- **Business Problem:**

Global giants like Amazon sell varieties of products in multiple categories and one of the important product category is Baby Products. As a part of current business opportunity, company needs to identify the customer feedback on variety of products and analyze the customer sentiments / satisfaction about the products purchased. This also helps in knowing which products are being preferred and in demand the most, by the customers and it can plan on keeping sufficient stocks on these in inventory. It can also be a great source of product improvement opportunities based on the customer feedback.

- **Background/History:**

All the product selling companies / platform strive to sale better quality products as well as are on the lookout to improve their products line up on an ongoing basis. This helps them with enhanced customer satisfaction, increase popularity / usage of the products and in turn results in maintaining market leadership positions, brand value and ultimately helps with increased sales and profit.

As a part of this project, we will using the Amazon Baby products reviews to analyze sentiments for the Baby products available for purchase on the Amazon website based on the reviews from the customers for variety of products. This helps us understand how the products are being received / welcomed or people are not happy with the purchases. It will also help us understand the improvement opportunities in the respective products for future improvements. And which products are being preferred by the customers. All of this information ultimately helps Amazon decide which products need to be maintained in the warehouse / inventory in sufficient quantity for the customers and which ones should be avoided from the future inventory and share the feedback accordingly with the product supply team.

- **Data Explanation (Data Prep/Data Dictionary/etc):**

Following features are present in the datasets:

- name : Product name purchased by the customer
- review : Review comments from the customer detailing their experience
- rating : Product ratings given by the customers based on their experience

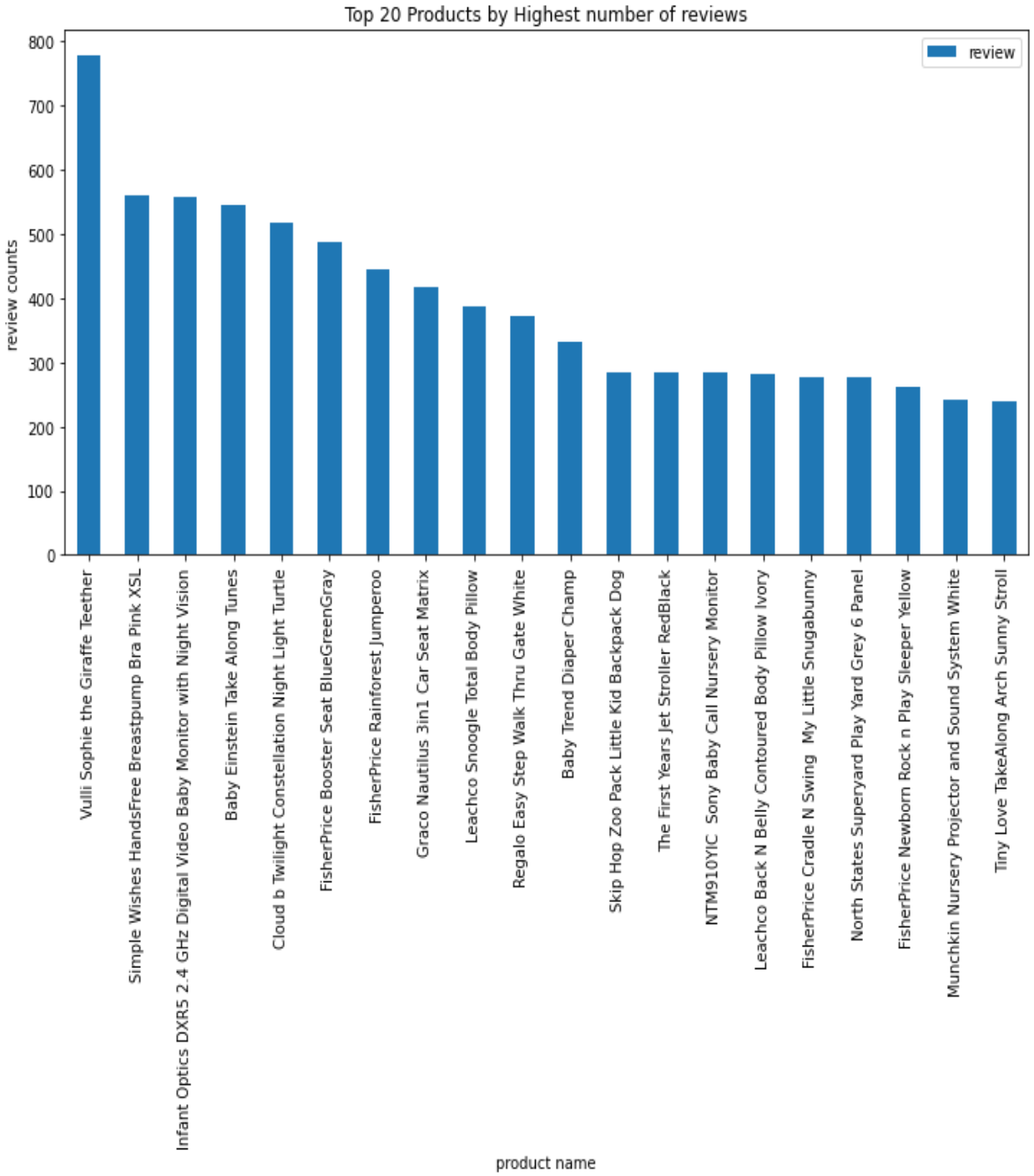
As a part of the final predictions, we will be using the customer review in predicting whether a given review comment is positive / neutral / negative. We will use rating column to verify the categorization / predictions.

