



SUMMARIZATION AND SENTIMENT ANALYSIS OF REVIEWS

CSC 522: Automated Learning & Data Analysis

P26

Aayush Indrapratap Singh

Diksha Paliwal

Bhavya Omprakash Agrawal

INTRODUCTION

DATA

SUMMARIZATION


SENTIMENT ANALYSIS

METHOD & OUTPUTS

CONCLUSION

INTRODUCTION

Electronics › Portable Audio & Video › Portable Speakers & Docks › Portable Bluetooth Speakers









Bose SoundLink Color II: Portable Bluetooth, Wireless Speaker with Microphone- Soft Black
[Visit the Bose Store](#)
★★★★★ 51,048 ratings | 1000+ answered questions
Amazon's Choice for "bluetooth speaker" above \$100

Price: **\$129.00**

Style: **Color Series**

Color: **Soft Black**

| | | |
|--|--|---|
|  \$115.93 |  \$129.00 |  \$189.91 |
|  \$179.82 |  \$129.00 |  \$129.00 |

Connectivity Technology Wireless, Bluetooth, NFC


Speaker Type Outdoor

Brand Bose

Model Name SoundLink Color II

Mounting Type Tabletop










Electronics › Portable Audio & Video › Portable Speakers & Docks › Portable Bluetooth Speakers



JBL FLIP 5, Waterproof Portable Bluetooth Speaker, Blue (New Model)
[Visit the JBL Store](#)
★★★★★ 41,361 ratings | 646 answered questions

Price: **\$129.95**

Color: **Blue**

| | | | | | | |
|---|---|---|---|---|---|---|
|  |  |  |  |  |  | |
|  |  |  | | | | |

Connectivity Technology Bluetooth

Speaker Type Portable Bluetooth Speakers

Brand JBL

Model Name JBL FLIP 5 - Blue

Recommended Uses For Product For Smartphones or Tablets

Faced such a situation before?

How much time have we lost in making such choices?

