Sentiment Analysis of the Speaker in Literary Works

Team 10

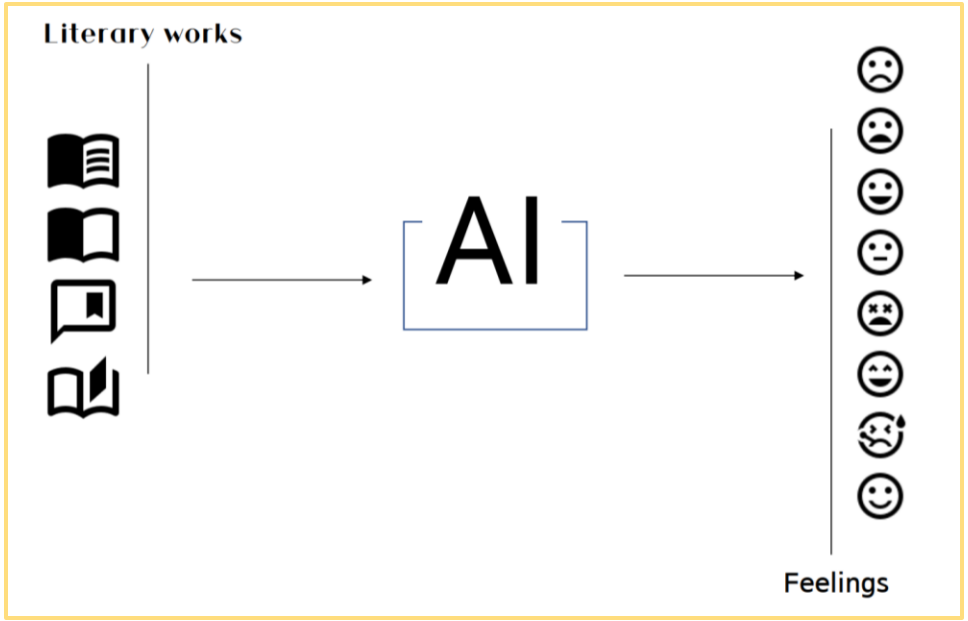
김서정
이광민
한건희
이현준

TABLE OF CONTENTS

Background

Topic

Sentiment classification
of the speaker
in literary works



Classify into various emotions

Pre-existing model

Our model

Social Media Texts

Tweets

Reveal the emotions well

Many researches on
social media emotion analysis

Training

Dataset

Emotion_final.csv (2.26 MB)

Detail | Compact | Column

Text	Emotion
21405 unique values	happy 33% sadness 29% Other (8165) 38%
i didnt feel humiliated	sadness
i can go from feeling so hopeless to so damned hopeful just from being around someone who cares and ...	sadness
im grabbing a minute to post i feel greedy wrong	anger
i am ever feeling nostalgic about the fireplace i will know that it is still on the property	love
i am feeling grouchy	anger
ive been feeling a little burdened lately wasnt sure why that was	sadness
ive been taking or	surprise

