

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## Customer Review Analysis with self-Learning Agent for Suggesting Solutions using Incremental Reward Deep Q Learning Model

Harish Bonthu<sup>1</sup>, Narendrasivam Karthikeyan<sup>2</sup>

Software Architect, Access Technology Solutions, USA<sup>1</sup>

Technical Architect, Access Technology Solutions, USA<sup>2</sup>

**ABSTRACT:** A large volume of product online reviews is generated from time to time, which contain rich information regarding customer requirements. These reviews help designers to make exhaustive analyses of competitors, which is one indispensable step in market-driven product design. How to extract critical opinionated sentences associated with some specific features from product online reviews has been investigated by the researches. However, these review sentences should be descriptive about general customer concerns and, at the same time, they are expected to be comparative to reflect similar customer feedback of different products. Hence, in this research, with the help of the techniques on sentiment analysis, opinionated sentences referring to a specific feature are first identified from product online reviews. We have proposed Incremental Reward Deep Q Learning Model (IRDQLM) as a classification model for finding the state of review. The results showed similarity value to the cosine function between the query review and answer given to query review for the customer. Moreover, it is observed that the proposed technique can able to reply all the reviews with a better similarity like a human response to the customer.

**KEYWORDS:** Customer, Online Reviews, Query Review, Requirements, Sentiment Analysis,

### I. INTRODUCTION

Manufacturers and marketers have long been focusing on product market performance, which benefits product marketing strategy development, product quality management decision making, and organizational performance improvement. Thus, manufacturers and marketers continually collect product data to analyze product performance in the market. Traditionally, the main data sources for product performance analysis have been manufacturers' internal data, reviews from experts, and off-line customer surveys [1]. Nowadays, millions of customers gain opportunities to compare similar products in online websites and pick their favorites in digital retailers, such as Amazon.com and Taobao.com [2,3]. Customer reviews in social media and electronic commerce Websites contain valuable electronic word information of products. In the first period, early customers do not have any information regarding past customers' reviews but only make their purchase decisions based on their quality expectations [4]. In the second period, late customers observe the average value of early customers' ratings and make their purchase decisions. On the other hand, sentiment analysis is a current research area where industrialists can know the opinions of people towards their product by analyzing the text written in form of reviews, blogs, survey response, etc. Sentiment Analysis is widely applied as voice of customers for applications that target marketing and customer service. Sentiment extractors in their most basic form categorize texts as either having a positive or negative or sometimes neutral sentiment. A common application of sentiment analysis is the automatic determination of whether an online review contains a positive or negative assessment [5,6].

This research paper is pre-arranged as follows. In section 2, numerous research papers on time data analysis using stock market data are reviewed. Detailed explanation about the proposed system is given in section 3. In addition, section 4 illustrates about the evaluation of the results and comparative analysis of the proposed system. The conclusion is made in section 5.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## II. LITERATURE REVIEW

To gain a general sketch outline regarding customer online reviews, various summarization approaches were innovated at the sentence level, the sentiment level or the product feature level.

Jin, J. *et al.*, [3] developed three greedy algorithms, to present suboptimal solutions for the optimization problems. To verify the availability of the developed approach, in particular, 4055 reviews of four popular mobile phones of different brands were obtained from Amazon.com. The names of the four products are represented as Product 1, Product 2, Product 3 and Product 4. The number of reviews of the four mobile phones is 905, 1108, 1088 and 954, respectively. With the help of the techniques on sentiment analysis, opinionated sentences referring to a specific feature are first identified from product online reviews. Then, information representativeness, information comparativeness and information diversity were investigated for the selection of a small number of representative yet comparative opinionated sentences. Potential research work can be extended in many directions, such as how to visualize these results in an interactive graphical user interface and how to compare products with the help of big opinionated product reviews in QFD, etc

Li, *Yet al.*, [4] developed a novel model to realize aspect-based sentiment summarization in an integrative way: composing the system with consistently designed feature extraction and clustering, collocation orientation disambiguation, and sentence sentiment strength calculation. Collocations of product features and opinion words are initially extracted through pattern-based bootstrapping. A novel confidence estimation method considering two measurements, Prevalence and Reliability, is exploited to assess both patterns and features. The obtained features are further clustered into aspects. Each cluster is assigned a weight based on arithmetic means of feature similarities and confidences. The orientations of dynamic sentiment ambiguous adjectives (DSAs) are then determined within opinion collocations. Finally, sentiment strengths of opinion clauses for each aspect are computed according to a set of fine-grained and stratified scoring formulae. Experimental results on a benchmark data set validates the effectiveness of the proposed model. However, the automatic determination of parameters such as iteration in bootstrapping, the weighting coefficients and the clustering thresholds, was also a crucial issue.

