International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 4, Issue 08, August-2018

# HYBRID TECHNIQUES (HMM RANKING AND FUZZY C-MEANS) USED IN WEB MINING FOR GENERATING EFFICIENT RULES

[1]Tripti Saxena, [2]Dr.Pratima Gautam

[1]*Asst Prof , Computer Science & Engineering, Bhopal, India*
[2]*Professor & head, Computer Science & Engineering, AISECT, Bhopal, India*

***Abstract—During the previous couple of years the World Wide Web has turned into the greatest and most famous method for communiqué & info distribution. It serve as a stage for exchange different sorts of info. The amount of info accessible on the web is expanding quickly with the hazardous development of the World Wide Web & the approach of e-Commerce. While users are provided with more information and service options, it has become more difficult for them to find the "right" or "interesting" information, the issue regularly known as info overload. It is notable that more than 80% of the time necessary for complete any real world data mining project is generally spent on information pre-processing. Information pre-processing lays the preparation for data mining. Web mining is to find out & concentrate valuable data from the World Wide Web. It includes the programmed disclosure of patterns from at-least one Web servers. In this paper, HMM Ranking and FCM clustering used for the generation of better rules to improve the purchase products from the particular website.***

***Keywords—Web Mining, Clustering, Data prepprocessing, WebPage Ranking, URLs and Products.***

## I. INTRODUCTION

Web is a collection of billion of documents. The web is very enormous, diverse, flexible, and dynamic. The World Wide Web keeps on becoming both in the gigantic amount of movement & the size & multifaceted nature of Web sites. It is difficult to identify the relevant information present in the web. The vast majority of the substance in the web are unstructured in nature, however almost no work manages unstructured & heterogeneous data on the Web. The rising field of web mining goes for finding & extracting significant information that is covered up in Web related information, specifically in text documents published on the Web. Data Mining includes the idea of extraction significant & useful info from huge amount of data. Web mining is a imperative region in data mining where, we separate the interesting pattern from the substances. By and large three sorts of data are dealt with web-site namely

1) Content
2) Structure
3) Log information

In view of these sorts of data the Web Mining comprises of 3 forms specifically Web Content Mining, Web structure Mining and Web Usage Mining [1] as appeared in fig1. Web content mining deals with the raw data that is available on the web. The web structure mining mainly deals with the structure of the web sites [2]. Web Usage mining involves mining the usage characteristics of the users of Web applications. It is in a semi organized organization so that it needs heaps of pre-processing & parsing before the genuine extraction of the required data. This paper gives the study of web mining methods. Data mining process comprise of numerous phases that is Domain Understanding, Data selection, Data pre-processing and cleaning, Pattern discovery, Interpretation and Reporting.

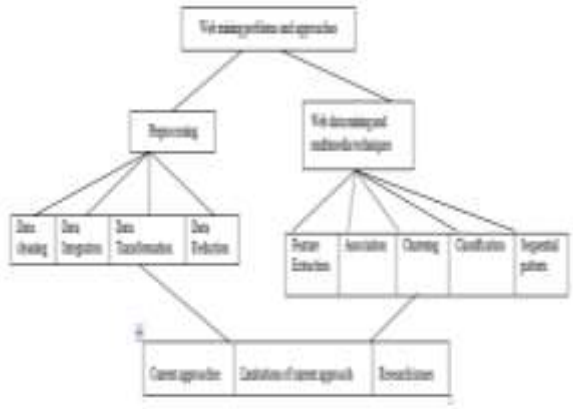

**Fig 1. Web Mining Techniques**

IJTIMES-2018@All rights reserved 

## II. WEB MINING PROBLEMS AND APPROACHES

Web mining is a method in data mining that naturally recovers extract & analyze the info from web. Yang and Wu et al, (2006) talks about the different issues to be tended in data mining. significant issues incorporate Automated Data Cleaning, Over Fitting, Under Fitting & Oversampling of data, Scaling up for high dimensional data, Mining sequence & time sequence info. A survey was conduct & set by k d pieces & various of the analysts recommended the essential employment for investigate as Scaling up Data Mining algo for gigantic data, mining text & automatic data refinement as the main issue talked about most elevated priorities. Different issues incorporates managing with lopsided data, mining data streams, link & network. Security in mining & distributed data mining also trapped the importance but not to as larger extent. A hotly debate technological problem is whether it is better to set-up a relational database arrangement or a multi-dimensional one. At long last, there is the subject of cost.

### 1) Data Pre-processing Techniques

Web log pre- processing is the initial step that is vital to enhance the productivity & value of the web info because almost 70% of the time is taken in pre-processing & these pre-pared info are given as an input to the next phases pattern detection & pattern analysis. There are numerous methods accessible for pre-processing since quite a while. Web log file assumes a huge part in pre-processing as the substance the client browse are recorded in these log files.

The info can be put away either at sever side, client side, on proxy servers & on operational databanks. Web Server Logs keeps up the past of page requests. Info about the request, client IP address, request date/time, page requested, HTTP code, bytes served, client agent, are put away. Proxy Server Logs a caching method which lies among client browsers & Web servers. It diminishes the heap time of Web pages & in addition the network traffic load at the server & client side & Browser Logs that can be changed or different JavaScript & Java applets can be used to collect client side data. Customer side accumulation scores over server-side gethering since it diminishes both the user & session identification problems. The advantages and disadvantages of log files and their behaviour are shown in the table given below. To enhance effectiveness & value of patterns mined & to avoid these loud & grimy data different pre-processing methods are accessible like Data cleaning, Data integration, Data transformation & data reduction.

| Log file | Advantages | Disadvantages | Behaviour relation |
|---|---|---|---|
| Client log file | Authentic and accurate | Modification, Collaboration | One to many. |
| Server log file | Reliable and accurate | Incomplete | Many to one |
| Proxy log file | Control efficiency of corporate access to the internet, log traffic. | Complex, unsecure. | Many to many. |
| Operational database | Simplicity and efficiency, accurate. | Vulnerability, maintenance | ---- |

## III. WEB LOG

Web server logs stores click stream info which can be helpful for mining purpose [7]. They are plain text (ASCII) documents which contain data about client Name, IP Address, Time Stamp, Access Request, URL that Referred, blunder codes (assuming any) & so forth & usually exist in the web servers. Generally there are four sorts of server logs: Transfer Log, Agent Log, Error Log and Referrer Log. The Transfer & the Agent Log are called as standard while the error & referrer log are viewed as discretionary as they may not be turned on. Every log entry records the traversal from one page to another, storing user IP number and all the related information [7]. A sample log taken from an astrology website is given below:

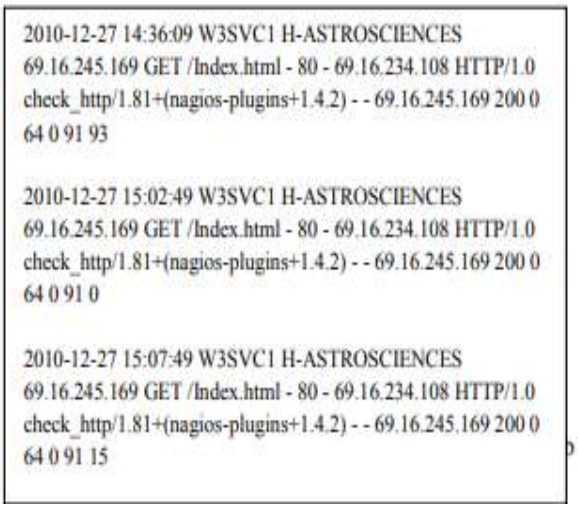


**Fig.2. Web log**

**a) Why Log Analysis is required?**

The mere complexity of the data volume paves way for the requirement of hybrid intelligent systems for intelligent information analysis and trend prediction [7]. But this complex data needs to be handled very carefully. For analyzing the data of web access logs of a website one needs to preprocess it. Preprocessing involves Cleaning, User/Session Identification and Transaction identification. An assortment of data cleaning algo have been proposed in this unique circumstance. Aside from making utilization of cleaning algo & afterward applying different mining algo to the scrubbed information, yet another route by which we can remove helpful info from this log data is by making utilization of automated log investigation tools. Web get to designs mined from Web logs are attractive & helpful information in practice [8]. This extricated information can be profoundly valuable if managed with carefully. It controls the expert in deciding the navigational pattern of the client i.e. which pages are much of the time went to by the client, after which page is he losing his enthusiasm for the site, which browser is being utilized by the people who get the website , the kind of errors that they get etc.
Logs if used legitimately can end up being exceptionally valuable in transforming the websites visitors into clients particularly if there should be an occurrence of a an e-commerce website. The personalization that is offered on different websites these days frequently makes utilization of this part of web usage mining.

