Brianna Murray

Topic:
Plants

Website:
Texas Tulips
`https://texas-tulips.com/?utm_source=local&utm_medium=organic&utm_campaign=gmb`

The topic was easy to settle on however, I struggled finding a website with enough links to generate a good corpus. My go to website calloways.com actually blocked my web crawler from being able to pull any links. All `link.get('href')` responses I got back were coming back blank or saying the urls were disconnected / private. I found that very odd. The other plant shops I like to frequent (ThePlantProject / Jade&Clover) had much smaller websites that either did not have links from other domains on them or did not have enough url pages to pull from. After doing some digging I found that the Texas Tulips website had enough links for me to build a strong corpus and had some external links as well.

The data I pulled from the Texas Tulips website is all about information relating to the field and experience. Things such as parking, cost, open season, and contact information.

Data Cleaning / Management
I started by putting all words into a .txt file after removing the newlines and tabs. Then I tokenized each sentence and wrote the tokenized sentences into a new file. After going through all files, I tokenized each word in the sentence and lowercase all words then removed stopwords and punctuation. I struggled with how to wrap my mind around managing each page of text individually while also going through all pages. To help with this I kept each page in their assigned .txt file until I needed them. However, when doing that the tokenized text would always be converted back into a string so I would have to tokenize it again before manipulating it again. Once I got to the step where term frequency needed to be determined I began saving everything into a dictionary which was easier to pass back and forth.

Knowledge Base Creation
I created this knowledge base as a python dictionary. I did this because it was the most comprehensive for me and allowed for much wiggle room in the early phases of the Chatbot knowledge base. When looking at the knowledge base you can see the top words all with their own list of responses. While all of the responses relate to the given word, not all of the responses will be relevant to the user if chosen at random. With that in mind, I plan on making subgroups of the responses so that the responses are more specific to the user's question. Also, I think there should be a few more keywords added to the knowledge base that would help the user with their Chatbot interactions (ex. Season and fee). In the Chatbot project, I plan on adding those keywords for better results if allowed in the project rubric.

I view the knowledge base I created for this project as the draft / rough outline of what I want the Chatbot knowledge base to contain. I do not feel as though the current knowledge base I have created is currently sufficient for the Chatbot project yet because there are many more facts /

responses I would like to add based on possible user questions. In order to broaden my perspective on the types of questions people may have for the chatbot I will survey some friends and family to gain more insight on how I can improve the knowledge base for the Chatbot.

Top 10 Terms
Tulips, Dallas, Texas, Spring, Contact, Picking, April, Visit, Field, Activities, Parking

Brianna Murray

Knowledge Base

```python
# mini knowledge base w/ basic facts to build out later
# I think the word season should be there
knowledgeBase = {
    'tulips': ['Tulips come in many colors such as purple, red, orange, yellow, white, and pink',
                'Tulip Care: Cut stems and place in fresh water',
                'There are 100 varieties of tulips on the grounds'
            ],
    'dallas': ['Address: North of Dallas at 10656 FM2931, Pilot Point TX 76258'
            ],
    'texas': ['Address: North of Dallas at 10656 FM2931, Pilot Point TX 76258'
            ],
    'spring': ['Open Season: During end of winter and early spring',
                'Open during February, March, and mid April'
            ],
    'contact': ['Address: North of Dallas at 10656 FM2931, Pilot Point TX 76258',
                'Email: info@texas-tulips.com',
                'Phone: 940-440-0232',
                'Bulb planting services. Email/Call for a quote.'
            ],
    'picking': ['Open Season: During end of winter and early spring',
                'You may pick tulips for $2.50/stem. All tulips will be wrapped and put in flower gel',
                'Picking baskets are provided on the field'
            ],
    'april': ['Open Season: During end of winter and early spring',
                'End Season: Mid April'
            ],
    'visit': ['Hours: 10AM - 8PM Everyday',
                'Address: North of Dallas at 10656 FM2931, Pilot Point TX 76258',
                'Free parking'
            ],
    'field': ['Hours: 10AM - 8PM Everyday',
                'Dogs are NOT allowed in the field',
                'Picnic tables and restrooms available',
                'No drones allowed'
            ],
    'accept': ['Tulip Picking Entrance Fee: $5/person',
                'Discounts for Veterans and Seniors: $7.50/person (includes three tulips & only during the week on business days)',
                'Accepted forms of payment: Visa, MasterCard, and Cash',
                'On Site Cafe',
                'Professional photographers allowed for $25 entrance fee and $5 per client'
            ],
    'parking': ['Parking Fee: FREE'
            ]
}
```

Brianna Murray

Sample Dialog

Hello! How may I help you today?

Where is the Texas Tulips field?

The address to Texas Tulips is North of Dallas at 10656 FM2931 Pilot Point, TX 76258

How else can I help you today?

Can I bring my dog?

Unfortunately, dogs are NOT allowed in the field (with the exception of service dogs).

How else can I help you today?

How much does it cost to enter?

The entrance fee is $5/person. There are also senior and veteran discounts. Would you like more information?

Sure

Veteran & Senior discount rate is only offered during the week for $7.50/person (including 3 tulips).

How else can I help you today?

How much does it cost to pick tulips?

Tulips are $2.50/stem.