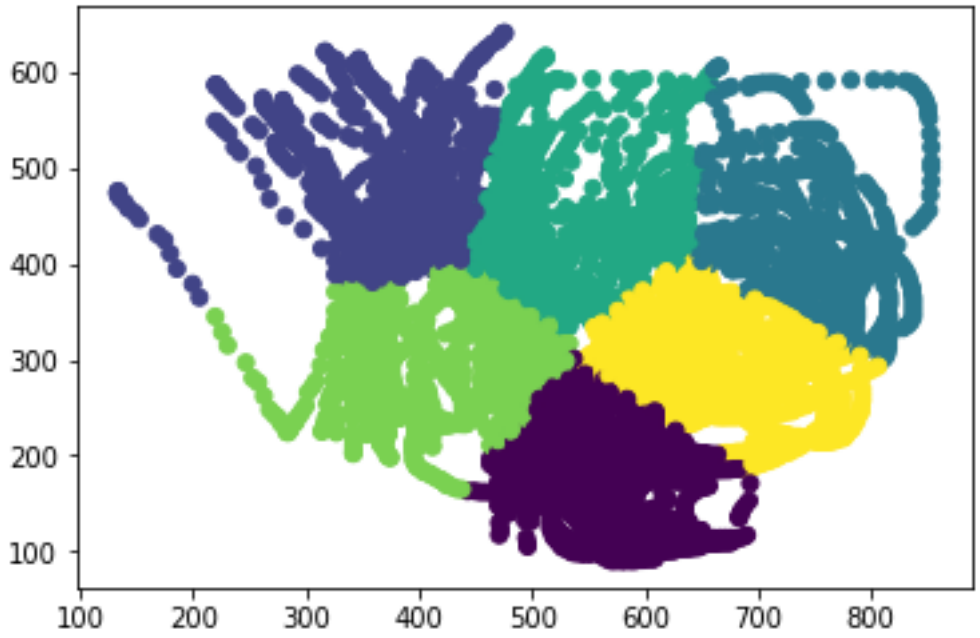
# Project 2:

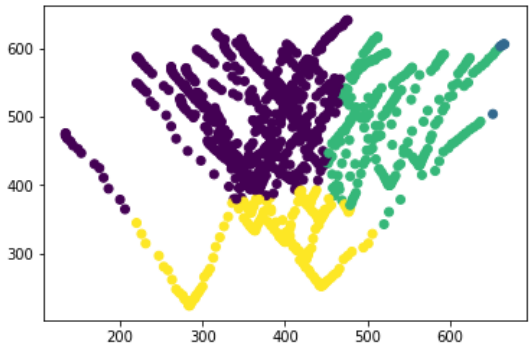
# Pen-gesture Recognition with Hidden Markov Models

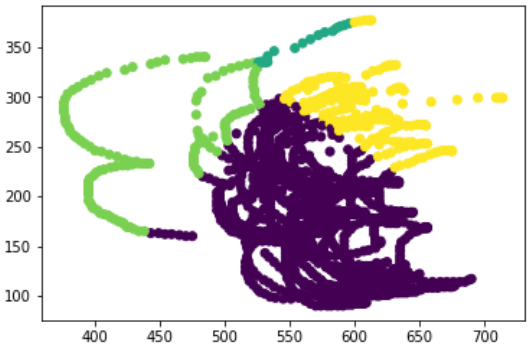
**(1) Data: project2-data.zip . This dataset has 5 files, one for each vowel of the alphabet. These are xml files and the format is self-explanatory. Take the odd indexed entries as training data and the even ones as testing data.**