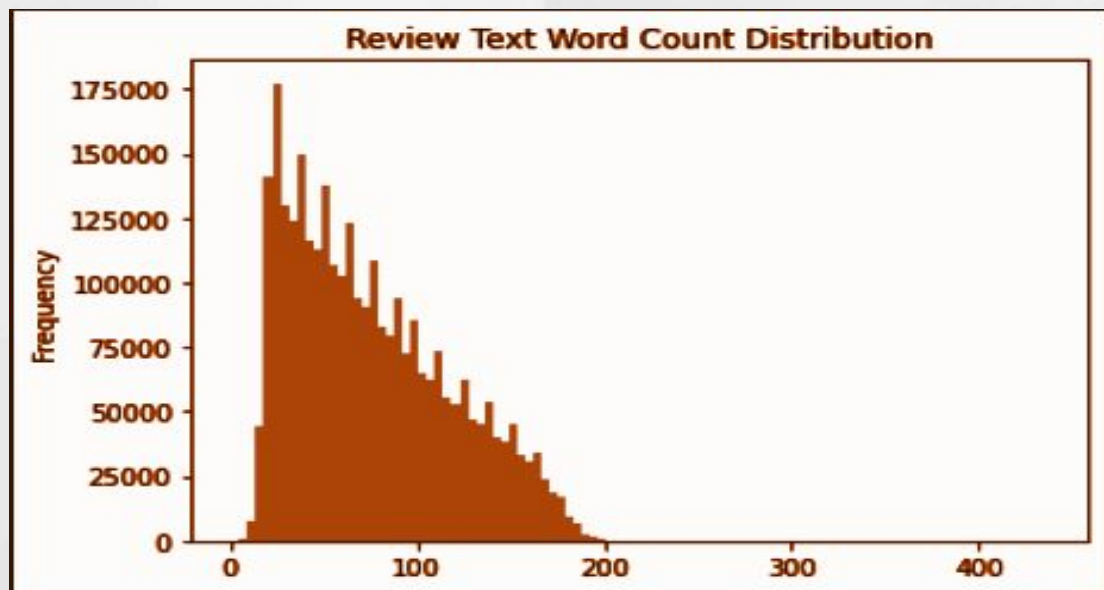
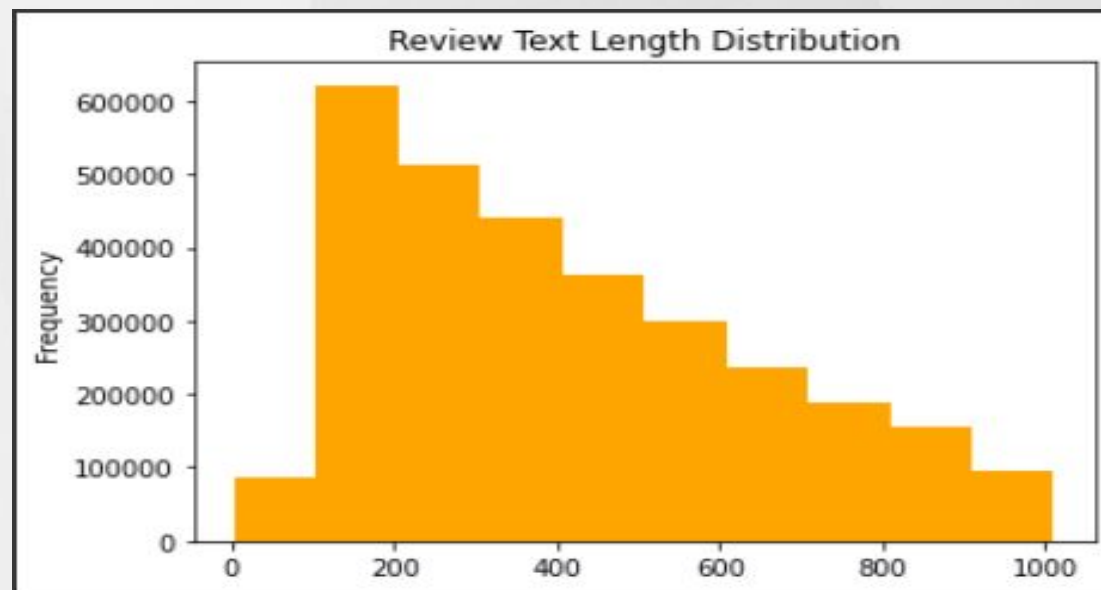
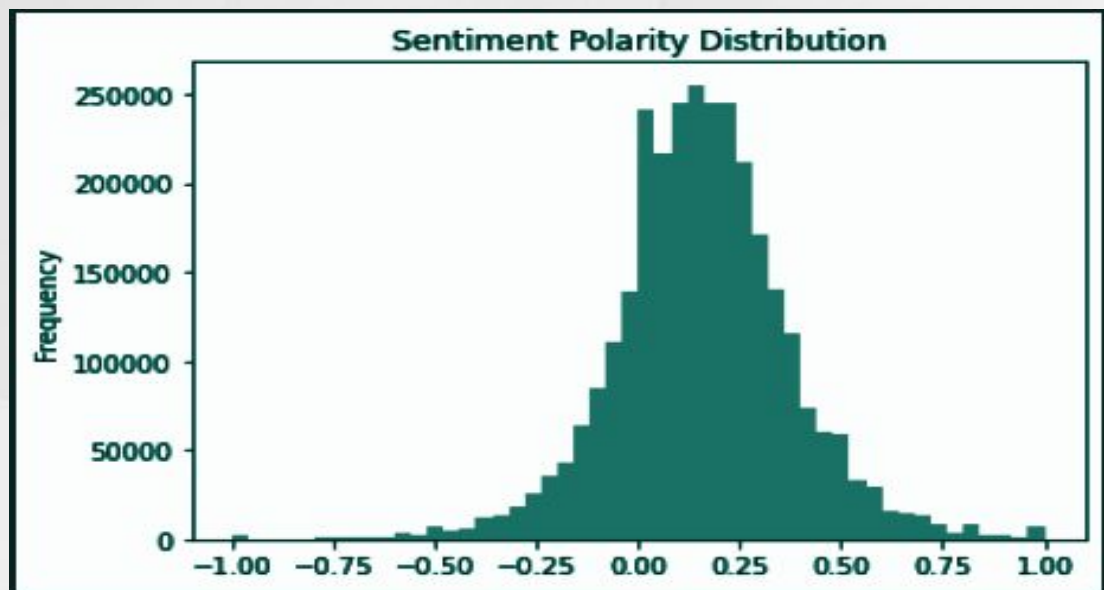
How much time have we spent in the Reviews Rabbithole?