Data

21,459

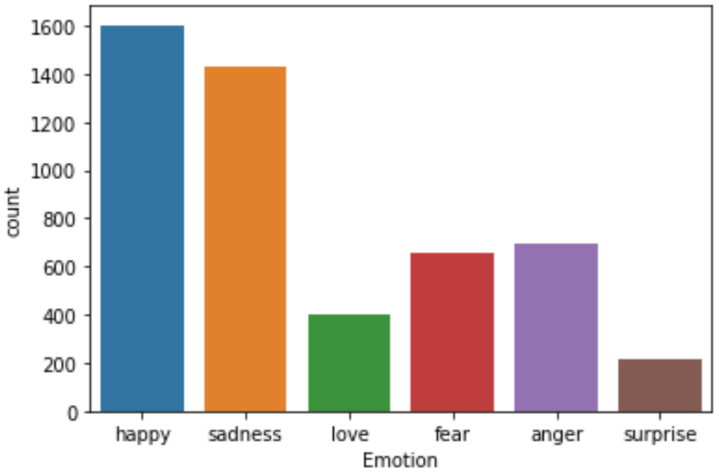
Emotions

Happy, Fear, Anger,

Sadness, Surprise, Love

Split data

- Train: 65%
- Validation: 15%
- Test: 20%



The amount of data each emotion

Baseline model - SVM

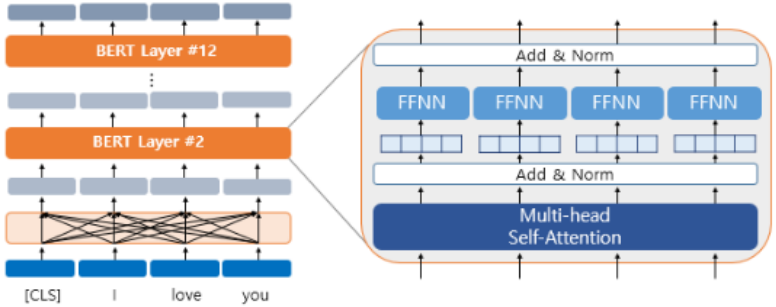
- Analyze data for classification and regression analysis.
- Kernel trick - efficiently perform a non-linear classification

Accuracy: 0.699
F1-Score: 0.672

Result

Accuracy: 0.699

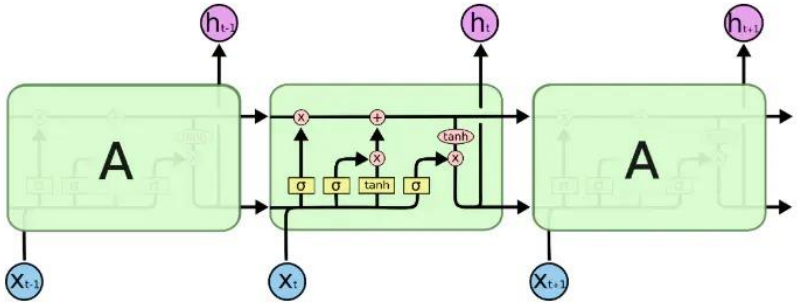
	precision	recall	f1-score	support
0	0.891	0.410	0.562	139
1	0.870	0.511	0.644	131
2	0.605	0.947	0.738	320
3	0.920	0.287	0.438	80
4	0.757	0.847	0.799	287
5	0.500	0.140	0.218	43
accuracy			0.699	1000
macro avg	0.757	0.524	0.567	1000
weighted avg	0.744	0.699	0.672	1000



LSTM

Which model
we should choose?

BERT



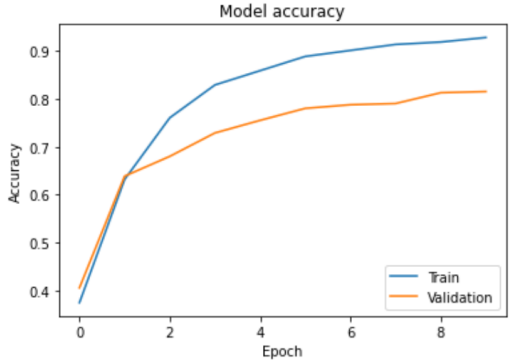
LSTM

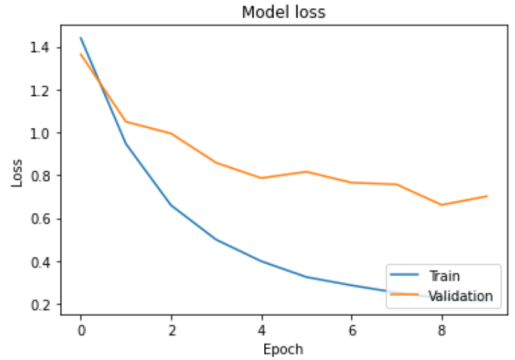
Structure

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, None, 16)	32000
lstm (LSTM)	(None, 64)	20736
dense (Dense)	(None, 128)	8320
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 6)	390

