Product Sentiment Analysis

- Pushkar Chougule
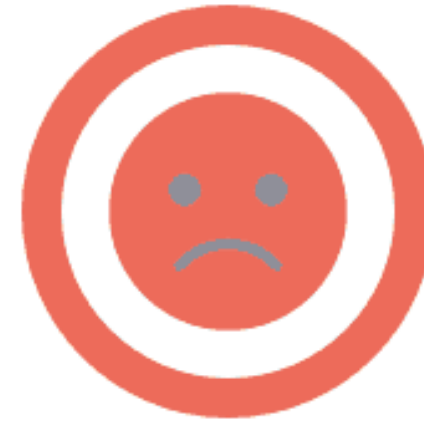
Content

- Business Problem and Background
- Data Explanation and Preparation
- Feature Analysis and Methods
- Machine Learning Models and Evaluation
- Conclusion
- Assumptions and Limitations
- Future Usage and Recommendations
- Ethical Assessment
- References

Business Problem and Background

- Baby Products

- Customer reviews Sentiment

- Maintain Popular Products in inventory

- Product Improvement opportunities

Data Explanation and Preparation

	name	review	rating
0	Planetwise Flannel Wipes	These flannel wipes are OK, but in my opinion ...	3
1	Planetwise Wipe Pouch	it came early and was not disappointed. i love...	5
2	Annas Dream Full Quilt with 2 Shams	Very soft and comfortable and warmer than it l...	5
3	Stop Pacifier Sucking without tears with Thumb...	This is a product well worth the purchase. I ...	5
4	Stop Pacifier Sucking without tears with Thumb...	All of my kids have cried non-stop when I trie...	5
5	Stop Pacifier Sucking without tears with Thumb...	When the Binky Fairy came to our house, we did...	5
6	A Tale of Baby's Days with Peter Rabbit	Lovely book, it's bound tightly so you may not...	4

