EDUCATIONAL DATA MINING AND PREDICTION OF LEARNING DISABILITIES

1Hanumathappa. M, Associate Professor, Department of Computer Science & Applications, Bangalore University, Bangalore-56

2Margaret Mary.T, Lecturer in MCA Dept., Sambhram Academy of Management Studies, Bangalore 97.

ABSTRACT

Data mining is a collection of techniques for efficient automated discovery of previously unknown, valid, novel, useful and understandable patterns in large databases. Conventionally, the information that is mined is denoted as a model of the semantic structure of the datasets. The model might be utilized for prediction and categorization of new data. In recent years the sizes of databases has increased rapidly. This has lead to a growing interest in the development of tools capable in the automatic extraction of knowledge from data. The term Data Mining or Knowledge Discovery in databases has been adopted for a field of research dealing with the automatic discovery of implicit information or knowledge within databases. The aim of this paper is to discuss about the educational data mining and its application. First we describe the techniques of education data mining and further we are discussing a case study, which is application of educational data mining for the prediction of learning disabilities in a child.

Index terms- Prediction, Clustering, Relationship mining, Distillation for human judgment.

1. Introduction

Educational data mining is emerging as a research area with a suite of computational and psychological methods and research approaches for understanding how students learn. New computer-supported interactive learning methods and tools—intelligent tutoring systems, simulations, games—have opened up opportunities to collect and analyze student data, to discover patterns and trends in those data, and to make new discoveries and test hypotheses about how students learn. Educational data mining generally emphasizes reducing learning into small components that can be analyzed and then influenced by software that adapts to the student [11]. Student learning data collected by learning systems are being explored to develop predictive models by applying educational data mining methods that classify data or find relationships. These models play a key role in building adaptive learning systems in which adaptations or interventions based on the model’s predictions can be used to change what students experience next or even to recommend outside academic services to support their learning [8].
An important and unique feature of educational data is that they are hierarchical. Data at the keystroke level, the answer level, the session level, the student level, the classroom level, the teacher level, and the school level are nested inside one another. Other important features are time, sequence, and context. Time is important to capture data, such as length of practice sessions or time to learn. Sequence represents how concepts build on one another and how practice and tutoring should be ordered. Context is important for explaining results and knowing where a model may or may not work. Methods for hierarchical data mining and longitudinal data modeling have been important developments in mining educational data.

Goal of educational data mining are:

- Predicting students’ future learning behavior by creating student models that incorporate such detailed information as students’ knowledge, motivation, meta cognition, and attitudes;
- Discovering or improving domain models that characterize the content to be learned and optimal instructional sequences;
- Studying the effects of different kinds of pedagogical support that can be provided by learning software;
- Advancing scientific knowledge about learning and learners through building computational models that incorporate models of the student, the domain, and the software’s pedagogy.

To accomplish these four goals, educational data mining research uses the five categories of technical methods described below.

1. **Prediction** entails developing a model that can infer a single aspect of the data (predicted variable) from some combination of other aspects of the data (predictor variables). Predictive models have been used for understanding what behaviors in an learning environment—participation in discussion forums, taking practice tests and the like—will predict which students might fail a class. Prediction shows promise in developing domain models, such as connecting procedures or facts with the specific sequence and amount of practice items that best teach them, and forecasting and understanding student educational outcomes, such as success on posttests after tutoring [2].

2. **Clustering** refers to finding data points that naturally group together and can be used to split a full dataset into categories. Examples of clustering applications are grouping students based on their learning difficulties and interaction patterns, such as how and how much they use tools in a learning management system, and grouping users for purposes of recommending actions and resources to similar users. Data as learning resources, student cognitive interviews, and postings in discussion forums can be analyzed using techniques for working with unstructured data to
extract characteristics of the data and then clustering the results. Clustering can be used in any domain that involves classifying, even to determine how much collaboration users exhibit based on postings in discussion forums [1].