## IV. RECOMMENDER SYSTEM

With the quick advancement of the World Wide Web, individuals would now be able to get & learning effectively through a wide range of online instruments, such as online forums and websites. Recommender systems have turned out to be tremendously common in recent years, & are utilized as a part of an assortment of utilizations like music, news, books, inquire about articles, seek inquiries,, social tags, & many products. The web has in this manner turn into a profitable & copious info source that significantly affects on clients lifestyles, especially their acquiring conduct.

Number of Internet users depends on information redeemed from the web to make their purchasing decisions. Even with the support of search engines, the number of retrieved documents is sometimes too large for users to obtain the desired information [10].

- **Collaborative Filtering Technique**

Method that is the most develop & most generally utilized for RS is collaborative filtering (CF). It relies only on opinions explicitly delivered by the users on items. The system recommends to the targeted customer products (or people), which have been evaluated by other people, whose interests are similar to the interest of the directed client. Collaborative filtering investigates methods for coordinating individuals with comparable interests & making suggestions on this premise. ordinarily, the work process of a collaborative filtering system is:

1. A client gives his or her inclinations by rating things of the system. These appraisals can be viewed as a unpleasant demonstration of the user enthusiasm for the compare space.
2. With comparative clients this present clients appraisals with other clients & discovers the people with most "comparable" tastes.
3. With comparative clients, the system prescribes things that the comparable clients evaluated exceptionally however not yet being appraised by this client.

- **Limitations**

Collaborative filtering method has some difficulties as follow:

**a) Cold start problem:** cold start difficulty comes when a novel client or things just enters the system. There are three kinds of cold start problems are: new user problem, new item problem and new system problem. **It is very difficult to provide recommendation in case of new user because we have very less information about user.**

**b) Sparsity problem: Sparsity h**as enormous impact on the nature of proposal. The vast majority of the client don't rate the majority of the things and the accessibility evaluations sparce it is the principal reason of data sparsity.

**c) Scalability:** Scalability is the property of system indicates its ability to handle increasing amount of information on web in well manner. With vast growth in information over internet, it is clear that the recommender systems are having so much data and thus it is a great challenge to handle it with growing demand.

**d) Over specialization:** over specialization problem means users are restricted to getting recommendations which is look like to those already defined in their profiles in some cases. It prevents user from getting new items and other options [10].

## V. WEB PAGE RANKING ALGORITHMS

The dimension of the www is developing quickly & at the mean time the quantity of enquiries can deal with has also developed incredibly. With expanding number of clients on the web, the quantity of questions submitted to the web indexes are additionally developing exponentially. In this way the search engines must have the capacity to process these inquiries proficiently. In this manner the web mining procedures are applied in order to take out only relevant documents from the databank & give expected data to the client. To display the reports in an arranged way, web page positioning techniques are connected which can orchestrate the archives arranged by their pertinence, significance and substance score & use web mining techniques to order them. The ranking algo can be proposed on various classification & some of them are

1) Link analysis algo
2) Personalized web search ranking algo
3) Page Segmentation algo

## VI. LITERATURE SURVEY

M.Sathya, Dr.P.Isakki et al. [2017] this paper consists of three phases. The first one is data preprocessing phase, which is the most important one because it makes the data with good quality. This should be possible by information cleaning, client ID & session distinguishing proof.. The next one is pattern discovery phase; in this the users' navigational pattern and rules are extracted by using apriori algorithm. Final one is pattern analysis phase, which is used to analyze and visualize the rules. The aim of this paper is to recognize the frequent connection from web log info by utilizing the Apriori algo. The aim of this paper is to identify the frequent link from web log data by using the Apriori algorithm [12].

Yonas Gashaw, Fang Liu et al. [2017] In this paper, three of the first association rule mining algo utilized for frequent pattern discovering namely, Eclat, Apriori, & FP-Growth inspected on three sets of transactional databases devised from server access log file. The comparison is made both in execution time and memory usage aspects. Unlike most previous research works, findings, in this paper, reveal that each of the algorithms has their own appropriateness and specificities that can best fit depending on the data size and support parameter thresholds [13].

Daminda Herath and Lakshman Jayaratne et al. [2017] in this paper, utilized for searching assets & for finding e-learner's navigation routes. Then collaborative filtering & content filtering are utilized to make personalized recommendations .

Bharat Patel, Varun Kakuste, Magdalini Eirinaki et al. [2017] in this work, our propose CaPaR, a Career Path Recommendation system, which tends to such weakness. Utilizing text mining & collaborative filtering methods first outputs the client's profile and resume, distinguishes the key abilities of the applicants & creates personalized job suggestions. Moreover, the system prescribes additional skills to student required for related employment opportunities, and learning assets for every expertise. In this way, the system not only allows its users to explore large amounts of information, but also expand their portfolio and resume to be able to advance their careers further. We explore and assess the different suggestion algo with true information gathered from the San Jose State University vocation focus site.

Yun-Young Hwang, Junghoon Park, Seoung Eun Park and Jungsun Yoon et al. [2017] This paper recommends the administrations & substance required by KOSEN individuals thinking about the local locations because of the idea of KOSEN. To do this, we use a big data processing solution and describe the analysis result [16].

Biswarup Bhattacharya, Iftikhar Burhanuddin, Abhilasha Sancheti‡ and Kushal Satya et al. [2017] our ran probes genuine time stamped client movement information, in the setting of prescribing reports to the client of a business examinations instrument/application & the outcomes are superior to the baselines. We additionally tuned certain parts of our model to land at streamlined outcomes.

Pratibha Sharma, Brahmdutt Bohra, Surendra Yadav et al. [2017] in this research, we will talk about the characterization of web information and its order and taking about the records (logs) maintained by the server. Server log documents are fundamentally the ASCII content records which contain the log record of users. The research work is a comparative investigation between web based log formats pre- bringing using two primary methods i.e. Apriori & FP Growth so that clients navigational behavior can be removed effortlessly & productively. To filter spam with traditional methodologies as dark white records (url, IP addresses, mailing data) is for all intents and purposes inconceivable. Utilization of content mining methodologies to a web logs can raise capability of a filtration of spam.

## VII. PROPOSED METHODOLOGY

The fuzzy c-means (FCM) algo is a clustering algo created via Dunn, & later on enhanced via Bezdek. It is helpful when the mandatory number of clusters is pre-decided; accordingly, the algo tries to put every of the data focuses to one of the groups. What makes FCM diverseis that it doesn't choose the supreme participation of an information point to a given group; rather, it figures the probability (the level of membership) that an information point will have a place with that bunch. Consequently, depending on the precision of the clustering that is required in practice, appropriate tolerance measures can be put in place. Since the absolute membership isn't computed, FCM can be to a great degree quick in light of the fact that the quantity of iterations required to accomplish a comparable to the required exactness.

