Adaptive Assessment Based on Machine Learning Technology

Irena Nančovska Šerbec, Alenka Žerovnik, Jože Rugelj

Faculty of Education, University of Ljubljana

Key words: adaptive assessment, regression trees

Abstract:

In the e-learning environment we use adaptive assessment based on machine learning technology called regression tree. In the paper we analyse the properties of the regression tree assessment.

1 Introduction

Educational assessment is the process of evaluation and documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs. In the e-learning environment we need assessment which is, besides being valid and reliable, quick and automated. These last two properties could be achieved by means of adaptive assessment.

The main idea of an efficient adaptive assessment is to start with a problem (exercise) of medium difficulty and then to move on to other problems, depending on the respective correctness of the response (answer). In case of correct answer, the next problem will be more difficult than the previous one. If the learner fails to answer the first question correctly, easier problems will be selected. This querying ends whenever sufficient information concerning the student’s knowledge state has been gathered. Obviously, such an adaptive procedure has the advantage that the assessment process is shortened and is more efficient, because it avoids to present the learner with a large number of problems. This is to a great extent analogous to teachers’ intuitive approach in traditional assessment settings.

Many research communities work on adaptability of learning objects in the e-learning environment. Their work is based on different psychological and mathematical theories such as Knowledge Space Theory (KST), Item Response Theory (IRT) in Computer Adaptive Testing (CAT), Multistage Testing (MST), and Educational Data Mining or Edumining (EDM).

Knowledge Space Theory (KST) is a theory in mathematical psychology invented by Doignon and Falmagne (1985). KST provides a mathematical framework for realizing adaptability in technology enhanced learning environments by formalizing prerequisites in learning and assessment [9]. KST offers a model for structuring domains of knowledge based on prerequisite relationships between content objects. This theory allows to be represented and efficiently assessed the knowledge state of an individual learner [7]. With its recent competency-based extensions, such as the competence performance approach [10], [11], there have been attempts to enhance this framework with an inclusion of the concepts that influence performance, such as knowledge or skills. In the frame of KST, we distinguish between deterministic, non-deterministic, and probabilistic assessment procedures [6].

Computerized adaptive testing (CAT) or tailored assessment is a method for administering tests that adapts to the examinee's ability level [1]. Compared to static tests, with a fixed set of items administered to all examinees, CAT requires fewer test items to arrive at equally accurate scores. CAT currently relies heavily on Item Response Theory (IRT) [1], [3]. One alternative is multistage testing (MST), also called computerized adaptive sequential testing (CAST). MST is a test administration method whereby examinees are routed to later tests on the basis of their performances on routing tests [2]. Other alternative are empirically based, nonparametric adaptive testing algorithms [8]. In the paper presents a nonparametric, regression-tree-based algorithm for adaptive testing [4], [12], [16].
From the e-learning point of view, data mining applications in e-learning could be divided into the following categories [5]:

1. Applications dealing with the assessment of students’ learning performance.
2. Applications that provide course adaptation and learning recommendations based on the students’ learning behavior.
3. Approaches dealing with the evaluation of learning material and educational Web-based courses.
4. Applications that involve feedback to both teachers and students of e-learning courses, based on the students’ learning behavior.
5. Developments for the detection of atypical students’ learning behavior.

We range our approach into categories 1 and 4. Machine learning methodology for generating regression trees is a data mining method, which means that it could be treated as a part educational data mining (EDM). In the next section we will present regression trees methodology as bases for assessment.

2 Regression trees methodology

Regression trees are generated or induced by using datasets that consist of validated statically generated tests with wide set of exercises. Through the process of building of regression trees we capture the knowledge structure of the statically tested students. Process of assessment is routed by regression tree structure and it is ended by student’s classification by reaching the tree leaf.

The procedure of generating regression tree from training set is called induction of the tree. We start with blank tree and whole set of training objects. Then on every step with the help of heuristic evaluation function we choose an attribute (exercise), which wasn't used jet [15]. If there is not enough training objects or if the data contains missing values, then it usually leads to overfitting. The result of that are large trees with a lot of unimportant branches. That's why it is important where to stop growing of the tree by using the procedures for pruning unimportant branches. In most of the algorithms today we can find a mechanism called n-fold cross validation, which is one of the mechanisms for preventing the overfitting of the model to the training set. The final model is built on a whole training set.

