ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Modified Latent Semantic Index based Product Catalog Pattern Matching on Ecommerce Data Portal Mining

[I]Dr R.S. Vetrivel, [II]M .Shanmuga Rajalakshmi, [III]P Jeyanthi Rani
[I]Professor, [II]Research Scholar, [III]Assistant Professor
[I,II,III]Computer Science, Subramanya College of Arts & Science, Palani

Abstract
Ecommerce is the electronic commerce which is used in the field of funds or data transaction over the electronic network. The business exchange trading is developed among the companies and customer for buying and selling of products through the internet.
Based on the provider taxonomy information, the existing system was developed for organizing the products from data providers into master taxonomy. TACI approach was scalable and applicable to huge data sets on the web. The Taxonomy-Aware Catalog Integration (TACI) approach was designed with text classifier to ensure that products close together in the provider taxonomy which remains close in the master taxonomy which made as structured prediction optimization problem. The catalog integration was improved by the leverages of structure of the taxonomies. However, this approach was failed to classify the unlabelled product categories.
To overcome these problems, we propose a technique as Modified Latent Semantic Index based on the product catalog pattern matching. Modified LSI technique is developed for extracting the required data from the Ecommerce data portal. The Semantic Relationship Extractor is used in the LSI to discover the products which is semantically close with user query. The Product catalog pattern matching is performed to classify the similar data corresponding to the category which is easily extract the search products of user. The Modified LSI technique is used to enhance the accuracy of search results which provide an indexing on complex semantic relationship speed up the search.

Key Terms
Ecommerce data portal, Data mining, Semantic search, Latent Semantic Index, Product catalog pattern matching.

I. Introduction

A. Semantic search in data mining

Data mining technique is used for extracting useful information from the web effectively with less time consumption. Data mining is developed for identifying correlations or patterns among number of fields in large database. The data mining permits to user for examining the data from various dimensions or dissimilar perspectives and summarize it.
Semantic search is implemented in data mining for removing uncertainty through web search which helps to find out data patterns from World Wide Web (WWW). Ontology is the term to describe as the types of resources and semantic relationships between resources in the Semantic web. Ontology database contains the set of concepts and categories in a required region.
For communication, sequence of symbols also used in data mining which is easily affect the meaning and behavior of communication. Due to incredible extension in quantity of data with terrific growth of number of web pages, traditional search engines are not efficient to discover the required data patterns in web. These search engines do not easy for people to maintain enormous data manually and it is more time consuming to extract the information. Due to some difficulties in traditional web search, the semantic web search is introduced in data mining to provide efficient results.
Semantic search is the process of communicating with sufficient meaning to provide results from large data sets. Semantic search is developed to extract the data from database and to perform data for reuse process everywhere. Semantic search is used to provide best web service which coordinates and organizes all the data over web in a disciplined manner.
Semantic search generally developed to analysis the data from the web which is understandable not only by humans, but also machine understandable. During whole life cycle of data mining, semantics are necessary to enhance for improving performance level.

B. Pattern Matching

Extracting of useful patterns from database consists of various methods as data pattern processing, information extraction and knowledge discovery. In the field of data mining, pattern matching is the one of the essential technique. Pattern matching handles the complexity of category mismatch to overcome local maxima and increase classification accuracy.
The pattern matching technique is introduced for performing efficient data classification. Pattern matching technique organizes the data into predefined groups. Based on highest relevance to the keywords in the query, the documents are classified which contains the similar meaning to data by generating the relationship between web sites. Pattern matching is used to enhance the effectiveness of extracted patterns for identifying the relevant and interesting information from the database.

C. Latent Semantic Index

Latent Semantic Index (LSI) plays a significant role in the data (document) indexing process. Latent semantic index is the process of comparing and extracting the hiding data information from the groups of document through the search engine. LSI is able to identify the conceptual content of the data patterns by establishing semantic relationships between terms.
LSI is employed to recognize the data patterns in the relationship between terms and concepts which is enclosed with database. The contents of a webpage are dragged by a search engine and the most common words and phrases are collated and identified as the keywords for the page.
LSI method is developed for examine the commercial data portal (data base) to perceive which other documents enclosed with some of those same keywords. LSI process is used to analysis the documents from the database which provide results as that have many words in common to be semantically close, and ones with few words in common to be semantically distant to user query.