- Iterations: In every iteration of the FCM algo, the following objective function J is minimized:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \delta_{ij} \parallel x_i - c_j \parallel^2$$

…….(1)

Here, N is the quantity of info focuses, C is the quantity of groups required, cj middle vector for bunch j, & δij is the level of participation for the ith information point xi in group j. The standard, k xi − cj k measures the likeness (or closeness) of the information guide xi toward the middle vector cj of bunch j. Note that, in every emphasis, the algo keeps up an inside vector for every one of the groups. These information focuses are figured as the weighted normal of the information focuses, where the weights are given by the degrees of enrollment.

- Degree of membership: For a given information point xi , the level of its enrollment to bunch j is computed as follows:

$$\delta_{ij} = \frac{1}{\sum_{k=1}^{C} \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

……..(2)

where, m is the fuzziness coefficient & the centre vector cj is computed as follows:

$$c_j = \frac{\sum_{i=1}^{N} \delta_{ij}^m . x_i}{\sum_{i=1}^{N} \delta_{ij}^m}$$

……..….(3)

In equation (3) above, δij is the estimation of the level of membership computed in the past emphasis. Note that toward the beginning of the algo, the level of participation for data point i to cluster j is initialised with a random value θij , $0 \leq \theta ij \leq 1$, such an extent is PC j δij = 1.

- Fuzziness coefficient: In condition (2) and (3) the fuzziness coefficient m, where $1 < m < \infty$, measures the resistance of the required clustering. This value decides how much the clusters can cover with each other. The higher the estimation of m the bigger the cover between bunches. In other words, the higher the fuzziness coefficient the algo utilizes, a bigger number of data points will fall inside a 'fuzzy' band where the level of participation is neither 0 nor 1, however some place in between.

- Termination condition: The required precision of the level of participation decides the quantity of cycles finished by the FCM algo. This measure of accuracy is calculated using the degree of membership from one iteration to the next, taking the largest of these values across all data points considering the greater part of the groups. If we represent the measure of precision between emphasis k & k + 1 with , we computed its value as takes after:

$$\epsilon = \Delta_i^N \Delta_j^C |\delta_{ij}^{k+1} - \delta_{ij}^k|$$

……..…….(4)

where, δ k ij and δ k+1 ij are separately the level of participation at cycle k and k + 1, and the administrator Δ, when provided a vector of qualities, restores the, largest value in that vector.

A Markov chain is useful when we need to compute a probability for a sequence of events that we can observe in the world. In many cases, however, the events we are interested in may not be specifically detectable in the world. the occasions we are keen on may not be specifically detectable on the planet. For instance, in we'll present the undertaking of grammatical form labeling, allocating labels like Noun and Verb to words. We didn't watch grammatical form labels on the planet; we saw words and needed to induce the right labels from the word arrangement. We call the grammatical form labels concealed on the grounds that they are not watched. A similar design comes up in discourse acknowledgment; all things considered we see acoustic occasions on the planet and need to derive the nearness of "concealed" words that are the basic causal wellspring of the acoustics. A concealed Markov show (HMM) enables us to discuss both watched occasions (like words that we find in the info) and shrouded occasions (like grammatical feature labels) we think of as causal factors in our probabilistic model.

In the proposed work, the pre-processing is performed in primary phase in the text mining. Data preprocessing is the most imperative piece of the whole information mining forms. The motivation behind information preprocessing is to transform the Web sign into some dependable, finish and precise sources to fulfill the need of information mining algo. Data preprocessing changes information into a configuration that will be more effortlessly & proficiently handled with the end goal of the client. The data from real world are often incomplete and inconsistent with the noise, the data preprocessing can improve the quality of the data for data mining. It not only can save a lot of time and space, but also are better able to play a role in decision making and forecasting. The fundamental undertaking of information preprocessing is to choose institutionalized information from the first log documents, arranged for client route design revelation calculation. The phase of information preprocessing incorporates information cleaning, client distinguishing proof and session ID. In data cleaning, the Web Log file is in text format then it is required to convert the file in database format and then clean the file. First, all the fields which are not required are removed and finally we will have the fields like date, time, client IP, URL access, Referrer and Browser used/ Access log files consist of large amounts of HTTP server information. Analyzing, this information is very slow and inefficient without an initial cleaning task. Every time a web browser downloads a HTML document on the internet the images are also downloaded and stored in the log file. This is on account of however a client does not expressly ask for designs that are on a website page, they are naturally downloaded because of HTML labels. The procedure of information cleaning is to expel insignificant information. All log passages with document name postfixes, for example, gif, JPEG, jpeg, GIF, jpg, JPG can be disposed of since they are unimportant.

In client and session recognizable proof, once HTTP log documents have been cleaned, following stage in the information preprocessing is the distinguishing proof of clients. Different methods for this are 1) By converting IP address to domain name. 2) The web server randomly assigns an ID to internet browser while it interfaces first time to the website. This is known as cookies. The web browser sends same ID back to web server effectively telling the web site that a specific user has returned. Cookies help the website developer to easily identifying individual visitors which results in a greater understanding of how the site is used.

A unique number is given to all the pages visited by users (p1,p2,..). The count, how many times the particular page is visited by a user is calculated. Also the count, how many times the page is visited by various users is calculated. To group the activities of a single user from the web log files is called a session. As long as user is connected to the website, it is called the session of that particular user. Most of the time, 30 minutes time-out was taken as a default session time-out. A session is a set of page references from one source site during one logical period. Historically a session would be identified by a user logging into a computer, performing work and then logging off. The login and logoff represent the logical start and end of the session. Then perform pattern mining calculation to manage the organized database and capture the database via the pattern mining algorithm for the age of as often as possible were given to design. These patterns are the series of item that stands up in massive amount of transaction. For Recommendation rule generation, apply Algorithm to the log data to capture the rundown of every from time to time utilized thing as sets. At that point, compute the Confidence and Support for lots URL of a website site. At that point on the idea of the successive frequent information item list we'll recommend a web page to the client this is the good in match of the patron query.

Proposed Algorithm:
Step:1 Start the process
Step:2 Initially apply preprocessing
Step:3 Perform data cleaning to remove extra attributes from the data
Step:4 Evaluate Session and customer identification
Step:5 Apply Processed Knowledge Base and Pattern mining calculation
Step:6 Generate rules using Recommendation rule generation
Step:7 Calculate ranking of the URLs and purchased products using HMM Ranking

Given a sequence of observations O, each observation an integer corresponding to the number of events, figure out the correct 'hidden' sequence Q. Let's be start with a formal meaning of a hidden Markov model, concentrating on how it varies from a Markov chain. An HMM is specified by the following components:
Q = q1q2 ...qN
a set of N states
A = a11a12 ...an1 ...ann
a transition probability matrix A, each ai j representing the probability of moving from state i to state j, s.t. Pn j=1 ai j = 1 ∀i
O = o1o2 ...oT
a arrangement of T perceptions, every one drawn from a vocabulary V = v1, v2,..., vV
B = bi(ot)
a arrangement of perception probabilities, also called emission probabilities, each communicating the probability of an observation ot being created from a state i
q0 ,qF
a exceptional start state and end (final) state that are not related with observations, together with transition probabilities a01a02 ...a0n out of the start state and a1Fa2F ...anF into the end state
Step:8 Generate PID for each products available
Step:9 Calculate PID counts for visited URLs and Purchased Products
Step:10 If Counts > Threshold
Apply Apriori Algo on choose data
Join Step: $C_k$ is created by joining $L_{k-1}$ with itself
Prune Step: Any (k-1)-itemset that isn't frequent can-not be a subset of a frequent k-itemset
$C_k$: Candidate itemset of size k
$L_k$: frequent itemset of size k
$L_1$ = {frequent items};
For ( k = 1; L k != ∅; k++) do begin
$C_{k+1}$ = candidates generated from L k;
For every exchange t in database do
Increment the count of all candidates in $C_{k+1}$ that are contained in t
$L_{k+1}$ = candidates in $C_{k+1}$ with min_support
End
Return $\cup_k L_k$;
Step:11 Evaluate Minimum Support Value (MSV) and Minimum Confidence Value (MCV)
Step:12 Now apply Fuzzy C Means over the data
Step:13 Calculate PID counts for visited URLs and Purchased Products
Step:14 Form two clusters of the generated URLs and Products
Step:15 Rule formation performed using Apriori Algorithm over the clusters
Step:16 End the process

