International Journal of Recent Engineering Research and Development (IJRERD)
ISSN: 2455-8761
www.ijrerd.com || Volume 09 – Issue 04 || Jul - Aug 2024 || PP. 42-47

Efficient Automated Evaluation System for Product Innovation

Pavithra[1], R. Kavitha[2], M. Sheela[3]

[1]*PG Scholar, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India*

[2,3]*Assistant Professor, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India*

Abstract: Text summarization is emerged as an important research area. 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 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 and rule generation. A supervised learning and deep hypergraph are predicting the result for user opinion. In our project summarize the text into following steps.

Keywords: Word of mouth, new product development, sentiment analysis, RNP.

1. Introduction

Integrating customers into new product development (NPD) is highly valued for gaining insights into customer needs, especially for incrementally new products (INPs), where understanding current market demands is essential for success. However, its effectiveness is debated for radically new products (RNPs), with some arguing that customers may not provide useful input for breakthrough innovations. For instance, Steve Jobs famously disregarded customer feedback, believing it was inadequate for pioneering products. Similarly, Seagate's decision to abandon the 3.5-inch drive market based on negative responses from existing customers illustrates the pitfalls of not considering potential new market segments, leading to lost opportunities as competitors seized the demand from emerging markets like laptops.

2. Literature Survey

Bansal, et. al; (2000),Word of mouth (WOM) has been one of theextensive methods for product development and there are literally many studies on learning moreabout the effects of various factors on marketing planning. Bone investigatedthe effect of WOM communications on product judgments. In addition, he studied the moderatingeffect of several situational, personal, and source characteristics in three different experiments. Thestudy indicated that WOM impacts on short-term and long-term judgments. This effect stated to bebigger when a consumer faces a disconfirmation experience and when an expert monitors the WOMcommunication.

Brown, et. al; (2007), reported the results of a two-stage study aimed at studying online WOM where a set of in-depth qualitative interviews followed by a social network analysis of a single online community. Combined, the results provided some evidence that individuals may behave as if Web sites themselves were primary "actors" in online social networks and that online communities could act as a social proxy for individual identification. They offered a conceptualization of online social networks, which could take the Web site into account as an actor,

Davis, et. al; (2008), investigated the effect of online WOM attributes and other related factors on e-commerce sales on a multi-product retail e-commerce firm. They reported that the introduction of online WOM on a retail e-commerce site could positively influence on product sales. They proposed and validated a conceptual model of online WOM and its effect on product sales and the effect of moderator variables such as promotion, product category and product views. In their survey, pure increase in volume or number of reviewer comments had no substantial impact on sales.

Miao, et. al: (2012), Electronic word of mouth (eWOM) is available to customers in various kinds of online consumer reviews used to help them make e-commerce purchasing decisions. Customers acknowledge that online consumer reviews could help them determine eWOM credibility and making purchasing decisions. This study used surveys and multiple regression analysis to generate an extended Elaboration Likelihood Model, which describes the relationship between customer expertise and involvement to acceptance and use of eWOM in making purchasing decisions.

3. Proposed System

The 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 considered 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.

4. System Architecture

The series you've provided outlines a workflow commonly used in data analysis and machine learning,particularly in the context of natural language processing (NLP) and text mining. Let's walk through each step with a brief explanation:

1. **Dataset:** This is the raw data collected for analysis. It could be in various forms, such as text, numbers, images, etc.
2. **Preprocessing:** Involves cleaning and preparing the data for analysis. This can include steps such as removing noise, handling missing values, normalizing data, tokenizing text, and converting categorical variables into numerical formats.
3. **ODA (Observational Data Analysis):** This step involves analyzing the data to uncover patterns, relationships, and trends without making causal inferences. It includes exploratory data analysis (EDA), visualizations, and descriptive statistics to understand the data better.
4. **Matrix Process:** This refers to organizing data into a matrix format, which is essential in many machine learning and data analysis tasks. In text mining, this might involve creating a term-document matrix (TDM) or a document-term matrix (DTM), where rows represent documents and columns represent terms.

