Journal of Information Technology and Digital World (ISSN: 2582-418X)
www.irojournals.com/itdw/

Performance Analysis of Machine Learning Algorithms on Product Reviews

Iwin Thanakumar Joseph.S

Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vijayawada, Andhrapradesh, India

E-mail: iwineee2006@gmail.com

Abstract

Digital reviews now have a significant impact on how consumers communicate globally and the effect of review on online purchases. Sentiment analysis is one of most common field for analysing and extracting information from data of text format from multiple platforms like Facebook, Amazon, Twitter, and others. This research aims to discuss and analyse the sentiments stated in Amazon product reviews using Support Vector Machine (SVM), Naive Bayes, Random Forest, and Decision Tree algorithms. The machine learning algorithms classify reviews as positive or negative. The performance of the algorithms is evaluated in terms of metrics such as accuracy, precision, recall and F score. As per the validated results, SVM performs better on accuracy rate compared to the other algorithms.

Keywords: Machine learning, Product review, Support Vector Machine, Performance Metrics

1. Introduction

In today's digital era, the retail market sector has expanded to include the purchase of products online as well as the ability for customers to offer their comments, suggestions, and recommendations. Customers of online shopping websites frequently have the option to give things a numerical rating as well as textual input regarding a purchased item. Machine learning is an area of logic that looks into the construction as well as the analysis of algorithms which benefit from data. Such calculations develop a model based on the sources of data and to make predictions or decisions. Machine learning allows for the application of many algorithms for a variety of tasks, including simple linear regression for prediction issues.

DOI: https://doi.org/10.36548/jitdw.2023.1.005
Received: 15-03-2023. received in revised form: 15-04-2023, accepted: 30-04-2023, published: 11-05-2023

Machine learning methods can be categorized into supervised, unsupervised, semi-supervised, and reinforcement learning categories technically. Supervised learning: For training machine learning models, supervised learning makes use of labelled data. The result in labelled data is previously known. The inputs and outputs of the model only need to be mapped. Unsupervised Learning: Unsupervised learning trains computers using unlabelled data. Data without labels lacks a constant output variable. The model gains knowledge from the data, identifies its patterns and features, and then outputs the results. Reinforcement Learning: Reinforcement Learning teaches a machine to respond appropriately and optimize its gains in a certain circumstance. It generates actions and rewards to use an agent as well as an environment. The agent seems to have a beginning and an ending. This learning strategy does not have a fixed goal variable.

Sentiment analysis refers to the process of computationally recognizing and classifying the sentiment of customers expressed in a text. The below figure shows the example for the positive and negative comments on product review.

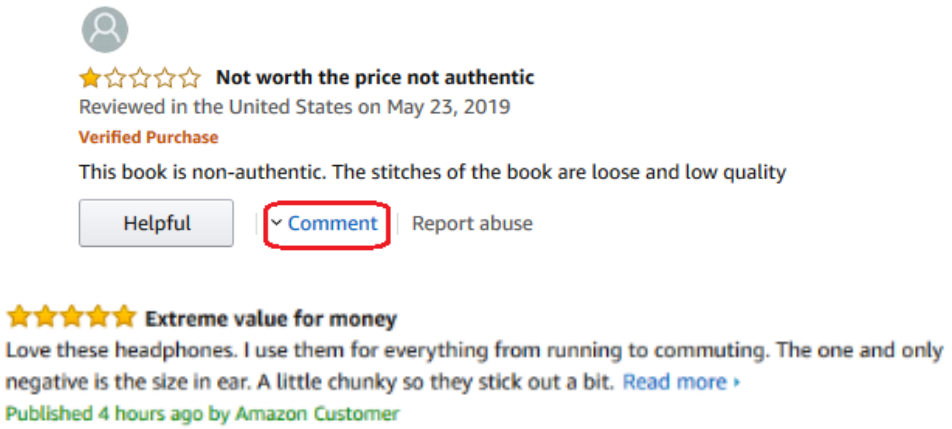

Figure 1a and 1b. Positive and negative review by the customer (https://www.sellerapp.com/blog/how-to-respond-to-negative-reviews/)

The below are the few methods for classifying sentiment analysis.

