Data mining for the Analysis of Content Interaction in web-based Learning and Training Systems

Saarthak Singh*

Student, Class 12th, Inspiration Public School, Haldwani, Uttrakhand, India

Email: saarthak.singh.2304@gmail.com

Abstract - The process of data mining is exploring databases for useful information and extracting it. Analysis how people interact with websites is the main goal of web usage mining, a subfield of data mining. The main, unobtrusive, and objective way to evaluate Web-based training and learning systems, namely how users engage with course materials, is web use mining. Many mining methods were developed with the classroom in mind, and we will showcase and explain them all. In order to prove effectiveness and improve instructional design as required, it is crucial to analyse and evaluate how learners behave in learning and training technology systems. This is particularly true when there are several interactive learning and training components accessible. We take a look at methods for figuring out what we want to learn, how we're going to learn it, and how our habits change as we go.

Keywords - Data mining, web use mining, educational technology, learning analytics, instructional design.

- X -

1. INTRODUCTION

In recent years, there has been a growing trend in utilising computer-supported learning and training environments to augment or substitute conventional methods of learning and training. The World-Wide Web has become the dominating platform for these ecosystems. The initial wave of web-based educational platforms achieved success by capitalising on their inherent advantage of providing convenient access to educational information. Lately, there has been a shift towards bolstering a broader array of educational activities, so strengthening the user's learning experience by improving interaction and engaging the learner more effectively. Skills-oriented training activities supplement traditional knowledge-based learning.

The key aspect of Web-based settings is the interaction between learners and material, which supports the acquisition of information and training of skills. Technically, interaction is a manifestation of learning activities and techniques. To assess learning and training behaviour, it is essential to analyse the interaction between individuals and the material in these settings.

The understanding of learning behaviour in learning and training contexts is still lacking. Unlike conventional classroom-based learning and training, the learner's personal decisions play a greater role in determining the learning techniques and behaviour in organising their learning and training. In addition, many instructional aspects are frequently accessible simultaneously, enabling proficient learners to select their preferred method of integrating resources and features. Therefore, the examination and assessment of learning and training behaviour has significant significance. Authors and instructional designers need to possess a comprehensive knowledge of successful and preferred learning styles and behaviour in order to enhance the design of learning content. Feedback on usage is necessary for instructors to enhance the delivery of Web-based instructional tools. Our goal is to support both summative and formative assessment. However, given the newness of the application, the integration of formative evaluation into an incremental design and development process is the more important element.

A framework for analysing and evaluating learning and training behaviour and interaction in Web-based educational systems must be capable of supporting a range of techniques:

- The identification and exploration of educational and instructional interactions derived from sources such as records of accessing the Web.
- The deliberate recording and display of interaction activity, apart from the logs of recorded access requests and interactions.
- Applying an analytical paradigm to the study and interpretation of student behaviour in a classroom context.

Direct observation and surveys were the usual methods of evaluating learning behaviour in traditional classroom settings. A new option, however, has emerged thanks to the proliferation of computer-supported learning and training environments, especially those that are Web-based. Users generate digital footprints of their actions and conduct within computer-assisted platforms. Web servers in Web-based applications automatically produce access logs to record user requests. Data mining, specifically in the context of Web use mining, may be utilised to extract and reveal hidden knowledge related to behaviour in Web-based settings.

Data mining is the process of uncovering, extracting, and analysing data from extensive databases. The objective of data mining is to transform information that is implicitly represented into an explicit form. Web use mining, in its specialised form, focuses on extracting user activity from access records of websites. Web use mining is a method of analysing user activity that is focused on observation and assessment. It is particularly useful for studying learning behaviour. This methodology has been supported by research studies. Although it has certain limits, it provides a non-invasive method of observation that can significantly contribute to dependable and precise evaluation outcomes for educational purposes.

This paper aims to provide a comprehensive overview of web usage mining techniques, explain their key applications, and showcase their benefits using a case study. Despite web usage mining's widespread use, we'll be concentrating on tactics developed for use in educational settings and with interactive material.