The algorithm for tree induction, which selects the attributes (problems) in the nodes, utilizes the high dependency among attributes (problems).

In KTS the knowledge state of an individual is identified with the subset of problems of the knowledge domain that this individual is capable of solving. In regression tree framework the knowledge state of the student is presented by the subset of correctly solved problems “lying” in the particular nodes on the path from the root to the leaf. Collection of knowledge states of a given knowledge domain is called a knowledge structure [7]. The knowledge structure could be captured by the tree.

3 Experimental settings

In the beginning, after the presentation of the thematic unit by a professor (in our case elementary mathematics “Expressions”, “Introduction into programming with programming language Pascal” or “Common knowledge about European Union” [4], [12], [16]), we carry out static testing of students with 20-30 questions or exercises. Each exercise is randomly selected from the set of items of the selected topic. We use items scored correct (associated with number of points) or incorrect. Questions of type: “choose correct answer” or “fulfill the answer” are solved for each domain.

For example, the domain of EU consists of web-inquire results. Static tests were sent to students and teaching staff on a program study of computer science at the Faculty of Education. Solved tests were anonymously stored in MySQL database. We have collected 120 instances. The test contained 20 questions, each of them marked with 5 points. The values of the attributes were numerical (except the attribute class which was descriptive) and they presented achieved points of an individual student for each individual question. The maximum number of points for the test was 100. Attribute Success was the rating of student into three classes as regards the points achieved on the test.
For modeling we use the open source software tool WEKA [14] and tree-induction algorithm J4.8 included in the WEKA.

After the generation of regression trees we realize the adaptive assessment system based on the tree structure. Table 1 presents the classification accuracy of the adaptive assessment system. The system is asking questions (exercises) in successive order, one after another. In the background, the algorithm is following the structure of regression tree. The question in the root of the tree is given to all students. After they answer particular question, the algorithm chooses the next one regarding to the correctness of the current answer. The testing is finished after the leaf of the regression tree is reached. It means that the student’s knowledge is successfully rated.

Table 1: Classification accuracy (in %) of regression tree on each domain (10-fold cross validation/whole training set)

<table>
<thead>
<tr>
<th>Domain/Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressions (math.)</td>
<td>76/100</td>
</tr>
<tr>
<td>Programming</td>
<td>95/98</td>
</tr>
<tr>
<td>EU knowledge</td>
<td>88/96</td>
</tr>
</tbody>
</table>

4 Conclusions

We modeled the students’ knowledge captured in the tests on knowledge about the selected thematic units. Regression trees are interesting models not only because they are predictors of learning outcomes but because of the transparency of their structure. As knowledge structures, they are interesting for teachers as feedback information about the learning behavior of their students. For students, they are interesting as paradigm for adaptive formative or self-assessment. Other possible purpose of such a tree could be directing students in the learning process with purpose to improve his/her learning outcome. Students classified in the left leaves in the tree structure, could be suggested to adopt topics which distinguish them from the right or better classified colleagues. One limitation of the tree-based approach is that the problem related to the topic in the root of the tree is given to all students [8]. Another limitation is that the tree should be built off-line, before the adaptive assessment and its structure doesn’t adapt to the knowledge shown through the adaptive assessments. Another limitation is that the students should answer all the questions in the nodes on the path from the root to the leaf, without possibility to omit some of them. Weakness of these models is that their classification accuracy is in average around 90%. Wrong classification can be especially
problematic in the cases of successful students with low self-confidence. Besides this, we can not predict how the rating will influence on the motivation of students.

Regression trees as basis of adaptive assessment could be related to the theories for adaptive assessment such as CAT and KTS. Our adaptive assessment system is in its initial testing phase and a lot of improvements are still needed. Constructed models are suitable mainly for experimental use.

References:


Author(s):

Dr. Irena Nančovska Šerbic, Alenka Žerovnik, Teaching Assistants and Prof., Dr. Jože Rugelj, Assoc. Prof.
University of Ljubljana, Faculty of Education, Kardeljeva ploščad 16, 1000 Ljubljana
e-mails: {irena.nancovska, alenka.zerovnik, joze.rugelj}@pef.uni-lj.si