SENTIMENT ANALYSIS ON AMAZON KINDLE STORE REVIEWS

Shryavani K[1],Abdul Razak M S[2],Dr. Roopa G M[3]
[P1] Assistant Professor, Department of Computer Science and Engineering, BIET, Davanagere.
[2] Assistant Professor, Department of Computer Science and Engineering, BIET, Davanagere.
[3] Associate Professor, Department of Computer Science and Engineering, BIET, Davanagere.
Email: shravani268@gmail.com[1],msabdulrazak@gmail.com[2], roopa.rgm@gmail.com[3]

Abstract—the world we see nowadays is becoming more digitized. In this digitalized world e-commerce is taking the ascendancy by making products available within the reach of customers where the customer doesn't have to go out of their house. Nowadays people are relying ononline products so the importance ofa reviewis going higher. For selecting a product, a customer needs to go through thousands of reviews to understand a product. But in this prosperous day of deep learning, going through thousands of reviews would be much easier if a model is used to polarize those reviews and learn from it. We used deep learning techniques on a large scale amazon dataset to polarize it and get satisfactory accuracy.

Keywords—e-commerce, deep learning, feature extraction, text classification, sentiment analysis

I. INTRODUCTION

As the commercial sites of the world are almost fully online platforms, people are trading products through different e-commerce websites. And for that reason reviewing products before buying is also a common scenario. Also nowadays, customers are more inclined towards the reviews to buy a product. So analyzing the data from those customer reviews to make the data more dynamic is an essential field nowadays. In this age of increasing machine learning and deep learning based algorithms, reading thousands of reviews to understand a product is rather time consuming where we can polarize a review on a particular category to understand its popularity among the buyers all over.

The objective of this paper is to categorize the positive and negative feedback of the customers over different products and build a supervised learning model to polarize large amounts of reviews. A study on amazon last year revealed more than 80% of online shoppers trust reviews as much as personal recommendations. Any online item with a large amount of positive reviews provides a powerful comment of the legitimacy of the item. Conversely, books, or any other online item, without reviews puts potential prospects in a state of distrust. Quite simply, more reviews look more convincing. People value the consent and experience of others and the review on a material is the only way to understand others' impression on the product. Opinions, collected from users' experiences regarding specific products or topics, straightforwardly influence future customer purchase decisions. Similarly, negative reviews often cause sales loss. For those understanding the feedback of customers and polarizing accordingly over a large amount of data is the goal. There are some similar works done over amazon dataset. In opinion mining over a small set of dataset of Amazon kindle product reviews to understand the polarized attitudes towards the product.

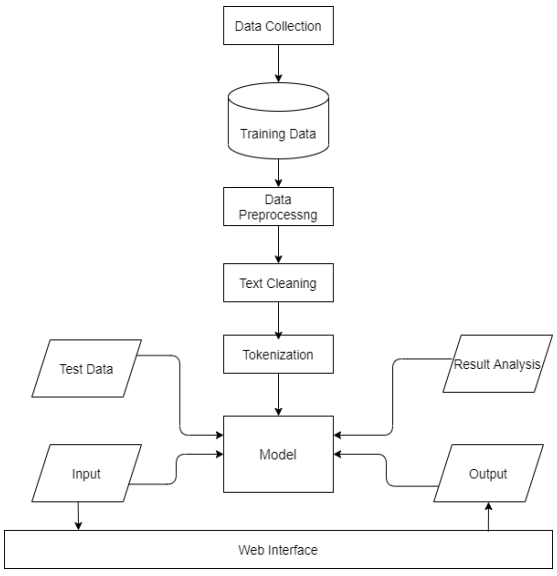
2. RELATED WORKS

Over a while, there are plenty of research papers related to previews analysis, sentiment analysis or opinion mining. Among such papers is the [5]Xu Yun el al from Stanford University applied existing supervised learning algorithms such as perceptron algorithm, naive bayes and supporting vector machine to predict a review's rating on Yelp's rating dataset. They used hold out cross validation using 70% data as the

training data and 30% data as the testing data. The author used different classifiers to determine the precision and recall values. In paper, [2] Maria Soledad Elli and Yi-Fan extracted sentiment from the reviews and analyzed the result to build up 1 a business model. They claimed that this tool gave them pretty high accuracy. They mainly used Multinomial Naive Bayesian(MNB) and support vector machines as the main classifiers. Callen Rain[3] proposed extending the current work in the field of natural language processing. Naive Bayesian and decision list classifiers were used to classify a given review as positive or negative.

Deep-learning neural networks is also popular in the area of sentiment analysis. Ronan Collobert[1] et al used a convolutional network for the semantic role labeling task with the goal of avoiding excessive task-specific feature engineering. On the other hand, in paper[4], the authors proposed using recursive neural networks to achieve a better understanding of compositionality in tasks such as sentiment detection.