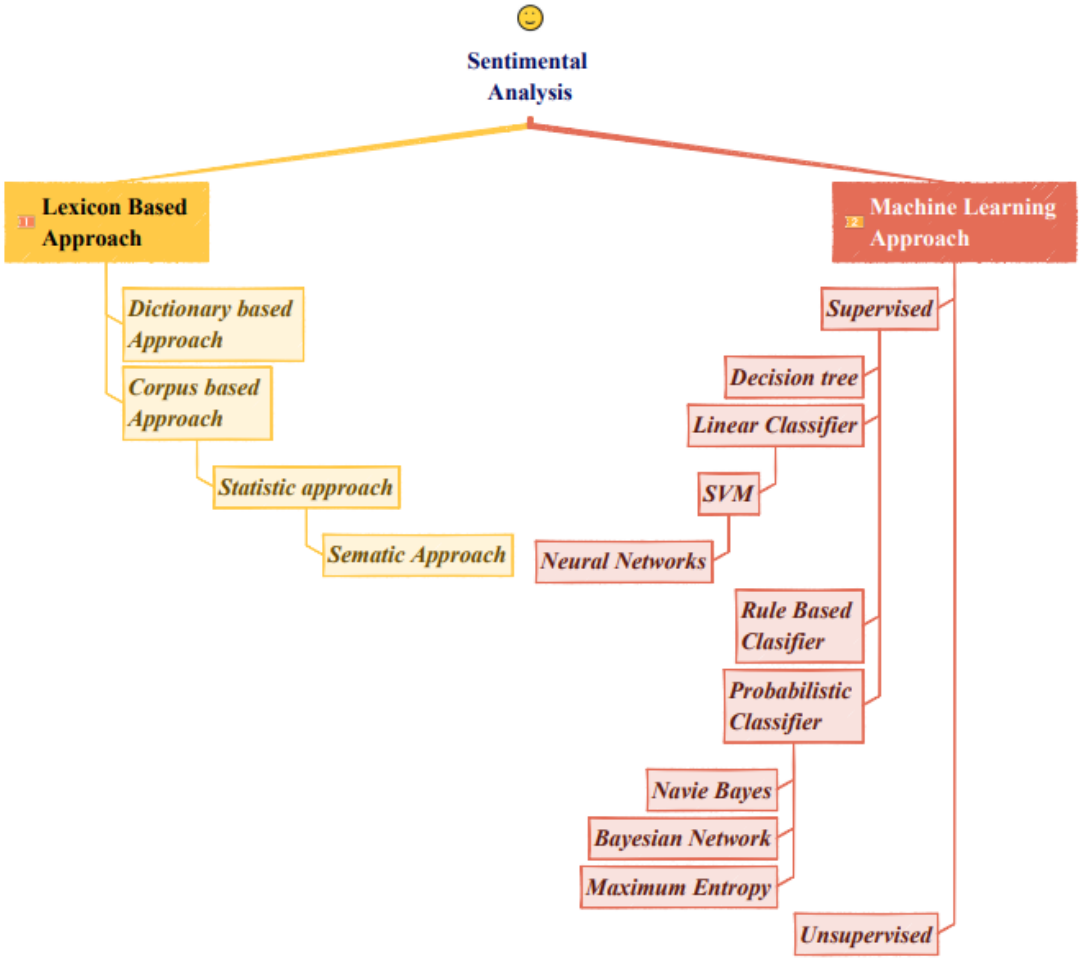

Figure 2. Different Methods of Sentimental Analysis

The purpose of this research is to classify customer reviews of various products into positive and negative feedback, and to develop a supervised learning approach to polarise a huge number of reviews. About 80% of online customers trusted reviews more than personal recommendations. Decision Tree, Naive bayes, Support Vector Machine, and Random Forest approaches are used for product analysis. These algorithms are evaluated for sentiment polarity, feature reduction, online site source, products categories, as well as set of review, completely in English, before being compared to one another. The focus of this work is on researching the sentiment classification of Amazon product online reviews [1][2]. Moreover, this research compares the model performance in terms of accuracy as well as other quality metrics.

2. Literature review

Jagdale et al.,[3] proposed a product review classification system, in which dataset was obtained from Amazon and reviews of cameras, laptops, phones, tablets, and other electronic devices used for video surveillance were included. Machine learning methods were used after pre-processing to categorise reviews as good or negative. This study concludes that classification of product reviews with the help of machine learning models yields better results. For camera reviews, accuracy of 98.17% was achieved using Naive Bayes, while Support Vector Machine achieves 93.54% accuracy score.

M. A. Shafin,et al .,[4] proposed using Natural Language Processing (NLP) to examine online customer reviews and produce a ratio of the positive as well as negative remarks made by prior customers. To perform the research, more than a thousand reviews and comments about the product were considered. Sentiment analysis was performed with a number of classification methods, including K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Decision Tree, Random Forest, and Logistic Regression. SVM outperformed the competition with highest accuracy of 88.81%.