The synonymy and polysemy issues are overcome by the LSI process by using the term based information retrieval. Based on the singular value decomposition SVD of the term document matrix, LSI is employed for mining accurate data while detecting the similarity between the documents. LSI is an automatic indexing process that analysis the both documents and terms into a low dimensional space by intent as semantic concepts in the document.

II. Literature Survey

1. A Generic Pattern Matching Approach for Multiple Events

In this paper [1], the author proposes a new technique as Generic Pattern Matching technique which related with validate similarity between two events. Generic Pattern Matching technique was developed to stores the numbers of events and for managing all the characteristics of events.

Patterns are compared with all stored patterns in the large data set. The performance and efficiency of events was enhanced by improving the several events through the pattern matching. Generic Pattern Matching technique was time independent. Hence 1 : n issue was sorted out. This technique retrieves accurate results efficiently without any approximate value.

2. A methodology for interactive mining and visual analysis of clinical event patterns using electronic health record data

In this paper [2], the author introduces a method for interactive pattern mining and visual analysis. This proposed method was implemented for ad hoc visual exploration of patterns which mined from retrospective clinical patient data. This method was developed with an easy-to-use graphical query interface.

The powerful event pattern mining techniques was used to discover the important intermediate events within an episode. The interactive visualization techniques aids to uncover the event patterns that most impact outcome and how those associations change over time. In the patient medical conditions, the sequences of events variations were investigated in proper manner through interactive pattern mining and visual analysis.

3. Protein Structure Prediction By Means Of Sequential Pattern Mining

Based on the data mining and pattern matching techniques, the author implements a new hybrid method in this paper [3] for predicting the protein structure. This method uses the sequential type of protein data to obtain the enhanced results. This proposed method was developed with Prifix Span algorithm, to discover the frequent pattern from the input sequence. The best set of candidate patterns was extracted by using of scoring function on the frequent patterns.

The author designs a technique to extract the structure of unknown proteins which is required for biological experiments which also assists to expert domain by discovering the other feature of proteins. Protein Structure Prediction helps to pharmacist for discovering new drugs.

4. Modeling Associated Protein-DNA Pattern Discovery with Unified Scores

In this paper [4], the author introduces a unified score to develop an effective pipeline for associated TFTFBS pattern discovery. The discovered patterns are used to disease analysis application which leads to promising subtype.

The TF-TFBS pattern discovery incorporated with one-sided motif discovery on both the TF and the TFBS sides. Based on the additivity of the scores, the effective pipeline was implemented for retrieving the TF and TFBS paired instances through the identification of core motifs on each side.

5. Pattern-based Topics for Document Modelling in Information Filtering

In this paper [5], the author proposes a new approach as Information filtering model of Maximum matched Pattern –based Topic Model (MPBTM). The following important characteristics are implemented in this approach. (i) Through the multiple topics, requirements of user are generated. (2) Patterns are used for representing the each topic (3) based on the topic models, patterns are generated and classified in terms of their statistical and taxonomic features (4) the most discriminative and representative patterns as Maximum Matched Patterns, MPBTM approach aids to evaluate the document relevance to the user's information needs through the extracting of irrelevant documents.

6. Pattern Matching with Flexible Wildcards

In this paper [6], the author presents a new approach as Pattern Matching with Flexible Wildcards (PMFW) which deals with the Ending Positions of Pattern (EPP), Number of All Matches (NAM) and One-off Condition (OOC) problems. This PMFW method permits to users for controlling the ranges of wildcards.

7. Mining and Matching Relationships from Interaction Contexts in a Social Manufacturing Paradigm

In the context of cross-enterprise social interaction, the author introduces a new proposed model in this paper [7] which assists to knowledge transferring and sharing. They sustain the integration of the resources and capabilities among different enterprise. The two-phase unified model was developed to contest the group level manufacturing demand-capability relationships from manufacturing service interaction contexts (MSICs).

8. An Analysis of Tuberculosis Data for Pattern Matching Using Data Mining Techniques

The author was analyzing and then comparing the K-means clustering and farthest first clustering algorithm in this paper [8]. This technique was implemented for analyzing the Tuberculosis data, for which algorithm gave better results than the other algorithm by using the decision table. In the field of tuberculosis, classification techniques are improved effectively by this proposed technique.