Total params: 69,702
Trainable params: 69,702
Non-trainable params: 0

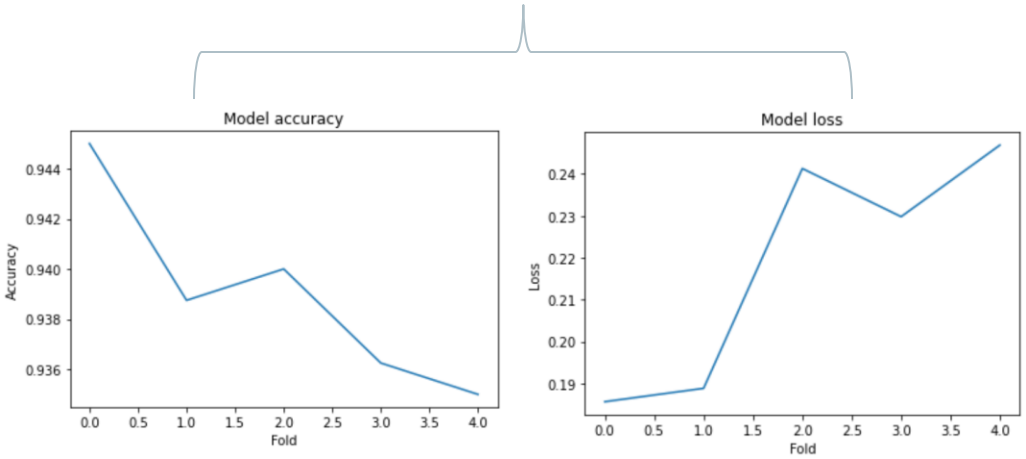
Training Result

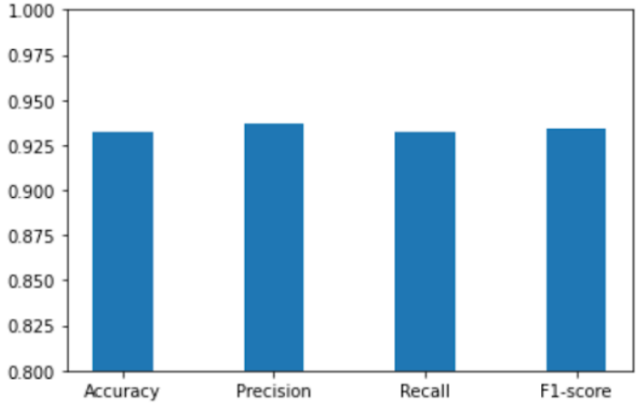




LSTM with K-fold cross validation

The result of each fold

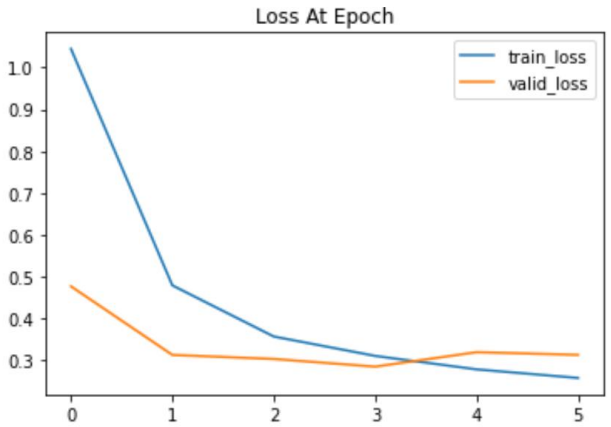




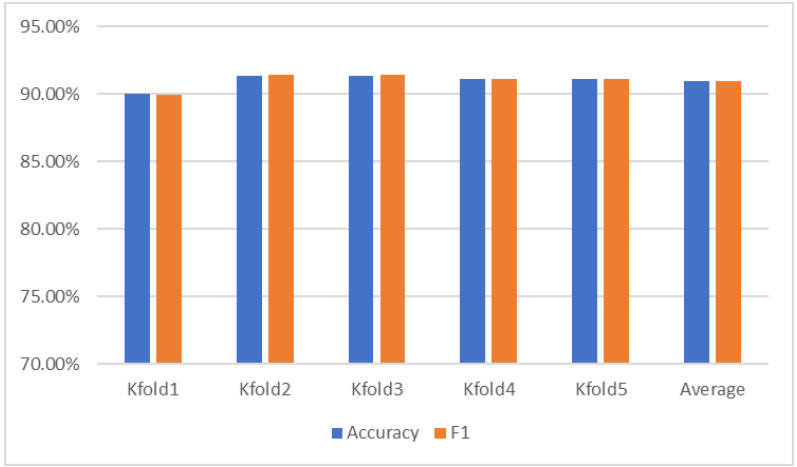
Accuracy	0.9325
Precision	0.9365
Recall	0.9325
F1-score	0.9337

BERT with K-fold cross validation

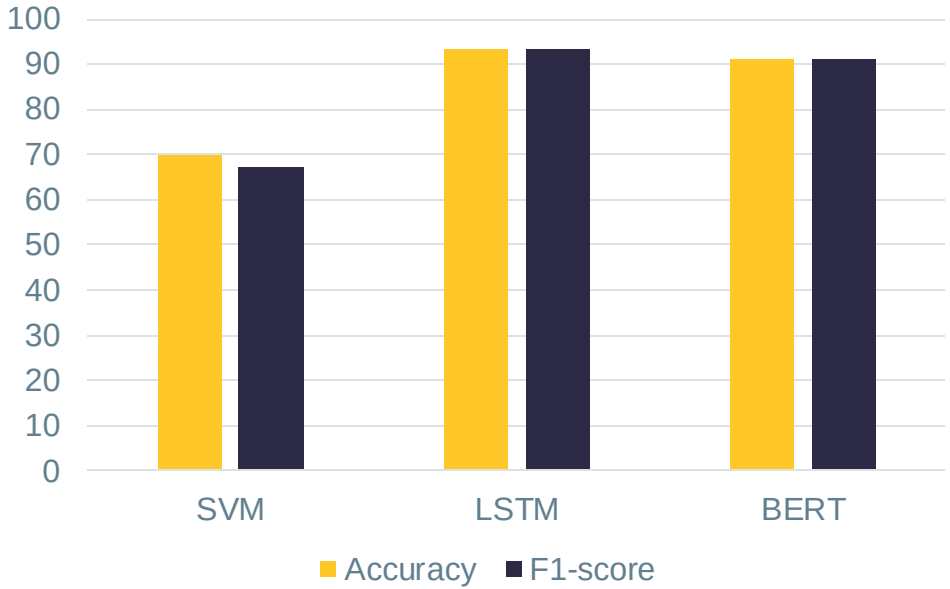
Training Result



The result of each fold



Comparison of LSTM and BERT



Create test dataset

How to increase the association between two types of text?

Create test dataset

- **Type**

  Poem, Novel, Lyric, and YouTube comment

- **Amount of data**

  Almost 30 data each type and each emotion

- **How to collect data**

  Team members crawled the data

- **Criteria of collecting data**

  Not appearing emotional words repeatedly

  Filtering curses, and correcting spelling

Dataset	Test	Poem	Novel	Lyric	YouTube
anger	599	30	30	30	30
fear	530	30	30	30	30
happy	1406	30	30	30	30
love	328	31	31	30	30
sadness	1253	33	33	31	30
surprise	176	11	11	1	20
Total	4292	165	165	152	170

Literary Work Test Data Set

Poem / Love

Shy one, shy one,
Shy one of my heart,
She moves in the
firelight
Pensively apart.

And shy as a rabbit,
Helpful and shy.
To an isle in the water
With her would I fly.