2. LITERATURE OF REVIEW

Dai N. (2022) this research focuses on using data mining and constructing a neural network model to develop a data interaction process model that is based on data mining and topology visualisation. This work conducts preprocessing data operations, including data filtering and cleaning of the gathered data. A conventional multichannel convolutional neural network (MCNN) is utilised in deep learning methodologies to notify pupils about their academic progress. Furthermore, the CNN's network architecture is optimised to enhance the model's performance. The CNN architecture requires careful tuning of its hyperparameters to create an ideal model that can efficiently process the data. This study presents a technique for visualising the network topology in unstable locations. It aims to solve the problem of not having a suitable approach for organising the network topology in specific places. This approach converts the challenge of network topology layout in an unstable zone into a circular topology diffusion problem within a convex polygon. This ensures a clear and logical relationship between the topology, significantly minimising gaps in the area and resulting in a more uniform and aesthetically pleasing layout. This work develops a real-time data interaction paradigm by utilising JSON format and database triggers, while ensuring reliable delivery through the use of message queues. A real-time data interaction solution is created by integrating the timer technique with the original key in a platform-based approach. The approach proposed in this work takes into account the real-time precision, security, and dependability of data exchange. It fulfils the platform's original and recently identified needs for data interaction.

Ijaz et al. (2020) Data mining is a method used to extract valuable patterns from many sources. Data mining is a crucial component in several fields such as marketing, electronic-commerce, corporate intelligence, healthcare, and social network analysis. Due to the progress made in these applications, several academics have shown their interest in developing data mining applications specifically for educational purposes. Educational data mining is a scientific field that focuses on analysing and extracting insights from various forms of data collected in educational settings. This paper examines several case studies that utilise data mining in educational settings. These systems and mining approaches are utilised for the purpose of collecting and analysing information. Assessing student performance becomes exceedingly difficult due to the vast quantity of data in educational databases. There is an urgent requirement in Pakistan to closely monitor and evaluate the academic achievement of students. There are two primary factors that have hindered the ability of previous systems to analyse student performance.

Currently, the study on contemporary assessment techniques is inadequate in analysing the suitable ways for assessing the development and performance of students in Pakistani institutions. The second reason is the lack of research on the characteristics that affect students' success in certain courses. Therefore, this proposal suggests doing a thorough examination of student performance evaluation by employing Data Mining techniques in order to enhance student achievements. The objective of this study is to enhance students' academic achievement by selecting the most appropriate characteristics through the application of Educational Data Mining (EDM) techniques.

Gushchina et al. (2019) The paper examines the utilisation of educational data mining tools to improve the efficiency of E-learning processes. It specifically emphasises the implementation of adaptive feedback, individual evaluation, and personalised attention to student profiles. The techniques employed in this study are cluster analysis to establish time thresholds for tasks, data visualisation to emphasise course alternatives and popular resources, and V-fold cross-checking to assess the relationship between a high activity % and academic success. These strategies aid in analysing student behaviour inside the E-learning system, gauging their level of engagement with learning materials, and assessing the calibre of instructional content. The objective is to enhance the whole educational experience for pupils.

Angeli et al. (2017) the report examines and elucidates fundamental inquiries about the use of data mining in educational technology classroom research. This paper presents two instances of data mining approaches; specifically association rules mining and fuzzy representations. These examples are drawn from research undertaken in Europe and Australia. Both of these studies investigate student learning, behaviours, and experiences in computer-supported classroom activities. The initial study employed association rules mining as a means to get deeper insights into the interaction patterns of learners with varying cognitive profiles when engaging with a simulation to resolve an issue. Association rules mining has been identified as a valuable technique for acquiring dependable data on learners' utilisation of the simulation and their proficiency in using it. The study demonstrates the application of data mining in enhancing the evaluation of educational software in the field of educational technology. The second research utilised the method of fuzzy representations to systematically analyse questionnaire data. The project exemplifies how educational technologists might employ data mining to guide and monitor the incorporation of technology in schools. The study's findings highlight the need of creating instructional data mining tools that effectively present results, information, explanations, comments, and suggestions to non-expert users in a relevant manner. Finally, matters of data privacy are resolved.

3. INTERACTION AND BEHAVIOUR

3.1 Interaction between Learning and Training

Supporting design and formative evaluation is the main goal of implementing Web usage mining. Learning and training systems cannot be developed without sufficient methodology. Both analytical models for assessments and a foundational base for instructional design may be provided by an abstract learning and training model. Particularly fundamental to this method is the idea of interaction.

Various ways may interaction be characterised or categorised. Students often divide their interactions into three types: those between students and material, those between students and instructors, and those between students themselves. Ohl argues that the connection between learners and material is vital. In particular, student behaviour in online classrooms is often indicative of how they engage with course materials. The reason for this is because making educational materials more accessible has always been the main goal of the Web. The term "interaction" has a heavy load of connotations when used in relation to the Internet. It is important in the field of education and also in the field of computers. The use of an abstract interaction model allows for the harmonisation of the different viewpoints.