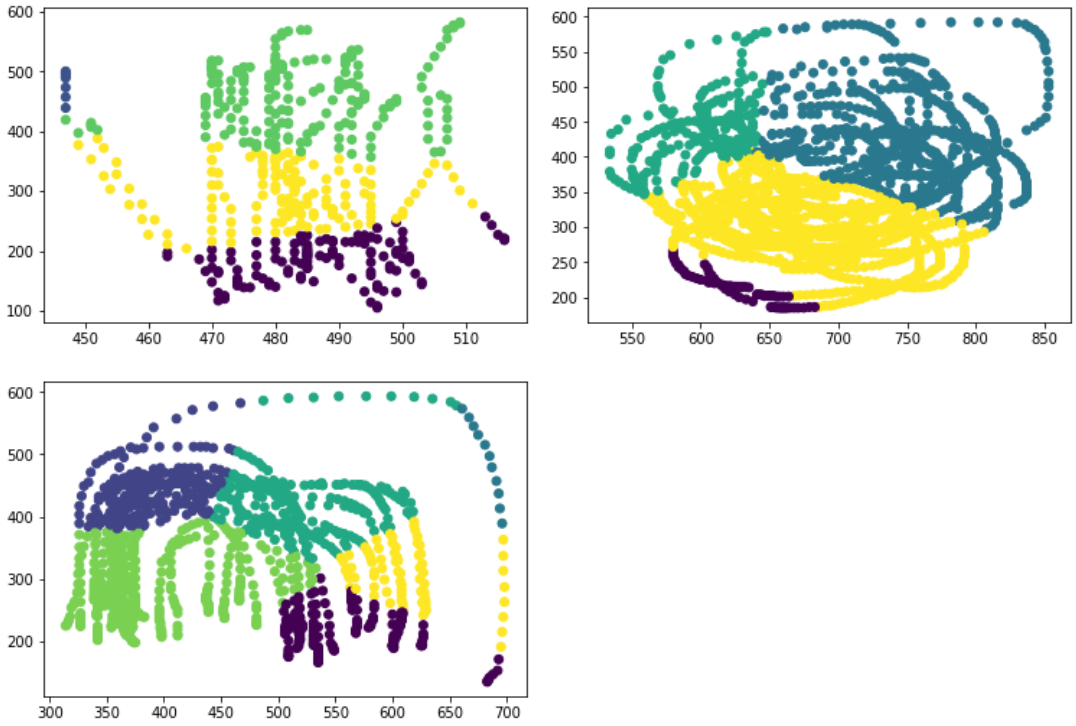
Parameters: observation num=6, state num=8




**(2) Implement a spatial clustering algorithm (e.g., K-Means) on the 2D training data.**






**(3) & (4)**

**Implement the standard Hidden Markov Model. The input will be the cluster index associated with each 2D point.**

**Train a separate HMM for each vowel (do separate clustering for each and then learn the HMM independently). Quantify accuracy against the training and testing set; for each candidate datum, compute its log-likelihood against each HMM and take the label of the HMM that gives the highest log-likelihood. Compute a confusion matrix of results. Describe your findings.**

The figure below shows A, B, PI for hmm mode of a vowel. Because there are five models, it is difficult to show all the parameters in this article, so only the first one is selected as an example. In the bottom of method “bw_train (self, observations, outline=0.1,eps=20)” of class

HMM I commented three lines which can print A,B,pi.

```
A [[0.79402953 0.00000043 0.05470125 0.01366891 0.0024812  0.07081732
0.02223622 0.04206514]
 [0.         0.03340632 0.82745538 0.00051706 0.00017497 0.00003765
0.0027043  0.13570432]
 [0.00539829 0.00000907 0.5476772  0.00011691 0.02739992 0.045551
0.00176135 0.37208625]
 [0.00000002 0.26792916 0.06240949 0.46380843 0.00002481 0.0002539
0.18958655 0.01598764]
 [0.5830748  0.00000125 0.15042163 0.00067941 0.06454844 0.04270947
0.04266561 0.11589939]
 [0.12201479 0.00022018 0.000024   0.30208049 0.00000897 0.0229681
0.55262699 0.00005648]
 [0.00000002 0.06891232 0.04255876 0.64488619 0.00000068 0.00464564
0.23895716 0.00003923]
 [0.09308825 0.00000106 0.75175298 0.02075969 0.08627401 0.04004369
0.00380003 0.00428028]]
B [[0.         0.00024029 0.         0.99975307 0.00000664 0.
]
 [0.         0.20035038 0.11700672 0.         0.6826429  0.        ]
 [0.         1.         0.         0.         0.         0.        ]
 [0.         0.00000025 0.00000001 0.         0.99999974 0.        ]
 [0.         0.95974379 0.         0.03984956 0.00040665 0.        ]
 [0.         0.18367087 0.         0.00159375 0.81473538 0.        ]
 [0.         0.00000044 0.         0.00000005 0.99999951 0.        ]
 [0.         0.99999903 0.         0.00000096 0.00000001 0.        ]]
pi [0.         0.00000005 0.96346697 0.         0.00000058 0.
0.         0.0365324 ]
```

