

IMDB Movie Review Sentiment Analysis using NLP

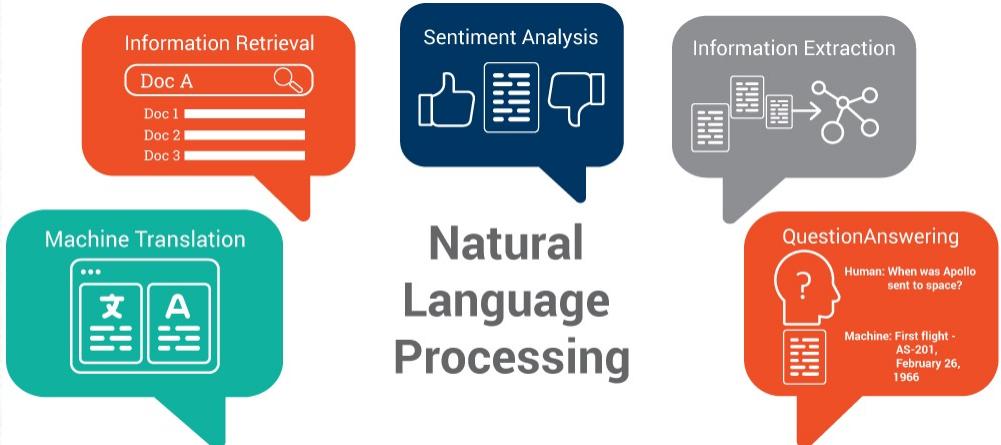
NLP Using BERT

- *Chandramouli Yalamanchili*

Abstract

- Natural Language Processing (NLP) has seen great progress in last two decades and is being integrated more and more into the day-to-day life.
- Rise of data generation & availability, reduction in compute cost, and advancement of technology have helped the growth in NLP domain, and it is already being used by several market segments.
- Sentiment analysis is the sub field of NLP that works on deriving the feeling or sentiment from the human language using NLP techniques.
- Goal for this project is to build a DNN model using pre-trained BERT models and see how efficiently this model can classify the sentiment attached to a movie review.

Domain Background – NLP & Sentiment Analysis?

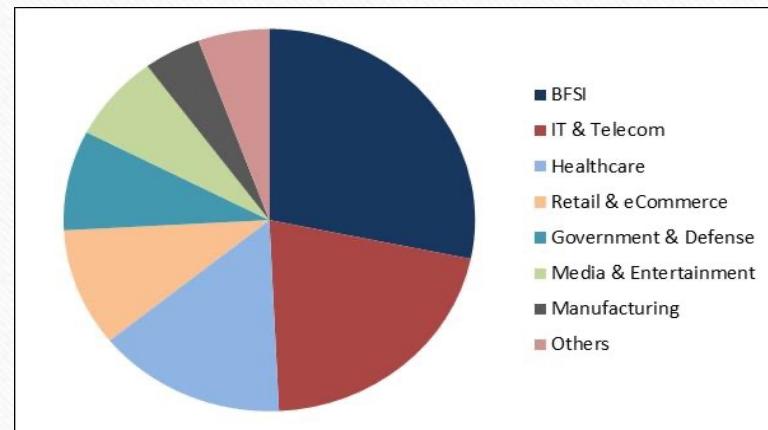
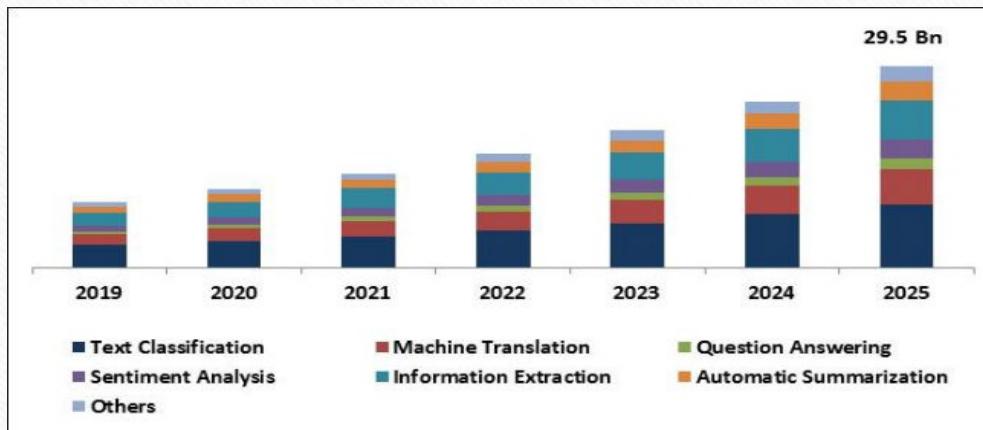


Natural Language Processing (NLP) is a subfield of Artificial Intelligence that helps machines to understand and communicate in same natural language that humans use.