3.2 Putting Interaction Into Practice

A learning technology system can't be complete without including training and learning. Two particular aspects of technology must be considered when planning a learning technology system: the general human-computer interface and the educational media and services that comprise it. Each of these two parts adds a new layer to the connection between training and learning, and all three of them use different ideas of interaction. It is essential to plan and execute the learner's engagement with the system that provides instructional resources and services.

- The learner's characteristics at the human-computer interface need to be described in terms of their behaviour, cognitive elements, learning objectives and tasks, and language factors.

- The system should be specified using a technically focused interaction language that is suitable for implementing educational services and media.

The concept of interaction is crucial, however its interpretation varies across different situations. The involvement of a student in educational systems may be seen at the level of execution, converted into human-computer interactions, and examined within the framework of learning and training interaction. The layered abstract interaction model for our design and assessment methodology is characterised by its reliance on Web use mining.

3.3 An Abstract Model for Interactions with Content

The term "interaction" has several possible applications. Several levels of abstraction are applicable to interaction and behaviour when classifying interactions.

- Interactions in learning and training include the ever-changing flow of information among students, teachers, and course materials.
- To put it simply, the human-computer interface (HCI) is the part of a computer system that allows users to interact with its software and its interface features.
- The processes and frameworks of technology engagement are referred to as multimedia and service interaction.

An abstract conceptual model that organises and creates conceptions of interaction consists of the three levels stated.

It is important to link the interactions that take place at the interface of a system's multimedia and services with those that take place during training and learning within the realm of instructional design [18]. When mining the web for information, it's important to think about how system-level interaction activities gleaned from web logs might be interpreted and used for training and education.

Conducting a formative assessment is our paramount objective. Since online education is still in its infancy, development methodologies that place an emphasis on user engagement and are supported by data mining are necessary. Nevertheless, aiding in the creation and implementation of incremental instructional design is the only goal. With the use of web usage mining, educational technology systems can keep tabs on students in real time. One possible use of this method is to identify people whose learning capacities are lacking by observing their behaviour. The hallmark of them is oftentimes erratic behaviour.

3.4 Identity and Data Learning Environment (IDLE)

We will talk about the Interactive Database Learning Environment (IDLE) to show how mining methods are unique to education and how beneficial it is to use these techniques to analyse and evaluate training and learning interactions. IDLE is a great resource for a database basics course that is offered to third-year undergraduates in computer science. The following is an outline of IDLE's educational goals:

- The IDLE system follows the principle of a virtual apprenticeship, where the student is portrayed as an apprentice who, under the guidance of a virtual master, is expected to do subject-specific tasks on their own.
- The goal of IDLE is to provide a realistic learning and training environment that seamlessly integrates several educational offers (such labs, tutorials, and lectures) into one platform, enhancing the practical experience.

The virtual lecture component of IDLE relies on the simultaneous presentation of audio and visual content. IDLE teaching services utilise animation and simulation techniques to effectively illustrate the practical components of the course material. The lab function is designed to facilitate the hands-on coursework, which includes tasks related to database building and programming, in the learning environment. The instructional and lab training offerings are particularly intriguing components for Web mining applications. These focus on enhancing abilities through the implementation of various interactive learning elements. Learners engage with IDLE to acquire knowledge and practise the skills of graphical database design and database programming. In this context, IDLE assumes the role of a virtual mentor, engaging with a virtual learner.

The IDLE environment has undergone development for several years. Due to the newness of the technology, a development method based on incremental prototyping was adopted. The system prototypes were created, utilised during the course, and assessed; the formative assessment findings were incorporated into subsequent enhancement and expansion phases. Web use mining has been crucial in facilitating this approach. It facilitated the identification and assessment of learner behaviour to enhance both the system and the instructional design. Furthermore, the ability to continuously observe the class has greatly assisted the instructor in managing the course.

4. MINING DATA AND WEB USAGE

Data mining aims to uncover and extract hidden knowledge from a database. This knowledge is categorised into rules and patterns to facilitate analysis and decision-making processes. Data mining is commonly employed in decision support systems for corporate purposes and as analytical tools in scientific applications. It may be utilised in a predictive manner for decision making, a generative manner for creating new or improved designs, and an explanatory manner for scientific study.