The dataset has total of 183,531 reviews for the 32,400+ unique Baby products listed. I came across null entries in both the product name (318 null values) and review (829 null values) columns. So, decided to drop these rows as these rows won't be helpful for analysis purpose.

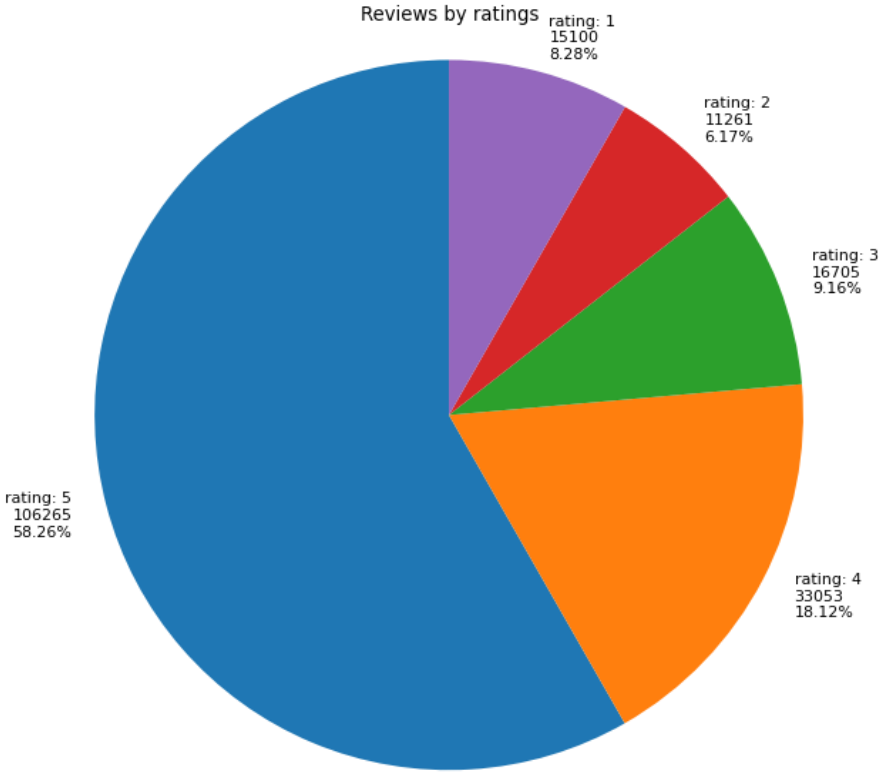
For the verification purpose, 4 star and 5 star rating will be treated as positive. 3 star rating is considered as neutral whereas 1 star and 2 star ratings are treated as negative. The dataset is heavily biased in terms of 5 star (106,265) and 4 star (33,053) reviews. For 3 star (16,705), 2 star (11,261) and 1 star (15,100) limited reviews are available. So, for Data preparation for the modeling phase, we will use data balancing techniques to ensure that all the categories have equal number of reviews available. This can help us in terms of balanced predictions.

