RW315
Assignment 5

Eaton Meyer Emmerich
16579070
16579070@sun.ac.za

Introduction

In this document we will investigate the use of a Hidden Markov Model (HMM) for classification of data in a sequence, the example data used will be data for a speech diphthong classification and also a signature classification. The data in the examples will be tested with a HMM with a number of states ranging from one to six.

Speech data example

The speech data used in this example was exactly the same as the data in assignment 4. From the confusion matrices obtained, we can calculate the percentage of errors made by the classification system as follows:

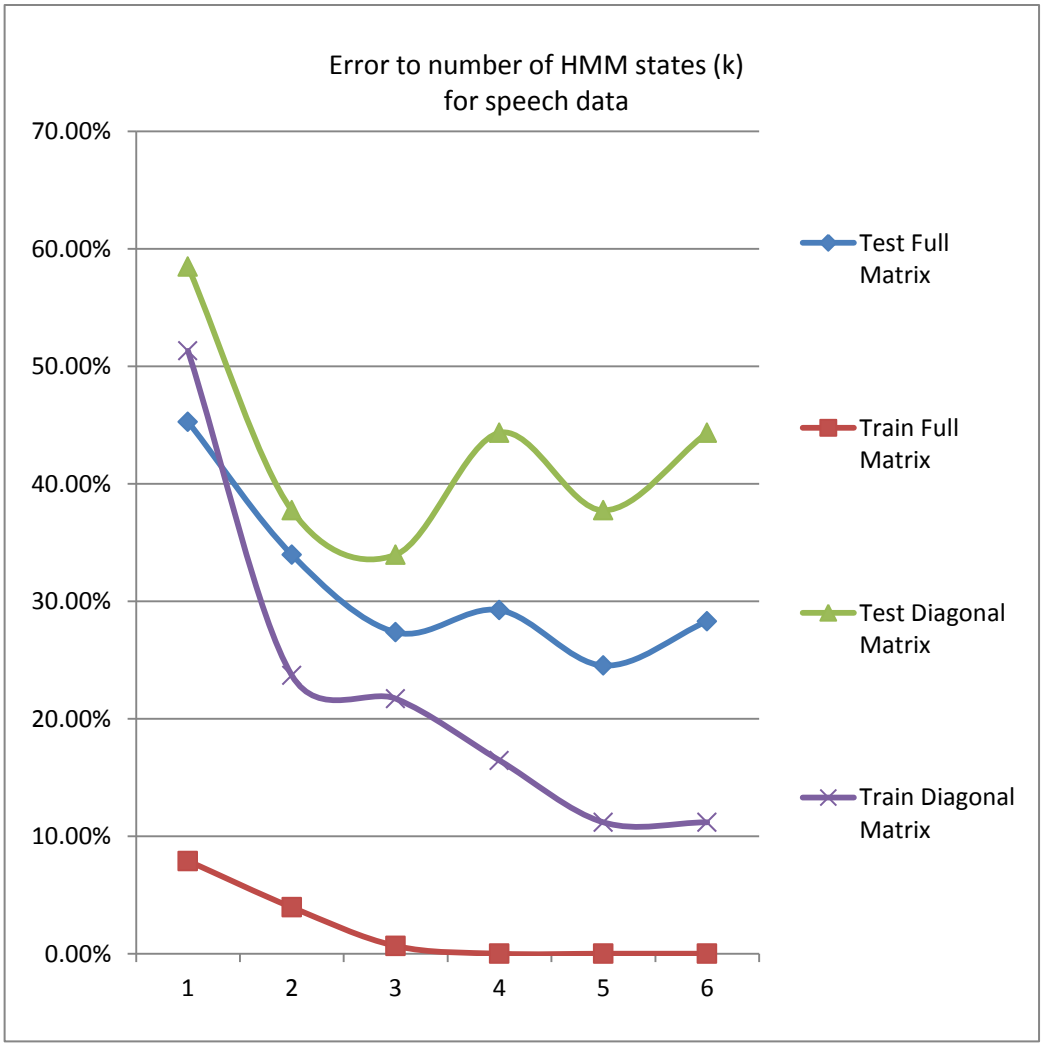

Figure 1

From **Figure 1** we can see the Full Matrix with the number of states in the HMM at 3, the system classifies the data the most correctly. This level of correctness is slightly more than the correctness found in the GMM method. But the correctness still remains at a relatively low value (65% correctness) because the amount of data used to train the models is still too small.

