Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 04, April 2021: 2233-2246

Research Article

Discovering a Wide Range of Features For Sentiment Quantification Using Naive Bayes Algorithm

Dr. Dhanalakshmi R[1], Ms. Abirami D[2], Ms. Sivagami S[3]

[1,2,3]Department of CSE, RMK Engineering College, Thiruvallur, India

Email: rdl.cse@rmkec.ac.in

Abstract

Today, there are several product reviews accessible via the web. Consumer reviews provide a large amount of informative details that is useful to both consumers and businesses. However, product reviews are often unorganized, causing problems with collection of data and management of knowledge. The aim of this document is to present a grading system for product features that spontaneously identifies key features of products based on online customer feedback, with the goal of enhancing the usability of the various reviews. The most significant product aspect is characterized by two findings. First, a decent number of customers typically remark on the most critical things, and second, customer views on the important considerations have a substantial impact based on their general opinions of the product in terms of aspects rating. Given the customer feedback, we use a shallow dependency parser to classify reviews of product aspects and a sentiment classifier to assess consumer opinions on these aspects ranking. Then, by simultaneously considering the frequency of the aspect as well as the effect of customer opinions provided to each element of their entire product review viewpoints, we create a probability-based aspect ranking algorithm to infer the value of aspects.

Keywords: *Machine Learning, Sentiment Quantification, Deep Learning, Lexical Features*

Introduction

Consumers can make better buying decisions by concentrating more on important aspects or functionality, and companies can boost aspect efficiency by focusing more on important features or aspect rankings. As a result, the product in this proposed scheme would recognize the essential feature from online user feedback[20]. The NPL tool will verify the important aspect ranking based on the product feedback provided by the customer, identify the sentiment, and finally decide the specific product ranking based on the ranking algorithm[13]. The field classification can be applied to a a huge array of real-life situations. Aspect levels are useful for a broad variety of real-life applications[6]. Efficiency survey for two activities: sentiment classification at the document level for review documents and extractive analysis for rating summary. Further taking into account the most crucial aspects of product is highly helpful when making product review choices.

Companies and individuals will concentrate on improving the consistency of product factor ranking and enhancing the product's credibility more effectively. This collection is publicly available by request. In the proposed approaches Experimental results have demonstrated the effectiveness[14]. Moreover, ranking of product aspects was used to facilitate two real applications, a study of document classification and extractive analysis at the document stage.

1.1 Related Work

The existing methods for sentiment quantification from the relevant literature are discussed in this section.

The paper [1] examines the effect of ratings on economic results such as retail prices, as well as how various factors influence social outcomes such as their perceived utility. To identify essential text features, this paper investigates various facets of the summary document, such as subjectivity, readability ratings, as well as the extent of spelling errors. Multiple characteristics at the reviewer level are also examined, such as the average usefulness of previous reviews and the reviewers' self-discovery steps, which are shown next to the article.

The main goal of the project [2] in sentiment prediction, is to instantly determine whether or not a textual component reflects a personal opinion towards a subject of interest. As a regular text categorization problem, one can construct a sentiment prediction response, but gathering defined data proves to be a barrier. Fortunately, prior details about the sentiment polarity of terms in a lexicon in the form of context experience is frequently accessible. The writers introduce a new semi-supervised sentiment analysis method that uses lexical previous information with unlabeled sentiment prediction examples.

In the internet, Online auction Websites are highly dynamic, fast changing, and complex as they involved potential buyers and sellers, as well as a vast variety of goods eligible for bidding. The paper [3] enacted a two-phase method for mining and summarising hot products for various auction sites in order to aid in item decision-making. The first step of this paper is to automatically retrieve the item's product characteristics and product attribute prices from the merchant's details, and the second step is to summarise and find the hot products using the collected details from different auctions.