B. K. Bhavitha et al.,[5] presented various machine learning methods that were employed in sentiment analysis and opinion mining. Emotional analysis combined with machine learning may be helpful in forecasting customer feedback and attitudes about recently launched products. The study provided a thorough overview of numerous machine learning approaches before comparing the effectiveness, benefits, and drawbacks of each technique. By applying supervised machine learning approaches, an accuracy of 85% was obtained, which was higher than that achieved using unsupervised learning techniques.

S. Dey et al.,[6] formulated a research analysis on sentiment of Amazon product evaluations, which also gives a comparison of SVM with Naive Bayes classifiers. After pre-processing, the models were trained using about 2250 features and nearly 6000 datasets. The system gives the SVM classifier a precision at 82.853%, recall at 82.884%, and a f1 score of 82.662%. Also, the Naive Bayes classifies the data with an accuracy, recall, and f1 score of 83.990%, 83.997%, and 83.993%, respectively. As a result, the model provides SVM and Naive Bayes accuracy values of 84% and 82.875%, respectively, which is proven to be superior to the traditional approaches. The support vector machine can polarise Amazon product reviews more accurately, according to the experimental findings.

A. Poornima et al.,[7] evaluated how well various machine learning algorithms perform sentiment analysis on Twitter data. The suggested approach employed phrase frequency to determine the sentence's sentiment polarity. In order to compare their effectiveness, the sentence classification algorithms Multinomial Naive Bayes, SVM, and Logistic Regression were used. Based on the findings, it can be concluded that when used in conjunction with the n-gram and bigram models, logistic regression has the highest accuracy. By using a bigram model, the accuracy of the logistic regression was about 86%. For Twitter sentiment analysis, logistic regression outperforms than other supervised machine learning methods.

3. Proposed Method

The model is developed by employing the sentiments of Amazon.com reviews to analyse the product performance. To utilise the dataset in a supervised learning model, it must be labelled. The method used to classify the Amazon datasets, goes through pre-processing, stage of extraction and classification, and concludes with the evaluation phase. The below figure shown the process flow of the proposed algorithm.

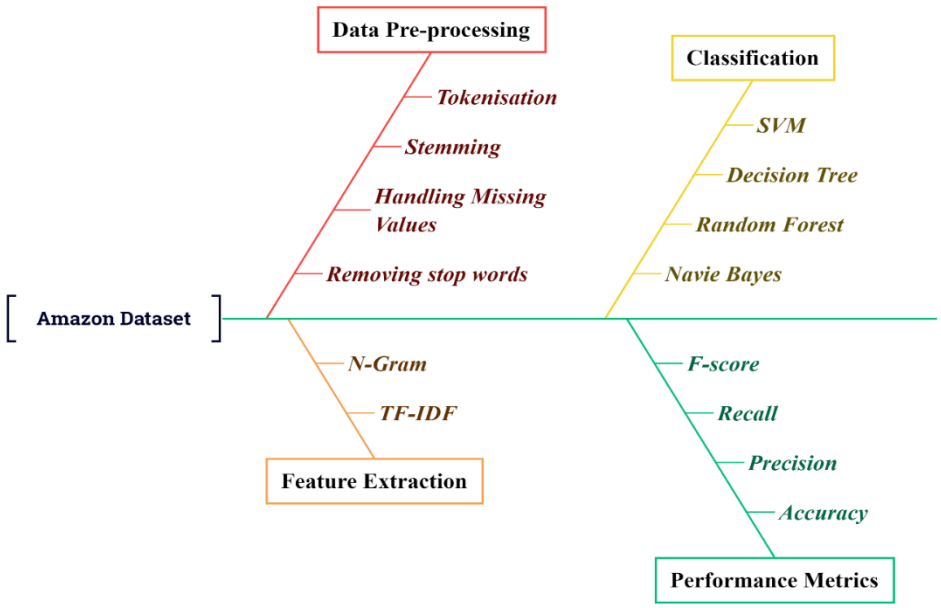

Figure 3. Proposed workflow

Dataset collection

The database consists of more than 34,000 product reviews for items sold by Amazon, such as the Fire TV Stick, Kindle, and more [8]. The suggested dataset has several columns with null values. The collected dataset is divided into a ratio of 80:20 for training and testing.

Data pre-processing

In sentiment analysis based on product review, pre-processing is essential. There are several subcategories of text pre-processing techniques, including tokenization, stemming, removing punctuation marks, and removing stop words. The proposed dataset consists of null values which are handled by this section.