INTRODUCTION

- Due to increased footfall on online sites, people seek for feedback from prior customers.
- The review summary can be useful for sellers, to enhance their services and products.
- Due of the length of reviews, buyers tend to skip long reviews and may miss about all the aspects of the review.
- Aim to provide users with concise and accurate product reviews.
- People will be able to get the information more quickly if they were provided with the summarized reviews.
- The project creates a summarization model using a dataset of Amazon reviews.
- Dataset: Amazon Reviews ([Link](#))

DATA

| RATING | TITLE | REVIEW |
|--------|---------------------------------------|---|
| 5 | Inspiring | I hope a lot of people hear this cd. We need more strong and positive vibes like this. Great vocals, fresh tunes, cross-cultural happiness. Her blues is from the gut. The pop sounds are catchy and mature. |
| 5 | The best soundtrack ever to anything. | I'm reading a lot of reviews saying that this is the best 'game soundtrack' and I figured that I'd write a review to disagree a bit. This in my opinino is Yasunori Mitsuda's ultimate masterpiece. The music is timeless and I'm been listening to it for years now and its beauty simply refuses to fade. The price tag on this is pretty staggering I must say, but if you are going to buy any cd for this much money, this is the only one that I feel would be worth every penny. |
| 5 | There's a reason for the price | There's a reason this CD is so expensive, even the version that's not an import. Some of the best music ever. I could listen to every track every minute of every day. That's about all i can say. |



PRE-PROCESSING

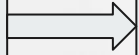
CONTRACTION

Don't



Do not

Can't



Can not

UNWANTED CHARACTERS

[_"\-;%()|+&=\$@\[\]/]

PUNCTUATIONS

STOPWORDS

The, And, An, Of, Is,
This, Their, It, In,
For, Was, But, Not,
What, Where,
Themselves

