Efficient Automatic Reviewing System for New Product Development

B.Pavithra, Mrs. R.Kavitha M.E
P.G Student, Assistant Professor
Department of CSE, A.R.J.College of Engineering and Technology,
Mannargudi, Tamilnadu.

ABSTRACT: **Text summarization is a process where the most salient features of a text are extracted and compiled into a short abstract of the original document.In this paper, both an unsupervised and a supervised method are proposed that are able to find aspect categoriebased on co-occurrenc frequencies.First, identify for each of the given categories a set of seed words containing the category word and any synonyms of that word.Next, as a natural language preprocessing step, both training and test data.Then it Construct the Co-occurrence matrix and co-occurrence digraph.Next, applying the Mine Association Rules. In the last step we predict categories for each unprocessed sentence.**

Keywords:**Radar,Bloodpressure,IOT**

I. INTRODUCTION

Customer integration into new product development (NPD) is a high-priority topic in the innovation management field. Integrating customers into NPD is generally believed to provide vital information about customer needs and requirements because innovations that satisfy customer needs are likely to sell. The value of integrating customers is widely accepted for incrementally new products (INPs), while it is less clear regarding radically new products (RNPs). Practitioners often claim that customers are too conservative and lack the necessary skills and knowledge to contribute to the successful development of RNPs. For example, Steve Jobs allegedly built his most innovative products by ignoring the voices of current and potential customers allegedly based on his belief that market research is unsuitable for RNPs. Seagate, a dominate player on the 5.25-inch drive market in the 1970s and 1980s, briefly looked into the 3.5-inch drives market in the early 1980s, asked their current desktop customers.

whether they found these drives attractive, received negative response, and abandoned the idea. They did not study the relevance of these drives for potential customers outside their current markets, such as laptop manufacturers, and it turned out that the latter customers found the 3.5-inch drives attractive, and thus Seagate left tremendous growth opportunities to their competitors.

The rapid growth of the Internet, users' ability to publish content has created active electronic communities that provide a wealth of product information. Consumers naturally gravitate to reading reviews in order to decide whether to buy a product. However, the high volume of reviews that are typically published for a single product makes it harder for individuals to locate the best reviews and understand the true underlying quality of a product based on the reviews.

Similarly, the manufacturer of a product wants to identify the reviews that influence the customer base, and examine the content of these reviews. In this project propose two ranking mechanisms for ranking product reviews: a consumer-oriented ranking mechanism ranks the reviews according to their expected helpfulness, and a manufacturer-oriented ranking mechanism ranks the reviews according to their expected effect on sales. Our ranking mechanism combines econometric analysis with text mining techniques in general, with subjectivity analysis in particular. It shows that subjectivity analysis can give useful clues about the helpfulness of a review and about its impact on sales. Our results can have several implications for the market design of online opinion forums.

II. EXISTING SYSTEM:

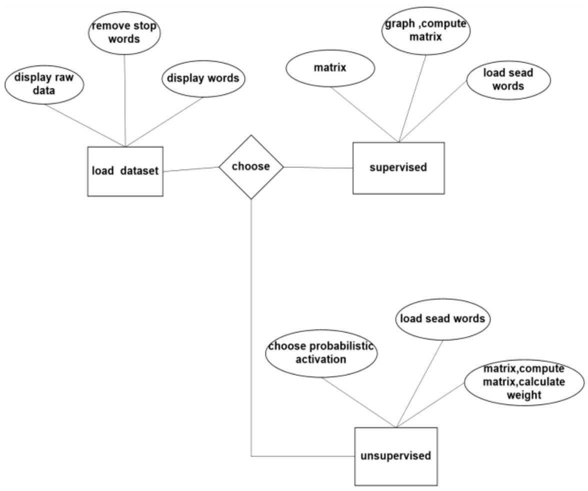
In existing approach, suggested that simultaneously and iteratively clusters product aspects and opinion words. Aspects/opinion words with high similarity are clustered together, and aspects/opinion words from different clusters are dissimilar. The similarity between two aspects/opinion words is measured by fusing both homogeneous similarity between the aspects/opinion words (content information), calculated by traditional approach, and similarity by their respective heterogeneous relationships they have with the opinion words/aspects (link information). Based on the product aspect categories and opinion word groups, a sentiment association set between the two groups is then constructed by identifying the strongest *n* sentiment links. Unfortunately, there were no quantitative.