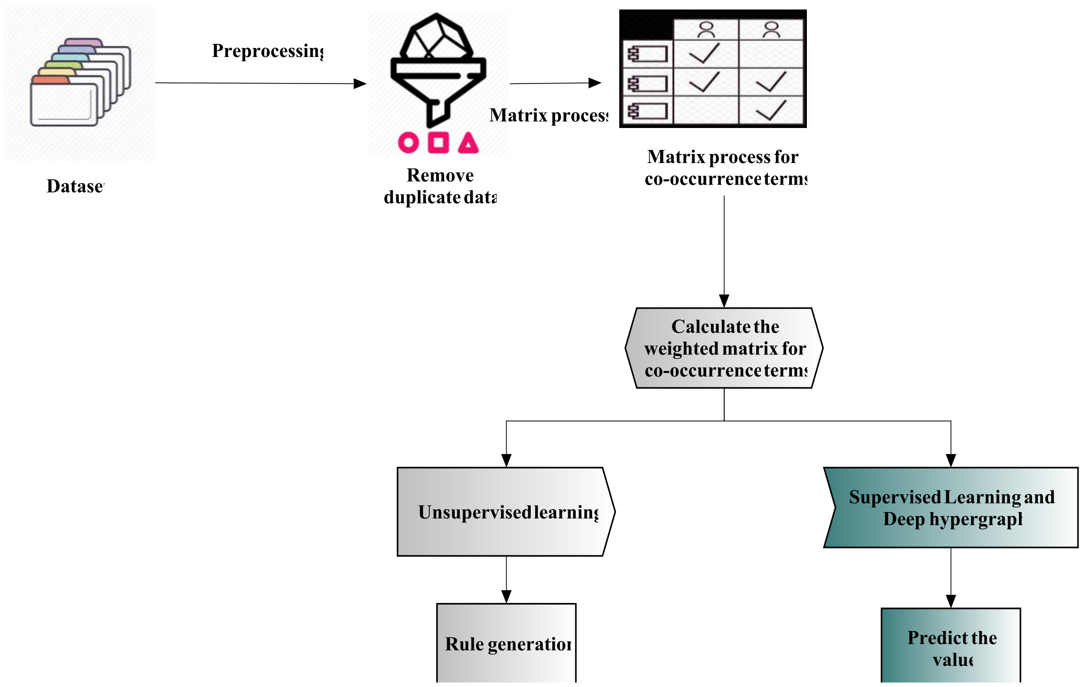

Figure 1: System architecture

5. **Remove Duplicate Data:** Duplicate data can skew results and reduce the efficiency of data processing. This step involves identifying and removing duplicate entries from the dataset to ensure the analysis is based on unique observations.
6. **Unsupervised Learning:** A type of machine learning where the algorithm is trained on unlabeled data. Common techniques include clustering (e.g., K-means), association rule learning, and dimensionality reduction (e.g., PCA). These methods help uncover hidden patterns and relationships in the data.
7. **Rule Generation:** Involves generating rules that describe the relationships or patterns found in the data. This is common in association rule learning, where rules like "If A, then B" are discovered (e.g., in market basket analysis).
8. **Matrix Process for Co-occurrence Terms:** This step focuses on creating a matrix that captures the co-occurrence of terms within the data. In text analysis, this could be a co-occurrence matrix that shows how often pairs of words appear together, which is useful for understanding semantic relationships.
9. **Calculate the Weighted Matrix for Co-occurrence Terms:** Involves assigning weights to the co-occurrence terms, which could be based on frequency, importance, or other metrics. The weighted matrix helps to prioritize or highlight significant relationships among the terms.
10. **Supervised Learning and Deep Hypergraph:** In supervised learning, the model is trained on labeled data to learn the mapping from inputs to outputs. A "deep hypergraph" may refer to a complex network structure that represents relationships in the data, where nodes and hyperedges connect multiple nodes. This can be used in deep learning models to capture intricate patterns.
11. **Predict the Value:** The final step involves using the trained model to predict outcomes or values based on new data. This can include predicting classifications, regressions, or other metrics depending on the problem at hand.

5. Use Case Diagram

The use case diagram for the efficient automated evaluation system for product innovation illustrates interactions between different actors and the system's functionalities. Data Analysts are responsible for collecting and preprocessing data, while Product Managers and Marketing Specialists use the system to analyze data, generate rules, and predict outcomes.Machine Learning Engineers develop and apply models to forecast product success, and Customers provide feedback that is incorporated into the system. The diagram highlights how each actor engages with use cases like data integration, rule generation, and predictive modeling, ensuring a streamlined process for product development and launch.