Novel / Surprise

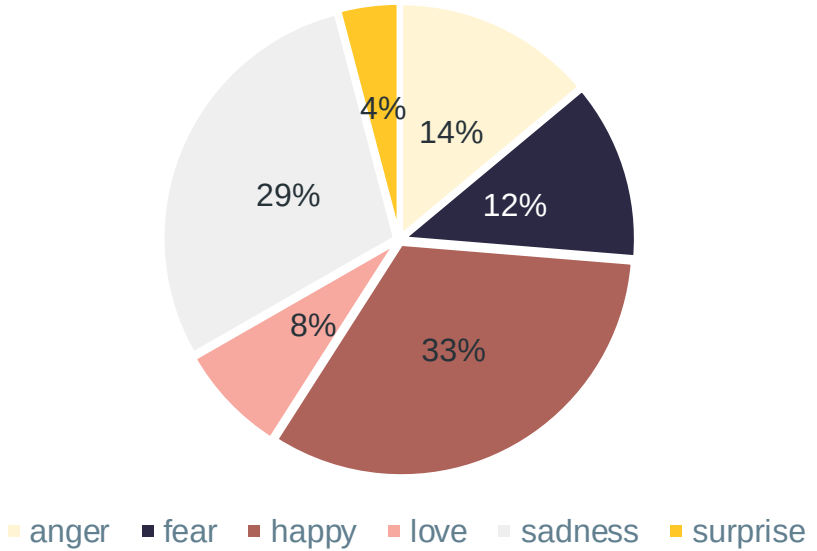
"My stars!" cried Aunt
Polly.
"Why, I can't believe it!"
She stared at the fence.
"You can work when
you want to, Tom
Sawyer!"

YouTube / Anger

You guys really have
the nerve to call him a
“devoted father?” What
a slap in the face to the
kids!

BERT Train Data set

Dataset	Train	Validation	Test	Total
anger	1,915	479	599	2,993
fear	1,698	424	530	2,652
happy	4,498	1,125	1,406	7,029
love	1,050	263	328	1,641
sadness	4,010	1,002	1,253	6,265
surprise	562	141	176	879
Total	13,733	3,434	4,292	21,459
Ratio	64%	16%	20%	100%