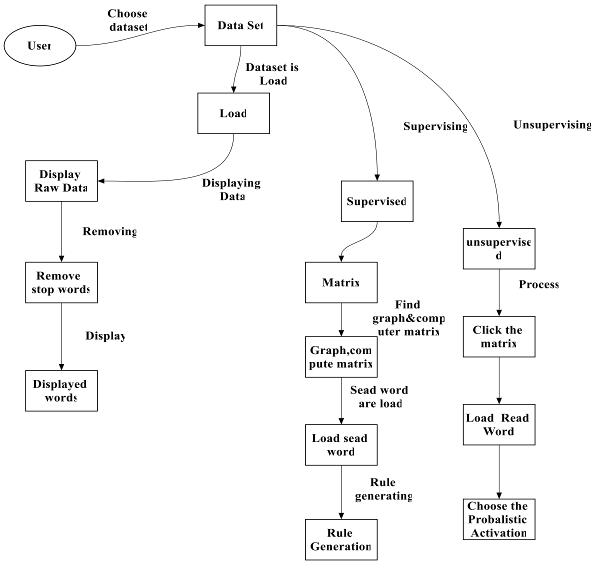
experimental results reported, specifically for implicit aspect identification

III. PROPOSED SYSTEM

In this proposed system, both an unsupervised and a supervised method are proposed that are able to find aspect categories based on co-occurrence frequencies. The unsupervised method uses spreading activation on a graph built from word co-occurrence frequencies in order to detect aspect categories. In our project summarize the text into following steps. (a) Identify category seed word sets. (b) Determine co-occurrence digraph. (c) Apply spreading activation. (d) Mine association rules. (e) Assign aspect categories. First, we identify for each of the given categories a set of seed words containing the category word and any synonyms of that word. Next, as a natural language preprocessing step, both training and test data. Then it Construct the Co-occurrence matrix and co-occurrence digraph. Next, Applying the Mine Association Rules. In addition to counting the co-occurrences of lemmas and aspect categories, the co-occurrences between grammatical dependencies and aspect categories are also counted. Similar to lemmas, low frequency dependencies are not taken into account to prevent over fitting, using the parameter α D. The added information provided by dependencies, may provide more accurate predictions, when it comes to category detection. A deep hypergraph model based on word embeddings clustering and ANN proposed which can capture the high-level features and reflect the high-order relations among samples. Propose an improved task specific hierarchical clustering algorithm based on density peaks searching for semantic clustering of word embeddings. Semantic units are detected with considering the central words, which maximally preserves original information of reviews for improving sentiment classification accuracy

IV.SYSTEM ARCHITECTURE





4.1 Data set collection

In this project, collect the sentiment related dataset. Social sentiment is a way of measuring the emotions behind social media mentions. It is a manner in which you can measure the tone of the conversation that's taking place is this person satisfied, happy, angry, or annoyed? It's not enough to know that something is trending. Sentiment adds context to social media

4.2 Preprocessing

Data preprocessing describes any type of processing performed on raw data to prepare it for another processing procedure. Commonly used as a preliminary data mining practice, data preprocessing transforms the data into a format that will be more easily and effectively processed for the purpose of the user.

4.3 Matrix process and calculate weighted for co-occurrence terms

To find the co-occurrence terms and form the matrix. Association rules are mined when a strong relation between a notional word and one of the aspect categories exists, with the strength of the relation being modeled using the co-occurrence frequency between category and notional word. The data network structure used for the spreading activation algorithm will consist of vertices that represent the notional words, and links between two vertices representing strictly positive co-occurrence frequency. Then calculated the weighted for user opinion.