In this paper, the aim is to apply natural language processing in order to train the model



on a dataset containing Amazon kindle store reviews. The trained model is tested on unseen new reviews. The novelty of the model is to analyse the test data (reviews) and classify them into three main categories: positive, neutral and negative based on the sentiment hidden in them.

III. METHODOLOGY

Amazon is one of the largest E-commerce sites as for that there are innumerous reviews that can be seen. We used data named Amazon Kindle Reviews which was provided in Kaggle. The dataset is a .csv file consisting of labeled data having the following columns -

"reviewerID": D of the reviewer
"asin": ID of the product
"reviewerName": name of the reviewer
"helpful": helpfulness rating of the review
"reviewText": text of the review
"overall": rating of the product
"summary": summary of the review
"reviewTime": time of the review (raw)
"unixreviewTime": unix timestamp

For the dataset we selected, it consists of more than 50,000 kindle book reviews. From the format used analyzing the review polarity we used review Text & Overall from it. We can see an overview of our methodology:

A. Data Collection

We acquired our dataset in .csv format that was already labeled from kaggle. As we have a large amount of reviews, manually analysing was quite impossible for us. Therefore we preprocessed our data and used elementary data analysis techniques to pre-process the datasets. As amazon reviews come in 5-star rating based generally 2 star ratings are either positive or negative. So we wrote a function that considers 5,4 and 3 ratings to be positive and 2,1 ratings to be negative reviews and proceed to the next step.

B. Pre-Processing

Tokenization: It is the process of separating a sequence of strings into individuals such as words, keywords, phrases, symbols and other elements known as tokens. Tokens can be individual words, phrases or even whole sentences. In the process of tokenization, some characters like punctuation marks are discarded. The tokens work as the input for different processes like parsing and text mining.

Stemming: Stemming is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers.

A stemming algorithm reduces the words “chocolates”, “chocolatey”, “choco” to the root word, “chocolate” and “retrieval”, “retrieved”, “retrieves” reduce to the stem “retrieve”. Stemming is an important part of the pipelining process in Natural language processing. The input to the stemmer is tokenized words.

Removing Stop Words: Stop words are those objects in a sentence which are not necessary in any sector in text mining. So we generally ignore these words to enhance the accuracy of the analysis. In different formats there are different stop words depending on the country, language etc. In English format there are several stop words.

C. Feature Extraction

Bag-of-Words: It is one of the most fundamental methods to transform tokens into a set of features. The BoW model is used in document classification, where each word is used as a feature for training the classifier.

For example, in a task of review based sentiment analysis, the presence of words like ‘fabulous’, ‘excellent’ indicates a positive review, while words like ‘annoying’, ‘poor’ point to a negative review .

There are 3 steps while creating a BoW model :

1. Text pre-processing
2. Creating the vocabulary
3. Creating matrix of features- Text Vectorisation

D. CNN Model

We use a Convolutional Neural Network (CNN) as they have proven to be successful at document classification problems. A conservative CNN configuration is used with 32 filters (parallel fields for processing words) and a kernel size of 8 with a rectified linear (‘relu’) activation function. This is followed by a pooling layer that reduces the output of the convolutional layer by half.

Next, the 2D output from the CNN part of the model is flattened to one long 2D vector to represent the ‘features’ extracted by the CNN. The back-end of the model is a standard Multilayer Perceptron layer to interpret the CNN features. The output layer uses a sigmoid activation function to output a value between 0 and 1 for the negative and positive sentiment in the review.

E. Train and Test Data

We are pretending that we are developing a system that can predict the sentiment of a textual book review as either positive or negative.This means that after the model is developed, we will need to make predictions on new textual reviews. This will require all of the same data preparation to be performed on those new reviews as is performed on the training data for the model.

We will ensure that this constraint is built into the evaluation of our models by splitting the training and test datasets prior to any data preparation. This means that any knowledge in the data in the test set that could help us better prepare the data (e.g. the words used) are unavailable in the preparation of data used for training the model.

That being said, we will use 80% train, 20% as a test of the data.

F. Evaluation Model

Evaluation is an integral part of the model development process.

It helps to find the best model that represents the data and how well the chosen model will work in the future.The output is predicted by analyzing the test data as input along with test data output and then the output is displayed.

Accuracy: Accuracy predicts how often the classifier makes the correct prediction. Accuracy is the ratio between the number of correct predictions and the total number of predictions.

G. Interface

A web interface is built to take input and display an output.

Flask web framework is used to build a web interface and other libraries are used to integrate the model.