9. Event log imperfection patterns for process mining towards a systematic Approach to cleaning event logs

In this paper [9], the author presents a technique to discuss the set of data quality issues in a systematic manner. The author considers the event log imperfection patterns for discovering the event logs by the systematic approach. The process mining analysis in this technique was used for extracting the event logs or encountered from raw data sources.

10. Quality-Aware Sub graph matching Over Inconsistent Probabilistic Graph Databases

The RDF data was described by inconsistent probabilistic RDF

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

graphs in this paper [10].This RDF graph includes the both inconsistencies and uncertainty. The author proposes an approach with two effective pruning methods as adaptive label pruning and quality score pruning. These methods were implemented for efficiently answer the QA-gMatch, which are filter out false alarms of sub graphs extremely.

11. Mining the entire Protein Data Bank for frequent spatially cohesive amino acid patterns

In this paper [11], the author introduces a new term as Frequent Spatially Cohesive Component sets (FreSCOs) for mining the protein data. For the huge data set of protein molecular structures, FreSCOs developed to extract the interesting and relevant new pattern class**.** A FreSCOs contains the combinations of two or three amino acids. FreSCOs signify the common building blocks which help in the stability of the protein structure.

12. Mining maximal frequent patterns in a single graph using inexact matching

In this paper [12], the author introduces MaxAFG algorithm for extracting the maximal frequent patterns in a single graph using inexact matching. The author uses MaxAFG algorithm for obtaining the details about search strategy and the similarity function. By using the inexact matching, the maximal patterns were identified which aids to minimize the amount of mined patterns

13. Multi-Core Processing of XML Twig Patterns

The author proposes a technique with Parallel Path Stack Algorithm (PPS) and Parallel Twig Stack Algorithm (PTS). The standard (sequential) Path Stack and Twig Stack algorithms were used for matching the XML query twig patterns in a parallel multi-threaded computing platform by using L-Stream representation scheme. These algorithms were implemented for limiting L-Stream processing to specific subtrees. According to the run time (to completion) compared PPS and PTS algorithms are compared by this technique.

14. Fuzzy Encoding Pattern for Stereo Matching Cost

In this paper [14], the author proposes a new Fuzzy Encoding Pattern which fuzzily encoding the relative orders between pixel pairs. For discovering more detailed information from a local structure, the relative orders were fuzzily encoded. The novel approach was introduced a pattern as matching function cost to operate under radiometric distortions variations between stereo correspondence images with high robust and accurate values.

15. Improving Fishing Pattern Detection from Satellite AIS Using Data Mining and Machine Learning

The author introduces a new approach as Satellite-based Automatic Information Systems (S-AIS) in this paper [15]. While considering of S-AIS data, the fishing activity was discovered by this proposed method for main fishing gear types as trawl, longline and purse seine. The author employed with three methods for identifying and extracting the fishing activities.

The Hidden Markov Model was implemented with vessel speed as observation variable for the trawlers. To longline, the data mining approach developed with algorithm from studies on the animal movement. We notice that vessel speed and operation time, for developing multi-layered filtering strategy which aids to discover the purse seiners. According to the S-AIS track data, automated methods are implemented in this paper, for discovering the potential fishing behavior from different gear types.

III. Methodology

Based on the Product Catalog Pattern Matching, the Modified Latent Semantic Index technique is developed for mining the data from Ecommerce data portal. This proposed technique is used to enhance the efficiency of mining which provides accurate search results.

The Taxonomy-Aware Catalog Integration (TACI) approach was introduced to resolve the difficulties in the categorization of commercial products from data providers into the master taxonomy. The product categorization was done by using the provider taxonomy information from the provider taxonomy. TACI approach was implemented based on a taxonomy-aware processing step that adjusts the results of a text-based classifier. Text-based classifier is used to ensure products which are close together in the provider taxonomy remain close in the master taxonomy. We consider this -intuition as structured prediction optimization problem. The leverages structure of taxonomies was introduced in this approach to improve the catalog integration**.** But, this exciting method reduces the catalog integration accuracy on uncertain featured products.

.The data integration task plays a vital role in online commercial portals and commerce search engines. The data integration task increases the complexity in the commercial portal and search engines for integration of products coming from multiple providers to their product catalogs.

