Pattern Recognition Assignment IV

Muhammad Arsath K F(CS16S035) and Srihari Maruthachalam (CS16S024) Group 17

November 1, 2017

1 Dynamic Time Warping(DTW)

1.1 Isolated Digits

Intuition

We segregated reference and test utterance MFCC files. We computed DTW distance across all reference and test utterance. We classify the test utterance based on the minimum DTW distance.

1.1.1 Confusion Matrix

		Target Classes		
		1	**9**	**z**
Predicted Classes	**1**	**14**	1	0
	9	1	**15**	1
	z	2	1	**16**

Table 1: Classification accuracy for Isolated Digits using DTW

Inference

- As expected, DTW technique performing decently on isolated digit utterance.
- The number of references should be chosen wisely. Otherwise, DTW distance computation becomes inefficient.
- ROC and DET curves are shown below on the following page.

1.2 Connected Digits

Intuition

For connected digits, we have to check for each digits DTW distance. To find the boundary of each digit in the utterance, the window size is varied from the quarter of the length of each digit utterance to four times of the same. We will classify each digit based on the minimum DTW distance. Estimated window size gives boundary for next digit.

1.2.1 Recognition of Connected Digits using DTW

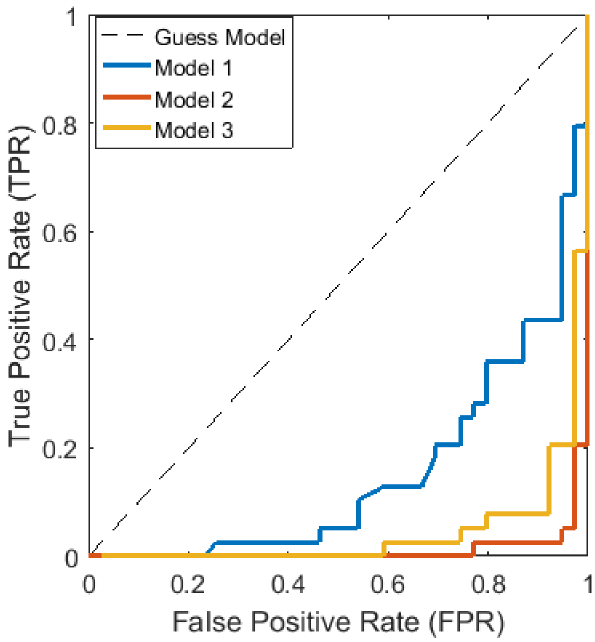
Target	**Obtained**
11z	111
1zz	1z1
91z	911
9zzz	911
z1	11

Target	**Obtained**
zz1	z11
19	91
1zzz1	1zzz1z11
991	911
z1z	z11

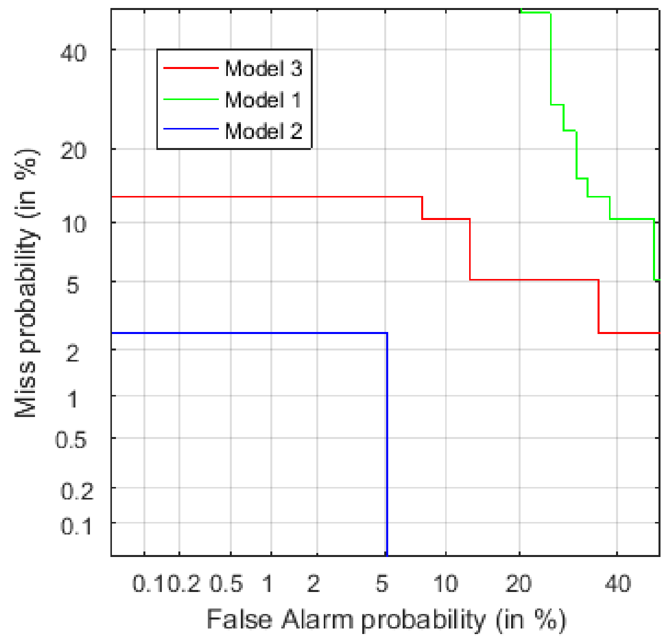
Target	**Obtained**
zz	z11
1z	11
91	11
99	11
z19	z11