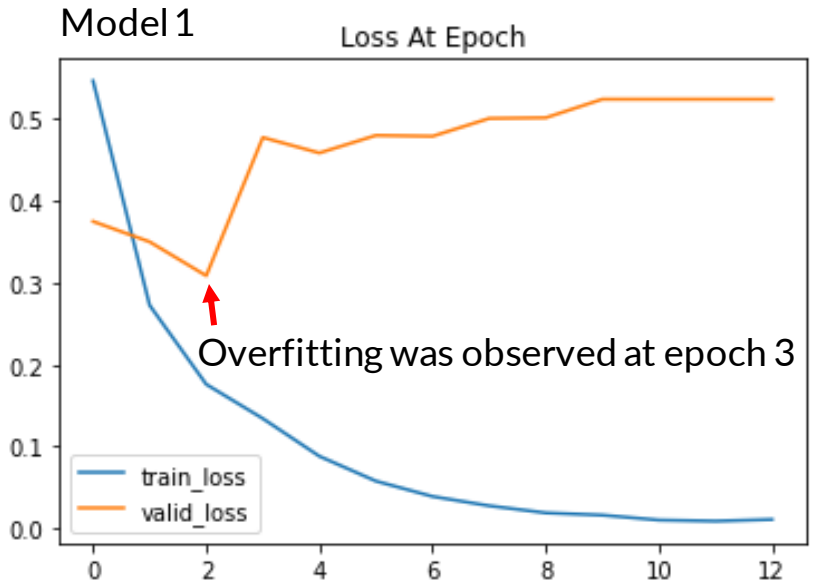
Our model - BERT

Train Bert with full data

Hyper Parameters

Learning rate = 1e-5, eps = 1e-8

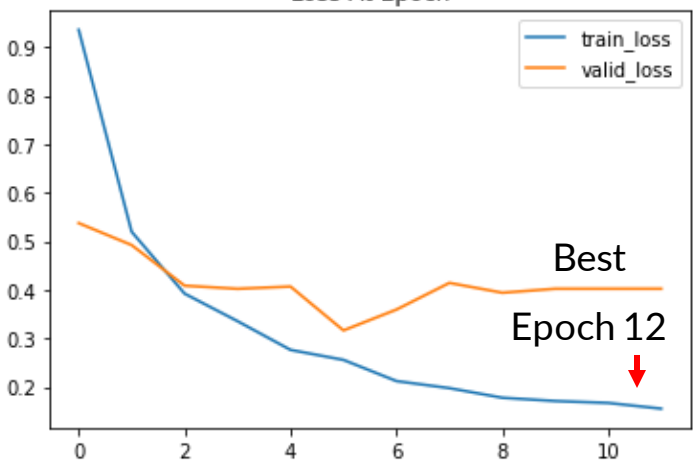
dropout = 0.1, epoch = 13, batch-size = 3