Counts by Ratings

	Counts
5	106265
4	33053
3	16705
1	15100
2	11261

Features with NULL Values
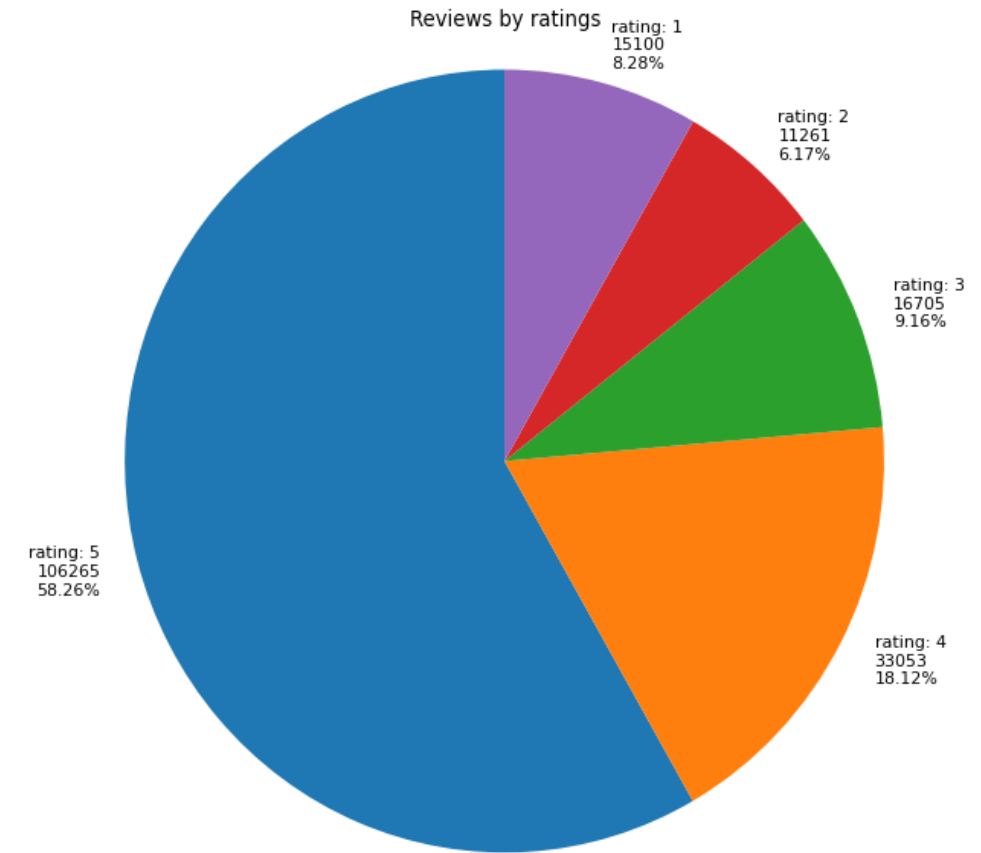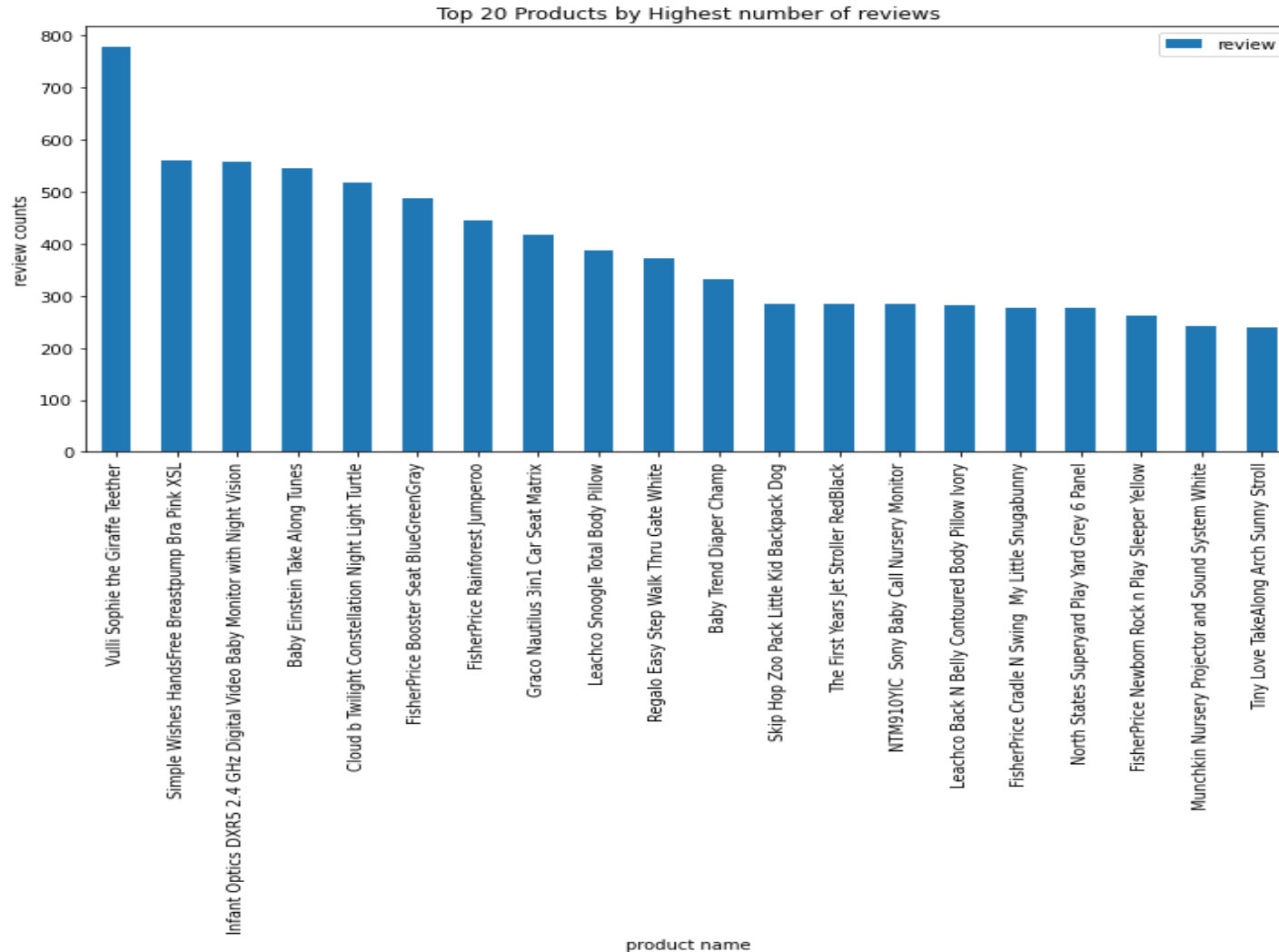
```
name        True
review      True
rating      False
dtype: bool
```

```
name        318
review      829
rating        0
dtype: int64
```