4.1 Web Usage Mining in the Educational Context

Web use mining is a specialised type of data mining that concentrates on examining user activity in systems that operate on the internet, as shown in Figure 1. In the context of the Web, the database refers to the access log generated by Web servers. A weblog item contains the requester (or the corresponding URL), the requested resource, and the date and time of the request. By understanding the content and purpose of the resource, we may infer the related actions based on the log entries. For example, the IDLE system allows lecture participation through the use of audio files and certain HTML sites. Web use mining focuses specifically on analysing actions and behaviour, distinguishing it from the broader scope of data mining. Web use mining has been extensively utilised in e-commerce applications to track and evaluate buying habits. The primary goal in the context of e-commerce is to establish effective customer relationship management. This is accomplished by optimising and personalising the shopping operations using the data obtained from use mining.

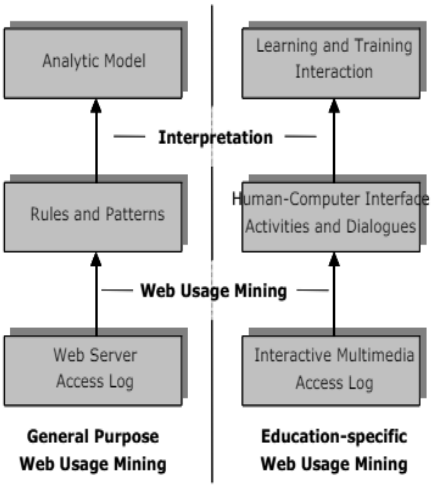

Figure 1: Web usage mining in the educational context

Web usage mining is proposed as a tool for designing and managing the relationship between the learner and the educational environment, aiming for an improved learning experience. Data mining techniques can be classified into three categories: basic statistics, static relationships, and dynamic relationships.

4.2 Data and Web Mining Techniques

Usage statistics are often not seen as data mining approaches, however primary quantitative usage metrics are routinely prioritised. But they are usually the starting point for evaluations. Basic metrics like total visit count and page-specific visit count are used by web-based systems to record utilisation. These metrics are essential for the tracking functionalities of most learning technology systems and web log analyzers.

Static relationships pertain to the connections between items of interest at a certain moment or during a specified timeframe.

- Methods for classification and prediction work hand in hand. Classification finds the labels for the classes, whereas prediction finds the functions with continuous values. In order to examine a specimen, a model is used. The results of this learning stage are then put into action. Predictions are often made using regression.

- Clustering involves the grouping of data elements that are comparable to each other. Unlike categorization, the class designations are not predetermined. The learning process is referred to as unsupervised in this context. Pattern recognition is a classic illustration.

- Association rules are intriguing correlations identified within a collection of data elements. An illustrative instance is the analysis of purchases, which can find pairings of items that are commonly bought together.

Dynamic connections pertain to alterations or evolving trends within the database being examined.

- When events are documented in a database over a certain time period, sequential pattern analysis is used. Events that tend to repeat themselves are retrieved. Trends in things like online shopping or web use are commonplace.

- The ability to evaluate patterns and variations in behaviour across time makes time series analysis an important tool for analysts.

There are essentially three stages to data mining methods: cleaning, extraction, and interpretation. When you're cleaning up, you may get rid of unnecessary entries like photographs. The aforementioned mining methods are essential to the extraction process since they allow for the identification of rules and patterns. The results of the extraction process, once evaluated and understood, must be interpreted in the context of the particular domain of application.

There are a number of limitations to web usage mining despite its many benefits. These mostly stem from the fact that web logs do not record every contact. To reduce the strain on the network caused by frequently downloaded content, web browsers use caching. Using data mining to create a visual depiction in this case would provide an inaccurate image. However, these problems may be circumvented if commercial websites use the dynamic page generation method.

4.3 Education-specific Web Usage Mining

Several factors must be taken into account while implementing data mining in the educational setting.

The primary emphasis in mining is on acquiring knowledge of behaviour. Several mining strategies proposed for the typical scenario are not sufficiently focused on the educational environment to obtain significant results. Therefore, certain domain-specific approaches, which are primarily modifications of common techniques, will be introduced. The interpretation of mining findings must be done within the framework of learning and training interaction. This may be achieved by developing an analytic model of interaction. Based on these factors, we have created the following strategies that are specifically designed for teaching.