## VIII. RESULT ANALYSIS

MATLAB makes utilization of in wide assortment of capacities, together with signal and image handling, correspondences, control configuration, test and size, financial modeling and analysis, computational science and parallel preparing. The present PC systems have tremendous registering force as customary CPU centers and also throughput-situated quickening agents, for example, pix processing units (GPUs). MATLAB programs are declarative and naturally express data-level parallelism as the language provides several high-level operators that work directly on arrays. Traditionally, MATLAB is used as programming language to write various types of simulations. It is used extensively to simulate and design systems in areas like control engineering, image processing and communications. These programs are typically long running and developers expend significant effort in trying to shorten their running times.

Rules Without Hmm

| | | |
|---|---|---|
| Rule #1: [8 10] --> 9 | | |
| Support = 0.126 | Confidence = 1 | Lift = 4.958 |
| Rule #2: [7 8] --> 9 | | |
| Support = 0.134 | Confidence = 0.941 | Lift = 4.666 |
| Rule #3: [5 10] --> 9 | | |
| Support = 0.126 | Confidence = 0.937 | Lift = 4.648 |
| Rule #4: [7 10] --> 9 | | |
| Support = 0.142 | Confidence = 0.894 | Lift = 4.436 |
| Rule #5: [5 9] --> 7 | | |

| |
|---|
| Support = 0.134 Confidence = 0.888 Lift = 3.777 |
| Rule #6: [8 9] --> 7 |
| Support = 0.134 Confidence = 0.888 Lift = 3.777 |
| Rule #7: [7 11] --> 9 |
| Support = 0.126 Confidence = 0.882 Lift = 4.375 |
| Rule #8: [9 11] --> 7 |
| Support = 0.126 Confidence = 0.882 Lift = 3.75 |
| Rule #9: [9 11] --> 10 |
| Support = 0.126 Confidence = 0.882 Lift = 3.387 |
| Rule #10: 9 --> 10 |
| Support = 0.176 Confidence = 0.875 Lift = 3.358 |
| Rule #11: [7 9] --> 10 |
| Support = 0.142 Confidence = 0.85 Lift = 3.262 |
| Rule #12: [5 7] --> 9 |
| Support = 0.134 Confidence = 0.842 Lift = 4.175 |
| Rule #13: 9 --> 7 |
| Support = 0.168 Confidence = 0.833 Lift = 3.541 |
| Rule #14: [5 9] --> 10 |
| Support = 0.126 Confidence = 0.833 Lift = 3.198 |
| Rule #15: [8 9] --> 10 |
| Support = 0.126 Confidence = 0.833 Lift = 3.198 |
| Rule #16: 15 --> 14 |
| Support = 0.159 Confidence = 0.826 Lift = 3.932 |
| Rule #17: 22 --> 17 |
| Support = 0.151 Confidence = 0.818 Lift = 4.056 |
| Rule #18: [9 10] --> 7 |
| Support = 0.142 Confidence = 0.809 Lift = 3.44 |
| Rule #19: [7 9] --> 5 |
| Support = 0.134 Confidence = 0.8 Lift = 3.173 |
| Rule #20: [7 9] --> 8 |
| Support = 0.134 Confidence = 0.8 Lift = 4.139 |
| Rule #21: 40 --> 41 |
| Support = 0.126 Confidence = 0.789 Lift = 4.47 |
| Rule #22: [10 11] --> 9 |
| Support = 0.126 Confidence = 0.789 Lift = 3.914 |
| Rule #23: 8 --> 9 |
| Support = 0.151 Confidence = 0.782 Lift = 3.880 |
| Rule #24: 16 --> 17 |
| Support = 0.151 Confidence = 0.782 Lift = 3.880 |
| Rule #25: 4 --> 3 |
| Support = 0.142 Confidence = 0.772 Lift = 2.786 |
| Rule #26: 14 --> 15 |
| Support = 0.159 Confidence = 0.76 Lift = 3.93 |
| Rule #27: 26 --> 27 |
| Support = 0.159 Confidence = 0.76 Lift = 3.478 |
| Rule #28: 9 --> 5 |
| Support = 0.151 Confidence = 0.75 Lift = 2.97 |
| Rule #29: 9 --> 8 |
| Support = 0.151 Confidence = 0.75 Lift = 3.880 |
| Rule #30: 17 --> 16 |
| Support = 0.151 Confidence = 0.75 Lift = 3.88 |
| Rule #31: 17 --> 22 |

| |
|---|
| Support = 0.151 Confidence = 0.75 Lift = 4.05 |
| Rule #32: [7 9] --> 11 |
| Support = 0.126 Confidence = 0.75 Lift = 2.97 |
| Rule #33: 8 --> 7 |
| Support = 0.142 Confidence = 0.739 Lift = 3.141 |
| Rule #34: 27 --> 26 |
| Support = 0.159 Confidence = 0.730 Lift = 3.478 |
| Rule #35: 6 --> 7 |
| Support = 0.126 Confidence = 0.714 Lift = 3.035 |
| Rule #36: 7 --> 9 |
| Support = 0.168 Confidence = 0.714 Lift = 3.54 |
| Rule #37: 41 --> 40 |
| Support = 0.126 Confidence = 0.714 Lift = 4.47 |
| Rule #38: [9 10] --> 5 |
| Support = 0.126 Confidence = 0.714 Lift = 2.83 |
| Rule #39: [9 10] --> 8 |
| Support = 0.126 Confidence = 0.714 Lift = 3.69 |
| Rule #40: [9 10] --> 11 |
| Support = 0.126 Confidence = 0.714 Lift = 2.83 |
| Rule #41: 2 --> 5 |
| Support = 0.142 Confidence = 0.708 Lift = 2.80 |
| Rule #42: 9 --> 11 |
| Support = 0.142 Confidence = 0.708 Lift = 2.80 |
| Rule #43: 13 --> 17 |
| Support = 0.142 Confidence = 0.708 Lift = 3.51 |
| Rule #44: 17 --> 13 |
| Support = 0.142 Confidence = 0.708 Lift = 3.512 |
| Rule #45: 9 --> [7 10] |
| Support = 0.142 Confidence = 0.708 Lift = 4.43 |
| Rule #46: 8 --> 5 |
| Support = 0.134 Confidence = 0.695 Lift = 2.75 |
| Rule #47: 8 --> [7 9] |
| Support = 0.134 Confidence = 0.695 Lift = 4.13 |
| Rule #48: 22 --> 13 |
| Support = 0.126 Confidence = 0.681 Lift = 3.380 |
| Rule #49: 22 --> 16 |
| Support = 0.126 Confidence = 0.681 Lift = 3.52 |
| Rule #50: 7 --> 5 |
| Support = 0.159 Confidence = 0.678 Lift = 2.69 |
| Rule #51: 7 --> 10 |
| Support = 0.159 Confidence = 0.678 Lift = 2.60 |
| Rule #52: 10 --> 9 |
| Support = 0.176 Confidence = 0.677 Lift = 3.35 |
| Rule #53: 9 --> [5 7] |
| Support = 0.134 Confidence = 0.666 Lift = 4.175 |
| Rule #54: 9 --> [7 8] |
| Support = 0.134 Confidence = 0.666 Lift = 4.666 |
| Rule #55: 8 --> 10 |
| Support = 0.126 Confidence = 0.652 Lift = 2.50 |
| Rule #56: 16 --> 22 |
| Support = 0.126 Confidence = 0.652 Lift = 3.52 |
| Rule #57: 8 --> [9 10] |