WORD EMBEDDINGS

GLOVE for Vector
Representation of
Words

TAGGING TOKENIZATION

Start and End Tokens

Splitting Sentences to
Words

TEXT TO SEQUENCE

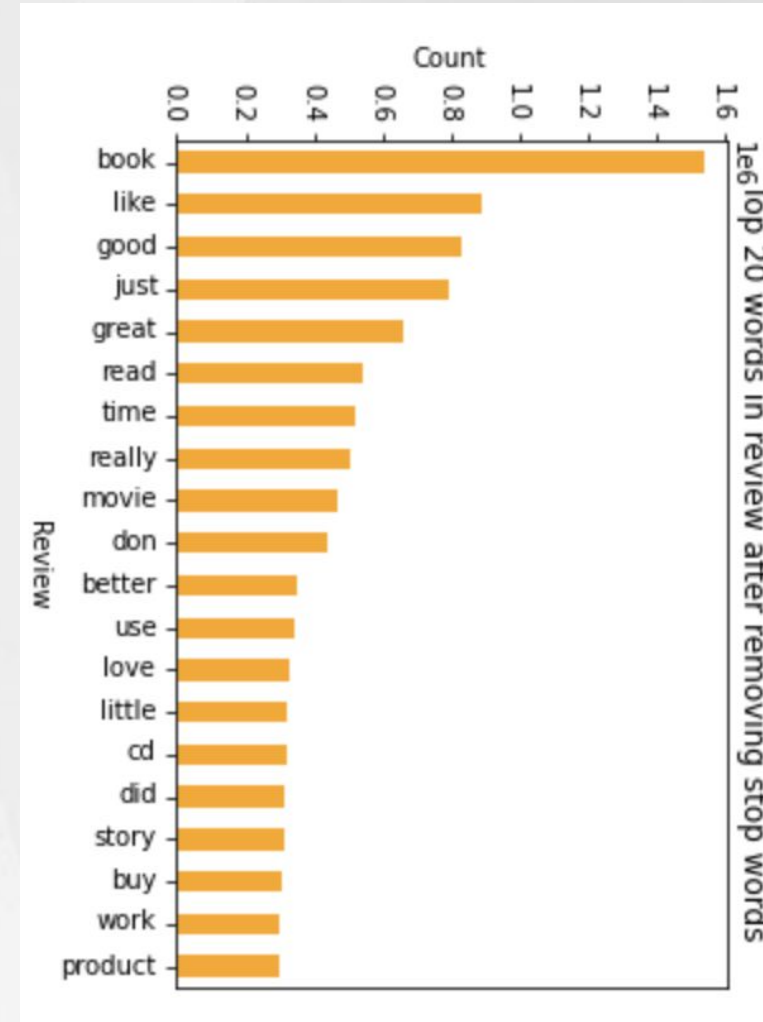
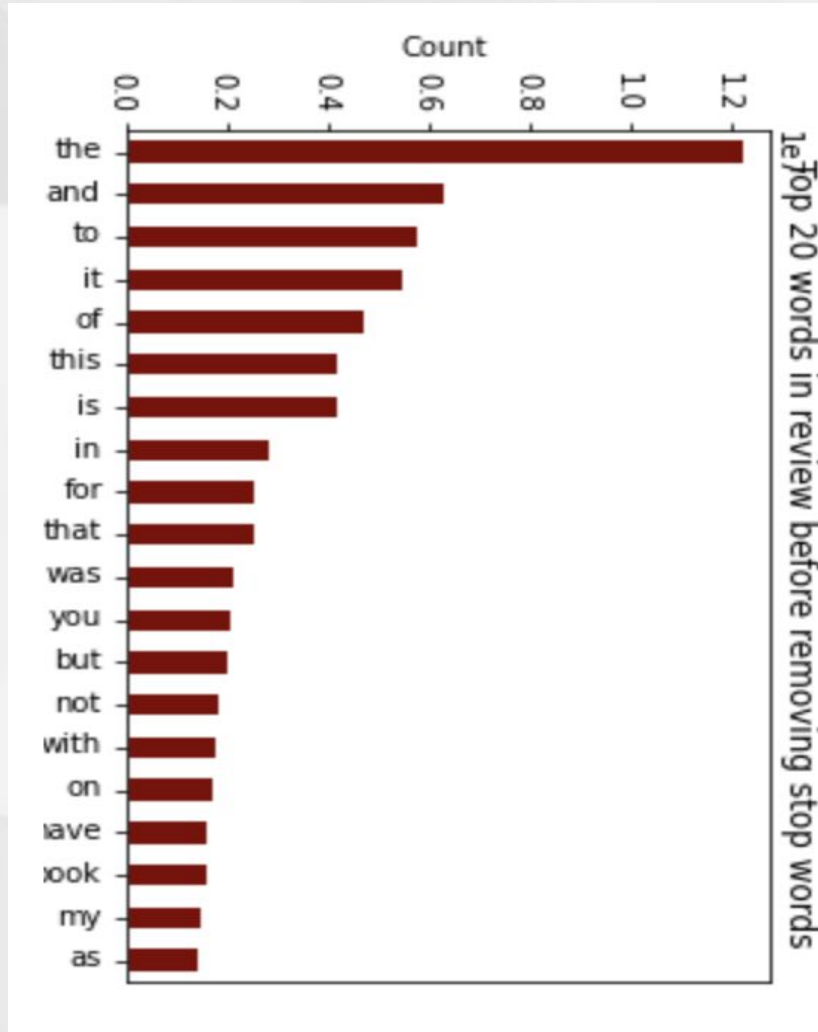
Mathematical Model