Yan, E. *Qet al.*, [7] required the user to focus on one element at a time, but StarCraft 2 is a game where players often control more than a hundred units simultaneously. The game interface provides an optional mechanism called “control groups” that allows players to select multiple units and assign them to a group in order to quickly recall previous selections of units. From an analysis of over 3,000 replays, we show that the usage of control groups is a key differentiator of individual players as well as players of different skill levels—novice users rarely use control groups while experts nearly always do. But players also behave differently in how they use their control groups, especially in time-pressured situations. While certain control group behaviors are common across all skill levels, expert players appear to be better at remaining composed and sustaining control group use in battle. We also qualitatively analyze discussions on web forums from players about how they use control groups to provide context about how such a simple interface mechanic has produced numerous ways of optimizing unit control. However, guidance for the interactive scenarios where the human abilities were taxed needed improvement.

Flory, L. *et al.*, [8] developed a new web personalization decision-support artifact for utility-sensitive customer review analysis in response to an increase in consumer use of the vast array of online reviews. Given the increasingly high volume of such reviews, automatic analyses of their quality have become imperative. Not surprisingly, the interest of researchers was attracted. However, prior approaches are insufficient to address the consumers’ need for non-burdensome sense making of online reviews. This research attempts to close this gap by developing a novel design science artifact, i.e. construct, architecture, algorithms and prototype to address the consumers’ need. The artifacts using a set of experiments and hypothesis tests. The results validate the effectiveness and efficiency of the proposed artifacts. The practical utility and relevance using real world pilot experiments were evaluated. The results performed adequately validating the solutions of review quality research. However, improvement in the IUSR Analyzer into an artifact effectively and efficiently needed to analyze the quality of online reviews written in multiple natural languages and was very valuable to the research and professional communities.

Tandon, *Uet al.*, [9] examined key predictors influencing customer satisfaction towards online retailing in India. The findings of the study reveal that perceived usefulness, perceived ease of use and website design have positive impact on customer satisfaction. Perceived time performance emerged as strongest predictor of customer satisfaction

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

indicating a strong positive association between perceived usefulness and customer satisfaction. The emergence of ease of purchase, ease of understanding, and ease of navigation also appeared as significant variables leading to customer satisfaction where ease of ordering had a significant but negative impact. This research will help online retailers in India to motivate new customers for online shopping and existing customers to extend it in their daily purchase. The scale of ease of use and perceived usefulness have been extended in emerging economies context. However, the developed method was an exhaustive model to improve website functionality, perceived ease of use, perceived usefulness and customer satisfaction of online retailing.

### III. PROPOSED METHOD

After reviewing the related works, a quite a number of research works have been attempted in recent years to extract customer preferences and product attributes, determine their rankings, and relate customer preferences with product attributes based on online customer reviews. However, the development of customer preference models based on the review of data obtained from online customer reviews, was used to determine the product attribute settings of new products. The limitations inferred from the existing methods such as are criticized and a proposed method aims to overcome such limitations there by contributing an efficient proposed method. The objectives of the proposed method are shown in figure 3.1.

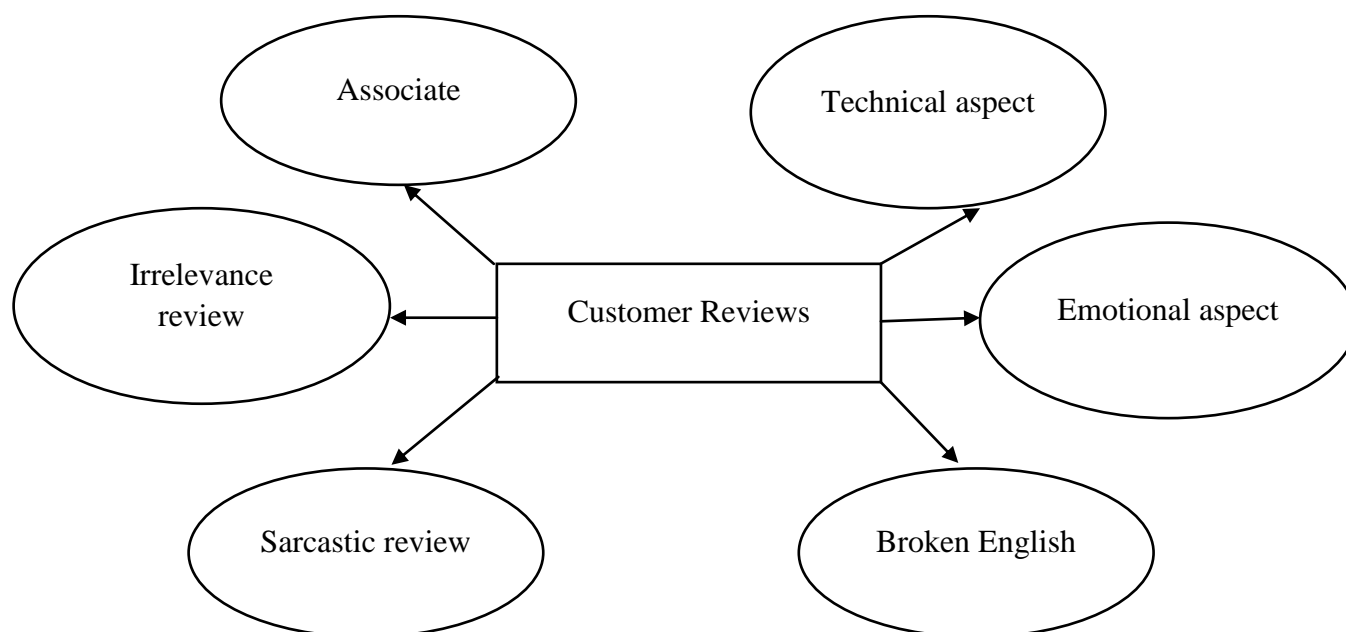


Figure 3.1: Objectives of the proposed methodology

The main objectives of the proposed method are shown in the block diagram. The research work mainly aims at replying automatically to the review given by the customers in the online sites [10-14]. The customer reviews obtained from the online sites are analyzed carefully and replying in turn to the customer based on the

- i) technical aspect such as product based reviews
- ii) emotional aspect such as anger, happiness
- iii) broken English reply (Review with grammatical mistakes).