- Session statistics rely on basic quantitative metrics to provide a quantitative summary of usage.
- The classification of sessions is determined by the classification approach used, with the aim of identifying learning goals.
- Behavioural patterns are a way to identify actions by expanding and applying the sequential pattern technique.
- Education makes use of time series, which are a byproduct of time series technique, to discover new approaches.

In the upcoming parts, we will demonstrate these methods by utilising the case study.

Extracting and making sense of data relies heavily on languages. It is necessary to document the behaviour of the interaction. Learners' interactions with course materials may be seen in web logs, which behave as behavioural representations. Nevertheless, a more theoretical vocabulary is required for a more thorough comprehension and analysis of these interactions as they pertain to interpretation and the interaction model. Interaction protocols for multimedia interfaces and services are considered to be at the intermediate level. A learning activity language that tracks user navigation and allows interpretation inside a training and learning interaction model may make the learning environment more interactive. The language used must faithfully depict the interaction's structure, which comprises nodes and their interconnections.

- Nodes symbolise the many resources in the system, including both static content items and active services. Nodes are resources that are given names and are categorised depending on the specific activities and subjects they are designed to assist.
- The user's actions, such browsing the system and choosing activities, are represented by the arcs connecting the nodes. Iterations, sequences, options, and concurrent activities are all examples of activity combinations that may be expressed.

The expression is given by the nodes (resources) Exercise1, Exercise2, and Exercise3.

(Exercise1; (Exercise2 or Exercise3))*

This shows that Exercise1 is iteratively addressed using the '*' iteration operator before either Exercise2 or Exercise3 is sequentially handled using the ';' sequence operator. This language may be utilised for both the creation and execution of tasks.

5. Session Statistics

In e-learning and e-training, a session is a basic idea. A learner's interactive behaviour throughout a time of active study is represented by a series of web log entries called a session.

Table 1: shows the IDLE Chapter 1 material's access statistics over the course of a week

Resource	Number of Requests
ch1-lectov.html	396
ch1-lect1.html	224
ch1-lect2.html	218
ch1-lect3.html	207
ch1-lect4.html	198

Fundamental quantitative data that might provide light on resource utilisation and time allocation for a particular learning activity are session statistics. Sessions, the basic unit of analysis, are often the basis of these metrics. It is possible to compare the outcomes to the instructional designer's expectations. The analytic model incorporates well-defined expectations. Other statistical measures may also be useful in this regard.For relevant data, simple metrics may be used to sort the total number of requests by time period or sum the statistics by resource. Table 1 shows that about half of the students who have seen the chapter summary pages have also seen the content for that chapter. Additionally, fewer students have finished the chapter than have begun it.

Take the IDLE system as an example; after removing superfluous inquiries, students typically submitted an average of 239 resource requests every term. Although it was surprised at first, more investigation showed that students access the course notes online during each session, so it's no surprise that they have the highest requests overall. The lab's interactive elements were a distant second. A large amount of effort will need to be devoted to the interactive lab components that students use to submit their answers to homework problems. However, these metrics mostly focus on the 'what' rather than the 'how' of resource utilisation.

5.1 Session Classification

Extracting and identifying the main learning goals and higher-level learning activities from a session record is the purpose of session categorization [25]. The typical learning session consists of the student focusing on no more than two primary tasks. Classification analysis of the most common user interactions with the system allows us to identify the primary learning goals.

Table 2: Classification of resources and activities

Class	URLs	Activity
Lectures	{ch1-lectov.html, ch1-lect1.html, ch1-lect2.html, … }	attending virtual lectures
Tutorials	{ch3-anim1.html, ch5-anim1.html, …}	participating in a virtual tutorial
Labs	{ch6-sql1.html, ch6-sql2.html, …}	practicing and training in a virtual lab
Downloads	{CourseNotes.pdf, ProjectSpec.pdf, …}	downloading resources
Look Up	{Schedule.html, Results.html, …}	look up of course-related information

Media resources used in the course One way to organise a website is to create many classes, C1,..., CN, with each class Ci containing {Ui1,..., UiM} of URLs. Our learning activity language's nodes and the categories to which they belong are being referred to here. Participating in a virtual lecture or doing lab assignments are examples of knowledge-level learning activities, and each class Ci corresponds to a certain kind of system-level interaction that makes these possible. An activity might be seen as indicative of a clear goal when a student spends a considerable amount of time on it or makes several requests for relevant materials and activities. A total of all page requests for all classes are added together. Class titles in the provided example, such "Lectures" or "Tutorials," indicate the connected activities with the resources listed in Table 2.