Table 2: Recognition Connected Digits using DTW

1.2.2 Analysis on Dynamic Time Warping on Isolated Digits



(a) ROC Curves for Dynamic Time Warping on Isolated Digits



(b) DET Curves for Dynamic Time Warping on Isolated Digits

Figure 1: Classification for Dynamic Time Warping on Isolated Digits

2 Hidden Markov Model(HMM)

2.1 TIDIGIT Dataset

Intuition and Preprocessing

As we provide with Discrete Hidden Markov Model(DHMM) code, we have to feed Vector Quantized MFCC feature vectors to the system to train the HMM model and test the utterance. In order to quantize the vectors, we clustered the MFCC feature vectors using K Means and obtained the corresponding index.

2.1.1 Isolated Digits

Intuition

Vector Quantized MFCC features are feed to DHMM code to generate the HMM model. For each vector quantized test utterance, we compare the likelihood from every HMM models generated for isolated digits. We classify the test utterance based on higher likelihood values.

We have to choose number of clusters to quantize the MFCC features and number of states to incorporate in the HMM model.

Confusion Matrix

		Target Classes		
		1	**9**	**z**
Predicted Classes	**1**	**17**	1	0
	9	0	**16**	0
	z	0	0	**17**

Table 3: Classification accuracy for Isolated Digits using HMM

Inference

- As we increase the number of clusters in K Means, accuracy was kept on increasing. Once we reached cluster count of 60, we obtained around cent percent accuracy.
- To infer, we kept on increase increasing the cluster count to 120, accuracy was stable for clusters more than 60.
- We tested the HMM model performance for universal MFCC features rather than allocated group MFCC features. The accuracy of the model was declined for cluster count of 60. As we increase the cluster count, we could get better results. For universal MFCC features, we obtained cent percent accuracy with cluster count of 100.
- For estimating the number of states for HMM, we started with the state count of 2. We obtained decent accuracy, yet not cent percent. For 3 states, we could get cent percent accuracy. As we increase the number of states, HMM model's performance was stable.
- Since we got cent percent accuracy, we are not showing ROC and DET curves.

2.1.2 Connected Digits

Intuition

From the HMM models generate for the Isolated digits, we can generate HMM models for connected digits. We introduce a new dummy state in between two HMM models. Transition probability from dummy state to next state is a unit and same is zero for dummy state to the same state.