- Tokenisation: Tokenisation is defined as a technique of dividing a text document to tokens, which are mostly small units of text. Words, chunks of word, as well as characters like punctuation are considered as tokens.
- Stemming: Stemming is a method for eliminating affixes from words to reveal their basic structure.
- Handling missing values: The dataset has been examined by the system to discover the missing value. Subsequently, missing values were substituted with the appropriate constant to fill in the blanks.
- Removal of stop words: Stop words are terms that are not necessary for text mining to divide a sentence into any certain category. These terms are often avoided in order to increase the model accuracy. These are a part of text cleaning [9].

Feature Extraction

One of the most difficult tasks is feature extraction since it requires for applying Natural Language Processing methods to recognize features. There are several methods for extracting features, including word embedding, N-gram, TF-IDF, Bag of Words (BOW), and NLP - supported features such as word count and noun count [10]. In this work, the effects of two characteristics are examined. Sentiment analysis using TF-IDF and N-gram.

Term Frequency – Inverse Document Frequency (TF-IDF) It is one of the most frequently used for information retrieval and is the representation of the significance of a word or phrase inside a given text. The frequency of a word within the corpus is usually used to balance the TF-IDF value, which helps to account that few words are used more frequently than overall words. The phrase "term frequency" refers to how frequently a particular term (t) appears in a text (doc) in relation to the overall number of words in the document. Each word is given its own vector as it transforms raw strings or datasets into vectors. Cosine similarity is used for retrieving features. Finally, it gives the numerical values of the full document.

N-gram: The continuous collection of ‘n’ elements from dataset of text is known as a "N-gram." Depending on the application, the objects might be characters, words, or base pairs. The

N-grams are usually gathered from a corpus of text. An effective N-gram model is capable of predicting the value of p(w|h), or the next phrase in the sentence.

Classification:

Support vector machine: The supervised (feed-me) machine learning method known as SVM may be applied to classification or regression problems. Regression predicts a continuous value, whereas classification predicts a label or group. Identifying the hyper-plane that differentiates the classes is how SVM performs classification.

Naive Bayes: The Naive Bayes classifier utilizes Bayes theorem which predicts membership probabilities of each class, including likelihood that certain record and data point corresponds to that class. The class with the highest probability is referred to as the most likely class.

Decision tree: Recursively, as the tree is constructed, decision tree algorithms divide the training set to smaller subgroups. A supported decision tree is an instructional technique in which the secondary tree reviews for the errors of the primary tree, and third tree corrects for the errors in both the initial and second trees, etc. Every leaf of a tree is labelled with a class. Predictions rely on the entire cluster of trees coming together to form the expectation.

Random Forest: A collecting learning algorithm for classification and regression is called Random Forest. Based on the combined judgement of those trees, Random Forest finally generates a huge number of decision trees.

4. Results and Discussion

Amazon online shopping product review remarks are collected from Kaggle. The dataset description, including the quantity of products, product ID, name, its brand review date etc., are shown in below figure.

id	name	asins	brand	categorie	keys	manufac	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.(	reviews.username
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	This proc	Kindle				Adapter
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	great for	very fast				truman
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	Inexpens	Beginner tablet for our 9 ye				DaveZ
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	4	http://re	I've had	Good!!!				Shacks
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	I bought	Fantastic Tablet for kids				explore42
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	This ama	Just what we expected				tklit
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	4	http://re	Great for	great e-reader tablet				Droi
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-1	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	I gave thi	Great for gifts				Kacy
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	Great as	Great for reading				Weebee
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	I love orc	Great and lightweight reade				RoboBob
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	4	http://re	Not easy	nice tablet for the price				tld2
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	Excellent	Excellent product				ralexander422
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	4	http://re	Wanted i	Great Value				RegE
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	Simply di	Excellect				TommyL
AVqkIhw	All-New	B01AHB9	Amazon	Electroni	84166710	Amazon	2017-01-2	2017-07-(	2017-06-07T09:04:0		TRUE		0	5	http://re	Got it as	Living It				Teacheeor

Figure 4. Sample dataset [8]

According to the word frequency, the comments are divided into three categories: positive, negative, and neutral. The reviews are pre-processed by removing stop words, stemming, identifying terms with high frequency, and removing punctuation. The review comments are classified using various Machine Learning (ML) algorithms. Precision, recall, F-score, and accuracy are the performance-related metrics. The validated results are shown in below in table 1.

