

SENTIMENT ANALYSIS



OUTLINE

Motivation

What? Why??



Method

Data, Preprocess, Model

Outcome

Result and challenges

Q&A

• What is Sentiment Analysis? •



Negative



terrible acting and mediocre story

Positive



What is Sentiment Analysis?

FINE-GRAINED

Negative

Positive



THREE-CLASS



BINARY



• Why is it useful? •

Collect Market Response



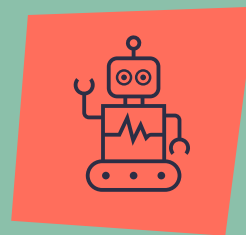
customer review,
social media comment

Decision-making



e.g. news/forums
for stock prediction

Basis for NLP Tasks





emotion analysis,
intent classification



But... even a three-year-old can do it?





Bae can I hang out with my
friends at the bar tonight?

No problem, you can go as
you like 😊

the guy will be in trouble if he fails to
capture the underlying sentiment

Cantonese version that better capture the meaning...

BB我今晚可摸耳同frd去bar呀

無問題啊，你鍾意啦 😊



• Tricky parts of sentiment analysis •

Double Negation

I do not dislike cabin cruisers.

Context Dependent

It is a fantastic _____ with two-hour battery!

Sarcasm

I'd really truly love going out in this weather!

OUTLINE

Motivation

What? Why??



Method

Data, Preprocess, Model

Outcome

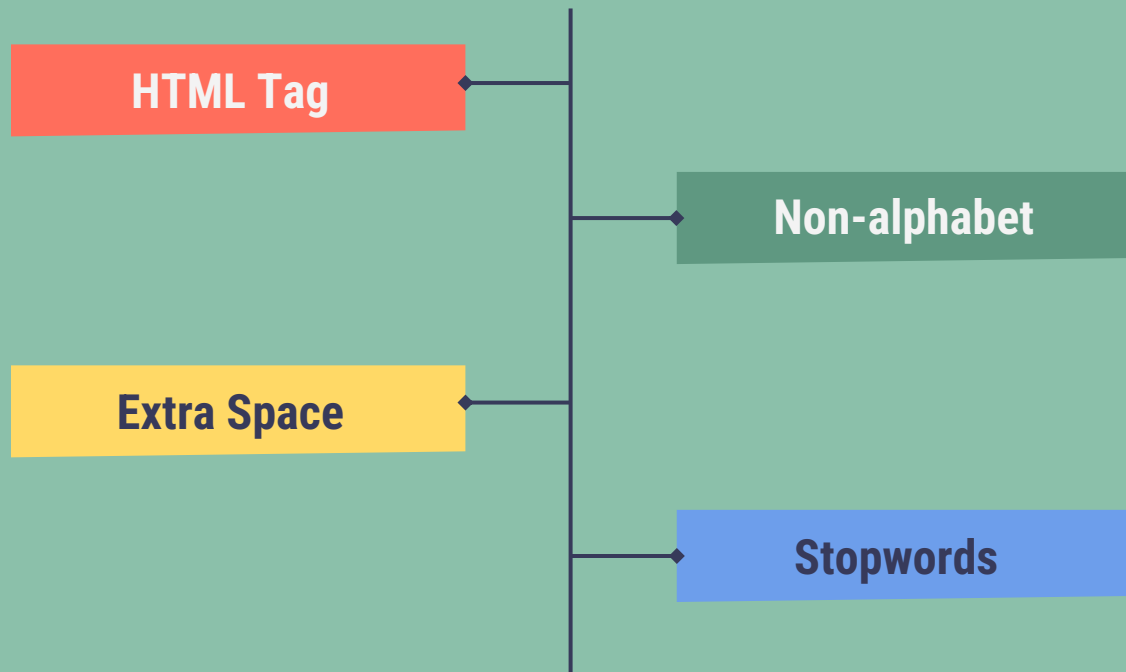
Result and challenges

Q&A

Datasets

	# class	Description
IMDB	2	movie review dataset, 50000 data
Financial News	3	financial news dataset from Kaggle, 5800 data
Stanford Sentiment Treebank	SST-2 : 2 SST-5 : 5	review dataset, ~10000 data