Review sites are a valuable resource for potential customers when making purchasing decisions. Even then, the overwhelming volume of available online reviews, as well as significant differences in online review quality, create a significant impediment to making constructive use of the ratings, as the most useful online reviews can be buried among a wide range of low feedback. The main objective of the document [4] is to build templates and predict the usefulness of a review algorithm, which is the foundation for finding the most helpful feedback for a given product in the online marketplace. They then demonstrate that the usefulness of reviews in modeling and predicting depends on three main factors: the writing style of the review in product, the reviewer’s expertise and the review's ability to be completed on time. These factors centered on the results, they were given a Simulation and estimation of usefulness using a nonlinear regression model. The suggested method is highly successful, according to the empirical research on the dataset on the IMDB movie reviews

In this paper [5], the authors have generated a probability-based paradigm in order to adapt knowledge fresh feature discovery and extraction wrappers. Wrapper adapting information refers to the process of automatically converting a prior gathered wrapper from a source Web site to a new, unknown web site for the purpose of extracting details. Characterization the ability of our system to find out about new or previously unknown characteristics from a new website is one of its strongest features. They have considered two varieties of data from the Web site to solve the wrapper adaptation issue of reviews online. The collection found in the previously taken in wrapper adaptation from the source online forum is the first type of detail. The second type of information is that which has been already collected or gathered from various websites. The authors used a Bayesian strategy to choose a collection of adaptive training examples knowledge in a wrapper for a previously unknown website. We introduce a model that allows the accompanying attributes parts in text to be examined in the new, never-before-seen website to solve the new attribute finding problem in adapting information. To find new properties in adapting information and their headers, a Bayesian learning method is created. To illustrate the utility of our system in knowledge extraction, we use the EM technique on a wide range of real Web pages.

Algorithm Descriptions

2.1 Algorithm

The intention of the Product Aspect Ranking algorithm is to discover important product attributes from a larger number of product reviews. In a product review, the overall opinion is a list of terms used to describe specific aspects of a product review. to figure out the product aspect's significance score. The features that are often mentioned are critical for consumers to make buying decisions. The views of consumers on particular aspects of the product have an impact on the general impressions of the commodity. In reviews, The Probabilistic Aspect Ranking Algorithm is used to measure the score after different factors are explored. The reports on the most critical aspects have a big influence on how people feel about the products overall. To get this general impression, we can measure the Overall score or review r from the weighted sum of opinions on a particular feature rating algorithm known as rkork. The view on the aspect ak and the significance weight rkw of the aspect ak is Ork m k=1. A higher rkw indicates that ak is more i Or is a vector of opinion on a particular element, and rw is a vector of weights. The Gaussian distribution is used to produce overall estimates and probabilities

Algorithm for probabilistic aspect ranking:

Input: A customer's perspective of a product corpus R, with each review r R having an overall rating Or and a vector of views.

Output: Importance scores m k k 1 v| = for all m aspects.

While not converged

Do Update | | 1 { } =r rw R

Update {,,} 2 s å m

End While Compute aspect importance scores m k k 1 v| = =1.

2.2 Sentiment

Opinion mining is the method of systematically defining, checking, quantifying, and studying affective states of a given product and subjective knowledge by using natural language processing, text analysis, and biometrics to identify and categorise opinions expressed in a part of text document. Whether the writer has a positive, negative, or neutral attitude toward a product or subject in the analysis in the product, It is the analytical task of determining what emotions an user is reflecting in text automatically. Sentiment analysis is used in a variety of settings, including smart businesses.

- Investigating media discussions about a topic
- Evaluating survey responses in terms of sentiment
- Determining whether product reviews are positive or negative are just a few examples of applications for detailed information.

The sentiment analysis algorithm is not perfect, and you will encounter errors in your output as with any automatic language analysis. In sentiment analysis, you won't be able to explain what makes a writer feel that way. It is, however, necessary for rapidly describing some regression estimates., particularly if you have a large amount of text to interpret.

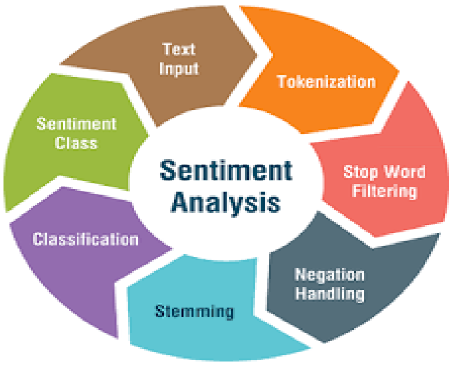

Figure 1. Sentiment Analysis

2.3 Working

There are several techniques to using a sentiment analysis algorithm. However, many methods share the same general concept:

1. Make or find a list of words in the review that are linked with either a heavy favorable or unfavorable sentiment.

2. Count how many favorable and unfavorable terms you will come up with. there are in the text.

3. Examine the review's favorable and unfavorable word balance. Positive sentiment is represented by a large number of favorable words and a small number of unfavorable words in the text document, whereas negative sentiment is reflected by a small number of favorable words and a larger number of unfavorable words.