Table 1. Performance metrics for Amazon Product review

ML Models	***Accuracy***	***Precision***	***Recall***	***F -score***
SVM	96.7	95.1	88.2	86.4
Naive Bayes	95.1	93.2	84.9	82.6
Decision tree	94.2	93.9	83.1	80.9
Random Forest	93.5	92.4	80.1	78.6

The performance graph is shown below:

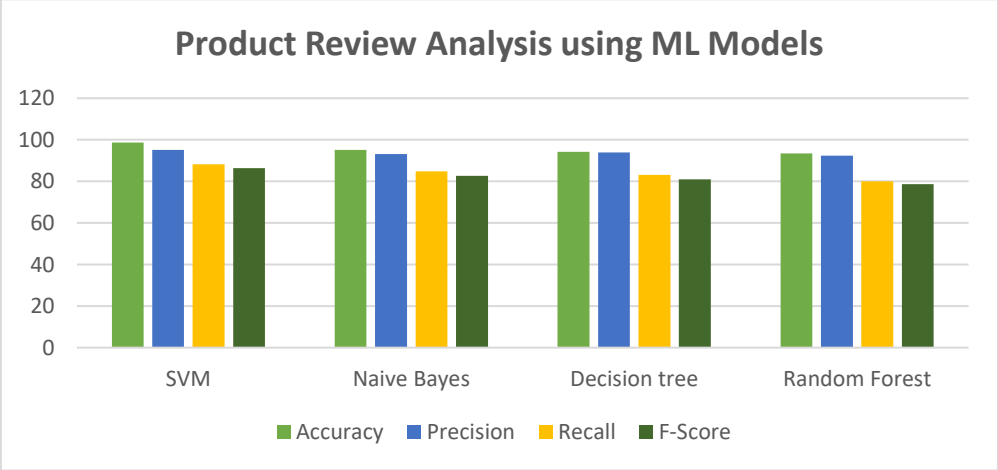

Figure 5. Performance Analysis of Various ML models

Based on the analysis of the collected datasets of customer reviews, SVM algorithm performs better with the accuracy rate of 96.7% and a recall score of 88.2% than the rest of the three algorithms.

5. Conclusion

Reviews are the valuable source of information where customers analyse before buying a product. This study examines how different machine learning algorithms analyse reviews for product evaluations using data from Amazon. The data are pre-processed and are examined using Machine Learning algorithms such as Support Vector Macine, Naive Bayes, Decision Tree, and Random Forest. Finally, the classification algorithm that achieves great accuracy and found to be better than the others, is the Support Vector Machine algorithm.

References:

[1] https://www.simplilearn.com/tutorials/machine-learning-tutorial/types-of-machine-learning

[2] https://www.sellerapp.com/blog/how-to-respond-to-negative-reviews/

[3] Jagdale, R.S., Shirsat, V.S., Deshmukh, S.N. (2019). Sentiment Analysis on Product Reviews Using Machine Learning Techniques. In: Mallick, P., Balas, V., Bhoi, A., Zobaa, A. (eds) Cognitive Informatics and Soft Computing. Advances in Intelligent Systems and Computing, vol 768. Springer, Singapore. https://doi.org/10.1007/978-981-13-0617-4_61

[4] M. A. Shafin, M. M. Hasan, M. R. Alam, M. A. Mithu, A. U. Nur and M. O. Faruk, "Product Review Sentiment Analysis by Using NLP and Machine Learning in Bangla Language," *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, DHAKA, Bangladesh, 2020, pp. 1-5, doi: 10.1109/ICCIT51783.2020.9392733.

[5] B. K. Bhavitha, A. P. Rodrigues and N. N. Chiplunkar, "Comparative study of machine learning techniques in sentimental analysis," *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India, 2017, pp. 216-221, doi: 10.1109/ICICCT.2017.7975191.

[6] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar and M. Dey, "A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews," *2020 International Conference on Contemporary Computing and Applications (IC3A)*, Lucknow, India, 2020, pp. 217-220, doi: 10.1109/IC3A48958.2020.233300.

[7] A.Poornima and K. S. Priya, "A Comparative Sentiment Analysis Of Sentence Embedding Using Machine Learning Techniques," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2020, pp. 493-496, doi: 10.1109/ICACCS48705.2020.9074312.

[8] https://www.kaggle.com/datasets/datafiniti/consumer-reviews-of-amazon-products

[9] https://michael-fuchs-python.netlify.app/2021/05/31/nlp-text-pre-processing-iii-pos-ner-and-normalization/#tokenization

[10] https://www.geeksforgeeks.org/sklearn-feature-extraction-with-tf-idf/