# Text Visualization

Creating Bag Of Words Model

# Text Visualization

"It is going to rain today"

"Today I am not going outside"

"I am going to watch the season premiere"

# Bag Of Words – Lower the sentences

"it is going to rain today"

"today i am not going outside"

"i am going to watch the season premiere"

# Bag Of Words – Tokenization

**Sentence 1**

it

is

going

to

rain

today

**Sentence 2**

today

i

am

not

going

outside

**Sentence 3**

i

am

going

to

watch

the

season

premiere

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| | |
| | |
| | |
| | |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
|  |  |
|  |  |
|  |  |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 1 |
| | |
| | |

**Sentence 1**

it

is

going

to

rain

today

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 1 |
| | |
| | |

**Sentence 2**

today

i

am

not

going

outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 2 |
| | |
| | |
| | |
| | |

**Sentence 2**

today

i

am

not

going

outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 2 |
| i | 1 |
| | |
| | |
| | |

**Sentence 2**

today

i

am

not

going

outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 2 |
| i | 1 |
| am | 1 |
| | |
| | |

**Sentence 2**
today
i
am
not
going
outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 1 |
| to | 1 |
| rain | 1 |
| today | 2 |
| i | 1 |
| am | 1 |
| not | 1 |
| | |

**Sentence 2**
today
i
am
not
going
outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 2 |
| to | 1 |
| rain | 1 |
| today | 2 |
| i | 1 |
| am | 1 |
| not | 1 |
|  |  |

**Sentence 2**
today
i
am
not
going
outside

# Bag Of Words – Histogram

| Word | Count |
|------|-------|
| it | 1 |
| is | 1 |
| going | 2 |
| to | 1 |
| rain | 1 |
| today | 2 |
| i | 1 |
| am | 1 |
| not | 1 |
| outside | 1 |

**Sentence 2**
today
i
am
not
going
outside

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | | |
| is | 1 | | |
| going | 2 | | |
| to | 1 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 1 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**

i

am

going

to

watch

the

season

premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | | |
| is | 1 | | |
| going | 2 | | |
| to | 1 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**

i

am

going

to

watch

the

season

premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | | |
| is | 1 | | |
| going | 3 | | |
| to | 1 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**
i
am
going
to
watch
the
season
premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|---|---|---|---|
| it | 1 | | |
| is | 1 | | |
| going | 3 | | |
| to | 2 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**

i

am

going

to

watch

the

season

premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|---|---|---|---|
| it | 1 | watch | 1 |
| is | 1 | | |
| going | 3 | | |
| to | 2 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**
i
am
going
to
watch
the
season
premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | | |
| to | 2 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**
i
am
going
to
watch
the
season
premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | season | 1 |
| to | 2 | | |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**

i

am

going

to

watch

the

season

premiere

# Bag Of Words – Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | season | 1 |
| to | 2 | premiere | 1 |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

**Sentence 3**

i
am
going
to
watch
the
season
premiere

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | season | 1 |
| to | 2 | premiere | 1 |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |
|      |       |      |       |

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | season | 1 |
| to | 2 | premiere | 1 |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| going | 3 | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

Maximum: going

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|---|---|---|---|
| it | 1 | watch | 1 |
| is | 1 | the | 1 |
| going | 3 | season | 1 |
| to | 2 | premiere | 1 |
| rain | 1 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| not | 1 | | |
| outside | 1 | | |

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| going | 3 | | |
| to | 2 | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

Maximum:  to

© Bijoyan Das

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| going | 3 | | |
| to | 2 | | |
| today | 2 | | |
| | | | |

Maximum: today

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| going | 3 | | |
| to | 2 | | |
| today | 2 | | |
| i | 2 | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

Maximum: i

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|---|---|---|---|
| going | 3 | | |
| to | 2 | | |
| today | 2 | | |
| i | 2 | | |
| am | 2 | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

Maximum:  am

# Bag Of Words – Sort the Histogram

| Word | Count | Word | Count |
|------|-------|------|-------|
| going | 3 | watch | 1 |
| to | 2 | the | 1 |
| today | 2 | season | 1 |
| i | 2 | premiere | 1 |
| am | 2 | | |
| it | 1 | | |
| is | 1 | | |
| rain | 1 | | |
| not | 1 | | |
| outside | 1 | | |

# Bag Of Words – Filter words

| Word | Count | Word | Count |
|---|---|---|---|
| going | 3 | watch | 1 |
| to | 2 | the | 1 |
| today | 2 | season | 1 |
| i | 2 | premiere | 1 |
| am | 2 | | |
| it | 1 | | |
| is | 1 | | |
| rain | 1 | | |
| not | 1 | | |
| outside | 1 | | |

10 most frequent words

# Bag Of Words – Filter words

| Word | Count |
|---|---|
| going | 3 |
| to | 2 |
| today | 2 |
| i | 2 |
| am | 2 |
| it | 1 |
| is | 1 |
| rain | 1 |
| not | 1 |
| outside | 1 |

# Bag Of Words – Creating the matrix

| Word | Count |
|------|-------|
| going | 3 |
| to | 2 |
| today | 2 |
| i | 2 |
| am | 2 |
| it | 1 |
| is | 1 |
| rain | 1 |
| not | 1 |
| outside | 1 |

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

© Bijoyan Das

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |

it is going to rain today

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | | | | | | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | | | | | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | | | | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | | | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 3 | | | | | | | | | | |

today i am not going outside

# Bag Of Words – Filter words

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 3 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |

i am going to watch the season premiere

# Bag Of Words – Model Ready

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 3 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |

# Text Visualization

"It is going to rain today"

"Today I am not going outside"

"I am going to watch the season premiere"

# Bag Of Words – Model Ready

| Words/Documents | going | to | today | i | am | it | is | rain | not | outside |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 3 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |