Amazon Reviews Sentiment Analysis using Neural Networks

Nitin Godi

INTRODUCTION

- **Sentiment analysis** (also known as **opinion mining** or **emotion AI**) is the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.[1]
- Feedbacks and reviews from the customers dictates majority of the executive decisions, business strategies, etc.
- *The problem statement is to classify the reviews collected by Amazon based on their sentiments.*

DATA DESCRIPTION

- The dataset used for the project is “Amazon Reviews for Sentiment Analysis”[2] obtained from Kaggle.
- Each observation has the following format, <__label__> <title>: <review>.
- These observations were extracted and stored in the form of pandas data frame using regular expressions.
- The data frame has 3599330 observations and two columns namely, ‘label’ and ‘text’.

EXPERIMENT SETUP

- Used 11 different pre-trained models.
- Selected 3 models out of 11.
- Models are Mobile Bert, Distil Roberta and Ensemble of the two models.

Mobile Bert

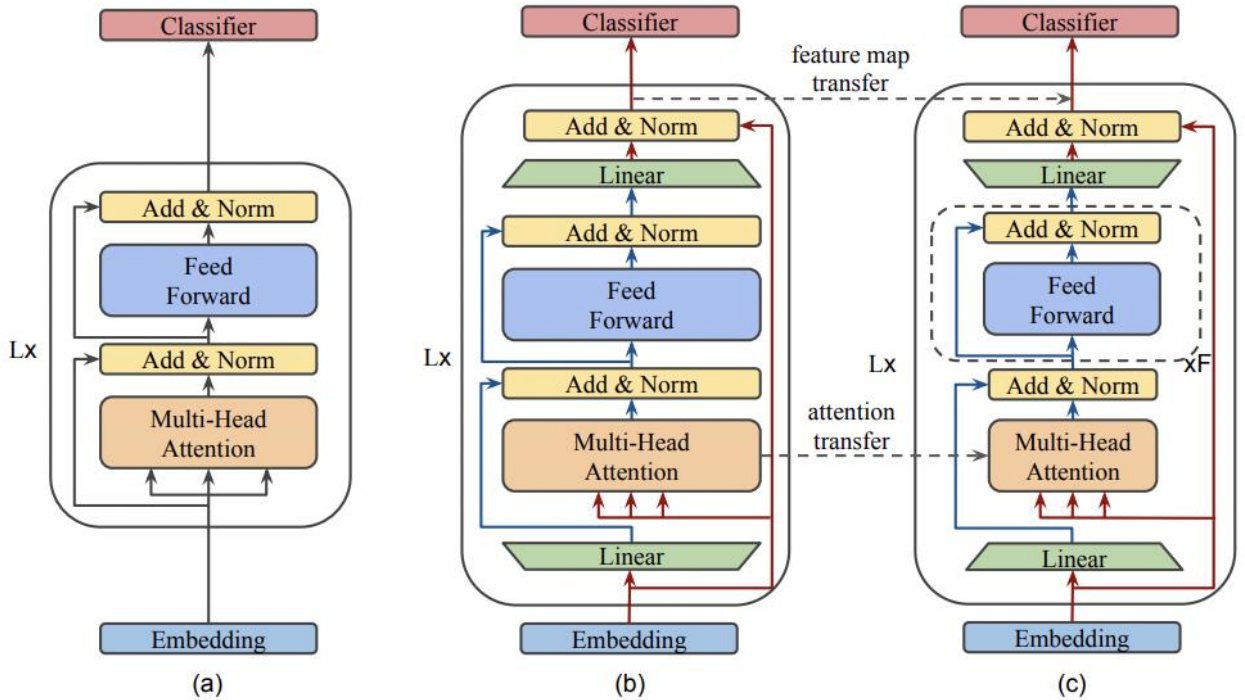


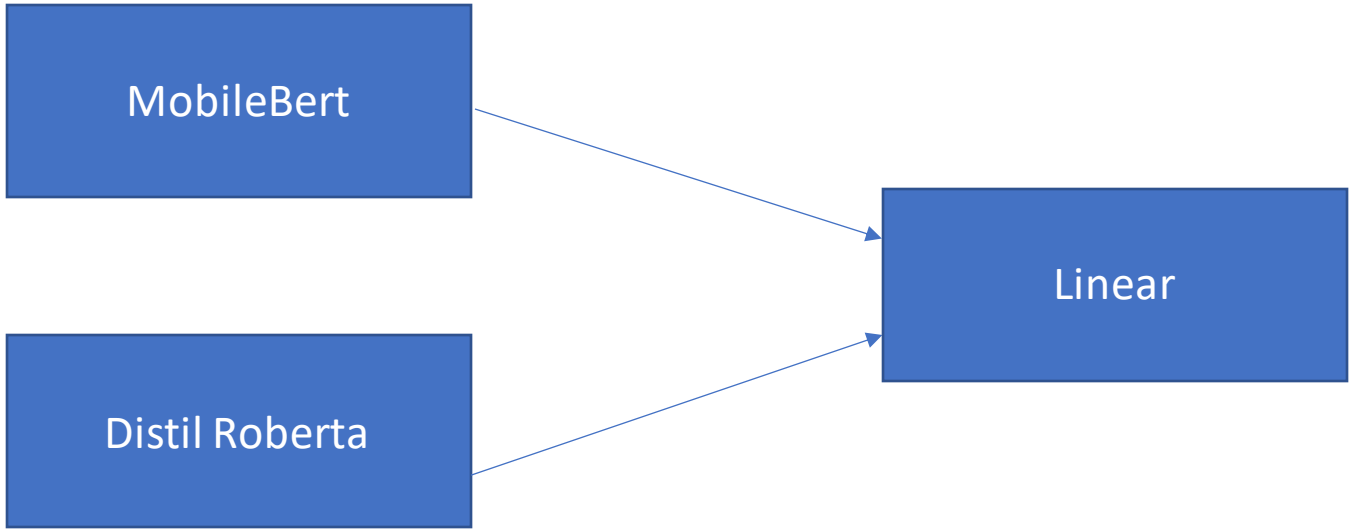
Figure 1: Illustration of three models: (a) BERT; (b) Inverted-Bottleneck BERT (IB-BERT); and (c) MobileBERT. In (b) and (c), red lines denote inter-block flows while blue lines intra-block flows. MobileBERT is trained by layer-to-layer imitating IB-BERT.

Distil Roberta

- Distill Roberta is also a pre-trained model which is a distilled version of Roberta model. Its training procedure is similar to Distil Bert.