Sentiment analysis, also known as opinion mining is a subfield of NLP that performs contextual analysis on the text with the goal of determining the emotional tone attached to the text.

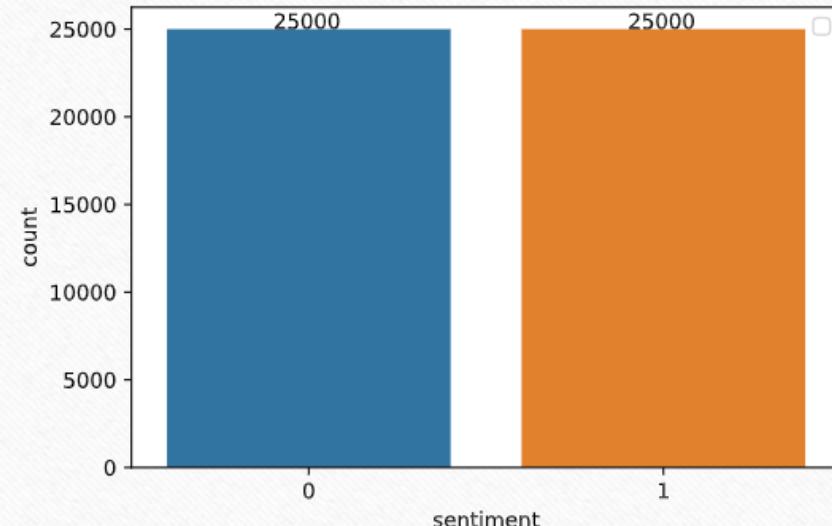
Domain Background – NLP Stats



- NLP market is expected to raise up to \$29.5 Billion by 2025 at global market level at the growth rate of 20.5 % cumulative annual growth rate.
- NLP is being used in several market segments, with banking & finance, IT & Telecom, and healthcare leading the market share.

Exploratory Data Analysis - Input dataset

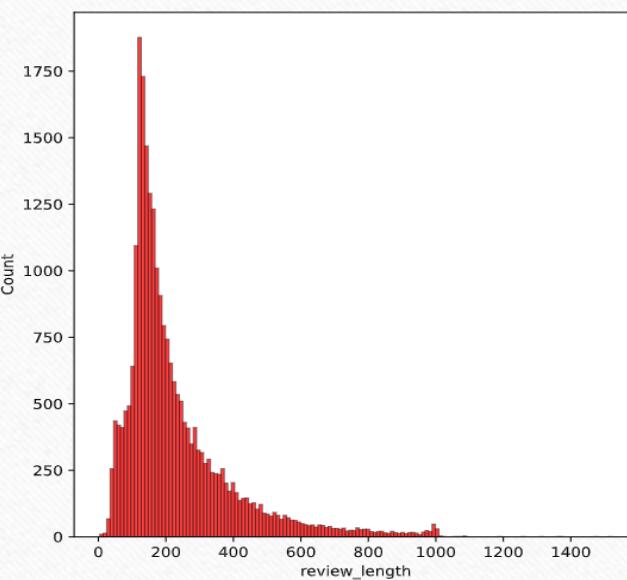
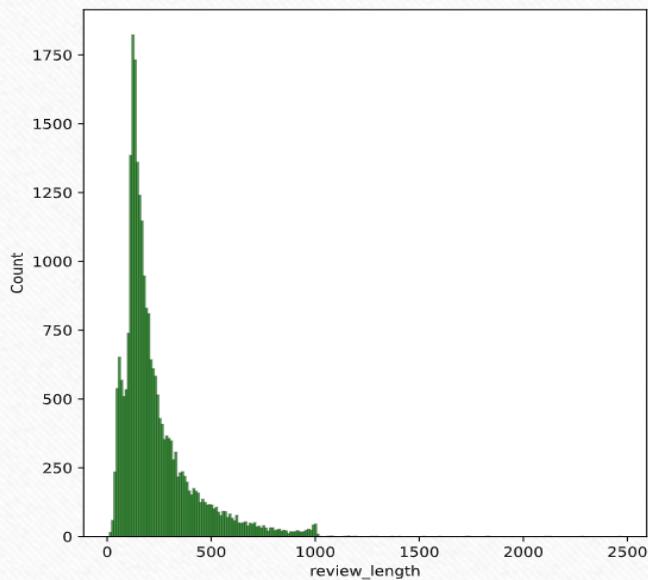
Parameter	Value
Shape of the dataset	(50000, 2)
Total number of reviews	50,000
Number of positive reviews	25,000
Number of negative reviews	25,000



Dataset Location:

<https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

Exploratory Data Analysis – Review Length Analysis



Parameter	Positive Review	Negative Review
Mean	232.85	299.46
Median	172	174
Minimum	10	4
Max	2470	1522

- As we can see the distribution of review length (number of words) look similar for both positive and negative reviews, most of the reviews are around 200 words.
- Positive reviews seem to have some outliers expanding the max review length close to 2500 words.

