MJRET
Open Access
Multidisciplinary Journal of Research in Engineering and Technology
Journal homepage: www.mjret.in
ISSN:2348 - 6953

A SURVEY ON EFFICIENT DATA MINING METHOD FOR FINDING COMPETITORS FROM LARGE UNSTRUCTURED E-COMMERCE DATA

Komal Ghadage[1], Dr. Sunil Rahod[2]
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering Lohgoan,
Savitribai Phule, Pune University
Pune, India

Abstract:*In this new era any competitive business and success is depending on the ability to make an item more attractive to the customer than the competition. A number of the questions are coming in this task such as first question is that: A) Who are main competitors of the given items? B) How to formulize and quantify competitiveness between items? And C) What are different features of an item that most affect its competitiveness? Solutions of these problems are available on many domains but limited amount of work has been carried out for this problems. Here, we are presenting formal definition of competitiveness between the two different items, which is based on the market segments they can both cover and which are validated for qualitatively and quantitatively. Finally from the various surveys, we find the conclusion of basic significance of competitiveness between two items on the basis of market segments.*

Keywords*: Competitiveness, Qualitatively, Quantitatively and Business.*

1. INTRODUCTION

The identification of competitors serves as an important fact in the various areas. In industrial organization economics, this involves the task of defining markets, which is the crucial for the regulatory and antitrust policy. In marketing, it supports the analysis of pricing policies, product design, development and positioning, communication strategy and distribution channels. Every company has competition, and potential entrepreneurs

ignore competitors at their own risk. If a company does not have an absolute monopoly on a vital product, there are competitors offering replacement products and services. In any business plan competitor analysis is the important requirement because (a) The organization's competitive position in "market space" is demonstrated, (b) Partners and readers of business plan assume it and (C) Assists you develop strategies to be competitive. The main objectives of the competitor analysis and execution are processes for identify the information needs and key competitor, gathering the relative information and interpreting that information.

The management and marketing community has been focus on the empirical methods for the analyze competitors [1]. The extensive research has been focus on the find examples of comparative expressions: "Article A is the better than the Article B" from different Websites or other text sources. The paradigm of competitiveness is mainly based on the following observation: The competitiveness between the two different factors depends on whether they compete for the attention and business of the same customer groups (i.e the same market segments).

For example, two restaurants that are exists in the different countries are obviously not competitive, since there is no overlap between their target groups. Consider the example shown in Figure 1. Figure shows that the competitiveness between the three items X, Y, and Z. Each item is mapped to the set of features that it can offer to a customer. Three features are considered in this example: A, B and C. Although this simple example only considers binary functions like available / unavailable. The actual formalization covers a much wider range, including binary, categorical, and numeric functions. Users are grouping with the preferences in terms of features. For example, customers in G2 are only interested in features B and C. The points X and Z are not competitive because they simply do not address the same customer groups. Y competes with both X (for groups G1 and g2) and Z (for G3).

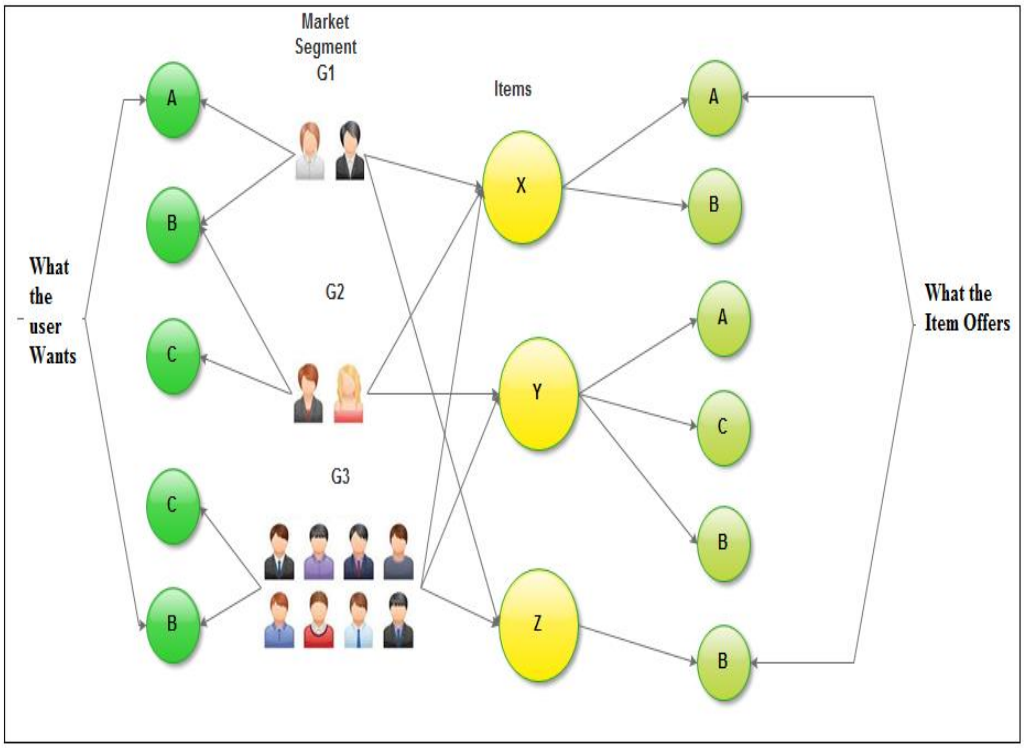

Fig.1 Competitiveness Paradigm

2. LITERATURE REVIEW

Here we discussed the literature review of existing techniques:

Data mining is a way of handling huge amount of information for mining competitors. In this paper authors present an efficient method for competitiveness in the larger review datasets. Here they describe the finding the top-k competitor's problem [1].