- The model comprises of 6 layers, 768 dimensions and 12 head resulting in 82million parameters.
- This model was developed for English language and is case sensitive.
- It is modeled using Masked Language Modeling technique.

Ensemble Model

- It is the ensemble of Distil Roberta and Mobile Bert Models.



RESULTS

Models	Accuracy	Precision	Recall
Distil Bert	0.9291	0.9289	0.9273
Electra	0.9323	0.9316	0.9269
Distil Roberta	0.9346	0.9350	0.9307
Mobile Bert	0.9296	0.9308	0.9283
Flaubert	0.8778	0.8921	0.876
Ensemble of Electra and Distil Bert	0.92894	0.9377	0.92373
Ensemble of Electra and Electra	0.92773	0.92568	0.93436
Ensemble of Mobile Bert and Electra	0.93487	0.93672	0.93655
Ensemble of Mobile Bert and Distil Roberta	0.94080	0.9462	0.9378
Ensemble of Electra and Distil Roberta	0.89364	0.89047	0.90275
Ensemble of Mobile Bert and Distil Bert	0.93199	0.93633	0.93197

CONCLUSION

- In the dataset description in Kaggle website, the benchmark is precision and recall scores of 0.916.
- 9 out of 11 models performed better than the benchmark.
- Ensemble of Mobile Bert and Distil Roberta models is the best model of the lot.

REFERENCES

1. https://en.wikipedia.org/wiki/Sentiment_analysis
2. https://www.kaggle.com/bittlingmayer/amazonreviews
3. Sun, Z., Yu, H., Song, X., Liu, R., Yang, Y., & Zhou, D. (2020). Mobilebert: A compact task-agnostic Bert for Resource-limited devices. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. https://doi.org/10.18653/v1/2020.acl-main.195
4. https://huggingface.co/distilroberta-base