The three figs at the top of the picture below are the confusion matrix for the first three training and testing. The fig at the bottom are the results for 10 training.

```
total train
a 0.95
e 1.0
i 1.0
o 0.95
u 1.0
[[19.  0.  0.  0.  1.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  0.  0. 19.  1.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 1.0
e 1.0
i 0.95
o 0.85
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 1.  0. 19.  0.  0.]
 [ 1.  0.  0. 17.  2.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 1.0
e 1.0
i 1.0
o 0.95
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  0.  0. 19.  1.]
 [ 0.  0.  0.  0. 20.]]
```

```
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9600000000000002
Correct classification of aeiou: 0.99
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.99
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.97
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.97
```

**Analysis:**

The mode can get a correct classification rate of nearly 0.98 for average, which means the mode is trained successfully. Observe the confusion matrix, in every category, the diagonal takes the biggest decision.

In the matrix, vowel “o” get a slightly lower score, I guess that because the vowel “o” is a circle, and the starting point of “o” is not always at the same place, we can write “o” counterclockwise and Clockwise, or even write it from the top, that make the mode difficult to decide which sequence is right.

I tried another way to dispose the data. I normalization five vowel, do k-means mode for each of them and then make HMM for them singly. The result is showed after analysis, which is not as good as train totally. In my opinion, it is because train k-means for each vowel will drop the category message between them, so the sequence may not be so classify as total training.

When make k-means, I use k-means++ to initialize the starting points, it can make the distance of each starting point be biggest, but different fist

chosen point can make different k-means results. However, according to ten consecutive trainings, the initial point selection has little impact on the HMM model.

```
ones train
a 1.0
e 1.0
i 0.1
o 0.3
u 0.7
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  5.  2. 11.  2.]
 [ 0.  0.  0.  6. 14.]
 [ 0.  0.  0.  6. 14.]]
```

```
ones train
a 1.0
e 1.0
i 0.65
o 0.65
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 13.  6.  1.]
 [ 0.  0.  0. 13.  7.]
 [ 0.  0.  0.  0. 20.]]
```

```
ones train
a 1.0
e 1.0
i 0.5
o 0.95
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 10.  8.  2.]
 [ 0.  0.  0. 19.  1.]
 [ 0.  0.  0.  0. 20.]]
```

```
Correct classification of a e i o u: 0.6199999999999999
Correct classification of a e i o u: 0.86
Correct classification of a e i o u: 0.89
Correct classification of a e i o u: 0.9199999999999999
Correct classification of a e i o u: 0.9
Correct classification of a e i o u: 0.7200000000000001
Correct classification of a e i o u: 0.7899999999999999
Correct classification of a e i o u: 0.85
Correct classification of a e i o u: 0.8800000000000001
Correct classification of a e i o u: 0.85
```

**(5) Vary the number of clusters in (2) and the number of hidden nodes (and their connectivity) in (3) and repeat the quantification. Do this for an additional three variants. Describe your findings.**