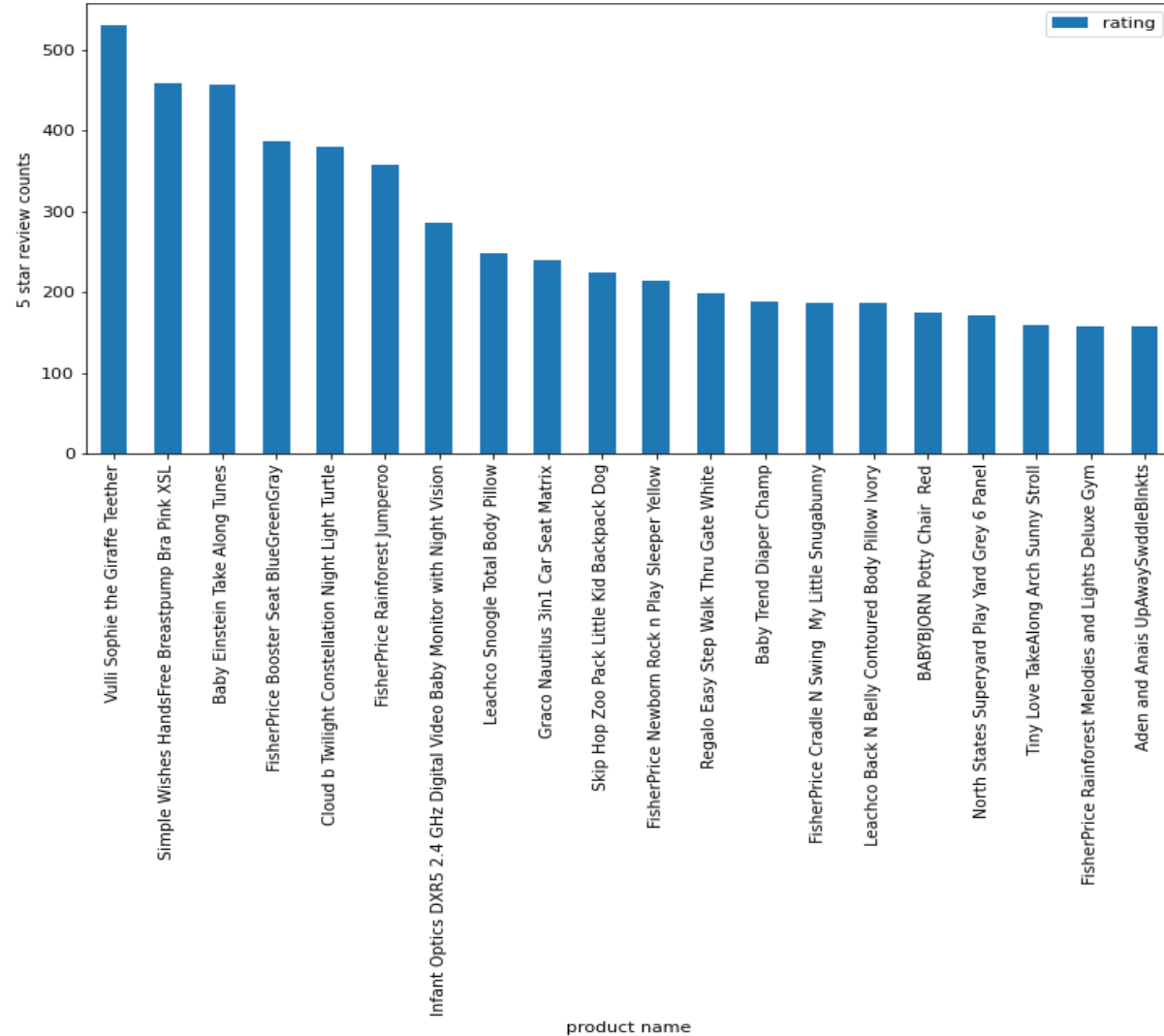Total number of reviews after cleanup : 182, 384

