

DATA ANALYSIS OF CONSUMER COMPLAINTS IN BANKING INDUSTRY USING HYBRID CLUSTERING AND SENTIMENT ANALYSIS

R.Navin Kumar M.C.A.,M.Phil.,¹, S.Sujithkumar²

¹Assistant Professor, Department of Computer Applications, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India.

²Final MCA, Department of Computer Applications, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India.

Email: ¹navinsoccer07@gmail.com, ²smartsujith88@gmail.com

Abstract:

A consumer's complaints present bank or reporting agency with an opportunity to identify and rectify specific problems with their current product or service. The banks that are receiving customer complaints filed against them will analyze the complaint data to provide results on where the most complaints are being filed, what products/ services produce the most useful complaints and other data. This project assists banks in identifying the location and types of errors for resolution, leading to increased customer satisfaction to drive revenue and profitability. This project finds a correlation between complaints, companies and consumers to refine company applications to better accommodate consumer needs using k-means clustering. In addition, using SVM classification, the complaints sentiment values are analyzed and classified into positive or negative reviews. The project is designed using R Studio. The objectives of this study is: a) To give the estimated sentiment prediction of the subject based on the text reviews/complaints sent by the customers. b) To carry out Sentiment analysis so that the review is judged as either positive or negative. c) To find Percentage of positive/negative reviews. d) To give exact sentiment numerical values for various words and so classification such as positive or negative should be accurate. e) To apply neural network

such that it helps to classify the given loan request details into one of the predefined applied loans.

Keywords: Sentiment Analysis, SVM Classification, Machine Learning, Consumer Reviews.

I. INTRODUCTION

As we're apprehensive that in moment's ultramodern period people are more into business, so entering a complaint from a consumer happens nearly every day. A consumer's complaints present bank or reporting agency with an occasion to identify and amend specific problems with their current product or service. Service complaints operation is a critical part of business operation.

A good complaint- operation strategy will affect in stylish client relationship outgrowth with minimum mortal- resource investment and so hope to find a correlation between complaints, companies, and consumers to upgrade company operations to more accommodate consumer requirements. Decreasingly companies are feting the value of a client complaint in that it's feedback on their experience, and an occasion to not only resolve a problem for that particular client but maybe also for a much larger number

of guests and that leads to ineluctable quantities of data that has to be anatomized and specific functions are used to total the analysis results.

Clustering is regarded as a pivotal unsupervised literacy problem, that tries to search for analogous structures among an unlabeled data set. These analogous structure are data sets, generally appertained to as clusters. the information within every cluster is similar (or close) to factors within its cluster, and is different to (or fresh from) corridor that belong to indispensable clusters.

The mining ways' thing is to descry the natural grouping of a data set. In hierarchical clustering, a treelike cluster structure (dendrogram) is created through recursive partitioning (divisive styles) or combining (agglomerative) of being clusters, whereas in k-means clustering divides a cluster of k points with reference to a centroid, which helps if we're apprehensive of the data points that are probable and affair applicable. We hope to find a correlation between complaints, companies and consumers to upgrade company operations to more accommodate consumer requirements using k- means clustering.

The number of studies has been conducted regarding the services to guests and their mindfulness. As similar, we've reviewed some of them. Kamakodi (2007) concluded that ultramodern day generation is told by the calculation features used by banks and so the banks study about factors impacting their preferences. Residence relocation, payment change and attainability banking grounded services are reasons enough to change bank.

Uppal and Kaur (2007) determined how consumer's mindfulness of web disciplines used by banks and gave some measures to make these operations more successful. They concluded that the limitation about moment's web sphere operation is spreading the mindfulness about the varied features offered. Mishra and Jain (2007) took up confines of consumer satisfaction in public and private banks.

The study addresses about how satisfaction is the foremost asset to the association, which provides unmatched competitive edge that helps achieving fidelity of a client. They also spoke how high position of client satisfaction leads to fidelity. The study observed ten factors and five areas of satisfaction for both public and private sector bank.