3. **Relationship mining** involves discovering relationships between variables in a dataset and encoding them as rules for later use.
   - **Association rule mining** can be used for finding student mistakes that co-occur, associating content with user types to build recommendations for content that is likely to be interesting, or for making changes to teaching approaches [9]. These techniques can be used to associate student activity, in a learning management system or discussion forums, with student grades or to investigate such questions as why students’ use of practice tests decreases over a semester of study.
   - **Sequential pattern mining** builds rules that capture the connections between occurrences of sequential events, for example, finding temporal sequences, such as student mistakes followed by help seeking. This could be used to detect events, such as students regressing to making errors in mechanics when they are writing with more complex and critical thinking techniques. Key educational applications of relationship mining include discovery of associations between student performance and course sequences and discovering which pedagogical strategies lead to more effective or robust learning. This latter area—called teaching analytics—is of growing importance and is intended to help researchers build automated systems that model how effective teachers operate by mining their use of educational systems.

4. **Distillation for human judgment** is a technique that involves depicting data in a way that enables a human to quickly identify or classify features of the data. This area of educational data mining improves machine-learning models because humans can identify patterns in, or features of, student learning actions, student behaviors, or data involving collaboration among students. This approach overlaps with visual data analytics (described in the third part of this section).

5. **Discovery with models** is a technique that involves using a validated model of a phenomenon (developed through prediction, clustering, or manual knowledge engineering) as a component in further analysis. A sample student activity discerned from the data was “map probing.” A model of map probing then was used within a second model of learning strategies and helped researchers study how the strategy varied across different experimental states. Discovery with models supports discovery of relationships between student behaviors and student characteristics or contextual variables, analysis of research questions across a wide variety of contexts, and integration of psychometric modeling frameworks into machine-learned models.

Using these techniques, educational data mining researchers can build models to answer such questions as:
2. Prediction of Learning Disability – A Education Data Mining Implementation Example

Learning disability is a general term that describes specific kinds of learning problems. Learning disabilities are formally defined in many ways in many countries. However, they usually contain three essential elements: a discrepancy clause, an exclusion clause, and an etiologic clause. The discrepancy clause states there is a significant disparity between aspects of specific functioning and general ability; the exclusion clause states the disparity is not primarily due to intellectual, physical, emotional, or environmental problems; and the etiologic clause speaks to causation involving genetic, biochemical, or neurological factors.

The most frequent clause used in determining whether a child has a learning disability is the difference between areas of functioning [5]. When a person shows a great disparity between those areas of functioning in which she or he does well and those in which considerable difficulty is experienced, this child is described as having a learning disability [7]. A learning disability can cause a child to have trouble in learning and using certain skills. The skills most often affected are: reading, writing, listening, speaking, reasoning and doing math [6]. Learning disabilities vary from child to child. One child with LD may not have the same kind of learning problems as another child with LD. There is no "cure" for learning disabilities [10]. They are life-long. However, children with LD can be high achievers and can be taught ways to get around the learning disability. With the right help, children with LD can and do learn successfully [7].

2.1 Data Mining Approach For Learning Disability Prediction

Various data mining techniques can be used to predict the learning disability. In this paper we are discussing about decision tree method, which can be used to predict the learning disability. The decision tree is a flow chart like structure, where each internal node denotes a test on an attribute, each branch of the tree represents an outcome of the test and each leaf node holds a class label[3][4][5]. The topmost node in a tree is the root node. Decision trees are powerful and popular tool for classification and prediction. It is a classifier in the form of a tree structure where each node is either a leaf node-indicates the value of the target attribute of examples or a decision node –specifies some test to be carried out on a single attribute-with one branch and sub tree for each possible outcome of the test.
Classifiers do not require any domain knowledge or parameter setting and therefore is appropriate for exploratory knowledge discovery. Decision tree can handle high dimensional data [4]. The learning and classification step of decision tree are simple and fast. A decision tree can be used to classify an example by starting at the root of the tree and moving through it until a leaf node, which provides the classification of the instance [12]. A tree building process starts by selecting an attribute to place at the root node and at each succeeding level the subset generated by proceeding levels are further partitioned until it reaches a relatively homogeneous terminal node or leaf node. The condition attribute, that induces most amount of entropy reduction and information gain are placed closer to the root node. Table 1 show the attributes, which can be used to construct the decision tree.