- **Analysis & Methods:**

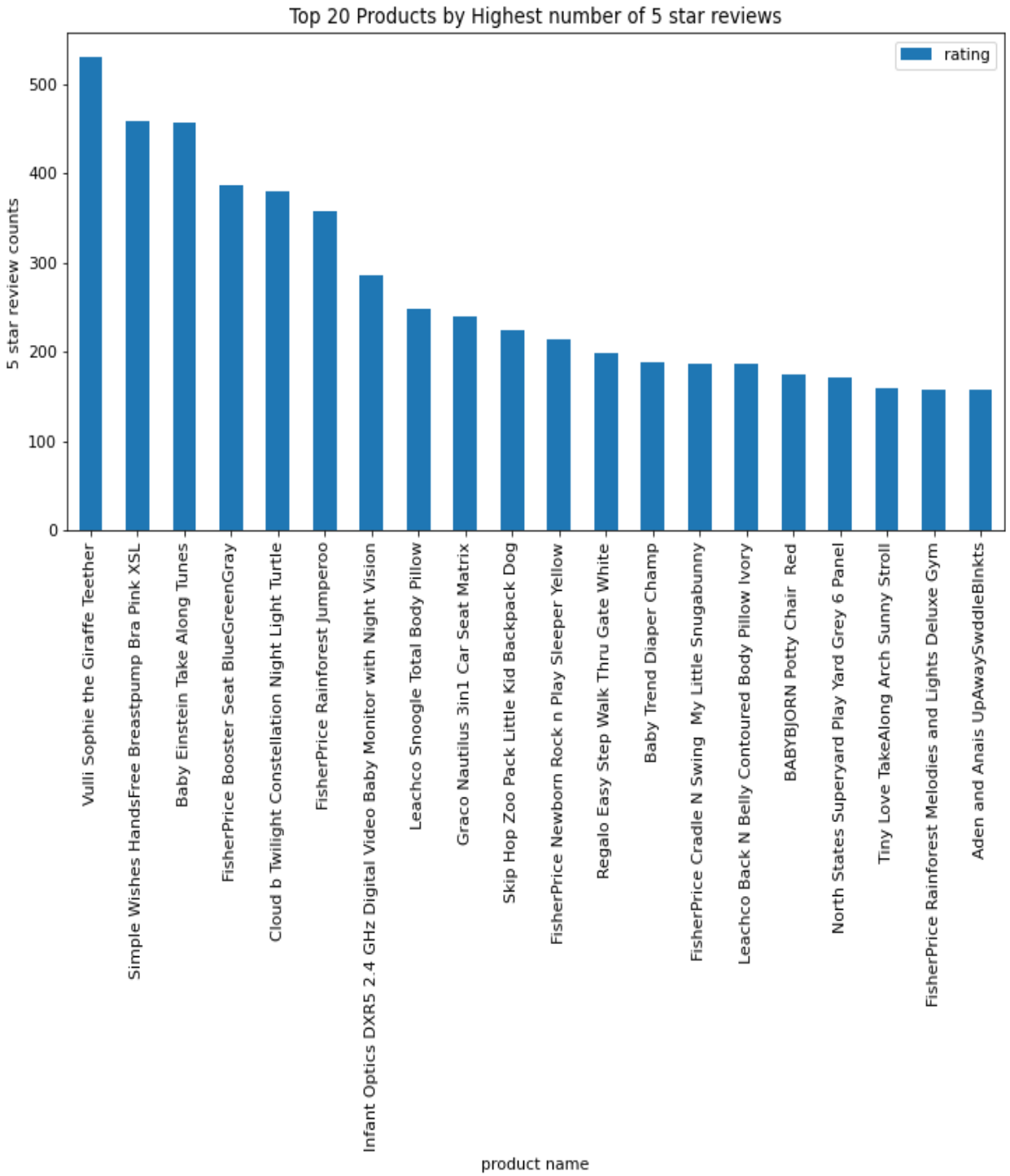
Below is the snapshot of the Top 20 products by the highest number of reviews. "Vulli Sophie the Giraffe Teether" is the top product with highest number of total review comments (779).