Total number of unique products after cleanup : 32, 305
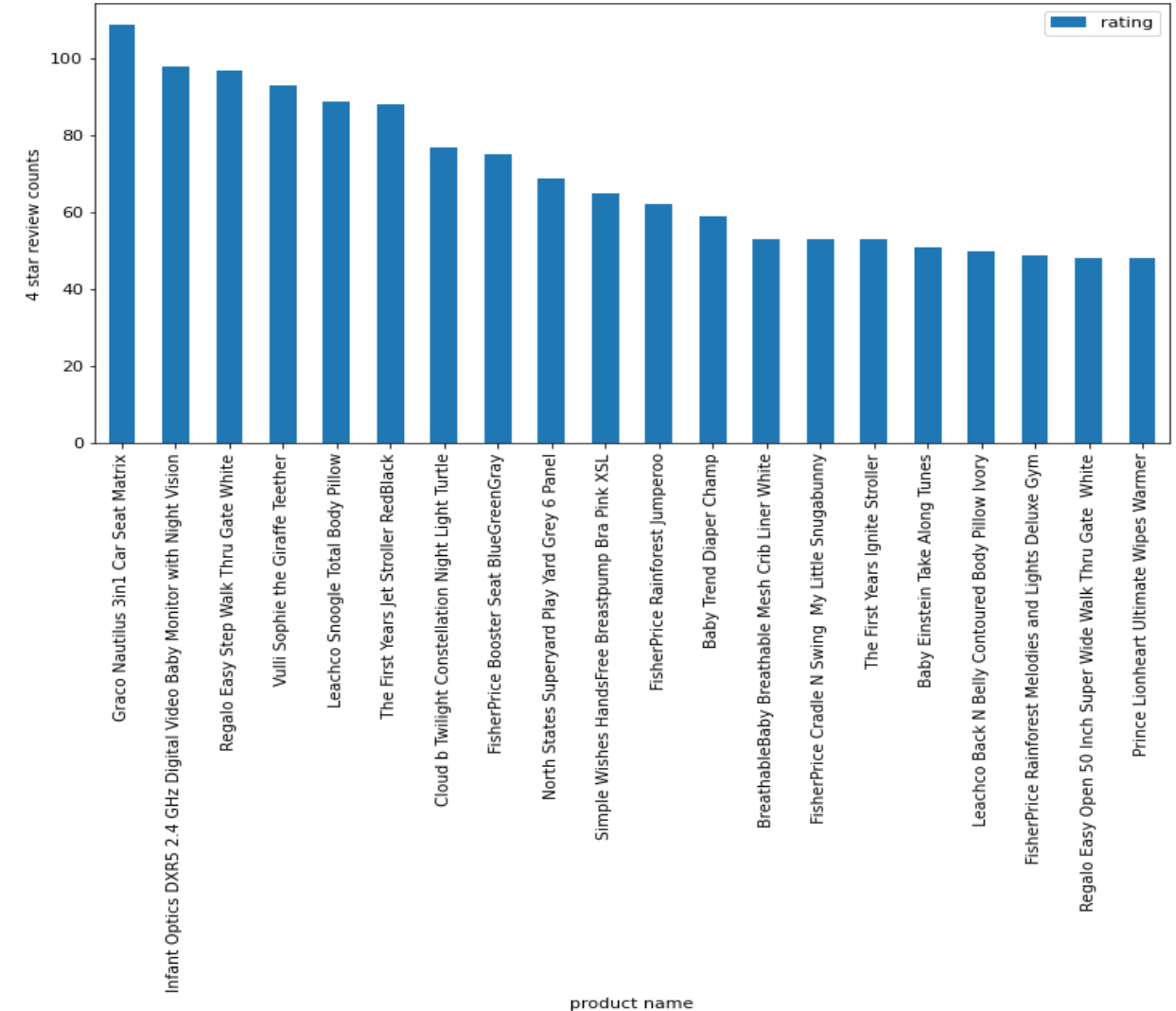
Features Analysis and Methods

Top 20 Products by Highest number of reviews

Reviews by ratings

Features Analysis and Methods (continued)

Features Analysis and Methods (continued)