Signature data example

The data in this example is the same data calculated for signatures in assignment 1. This data is split up so the first 3 signatures data is used as the training data and the remaining data is used for the testing data. With this we calculate the percentage of errors (just as in the ***Speech data example***) as follows:

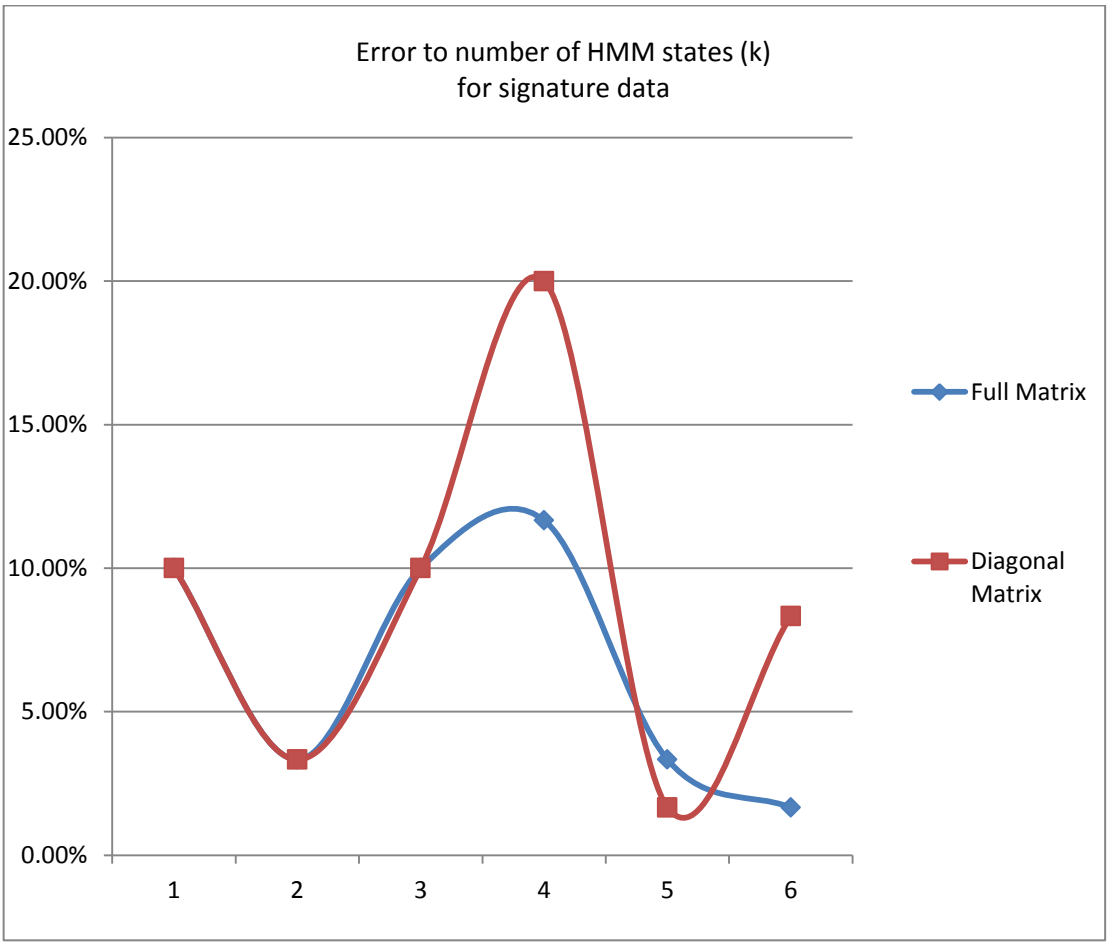

Figure 2

The most accurate classification system is the Diagonal Matrix with five HMM states. It is possible to say that the number of states necessary to accurately distinguish the five signatures is proportional to the number of signatures.

Conclusion

The performance for the HMM is better than the GMM for both the speech and signature data. Although the signature data has a smaller performance boost in the HMM than the speech data, this can be due to the fact that the speech data is a nx16 (n between 4 and 14) dimensional data block that is used for classification, whereas the signature data is a big nx2 (n between 200 and 600) dimensional data block where the classification process needs to fit the entire n-dimension into the system to get a likelihood value.