**EXPLORATORY VISUAL SEQUENCE MINING BASED ON PATTERN-GROWTH**

**ABSTRACT:**

Continuous Enforcement Mines have found applications in a number of different fields. Due to the unconscious nature of the problem, both the main challenges arise. Firstly, there are a number of formats that are already in the existing algorithms, which are useless for the benefit of many users Perspective. Secondly, since databases are growing, they are available at the mining computational cost of large scale formats. So, there's a need mining approaches that can focus on search towards the direction of interest. This work copes with this problem Interactive visualization with continuous mines with a mine in order to create a "transparent box" functional model. We propose a Novel Approach for Interactive Display Row Mines, which allows the user to guide a system-development algorithm Suitable by a powerful visual interface. Our approach (1) introduces the possibility of using local barriers mining process; (2) methods allow visualization to be cut, and (3) the mining product algorithm towards the direction of interest. The use of local controls significantly improves users' ability to search advanced search no need to restart calculations. We present our approach using two event line data; one that covers the web page visits and other scenes of individuals.

**ARCHITECTURE:**

****

**EXISTING SYSTEM:**

There are, however, two main challenges that need to be addressed before sequential pattern mining can be fully utilized. The first challenge is based on the vast number of possible patterns. State-of-the- art algorithms may extract too many patterns, many of which may be of lesser significance or even irrelevant for the current analysis. This aspect makes it difficult for the user to grasp, and consequently use, the multitude of obtained patterns. The second challenge is the computational complexity involved in pattern identification, as mining large number of patterns is computationally very expensive. One approach to tackling these problems is to introduce constraints and promising results have been shown in many applications. These two challenges are the motivation behind several interactive systems which allow the user to define constraints to increase the effectiveness and efficiency of the mining process. However, the actual mining algorithms in these systems then operate as a black box, and the user only gets to interact with the resulting patterns and not with the pattern generation.

**PROPOSED SYSTEM:**

We propose a novel exploratory event sequence mining approach based on the pattern-growth methodology. The main contributions of the approach are the following. User-steered pattern mining. The proposed approach enables the entirely interactive mining of patterns by giving control to the user to steer the mining algorithm to directions of interest to the specific task. This is achieved by allowing the user to: (1) choose which sequence patterns to grow during the mining process, and (2) dynamically apply local constraints. Support for local constraints. The presented approach introduces the notion of ‘local constraints’ in the mining process by allowing a user to apply a number of different types of constraints on subsets of the search space. Stepwise visualization of patterns. Patterns are stepwise visualized in two views. A pattern tree view visualizing the frequent subsequences being built and an event sequence view displaying selected patterns in the context of the event sequences they appear in.

**ALGORITHM:**

**PREFIXSPAN**

**Prefix Span** is an algorithm for discovering sequential patterns in sequence databases. The input of **Prefix Span** is a sequence database and a user-specified threshold named minsup (a value in [0,1] representing a percentage). A **sequence database** is a set of sequences where each sequence is a list of itemsets. An itemset is an unordered set of distinct items. For example, the table shown below contains four sequences. The first sequence, named S1, contains 5 itemsets. It means that item 1 was followed by items 1 2 and 3 at the same time, which were followed by 1 and 3, followed by 4, and followed by 3 and 6. It is assumed that items in an itemset are sorted in lexicographical order. Note that it is assumed that no items appear twice in the same itemset and that items in an itemset are lexically ordered.



**MODULES:**

1. **CONTENT MAKING**

The main phase to initiate the process is to make the content that is available to user. Admin is the person to adding such contents with the search keywords and titles that are required to visible to users. The admin is entering the details and content and link to the other pages are shown. The total link, content, title, keywords are all stored in the database and its activities are noted. The contents are then, listed to the users in order to search of users query.

1. **USER INTENTION**

Users are authorized in this module, users are giving input for searching particular details and they are getting the details for what they are given in input. The related details are also shown to user under the required result in order to make a chain. Suggestions are given to users if they entered in to particular link. If users are interested to move further they can use the link to move and make the tree live. The suggestions are given based on the topic that user is entering.