Features Analysis and Methods
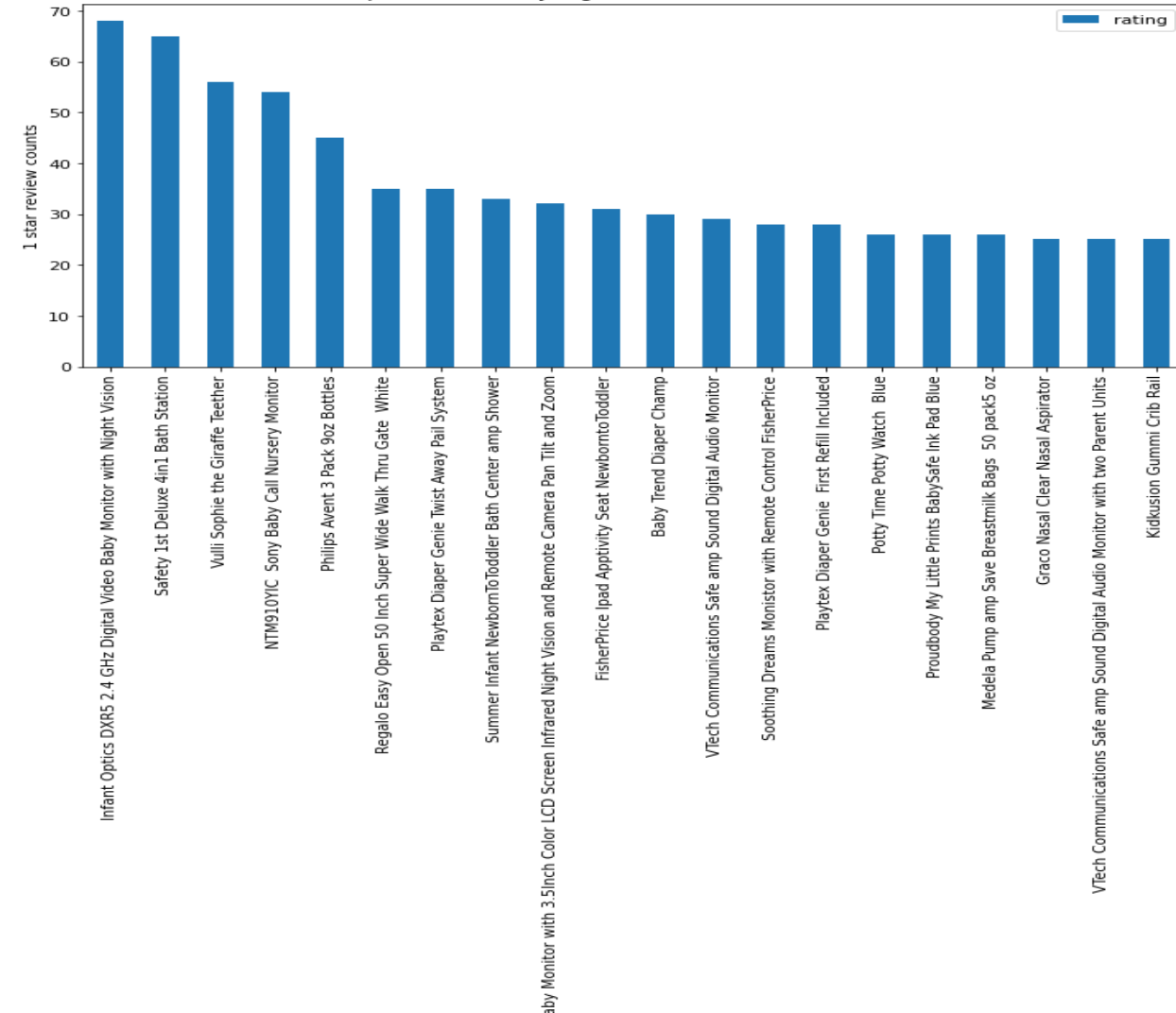
Machine Learning Models and Evaluation

MLP Classifiers (sci-kit learn)

Artificial Neural Network (Keras)

TextBlob (sentiment analysis)

Recurrent Neural Network (LSTM)

Machine Learning Models and Evaluation

Tf-Idf Vectorizer

sklearn MLP accuracy score using TfIdf vectorizer is: 72.89 %

Keras neural network accuracy score using Tfidf vectorizer is: 72.94 %

TextBlob (sentiment analysis)