Exploratory Data Analysis – WordCloud Plots



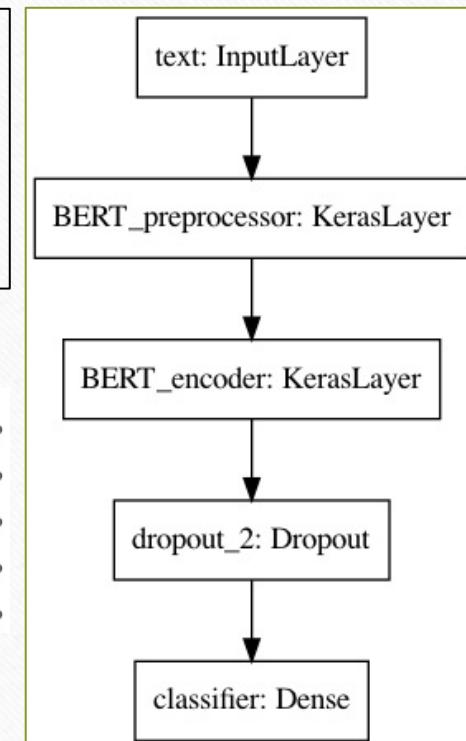
WordCloud plots showing the frequent words from both the positive reviews and negative reviews.

Modeling - DNN using BERT model layers

- We are using DNN model with 5 layers total, including input text layer.
- We have used BERT pre-trained pre-processing model and encoder transformation model as the layers in the model along with one dropout and one dense layer.
- We have trained the model for 5 Epochs using Adam optimizer for fine tuning, and we have used the binary cross loss function.

Model: "model_2"			
Layer (type)	Output Shape	Param #	Connected to
text (InputLayer)	(None,)	0	
BERT_preprocessor (KerasLayer)	{'input_word_ids': (0	text[0][0]	
BERT_encoder (KerasLayer)	{'sequence_output': 28763649	BERT_preprocessor[0][0] BERT_preprocessor[0][1] BERT_preprocessor[0][2]	
dropout_2 (Dropout)	(None, 512)	0	BERT_encoder[0][5]
classifier (Dense)	(None, 1)	513	dropout_2[0][0]
Total params: 28,764,162			
Trainable params: 28,764,161			
Non-trainable params: 1			

```
Epoch 1/5  
1050/1050 [=====] - 7114s 7s/step  
Epoch 2/5  
1050/1050 [=====] - 6929s 7s/step  
Epoch 3/5  
1050/1050 [=====] - 6724s 6s/step  
Epoch 4/5  
1050/1050 [=====] - 6714s 6s/step  
Epoch 5/5  
1050/1050 [=====] - 6959s 7s/step
```



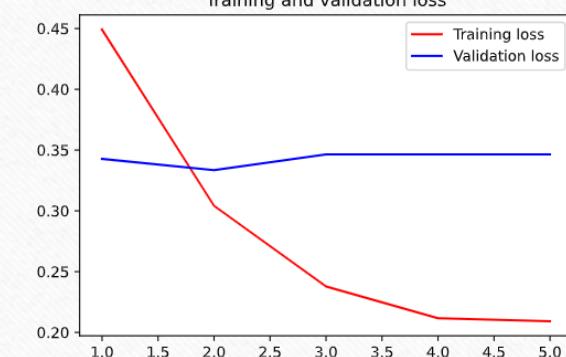
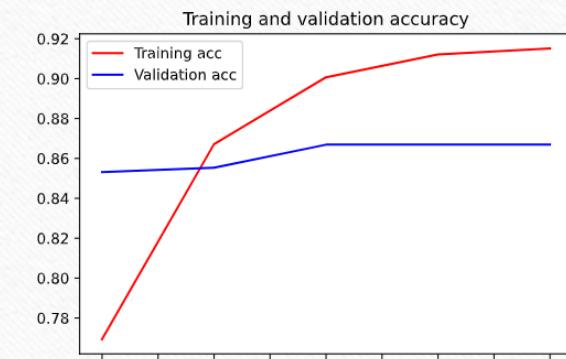
Modeling – BERT Model performance