It consists of two steps, the first of which is the most time-consuming: creating or defining a word list. While you can also use pre-existing lexicon features to save time, the second is that if your text is describing a particular subject in the algorithm, you may need to add to or alter it.

"Sick" is an example of a word that can have a positive, negative, or neutral consequence depends on what it's used to describe in sentimental analysis. If you're talking about a pet store that buys a lot of sick animals, the sentiment is likely to be negative. If you're talking about a skateboarding instructor who taught you how to do sick flips on a skateboard, the sentiment is now almost certainly positive.

2.4 Naive Bayes

A probability is a number between zero and one, with zero indicating that it will never take place and one indicating that it is unavoidable, and it is represented in mathematics as a number between 0 and 1. A prior distribution is a type of probability that is affected by external variables such as data or conditions. For example, you might go out with your friends (a possibility), but the weather might influence your decision — if it's raining heavily, you might not want to go out. As a result, whether or not you go out is dependent on the weather. Assume the likelihood of you going out regardless of what happens is A, and the likelihood of bad climate is B. To put it mathematically, based on the climate, the conditional likelihood of you going out, is p(A|B), or "the likelihood of A given B." The probability of both occurrences happening at the same time is known as the conjoint probability. The probability of you coming out in bad weather in the preceding case is p(A and B, You may recall (or have some absolutely no recollection from high school mathematics) that the probability of A and B is the multiple of the probability that is distinct of A and the probability that is distinct of B if both probabilities are distinct of each other. so p(A and B) = p(A)p(B). p(A) is an unique case of p(A|B) since we just observed that A is not independent of B. If it rains, you are less likely to go out, so p(A|B) p (A). p(A and B) = p(A|B)p, to put it another way, is a broader mathematical term (B). Because A and B can be any case, their probability of being together is commutative:

p(A and B) = p(A and B) (B and A)

If we replace the equations with the following:

p(A|B)p(B) = p(B|A)p(A)

Bayes' Theorem (also refered as Bayes' Law or Bayes' Rule) include this.

2.5 Advantages And Drawbacks Of Naive Bayes

Advantages of naïve bayes: Predicting the sample data set's class is simple and fast. It's also excellent the ability to predict different groups. A Naive Bayes classifier outperforms other models like logistic regression when the assumption of independence is correct, and less training data is needed. With categorical input variables, it works better than with quantitative input variables(s). A standard scheme is used for numerical variables (bell curve, which is a high assumption).

Drawbacks of naïve bayes : The model will give a probability of 0 (zero) if a categorical variable in the test data set has a category that was not present in the training data set. "Zero Frequency" is the term for this. We'll use the smoothing method for solving this. One of the most fundamental is Laplace estimation. The probability outputs from predict probe, on the other hand, should be treated with caution because naive Bayes is a poor estimator. Another flaw in Naive Bayes is the Independent predictors are believed. In practice, obtaining a completely independent set of predictors is virtually impossible.

The Naive Bayes algorithm is a processing method that uses the Bayes'Theorem algorithm and assumes predictor independence. A Naive Bayes classification in simple terms of algorithm, assumes that the existence of an important component in a class is unrelated to the presence of any other function in the algorithm. For example, if a fruit is orange in color, round, and about 3 inches in diameter, it is called an orange. Even if these features are dependent on one another or on the presence of other feature, they all provide to the chance of this fruit being an orange, which is why it is called "Naive." Naive Bayes models are simple to construct and extremely beneficial when dealing with huge data sets. Naive Bayes is said to be far superior to the simplest classification methods due to its simplicity.