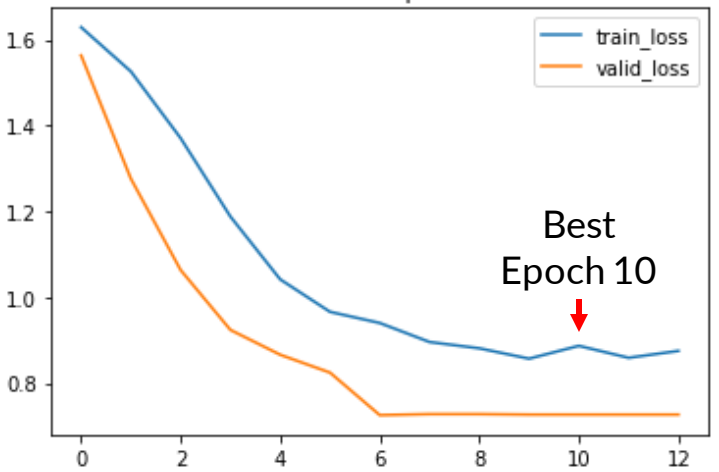
Our model - BERT

Change Hyper Parameters
Try to avoid overfitting

Model 2 Dropout **0.3**



Model 3 Dropout **0.5**

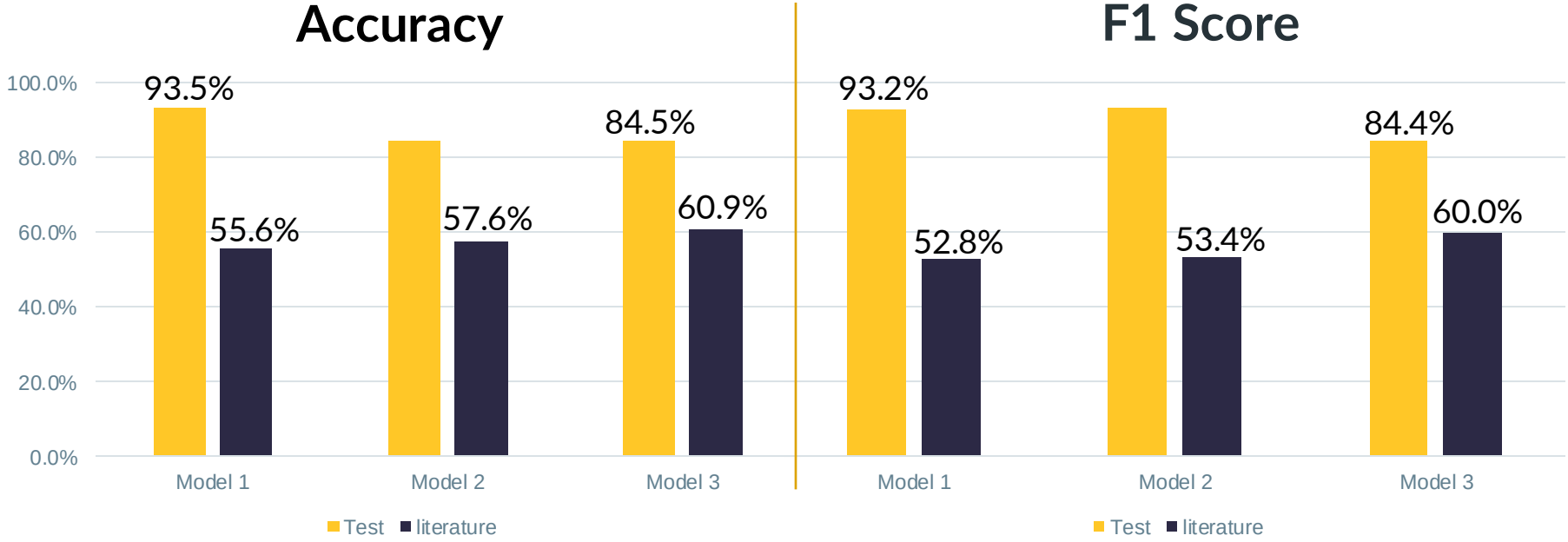


Train loss and validation loss show decreasing trend and converge to a certain value.
In model2 and model3 it seems that **overfitting is highly unlikely.**

Result

Model		test	lyric	poem	Youtube	novel
Model1	Accuracy	93.2%	40.13%	58.79%	67.65%	54.36%
	F1 Score	93.18%	37.22%	55.32%	65.66%	52.93%
Model2	Accuracy	93.5%	48.68%	56.36%	68.82%	55.90%
	F1 Score	93.53%	37.22%	55.79%	66.16%	54.38%
Model3	Accuracy	84.5%	44.74%	69.70%	71.76%	56.92%
	F1 Score	84.38%	43.84%	69.34%	71.09%	56.03%