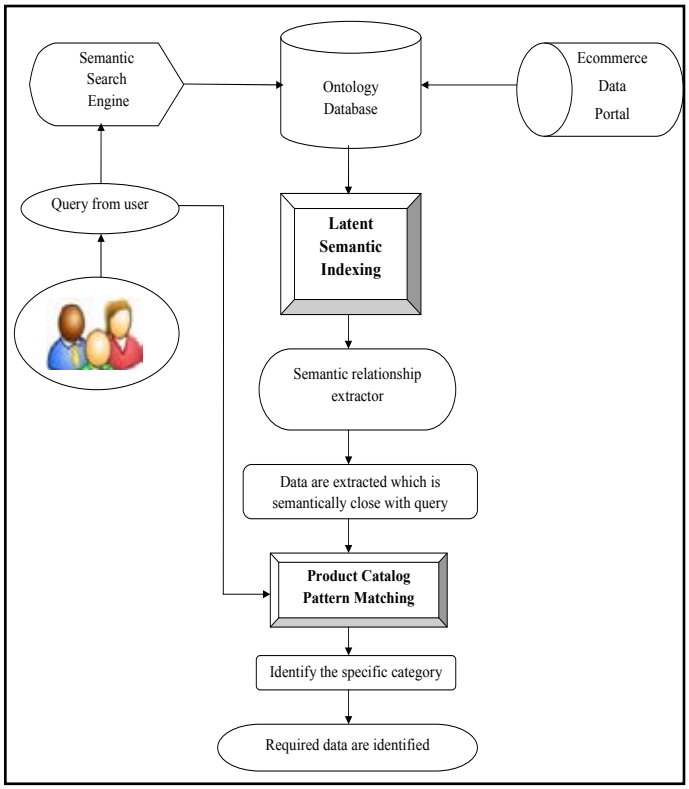

Fig. 3.1 : Architecture for Modified Latent Semantic Index based product catalog pattern matching on ecommerce data portal mining

The semantic search is used in commercial data portal mining to increase the search accuracy by understanding user (searcher) intention and contextual meaning of data from the document in the commercial portal which computes more relevant results. The semantic search consider context of search, location, intent,

variation of products, synonyms, generalized and specialized data patterns.

From the Figure 3.1,we analysis the Latent Semantic Index technique. According to the product catalog pattern matching, the Modified LSI technique is introduced in this proposed work for mining data from the ecommerce data portal. This work represented for semantic web search which input is considered as a query (Q,k), Q is the desired resource type and k is the set of keywords.

Initially, the input is given to the semantic search engine. Here, the ontology database is assists user to organize the related (close) meaning data items to query which is taken from Ecommerce data portal. LSI is used to identify the possible semantic relationships between resources to the query. Semantic Relationship Extractor is performed with LSI for extract the data (products) which is semantically close with query. The Product catalog pattern matching compares the both similar data and query which helps to identify the document folder that contains those similar data. Finally the required data is extracted from the Ecommerce portal database.

A. Modules

a) Modified Latent semantic index to identify similar data
b) Product Catalog Pattern Matching
c) Ecommerce data portal mining

B. Modules Description

1. Modified Latent semantic indexing to identify similar data

Latent semantic indexing is implemented for estimating the relationship between set of patterns and documents by providing the results which related to the concepts of patterns and documents. LSI is used for producing the results that are close in meaning which is considered as similar pieces of query.

The Latent semantic index is developed with singular value decomposition (SVD) by the term based document matrix. First step in the LSI, obtain the term based document matrix with the data from Ecommerce data portal database. The Term based document matrix illustrate the appearance of data patterns (terms) in documents. From the matrix we assume rows as data patterns and columns as documents from the database.

The singular value decomposition (SVD) is employed for minimizing the number of rows while preserve the similarity structure among columns. The singular value decomposition (SVD) matrix containing the patterns as rows corresponds to unique data and columns corresponds to each paragraph is constructed from a large piece of text.

Through the analysis of semantic relationship between data pattern and documents by SVD, term based document matrix indicates the number of related data are appear in each document. Data patterns are then compared by taking the cosine of the angle between the two data formed by any two rows. i) Values close to 1 represent very similar data. ii) Values close to 0 represent very dissimilar data.

2. Identify specific category by product catalog pattern matching

The provider product taxonomy is different from the master taxonomy, but in most cases, there is still a powerful signal coming from the provider classification. The products that are in nearby categories in the provider taxonomy should be classified into nearby categories in the master taxonomy.