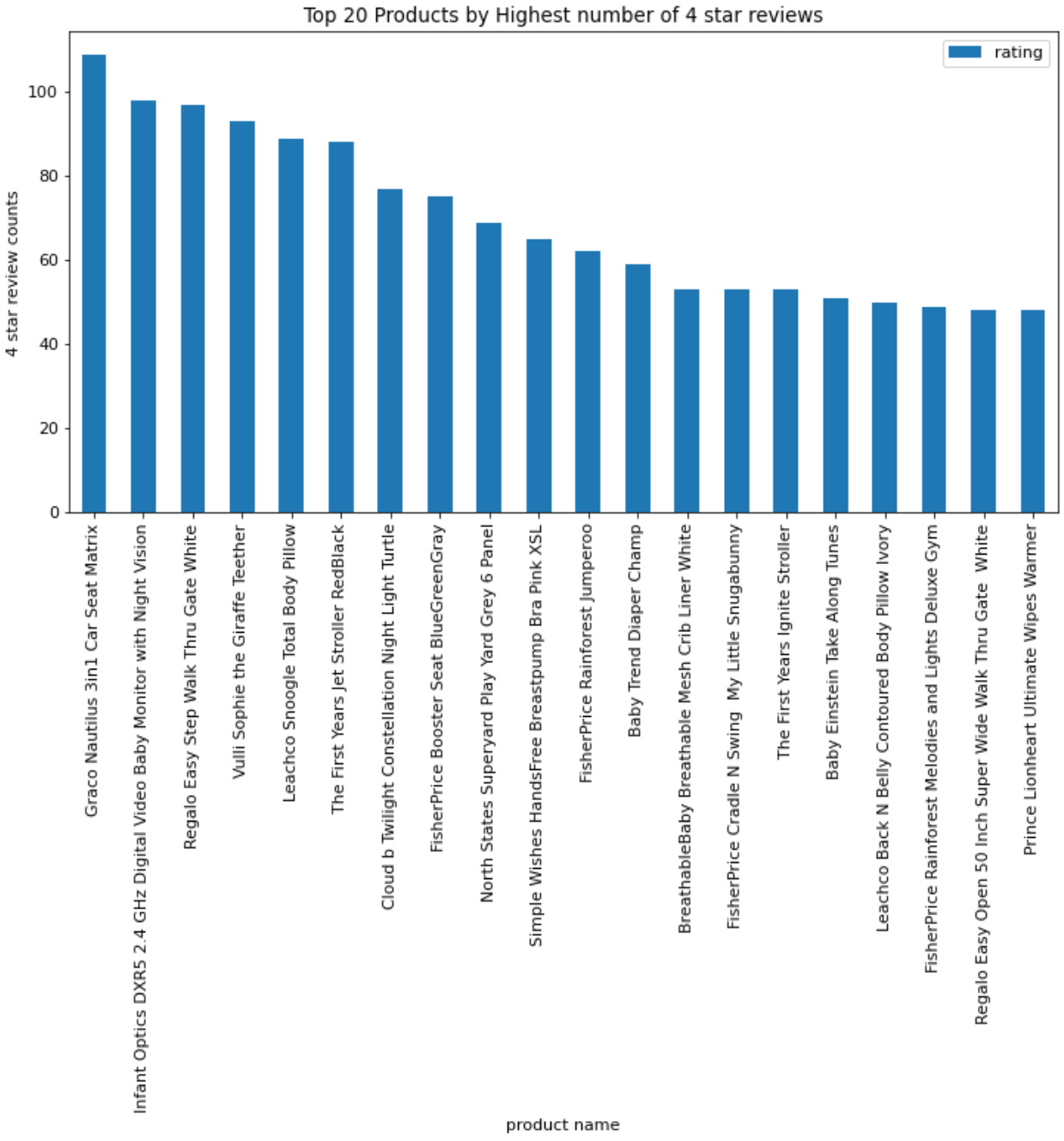
Below is the snapshot of distribution of reviews ratings by the rating from 1 star through 5 star ratings. 5 star ratings occupy the largest portion around 58% in the given reviews ratings.