Input words as series
of numbers

PADDING

Equal Length Inputs

HOW STOPWORDS AFFECT WORD COUNT



SUMMARIZATION



SUMMARIZATION

EXTRACTIVE
SUMMARIZATION

- ★ Extract sentences from the Text

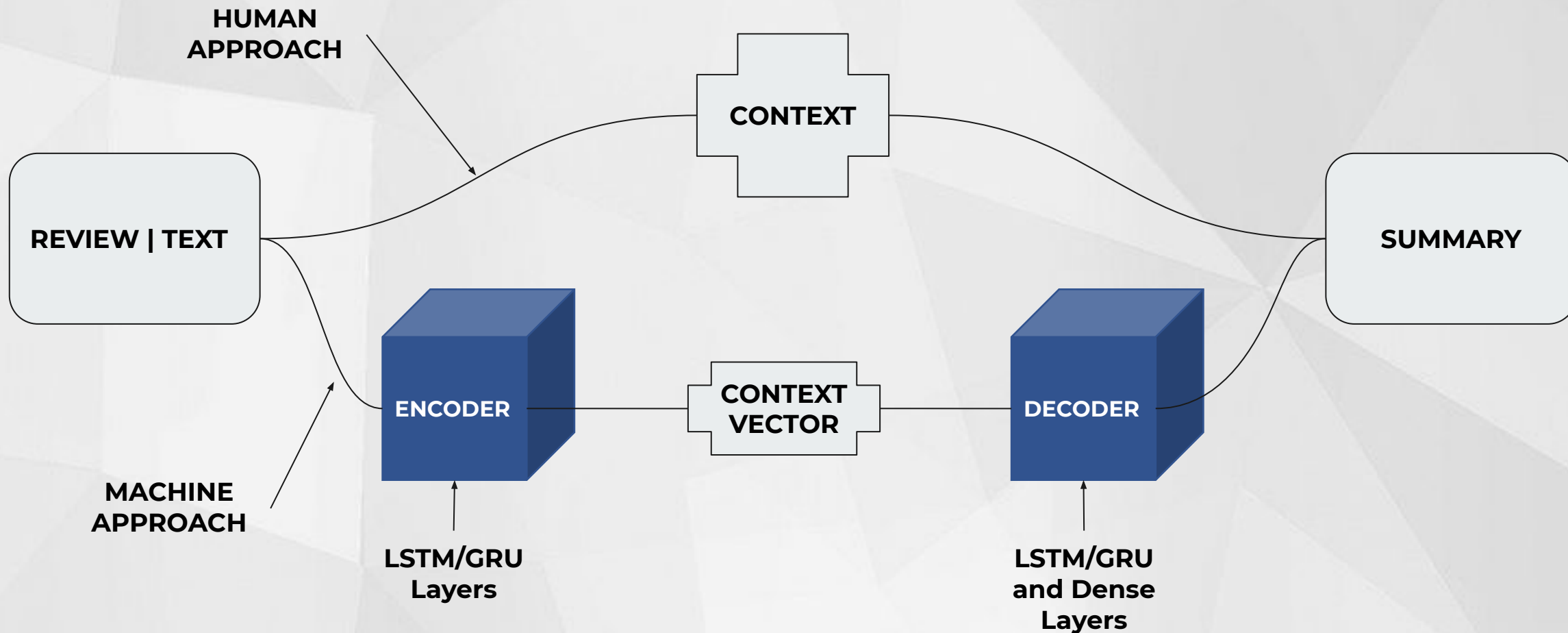
Selecting sentences and phrases from the text based on frequency.

