A Lexicon based Unsupervised Model to Evaluate Product Ratings Vs Reviews

Mahapara G
REVA Academy of Corporate Excellence,
REVA University
Bengaluru, India
mahapara.ba05@reva.edu.in

Taiba N
REVA Academy of Corporate Excellence, REVA University
Bengaluru, India
taiba.ba05@reva.edu.in

Ramamani V
REVA Academy of Corporate Excellence, REVA University
Bengaluru, India
ramamani.ba05@reva.edu.in

Abstract — E-Commerce has emerged as the new paradigm for purchase. E-commerce shoppers would like to look at customer review as a reliable source of information. These e-commerce websites provide features for customer to write the product reviews and scores the product from 1 to 5 or it's commonly referred to as star rating.

But sometimes it is seen that there are inconsistencies between the star ratings and the reviews. Because of that it is necessary to validate the star rating versus the reviews. There may also be cases where the customer would have given the review without rating the product. In such cases, we would also like to predict the star rating for a given review.

In this work we would like to propose a framework based on text analytics using unsupervised sentiment analysis [4][16] and KNN based regression [5] to provide the star rating for the reviews and validate the ratings with respect to the reviews. We have taken 2000 product reviews from Amazon and applied this methodology. The results found was encouraging and we would like to apply other techniques to validate and do deeper analysis to establish the actual star rating

Keywords — Sentiment analysis; Natural Language Processing; Text Mining; Corpus; Lexicon; Unsupervised method

I. INTRODUCTION

Today, almost every web page in any e-commerce portal contains a section for the users to give their comments about products or services, and share it with friends on social media sites like Facebook, Twitter or Pinterest [1]. Mining these opinions provide information for understanding collective human behavior. Specifically, what they perceive as the product/service from the provider and what the rating is. Generally, the perception is given by review comments and the rating is done by a star rating system. Since the review comments and start rating are subjective, there is a possibility for inconsistency in the perception. Hence it becomes necessary to validate the star rating versus the reviews. In addition, sometimes only a review is provided without a start rating, which may require prediction of star rating from these reviews. In this work, an attempt has been made to develop a methodology for predicting the start rating based on the reviews.

II. LITERATURE SURVEY

Purchase decisions can be better when opinions/reviews about products are and analyzed and used. Similarly, reviewing customer feedback help in revising the sales strategies to improve the sales that ultimately benefit the revenue of the company. There have been various studies on product reviews that to analyze the customer reviews given on a product. Sentiment analysis can be performed using both machine learning and lexicon-based approaches [3][4][5].

Jagdale et.al [13] in their paper 'Sentiment Analysis on Product Reviews Using Machine Learning Techniques' use machine learning techniques to review the products to great effect. They applied machine learning algorithms to classify reviews that are positive or negative. This paper concludes that, Machine Learning Techniques gives best results to classify the Products Reviews.

Irel et.al [14] presented a framework to analyze online product reviews. A series of customer needs were identified using machine generated data. Huge amount of qualitative data were distilled into quantitative insights to get inferences on product features using this framework. This was useful to the product designers in making more informed decisions. This framework revealed good insights not only on online product reviews and design theory but also on the methodology and data analytics. Reviews from Amazon website were used to validate the effectiveness of the proposed framework. This framework demonstrated a good statistical approach for analyzing online product reviews[2].

Vyas et.al [15] in their paper ccomprehensively talk about the techniques used in sentiment analysis. They describe how both Machine learning and lexicon-based approaches can be effectively used for sentiment analysis. This paper also discusses the evaluation parameters for both techniques.

Star ratings gives the aggregated rating for the product. Generally, star ratings complement the textual sentiments given in the review comments. But, sometimes they may differ from the sentiment that is expressed in the review.

Shah et.al [10] proposed a very unique approach by performing abstract-level sentimental analysis. They used n-gram classification along with POS tagging for customer reviews. The classification obtained was used as entropy for machine learning algorithm. Their study leveraged this methodology with promising outcomes. They used two algorithms - MaxEnt model and Naïve Bayes classifier to improved the accuracy.

Prabakaran et.al [11] worked on a statistical credibility scoring mechanism that help in identifying spam reviews. Their study consisted of three components: duplicate reviews detection, detection of anomaly in review count and rating,

and incentivized reviews detection. Credibility of the product reviews were effectively identified using these three methodologies that complement each other without requiring significant computational resources.

