Readme

Please read this file before running the program

1. Python Shell:
This program is running using python3.8.5.

2. Imported Package:
This program only uses NumPy
numpy==1.16.4

3. Data
Train.txt
Test.txt
Test_gold.txt

4. Running Time:
The total running time of this program is about 10 seconds.

5. Code Structure
Questions are written in class and function
All the functions have been merged into the "main()" function

6. Meaning of each function:
    __init__: I will initialize sets to store my data. This function will include:
        self.train = train_data
        self.test = test_data
        self.label = test_label
        self.total = None
        self.root = None
        self.tran = None
        self.pb = 0
        self.ps = 0
        self.dict_list = None
        self.whole = None
            these are all variables to obtain the main information.

    def compute_start_prob(self):
        we can get self.pb and self.ps

    def compute_transition(self):
        we can get the transition matrix's transpose

    def compute_emission(self):
        the emission matrix is presented as four Dict(after Laplace smoothing)

```
def viterbi_decoding(self):
    I use two matrixes to contain: 1). The prob 2). The Last positions
    These two can help to decoding.

def word_segmentation(self):
    use str 切片, we can easily do that, and we write it into the txt file.


Def main():
In this function, we call the first five function, and also calculate the F1-score.
```

7. Result
The precision is: 75.62%, The recall is: 76.41%, the f1 score is: 0.76013.

Thank you for reading!