4.4 Unsupervised learning

The proposed unsupervised method (called the spreading activation method) uses co-occurrence association rule mining in a similar way as, by learning relevant rules between notional words, defined as the words in the sentence after removing stop words and low frequency words, and the considered categories. This enables the algorithm to imply a category based on

4.5 Supervised learning and deep hypergraph

The supervised method (called the probabilistic activation method) employs co-occurrence association rule mining to detect categories. It borrows the idea from to count co-occurrence frequencies between lemmas and the annotated categories of a sentence. However, low frequency words are not taken into account in order to prevent over fitting. For the supervised method we use the training set to learn the parameters and co-occurrence frequencies, after which we evaluate the method on the test set. The training set provides the annotated categories of each sentence. The training set is then evaluated for a range of values of thresholds. In addition to counting the co-occurrences of lemmas and aspect categories, the co-occurrences between grammatical dependencies and aspect categories are also counted. Similar to lemmas, low frequency dependencies are not taken into account to prevent over fitting, using the parameter α D. The added information provided by dependencies, may provide more accurate predictions, when it comes to category detection

V. CONCLUSION

Unsupervised method uses spreading activation over a graph built from word co-occurrence data, enabling the use of both direct and indirect relations between words. May be this results in every word having an activation value for each category that represents how likely it is to imply that category. In supervised method, use the training set to learn the parameters and co-occurrence frequencies, after which evaluate the method on the test set. To see the impact the dependency indicators have, this method is executed separately for the dependency indicators, lemma indicators and a combined version where both lemma and dependency indicators are used, and evaluated on the test set.

VI.REFERENCES

[1] P. F. Bone, "Word-of-mouth effects on short-term and long-term productjudgments," J. Bus. Res., vol. 32, no. 3, pp. 213–223, 1995.

[2] R. Feldman, "Techniques and applications for sentiment analysis,"Commun. ACM, vol. 56, no. 4, pp. 82–89, 2013.

[3] S. Sen and D. Lerman, "Why are you telling me this? An examinationinto negative consumer reviews on the Web," J. Interact. Marketing,vol. 21, no. 4, pp. 76–94, 2007.

[4] B. Bickart and R. M. Shindler, "Internet forums as influential sourcesof consumer information," J. Consum. Res., vol. 15, no. 3, pp. 31–40,2001.

[5] D. Smith, S. Menon, and K. Sivakumar, "Online peer and editorialrecommendations, trust, and choice in virtual markets," J. Interact.Marketing, vol. 19, no. 3, pp. 15–37, 2005.

[6] M. Trusov, R. E. Bucklin, and K. Pauwels, "Effects of word-of-mouthversus traditional marketing: Findings from an Internet social networkingsite," J. Marketing, vol. 73, no. 5, pp. 90–102, 2009.

[7] M. T. Adjei, S. M. Noble, and C. H. Noble, "The influence ofC2C communications in online brand communities on customer purchasebehavior," J. Acad. Marketing Sci., vol. 38, no. 5, pp. 634–653,2010.

[8] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Found.Trends Inf. Retrieval, vol. 2, nos. 1–2, pp. 1–135, 2008.

[9] C.-L. Liu, W.-H. Hsaio, C.-H. Lee, G.-C. Lu, and E. Jou, "Movie ratingand review summarization in mobile environment," IEEE Trans. Syst.,Man, Cybern. C, Appl. Rev., vol. 42, no. 3, pp. 397–407, May 2012.

[10] M. Pontikiet al., "SemEval-2014 Task 4: Aspect based sentiment analysis,"in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin,Ireland, 2014, pp. 27–35.

[11] S. Kiritchenko, X. Zhu, C. Cherry, and S. M. Mohammad, "NRCCananda-2014: Detecting aspects and sentiment in customer reviews,"in Proc. 8th Int. Workshop Semantic