| |
|---|
| Support = 0.126 Confidence = 0.652 Lift = 3.69 |
| Rule #58: 5 --> 7 |
| Support = 0.159 Confidence = 0.633 Lift = 2.69 |
| Rule #59: 11 --> 10 |
| Support = 0.159 Confidence = 0.633 Lift = 2.43 |
| Rule #60: 2 --> 11 |
| Support = 0.126 Confidence = 0.625 Lift = 2.479 |
| Rule #61: 13 --> 22 |
| Support = 0.126 Confidence = 0.625 Lift = 3.38 |
| Rule #62: 9 --> [5 10] |
| Support = 0.126 Confidence = 0.625 Lift = 4.648 |
| Rule #63: 9 --> [7 11] |
| Support = 0.126 Confidence = 0.625 Lift = 4.375 |
| Rule #64: 9 --> [8 10] |
| Support = 0.126 Confidence = 0.625 Lift = 4.958 |
| Rule #65: 9 --> [10 11] |
| Support = 0.126 Confidence = 0.625 Lift = 3.914 |
| Rule #66: 10 --> 7 |
| Support = 0.159 Confidence = 0.612 Lift = 2.604 |
| Rule #67: 10 --> 11 |
| Support = 0.159 Confidence = 0.61 Lift = 2.43 |
| Rule #68: 7 --> 8 |
| Support = 0.142 Confidence = 0.607 Lift = 3.14 |
| Rule #69: 7 --> 11 |
| Support = 0.142 Confidence = 0.607 Lift = 2.40 |
| Rule #70: 7 --> [9 10] |
| Support = 0.142 Confidence = 0.607 Lift = 3.44 |
| Rule #71: 5 --> 9 |
| Support = 0.151 Confidence = 0.6 Lift = 2.97 |
| Rule #72: 7 --> [5 9] |
| Support = 0.134 Confidence = 0.571 Lift = 3.77 |
| Rule #73: 7 --> [8 9] |
| Support = 0.134 Confidence = 0.571 Lift = 3.77 |
| Rule #74: 5 --> 2 |
| Support = 0.142 Confidence = 0.566 Lift = 2.80 |
| Rule #75: 5 --> 11 |
| Support = 0.142 Confidence = 0.566 Lift = 2.24 |
| Rule #76: 11 --> 5 |
| Support = 0.142 Confidence = 0.566 Lift = 2.247 |
| Rule #77: 11 --> 7 |
| Support = 0.142 Confidence = 0.566 Lift = 2.40 |
| Rule #78: 11 --> 9 |
| Support = 0.142 Confidence = 0.566 Lift = 2.809 |
| Rule #79: 10 --> [7 9] |
| Support = 0.142 Confidence = 0.548 Lift = 3.262 |
| Rule #80: 7 --> 6 |
| Support = 0.126 Confidence = 0.535 Lift = 3.035 |
| Rule #81: 7 --> [9 11] |
| Support = 0.126 Confidence = 0.535 Lift = 3.75 |
| Rule #82: 5 --> 8 |
| Support = 0.134 Confidence = 0.533 Lift = 2.75 |
| Rule #83: 5 --> 10 |

| |
|---|
| Support = 0.134 Confidence = 0.533 Lift = 2.047 |
| Rule #84: 5 --> [7 9] |
| Support = 0.134 Confidence = 0.533 Lift = 3.173 |
| Rule #85: 10 --> 5 |
| Support = 0.134 Confidence = 0.516 Lift = 2.04 |
| Rule #86: 3 --> 4 |
| Support = 0.142 Confidence = 0.515 Lift = 2.78 |
| Rule #87: 11 --> 2 |
| Support = 0.126 Confidence = 0.5 Lift = 2.479 |
| Rule #88: 5 --> 3 |
| Support = 0.1260 Confidence = 0.5 Lift = 1.803 |
| Rule #89: 5 --> [9 10] |
| Support = 0.1260 Confidence = 0.5 Lift = 2.833 |
| Rule #90: 11 --> [7 9] |
| Support = 0.1260 Confidence = 0.5 Lift = 2.975 |
| Rule #91: 11 --> [9 10] |
| Support = 0.126 Confidence = 0.5 Lift = 2.833 |
| Rule #92: 10 --> 8 |
| Support = 0.126 Confidence = 0.483 Lift = 2.50 |
| Rule #93: 10 --> [5 9] |
| Support = 0.126 Confidence = 0.483 Lift = 3.198 |
| Rule #94: 10 --> [8 9] |
| Support = 0.126 Confidence = 0.483 Lift = 3.198 |
| Rule #95: 10 --> [9 11] |
| Support = 0.126 Confidence = 0.483 Lift = 3.38 |
| Rule #96: 3 --> 5 |
| Support = 0.126 Confidence = 0.454 Lift = 1.803 |
| |
| **Final Rules After HMM :** |
| |
| Rule #1: [5 10] --> 9 |
| Support = 0.126 Confidence = 0.93 Lift = 4.64 |
| Rule #2: [9 11] --> 10 |
| Support = 0.126 Confidence = 0.882 Lift = 3.38 |
| Rule #3: 9 --> 10 |
| Support = 0.176 Confidence = 0.875 Lift = 3.35 |
| Rule #4: [5 9] --> 10 |
| Support = 0.126 Confidence = 0.833 Lift = 3.198 |
| Rule #5: 22 --> 17 |
| Support = 0.151 Confidence = 0.818 Lift = 4.056 |
| Rule #6: [10 11] --> 9 |
| Support = 0.126 Confidence = 0.789 Lift = 3.914 |
| Rule #7: 26 --> 27 |
| Support = 0.159 Confidence = 0.76 Lift = 3.478 |
| Rule #8: 9 --> 5 |
| Support = 0.151 Confidence = 0.75 Lift = 2.975 |
| Rule #9: 17 --> 22 |
| Support = 0.151 Confidence = 0.75 Lift = 4.056 |
| Rule #10: 27 --> 26 |
| Support = 0.159 Confidence = 0.730 Lift = 3.47 |
| Rule #11: [9 10] --> 5 |
| Support = 0.126 Confidence = 0.714 Lift = 2.83 |