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

- iv) Sarcastic review from the customer
- v) Irrelevance review on the product which is not at all related to that product
- vi) Associates which are related to the product.

Considering all these objectives, a reply to the customer is processed that is based on the perspectives of the person IQ level. To implement all these objectives in the proposed methodology, an efficient model is proposed and is shown in the figure 3.2 An overview of the proposed methodology block diagram is shown in the figure 3.2. The steps are as follows,

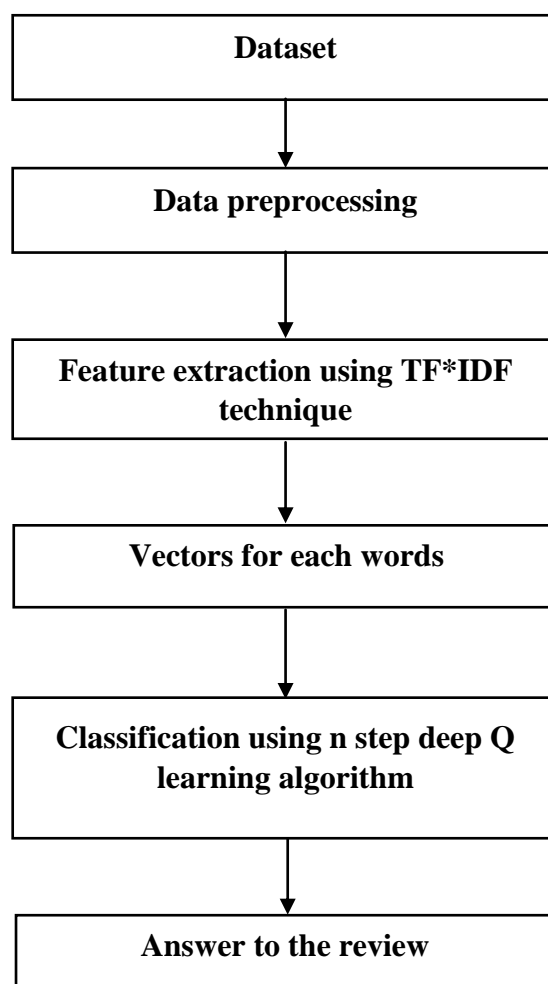


Figure 3.2: Block diagram of the proposed methodology

### 3.1 Data Collection

The online reviews of the sample products are found from the website and a web scraping software is used to collect the contents. Based on a fixed-time-period strategy, the collected online reviews are divided according to time period and are stored in different Excel files. Opinion mining is then employed to conduct sentiment analysis to extract customer preferences of the sample products and compute the corresponding sentiment scores. The process of opinion

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

mining adopted in this study mainly involves the following six steps. First, the pre-processing of the unstructured texts is conducted to remove noise such as HTML characters, punctuation and stop words. Second, Part-Of-Speech (POS) tagging is applied to tag the opinion words in the text as either nouns, adverbs, verbs or adjectives [15-17]. In this study, nouns are considered as customer preferences, whereas adverbs and adjectives are treated as the corresponding sentiment expressions. Third, relevant sentiment expressions of customer preferences are extracted from online reviews. Fourth, incorrect features and redundant features are eliminated using feature extraction. Similar to previous studies, elementary product features are represented by noun phrases, and sentiment indicators are represented by adjective or adverb. Therefore, nouns are extracted to form a sentences library as the basic transaction documents. High-frequency feature word ranking by association rules. The association rules detect frequent itemsets for extracting product feature words. From the basic transaction documents, those frequent itemsets meeting the minimum support algorithm are considered as the candidate words.

## 3.2 Online Data Preprocessing

The pre-processing of the unstructured texts is conducted to remove noise such as HTML characters, punctuation and stop words. Transforming text into something an algorithm can digest is a complicated process [18]. In this section, the steps involved in text processing are as follows:

- (i) **Tokenization:** Tokenization includes conversion of sentences to words also called as tokens before transforming it into vectors. The unnecessary tokens are easy to filter, where a document is transformed into paragraphs or sentences into words. In our research work, the online reviews are tokenizing into words.
- (ii) **Removal of Unnecessary tags and Punctuation:** The next step is to remove punctuation, as the punctuation don't do any extra information while treating text data. Therefore, removing all instances will help to reduce the size of the training data.
- (iii) **Removing stop words:** Frequent words such as "the", "is", etc. that do not have specific semantic. The search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. Those words that appear sparsely in the appraisal relationship will be filtered out. The list of words that are not to be added is called a stop list.
- (iv) **Stemming:** The words are reduced to a root by removing inflection through dropping unnecessary characters, usually a suffix. Stemming is the process of reducing a word to its word stem that affixes to suffixes and prefixes or to the roots of words known as a lemma. Stemming is important in Natural Language Understanding and Natural Language Processing.
- (v) **Lemmatization:** Another approach to remove inflection by determining the part of speech and utilizing detailed database of the language. Thus stemming & lemmatization help to reduce a common base form or root word. The filtered words are now undergone for feature extraction process in order to increase the proportion to the number of times a word appears in the document based on the frequency offset.

