

List five types of text data that are not already covered by the list on slides 16-18 of the lecture. For each of the data types, suggest what a question that could be analyzed based on a collection of such data.

### 1.1

01. Business and Personal Email/ Personal Chat
02. Course/Classroom Lectures
03. Digital Advertisement
04. Documentation of a Software/Product
05. Public speech of program
06. News and movie script

**Exercise 1.2:** Chatterbots. Perform a conversation with an online chatterbot, such as Mitsuku or another chatterbot. Try to make the bot reveal that it is a bot. Report the name and web address of the chatterbot and your conversation with the bot in your answer.

Bot: Hi there, I'm Kuki

Bot : We could start by getting to know each other if you like . What is your name?

Me : Ahmad Sharif

Bot: How old are you?

Me: 32

Bot: So you were born in 1990. What are the advantages to being 32 years old?

Me : I born in a time when Internet is not common like today

Bot: How often do you go back to see what it looked like?

Me:: 2/3 times in a month\*\*\*

Bot: Sorry, I don't really understand what you are asking me to calculate.

\*\*\*\*\* This response helped me to detect it is a bot

01.3

Exercise 1

$$\prod_{i=1}^n \frac{p(w_i / \text{Left})}{p(w_i / \text{Right})}$$

$$= \frac{p(w_1 \cap \text{Left}) / p(w_1)}{p(w_1 \cap \text{Right}) / p(w_1)}$$

$$= \frac{p(w_1 \cap \text{Left})}{p(w_1 \cap \text{Right})}$$

$$= \frac{p(w_1)}{p(w_1)} = 1.$$

```
In [11]: # Exercise 1.4: Python basics, part 1.  
# Ahmad Sharif  
# Student ID : K 436765
```

```
In [7]: from scipy.stats import multivariate_normal  
import numpy as np
```

```
Out[7]: 0.0013717986904768747
```

```
In [8]: multivariate_normal.pdf(x = [2,2,2], mean=[1,3,5], cov=[[4,2,1],  
[2,5,2],  
[1,2,3]])
```

```
Out[8]: 0.0013717986904768747
```

```
In [9]: multivariate_normal.pdf(x = [1,4,3], mean=[1,3,5], cov=[[4,2,1],  
[2,5,2],  
[1,2,3]])
```

```
Out[9]: 0.0026090337298477584
```

```
In [10]: multivariate_normal.pdf(x = [1,1,5], mean=[1,3,5], cov=[[4,2,1],  
[2,5,2],  
[1,2,3]])
```

```
Out[10]: 0.005724150877228364
```