| |
|---|
| Rule #12: [9 10] --> 11 |
| Support = 0.126 Confidence = 0.714 Lift = 2.833 |
| Rule #13: 9 --> 11 |
| Support = 0.142 Confidence = 0.708 Lift = 2.809 |
| Rule #14: 10 --> 9 |
| Support = 0.176 Confidence = 0.677 Lift = 3.358 |
| Rule #15: 11 --> 10 |
| Support = 0.159 Confidence = 0.633 Lift = 2.431 |
| Rule #16: 9 --> [5 10] |
| Support = 0.126 Confidence = 0.625 Lift = 4.648 |
| Rule #17: 9 --> [10 11] |
| Support = 0.126 Confidence = 0.625 Lift = 3.914 |
| Rule #18: 10 --> 11 |
| Support = 0.159 Confidence = 0.612 Lift = 2.431 |
| Rule #19: 5 --> 9 |
| Support = 0.151 Confidence = 0.6 Lift = 2.975 |
| Rule #20: 5 --> 11 |
| Support = 0.142 Confidence = 0.566 Lift = 2.247 |
| Rule #21: 11 --> 5 |
| Support = 0.142 Confidence = 0.566 Lift = 2.24 |
| Rule #22: 11 --> 9 |
| Support = 0.142 Confidence = 0.566 Lift = 2.809 |
| Rule #23: 5 --> 10 |
| Support = 0.134 Confidence = 0.533 Lift = 2.047 |
| Rule #24: 10 --> 5 |
| Support = 0.134 Confidence = 0.516 Lift = 2.04 |
| Rule #25: 5 --> 3 |
| Support = 0.126 Confidence = 0.5 Lift = 1.803 |
| Rule #26: 5 --> [9 10] |
| Support = 0.12605 Confidence = 0.5 Lift = 2.833 |
| Rule #27: 11 --> [9 10] |
| Support = 0.126 Confidence = 0.5 Lift = 2.833 |
| Rule #28: 10 --> [5 9] |
| Support = 0.126 Confidence = 0.483 Lift = 3.198 |
| Rule #29: 10 --> [9 11] |
| Support = 0.126 Confidence = 0.483 Lift = 3.387 |
| Rule #30: 3 --> 5 |
| Support = 0.126 Confidence = 0.454 Lift = 1.803 |

**Final Rules Without FCM:**

| |
|---|
| Rule #1: [8 10] --> 9 |
| Support = 0.126 Confidence = 1 Lift = 4.958 |
| Rule #2: [7 8] --> 9 |
| Support = 0.134 Confidence = 0.941 Lift = 4.66 |
| Rule #3: [5 10] --> 9 |
| Support = 0.126 Confidence = 0.937 Lift = 4.648 |
| Rule #4: [7 10] --> 9 |
| Support = 0.142 Confidence = 0.894 Lift = 4.436 |
| Rule #5: [5 9] --> 7 |
| Support = 0.134 Confidence = 0.888 Lift = 3.77 |
| Rule #6: [8 9] --> 7 |

| |
|---|
| Support = 0.134 Confidence = 0.888 Lift = 3.77 |
| Rule #7: [7 11] --> 9 |
| Support = 0.126 Confidence = 0.882 Lift = 4.375 |
| Rule #8: [9 11] --> 7 |
| Support = 0.126 Confidence = 0.882 Lift = 3.75 |
| Rule #9: [9 11] --> 10 |
| Support = 0.126 Confidence = 0.882 Lift = 3.387 |
| Rule #10: 9 --> 10 |
| Support = 0.176 Confidence = 0.875 Lift = 3.358 |
| Rule #11: [7 9] --> 10 |
| Support = 0.142 Confidence = 0.85 Lift = 3.262 |
| Rule #12: [5 7] --> 9 |
| Support = 0.134 Confidence = 0.842 Lift = 4.175 |
| Rule #13: 9 --> 7 |
| Support = 0.168 Confidence = 0.833 Lift = 3.541 |
| Rule #14: [5 9] --> 10 |
| Support = 0.126 Confidence = 0.833 Lift = 3.198 |
| Rule #15: [8 9] --> 10 |
| Support = 0.126 Confidence = 0.833 Lift = 3.198 |
| Rule #16: 15 --> 14 |
| Support = 0.159 Confidence = 0.826 Lift = 3.93 |
| Rule #17: 22 --> 17 |
| Support = 0.151 Confidence = 0.818 Lift = 4.05 |
| Rule #18: [9 10] --> 7 |
| Support = 0.142 Confidence = 0.809 Lift = 3.44 |
| Rule #19: [7 9] --> 5 |
| Support = 0.134 Confidence = 0.8 Lift = 3.173 |
| Rule #20: [7 9] --> 8 |
| Support = 0.134 Confidence = 0.8 Lift = 4.13 |
| Rule #21: 40 --> 41 |
| Support = 0.126 Confidence = 0.78 Lift = 4.473 |
| Rule #22: [10 11] --> 9 |
| Support = 0.126 Confidence = 0.789 Lift = 3.91 |
| Rule #23: 8 --> 9 |
| Support = 0.151 Confidence = 0.782 Lift = 3.88 |
| Rule #24: 16 --> 17 |
| Support = 0.151 Confidence = 0.782 Lift = 3.88 |
| Rule #25: 4 --> 3 |
| Support = 0.142 Confidence = 0.77 Lift = 2.786 |
| Rule #26: 14 --> 15 |
| Support = 0.159 Confidence = 0.76 Lift = 3.93 |
| Rule #27: 26 --> 27 |
| Support = 0.159 Confidence = 0.76 Lift = 3.47 |
| Rule #28: 9 --> 5 |
| Support = 0.151 Confidence = 0.75 Lift = 2.975 |
| Rule #29: 9 --> 8 |
| Support = 0.151 Confidence = 0.75 Lift = 3.880 |
| Rule #30: 17 --> 16 |
| Support = 0.151 Confidence = 0.75 Lift = 3.880 |
| Rule #31: 17 --> 22 |
| Support = 0.151 Confidence = 0.75 Lift = 4.05 |
| Rule #32: [7 9] --> 11 |

| |
|---|
| Support = 0.126 Confidence = 0.75 Lift = 2.975 |
| Rule #33: 8 --> 7 |
| Support = 0.142 Confidence = 0.739 Lift = 3.14 |
| Rule #34: 27 --> 26 |
| Support = 0.159 Confidence = 0.730 Lift = 3.47 |
| Rule #35: 6 --> 7 |
| Support = 0.126 Confidence = 0.714 Lift = 3.035 |
| Rule #36: 7 --> 9 |
| Support = 0.168 Confidence = 0.714 Lift = 3.54 |
| Rule #37: 41 --> 40 |
| Support = 0.126 Confidence = 0.714 Lift = 4.47 |
| Rule #38: [9 10] --> 5 |
| Support = 0.126 Confidence = 0.714 Lift = 2.833 |
| Rule #39: [9 10] --> 8 |
| Support = 0.126 Confidence = 0.7142 Lift = 3.695 |
| Rule #40: [9 10] --> 11 |
| Support = 0.126 Confidence = 0.714 Lift = 2.8333 |
| Rule #41: 2 --> 5 |
| Support = 0.1428 Confidence = 0.708 Lift = 2.80 |
| Rule #42: 9 --> 11 |
| Support = 0.142 Confidence = 0.708 Lift = 2.809 |
| Rule #43: 13 --> 17 |
| Support = 0.142 Confidence = 0.708 Lift = 3.5122 |
| Rule #44: 17 --> 13 |
| Support = 0.142 Confidence = 0.708 Lift = 3.51 |
| Rule #45: 9 --> [7 10] |
| Support = 0.142 Confidence = 0.708 Lift = 4.43 |
| Rule #46: 8 --> 5 |
| Support = 0.134 Confidence = 0.695 Lift = 2.759 |
| Rule #47: 8 --> [7 9] |
| Support = 0.1344 Confidence = 0.695 Lift = 4.139 |
| Rule #48: 22 --> 13 |
| Support = 0.126 Confidence = 0.681 Lift = 3.3807 |
| Rule #49: 22 --> 16 |
| Support = 0.126 Confidence = 0.681 Lift = 3.5277 |
| Rule #50: 7 --> 5 |
| Support = 0.159 Confidence = 0.678 Lift = 2.691 |
| Rule #51: 7 --> 10 |
| Support = 0.159 Confidence = 0.678 Lift = 2.6048 |
| Rule #52: 10 --> 9 |
| Support = 0.17647 Confidence = 0.677 Lift = 3.35 |
| Rule #53: 9 --> [5 7] |
| Support = 0.134 Confidence = 0.66 Lift = 4.1754 |
| Rule #54: 9 --> [7 8] |
| Support = 0.134 Confidence = 0.666 Lift = 4.6667 |
| Rule #55: 8 --> 10 |
| Support = 0.126 Confidence = 0.652 Lift = 2.5035 |
| Rule #56: 16 --> 22 |
| Support = 0.126 Confidence = 0.652 Lift = 3.527 |
| Rule #57: 8 --> [9 10] |
| Support = 0.126 Confidence = 0.652 Lift = 3.6957 |
| Rule #58: 5 --> 7 |