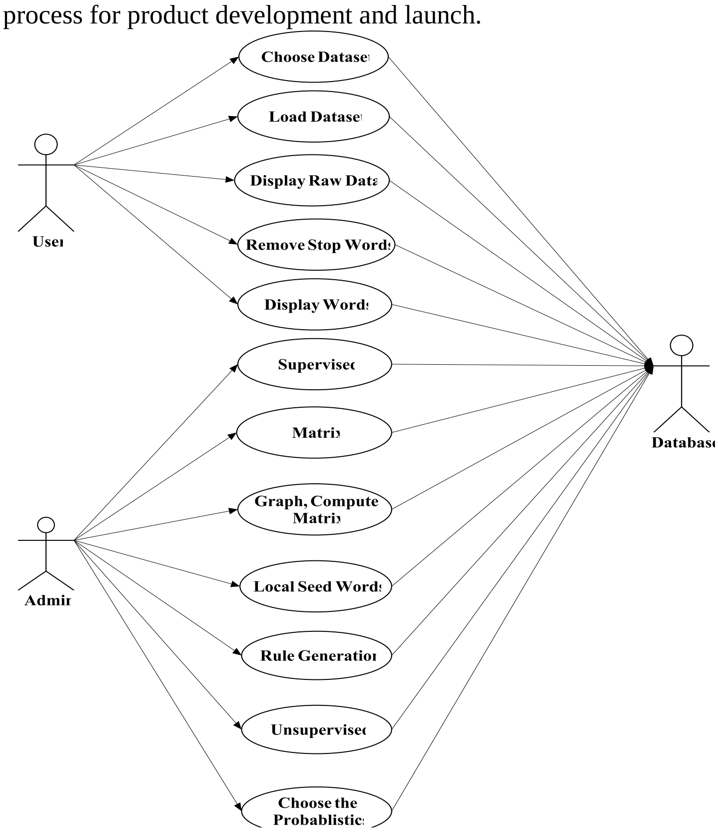

Figure 2: Use Case Diagram

6. System Implementation

1) Dataset Collection:

The dataset collection phase involves gathering diverse and comprehensive data from various sources. This includes customer feedback, market trends, competitor analysis, and internal performance metrics. The goal is to compile a rich and relevant dataset that provides a holistic view of customer needs, market dynamics, and potential opportunities for new product development.

2) Preprocessing:

During the preprocessing phase, the collected data is cleaned and organized to ensure accuracy and consistency. This involves removing duplicates, handling missing values, normalizing data formats, and structuring data into a format suitable for analysis. Preprocessing prepares the data for further analytical processes by ensuring it is free from errors and inconsistencies.

3) Matrix Process and Calculation of Weighted Co-occurrence Terms:

In this stage, the data is transformed into matrices, such as Term-Document Matrices, to facilitate detailed analysis. The process includes calculating the weighted co-occurrence terms, which quantifies the strength and significance of relationships between terms or features. This helps in identifying key patterns, associations, and trends within the data.

4) Unsupervised Learning:

Unsupervised learning techniques are applied to discover hidden patterns and structures within the data. Methods such as clustering and dimensionality reduction are used to group similar data points and uncover underlying trends. This stage provides insights into customer segments, emerging trends, and potential areas for innovation that may not be immediately apparent.

5) Supervised Learning and Deep Hypergraph Analysis:

The final stage involves applying supervised learning algorithms and deep hypergraph analysis to make predictions and validate hypotheses. Supervised learning models are trained on labeled data to forecast customer preferences, sales outcomes, and market responses. Deep hypergraph analysis captures complex relationships within the data, enhancing the accuracy of predictions and providing a nuanced understanding of the data's interdependencies.

6. Result

The framework for an efficient automated evaluation system for product innovationintegrates data collection, preprocessing, unsupervised learning, and predictive modeling to streamline the identification, development, and launch of new products. It leverages diverse data sources, cleans and organizes data into useful matrices, and applies machine learning techniques to uncover market trends, segment customers, and predict sales outcomes. This process eliminates duplicates, prioritizes important features through weighted matrices, and uses advanced models like deep hypergraphs to capture complex relationships. The result is a data-driven approach that enhances decision-making, reduces time to market, and increases the likelihood of successful product launches.

