Text Summarization

A NLP based approach

Text Summarization -Techniques

- A simple Natural Language Processing based approach
- A Deep NLP based approach

Text Summarization -Techniques

- A simple Natural Language Processing based approach ⬅
- A Deep NLP based approach

Text Summarization – Sample paragraph

Thank you all so very much. Thank you to the Academy. Thank you to all of you in this room. I have to congratulate the other incredible nominees this year. *The Revenant* was the product of the tireless efforts of an unbelievable cast and crew.

Text Summarization – Tokenization

1. Thank you all so very much.
2. Thank you to the Academy.
3. Thank you to all of you in this room.
4. I have to congratulate the other incredible nominees this year.
5. *The Revenant* was the product of the tireless efforts of an unbelievable cast and crew.

Text Summarization – Preprocess

1. thank you all so very much
2. thank you to the academy
3. thank you to all of you in this room
4. i have to congratulate the other incredible nominees this year
5. *the revenant* was the product of the tireless efforts of an unbelievable cast and crew

Text Summarization – Histogram

Word	Count	Word	Count	Word	Count
thank	3	in	1	revenant	1
you	4	this	2	was	1
all	2	room	1	product	1
so	1	i	1	tireless	1
very	1	have	1	efforts	1
much	1	congratulate	1	an	1
to	3	other	1	unbelievable	1
the	5	incredible	1	cast	1
academy	1	nominees	1	and	1
of	3	year	1	crew	3

Text Summarization – Weighted Histogram

Maximum is 5

Word	Weight	Word	Weight	Word	Weight
thank	3/5	in	1/5	revenant	1/5
you	4/5	this	2/5	was	1/5
all	2/5	room	1/5	product	1/5
so	1/5	i	1/5	tireless	1/5
very	1/5	have	1/5	efforts	1/5
much	1/5	congratulate	1/5	an	1/5
to	3/5	other	1/5	unbelievable	1/5
the	5/5	incredible	1/5	cast	1/5
academy	1/5	nominees	1/5	and	1/5
of	3/5	year	1/5	crew	1/5

Text Summarization – Weighted Histogram

Word	Weight	Word	Weight	Word	Weight
thank	0.5	in	0.2	revenant	0.2
you	0.8	this	0.4	was	0.2
all	0.4	room	0.2	product	0.2
so	0.2	i	0.2	tireless	0.2
very	0.2	have	0.2	efforts	0.2
much	0.2	congratulate	0.2	an	0.2
to	0.6	other	0.2	unbelievable	0.2
the	1	incredible	0.2	cast	0.2
academy	0.2	nominees	0.2	and	0.2
of	0.6	year	0.2	crew	0.2

Text Summarization – Sentence Scores

1. thank you all so very much
2. thank you to the academy
3. thank you to all of you in this room
4. i have to congratulate the other incredible nominees this year
5. *the revenant* was the product of the tireless efforts of an unbelievable cast and crew

Text Summarization – Sentence Scores

thank

you

all

so

very

much

Text Summarization – Weighted Histogram

Word	Weight	Word	Weight	Word	Weight
thank	0.5	in	0.2	revenant	0.2
you	0.8	this	0.4	was	0.2
all	0.4	room	0.2	product	0.2
so	0.2	i	0.2	tireless	0.2
very	0.2	have	0.2	efforts	0.2
much	0.2	congratulate	0.2	an	0.2
to	0.6	other	0.2	unbelievable	0.2
the	1	incredible	0.2	cast	0.2
academy	0.2	nominees	0.2	and	0.2
of	0.6	year	0.2	crew	0.2

Text Summarization – Sentence Scores

thank	0.5
you	0.8
all	0.4
so	0.2
very	0.2
much	0.2

Text Summarization – Sentence Scores

thank	0.5
you	0.8
all	0.4
so	0.2
very	0.2
much	0.2
	2.3

Text Summarization – Sentence Scores

Sentence	Score
thank you all so very much	2.3

Text Summarization – Sentence Scores

Sentence	Score
thank you all so very much	2.3
thank you to the academy	3.1

Text Summarization – Sentence Scores

Sentence	Score
thank you all so very much	2.3
thank you to the academy	3.1
thank you to all of you in this room	4.3

Text Summarization – Sentence Scores

Sentence	Score
thank you all so very much	2.3
thank you to the academy	3.1
thank you to all of you in this room	4.3
i have to congratulate the other incredible nominees this year	3.4

Text Summarization – Sentence Scores

Sentence	Score
thank you all so very much	2.3
thank you to the academy	3.1
thank you to all of you in this room	4.3
i have to congratulate the other incredible nominees this year	3.4
the revenant was the product of the tireless efforts of an unbelievable cast and crew	6.2

Text Summarization – Sort by Score

Sentence	Score
the revenant was the product of the tireless efforts of an unbelievable cast and crew	6.2
thank you to all of you in this room	4.3
i have to congratulate the other incredible nominees this year	3.4
thank you to the academy	3.1
thank you all so very much	2.3

© Bijoyan Das

Text Summarization – Pickle N Largest

Sentence	Score
the revenant was the product of the tireless efforts of an unbelievable cast and crew	6.2
thank you to all of you in this room	4.3

Text Summarization – Summary done

The Revenant was the product of the tireless efforts of an unbelievable cast and crew.

Thank you to all of you in this room.