In this paper [2], they describe the Bing Liu's aspect based opinion mining technique. This technique applied on tourism domain. Using this technique discovers consumer preferences about the tourism products for hotels and restaurants. The result shows that important information available on web sites about customer preferences and it is accessed by the opinion mining approach.

In this paper [3], they propose techniques for business identify the competitors is a very important. To signify competitor relationships, in this paper propose a method that uses machine learning techniques and graph theoretic measures. Here they evaluate an approach to use corporate citations in online news to create an intercompany network whose structural features serve to derive competitive relationships between companies.

In [4] this paper they propose techniques for the mining competitors automatically from the Web. Here the CoMiner algorithm is proposed for mining all information related to competitors, competitors' strength and competing domains. The algorithm is used to conduct a web scale mining in a domain independent manner.

In this paper [5], they describe the techniques of mining the competitive information from the web. In this paper a Cminer algorithm is proposed. This algorithm first access a set of comparative candidates of the input entity and after that ranks according to comparability and then finally extracting different competitive fields.

In this paper [6], the author proposes a novel graphical model. Using this model to access and visualize the relationship between the customer reviews and products. The result shows that the proposed method extracts comparative relations more accurately.

The outcome of the above study is summarized in the table below.

Sr. No	**Paper Name**	**Author**	**Method Proposed**	**Limitations**
1.	Mining Competitors from Large Unstructured Datasets	George Valkanas, , and Dimitrios Gunopulos	Present efficient method for competitiveness in large review datasets	Dependency on transactional data.
2.	Identifying customer preferences about tourism products using an aspect-based opinion mining approach	Marrese Taylor, and Y. Matsuo	Describes the Bing Liu's aspect based opinion mining technique. This technique applied on tourism domain.	the algorithms were only capable of extracting 35% of the explicit aspect expressions. Less Accurate.
3.	Mining competitor relationships from online news: A network-based approach	Z. Ma, G. Pant, and O. R. L. Sheng	They propose company citations in online news for creating an intercompany network and structural attributes are used for infer competitor relationships between two companies.	Performance of this method is not good.
4.	Competitor mining with the web	S. Bao, and Y. Cao	Focuses on problem in the mining competitors from the different websited automatically. Proposes a CoMiner algorithm.	It only utilizes the most reliable direct competitive information

5.	Cominer: An effectivealgorithm for mining competitors from the web	R. Li, S. Bao, J. Wang, Y. Yu, and Y. Cao	Cminer algorithm is proposed. This algorithm first access a set of comparative candidates of the input entity and after that ranks according to comparability and then finally extracts the competitive fields.	Due to space limitation they only present 28 entity results.
6.	Mining comparative opinions from customer reviews for competitive intelligence	K. Xu, S. S. Liao, J. Li, and Y. Song	Proposes a novel graphical model. Using this model to access and visualize relationship between products from customer reviews.	The requirement of manually compiling rules makes this method difficult to adapt to new domains.

Table 1: Comparative Analysis

3. TAXANOMY CHART

Theme	**Mining Competitors from Large Unstructured Datasets**	**Mining Competitors from Large Unstructured Datasets**	**Identifying customer preferences about tourism products using an aspect-based opinion mining approach**	**Mining competitor relationships from online news: A network-based approach**	**Competitor mining with the web**
Competitiveness	✓	✓	✓	✓	✓
Nearest Neighbor (NN) Approach	✓	✗	✗	✓	✗
Collaborative filtering (CF)	✓	✗	✓	✗	✓

Naive algorithm	✓	✗	✓	✗	✓
Mining Algorithms	✓	✓	✗	✓	✓

Table 3: Taxonomy Chart

4. DISCUSSION

Data mining is a way of handling huge amount of information for mining competitors. It reviews information about client opinion and interest in producing the products. For competitive products, it's very difficult to analyze different reviews on different websites. Competitive intelligence first classifies the potential risk and opportunities to gather contextual information to help the manager make tactical decisions for an organization. Data Mining is important for finding examples, assessing and disclosing learning, etc. in different business areas. Machine learning is widely used as part of various applications. Every business-related application uses information mining systems. For the improving such business or to give the customer a suitable competitor, the help of web mining systems is required. Competitive degradation is one such approach to inspecting competitors for the preferred items.

5. CONCLUSION & FUTURE SCOPE

Research has demonstrated the strategic importance of identifying and monitoring a firm's competitors. This research has focuses on the mining comparatives expressions of the items e.g. "Item A is better than the Item B" from the different websites. Such expression indicates the competitiveness. Here we study techniques for formalization of the competitiveness between the two items, which is depends on the marketing segments. This technique assumes that the user requirements are uniformly distributed within the value space of each feature. This approach is based on the assumption that such comparative evidence can be found in abundance in the available data.

6. ACKNOWLEGEMENT

The authors would like to thank the researchers as well as publishers for making their resources available and teacher's for their guidance. We are also thankful to the reviewer for their valuable suggestions.

REFERENCES

[1] George Valkanas, and Dimitrios Gunopulos, "Mining competitors from large unstructured datasets", 2016.

[2] E. Marrese Taylor, and Y. Matsuo, "Identifying Customer Preferences About Tourism Products Using An Aspect-Based Opinion Mining Approach," 2013.

[3] Z. Ma, and O. R. L. Sheng, "Mining competitor relationships from online news: A network-based approach," 2011.

[4] S. Bao, R. Li, Y. Yu, and Y. Cao, "Competitor mining with the web," 2008.

[5] R. Li, S. Bao, J. Wang, Y. Yu, and Y. Cao, "Cominer: An effectivealgorithm for mining competitors from the web," 2006.

[6] K. Xu, S. S. Liao, J. Li, and Y. Song, "Mining comparative opinions from customer reviews for competitive intelligence", 2011.