• Preprocessing •

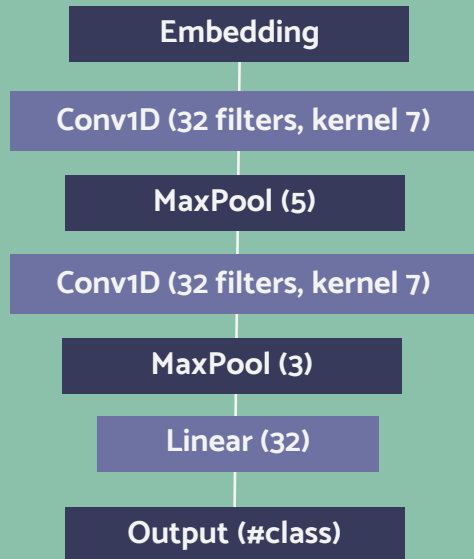


Models

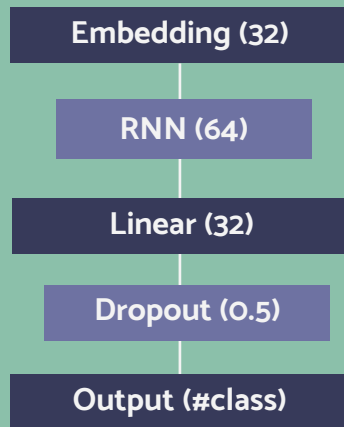


MODELS

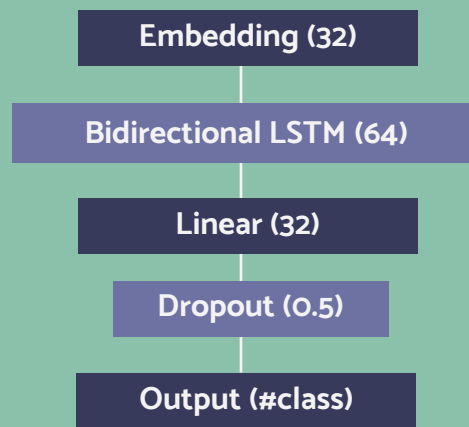
CNN



RNN



LSTM



Models

BERT

BERT-Base:
L-12_H-768_A-12

S-BERT

BERT-Tiny:
L-2_H-128_A-2

BERT Preprocess

pre-trained (S)BERT Layer

Linear (64)

Dropout (0.2)

Output (#class)

• OUTLINE •

Motivation

What? Why??