The categorization of products from data provider into master taxonomy is consider as a problem due to the make use of the provider taxonomy information. The Product Catalog Pattern matching is used to enhance the catalog integration accuracy even on uncertain product classification.

The extraction of data from multiple data providers into a single product catalog is the significant task in the every commercial portal and search engines. The Product catalog pattern matching compares the both similar data from the semantic relationship extractor and query which identify results as similar data contains the document folder.

To find the most relevant information, user requesting information on a particular topic is searched among the thousands of results. Based on user request, web document mining performs to collect, categorize, manage and offers best probable information in the web. This process eliminates the noise and to improves the accuracy in information retrieval on the Web.

3. Ecommerce data portal mining

The growing of web portals in the internet which provides a user experience centered on online shopping. This includes e-commerce sites as amazons and shopping.com and commerce search engines such as Google product search and bing shopping.

Ecommerce data portal contains the all the information about products. The required data as a product is extracted by LSI technique from the Ecommerce data portal. The ontology database is used in this proposed work for find outs the related data to the query from user.

Input:	Database ‘D’, Number of queries ‘Q’
Output:	Mining of Ecommerce data
Step 1:	**Begin**
Step 2:	Enter the required text with keyword as query in the Semantic search engine
Step 3:	Latent Semantic Indexing find outs the possible semantic relationships between target resources
Step 4:	Semantic Relationship Extractor discover the data which is semantically close with query
Step 5:	Product catalog pattern matching compares the both similar data and query
Step 6:	Identify the document folder that contains those similar data
Step 7:	The required data is extracted from the Ecommerce portal database
Step 8:	**End**

Fig. 3.2 : Algorithm for mining the Ecommerce data by using Latent Semantic index based product catalog pattern matching

IV. Experimental Evaluation

In this paper worked on performance evaluation in terms of Execution time, Accuracy of results. The modified Latent semantic index is developed based on product catalog pattern matching for data mining from commercial port through Java. The performance metric to evaluates and analyze the value in java environment simulations. This technique improves the effectiveness of data mining from commercial portal. The performance measures of the proposed work are analyzed with following metrics:

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

- Execution time
- Accuracy of results
- Pattern matching efficiency

1. Execution time

The Execution time is defined as the amount of time taken for the extracting the required data from the Ecommerce data portal by using the Modified Latent Semantic Index (LSI) technique. Execution time is measured in terms of milli seconds (ms).

Table 4.1: Tabulation for Execution time

Number of Products	Execution time (ms)	
	Taxonomy-Aware Catalog Integration (TACI) approach	Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching
10	55	45
15	60	54
20	65	59
25	71	63
30	80	68

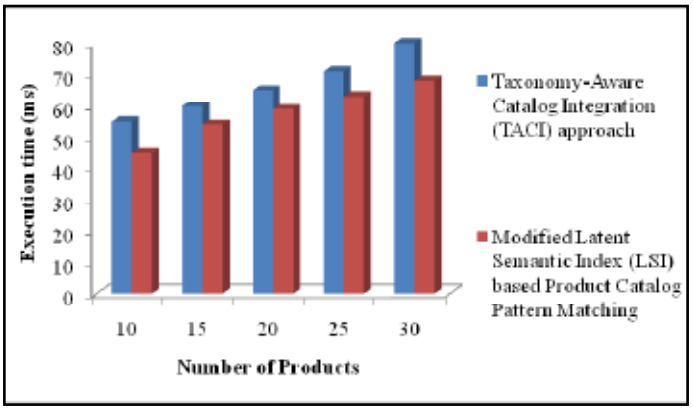

Fig. 4.1: Measure of Execution time

Figure 4.1 demonstrates the Execution time. The Products are taken for the experimental consideration is varied from 10 to 30. From the figure X axis represents the Number of products whereas Y axis denotes Execution time using Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique. From the figure it is clearly evident that the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique improves the Execution time than the Existing Taxonomy-Aware Catalog Integration (TACI) approach. Hence, the Execution time is reduced to 13% by the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique than the existing Taxonomy-Aware Catalog Integration (TACI) approach.

2. Accuracy of results

Accuracy of search results is measured as the ratios of the number of required products are discovered to user which is related to keywords. Semantic search process investigates to provide better search accuracy by accepting searcher decided. The dependant meaning of conditions as occur in the searchable data space on the web to make additional relevant results.