## 3.3 Feature Extraction using TF-IDF

In text processing, words of the text represent discrete, categorical features. The mapping from textual data to real valued vectors is called feature extraction. One of the simplest techniques to numerically represent text is Bag of Words. The list of unique words in the text corpus called vocabulary represents each sentence or document as a vector with each word represented as 1 for present and 0 for absent from the vocabulary [19]. Another representation can be count the number of times each word appears in a document. The most popular approach is using the Term Frequency-Inverse Document Frequency (TF-IDF) technique. The TF-IDF value increases in proportion to the number of times a word appears in the document but is often offset by the frequency of the word in the corpus, which helps to adjust with respect to the fact that some words appear more frequently in general. TF-IDF use two statistical methods, first is Term Frequency and the other is Inverse Document Frequency. Term frequency refers to the total number of times a given term  $t$  appears in the document  $doc$  against the total number of all words in the document and The inverse document frequency measure of how much information the word provides. It measures the weight of a given word in the entire document. IDF show how common or rare a given word is across all documents.  $TF*IDF$  do not convert directly raw data into useful features. Firstly, it converts raw strings or dataset into vectors and each word has its own vector. Then

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

we'll use a particular technique for retrieving the feature like Cosine Similarity which works on vectors, etc. TF\*IDF provides numeric values of the entire document and these values is next processed for the classification.

### 3.4 Classification using Incremental Reward Deep Q learning model

Similar to traditional multi-step methods, such as IRDQLM at each update step, the proposed IRDQLM also uses multi-step rewards, which exploit longterm high-reward strategies from past experience and improve the efficiency of the selected reward signal. This makes the Reinforcement Learning agent more data-efficient and can help to reduce the number of catastrophic actions. Moreover, similar to one-step Temporal Difference methods, in the proposed method, the time interval between the estimated state-action pairs are one-step [20]. The training data generated by the new n-stepwise Q-learning can be expressed as

$$\hat{e}_t = (S_t | \theta_t, a_t | \theta_t, r_t | \theta_t, S_{t+1} | \theta_t, r_{t+1} | \theta_{t+1}, r_{t+2} | \theta_{t+2}, \dots, \theta_{t+n-1} | \theta_{t+n-1}) \quad (1)$$

The data tuple generated by the proposed new IRDQLM, the future state is  $S_{t+1}$ , which is produced under condition  $\theta_t$ . Therefore, the data within  $\hat{e}_t$  are more strongly correlated than traditional incremental methods. This illustrates that the new multi-incremental reward Q-learning is more suitable for improving the performance of the classic Deep Learning by combining experience replay. The loss function is

$$L(\theta) = E_{(S, a, \hat{r}^{(n)}, S') \sim U(D)} [(\hat{Y}^{nDQM} - \hat{Q}(S, a; \theta))^2] \quad (2)$$

where  $\hat{Y}^{nDQM} = \hat{r}^n + \gamma \max_a \hat{Q}(s', a; \theta^-)$ . Unlike DQM, in the IRDQLM, each training data  $\hat{e}_t = (S_t, a_t, \hat{r}^{(n)}, S_{t+1})$  contains multi steps reward signals. When an agent executes a catastrophic action and obtains a negative reward, n data tuples contain the negative reward instead of one. The agent execution obtains an answer to the review commented by the customer. The multiple levels of the reviews are also being examined and a fair review will be replied to the customer based on the IQ level of the customer. Algorithm for the IRDQLM is as shown below.

```

Initialize experiences replay memory D
Initialize parameters  $\theta$  randomly
Initialize state-action value function  $\hat{Q}$ 
Initialize the parameters of target network  $\theta^- = \theta$ 
Initialize the hyper parameter  $n$ 
For episode = 1 to M do
    Initialize sequence  $S_1 = \{x_1\}$  and preprocessed sequence  $\varphi_1 = \varphi(S_1)T \leftarrow \infty$ 
    For t = 1, 2 . . . do
        If t < T then
            With probability  $\epsilon$  select a random action as review otherwise select
             $a_t = \arg \max_a \hat{Q}(\varphi(S_t), a; \theta)$ 
            Execute action  $a_t$  in the emulator and observe reward  $r_t$  and
            Answer  $x_{t+1}$ 
            Set  $s_{t+1} = S_t, a_t, x_{t+1}$  and preprocess  $\varphi_{t+1} = \varphi(S_{t+1})$ 
            If  $s_{t+1}$  is terminal then
                T  $\leftarrow$  t + 1
                 $\tau \leftarrow$  t - n + 1

            If  $\tau \geq 0$  then
                If  $\tau + n < T$  then
                     $\hat{r}_\tau = \frac{1}{n} \sum_{k=\tau}^{k=\tau+n-1} r_k$ 
                Else
                     $\hat{r}_\tau = \frac{1}{T-\tau} \sum_{k=\tau}^{k=T-1} r_k$ 