Kumar et.al [12] in their paper used the lexicon-based approach to derive star rating for the reviews. They categorized the reviews into five subcategories, i.e., Excellent, Good, Neutral, Bad and Worst, instead of classifying them as positive, negative or neutral.

Amazon as a prime e-commerce website attracts product reviews and ratings on a vast variety of the products it sells and there have been lot of research and study conducted on these reviews to gain knowledge on the trustworthiness, effectiveness and usefulness of these reviews.

Star Level	General Meaning
⭐	I hate it.
⭐⭐	I don't like it.
⭐⭐⭐	It's okay.
⭐⭐⭐⭐	I like it.
⭐⭐⭐⭐⭐	I love it.

Figure 1. Rating System for Amazon.com

Haque et.al [7] in analyzed large-scale amazon dataset using supervised learning method to polarize it and got satisfactory accuracy.

Singh et.al [9]: Customers ratings and reviews were related to discover how useful or good the product is. Customers' ratings were analyzed based on different categories to get insights as to which is a good product and problems associated with non-performing products. They used Apache Hadoop as a framework to assess various categories of Amazon Product Reviews with a big datasets containing around 144 million reviews.

Jamshidi et.al [8] in their paper worked on detecting incentivized reviews and characterizing them into two primary categories of Amazon products. They have described a new method to identify explicitly incentivized reviews (EIRs). They have also demonstrated that a classifier that is trained by EIRs (without explicit keywords) can accurately detect other EIRs and implicitly incentivized reviews. Overall, their analysis sheds light on the impact of EIRs on Amazon products and users.

From the literature review, it has been found that very little work is available in predicting the star rating from the reviews in Amazon. In this work we propose an unsupervised, lexicon-based approach to classify the sentiments and predict the star rating based on the sentiment score.

III. RESEARCH METHODOLOGY

A. Data Collection

For our work we collected the reviews for "One Plus 7pro" mobile phone. A total of 2000 reviews along with the respective star rating given by the customers who have bought this Phone through Amazon where collected. Our main criterion for selecting this product was that it had comparatively large number of product reviews. We scraped the data using Parse hub tool to get the desired dataset.

B. Data Exploration

Data used for this study is the Amazon Product dataset consisting of Customer Reviews for "One Plus 7pro" mobile phones
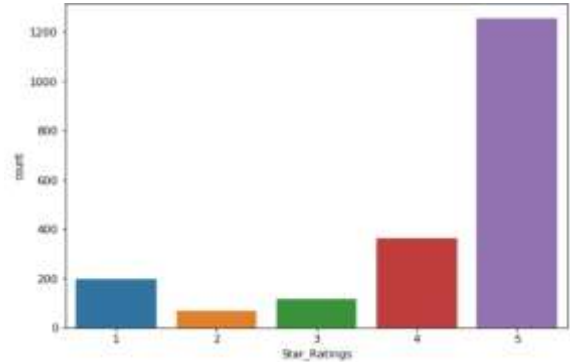

Figure 2. Frequency of Star ratings

From the cursory look at the data collected, we found the product to be quite popular among the masses. Over 81% reviews were positive (4 or 5 out of 5 stars). But sometimes, the ratings given may be misleading. Actual sentiment of the review may differ from the star rating given. Hence our aim is to compare the star ratings given with the actual sentiment of the review and find out if there is a major mismatch between the two. Working towards this goal, we analyzed the star ratings associated with the reviews.

C. Data Pre-processing

We performed pre-processing steps before applying sentiment analysis methods. We categorized the reviews as "Positive" whose star ratings are">3" and "Negative" for those star ratings are "<3". The below table shows the Rating categories and sample results.

Reviews	Star_Ratings	Rating _Category
Best Display in Market!	5	Positive
Awesome purchase	5	Positive
Overpriced to the cheap quality	2	Negative

Table 1. Reviews along with the Star-ratings category

Ratings_Category	Actual_Count
Positive	1616
Negative	384

Table 2. Star-ratings category

Figure 3: Word cloud of Positive and Nagative words

D. Lexicon-based classification

In our work we applied a lexicon-based approach for sentiment classification [6][17]. AFINN- 111 is a database of words with polarity attached to each word. The Valence is numbered between minus five (negative) and plus five (positive). Sentiment score was calculated for each review with the help of this.

Reviews	Star_Ratings	Sentiment_score
Best Display in Market!	5	3
Awesome purchase	5	4
Overpriced to the cheap quality	2	0

Table 4. Product reviews with their respective star ratings, sentiment score

The accuracy we got for this method was 62%.