II. LITERATURE REVIEW

Kamakodi (2007) concluded that ultramodern day generation is told by the calculation features used by banks and so the banks study about factors impacting their preferences. Residence relocation, payment change and attainability banking grounded services are reasons enough to change bank.

Uppal and Kaur (2007) determined how consumer's mindfulness of web disciplines used by banks and gave some measures to make these operations more successful. They concluded that the limitation about moment's web sphere operation is spreading the mindfulness about the varied features offered.

Mishra and Jain (2007) took up confines of consumer satisfaction in public and private banks. The study addresses about how satisfaction is the foremost asset to the association, which provides unmatched competitive edge that helps achieving fidelity of a client. They also spoke how high position of client satisfaction leads to fidelity. The study observed ten factors and five areas of satisfaction for both public and private sector bank.

Jain and Jain (2006) demonstrated that the banking sector, both private and public have suffered radical as well as revolutionary changed due to the liberalization act of 1991. Retail banking is the consumer preferred choice which articulates itself responses entered from 200 guests of HDFC bank, ICICI bank and some other banks in the megacity of Varanasi, Uttar Pradesh and he looked upon the schemes offered by the banks, quantized satisfaction in different types of services, prospects about these schemes and the height of segmentation among the services offered.

Singh (2006) discusses CRM approaches in colorful banks. He emphasized on how the operation targets guests in order to gain sapience and gives out value added services and products. Web as handed a smooth stoner experience, giving access to the colorful features used by the guests thereby achieving client satisfaction. Operation has to strive to insure end to end delivery and insure client satisfaction which is essential to the banks in terms of maintaining high respects and fidelity attained from guests.

Bhaskar (2004) reckoned that expansion of banking is directly commensurable to the quality of services handed by the banks and satisfaction is regarded largely as guests feedback is the only thing to lean on, when it comes to the largely competitive banking assiduity. Arguably, India's banking assiduity is largely thriving and depends heavily on client morale and fidelity.

Likewise, Hasanbanu (2004) stated how the pastoral India is ignorant about colorful schemes and benefits offered by the banks in order to insure fiscal weal. The maturity of pastoral population is inapproachable to the web sphere services of the banks and continue to prefer original moneylenders charging high interest rates. The study was conclusive and grounded on the data handed by the RBI, still, it's grounded on the questionnaire and checks.

Although Singh (2004) prodded about the reality of banks in terms of furnishing client support and plant out that the guests are told by the banks position and the tiniest detail of the banking details including the banking interest rates as well as stations and client support handed by the help.

Like all service diligence, the services handed within the hospitality sector have several effects in common which distinguish them from the products offered by manufacturing and other marketable sectors. These services are fairly impalpable, and are characterised by contemporaneous product and consumption. Therefore, it's delicate to observe hospitality services in advance and indeed harder to 'try before you buy'. The provision of hospitality

services is frequently immediate and robotic, counting on existential aspects for consumer evaluation.

For all of these reasons, achieving a situation of zero blights is delicate, maybe indeed insolvable. While seeking toward reducing the number of service failures remains an important thing, it's also important for hospitality and other service diligence to probe the stylish styles for recovering from service failure events. One way to deal with service failure events is to give an explanation to the displeased consumer as a way of dwindling negative comprehensions of the establishment. Despite the significance of explanations as a way to alleviate the negative goods of a service failure, fairly little exploration has been conducted into what works best and under what circumstances.

This paper documents a script- grounded study aimed at probing the part of accounts, or explanations, in the operation of a hospitality service failure event. Two different types of explanations linked from the literature, defense and referential, were examined in an experimental exploration design, which also manipulated the inflexibility of the failure and the thoroughness of the account. The end was to gain sapience into the effectiveness of different types of explanations and, more specifically, to examine how the impact of explanations on client satisfaction and intended fidelity varies with content (e.g., account thoroughness) and contextual (e.g., failure inflexibility) variables. Service failure, or service breakdown, can be defined as that service which doesn't meet the client's prospects.