```



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

Store transition  $(\varphi_\tau, a_\tau, \hat{r}_\tau, \varphi_{\tau+1})$  in D  
 Sample random mini-batches of transition  $(\varphi_\tau, a_\tau, \hat{r}_\tau, \varphi_{\tau+1})$  from D  
 Set  $y_i = \begin{cases} \hat{r}_i & \text{if } s_{i+1} \text{ is terminal} \\ \hat{r}_i + \gamma \max_a \hat{Q}(\varphi_{i+1}, a; \theta^-) & \text{otherwise} \end{cases}$   
 Perform a gradient descent step on  $(y_i - \hat{Q}(\varphi_{i+1}, a; \theta^-))^2$  with respect  
 To the network parameters  $\theta$   
 Replace target parameters  $\theta^- \leftarrow \theta$  every C step  
 Until  $\tau = T - 1$

End

The IRDQLM is used to illustrate the generality of the new incremental reward steps and update is the same as for IRDQLM, but the target IRDQLM is replaced

Where

$$\hat{y}^{IRDQLM} = \hat{r}^n + \gamma \widehat{Q}(s' \arg \max \hat{Q}(s', a; \theta^-)). \quad (3)$$

Moreover, it is clear that when  $n = 1$ , the proposed new DRL algorithms DQM and IRDQLM are equivalent to DQM and IRDQLM, respectively. The DENFIS approach that performs the dynamic modelling of customer preferences needed improvement with the adaptive determination of DENFIS parameter. By considering all the drawbacks of the developed methodologies, a novel approach is developed various parameters to analyze the customer review as either positive or negative and replying the consumer is introduced in the proposed IRDQLM model.

## IV. RESULTS AND DISCUSSION

In this research, customer reviews are analyzed with comprehensive and reasonable artificial learning technique and given the suitable solution. The proposed solution is executed through the system with python 3 in Windows 10, i7 core processor, 16 GB RAM, 6 GB 2080Ti NVIDIA GTX edition GPU environment. The proposed solution has been executed for a data collected from amazon for 12 different categories with each containing 10 products and their customer reviews for a year period of time ranging from January, 2015 to December, 2015 as shown in Table 4. 1.

Table 4.1: Customer reviews for a year period of time ranging from January, 2015 to December, 2015

Products	Number of Reviews
Arts	27,980
Beauty	252,058
Books	12,886,488
Cell_phones & Accessories	78,930
Clothing & Accessories	581,933
Home & Kitchen	991,794
Jewelry	58,621
Movies & TV	7,850,072
Musical Instruments	85,445
Office Products	188,234
Pet Supplies	217,230
Watches	68,356

Since, it is complex to detail all the scenario in customer review for all the collected data the different aspect of a single product review has been detailed as a case study, the reviews about Samsung Galaxy S6 edge plus mobile has been taken and portrayed in different aspects of customer review and their reply.

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

### Case I: technical aspect such as product based reviews

In this case, the customer will review about the technical features and its efficiency of the product in terms of a detailed technology either in positive or a critic way. In these cases the proposed solution has to understand the technical aspects said by the customer and find the corresponding solution from the description of the product. For an instance, the review and reply can be as follows:

#### Review:

*I've been using the phone for a month and I can honestly say I'm in love with the phone. After the galaxy S3 I vowed never to buy another Samsung phone, but after watching so many excellent reviews of the phone I decided to buy it.*

#### Pros:

- *Probably the best screen on any smart phone. 2K, super AMOLED, the curve looks gorgeous, ridiculously sharp, uses very little battery when reading ebooks (black background)*
- *Probably the Camera on a smart phone. The pictures are sharp, detailed, the auto focus is fast, pictures are taken instantly and the LOW LIGHT IMAGES ARE BRILLIANT. I can't stress this enough. The OIS and the ability to shoot wide open really helps. As a comparison it's better than the iPhone 6s by quite a bit and a little bit ahead of the 6S plus*
- *Great front facing camera. 5MP but anyone who knows anything about cameras knows that megapixels don't make the camera. This takes brilliant photos, and the image processing is superb.*
- *The phone looks gorgeous. The Gold colour is more of a champagne colour but it is stunning. Pure vanity, but this is easily the best looking phone in production today.*
- *Specs are off the charts. This thing is will keep going for quite a few years without any trouble.*
- *Touchwiz is much lighter than it has been.*
- *Wireless Charging*
- *Quick charging (I mean really quick. A full charge in just over an hour)*
- *Fingerprint sensor. It's really quick and accurate. Not as fast as the 6S or the Nexus 6P but it's more than quick enough. You'll never be left waiting for it to unlock or having to hit it multiple times to unlock.*