**(5-1)** N (observation num & clustering num of k-means) =3

State num=3

```
total train
a 1.0
e 1.0
i 0.0
o 0.75
u 0.85
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0.  0.  0. 20.]
 [ 0.  0.  0. 15.  5.]
 [ 1.  0.  2.  0. 17.]]
```

```
total train
a 1.0
e 1.0
i 1.0
o 0.75
u 0.7
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  0.  0. 15.  5.]
 [ 3.  0.  3.  0. 14.]]
```

```
total train
a 1.0
e 1.0
i 0.85
o 0.75
u 0.85
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 17.  0.  3.]
 [ 0.  5.  0. 15.  0.]
 [ 3.  0.  0.  0. 17.]]
```

```
Correct classification of aeiou: 0.72
Correct classification of aeiou: 0.89
Correct classification of aeiou: 0.89
Correct classification of aeiou: 0.74
Correct classification of aeiou: 0.9199999999999999
Correct classification of aeiou: 0.89
Correct classification of aeiou: 0.79
Correct classification of aeiou: 0.72
Correct classification of aeiou: 0.9099999999999999
Correct classification of aeiou: 0.8699999999999999
```

**(5-2)** N (observation num & clustering num of k-means) =4

State num=3

```
total train
a 1.0
e 1.0
i 1.0
o 0.75
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  3.  0. 15.  2.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 1.0
e 1.0
i 1.0
o 0.9
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  0.  0. 18.  2.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 1.0
e 1.0
i 1.0
o 0.9
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  0.  0. 18.  2.]
 [ 0.  0.  0.  0. 20.]]
```

```
Correct classification of aeiou: 0.95
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9400000000000001
Correct classification of aeiou: 0.9
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.95
Correct classification of aeiou: 0.9199999999999999
```

**(5-3)** N (observation num & clustering num of k-means) =12

State num=14

```
total train
a 1.0
e 1.0
i 1.0
o 0.85
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 1.  0.  0. 17.  2.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 1.0
e 1.0
i 1.0
o 0.95
u 1.0
[[20.  0.  0.  0.  0.]
 [ 0. 20.  0.  0.  0.]
 [ 0.  0. 20.  0.  0.]
 [ 0.  1.  0. 19.  0.]
 [ 0.  0.  0.  0. 20.]]
```

```
total train
a 0.9
e 0.95
i 0.9
o 0.85
u 1.0
[[18.  0.  0.  2.  0.]
 [ 1. 19.  0.  0.  0.]
 [ 0.  1. 18.  0.  1.]
 [ 0.  0.  0. 17.  3.]
 [ 0.  0.  0.  0. 20.]]
```

```
Correct classification of aeiou: 0.97
Correct classification of aeiou: 0.99
Correct classification of aeiou: 0.9199999999999999
Correct classification of aeiou: 0.96
Correct classification of aeiou: 0.9800000000000001
Correct classification of aeiou: 0.76
Correct classification of aeiou: 0.95
Correct classification of aeiou: 0.97
Correct classification of aeiou: 0.9400000000000001
Correct classification of aeiou: 0.97
```