IV. RESULTS

Fig : Home Page

The user will enter the command which he wants to analyze and if user click on the Kindle eBooks button then it redirect to the screen shown in below fig:

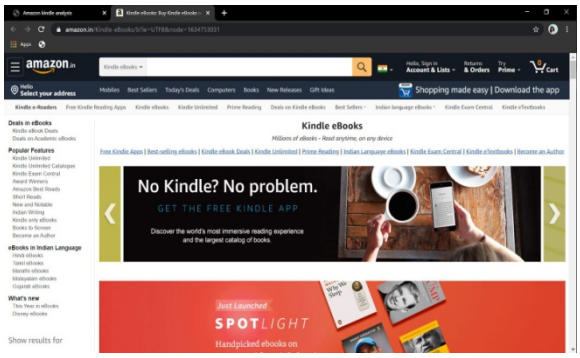


Fig : Kindle ebooks Page

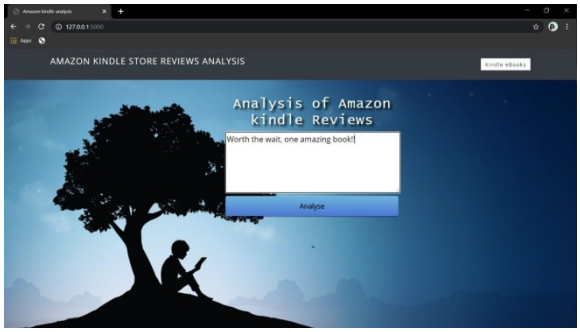


Fig : Review Analysis Page

The user enters the review in the review box and presses the analyze button to know the sentiment of the review.

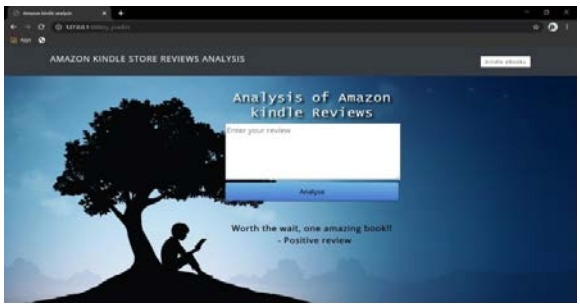


Fig : Output Page

The user gets the output as the **Positive review**.

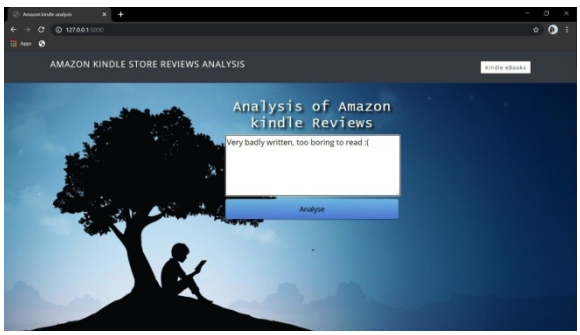


Fig : Review Analysis Page

The user enters the review in the review box and presses the analyze button to know the sentiment of the review.

Fig : Output Page

The user gets the output as the **Negative review**.

V. CONCLUSION AND FUTURE WORKS

It is completely impossible to use only raw text as input for making predictions. Hence,we saw that the pre-processing step played a major role in the complete process of NLP. To get better results, accuracy and make the machine take all the text as tokens, pre-processing of data is to be done carefully looking at the type of contents present in it.The most important thing is to be able to extract the relevant features from the given source of data. This kind of data can often come as a good complementary source in order to extract more learning features and increase the predictive power of the models. And the user is able to predict that the given comment is positive or negative.

In future, the work can be extended to perform multi-class classification of reviews which will provide a delineated nature of review to the consumer, hence better judgment of the product. It can also be used to predict the rating of a product from the review. This will provide users with a reliable rating because sometimes the rating received by the product and the sentiment of the review do not provide justice to each other. The proposed extension of work will be very beneficial for the e-commerce industry as it will augment user satisfaction and trust.

REFERENCES

[1] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa. Natural language processing (almost) from scratch.

Journal of Machine Learning Research, 12(Aug):2493–2537, 2011.
[2] M. S. Elli and Y.-F. Wang. Amazon reviews, business analytics with sentiment analysis.
[3] C. Rain. Sentiment analysis in amazon reviews using probabilistic machine learning. Swarthmore College, 2013.
[4] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 conference on empirical methods in natural language processing, pages 1631–1642, 2013.
[5] Y. Xu, X. Wu, and Q. Wang. Sentiment analysis of yelps ratings based on text reviews, 2015.
[6] Bhatt, Aashutosh, et al. "Amazon Review Classification and Sentiment Analysis." International Journal of Computer Science and Information Technologies 6.6 (2015): 5107-5110. [7]Chen, Weikang, Chihhung Lin, and Yi-Shu Tai."Text-Based Rating Predictions on Amazon Health & Personal Care Product Review." (2015) [8]Shaikh, Tahura, and DeepaDeshpande. "Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews.",(2016)
[9]Nasr, Mona Mohamed, Essam Mohamed Shaaban, and Ahmed Mostafa Hafez. "Building Sentiment analysis Model using Graphlab." IJSER., [10]Text mining for yelp dataset challenge; Mingshan Wang; University of California San Diego, (2017)