#### CONS (And there are a few)

- *Slippery as all hell. You don't feel confident holding the phone in your hand and it slips out of your pocket when you sit down if you're not careful.*
- *No expandable storage. There's just no excuse for it.*
- *No water proofing. At this price and in today's day and age. Come on Samsung.*
- *The single bottom speaker is average.*
- *No USB type C. For a late 2015 release this should have been there.*
- *Price. On the expensive side.*
- *The usual slow updates. If you want the latest version of Android the Nexus 6P is what you're looking for.*

*Reply: Thanks and appreciation for the pros you given about display, camera and charging. We value the cons such as slippery, storage, water proofing, speaker, USB type, price and software updates. The product was finished with smooth glass design for an efficient look in a compatible size which suits in hand and minimize slippery. And the product comes with a 128 GB internal memory and cloud backup feature which minimizes the need of external memory. Product is developed with IPS4 splash proof design. Moreover a monolithic speaker is given to avoid stereo effect, noise and low battery power. Even though without USB Type-C we given an efficient fast charging and universal adaption to all device connections effectively. We think the product is cost-efficient when compared to the similar products in market. As per the software agreement with google it will be made available to you as early as possible when they release. We will intimate all your concerns towards the producer as a official note and hope for the latest version of this product to be valued satisfactory by you.*



## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

### Case II: emotional aspect such as anger, happiness

In case of emotional aspect, the customer will express his anger, joy, happiness or any other feelings through text and its describing expression words. For this situation the proposed solution will identify the emotion of the customer and it will give the corresponding solution by taking part in his/her feelings and making the presence of the customer's review has been listened with a token of importance as shown below.

*Review: Delivery of the phone was as promised. open the box and you can see a slim shinny beautiful device. I bought the gold color that gives an exceptionally premium look. Design is unparalleled, light weight, superslim. Edge display has its own functions like night clock, battery backup is a day on normal usage with 3g on, call quality superb. One needs a phone case or cover as the phone is quite slippery. All in all its a head turner and if you have the money go buy it or wait till march for the S7 series.*

*Reply: Thanks for the appreciation about delivery. Hope you enjoy spending more time with your beautiful device. Hold it grip till you get a good-looking phone case delivery for it. You can find some attractive phone cases from our accessories section.*

### Case III: broken English review (Review with grammatical mistakes).

In scenario such as a broken english review, the customer may use an abbreviation or a grammatically wrong sentence or a colloquial word in the sentence. In such cases the proposed technique will understand the keywords the customer described and try to relate it for corresponding reply to the query as follows.

*Review: but in a nutshell beautiful device crafted with premium material as saying goes beauty lies in the eyes of the beholder truly an amazing gadget and total paisa vasool*

*Reply: We think the real beauty lies in the device is found by the eyes of the beholder. Its good to hear such appreciation about our design and cost efficiency from you.*

### Case IV: Sarcastic review from the customer

In sarcastic reviews, the customer will praise the product in negative way. The customer reviews the product in a degrading or with the intention to harm the product, brand, seller or the e-commerce website. In this case the proposed technique should understand the contextual meaning of the sentence and reply in such a way that it will be legally as well as ethically correct to the corresponding situation like given below.

*Review: Huh... We are bound to pay for only it's brand value ???! Huawei honar 7, mate S , even the one plus 2 is much more better in all aspects.. So think AGAIN before going for it.. Those who already bought this may be forced to love this as they paid such a huge amount..*

*Reply: We think the brand value lies in the quality of product. May different products also perform better based on their aspect quality. But we can assure our customers who bought product from us are not bound under any forceful circumstances and they can attain their expectations over the purchase by utilizing of return and warranty policies.*

### Case V: Irrelevance review on the product which is not at all related to that product

A customer review can be categorized to be irrelevance if it doesn't match any of the terms about the product or seller or the service provided by the e-commerce agent. It can be understood by the proposed solution by measuring its cosine similarity with the known details about the product and its seller. These kinds of irrelevant reviews will be automatically ignored by the system and it will not give any reply.

*Review: Kay huaa*

*Reply: - (Ignored)*

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

For example in the above situation, customer reviews a content in his/her colloquial language which is nowhere related to the product given. This was identified by the proposed technique and avoided replying to that review.

**Case VI:** Associates which are related to the product.

In this case, the review will be made towards the associates or delagates which are in official relation with the e-commerce selling agent. It can be a manufacturer, seller, dealer , delivery agent or even the e-commerce selling agent itself by one way or other related to the selling chain. Here, the proposed technique should understand the issue in customer’s purchase and find the corresponding source where the issue happened. Then it should suggest the corresponding ethical and legal remedy for all the persons involved in the purchase as like given below.