Table 3: Classification of sessions

Class	Percentage of Requests
Lectures	33
Labs	26
Tutorials	21
Download	12
Look Up	8

Each session generates a ranking Ci1 ≤ … ≤ CiN of the key learning objectives represented by the learning activity courses. This ranking is determined by the amount of requests for each class, providing us with valuable information about the students' learning goals and how they are being pursued. The example activities are classified accordingly.

Lectures, labs, tutorials, and downloads are all available. Refer to the rating in Table 3.

This session categorization may be extrapolated to encompass all sessions of an individual student or to groups of learners. It provides a teacher with a comprehensive overview of the utilisation of the technology.

The approach is applicable in an iterative evaluation process. Initial classifications may initially lack specificity and may be further improved to provide more detailed and specific classifications, which can identify more precise activities, objectives, and goals. This refinement process leads to more comprehensive and useful analytical findings. In our example, we first distinguished learning activities based on their kind. However, we found it beneficial to further classify them by themes, such as chapters. This resulted in subclasses such as LecturesCh1, LecturesCh2, etc., or TutorialsCh1, TutorialsCh2, etc. This represents a conventional method of employing data mining, known as a drill-down technique, to obtain a more comprehensive understanding of learner activities.

6. BEHAVIOURAL PATTERNS

The objective identification achieved by session categorization, as discussed in the preceding section, is a conceptual technique that disregards the temporal aspect, namely the order in which certain activities occur. Frequently, nevertheless, it is essential to examine interactions at a more detailed level in order to thoroughly analyse learning activities.

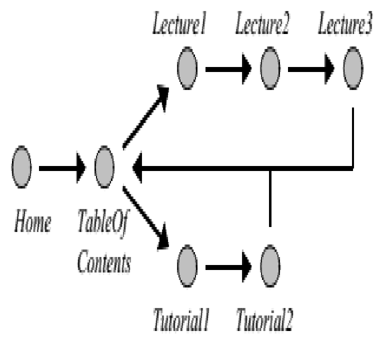

Figure 2: A graph depicting a pattern of behaviour

Finding patterns of interactions between behaviours in the log file is the main objective of the behavioural pattern mining method. Students' irrelevant page views or brief system exits are examples of irrelevant actions that may be ignored. Sequential patterns may be seen in the filtered sequences. The sequences go through threshold control, which acts as an extra filter to exclude really out-of-the-ordinary ones, in order to find out the patterns that learners follow.

More than merely sequences of acts are involved in behavioural patterns. Students work on many aspects of the course at once, repeat sections, or make selections from a list of options. Nodes and arcs abstract the navigation infrastructure and interactive elements, representing the course topology in a model. The behavioural patterns are built upon this paradigm. One way to explain navigating these topologies is via the expressions connected to the topology's links, which are part of the learning activity language. When students engage with a multimedia learning system, they are exhibiting a behavioural pattern, which is a language used to describe the learning process.

This code snippet loads the first lesson's tutorial and then loads the second lesson's tutorial and so on.

Lecture1 and Tutorial1 are considered activities in this case. The choice between tutorials and lectures is indicated by the | operator in the phrase. The * operator allows for several repetitions of the instruction before going on to the following lesson. Figure 2 shows another way to visually represent these emotions, and it shows

Back to the main page, where you can find the table of contents and links to the following courses: (Lecture 1; Lecture 2; Lecture 3 | Tutorial 1; Tutorial 2)*

6.1 While doing the language practice

The terminology employed to articulate these patterns is an essential component of the analytical model utilised to evaluate mining outcomes. Hence, it is imperative to establish a correlation between these behavioural patterns and the sequential patterns derived from the Web log. As previously stated, activities can be linked to the transitions between the nodes (URLs) recorded in the log. The expressions of behavioural patterns can serve as either a design instrument, reflecting the instructor's intended use, or as an assessment instrument, abstracting the learner's conduct. These patterns are obtained through the process of Web use mining. There are two distinct uses that may be identified:

The goal is to see whether the teacher's abstract behavioural patterns line up with the actual sequential patterns of learner-content interactions. The only thing needed to put this mining approach into action is the ability to extract consecutive patterns. For instance, we found that 85 percent of the time, the expected pattern of behaviour for a certain lab feature (as described by the teacher) was really seen in the lab (as shown in the web logs as a series of events). When comparing behavioural patterns to sequential ones, it's helpful to look at things like the number of allowed options and repetitions. Here, we used a simulated connection to see whether there was a correlation between sequence patterns and behaviour. In order for a sequential pattern to match or mimic a behavioural pattern, it may take the following forms:

Repetitions of P are permitted if P* is supplied. Choices between P1 and P2 can be made if P1 | P2 is specified.