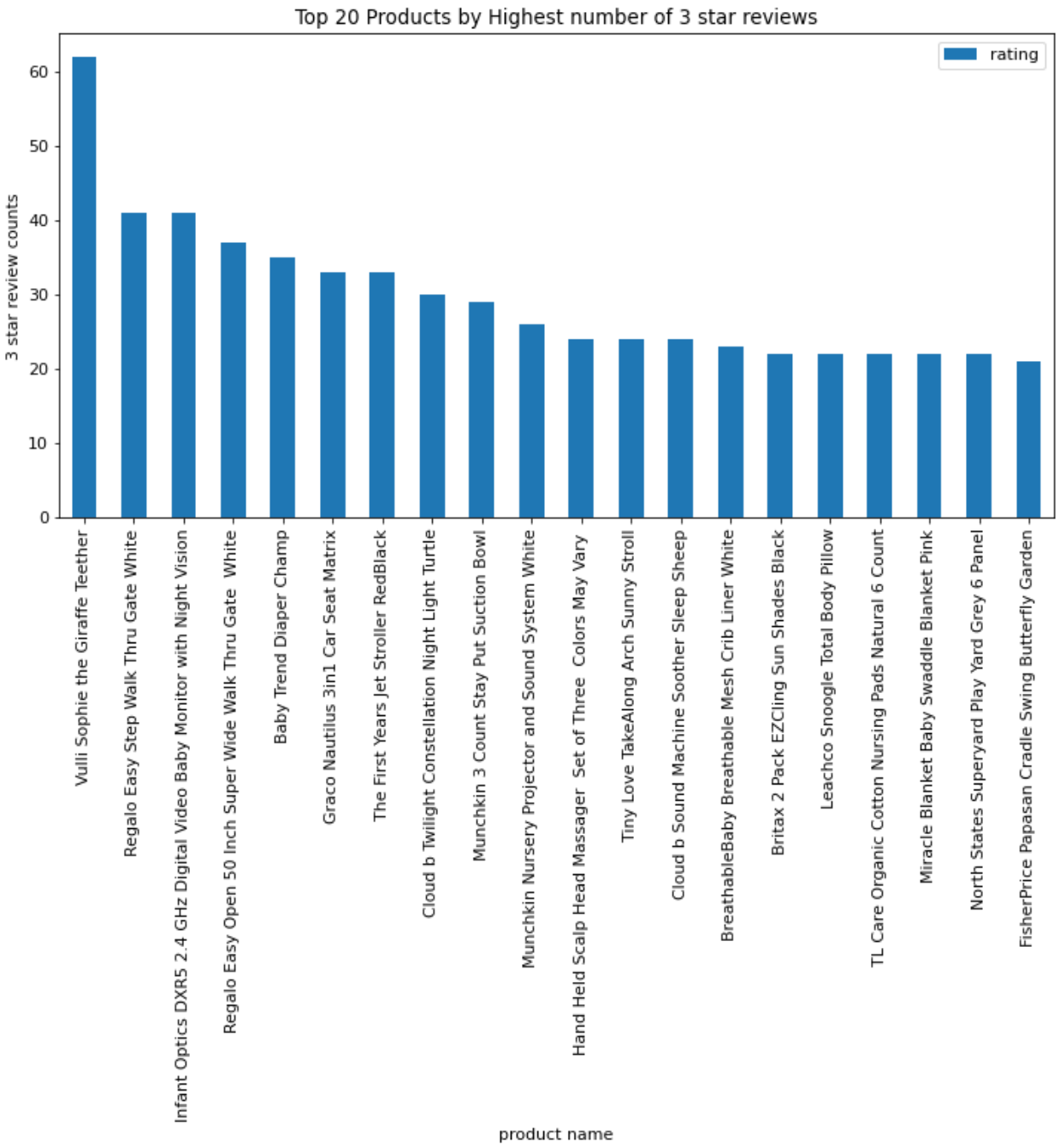
Snapshot of Top 20 products with highest number of 5 Star rating. As observed, "Vulli Sophie the Giraffe Teether" has highest number of 5 rating reviews.



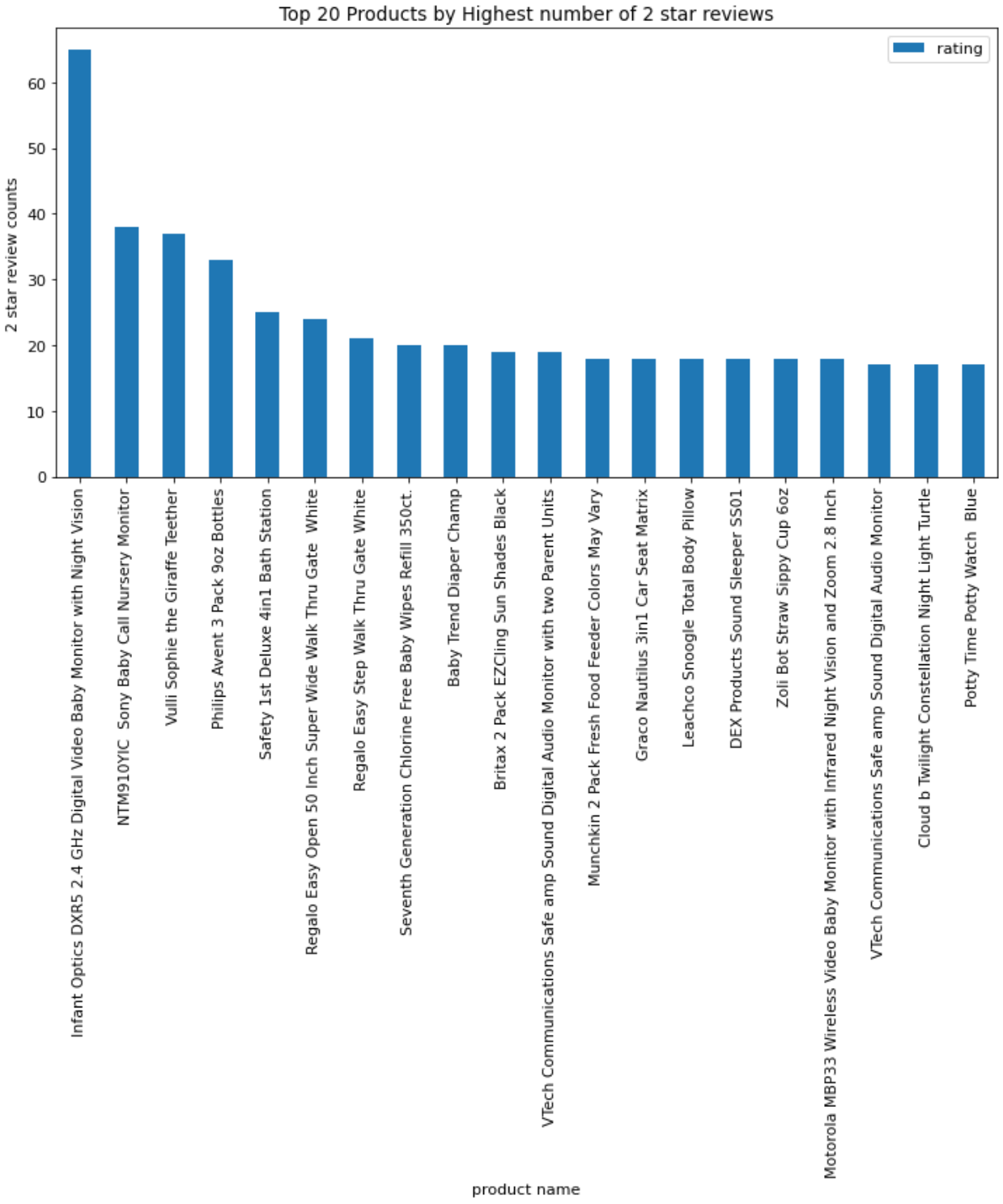
Snapshot of Top 20 products with highest number of 4 Star rating. As observed, "Graco Nautilus 3in1 Car Seat Matrix" has highest number of 4 rating reviews.