*Review: I have been a regular customer of Amazon and already I gave bussiness of Rs.5 lac to them but this is very unfortunate that they are having such dealer who deliver products which is defective . Phone is not consuming charge and there is some problem in it's headphone as well . Sound is not clear it is having some echo*

*Reply: Your inconvenience from our dealer is being sincerely regretted. We value all the purchases from amazon with our highest concern and care as possible, the same will be inherited through all our dealers also. We would like you to contact our customer care executive through an official call with your bill copy and issues you are facing. So that we can help you to claim the replacement or warranty terms as per the regulations defined from our dealer.*

The above seen reviews and reply cannot be evaluated manually for all the reviews in practical situations. These can be evaluated successfully with the help of cosine similarity between the review and the reply which shows the relevancy in the conversation. Cosine similarity can be measured with the help of following equation...

$$\text{Similarity}(\text{Re view}, \text{Re ply}) = \frac{\text{Re view} \cdot \text{Re ply}}{\| \text{Re view} \| * \| \text{Re ply} \|} \dots\dots(3)$$

Table 4.2: The average similarity values of reviews and reply in category wise

Products	Number of Reviews	Q Learning	Deep Q Learning	Incremental Reward based Deep Q Learning
Arts	27,980	0.792	0.845	0.923
Beauty	252,058	0.771	0.821	0.912
Books	12,886,488	0.787	0.865	0.945
Cell_phones & Accessories	78,930	0.773	0.882	0.975
Clothing & Accessories	581,933	0.792	0.867	0.943
Home & Kitchen	991,794	0.797	0.843	0.922
Jewelry	58,621	0.784	0.851	0.937
Movies & TV	7,850,072	0.795	0.868	0.954
Musical Instruments	85,445	0.782	0.854	0.968
Office Products	188,234	0.789	0.841	0.915
Pet Supplies	217,230	0.786	0.847	0.917
Watches	68,356	0.787	0.849	0.913

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

In the table 4.2 the average similarity values of reviews and reply in category wise is shown and been plotted as a graph in figure. Moreover, it is observed that the proposed technique can able to reply all the reviews with a better similarity like a human response to the customer.

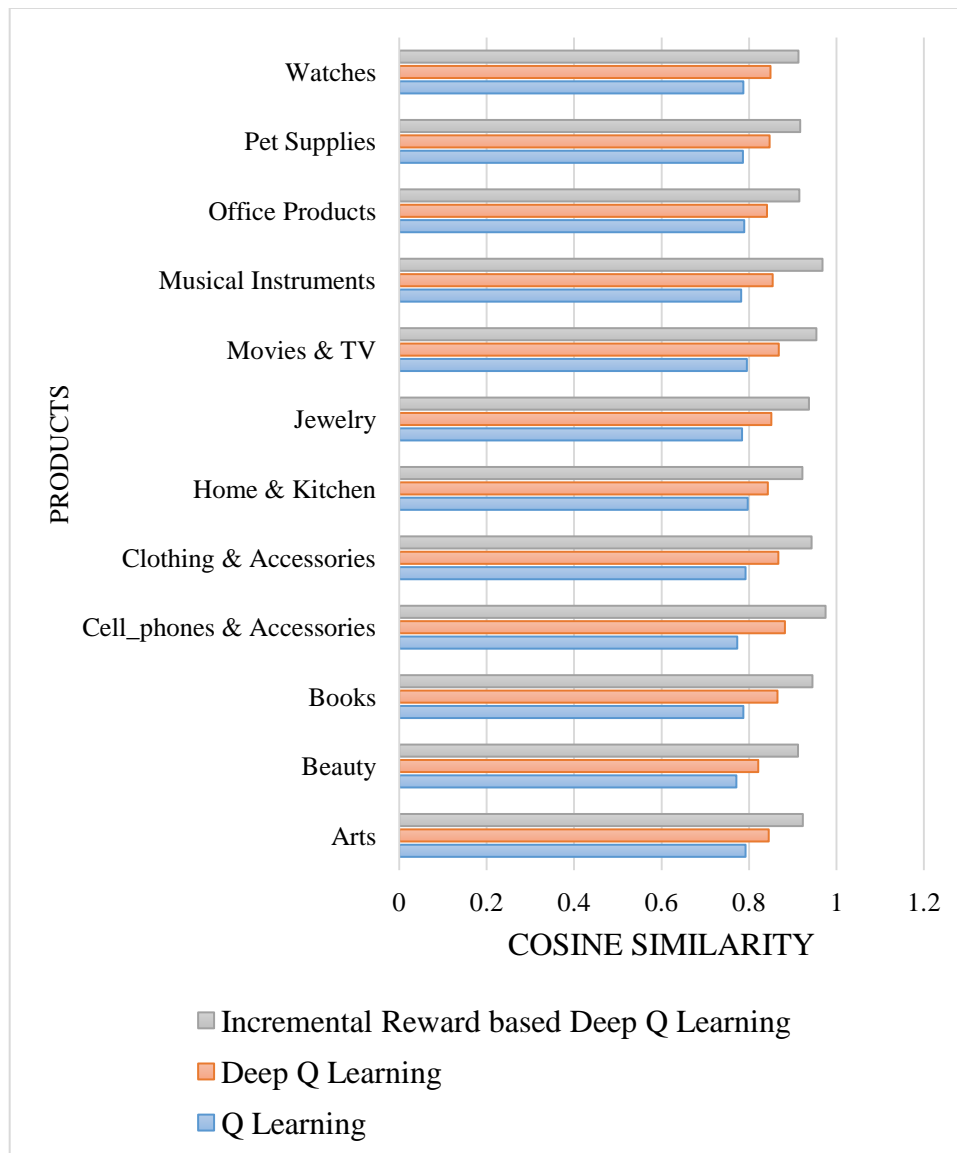


Figure 4.1: Customer Reviews with a higher relevancy to the customer expectations