1. **MINING INFROMATION**

Administrator needed to analysis the base of user action performs. For instance user search and enter into one particular option and suggestions are given to them later they can choose what they are interested or they can return to home. The linked details were given in database and that can form the tree structured data. Based on the given data admin can analysis what they required to do with that particular details.

1. **GRAPH REPRESENTATIONS**

The graph representation means that the analysis of the data that they are selected according to the navigation to other page. The graphs may vary to show that pie graph, bar graph or line graph in data that are processed to bring the problem to better understand. The solutions are represented in pictorial format than the written format so that the understand of particulars will be better. And comparisons are made easy to make decision.

**CONCLUSION:**

 The main contribution of the proposed work is an interactive sequence mining approach that allows a user to progressively refine constraints while pattern sequences are being built, enhancing in this way user exploration and control over the search for interesting patterns. This contrasts with existing interactive sequential pattern mining systems that mostly offer the possibility of setting constraints at the start of the mining process, using then different visualization techniques to explore the resulting patterns. Consequently, the latter tend to treat the mining process as a black box while our approach and prototype system, ELOQUENCE, attempts to open the box, reveal the process and allow a user to intervene and steer it. Additional key strengths of ELOQUENCE are the following. First, it combines two visual views, pattern tree and event sequence view, providing in this way additional context to the mining process by revealing how a selected pattern appears in the data. Second, different types of constraints are supported such as ontology level or gap constraints, and data filters. The practical usefulness of these features is demonstrated by two example use cases.

**FUTURE WORK:**

Several interesting problems merit further research. First, we would like to investigate how our proposed interactive “transparent box” approach can be incorporated in other sequence mining algorithms. It would also be interesting to closely examine how the pattern-growth approach can be extended to mine soft sequential patterns and which type of constraints and visualization techniques could be used to guide the search for such patterns. In the current status of ELOQUENCE, pattern support is computed based on the first match of the pattern in a sequence. A future step would be to extend this to also take into account the number of times a pattern appears within a sequence. Furthermore, more research is required to find ways to visualize patterns that minimize visual clutter. This problem is particularly relevant when big datasets are analyzed. Possibilities may include to represent sequential patterns in more expressive languages or to allow the user to define non trivial pattern ranking criteria. Finally, a formal usability evaluation of the system described here has not yet been performed, though we plan for it as a next step in our work.

**REQUIREMENT ANALYSIS**

The project involved analyzing the design of few applications so as to make the application more users friendly. To do so, it was really important to keep the navigations from one screen to the other well ordered and at the same time reducing the amount of typing the user needs to do. In order to make the application more accessible, the browser version had to be chosen so that it is compatible with most of the Browsers.

**REQUIREMENT SPECIFICATION**

**Functional Requirements**

* Graphical User interface with the User.

**Software Requirements**

For developing the application the following are the Software Requirements:

1. Python
2. Django
3. Mysql
4. Wampserver

**Operating Systems supported**

1. Windows 7
2. Windows XP
3. Windows 8

**Technologies and Languages used to Develop**

1. Python

**Debugger and Emulator**

* Any Browser (Particularly Chrome)

**Hardware Requirements**

For developing the application the following are the Hardware Requirements:

* Processor: Pentium IV or higher
* RAM: 256 MB
* Space on Hard Disk: minimum 512MB

**CONCLUSION:**

The main contribution of the proposed work is an interactive sequence mining approach that allows a user to progressively refine constraints while pattern sequences are being built, enhancing in this way user exploration and control over the search for interesting patterns. This contrasts with existing interactive sequential pattern mining systems [3]– [5] that mostly offer the possibility of setting constraints at the start of the mining process, using then different visualization techniques to explore the resulting patterns. Consequently, the latter tend to treat the mining process as a black box while our approach and prototype system, ELOQUENCE, attempts to open the box, reveal the process and allow a user to intervene and steer it.