CS669: Pattern Recognition

Programming Assignment 3

Instructor: Dileep A.D.

By-

Group 12

Ayush Garg (B13111)

Himanshu Singal (B13312)

Hitesh Tarani (B13139)

Part-I: Introduction

Dynamic Time Warping (DTW)

In time series analysis, dynamic time warping (DTW) is an algorithm for measuring similarity between two temporal sequences which may vary in time or speed. For instance, similarities in walking patterns could be detected using DTW, even if one person was walking faster than the other, or if there were accelerations and decelerations during the course of an observation. DTW has been applied to temporal sequences of video, audio, and graphics data — indeed, any data which can be turned into a linear sequence can be analyzed with DTW. A well-known application has been automatic speech recognition, to cope with different speaking speeds.

Hidden Markov Model (HMM)

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. A HMM can be presented as the simplest dynamic Bayesian network. Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics.

Part-II: Experiments and Observations

1) DTW

a) Image dataset

K (Nearest Neighbors)	Confusion matrix	Efficiency
2	62 27 1 4 78 0 17 23 37	71.084
4	57 31 2 3 79 0 14 17 46	73.092
8	48 41 1 1 81 0 8 12 57	74.7
16	43 45 2 1 81 0 6 15 56	72.289
32	42 46 2 0 82 0 6 14 57	72.691

b) Digit Data

K (Nearest Neighbors)	Confusion matrix	Efficiency
2	40 0 0 0 40 0 0 0 40	100
4	39 0 1 0 40 0 0 0 40	99.167
8	40 0 0 0 40 0 0 0 40	100
16	40 0 0 0 40 0 0 0 40	100
32	40 0 0 0 40 0 0 0 40	100

The accuracy in image data is lower than the audio data the reason being we assumed the block of images to be a sequence which is not true while in case of audio signal was a function of time so they were forming a unidirectional signal. The audio data is well distinguishable due to the presence of silence between different digits hence we are able to form such a good model to distinguish almost perfectly between the digits.

2) HMM

a) Image Dataset

N	M	Confusion Matrix	Efficiency
2	32	78 6 6 29 59 2 30 0 60	72.96
4	8	64 22 4 30 58 2 34 0 56	66.93
4	32	82 5 3 40 48 2 28 0 62	71.11
8	32	79 7 4 24 66 0 21 0 69	79.25
16	32	77 8 5 25 65 0 23 1 66	73.33
32	32	72 13 5 26 64 0 23 0 67	75.18
32	64	77 10 3 26 64 0 23 0 67	77.03

b) Digit Data

N	M	Confusion Matrix	Efficiency
5	16	38 0 2 0 29 11 0 0 40	89.16
6	16	40 0 0 0 29 11 0 1 39	90
7	16	38 1 1 0 32 8 0 0 40	91.67
7	32	37 1 2 1 37 2 0 0 40	95
8	32	39 0 1 0 39 1 0 0 40	98.333

In, HMM we observe that if we increase both N, M it doesn't increase the accuracy much, but if we increase only M keeping N fixed it increases the accuracy because when the number of clusters are more the classification is better.

Also, in case of image dataset we get slightly better results as compared to the DTW model as in that case we considered each block to be coming from a left to right sequence but in case of HMM we have considered an ergodic model for the same which is closer to the reality.