Accuracy of the result is measured in terms of percentage (%).

Table 4.2: Tabulation for Accuracy of results

Number of Products	Accuracy of results (%)	
	Taxonomy-Aware Catalog Integration (TACI) approach	Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching
10	25	30
15	30	36
20	35	42
25	40	48
30	45	56

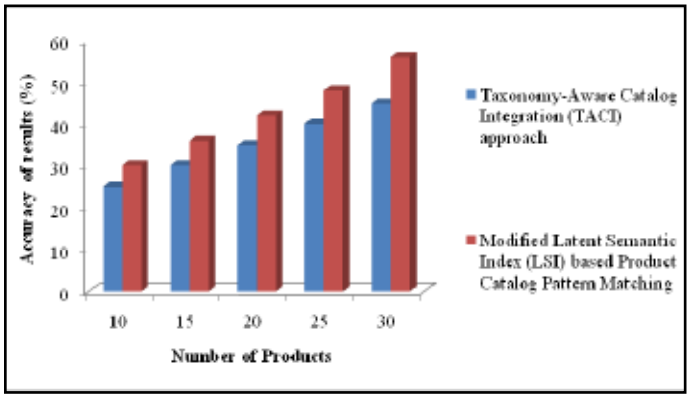

Fig 4.2: Measurement for Accuracy of results

Figure 4.2 demonstrates the Accuracy of results. The Products are taken for the experimental consideration is varied from 10 to 30. From the figure X axis represents the Number of products whereas Y axis denotes Accuracy of results using Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique. From the figure it is clearly evident that the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique improves the Accuracy of results than the Existing Taxonomy-Aware Catalog Integration (TACI) approach. Hence, the Accuracy of results is improved up to 21% by the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique than the existing Taxonomy-Aware Catalog Integration (TACI) approach.

3. Pattern Matching Efficiency

Pattern matching efficiency is defined as the ratio of the number of products to which are matching with related products in respective category. The Pattern matching efficiency is measured in terms of percentage (%).

Table 4.3: Tabulation for Pattern matching efficiency

Number of Products	Pattern Matching Efficiency (%)	
	Taxonomy-Aware Catalog Integration (TACI) approach	Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching
10	30	38
15	35	42
20	38	49
25	42	54
30	50	62

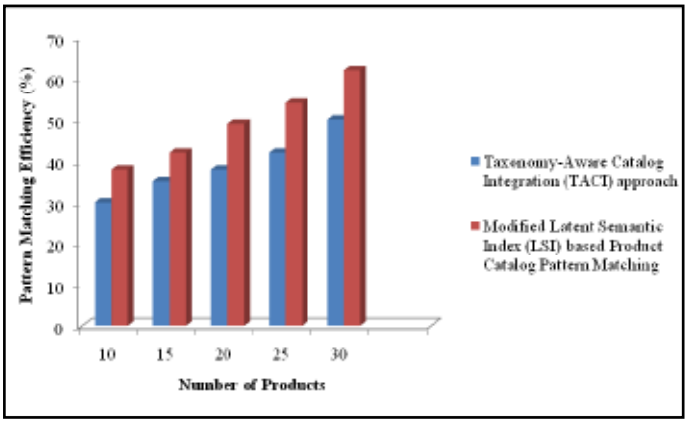

Fig. 4.3: Measurement for Pattern matching efficiency

Figure 4.3 demonstrates the Pattern Matching Efficiency. The Products are taken for the experimental consideration is varied from 10 to 30. From the figure X axis represents the Number of products whereas Y axis denotes Pattern Matching Efficiency using Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique. From the figure it is clearly evident that the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique improves the Pattern Matching Efficiency than the Existing Taxonomy-Aware Catalog Integration (TACI) approach. Hence, the Pattern Matching Efficiency is improved up to 26% by the proposed Modified Latent Semantic Index (LSI) based Product Catalog Pattern Matching technique than the existing Taxonomy-Aware Catalog Integration (TACI) approach.

V. Conclusion & Future Scope

1. Conclusion

For mining of Ecommerce data, Modified Latent Semantic index technique is introduced in this paper which is based on product catalog pattern matching. The LSI technique is developed for identifying the possible semantic relationship between the resources to the query. The main goal of this proposed technique is extracting the required data efficiently through of semantic relationship analysis. By the product catalog pattern matching, accuracy of pattern classification is improved for extracting the data from Ecommerce data portal. The Modified Latent Semantic Index (LSI) technique is developed to minimize the execution time required for classification.