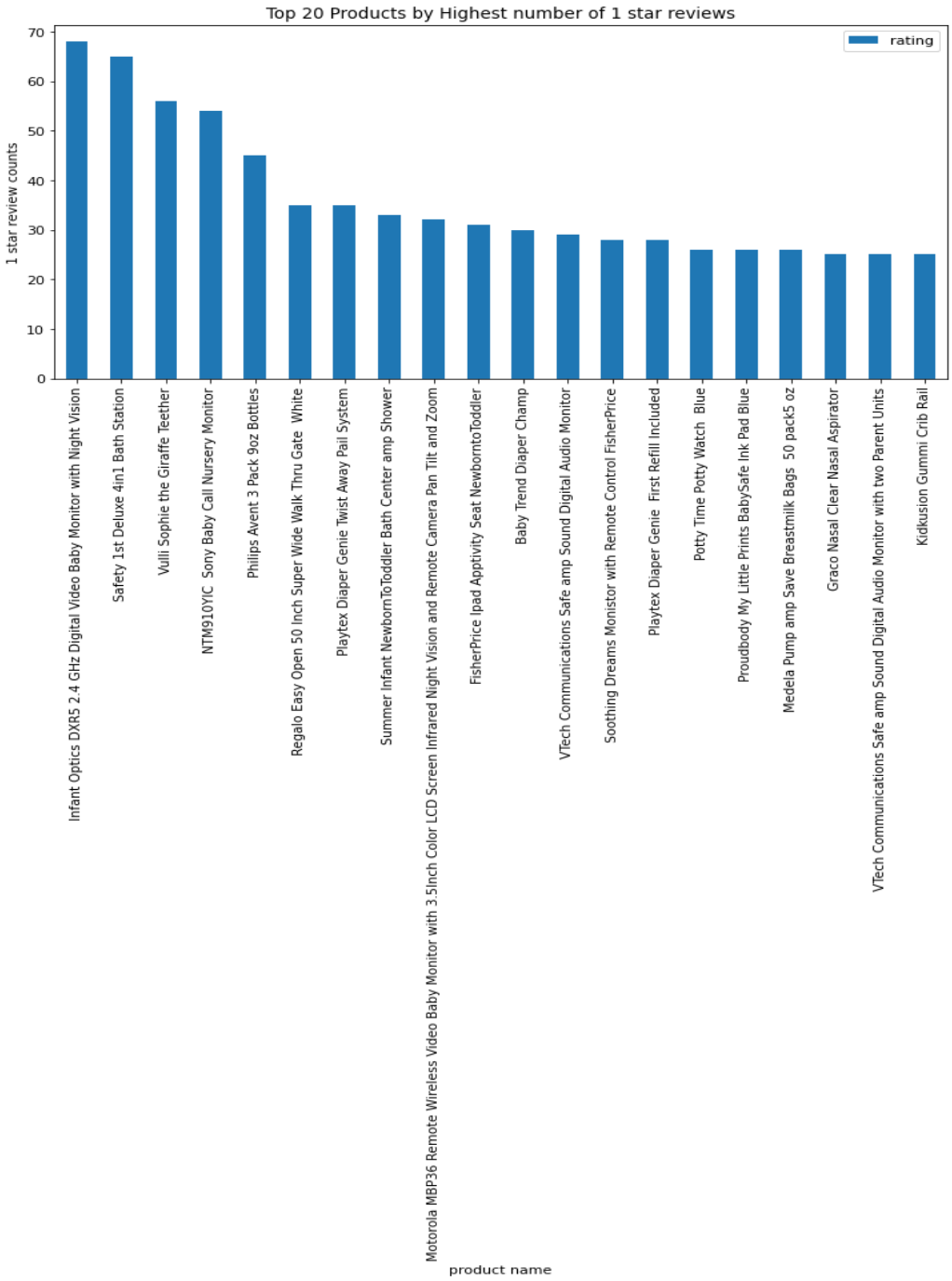
Snapshot of Top 20 products with highest number of 3 Star rating. As observed, "Vulli Sophie the Giraffe Teether" has highest number of 3 star rating.