**Analysis:**

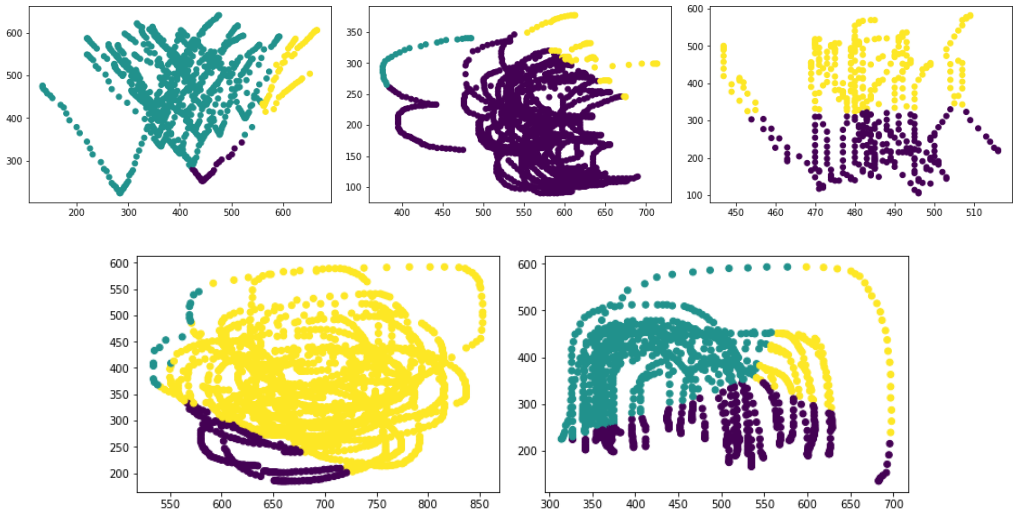
In the process of constant adjustment of parameters, I found a lot of interesting changes and characteristics, but unfortunately, the title requires only three additional cases. Here, I chose three groups of parameters that I think can best reflect the law for analysis.

(1) In (5-1), the observation num is 3 and the state num is 3 as well, the results take an average of 0.83, and in (5-2), It is really strange that only the observation num went from 3 to 4, the correct classification rate has been greatly improved.

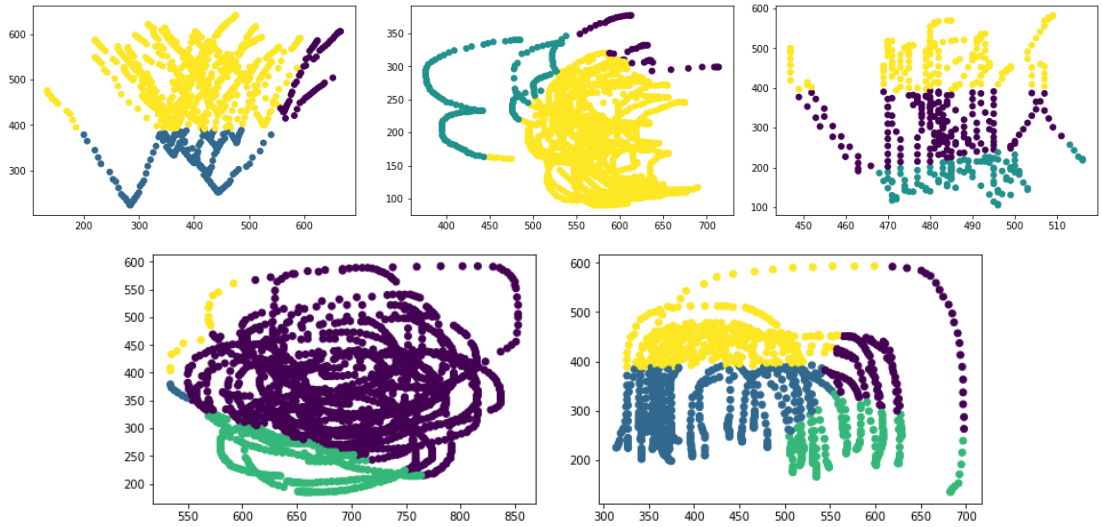
Let's observe the figs below, in 3 observation num, vowel "a" transfer from green to green and "e" transfer from purple to purple, which have distinct own transfer processes. But "i" and "o" may be considered as a subsequence of "u" or "e". In 4 observation num, each vowel has a much more complex sequence than in 3, so the result can be much better.

There is another problem is overfitting, the hidden state and observation state is too less, so the mode may consider that only a immobilization sequence can possibly be a letter, which loss the generalization ability.

**3 observation num:**



**4 observation num:**



(2) When increasing the num of hidden state and observation state, the correct classification rate increase, when it touch a threshold value, the rate be the biggest and even decrease while the num increase. In (5-3) the num of observation and hidden is 12 and 14, which did not bring better effect, it is because that the data is not as much as training a mode with too many states.