README

1. Environment: You should have python 3, latest version of keras, tensorflow 1.7,R studio installed in your computer.

2. Check the files in you file folder. You should have(at least): CNN1D.py, HMM.py, MA0007.2.sites, RNN.py, one_hot_encoding_CNN.py
, split.py.

3. Run in your command line:

```
$python3 split.py
```

You will get 6 files: human_perf_trainset.txt, human_perf_validation.txt, human_perf_test.txt are training, validation, test data set for perfect data; human_imperf_trainset.txt, human_imperf_validation.txt, human_imperf_test.txt are training, validation, test data set for imperfect data.

4. Run in your command line:

```
$python3 CNN1d.py human_perf_trainset.txt human_perf_validation.txt human_perf_test.txt 20

$python3 CNN1d.py human_imperf_trainset.txt human_imperf_validation.txt human_imperf_test.txt 20
```

This is to train, validation, test CNN model. The first three terms are training, validation, test data set. The last term is the number of epochs you want the program to run. Run both imperfect data and perfect data to compare. Usually, 20 is enough for CNN.

5. Run in your command line:

```
$python3 HMM.py    hmm.py human_perf_trainset.txt human_perf_test.txt

$python3 HMM.py    hmm.py human_imperf_trainset.txt human_imperf_test.txt
```

The first term is the training set and the second term is the test set. You will compare the difference between the imperfect and perfect data set.

6. Run in your command line, you will see the result of RNN:

```
$python3 HMM.py    RNN.py human_perf_trainset.txt human_perf_test.txt

$python3 HMM.py    RNN.py human_imperf_trainset.txt human_imperf_test.txt
```

7. Open your R studio, run plot.r, you will see the plots that shown in the report.