

CS-772: Assignment #4

Harsh Peswani 203050043

Pranjal Saini 203050014

Shubham Nemani 203050011

Team id - 47

Problem statement

The task is of sentiment analysis in which we train the model to classify the reviews into one of the 5 classes of ratings(1 to 5). 5 being the most positive review and 1 being most negative.

Architecture

- **Transformer Used** – Distill BERT (it is fintuned version of original BERT transformer model with less number of parameters)
- **Embedding dim** – 512
- **Vocab Size** – 30522
- **Model Encoder** – Our Distill BERT model Encoder consist of 6 transformer blocks

Libraries used :

Pandas

Numpy

Torch

String

Nltk

DistilBertForSequenceClassification

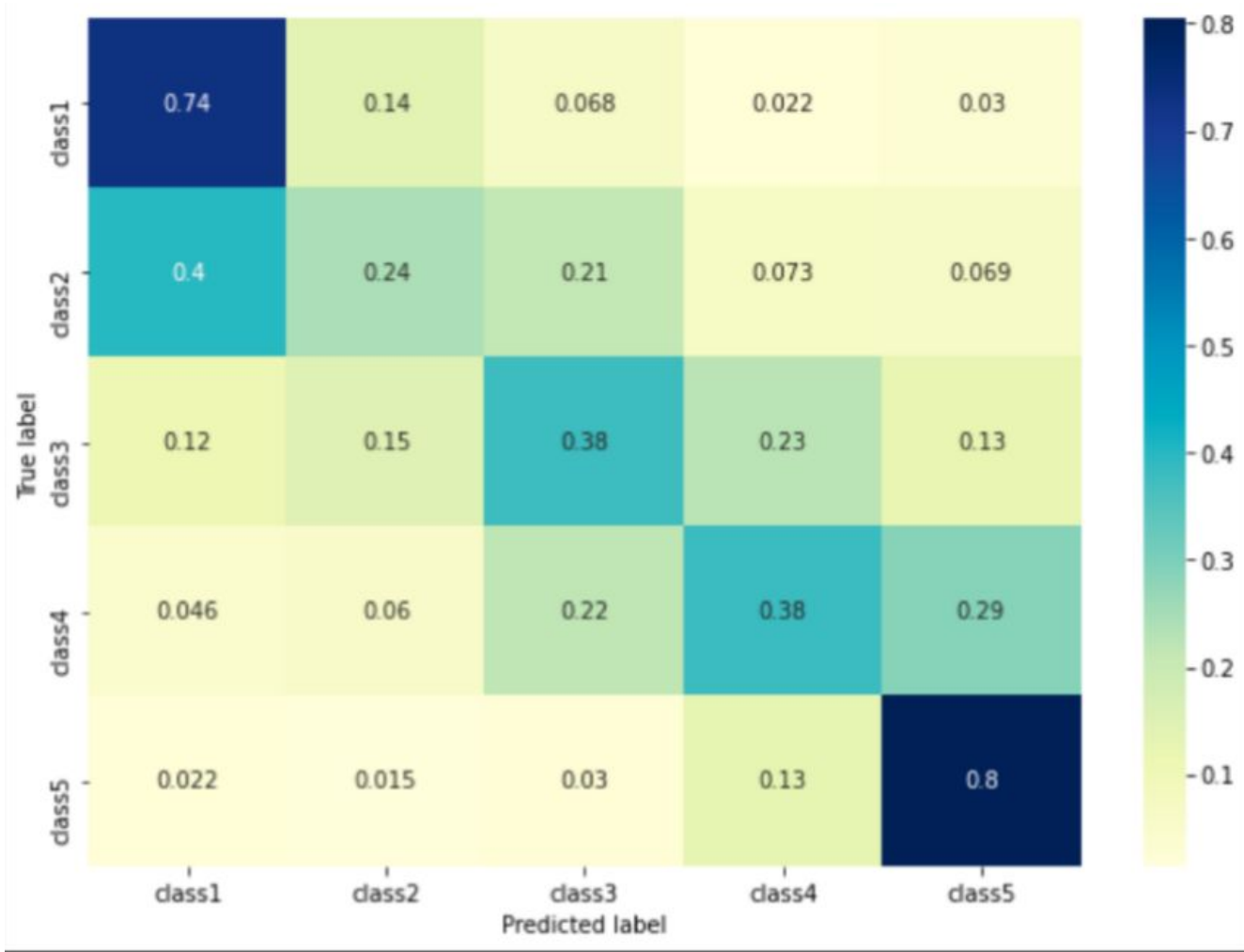
DistilBertTokenizer

WeightedRandomSampler

Results

Model Type	With Pre-trained Embeddings (Accuracy/Precision/Recall/ F1-Score)	Without Pre-trained Embeddings (Accuracy/Precision/Recall/ F1-Score)
RNN	0.58/0.12/0.20/0.15	0.58/0.12/0.20/0.15
LSTM	0.67/0.52/0.47/0.45	0.65/0.44/0.43/0.43
Bi-LSTM	0.68/0.51/0.46/0.46	0.68/0.49/0.44/0.45
GRU	0.67/0.50/0.45/0.45	0.65/0.47/0.45/0.45
Bi-GRU	0.67/0.52/0.44/0.44	0.67/0.49/0.44/0.44
Distill BERT	0.68/0.53/0.53/0.52	0.69/0.54/0.53/0.53

