0.9619	0.7354
	Recall	0.8277	0.9357	0.7362	0.9909	0.7086

Method4	Precision	0.9930	0.9006	0.8498	0.9725	0.9580
	Recall	0.9559	0.9655	0.9852	0.8299	0.9416

Appendix