Snapshot of Top 20 products with highest number of 2 Star rating. As observed, "Infant Optics DXRS 2.4 GHz Digital Video Baby Monitor with Night Vision" has highest number of 2 star rating. But the same product also has good number of 5 star and 4 star ratings as seen previously.



Snapshot of Top 20 products with highest number of 1 Star rating. As observed, "Infant Optics DXRS 2.4 GHz Digital Video Baby Monitor with Night Vision" has highest number of 2 star rating. But the same product also has good number of 5 star and 4 star ratings as seen previously.



Modeling and Evaluation:

Using Tf-idf vectorizer:

MLP Classifier model:

	precision	recall	f1-score	support
0	0.72	0.71	0.71	5499
1	0.69	0.77	0.73	5640
2	0.78	0.71	0.75	5753
accuracy			0.73	16892
macro avg	0.73	0.73	0.73	16892
weighted avg	0.73	0.73	0.73	16892

sklearn MLP accuracy score using TfIdf vectorizer is: 72.89 %

Keras Sequential Neural Network:

	precision	recall	f1-score	support
0	0.73	0.71	0.72	5499
1	0.68	0.78	0.73	5640
2	0.79	0.70	0.74	5753
accuracy			0.73	16892
macro avg	0.73	0.73	0.73	16892
weighted avg	0.73	0.73	0.73	16892

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

Using Count vectorizer:

MLP Classifier model:

	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

sklearn MLP accuracy score using Count vectorizer is: 73.24 %

Keras Sequential Neural Network:

	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 %

Count Vectorizer Using n-grams (2-gram):

MLP Classifier model:

	precision	recall	f1-score	support
0	0.74	0.75	0.75	5499
1	0.75	0.73	0.74	5640
2	0.78	0.79	0.79	5753
accuracy			0.76	16892
macro avg	0.76	0.76	0.76	16892
weighted avg	0.76	0.76	0.76	16892

sklearn MLP accuracy score using Count vectorizer is: 75.78 %

Keras Sequential Neural Network:

	precision	recall	f1-score	support
0	0.72	0.79	0.75	5499
1	0.81	0.71	0.76	5640
2	0.78	0.80	0.79	5753
accuracy			0.77	16892
macro avg	0.77	0.77	0.77	16892
weighted avg	0.77	0.77	0.77	16892