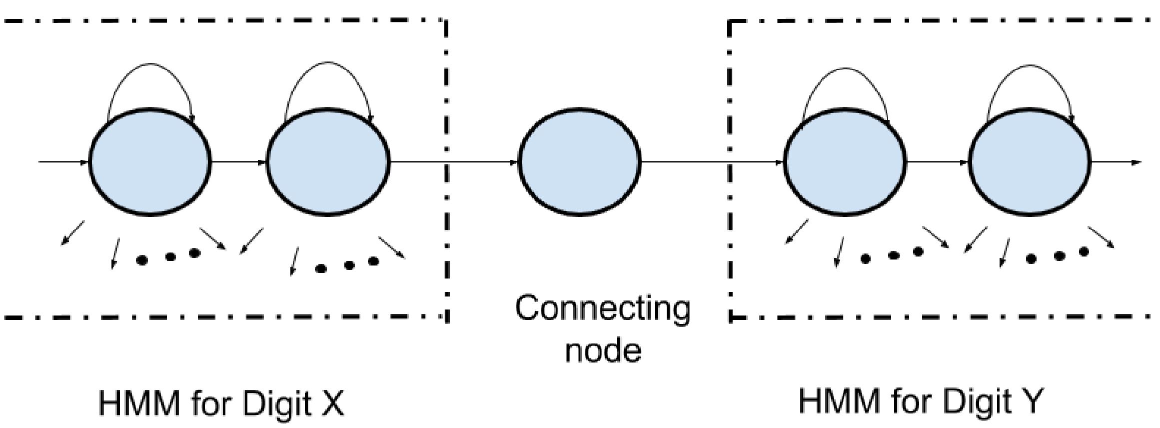


Figure 2: Introduction of new dummy states in between two HMM models

Recognition of Connected Digits using HMM

Target	Obtained
11z	z1z
1zz	1z9
91z	z1z
9zzz	z9z
z1	z9

Target	Obtained
zz1	z99
19	19
1zzz	1zz
991	91z
z1z	z9z

Target	Obtained
zz	z99
1z	9z
91	z1
99	z99
z19	z11

Table 4: Recognition of Connected Digits using HMM for Known Utterance

Obtained
11z
1zz
91z
91z

Obtained
z19
19z
z1z
99z

Obtained
z9z
z99
91z
19z

Table 5: Recognition of Connected Digits using HMM for Unknown Utterance

Inference

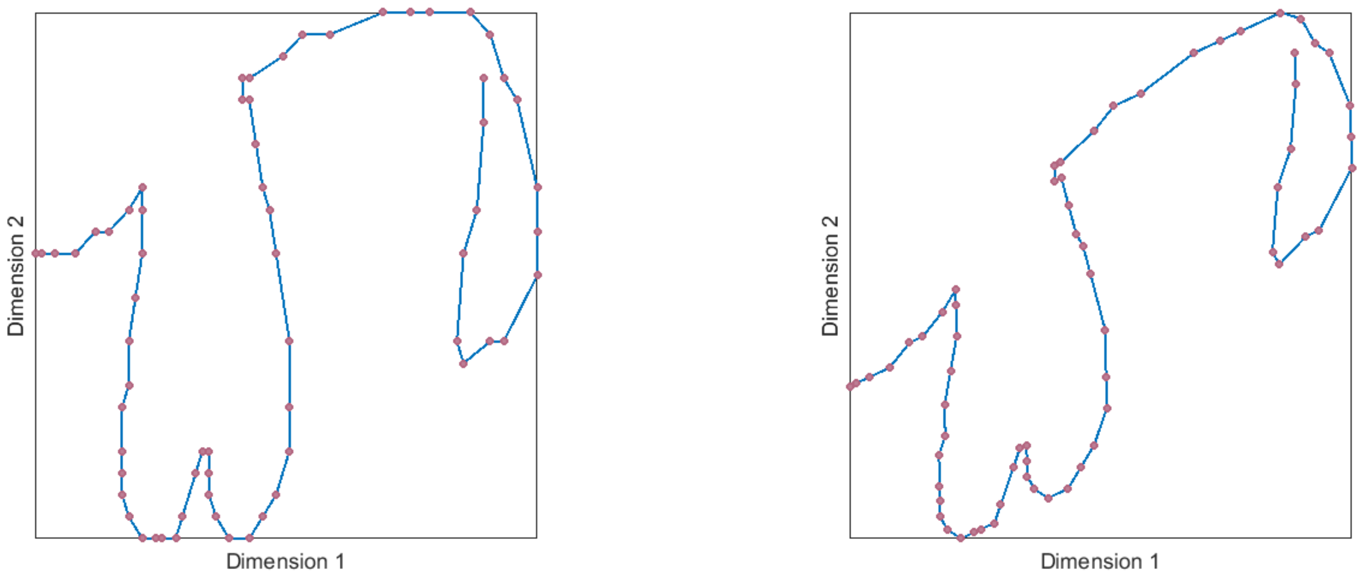
- We have shown the connected digits recognition which is best we obtained in the various configuration of a number of K Means clustering and number of states in the HMM. As we can see that accuracy of connected digits utterance is not cent percent. The model produces various kinds of errors say, modification, deletion, and insertion.
- We tried connected digits recognition with the vector quantization from group-specific MFCC features and universal MFCC features. We found that model performs better in group-specific vector quantization when a number of clusters are less and model performs better in universal vector quantization when the number of clusters is higher.
- One of the major observation is that, even though connected digits recognition performs poorly in most of the configuration, the first digit of the utterance is recognized properly.