Method

Data, Preprocess, Model

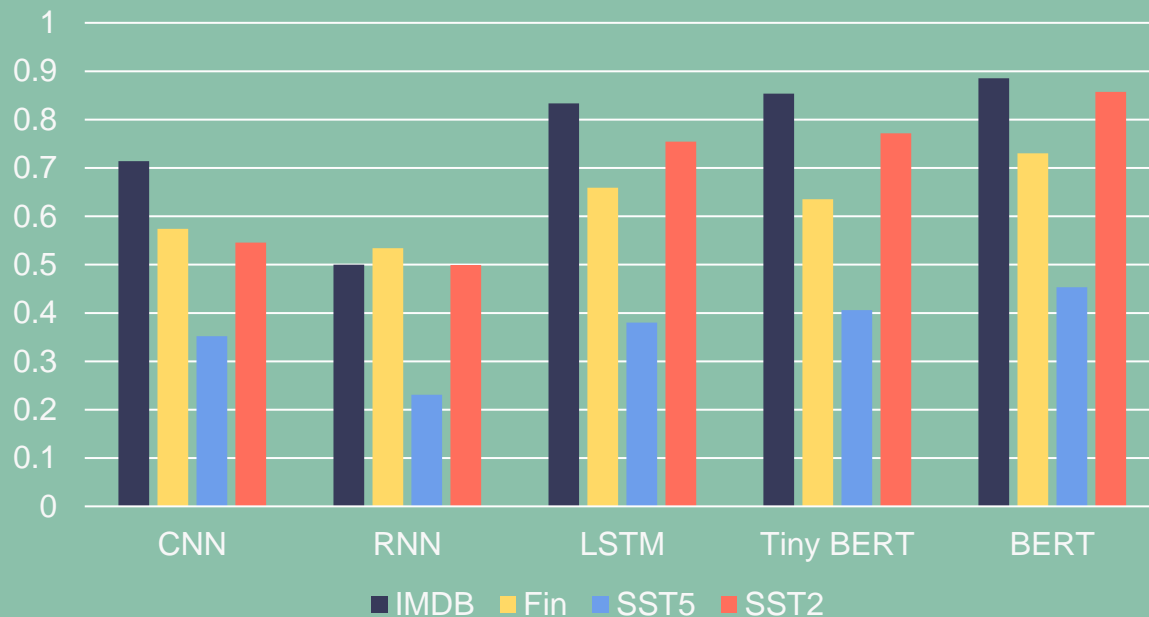
Outcome

Result and challenges

Q&A



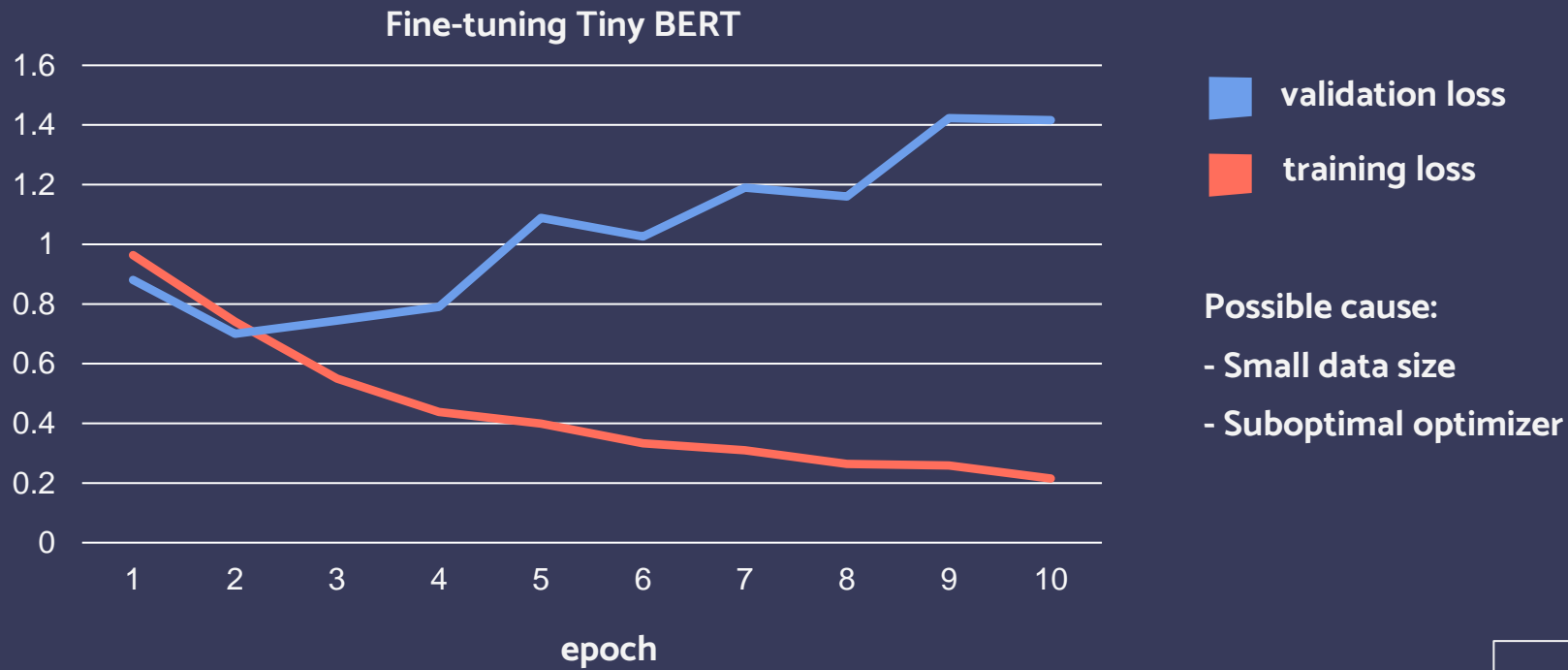
Model Performance



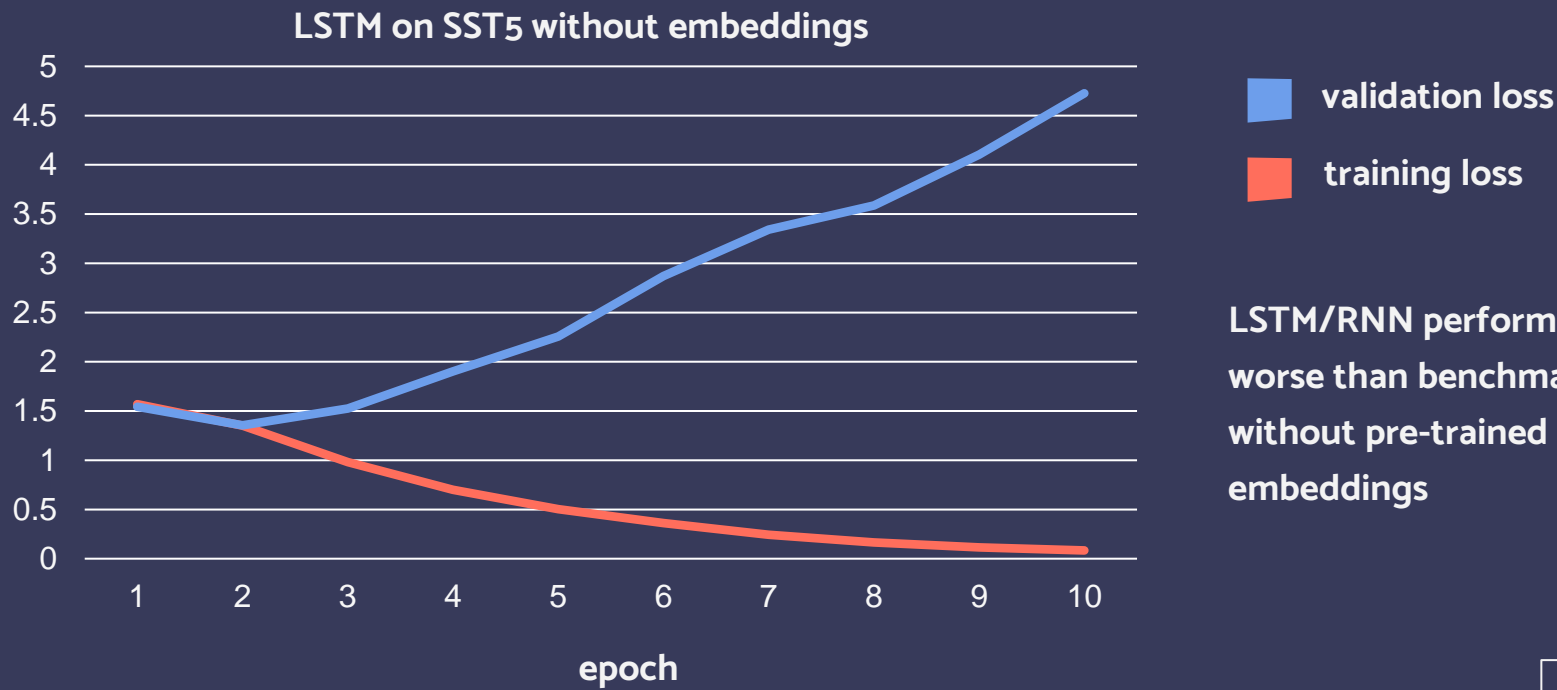
Consistent Rank:
BERT > Tiny BERT > LSTM
> CNN > RNN

Best Performance on
binary datasets

Challenge – Poor Generalization



• Challenge – Lack of Embeddings •



LSTM/RNN performs worse than benchmarks without pre-trained embeddings

References

Slides template from Slidesgo and illustrations from FreePik

Sentiment Analysis Examples: https://en.wikipedia.org/wiki/Sentiment_analysis

Financial Sentiment Analysis Dataset: <https://www.kaggle.com/datasets/sbhatti/financial-sentiment-analysis>

Stanford Sentiment Treebank Dataset: <https://nlp.stanford.edu/sentiment/index.html#download>

IMDB Movie Review Dataset: https://ai.stanford.edu/~amaas/data/sentiment/aclImdb_v1.tar.gz

Q & A