| |
|---|
| Support = 0.159 Confidence = 0.633 Lift = 2.6917 |
| Rule #59: 11 --> 10 |
| Support = 0.1596 Confidence = 0.633 Lift = 2.43 |
| Rule #60: 2 --> 11 |
| Support = 0.126 Confidence = 0.625 Lift = 2.479 |
| Rule #61: 13 --> 22 |
| Support = 0.126 Confidence = 0.625 Lift = 3.38 |
| Rule #62: 9 --> [5 10] |
| Support = 0.126 Confidence = 0.625 Lift = 4.64 |
| Rule #63: 9 --> [7 11] |
| Support = 0.126 Confidence = 0.625 Lift = 4.375 |
| Rule #64: 9 --> [8 10] |
| Support = 0.126 Confidence = 0.625 Lift = 4.958 |
| Rule #65: 9 --> [10 11] |
| Support = 0.126 Confidence = 0.625 Lift = 3.91 |
| Rule #66: 10 --> 7 |
| Support = 0.159 Confidence = 0.612 Lift = 2.60 |
| Rule #67: 10 --> 11 |
| Support = 0.159 Confidence = 0.612 Lift = 2.43 |
| Rule #68: 7 --> 8 |
| Support = 0.142 Confidence = 0.607 Lift = 3.141 |
| Rule #69: 7 --> 11 |
| Support = 0.142 Confidence = 0.607 Lift = 2.40 |
| Rule #70: 7 --> [9 10] |
| Support = 0.142 Confidence = 0.607 Lift = 3.44 |
| Rule #71: 5 --> 9 |
| Support = 0.151 Confidence = 0.6 Lift = 2.975 |
| Rule #72: 7 --> [5 9] |
| Support = 0.134 Confidence = 0.571 Lift = 3.77 |
| Rule #73: 7 --> [8 9] |
| Support = 0.134 Confidence = 0.571 Lift = 3.77 |
| Rule #74: 5 --> 2 |
| Support = 0.142 Confidence = 0.566 Lift = 2.80 |
| Rule #75: 5 --> 11 |
| Support = 0.142 Confidence = 0.566 Lift = 2.247 |
| Rule #76: 11 --> 5 |
| Support = 0.142 Confidence = 0.5666 Lift = 2.24 |
| Rule #77: 11 --> 7 |
| Support = 0.142 Confidence = 0.566 Lift = 2.40 |
| Rule #78: 11 --> 9 |
| Support = 0.142 Confidence = 0.566 Lift = 2.809 |
| Rule #79: 10 --> [7 9] |
| Support = 0.142 Confidence = 0.548 Lift = 3.26 |
| Rule #80: 7 --> 6 |
| Support = 0.126 Confidence = 0.535 Lift = 3.03 |
| Rule #81: 7 --> [9 11] |
| Support = 0.126 Confidence = 0.535 Lift = 3.75 |
| Rule #82: 5 --> 8 |
| Support = 0.134 Confidence = 0.533 Lift = 2.75 |
| Rule #83: 5 --> 10 |
| Support = 0.134 Confidence = 0.533 Lift = 2.047 |
| Rule #84: 5 --> [7 9] |

| |
|---|
| Support = 0.134 Confidence = 0.533 Lift = 3.173 |
| Rule #85: 10 --> 5 |
| Support = 0.134 Confidence = 0.516 Lift = 2.04 |
| Rule #86: 3 --> 4 |
| Support = 0.142 Confidence = 0.515 Lift = 2.78 |
| Rule #87: 11 --> 2 |
| Support = 0.126 Confidence = 0.5 Lift = 2.479 |
| Rule #88: 5 --> 3 |
| Support = 0.126 Confidence = 0.5 Lift = 1.803 |
| Rule #89: 5 --> [9 10] |
| Support = 0.126 Confidence = 0.5 Lift = 2.8333 |
| Rule #90: 11 --> [7 9] |
| Support = 0.126 Confidence = 0.5 Lift = 2.975 |
| Rule #91: 11 --> [9 10] |
| Support = 0.126 Confidence = 0.5 Lift = 2.8333 |
| Rule #92: 10 --> 8 |
| Support = 0.126 Confidence = 0.483 Lift = 2.50 |
| Rule #93: 10 --> [5 9] |
| Support = 0.126 Confidence = 0.483 Lift = 3.198 |
| Rule #94: 10 --> [8 9] |
| Support = 0.126 Confidence = 0.483 Lift = 3.1989 |
| Rule #95: 10 --> [9 11] |
| Support = 0.126 Confidence = 0.483 Lift = 3.3871 |
| Rule #96: 3 --> 5 |
| Support = 0.126 Confidence = 0.454 Lift = 1.803 |

**Final Rules for First Cluster After FCM :**

| |
|---|
| Rule #1: [7 8] --> 9 |
| Support = 0.134 Confidence = 0.941 Lift = 4.666 |
| Rule #2: [5 9] --> 7 |
| Support = 0.1344 Confidence = 0.888 Lift = 3.777 |
| Rule #3: [8 9] --> 7 |
| Support = 0.134 Confidence = 0.888 Lift = 3.777 |
| Rule #4: [7 11] --> 9 |
| Support = 0.126 Confidence = 0.882 Lift = 4.375 |
| Rule #5: [9 11] --> 7 |
| Support = 0.126 Confidence = 0.882 Lift = 3.75 |
| Rule #6: [5 7] --> 9 |
| Support = 0.134 Confidence = 0.8421 Lift = 4.175 |
| Rule #7: 9 --> 7 |
| Support = 0.168 Confidence = 0.833 Lift = 3.5417 |
| Rule #8: [7 9] --> 5 |
| Support = 0.1344 Confidence = 0.8 Lift = 3.173 |
| Rule #9: [7 9] --> 8 |
| Support = 0.13445 Confidence = 0.8 Lift = 4.13 |
| Rule #10: 8 --> 9 |
| Support = 0.151 Confidence = 0.782 Lift = 3.8804 |
| Rule #11: 4 --> 3 |
| Support = 0.142 Confidence = 0.772 Lift = 2.786 |
| Rule #12: 9 --> 5 |
| Support = 0.1512 Confidence = 0.75 Lift = 2.975 |