References

[1] *R.Pradeepa, L. Sharmila, " A Generic Pattern Matching Approach for Multiple Events", International Journal of Advanced Research in Computer and Communication Engineering, Volume- 3, Issue 11, November 2014,Pages 8486-8488*

[2] *David Gotz , Fei Wang, Adam Perer ",A methodology for interactive mining and visual analysis of clinical event patterns using electronic health record data", ELSEVIER, Journal of Biomedical Informatics,Volume-48 ,2014,Pages 148–159*

[3] *Maral AzizI, Mohammad Saniee Abade, "Protein Structure Prediction By Means Of Sequential Pattern Mining", International Journal of Artificial Intelligence & Applications (IJAIA) ,Volume- 6, No. 4, July 2015,Pages 31-42*

[4] *Tak-Ming Chan, Leung-Yau Lo, Ho-Yin Sze-To, Kwong-Sak Leung, Xinshu Xiao, and Man-Hon Wong ," Modeling Associated Protein-DNA Pattern Discovery with Unified Scores", IEEE/ACM Transactions On Computational Biology And Bioinformatics, Volume-10, No. 3, May/June 2013, Pages 696-707*

[5] *Yang Gao,Yue Xu , Yuefeng Li ",Pattern-based Topics for Document Modelling in Information Filtering", IEEE Transactions on Knowledge and Data Engineering, 2015,Volume-27,Issue:6,Pages 1629-1642*

[6] *Xindong Wu, Ji-Peng Qiang and Fei Xie "Pattern Matching with Flexible Wildcards", Ournal Of Computer Science And Technology, Volume- 29, Issue: 5, Sept. 2014,Pages 740–750*

[7] *Jiewu Leng , Pingyu Jiang "Mining and Matching Relationships From Interaction Contexts in a Social Manufacturing Paradigm", IEEE Transactions On Systems, Man, And Cybernetics: Systems, Volume-47, Issue:2, 2017, Pages 276-288*

[8] *Ravinder Kaur, Anshu Sharma,"An Analysis of Tuberculosis Data for Pattern Matching Using Data Mining Techniques", International Research Journal of Engineering and Technology (IRJET) ,Volume: 03 Issue: 03 ,Mar-2016,Pages 679-685*

[9] *S. Suriadi , R.Andrews, A.H.M.terHofstede , M.T.Wynn, "Event log imperfection patterns for process mining towards a systematic approach to cleaning event logs",ELSEVIER, Information Systems, Volume 64, March 2017, Pages 132–150*

[10] *Xiang Lian, Lei Chen, Guoren Wang ,"Quality-Aware Subgraph Matching Over Inconsistent Probabilistic Graph Databases",IEEE Transaction On Knowledge And Data Engineering, Vol. 28,Issue 6,2016,Pages 1560-1574*

[11] *Pieter Meysman, Cheng Zhou, Boris Cule, Bart Goethals and Kris Laukens,"Mining the entire Protein Data Bank for frequent spatially cohesive amino acid patterns", Meysman et al. BioData Mining ,Volume 8, Issue 4 ,January 2015,Pages 1-15*

[12] *M. Flores-Garrido , J.A. Carrasco-Ochoa, J.F. Martínez-Trinidad ,"Mining maximal frequent patterns in a single graph using inexact Matching" ,ELSEVIER, Knowledge-Based Systems,Volume 66,August 2014,Pages 166-177*

[13] *Lila Shnaiderman and Oded Shmueli ,"Multi-Core Processing of XML Twig Patterns", IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 4, April 2015, Pages 1057-1070*

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

[14] Vinh Quang Dinh, Vinh Dinh Nguyen, Hau Van Nguyen, Jae Wook Jeon ,"Fuzzy Encoding Pattern for Stereo Matching Cost" ,IEEE Transactions on Circuits and Systems for Video Technology, Volume 26,Issue 7,2016,Pages 1215-1228

[15] Erico N. de Souza, Kristina Boerder, Stan Matwin, Boris Worm, " Improving Fishing Pattern Detection from Satellite AIS Using Data Mining and Machine Learning", PLOS ONE, Volume 11,Issue 7,July 2016,Pages 1-20