The Bayes theorem helps in computing the posterior probability of P, which is P(c|x), by multiplying P(c), P(x), and P(x|c).as well as the following equation:

Likelihood   Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

Posterior Probability   Predictor Prior Probability

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 2 Naive bayes algorithm

• The probability in retrospect of class (c, target) given predictor is P(c|x) (x, attributes).

• P(c) is the class probability in the past.

• The likelihood is P(x|c), which is the probability of a predictor given a class.

• P(x) is the predictor's probability in the past.

2.6 The Proposed Framework

Using aspects ranking, a product aspect ranking framework in reviews is proposed for easily identifying key aspects of goods from consumer reviews on the internet It's our assumption that the following characteristics of a product's most critical features are widely discussed in consumer reviews of ranking, and that consumers' viewpoints on these aspects have a substantial impact on their overall opinions on the product in aspects ranking. A obvious frequency-based solution is emphasize the aspects ranking of the solution that process widely discussed in the online consumer reviews of product. In the Consumers' opinions on the product's aspects that are frequently encountered ranking, on either hand, doesn't have an influence on the overall opinions of the product, nor from their purchasing decisions. This methodology only assumes that the overall assessment is based on specific opinions on numerous perspectives, and it cannot precisely classify the relationship between an individual opinions and the overall rating opinion. As a consequence, we use another methodology and propose product aspects as part of a quality aspect ranking approach to infer product importance. On particular aspects over their viewer ratings on the product ranking is to calculate the number of incidents where their views on specific aspects and overall ratings are consistent, and then rank the aspects in the product based on the number of accurate instances.

Architecture

The architect is the one who dictates where even the line between software architecture design (architectural design) and intricate architecture design should really be stretched (non-architectural design).

2.7 Application layer

The user-oriented services in the application layer remain responsible for regulating user interaction with the system and are constituted up of numerous components in the presentation layer. This layer provides a common functionality to the system's core business logic, which itself is encapsulated in the system's business services.

2.8 Business layer

Business layered services encapsulate the contract manufacturing logic of services and integrate the system's core functionality in a framework. They typically comprise of necessary elements that are delegated to the business layer but might expose service interfaces that other callers can use. This layer's command pattern should be taken into consideration when planning these objects in the business layer.

2.9 Data link layer

Data link layer services provide access to the data within the constraints of the layered system, as well as data exposed by certain back-end systems in the application layer , which is frequently accessed via data services. The data link layer presents data to the business layered architecture through generic methods that the data layer's business services appear to find useful.

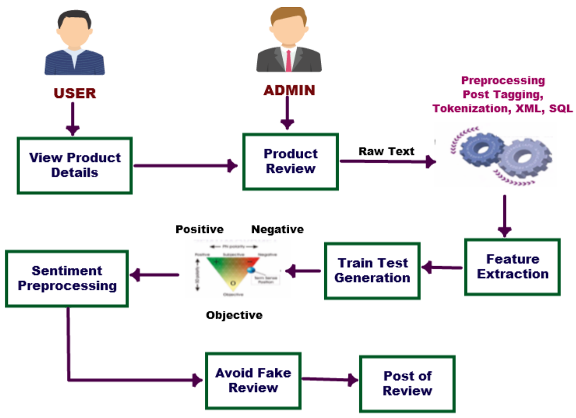

Figure 3. System Architecture

Data Modules

3.1 Product Aspect Identification

Consumer reviews in aspects are translated into various on a variety of forum blogs , which is shown. Online consumers are obliged to give the item in the review an overall rating, in the free text process on websites, explain concise comments and opinions on some product aspects, and write a paragraph of detailed review such as CNet. com. In the overall correlation was obtained by Viewpoints.com. just make a request in over-all process Other websites, like Reevoo.com, by instead require a product's overall rating of product item and a few descriptive positive and negative review perspectives on the critical features. In conclusion of system, a consumer reviews should include pros and cons criticisms, free of text reviews, or both., in combination with a high rating.We explore the potential aspects product's Pros and Cons by extract the frequently noun terms of words in the reviews. Aspects are traditionally noun phrases or nouns, as shown in the prior reports, and we can extract incredibly efficient aspect statistics by extractly constant noun terms from the product's Pros and Cons in the text. A effective approach in free text process reviews we can identify was to use a current approach to identifying aspects Proposed aspects is one of the most well-known techniques currently in use.