| | |
|---|---|
| Rule #13: 9 --> 8 | |
| Support = 0.1512 | Confidence = 0.75 Lift = 3.88 |
| Rule #14: [7 9] --> 11 | |
| Support = 0.12605 | Confidence = 0.75 Lift = 2.97 |
| Rule #15: 8 --> 7 | |
| Support = 0.14286 | Confidence = 0.739 Lift = 3.14 |
| Rule #16: 6 --> 7 | |
| Support = 0.126 | Confidence = 0.714 Lift = 3.035 |
| Rule #17: 7 --> 9 | |
| Support = 0.168 | Confidence = 0.714 Lift = 3.541 |
| Rule #18: 2 --> 5 | |
| Support = 0.142 | Confidence = 0.708 Lift = 2.809 |
| Rule #19: 9 --> 11 | |
| Support = 0.142 | Confidence = 0.708 Lift = 2.809 |
| Rule #20: 8 --> 5 | |
| Support = 0.134 | Confidence = 0.695 Lift = 2.7594 |
| Rule #21: 8 --> [7 9] | |
| Support = 0.134 | Confidence = 0.695 Lift = 4.1391 |
| Rule #22: 7 --> 5 | |
| Support = 0.1596 | Confidence = 0.678 Lift = 2.691 |
| Rule #23: 9 --> [5 7] | |
| Support = 0.134 | Confidence = 0.6666 Lift = 4.175 |
| Rule #24: 9 --> [7 8] | |
| Support = 0.134 | Confidence = 0.666 Lift = 4.666 |
| Rule #25: 5 --> 7 | |
| Support = 0.159 | Confidence = 0.633 Lift = 2.691 |
| Rule #26: 2 --> 11 | |
| Support = 0.126 | Confidence = 0.625 Lift = 2.479 |
| Rule #27: 9 --> [7 11] | |
| Support = 0.126 | Confidence = 0.625 Lift = 4.375 |
| Rule #28: 7 --> 8 | |
| Support = 0.142 | Confidence = 0.607 Lift = 3.14 |
| Rule #29: 7 --> 11 | |
| Support = 0.142 | Confidence = 0.607 Lift = 2.408 |
| Rule #30: 5 --> 9 | |
| Support = 0.15126 | Confidence = 0.6 Lift = 2.975 |
| Rule #31: 7 --> [5 9] | |
| Support = 0.134 | Confidence = 0.571 Lift = 3.7778 |
| Rule #32: 7 --> [8 9] | |
| Support = 0.13445 | Confidence = 0.571 Lift = 3.77 |
| Rule #33: 5 --> 2 | |
| Support = 0.1428 | Confidence = 0.566 Lift = 2.809 |
| Rule #34: 5 --> 11 | |
| Support = 0.1428 | Confidence = 0.566 Lift = 2.247 |
| Rule #35: 11 --> 5 | |
| Support = 0.1428 | Confidence = 0.566 Lift = 2.247 |
| Rule #36: 11 --> 7 | |
| Support = 0.1428 | Confidence = 0.566 Lift = 2.408 |
| Rule #37: 11 --> 9 | |
| Support = 0.1428 | Confidence = 0.566 Lift = 2.809 |
| Rule #38: 7 --> 6 | |
| Support = 0.126 | Confidence = 0.535 Lift = 3.035 |

| | | |
|---|---|---|
| Rule #39: 7 --> [9 11] | | |
| Support = 0.126 | Confidence = 0.535 | Lift = 3.7 |
| Rule #40: 5 --> 8 | | |
| Support = 0.134 | Confidence = 0.533 | Lift = 2.759 |
| Rule #41: 5 --> [7 9] | | |
| Support = 0.1344 | Confidence = 0.533 | Lift = 3.173 |
| Rule #42: 3 --> 4 | | |
| Support = 0.142 | Confidence = 0.515 | Lift = 2.78 |
| Rule #43: 11 --> 2 | | |
| Support = 0.126 | Confidence = 0.5 | Lift = 2.4792 |
| Rule #44: 5 --> 3 | | |
| Support = 0.126 | Confidence = 0.5 | Lift = 1.803 |
| Rule #45: 11 --> [7 9] | | |
| Support = 0.126 | Confidence = 0.5 | Lift = 2.975 |
| Rule #46: 3 --> 5 | | |
| Support = 0.126 | Confidence = 0.454 | Lift = 1.803 |

## Conclusion

A vital assignment in any information mining application is the formation of a reasonable target informational collection to which information mining and factual calculations can be connected. This is especially vital in Web utilization mining because of the qualities of snap stream information and its relationship to other related information gathered from various sources and over numerous channels. The information planning process is frequently the most tedious and computationally escalated advance in the Web utilization mining process, and regularly requires the utilization of unique

## References

[1] Rajni Pamnani, Pramila Chawan 1 Qingtian Han, Xiaoyan Gao, “Web Usage Mining: A Research Area In Web Mining”.
[2] Chidansh Amitkumar Bhatt · Mohan S. Kankanhalli, “Multimedia Data Mining: State Of The Art And Challenges” Published Online: 16 November 2010© Springer Science+Business Media, LLC 2010.
[3] Wenguo Wu, “Study On Web Mining Algorithm Based On Usage Mining”, Computer- Aided Industrial Design And Conceptual Design, 2008. CAID/CD 2008. 9th International Conference On 22-25 Nov.2008.
[4] J Vellingiri, S.Chenthur Pandian, “A Survey on Web Usage Mining”, Global Journal of Computer Science and Technology .Volume 11 Issue 4 Version 1.0 March 2011.
[5] Ms. Dipa Dixit, Fr.CRIT, Vashi, M Kiruthika,” PREPROCESSING OF WEB LOGS”, (IJCSE) International Journal on Computer Science And Engineering, Vol. 02, No. 07, 2010, 2447-2452
[6] Brijendra Singh, Hemant Kumar Singh,”WEB DATA MINING RESEARCH: A SURVEY”, 978-1-4244-5967-4/10/$26.00 ©2010 IEEE.
[7] Xiaozhe Wang, Ajith Abraham, Kate A. Smith: Intelligent web traffic mining and analysis, ELSEVIER Journal of Network and Computing Applications, 2004.
[8] Jian Pei, Jiawei Han, Behzad Mortazavi-asl, and Hua Zhu: Mining Access Patterns Efficiently from Web Logs.
[9] Makvana, Kamlesh, Pinal Shah, and Parth Shah. "A novel approach to personalize web search through user profiling and query reformulation." Data Mining and Intelligent Computing (ICDMIC), 2014 International Conference on. IEEE, 2014.
[10] Khoi-Nguyen Tran, “Semantic Web Mining”, School of Computer Science, the Australian National University, Canberra, ACT 0200, Australia.
[11] Neelam Duhan, A. K. Sharma, Komal Kumar Bhatia, “Page Ranking Algorithms: A Survey” Advance Computing Conference, 2009. IACC 2009 IEEE International.
[12] M.Sathya, Dr.P.Isakki @ Devi “Apriori Algorithm on Web Logs for Mining Frequent Link” 2017 IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNIQUES IN CONTROL, OPTIMIZATION AND SIGNAL PROCESSING.
[13] Yonas Gashaw, Fang Liu “Performance Evaluation of Frequent Pattern Mining Algorithms using Web Log Data for Web Usage Mining” 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017).
[14] Daminda Herath and Lakshman Jayaratne “A Personalized Web Content Recommendation System for E-Learners in E-Learning Environment” 2017 National Information Technology Conference (NITC), 13-15 September, Colombo, Sri Lanka.

[15] Bharat Patel, Varun Kakuste, Magdalini Eirinaki "CaPaR: A Career Path Recommendation Framework" 2017 IEEE Third International Conference on Big Data Computing Service and Applications

[16] Yun-Young Hwang, Junghoon Park, Seoung Eun Park and Jungsun Yoon "The mechanism of personalized service recommendation for the academic field" 2017, IEEE.

[17] Biswarup Bhattacharya, Iftikhar Burhanuddin, Abhilasha Sancheti‡ and Kushal Satya "Intent-Aware Contextual Recommendation System" 2017 IEEE International Conference on Data Mining Workshops.

[18] Pratibha Sharma, Brahmdutt Bohra, Surendra Yadav "Comparative Analysis of Web-Mining Approaches for Efficient Mining of Server Log Formats" 978-1-5090-1489-7/16/$31.00 ©2016 IEEE.