* You can get the link of slides at E3 “Hands-on 3”
* We will start the course at 9:00.

109-2
Natural Language Processing

Lab - 3

Yun-Zhu Song , Yi-Syuan Chen

Colab Link HackMD Link

Announcement

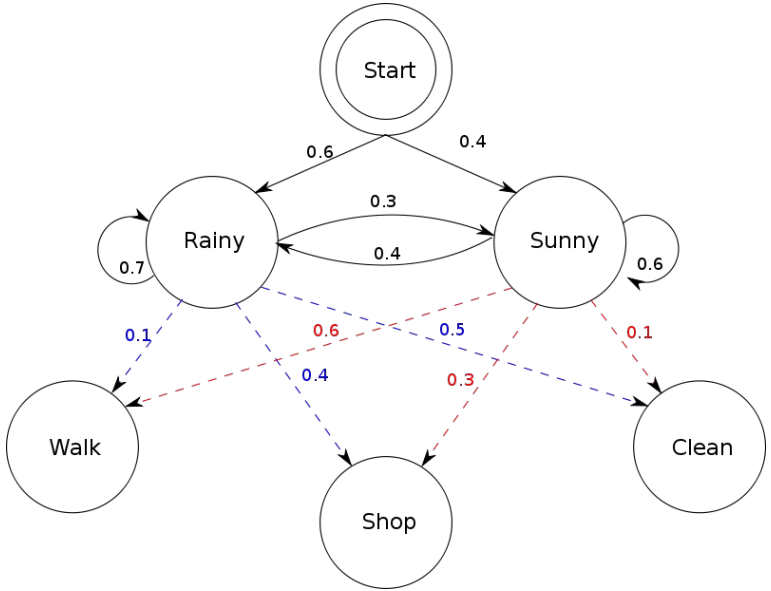
期末專題組隊

Topics

- Hidden Markov Models
  - Build a HMM
  - Sample from the HMM

Concept

- Weather Guessing Game
- You and your partner are in different place.
- Your partner will do different activities according to the weather.
  - Activities : walk / shop / clean
  - Weather : rainy / sunny
- You only know your partner's activity.
  - Hidden variables : weather
  - Observed data : activity
- Then we have several problems
  - Evaluation problem
  - Decoding problem
  - Learning problem



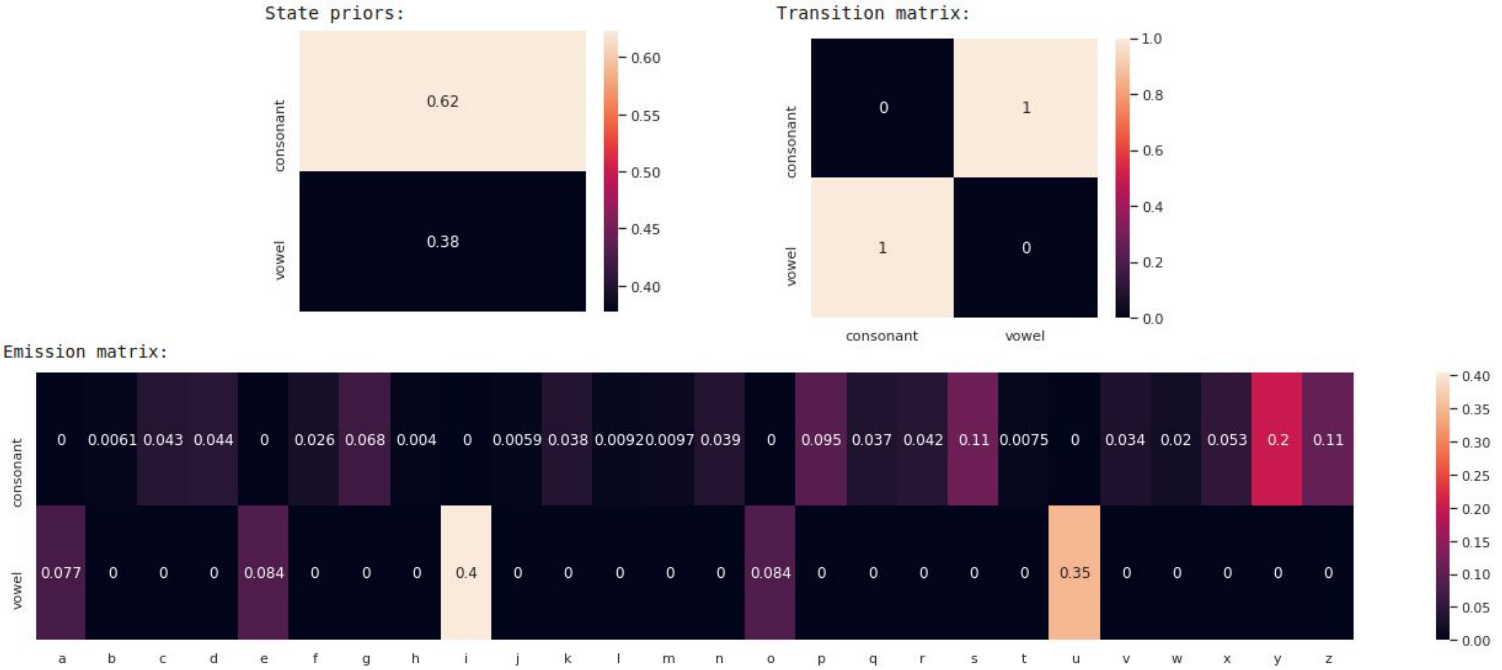
Problem

- Create fake words with HMM !
  - 2 hidden variables : consonants and vowels
  - 24 observations : a ~ z
- Problem 1 : build a HMM model
  - Set the model parameters with some constraints
    - For the first character, it is **more likely** to be a consonant.
    - In consonant state, it should **only emit** consonant observation, and vice versa.
    - The fake words should **not contain** consecutive consonants or vowels.
  - Plot the model parameters
- Problem 2 : Sample fake words from the HMM
  - Create 4 fake words with 5 characters

Requirements

Your parameters will be different due to randomness. Just make sure you fulfill the constraint.

- Problem 1 : build the HMM model
  - Please set and plot the model parameters.



Requirements

- Problem 2 : Sample fake words from the HMM
  - Please generate 4 fake words with 5 characters.
  - Print the characters and corresponding hidden state as following

```
x: geviz
z: [0, 1, 0, 1, 0]
x: uqeqa
z: [1, 0, 1, 0, 1]
x: weveb
z: [0, 1, 0, 1, 0]
x: asaxu
z: [1, 0, 1, 0, 1]
```

Submission

- Deadline
  - Submit files to E3 before **11:59 AM (before the class)** today.
  - Accept late submissions before **11:59 PM** today (80% of original scores).
  - No grades after **11:59 PM** today.
- Format
  - Lab_3_<StudentID>.ipynb (ex: Lab_3_1234567.ipynb)
  - Make sure the .ipynb file contains the **correct execution results**.

Grading Policy

- Correctness
  - We have 2 problems, and **the points are 50/50.**
  - Use **“Runtime/Restart and run all”** to check your execution results on Colab.
- Clarity
  - Write your code with clear logic, and documents it properly.