3.2 Product Aspect Ranking

We incorporate the characteristics of the proposed Product Aspect framework in this aspect ranking product. Three elements comprise the pipeline's overview: aspect identification of product, sentiment classification of product, and aspect ranking of product. First we recognize by aspects ranking in reviews of a product, and next investigate consumer opinions on aspects using only a sentimental classifier in the ranking system. We devise a probabilistic of aspect ranking algorithm is to identify the most influential aspects rank of a product's ranking and consequence of the consumer ratings on each aspect ranking on their opinions of product. a coalition of employee recommendations for a product or service. Investors voice their opinions in numerous aspect of products in each review, afterwards assign an rating, which should be a Additional levels of general opinion in the analysis are represented by a specific factor. A product's min and max ratings are the product's minimum and maximum ratings, respectively. It's worth remembering that existing customers from different websites might also have different rating distributions. In contrast, some Websites' ratings might become significantly lighter than others. Indeed, different online Websites should provide a various bell curve.

3.3 Probabilistic Aspect Ranking

We proposing a probabilistic aspect ranking algorithm in this module recognize consumer reviews in essential aspects of a product. Important aspects, in general, have the following process: prominently featured in consumer reviews item; and the opinions of product by consumers on these aspects ranking have a significant have an impact on overall opinion including its product. The general consensus of product is an aggregation facets ordered on product review, having different aspects of the product participating individually to the aggregation of viewpoints. opinions on important or insignificant aspects ranking have a strong or flimsy sway on generation of overall products opinions . That opinions on most important or minor factors have a strong or Very weak influence on the generation of overall products. We formulate overall rating as either a model for such aggregation. Otherwise, r is generated in each review based on weighted value of perspectives in various elements, The importance weight reflects the value placed on Larger indicating is more relevant, and vise - versa, as in matrix.

3.4 Extractive Product Review

There is a lot of online consumer Blog reviews have been published. In extractive, the particular product of review. The extractive reviews, and from the other hand, are discombobulated. It is impractical for a user or a corporate to interpret an overview of online consumer product reviews. As a consequence. In the extraction, there are different sorts of overall review summarization methods: abstractive and extractive summarization. The objective of abstract sweeping generalization is to develop a better understanding of big elements in source product and then implement those ideas in natural language system. It engages linguistic strategies to investigate Before establishing a new, shortcoming that conveys the basic information of the original text file in reviews, analyse and interpret the text.

3.5 Sentiment Classification

In literature, the task analyse the feelings expressed on the different aspects classifier is referred to as aspect-level of sentiment classification. There are two types of exiting methods classifiers: learning of supervised approaches and lexicon-bases approaches, are usually unsupervised learning. To achieve sentiment orientation on all aspect ranking, the lexicon-based technique uses a sentimental lexicon of product that consider number of phrases, idioms, and sentences. While these sentiment methods are simple to use, their effectiveness is highly dependent on quality of the sentimental lexicon characteristics. The supervised algorithm techniques, on the other hand, train a sentiment classification analysis involves corpus of training. Classification model is used to predicts how all element will be classified in terms of sentiment. The system may use a variety of learning-based classification models, Support Vector Machine and Naive Bayes algorithm, model are some examples., on the other hand, is a labor-intensive process and time-consumer process. Pros and Cons reviews of the product were used to categorise positively and negatively viewpoints on the ranking algorithm in this process.