textblob accuracy score is: 42.75 %
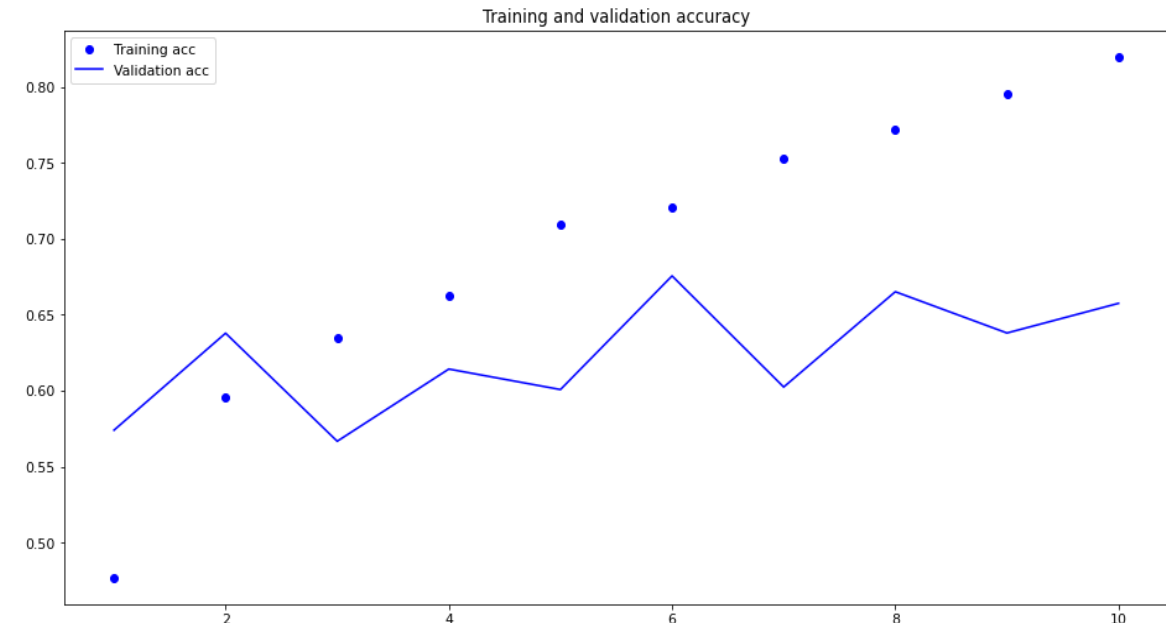
Recurrent Neural Network (LSTM)

Count Vectorizer

sklearn MLP accuracy score using Count vectorizer is: 73.24 %

Keras neural network accuracy score using Count vectorizer is: 74.48 %

Training and validation accuracy

Machine Learning Models and Evaluation (Classification reports – using Count Vectorizer)

MLP Classifiers

Artificial Neural Network

sklearn MLP accuracy score using Count vectorizer is: 73.24 %

	precision	recall	f1-score	support
0	0.76	0.70	0.73	5499
1	0.70	0.77	0.73	5640
2	0.78	0.76	0.77	5753
accuracy			0.74	16892
macro avg	0.75	0.74	0.74	16892
weighted avg	0.75	0.74	0.74	16892

Keras neural network accuracy score using Count vectorizer is: 74.48 %

	precision	recall	f1-score	support
0	0.74	0.71	0.73	5499
1	0.67	0.80	0.73	5640
2	0.81	0.70	0.75	5753
accuracy			0.73	16892
macro avg	0.74	0.73	0.73	16892
weighted avg	0.74	0.73	0.73	16892

Conclusion

❖ Most popular product in terms of total number of reviews is "Vulli Sophie Girrafe Teether"

❖ Majority of the product reviews fall in the positive reviews, which means most of products are working well for the customers, where customer reviews available

❖ "Vulli Sophie Girrafe Teether" is also one of the top 5-star rated – popular product

❖ "Graco Nautilius 3in1 Car seat" is another popular product in terms of positive ratings

❖ "Digital Video Baby monitor with Light Night vision" product had good number of positive reviews as well as neutral and negative ratings. So, it appears to be a product needing some improvements based on customer reviews.

❖ "Sony Baby call Nursery Monitor" and "Safety 1st Deluxe 4 in 1 Bath station" were the other products with the high number of negative review comments

Assumptions / Limitations

Assumption: all details captured correctly and no inaccuracies in ratings details listed in the dataset

Limitation : Majority of reviews (76%) were positive and about smaller portion was neutral (9%) or negative (15%). Capture additional reviews for better data balancing.

Future Usage / Recommendations / Ethical Assessment

Future use: The modeling approach can be reused for individual product level customer sentiments analysis or for capturing the customer sentiments for other product categories.

Ethical Assessment : No personal data collected for sentiment analysis. Any potential personalized information to be cleaned out prior to processing.

References

- **Data:** https://www.kaggle.com/sameersmahajan/reviews-of-amazon-baby-products

- Albon, Chris (2018) : **Machine Learning with Python Cookbook** (O'Reilly Media Inc.)

- Bengfort, B., Bilbro, R. & Ojeda, T. (2018) : **Applied Text Analysis with Python** (O'Reilly Media Inc.)

- Chollet, Francois (2018) : **Deep Learning with Python** (Manning Publications)

- MLP classifier in sklearn library : https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html

- Textblob sentiment analysis : https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524

THANK YOU