2.2 Online Handwritten Dataset

2.2.1 Up-sampling of Data

Intuition and Inference

We are provided with around 100 samples per each online handwritten character to build the HMM models. Since we got raw data, we had an intuition to up-sample the raw data and extract the features. The idea is to rotate the character using rotation matrix as rotation covers variants of the same character.



(a) Given Original Character (b) Rotated Character

Figure 3: Up-sampling of the Character Data

2.2.2 Feature Extraction

Once we upsampled the given online handwritten characters, we extracted gradient and curvature features.

Gradient Feature

We extracted gradient features from the online handwritten data using the following formula,

$$k = \frac{\sum_{i=1}^{2} i * (x_{i+1} - x_{i-1})}{2 * \sum_{i=1}^{2} i^2}$$

Curvature Feature

We extracted curvature features from the online handwritten data using the following formula,

$$k = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{\frac{3}{2}}}$$

Data Normalization

Once we obtained the gradient and curvature features, we normalized the features.

2.2.3 Confusion Matrix for Online Handwritten Isolated Characters

		Target Classes		
		chA	**dA**	**lA**
Predicted Classes	**chA**	**23**	11	4
	dA	6	**17**	1
	lA	1	2	**25**

Table 6: Classification accuracy for Isolated Digits using HMM

Inference

- We build the HMM models from the normalized feature vectors. We also attempted to build the HMM models from unnormalized feature vectors. But, performances of both the models are almost same.
- As we increase the number of clusters in K Means, accuracy was kept on increasing. Once we reached cluster count of 40, we got our best accuracy which is around 75%. On the other hand, when we increase the cluster count from 40, accuracy was declined.
- For estimating the number of states for HMM, we started with the state count of 2. As we increase the number of states, HMM model's performance was improving. We got our best accuracy which is around 75% on the state count of 5.
- ROC and DET curves are shown above.

2.2.4 Online Handwritten Connected Characters Recognition

Intuition

From the HMM models generate for the online isolated characters, we can generate HMM models for connected online characters. We introduce a new dummy state in between two HMM models. Transition probability from dummy state to next state is the unit and same is zero for dummy state to the same state.

Target	**Obtained**
chA lA lA	chA chA chA
chA chA lA	chA chA chA
chA chA dA	chA chA dA

Table 7: Recognition of Online Connected Characters using HMM

2.2.5 Inference

- We have shown the online connected characters recognition which is best we obtained in the various configuration of the number of K Means clustering and number of states in the HMM. As we can see that accuracy of online connected characters recognition was poor. The model produces various kinds of errors say, modification, deletion, and insertion.
- Most of the times, the first character is recognized properly. One of the reasons may be that the starting portion of the consecutive characters is not same as training. Up-sampling of data helped to improve the model performance to some extent.

3 References

- Bishop, Christopher M., "Pattern Recognition and Machine Learning", Information Science and Statistics, 2006.
- R.O.Duda, P.E.Hart and D.G.Stork, "Pattern Classification", John Wiley, 2001.
- M. Pastor, A. Toselli, and E. Vidal, "Writing speed normalization for on-line handwritten text recognition", Eighth International Conference on Document Analysis and Recognition, 2005.
- M. Hassan Shirali-Shahreza, Sajad Shirali-Shahreza, "Effect of MFCC normalization on vector quantization based speaker identification", IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2010.
- Martin, A and R. Doddington, George and Kamm, Terri and Ordowski, Mark and Przybocki, Mark, "The DET curve in assessment of decision task performance", Fifth European Conference on Speech Communication and Technology, EUROSPEECH 1997.
- https://www.mathworks.com/
- NIST DETWare