Confusion matrix - Prev Bi-LSTM model



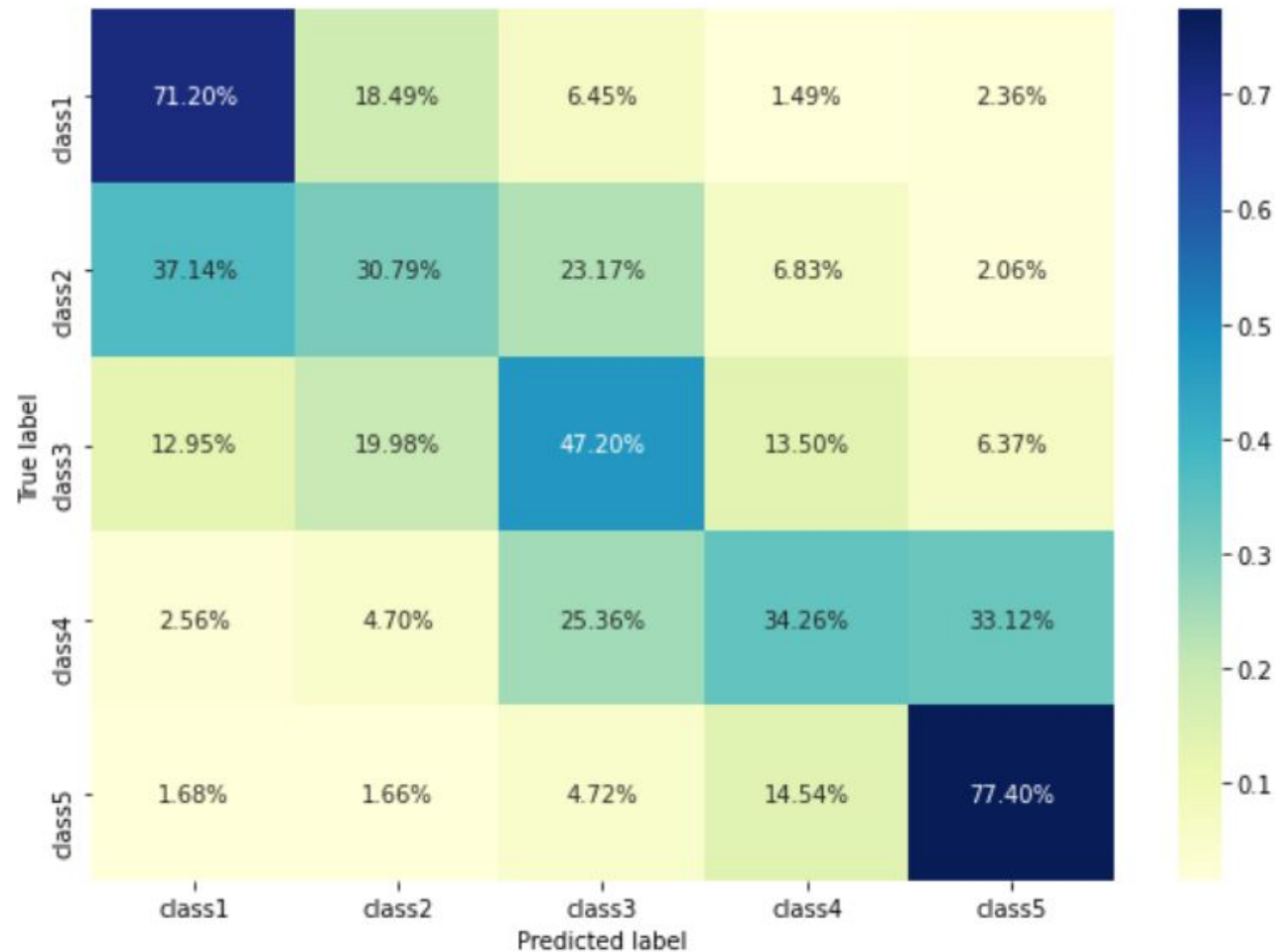
Here , the main confusion is between class 1 and class 2 .

Similarly class 4 has high confusion with class 3 and class 5.

To overcome these , we tried some techniques

1. We increased the word embedding size from 512 to 1024 in transformer , confusion decreased between class1 and class 2
2. We also tried training model with bigrams and trigrams , the confusion for class 3 decreased i.e. around 47% predicted true , but then a slight confusion arised for classe 1 and class 5

Confusion Matrix for transformer



Conclusion

We are able to decrease the confusion between the classes 2, 3 and 4 at little cost of confusion in classes 1 and 5

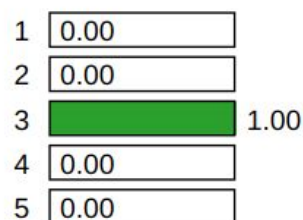
Qualitative Analysis

- **Sentence :**

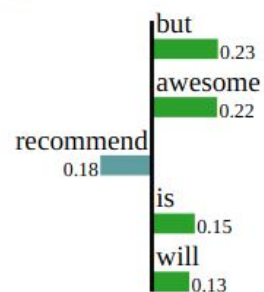
'This camera is awesome but i will not recommend it to anybody'

- Predicted – 3

Prediction probabilities

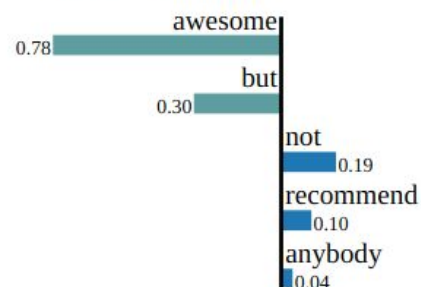


NOT 3



3

NOT 1



1

NOT 4



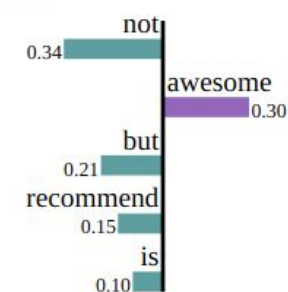
4

NOT 2



2

NOT 5



5

Text with highlighted words

This camera is awesome but i will not recommend it to anybody