There are several responses that guests may have to service failure or poor service reclamations. The most extensively delved are satisfaction responses, emotional responses (similar as wrathfulness, disappointment or remorse), and behavioral responses (similar as complaining, exiting or switching actions). Satisfaction is a extensively delved consumer evaluation because it's of the utmost Significance to enterprises, in order to insure sustainability. Another important behavioral consequence of

service failure relates to the repatronage of particular services. The long-term fidelity of guests is likely to be affected by service failure, especially where there's occasion to change service providers (that is, engage in switching geste). Indeed, Keaveney (1995) plant that the most frequent reason why people switched enterprises was the experience of service failures. In recent exploration, Ok and his associates (2005) demonstrated, in a eatery setting, that the manner in which a service failure event is managed eventually affects behavioral intentions including the liability of reprise visit.

The action taken by a hospitality establishment following a service failure can constitute a service recovery or, if shy, an alternate failure. The term 'service recovery' refers to the action taken by the service association in response to a client's expression of dissatisfaction with some aspect of the service. It's the process of dealing with a service failure situation with the points of returning the client to a state of satisfaction. Service recovery ways generally involve attempts to amend the service breakdown but may also number furnishing guests with explanations about the service failure, .apologizing, making offers of compensation, and being gracious in the process (see for illustration, Blodgett, Hill & Tax, 1997; Butcher & Heffermen, 2006; Karatepe, 2006; Matilla, 1999; Sparks & Callan, 1996). This paper seeks to probe the part of explanations in managing a service failure event.

III. PROPOSED METHODOLOGY

In existing system, K-Means clustering of bank customer data is used to analyze and group the customers. However, from their reviews/complaints the seriousness of the complaint type could not be identified. The sentiment values could not be found out to measure the exactness of the complaints. Only the customers can be grouped into three, four and more groups based on their similarity among the complaint records.

The existing system does not give the estimated sentiment prediction of the subject based on the text reviews/complaints sent by the customers.

- Sentiment analysis is not carried out so that the review is judged as either positive or negative.
- Percentage of positive/negative reviews could not be found out.
- Words are not given with exact sentiment numerical values and so classification such as positive or negative is not accurate.

In proposed system, all the existing methodology is carried out. The proposed approach takes input from the data set created by accumulating all the text messages send by the customers. All the messages may be from different banking reviews like loan processing delay, feedbacks etc. It also gives a highly efficient method of finding the sentiment of the customer by analyzing the text reviews and also processing emoticons. Using sentiment values for all words, the overall sentiment value for the paragraphs are calculated and so the polarity among the reviews is found out.

This project work focus on analyzing and predicting if the given loan will be fully repaid or not using Artificial Neural Network. This data set fall under the complaints of Annual income, Interest rate, Loan status, Duration term, Verification status etc.,. Prediction is made for better accuracy and for the practical users. In this paper prediction is made for if the loan will be fully repaid or not, and also analyzing what kind of situation does a customer is unable to pay loan. This prediction will lead to know more about the customer whether the loan can be given to them or not.

IV. FINDINGS

- The proposed system gives the estimated sentiment prediction of the subject based on the text reviews/complaints sent by the customers.
- Sentiment analysis is carried out so that the review is judged as either positive or negative.
- Percentage of positive/negative reviews can be found out.
- Various words are given with exact sentiment numerical values and so classification such as positive or negative is accurate.
- Neural network helps to classify the given loan request details into one of the predefined applied loans.

