

# Project 1 Implementation Report

Hanzhang Yin

October 19, 2024

For my implementation, I used two different strategy to address the author prediction task.

## Log-likelihood Calculation:

### Zhu Shuzhen's Markov Matrix:

$$\begin{bmatrix} \frac{3}{10} & \frac{47}{190} & \frac{29}{190} & \frac{3}{10} \\ \frac{53}{156} & \frac{1}{4} & \frac{11}{78} & \frac{7}{26} \\ \frac{35}{103} & \frac{27}{103} & \frac{15}{103} & \frac{26}{103} \\ \frac{55}{199} & \frac{56}{199} & \frac{27}{199} & \frac{61}{199} \end{bmatrix}$$

### Du Fu's Markov Matrix:

$$\begin{bmatrix} \frac{1}{88} & \frac{29}{103} & \frac{46}{309} & \frac{73}{309} \\ \frac{279}{27} & \frac{279}{41} & \frac{279}{1} & \frac{93}{49} \\ \frac{80}{69} & \frac{160}{37} & \frac{10}{21} & \frac{160}{51} \\ \frac{236}{236} & \frac{118}{118} & \frac{118}{118} & \frac{236}{236} \end{bmatrix}$$

### Test Case for Zhu Shuzhen:

#### Total Log-Likelihood for Zhu Shuzhen:

$$\begin{aligned} & 3 \log \left( \frac{35}{103} \right) + 5 \log \left( \frac{53}{156} \right) + 4 \log \left( \frac{61}{199} \right) + 3 \log \left( \frac{3}{10} \right) \\ & + 6 \log \left( \frac{56}{199} \right) + 3 \log \left( \frac{55}{199} \right) + 2 \log \left( \frac{7}{26} \right) + 5 \log \left( \frac{26}{103} \right) \\ & + 5 \log \left( \frac{1}{4} \right) + 3 \log \left( \frac{47}{190} \right) + 4 \log \left( \frac{29}{190} \right) + 2 \log \left( \frac{15}{103} \right) \\ & + 2 \log \left( \frac{11}{78} \right) + \log \left( \frac{27}{199} \right) \end{aligned}$$

### Total Log-Likelihood for Du Fu:

$$\begin{aligned} & 3 \log \left( \frac{27}{80} \right) + \log \left( \frac{1}{3} \right) + 5 \log \left( \frac{88}{279} \right) + 6 \log \left( \frac{37}{118} \right) \\ & + 5 \log \left( \frac{49}{160} \right) + 3 \log \left( \frac{69}{236} \right) + 3 \log \left( \frac{29}{103} \right) + 5 \log \left( \frac{73}{279} \right) \\ & + 2 \log \left( \frac{23}{93} \right) + 2 \log \left( \frac{73}{309} \right) + 4 \log \left( \frac{51}{236} \right) + \log \left( \frac{21}{118} \right) \\ & + 2 \log \left( \frac{49}{279} \right) + 4 \log \left( \frac{46}{309} \right) + 2 \log \left( \frac{1}{10} \right) \end{aligned}$$

### Predicted Author for Zhu Shuzhen's Test Tones:

Zhu Shuzhen

### Test Case for Zhu Shuzhen:

#### Total Log-Likelihood for Zhu Shuzhen:

$$\begin{aligned} & 10 \log \left( \frac{35}{103} \right) + 9 \log \left( \frac{53}{156} \right) + 5 \log \left( \frac{61}{199} \right) + 19 \log \left( \frac{3}{10} \right) \\ & + 6 \log \left( \frac{56}{199} \right) + 7 \log \left( \frac{55}{199} \right) + 9 \log \left( \frac{7}{26} \right) + 4 \log \left( \frac{27}{103} \right) \\ & + 3 \log \left( \frac{26}{103} \right) + 3 \log \left( \frac{1}{4} \right) + 18 \log \left( \frac{47}{190} \right) + 2 \log \left( \frac{29}{190} \right) \\ & + \log \left( \frac{15}{103} \right) + 7 \log \left( \frac{11}{78} \right) + 5 \log \left( \frac{27}{199} \right) \end{aligned}$$

### Total Log-Likelihood for Du Fu:

$$\begin{aligned} & 10 \log \left( \frac{27}{80} \right) + 11 \log \left( \frac{1}{3} \right) + 9 \log \left( \frac{88}{279} \right) + 6 \log \left( \frac{37}{118} \right) \\ & + 3 \log \left( \frac{49}{160} \right) + 7 \log \left( \frac{69}{236} \right) + 18 \log \left( \frac{29}{103} \right) + 3 \log \left( \frac{73}{279} \right) \\ & + 4 \log \left( \frac{41}{160} \right) + 9 \log \left( \frac{23}{93} \right) + 8 \log \left( \frac{73}{309} \right) + 5 \log \left( \frac{51}{236} \right) \\ & + 5 \log \left( \frac{21}{118} \right) + 7 \log \left( \frac{49}{279} \right) + 2 \log \left( \frac{46}{309} \right) + \log \left( \frac{1}{10} \right) \end{aligned}$$

### Predicted Author for Du Fu's Test Tones:

Du Fu

## **Euclidean Distance and Consine Similarity Calculation:**

### **Predicted Author for Zhu Shuzhen's Test Tones:**

- Cosine Similarity with Zhu Shuzhen: 0.9789041921911098
- Cosine Similarity with Du Fu: 0.9729092716043023
- Euclidean Distance to Zhu Shuzhen: 0.10495126147082444
- Euclidean Distance to Du Fu: 0.11872876556336236
- Cosine Similarity Prediction: Zhu Shuzhen
- Euclidean Distance Prediction: Zhu Shuzhen

### **Predicted Author for Du Fu's Test Tones:**

- Cosine Similarity with Zhu Shuzhen: 0.9598323728518315
- Cosine Similarity with Du Fu: 0.9755859048660054
- Euclidean Distance to Zhu Shuzhen: 0.1482684467244978
- Euclidean Distance to Du Fu: 0.11595078252667476
- Cosine Similarity Prediction: Du Fu
- Euclidean Distance Prediction: Du Fu

## **Baseline Method - Infinity Vector Norm Calculation:**

### **Baseline Infinity Norm Prediction for Zhu Shuzhen's test tones:**

- Infinity Norm Difference with Zhu Shuzhen: 0.07692068226362686
- Infinity Norm Difference with Du Fu: 0.09668192553969868
- Baseline Infinity Norm Prediction: Zhu Shuzhen

### **Baseline Infinity Norm Prediction for Du Fu's test tones:**

- Infinity Norm Difference with Zhu Shuzhen: 0.12762972269326903
- Infinity Norm Difference with Du Fu: 0.08415719396641286
- Baseline Infinity Norm Prediction: Du Fu