- ★ Extract meaning from the Text

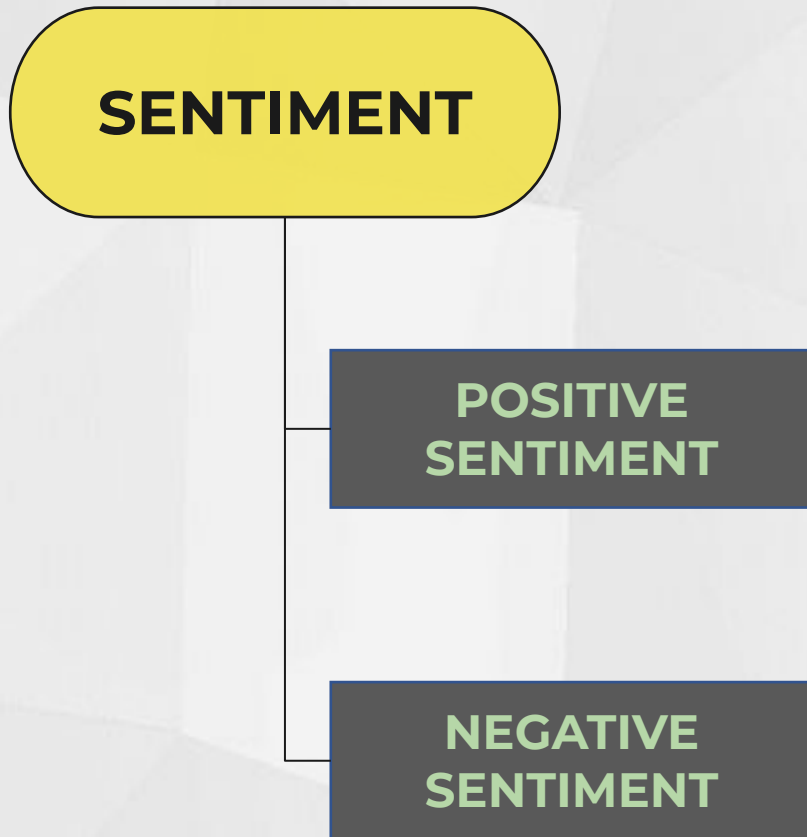
ABSTRACTIVE
SUMMARIZATION

Generating new sentences or phrases to capture the meaning of the text.

METHOD SUMMARISATION



SENTIMENT ANALYSIS



★ Positive Words/Phrases → Positive Sentiment

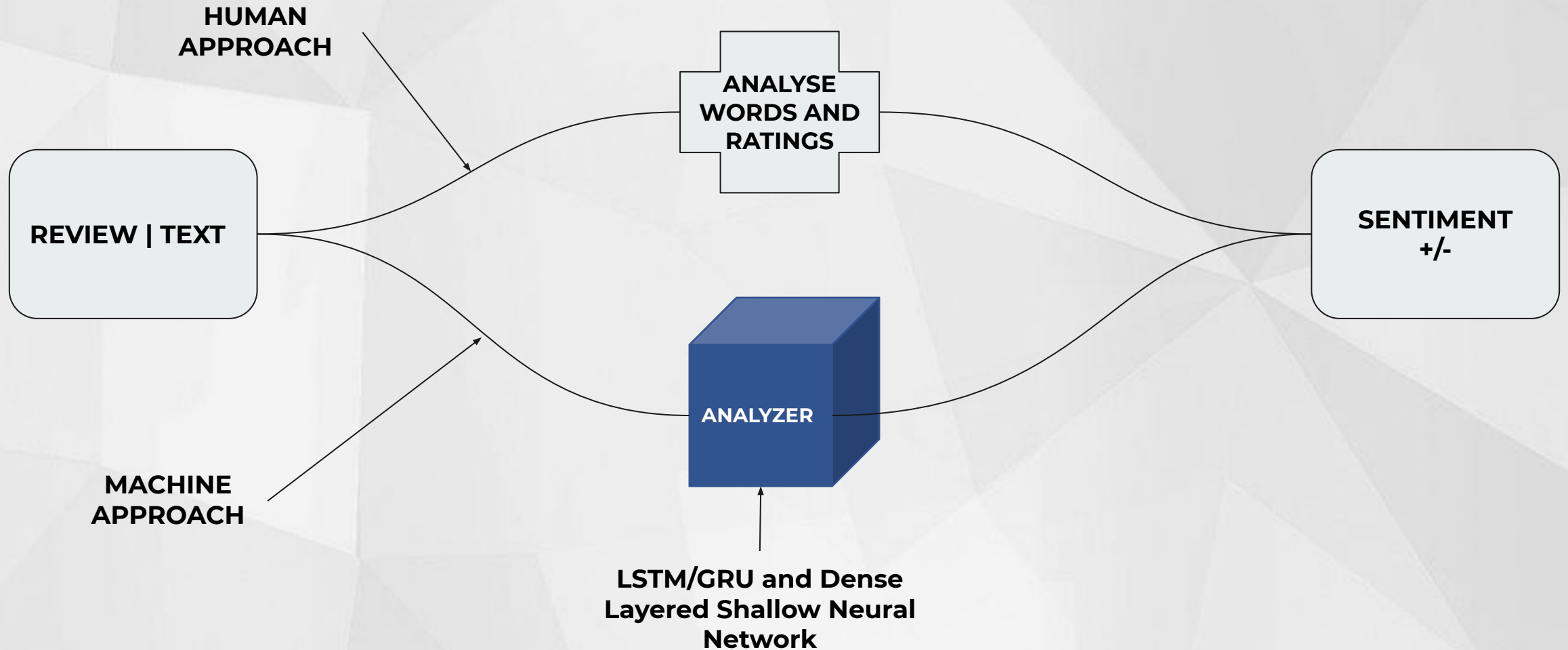
.... the **best** Bluetooth speaker you can buy for \$150.
It's **good** looking, rugged and most importantly sounds **GREAT!!!**
The Mic works really **well**, as I take daily meeting on it without an issue