Keras neural network accuracy score using Count vectorizer is: 76.75 %

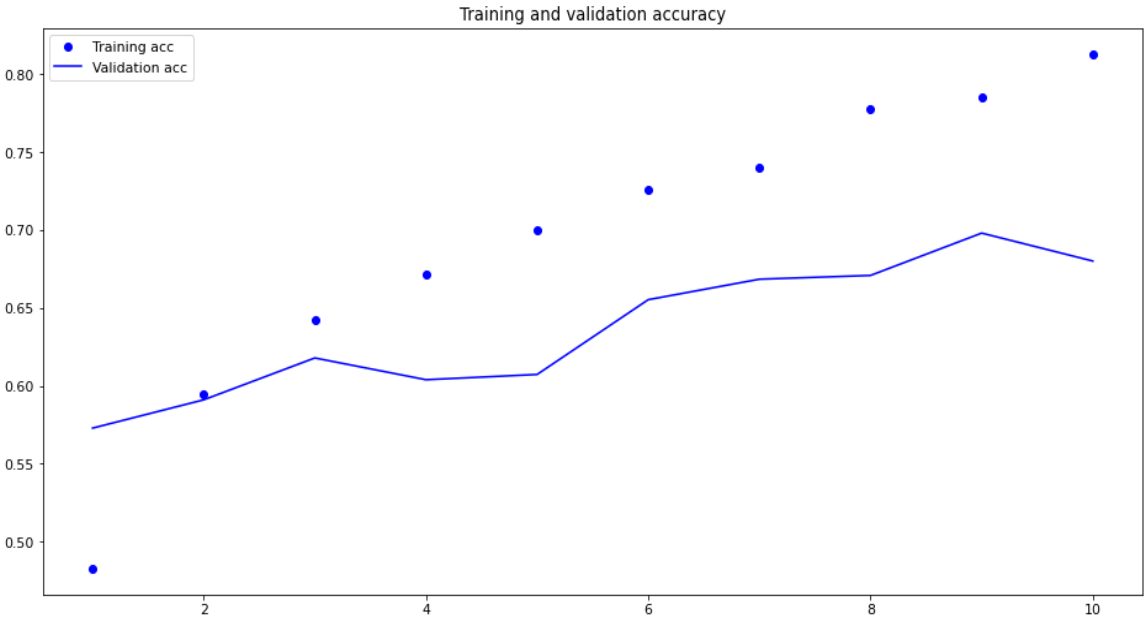
TextBlob sentiment analysis:

	precision	recall	f1-score	support
negative	0.62	0.33	0.43	22522
neutral	0.29	0.02	0.03	22522
positive	0.39	0.93	0.55	22522
accuracy			0.43	67566
macro avg	0.43	0.43	0.34	67566
weighted avg	0.43	0.43	0.34	67566

textblob accuracy score is: 42.75 %

Recurrent Neural Network (LSTM):

For the LSTM models, while the training accuracy goes on increasing up to 80%, the validations accuracy still hovers around 68% - 69% range. Also, LSTM models took longer time for the processing and execution as compared to Keras Sequential Artificial Neural Networks which had better accuracy above 76% using n-grams vectorizer.



MultiNomial Naïve Bayes (MultinomialNB):

```
Accuracy score of model with n-gram range of (1, 1): 0.64
Accuracy score of model with n-gram range of (1, 2): 0.72
Accuracy score of model with n-gram range of (1, 3): 0.72
Accuracy score of model with n-gram range of (1, 4): 0.72
```

- **Conclusion:**

Looking at the above analysis, we can notice that among all of the methods like MLP classifiers, Sequential Artificial Neural Network, Recurrent Neural networks and TextBlob, Sequential Neural networks using Count Vectorizer method worked better in terms of predictions.

- **Assumptions:**

Assuming that given data has all the ratings and corresponding reviews captured correctly against given set of products. If there are minor inconsistencies, then it could affect the prediction accuracy.

- **Limitations / Challenges:**

The reviews available have majority of the reviews (about 76.5%) marked as positive / favorable reviews (5 or 4 stars) as compared to neutral (9%) and negative categories (14.5%). So, these may need to cleaned up and balanced dataset needs to be prepared so that we can analyze and predict the future reviews / comments appropriately. Also, having larger number of negative and neutral categories

- **Future Uses/Additional Applications / Recommendations:**

The approach / models being created as a part of this opportunity can be partially utilized or concepts used for analyzing the Customer Sentiments for each of the individual products. The similar models can also be used for sentiment analysis for the sentiment analysis for other product categories or twitter sentiments about a specific topic or companies.

- **Implementation Plan:**

Since the outcome variable will be indicative of 3 sentiment categories – Text Classification models like sklearn Multi-Layer Perceptron (MLP) Classifier, Keras Sequential Neural Networks, TextBlob (sentiment analysis library) and Recurrent Neural Network (RNN) like LSTM will be used for the final predictions. We can train these models and then generate the final predictions. The predictions can be validated using classification reports to determine the model accuracy.

- **Ethical Assessment:**

Ensuring the privacy of the customers whose data is being utilized for analysis is of utmost importance. Hence, we will need to ensure that the dataset does not contain any personally identifiable information like names from the review comments. The current dataset does not contain such personal information and has review texts, product names and ratings.

- **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
- Using n-gram vectorizer: https://towardsdatascience.com/leveraging-n-grams-to-extract-context-from-text-bdc576b47049