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Attribute</th>
<th>Signs. &amp; Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DR</td>
<td>Difficulty with Reading</td>
</tr>
<tr>
<td>2</td>
<td>DS</td>
<td>Difficulty with Spelling</td>
</tr>
<tr>
<td>3</td>
<td>DH</td>
<td>Difficulty with Handwriting</td>
</tr>
<tr>
<td>4</td>
<td>DWE</td>
<td>Difficulty with Written Expression</td>
</tr>
<tr>
<td>5</td>
<td>DBA</td>
<td>Difficulty with Basic Arithmetic Skills</td>
</tr>
<tr>
<td>6</td>
<td>DHA</td>
<td>Difficulty with Higher Arithmetic Skills</td>
</tr>
<tr>
<td>7</td>
<td>DA</td>
<td>Difficulty with Attention</td>
</tr>
<tr>
<td>8</td>
<td>ED</td>
<td>Easily Distracted</td>
</tr>
<tr>
<td>9</td>
<td>DM</td>
<td>Difficulty with Memory</td>
</tr>
<tr>
<td>10</td>
<td>LM</td>
<td>Lack of Motivation</td>
</tr>
<tr>
<td>11</td>
<td>DSS</td>
<td>Difficulty with Study Skills</td>
</tr>
<tr>
<td>12</td>
<td>DNS</td>
<td>Does Not Like School</td>
</tr>
<tr>
<td>13</td>
<td>DLL</td>
<td>Difficulty in Learning a Language</td>
</tr>
<tr>
<td>14</td>
<td>DLS</td>
<td>Difficulty in Learning a Subject</td>
</tr>
<tr>
<td>15</td>
<td>STL</td>
<td>Slow To Learn</td>
</tr>
<tr>
<td>16</td>
<td>RG</td>
<td>Repeated a Grade</td>
</tr>
</tbody>
</table>

**Table 1:- List of Attributes**

2.2 Results from decision tree

Decision tree induction is one of the simplest, and yet most successful forms of learning algorithm. It serves as a good introduction to the area of inductive learning, and easy to implement [13]. A decision tree takes as input an object or situation described by a set of attributes and returns a decision. A divide and conquer approach to the problem of learning from a set of independent instances leads naturally to a style of representation called decision tree. The basic idea behind the decision tree learning algorithm is to test the most important attribute first;
by most important we mean the one that makes the most difference to the classification of an example. That way, we get to the correct classification with a small number of tests, meaning that all paths in the tree will be short and the tree as a whole will be small [13].

Initially we evaluate the worth of an attribute by measuring the information gain ratio with respect to the class. Attributes are then ranked by their individual evaluations by using in conjunction with gain ratio, entropy, etc. The decision tree formed based on the data set will look like as in figure 1 below.

![Figure 1: Learning Disability Decision Tree](image)

### 3. Conclusion

Now, with advances in mining techniques gives the possibilities to harness the power of feedback loops at the level of individual teachers and students. Measuring and making visible students’ learning and assessment activities open up the possibility for students to develop skills in monitoring their own learning and to see directly how their effort improves their success. This kind of system helps the teachers to predict the performance of the student at early stage. Teachers gain views into students’ performance that help them adapt their teaching or initiate interventions in the form of tutoring, tailored assignments, and the like. Such kinds of systems enable educators to quickly see the effectiveness of their adaptations and interventions, providing feedback for continuous improvement.

### References