★ Negative Words/Phrases → Negative Sentiment

the "user manual" has **no instructions** for setting up bluetooth.
Second **problem**, and **deal killer**, it sounds **awful**, like it uses a toilet paper tube for a speaker: **scratchy, rumbly, distorted**. This is true at any volume level. Third **problem**, the volume adjustment **doesn't work** very well. Ten minutes was enough; back to Amazon it goes!

METHOD

SENTIMENT ANALYSIS



OUTPUT

Original Review: These tracks are not the "original" versions but are re-recorded versions. So, whether the tracks are "remastered" or not is irrelevant.

Original summary: caution

Predicted summary: not as good as the original version

Sentiment: Negative

Original Review: In a world where media is dominant this movie is so refreshing as it chronicles the agony of a person who yearns for God. Watch it! 0:)

Original summary: divinely wonderful

Predicted summary: a must have

Sentiment: Positive

Original Review: While I believe in discrimination, I certainly can't believe what is in this book. It is an unsubstantiated smear campaign to rally the troupes in higher ed to claw again for the preservation of academic freedom- a rally which falls flat.

Original summary: yet another unsubstantiated case

Predicted summary: a biased book

Sentiment: Negative

EVALUATION METRICS

SUMMARIZATION

BLEU : Bilingual Evaluation Understudy

-Evaluates the quality of text translated by a machine from one natural language to another. Used in tasks such as text generation, paraphrase generation, and text summarization.

Expected BLEU score value for a good Summarization model:

50-60 : implies Very high quality, adequate, and fluent translations

(Source* - Google's AutoML documentation)

Our BLEU score:

54.27

SENTIMENT ANALYSIS

Accuracy

- Provides the proportion of total number of reviews with correct sentiment detected.
- Provides the performance of the model.

Expected Accuracy for a sentiment Analysis model:

- Higher accuracy shows that the model extracts better and accurate sentiment of the reviews.

Our Accuracy:

80.7%

CONCLUSION

- ❖ The Extractive model might generate the summary with higher accuracy as compared to the Abstractive model, but the abstractive model captures the context of the text and generates new sentences and phrases which efficiently recreates the human-like approach to summarisation.
- ❖ For our application, Abstractive Summarization is a better choice, as the motive is to extract the context of the large reviews in a few words in order to save time and efforts.

FUTURE SCOPE

- ❖ Use of grouped data for specific products to generate the summary for each kind of product.
- ❖ Increasing the hardware resources and fine tuning the model to increase the performance of the model and provide higher accuracy.
- ❖ As data is user generated, it has typos and errors, the generated summaries could be improved if these issues can be fixed in the data.



THANKYOU!