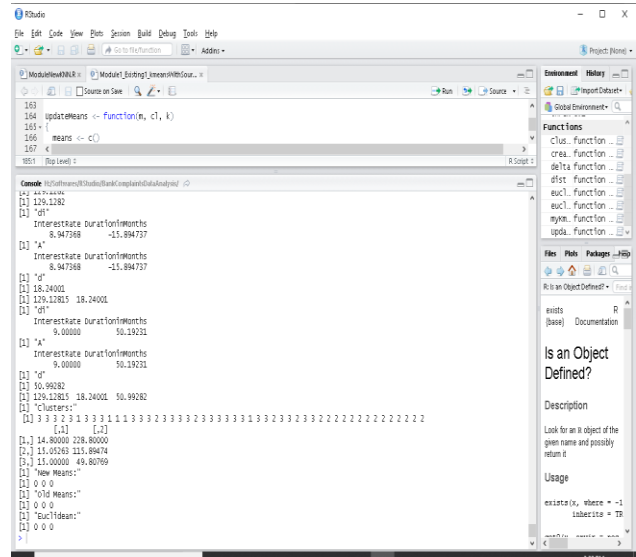


FIG 4.2 RECORDS GROUPED IN CLUSTERS

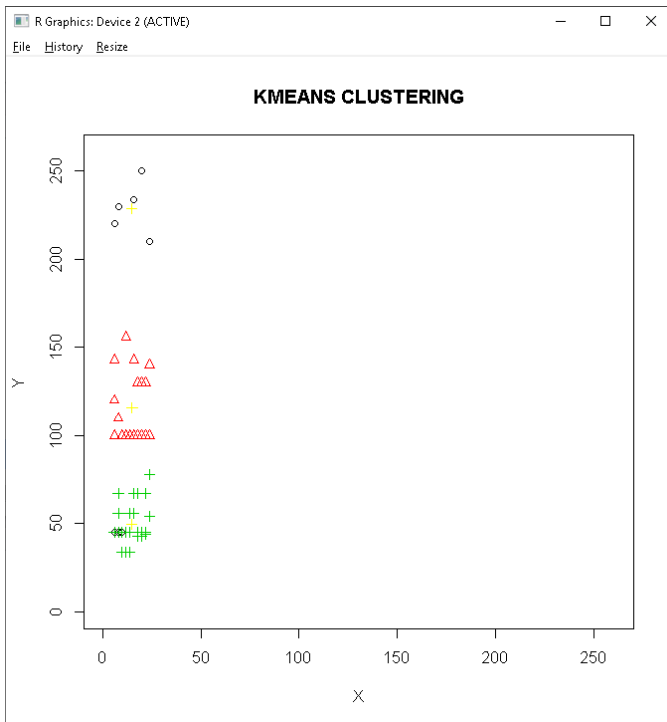


FIG 4.1 KMEANS CLUSTERING SAMPLE

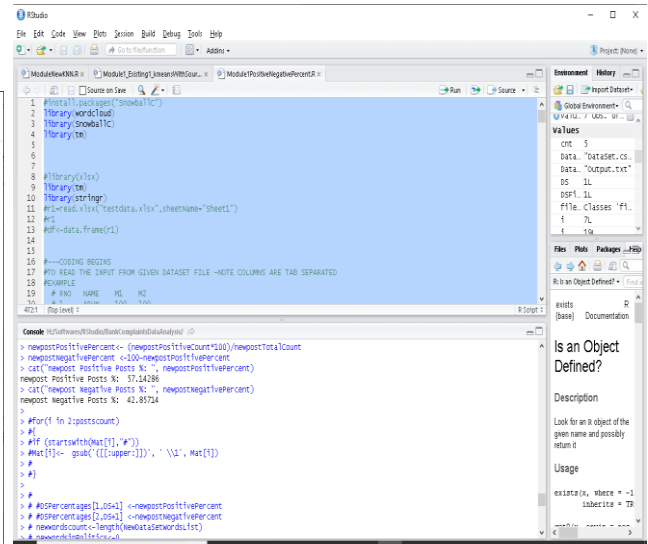


FIG 4.3 POSITIVE/NEGATIVE POST PERCENTAGE