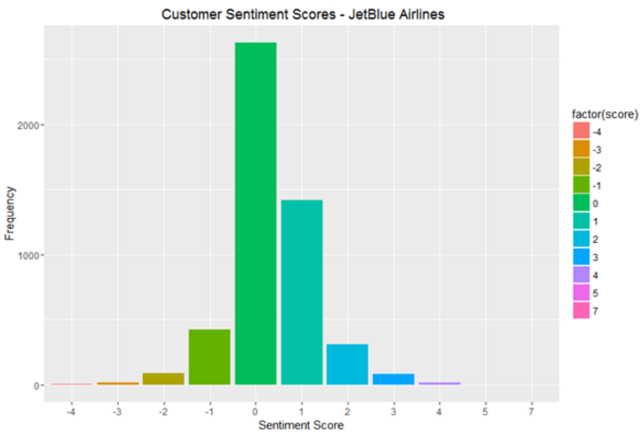

Figure 6. Sentiment Scores

3.6 Consumer

The primary aim of document-level sentiment analysis is to determine the overall impression of a product review report. A report for analysis usually reflects a variety of viewpoints on various features in a particular product. Divergent views on various topics can exist and have varying degrees of effect on the consumer's overall impressions of the text under analysis. Some features, such as dependability and ease of use, are praised in a sample iPhone 4 review paper, whereas others, such as the touch screen and music playback, are criticized. After that, it bestows a high overall rating on the iPhone 4 (a optimistic statement) because all of the key features have positive feedback. As a result, identifying important elements in the review document would naturally aid in the calculation of their overall impressions on samples prior. This finding prompts us to extend the results of the aspect ranking algorithm to the sentiment classification of product reviews in review papers.

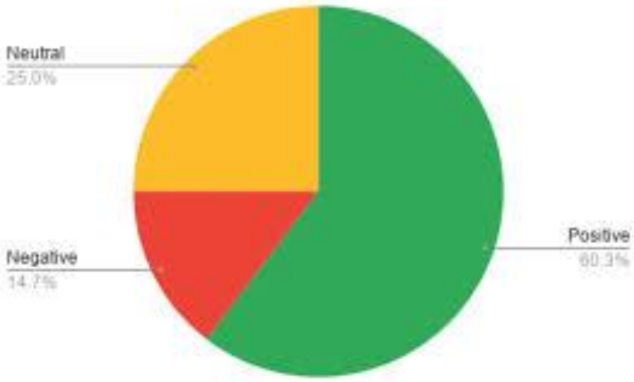

Figure 7. Probability of rating

Conclusion

In this article, we proposed a product aspect rating to assess the essential aspects of products from different customer feedback. Aspect sentiment classification, product aspect identification, and aspect ranking are the three main components of the framework. To begin, we used product Cons and Pros reviews to improve product aspect classification and sentimental analysis in free-text reviews. After that, we developed a probabilistic aspect ranking algorithm based on a huge amount of web reviewers to rate the value of different attributes of a product. Finally, the product features are rated in order of their relative value to the overall product. We carried out rigorous tests to assess the proposed framework components in a systematic manner. There are 95,560 customer feedbacks of 22 widely used items in eight areas in the experimental collection. This collection is obtainable to the public upon request, and it has two real-world applications: extractive review summarization and document-level sentimental analysis.

References

[1] Anindya Ghose, Panagiotis G. Ipeirotis, “Estimating the Helpfulness and Economic Impact of Product” IEEE transactions on knowledge and data engineering, Volume: 23, Issue: 10, Oct. 2011 pp: 1498 – 1512

[2] Vikas Sindhwani and Prem Melville, ” Document-Word Co-Regularization for Semi-supervised Sentiment Analysis”, 2008 Eighth IEEE International Conference on Data Mining.

[3] Prem Melville, Vikas Sindhwani, “Recommender Systems” Springer Reference 2017