It is necessary to adhere to sequences and simultaneous utilisation of resources, if necessary. Alternative interpretations that would relax the limitation of the simulation might potentially permit departures from the prescribed trajectory.

The extraction of real behaviour involves capturing behavioural patterns. The challenge with this method is in the absence of a singular solution. It is challenging to detect specific iterations and concurrent use, even if the overall topology and navigation linkages are known. The initial phase involves extracting the pattern, followed by determining the pattern's support based on the class. Two patterns can be compared by calculating their distance using the sequence alignment approach, which measures the amount of divergence from the shared route.

Although this approach yields intriguing outcomes, its full potential remains untapped; further investigation is necessary in this area. Greater levels of achievement can be attained by including factors such as the duration of each activity and other relevant attributes into the assessment [26]. The duration of an activity can provide insight into the true utilisation of a resource, rather than only relying on assumptions based on its planned utilisation.

6.2 Time Series

Time series refer to a collection of measurements taken at regular intervals over a certain duration. These metrics can encompass outcomes derived from any of the mining approaches previously discussed. The objective is to identify alterations in learning behaviour, which frequently indicate the overall learning approach during the course's duration. There are two significant reasons for this.

- Firstly, the instructional designer may have intended for a change to occur, and it is necessary to verify if the change has really taken place. One instance that illustrates this situation is the assessment of scaffolding characteristics using behavioural pattern analysis. The gradual reduction of scaffolds, which are tools that assist students in developing independence and proficiency in a subject, is a crucial aspect that is anticipated to occur in a successful scaffolding approach.

- Furthermore, it is crucial to identify unforeseen alterations. Web mining techniques may be used to evaluate time series data, enabling the discovery and continuous monitoring of changes in student behaviour. In addition to changes in conduct, the case study has also noticed changes in learning processes. Initial patterns frequently exhibit solitary objective use, however subsequent patterns demonstrate a simultaneous, integrated utilisation of diverse educational offerings. Time series data on usage patterns can depict the progression of student learning based on Web logs. Initially, this adjustment in our case study came as a surprise.

Table 4: Time series (by week) based on the results of session classification

	Lecture	Tutorial	Lab	Download	Look Up
Week 1	35	22	13	19	11
Week 2	29	22	33	9	7
Week 3	20	27	40	5	8
Week 4	42	33	10	9	6

The data shown in Table 4 demonstrates the shift in behavioural trends throughout a four-week period, namely the initial four weeks of the academic semester. At first, learners mostly concentrated on online lectures and downloaded course material. During the following weeks, there was an increased utilisation of the tutorial and lab components, indicating the commencement of practical sessions and course materials in weeks two and three of the term, as observed in the case study.

7. CONCLUSIONS

Web usage mining is a valuable tool for understanding and evaluating Web-based learning and training technology systems. It provides access to latent knowledge hidden in access logs and can be used in explanatory, predictive, and generative styles. Explanatory techniques help understand student learning in a novel environment, while predictive techniques confirm expectations and validate designs. Generative styles help improve design by comparing instructor expectations with actual behavior. Web usage mining can also identify learner types, such as Kolb's Learning Style Inventory, and provide immediate feedback for instructors. It can also be used to monitor students throughout the term, providing immediate feedback. Education-specific mining techniques can help improve instructional design and confirm delivery-related decisions, enabling instructors to provide quality instruction and improve learning experiences. They also allow instructors to maintain relationships with learners and react to unforeseen events and necessary changes.

REFERENCES

1. Sims, R. Interactive learning as "emerging" technology: A reassessment of interactive and instructional design strategies. *Australian Journal of Educational Technology*, 13(1), pp. 68-84, 1997.

2. Weston, T.J. & Barker, L. Designing, Implementing, and Evaluating Web- Based Learning Modules for University Students. *Educational Technology*, 41(4), pp. 15-22, 2001.

3. Northrup, P. A Framework for Designing Interactivity into Web-based Instruction. *Educational Technology*, 41(2), pp. 31-39, 2001.