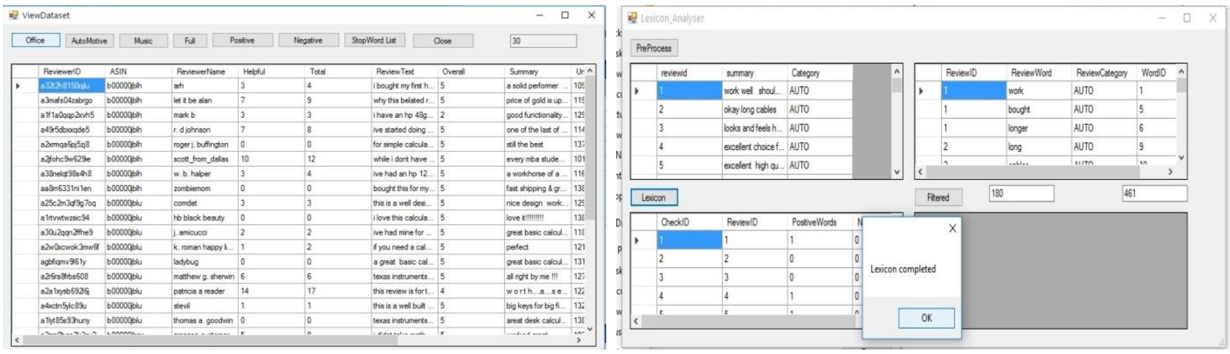

Figure 3: Dataset and Preprocessing

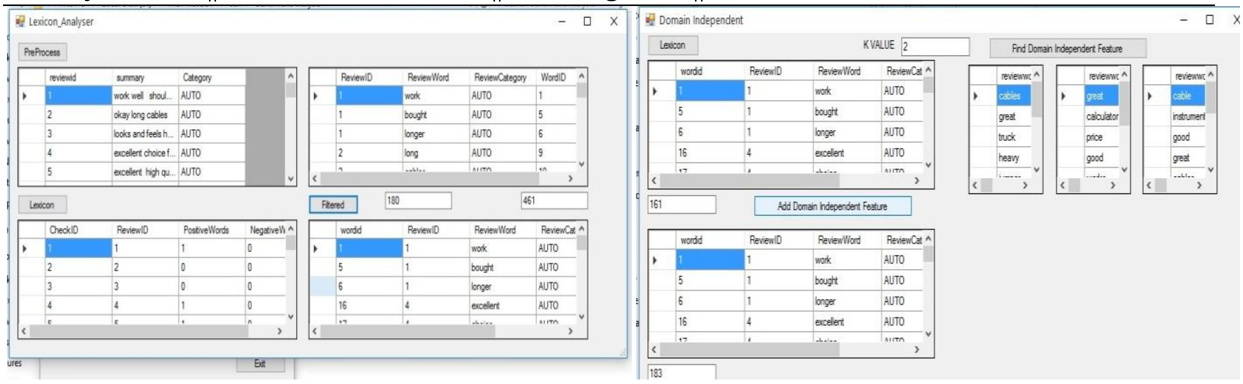

Figure 4: Analysis and Domain Independent

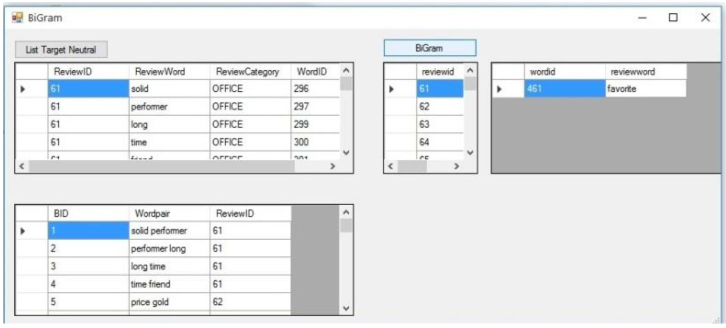

Figure 5: Bigram

7. Conclusion

The importance of integrating advanced data analytics and machine learning techniques to streamline and enhance the product development process. The proposed system leverages a multi-stage framework to collect and analyze customer feedback and market data efficiently. By systematically applying Dataset Collection, Preprocessing, Matrix Processing and Calculation of Weighted Co-occurrence Terms, Unsupervised Learning, and Supervised Learning with Deep Hypergraph Analysis, the system transforms raw data into actionable insights.The efficient automatic reviewing system provides a robust mechanism for identifying customer needs, emerging trends, and market opportunities. The integration of these advanced methodologies ensures that new products are developed with a strong alignment to market demands and customer expectations. By automating the review process, the system significantly reduces the time and effort required for manual data analysis while improving accuracy and relevance in decision-making.Overall, the system facilitates a data-driven approach to new product development, enhancing innovation, reducing risks, and optimizing the product launch process. This strategic integration of customer insights and predictive analytics positions organizations to more effectively meet market demands and achieve competitive advantage.

8. References

[1]. P. F. Bone, “Word-of-mouth effects on short-term and long-term product judgments,” 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 examination into 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 sources of consumer information,” J. Consum. Res., vol. 15, no. 3, pp. 31–40, 2001.

[5]. D. Smith, S. Menon, and K. Sivakumar, “Online peer and editorial recommendations, 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-mouth versus traditional marketing: Findings from an Internet social networking site,” J. Marketing, vol. 73, no. 5, pp. 90–102, 2009.

[7]. M T. Adjei, S. M. Noble, and C. H. Noble, “The influence ofC 2C communications in online brand communities on customer purchase behavior,” 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 rating and 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. (Sem Eval), Dublin, Ireland, 2014, pp. 27–35.