[4] Yang Liu, Xiangji Huang, Aijun An, Xiaohui Yu, “Modeling and Predicting the Helpfulness of Online Reviews”, IEEE International Conference on Data Mining 2008
[5] Tak-Lam Wong, Wai Lam, “A probabilistic approach for adapting information extraction wrappers and discovering new attributes”, IEEE International Conference on Data Mining, April 2005
[6] Dhanalakshmi R, Sethukarasi T, “ A hybrid approach for mismatch data reduction in datasets and guide data mining”, International Journal of Networks, Software Tools and Applications, Cluster Computing, Springer, Vol. 22, Issue No. 5 (2017) pp. 10605–10614.
[7] Dhanalakshimi R, Geetha C, Sethukarasi T, “ Monitoring and Detecting Disease in Human Adults Using Fuzzy Decision Tree and Random Forest Algorithm”, International Journal of Recent Technology and Engineering, Vol. 07, Issue No. 5S2 (2019) pp. 93 – 99.
[8] Jyotsna Priya C, Deepti Rosini R V, Likitha M, Dhanalakshmi R, “ Credit Card Transaction Using Fingerprint Recognisation And Two Step Verification”, International Research Journal of Engineering and Technology , Vol. 06, Issue No. 03 (2019) pp. 2271 - 2273.
[9] Sudarvizhi R , Jasmine Gilda A , Dhanalakshmi R, “An Evaluation of a Computational Method to Extract Spike Based Features from EEG Signals to Analyse the Epileptic Seizures”, European Journal of Molecular & Clinical Medicine, Vol. 7, Issue No. 04 (2020) pp 2067 – 2076, ISSN 2515-8260.
[10] Anitha K, Dhanalakshmi R , Naresh K, Rukmani Devi D, “ Hyperbolic Hopfield Neural Networks for Image Classification in Content Based Image Retrieval” , International Journal of Wavelets, Multiresolution and Information Processing International Journal of Wavelets, Multiresolution and Information Processing, (2020)
[11] Sarala B, Rukmanidevi S, Bhargava D.S, Dhanalakshmi R, Kalpana A V, “A survey on management of spectrum scarcity issues and research challenges involved in finding creative techniques and solutions to data insufficiency tribulations – Cognitive Radio”, International Journal of Gedrag & Organisatie Review, Vol. 33,Issue No. 04 (2020) pp. 522 – 538, ISSN 0921 – 5077.
[12] Dhanalakshmi R, Anitha K, Rukmani Devi D, Sethukarasi T, “ Association rule generation and classification with fuzzy influence rule based on information mass value”, Journal of ambient intelligence and humanized computing, Springer, (2020) ISSN 1868 – 5137.
[13] B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1- 167, 2012.
[14] H. U. Khan and A. Daud, "Using Machine Learning Techniques for Subjectivity Analysis based on Lexical and Nonlexical Features," International Arab Journal of Information Technology (IAJIT), vol. 14, no. 4, 2017.
[15] M. Marchetti-Bowick and N. Chambers, "Learning for microblogs with distant supervision: Political forecasting with twitter," in Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, 2012, pp. 603-612: Association for Computational Linguistics.
[16] G. Amato et al., "How data mining and machine learning evolved from relational data base to data science," in A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years: Springer, 2018, pp. 287-306.
[17] J.-W. Bi, Y. Liu, and Z.-P. Fan, "Crowd Intelligence: Conducting Asymmetric Impact-Performance Analysis Based on Online Reviews," IEEE Intelligent Systems, vol. 35, no. 2, pp. 92-98, 2020.

[18] J.-W. Bi, Y. Liu, Z.-P. Fan, and J. Zhang, "Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews," Tourism Management, vol. 70, pp. 460-478, 2019.
[19] J.-W. Bi, Y. Liu, Z.-P. Fan, and E. Cambria, "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model," International Journal of Production Research, vol. 57, no. 22, pp. 7068- 7088, 2019.
[20] D. Zhang, C. Wu, and J. Liu, "Ranking products with online reviews: A novel method based on hesitant fuzzy set and sentiment word framework," Journal of the Operational Research Society, vol. 71, no. 3, pp. 528-542, 2020.
[21] M. A. Qureshi, C. O'Riordan, and G. Pasi, "Clustering with error-estimation for monitoring reputation of companies on twitter," in Asia Information Retrieval Symposium, 2013, pp. 170-180: Springer.
[22] R. Alaíz-Rodríguez, A. Guerrero-Curieses, and J. Cid-Sueiro, "Class and subclass probability re-estimation to adapt a classifier in the presence of concept drift," Neurocomputing, vol. 74, no. 16, pp. 2614-2623, 2011.
[23] W. Gao and F. Sebastiani, "From classification to quantification in tweet sentiment analysis," Social Network Analysis and Mining, vol. 6, no. 1, p. 19, 2016.
[24] A. Esuli and F. Sebastiani, "Optimizing text quantifiers for multivariate loss functions," ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 9, no. 4, pp. 1-27, 2015
[25] Chau Vo, Dung Duong, Duy Nguyen, Tru Cao, "From Helpfulness Prediction to Helpful Review Retrieval for Online Product Reviews", Proceedings of the Ninth International Symposium on Information and Communication TechnologyDecember 2018 Pages 38–45