Performance Metric	Value
Accuracy	87%
Loss	34%

```
input: this is such an amazing movie!
input: The movie was great!
input: The movie was unrealistic.
input: Casting was horrible in this movie!
input: The movie was terrible...
input: I would love to watch this movie again!
input: There is nothing special about this film.
```

```
; score: 0.998134
; score: 0.989870
; score: 0.059322
; score: 0.003026
; score: 0.002388
; score: 0.976245
; score: 0.066461
```

- Accuracy of 87% reflects that the model has done a decent job in classification of movie review.
- We can also see the score generated by model for random texts, and the scores look good.
- Based on the accuracy and loss plots, we can see that the model is overfitting as the validation loss is not quite following the training loss. Training the model further for few more Epochs might improve the model's performance.



Conclusion

- There has been a rapid growth last couple of decades in NLP domain due to increase in availability of electronic data, reduction in cost for compute power, and the advent of internet.
- With the advancements in the space of NLP and the technology overall, we now have capability to consume the real time data and then execute it through the NLP models to produce insights about customer's sentiment in real time helping businesses to take the customer engagement and satisfaction to the next level
- In this project we have taken the IMDB reviews dataset that had 50,000 reviews and used it to train the Keras model with BERT pre-trained models. We have established that the BERT models have done a good job in classifying the movie review as positive or negative with 87% of accuracy score. We believe that allowing more time for training the model would yield even better classification outcomes.
- We believe the recent advancements in the form of BERT models is the next step in the evolution for NLP as these models come packed with lot of features, they also come in several processing levels making these models accessible to small scale applications with low compute powers as well.

References

1. Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc. Retrieved on 06/02 from <https://surface.syr.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1019&context=cnlp>
2. Vishal A. Kharde, S.S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques" International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11, April 2016. Retrieved from <https://arxiv.org/pdf/1601.06971.pdf>
3. Prabowo, R., & Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics*, 3(2), 143-157. Retrieved from.
4. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. J. (2011, June). Sentiment analysis of twitter data. In Proceedings of the workshop on language in social media (LSM 2011) (pp. 30-38). Retrieved on 06/02 from <https://www.aclweb.org/anthology/W11-0705.pdf>
5. Lin, C., & He, Y. (2009, November). Joint sentiment/topic model for sentiment analysis. In Proceedings of the 18th ACM conference on Information and knowledge management (pp. 375-384). Retrieved on 06/02 from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.163.5917&rep=rep1&type=pdf>
6. Whitelaw, C., Garg, N., & Argamon, S. (2005, October). Using appraisal groups for sentiment analysis. In Proceedings of the 14th ACM international conference on Information and knowledge management (pp. 625-631).
7. Maas, A., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011, June). Learning word vectors for sentiment analysis. In Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies (pp. 142-150). Retrieved on 06/02 from <https://www.aclweb.org/anthology/P11-1015.pdf>.
8. Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*, 30(4), 330-338. Retrieved on 06/02 from <https://www.sciencedirect.com/science/article/pii/S1018363916300071>
9. IMDB Dataset of 50K Movie Reviews. Retrieved on 06/02 from <https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>
10. Overview about the sentiment analysis
 - Overview about sentiment analysis. Retrieved on 06/02 from <https://monkeylearn.com/sentiment-analysis/>
 - Use cases for sentiment analysis. Retrieved on 06/02 from <https://www.whoson.com/customer-service/top-ten-benefits-of-sentiment-analysis/>
11. What is NLP and why is it important? Retrieved on 06/02 from <https://www.analyticsinsight.net/what-is-nlp-and-why-is-it-important/>
12. Why Sentiment Analysis is the major game changer for patient care, by Spin-Strategy. Retrieved on 06/02 from <https://www.spinalyticsandstrategy.com/blog/why-sentiment-analysis-major-game-changer-for-patient-care/>
13. Top 5 Semantic Technology Trends to look for in 2017, by Milena Yankova January 2017. Retrieved on 06/02 from <https://www.ontotext.com/blog/top-5-semantic-technology-trends-2017/>
14. Natural Language Processing Market Size, December 2019. Retrieved on 06/02 from <https://www.kbvresearch.com/natural-language-processing-market/>
15. Top NLP Algorithms and concepts – Igor Bobriakov, December 2019. Retrieved on 06/02 from <https://www.datasciencecentral.com/profiles/blogs/top-nlp-algorithms-amp-concepts>