4. Oliver, M. (ed.) Special Issue on Evaluation of Learning Technology.

5. *Educational Technology & Society*, 3(4), 2000.

6. Chang, G., Healey, M.J., McHugh, J.A.M., & Wang, J.T.L. (2001). *Mining the World Wide Web - An Information Search Approach*. Kluwer Academic Publishers: Boston, 2001.

7. Scime, A. *Web Mining: Applications and Techniques*. Idea Group Inc.: Hershey, 2004.

8. Spiliopoulou, M. Web Usage Mining for Web Site Evaluation.

9. *Communications of the ACM*, Vol. 43(8), pp. 127-134, 2000.

10. Zaïane, O.R. Web Usage Mining for a Better Web-based Learning Environment. *Proc. Conference on Advanced Technology for Education CATE'2001*, pp. 60-64, 2001.

11. Pahl, C. & Donnellan, D. Data Mining for the Evaluation of Web-based Teaching and Learning Environments. *Proc. E-Learn 2002 World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education*. AACE, 2002.

12. Zaïane, O.R. & Luo, J. Towards Evaluating Learners' Behaviour in a Web- Based Distance Learning Environment. *Proc. International Conference on Advanced Learning Technologies ICALT'01*. IEEE, 2001.

13. Xu, L., Pahl, C., & Donnellan, D. An evaluation Technique for Content Interaction in Web-based Teaching and Learning Environments. In *Proceedings International Conference on Advanced Learning Technologies ICALT 2003*. IEEE Press, 2003.

14. Moore, M.G. Three types of interaction. *The American Journal of Distance Education*, 3(2), pp. 1-6, 1992.

15. Ohl, T.M. An Interaction-Centric Learning Model. *Journal of Educational Multimedia and Hypermedia*, 10(4), pp. 311-332, 2001.

16. Pahl, C. A Conceptual Architecture for Interactive Educational Multimedia. In J. Ma (ed). *Web-based Intelligent e-Learning Systems: Technologies and Applications*. Idea Group Inc.: Hershey, 2005. (in press)

17. Jonassen, D.H., & Mandl, H. (eds) *Designing Hypermedia for Learning*. Springer-Verlag: Berlin, 1990.

18. Dix, A., Finlay, J., Abowd, G., & Beale, R. *Human-Computer Interaction*. Prentice Hall: London, 1993.

19. Elsom-Cook, M. *Principles of Interactive Multimedia*. McGraw-Hill: London, 2001

20. Hirami, A. The Design and Sequencing of e-Learning Interactions: A Grounded Approach. *International Journal on E-Learning*, 1(1), 2002.

21. Pahl, C., Barrett, R., & Kenny, C. Supporting Active Database Learning and Training through Interactive Multimedia. *Proc. International Conference on Innovation and Technology in Computer Science Education ITiCSE'04*. ACM, 2004.

22. Agrawal, R. & Srikant, R. Mining Sequential Patterns. Proc. 11th

23. *International Conference on Data Engineering ICDE*, Taipei, Taiwan, 1995.

24. Cooley, R., Tan, P.-N., & Srivastava, J. Discovery of Interesting Usage Patterns from Web Data. *Proc. Web-based Knowledge Discovery and Data Mining WEBKDD'99*, pp. 163-182. London: Springer-Verlag, 1999.

25. Hay, B., Wets, G., & Vanhoof, K. Clustering navigation patterns on a website using a Sequence Alignment Method. *Proc. 7th International Joint Conference on Artificial Intelligence IJCAI'2001*, 2001.

26. Mullier, D., Hobbs, D., & Moore, D. Identifying and Using Hypermedia Browsing Patterns. *Journal of Educational Multimedia and Hypermedia*, 11(1), pp. 31-50, 2002.

27. Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations* 1(2), pp. 12-23, 2000.

28. Donnellan, D. *User Session Classification Tool for the Analysis of Web Server Logs*. M.Sc. Dissertation. Dublin City University. School of Computing, 2002.

29. Xu, L. *Analysis of Behavioural Patterns using Web Mining*. M.Sc. Dissertation. Dublin City University, School of Computing, 2003.

30. Pahl, C. An Evaluation of Scaffolding for Virtual Interactive Tutorials. Proc. 7th E-Learn 2002 Conference, Montreal, Canada. AACE. 2002.

Corresponding Author

Saarthak Singh*

Student, Class 12th, Inspiration Public School, Haldwani, Uttrakhand, India

Email: saarthak.singh.2304@gmail.com