Based on the above observations, it can be understood that IRDQLM is capable of replying all the customer reviews with a higher relevancy to the customer thought and comprehensible enough for the customer to understand the reply given by IRDQLM.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## V. CONCLUSION

The problem of information summarization from large textual data, also incorporating the sentiment towards the summary information. For that purpose, new concept based IRDQLM in order to answer for the customer who has replied a review for the product in the online site. The textual feature extraction process extracts sentimental and textual data sources. Sentimental causal rules have many practical applications such as extracting summary information from customer reviews to understand the causality among different aspects of products and sentiment towards these aspects for marketing purposes. The proposed methodology in practice enables automatic processing replying for the review for the customer based on the IQ level of the customer. The future work includes the analysis of images in the customer review and an effective response is given after analyzing the customer review.

## REFERENCES

1. Smith, A.N., Fischer, E. and Yongjian, C., 2012. How does brand-related user-generated content differ across YouTube, Facebook, and Twitter?. *Journal of interactive marketing*, 26(2), pp.102-113.
2. Dong, J., Li, H. and Zhang, X., 2014, November. Classification of customer satisfaction attributes: an application of online hotel review analysis. In *Conference on e-Business, e-Services and e-Society* (pp. 238-250). Springer, Berlin, Heidelberg.
3. Jin, J., Ji, P. and Gu, R., 2016. Identifying comparative customer requirements from product online reviews for competitor analysis. *Engineering Applications of Artificial Intelligence*, 49, pp.61-73
4. Li, Y., Qin, Z., Xu, W. and Guo, J., 2015. A holistic model of mining product aspects and associated sentiments from online reviews. *Multimedia Tools and Applications*, 74(23), pp.10177-10194.
5. De Giovanni, P. and Zaccour, G., 2014. A two-period game of a closed-loop supply chain. *European Journal of Operational Research*, 232(1), pp.22-40.
6. Farhadloo, M., Patterson, R.A. and Rolland, E., 2016. Modeling customer satisfaction from unstructured data using a Bayesian approach. *Decision Support Systems*, 90, pp.1-11.
7. Yan, E.Q., Huang, J. and Cheung, G.K., 2015, April. Masters of control: Behavioral patterns of simultaneous unit group manipulation in starcraft 2. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 3711-3720). ACM.
8. Flory, L., Osei-Bryson, K.M. and Thomas, M., 2017. A new web personalization decision-support artifact for utility-sensitive customer review analysis. *Decision Support Systems*, 94, pp.85-96.
9. Tandon, U., Kiran, R. and Sah, A.N., 2016. Analysing the complexities of website functionality, perceived ease of use and perceived usefulness on customer satisfaction of online shoppers in India. *International Journal of Electronic Marketing and Retailing*, 7(2), pp.115-140.
10. Elish, M.O., 2014. A comparative study of fault density prediction in aspect-oriented systems using MLP, RBF, KNN, RT, DENFIS and SVR models. *Artificial Intelligence Review*, 42(4), pp.695-703.
11. Clavel, C. and Callejas, Z., 2015. Sentiment analysis: from opinion mining to human-agent interaction. *IEEE Transactions on affective computing*, 7(1), pp.74-93.
12. Nahm, Y.E., 2013. A novel approach to prioritize customer requirements in QFD based on customer satisfaction function for customer-oriented product design. *Journal of Mechanical Science and Technology*, 27(12), pp.3765-3777.
13. Clavel, C. and Callejas, Z., 2015. Sentiment analysis: from opinion mining to human-agent interaction. *IEEE Transactions on affective computing*, 7(1), pp.74-93.
14. Yang, J.B., Wang, Y.M., Xu, D.L., Chin, K.S. and Chatton, L., 2012. Belief rule-based methodology for mapping consumer preferences and setting product targets. *Expert Systems with Applications*, 39(5), pp.4749-4759.
15. Decker, R. and Trusov, M., 2010. Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing*, 27(4), pp.293-307.
16. Dehkharghani, R., Mercan, H., Javeed, A. and Saygin, Y., 2014. Sentimental causal rule discovery from Twitter. *Expert Systems with Applications*, 41(10), pp.4950-4958.
17. Chen, A., Lu, Y. and Wang, B., 2017. Customers' purchase decision-making process in social commerce: a social learning perspective. *International Journal of Information Management*, 37(6), pp.627-638.
18. Uysal, A.K. and Gunal, S., 2014. The impact of preprocessing on text classification. *Information Processing & Management*, 50(1), pp.104-112.
19. Zhang, W., Yoshida, T. and Tang, X., 2011. A comparative study of TF\* IDF, LSI and multi-words for text classification. *Expert Systems with Applications*, 38(3), pp.2758-2765.
20. Zhao, T. and Eskenazi, M., 2016. Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning. *arXiv preprint arXiv:1606.02560*.