Comparison the result



The best hyper-parameter about Literary works

Model 3 dropout = **0.5**, epoch = 10

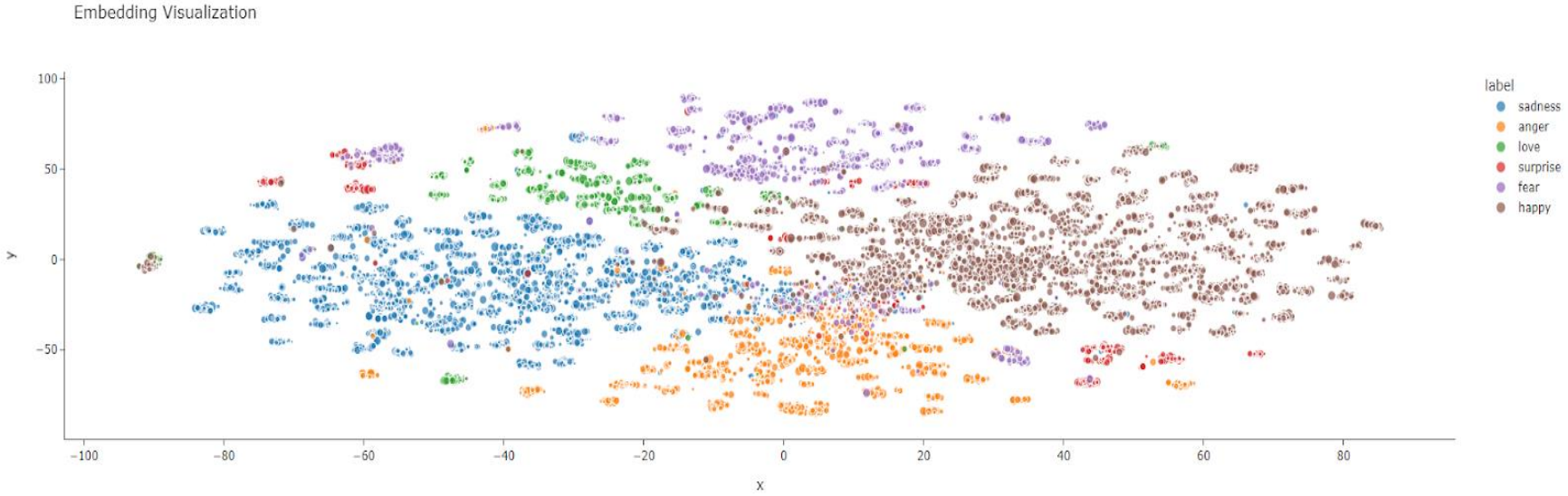
Our model – BERT Detail Table

	test	lyric	poem	Youtube	novel
anger	93.2%	43.3%	70.0%	70.0%	63.2%
fear	89.6%	23.3%	60.0%	73.3%	72.4%
happy	93.8%	76.7%	83.3%	100.0%	53.1%
love	89.0%	13.3%	12.9%	20.0%	37.5%
sadness	96.4%	41.9%	78.8%	80.0%	80.6%
surprise	83.0%	100.0%	27.3%	60.0%	21.2%
Total	93.2%	40.1%	58.8%	67.6%	54.4%
F1 Score	93.18%	37.22%	55.32%	65.66%	52.93%

	test	lyric	poem	Youtube	novel
anger	82.6%	26.7%	66.7%	63.3%	63.2%
fear	79.4%	23.3%	53.3%	66.7%	51.7%
happy	87.1%	76.7%	76.7%	93.3%	59.4%
love	57.6%	43.3%	58.1%	43.3%	46.9%
sadness	93.1%	51.6%	90.9%	93.3%	90.3%
surprise	75.0%	100.0%	72.7%	70.0%	30.3%
Total	84.5%	44.7%	69.7%	71.8%	56.9%
F1 Score	84.38%	43.84%	69.34%	71.09%	56.03%

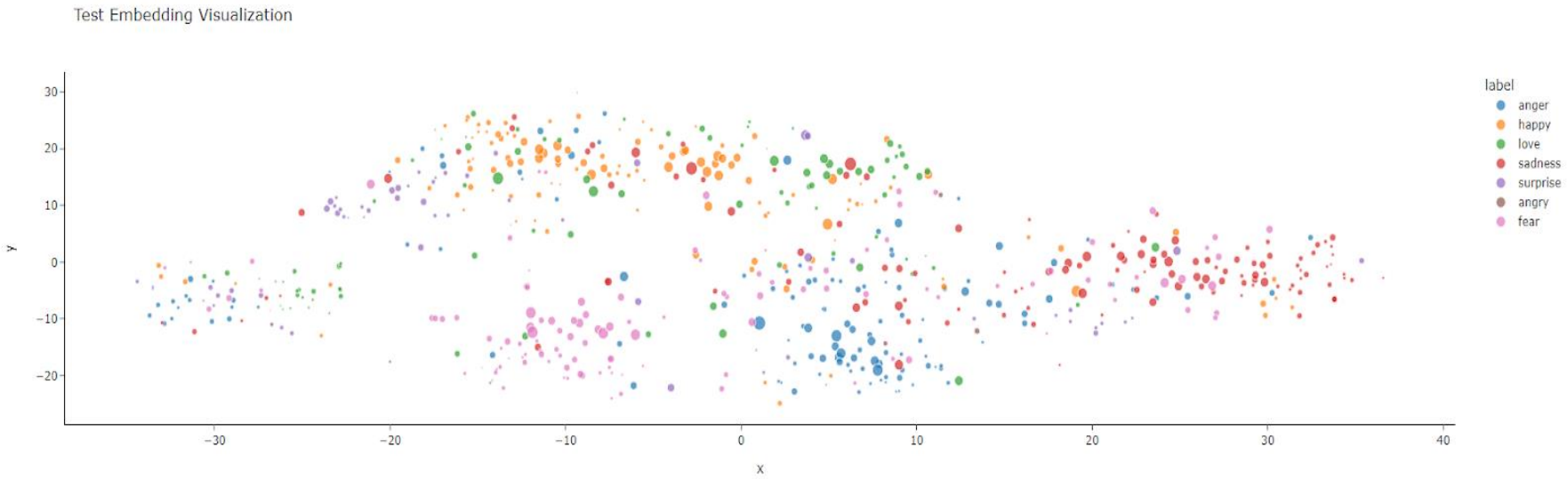
Result – Visualization

Scatter plot with train and validation data label clusters



Result – Visualization

Scatter plot with test data label clusters



THANKS