This prompted us to use Machine learning technics like KNN to see if we could analyze the sentiment of the review text better than what we had achieved so far with the unsupervised method.

E. Machine Learning Algorithms

K-Nearest-Neighbours (KNN) [5] is a simple supervised classification algorithm. This is non-parametric and simple, but effective in many cases. This is one of the most popular classifiers for pattern recognition because of its efficient results and its simplicity. We used this model to analyze review Vs ratings and the performance of this model was analyzed based on various performance analysis metrics like precision, recall and accuracy.

The accuracy we achieved was 66%.

IV. CONCLUSION

The propose of our study was to build a framework based on text analytics using unsupervised sentiment analysis and Machine leaning model to provide the star rating for the reviews and validate the ratings with respect to the reviews. With the lexicon-based approach we got 62% accuracy while with Machine learning Algorithm (KNN) we got accuracy 66%. This shows that there exists an inconsistency between the sentiment score of the Reviews w.r.t to the star ratings.

Model	Accuracy
Lexicon based model	62%
KNN model	66%

Table 3. The results Obtained from the Lexicon based KNN Algorithm

Using various ML techniques like Random Forest, Support Vector machine (SVM) and other Deep Learning models helps organizations to analyse and validate the customer reviews vs rating more accurately, helping the product managers to understand the online customers behaviours better.

REFERENCES

[1] Al-Ghalibi, M., Al-Azzawi, A., & Lawonn, K. (2019, March). NLP based sentiment analysis for Twitter's opinion mining and visualization. In Eleventh International Conference on Machine Vision (ICMV 2018) (Vol. 11041, p. 110412A). International Society for Optics and Photonics.

[2] Khan, M. T., Durrani, M., Ali, A., Inayat, I., Khalid, S., & Khan, K. H. (2016). Sentiment analysis and the complex natural language. Complex Adaptive Systems Modeling, 4(1), 1-19.

[3] Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. Journal of Big Data, 2(1), 5.

[4] Yadav, V., & Elchuri, H. (2013, June). Serendio: Simple and Practical lexicon based approach to Sentiment Analysis. In Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013) (pp. 543-548).

[5] Hota, S., & Phatak, S. (2018). KNN classifier-based approach for multi-class sentiment analysis of twitter data. International Journal of Engineering and Technology.

[6] Kolchyna, O., Souza, T. T., Treleaven, P., & Aste, T. (2015). Twitter sentiment analysis: Lexicon method, machine learning method and their combination. arXiv preprint arXiv:1507.00955.

[7] Haque, T. U., Saber, N. N., & Shah, F. M. (2018, May). Sentiment analysis on large scale Amazon product reviews. In 2018 IEEE International Conference on Innovative Research and Development (ICIRD) (pp. 1-6). IEEE.

[8] Jamshidi, S., Rejaie, R., & Li, J. (2018, August). Trojan Horses in Amazon's Castle: Understanding the Incentivized Online Reviews. In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 335-342). IEEE.

[9] Singh, A. P., & Singh, G. (2019). Analysis of Amazon Product Reviews Using Big Data-Apache Pig Tool. International Journal of Information Engineering and Electronic Business, 11(1), 11.

[10] Shah, A., Kothari, K., Thakkar, U., & Khara, S. (2020). User Review Classification and Star Rating Prediction by Sentimental Analysis and Machine Learning Classifiers. In Information and Communication Technology for Sustainable Development (pp. 279-288). Springer, Singapore.

[11] Prabakaran, A., & Chen, M. (2019, February). Product review credibility analysis. In 2019 International Conference on Computing, Networking and Communications (ICNC) (pp. 11-15). IEEE.

[12] Kumar, S., & Kumar, K. (2018, March). LSRC: Lexicon star rating system over cloud. In 2018 4th International Conference on Recent Advances in Information Technology (RAIT) (pp. 1-6). IEEE.

[13] Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning

techniques. In Cognitive Informatics and Soft Computing (pp. 639-647). Springer, Singapore.

[14] Irel, R., & Liu, A. (2018). Application of data analytics for product design: Sentiment analysis of online product reviews. CIRP Journal of Manufacturing Science and Technology, 23, 128-144.

[15] Vyas, V., & Uma, V. (2019). Approaches to sentiment analysis on product reviews. In Sentiment Analysis and Knowledge Discovery in Contemporary Business (pp. 15-30). IGI Global.

[16] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.

[17] Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.