Assessing mathematics learning achievement using hybrid rough set classifiers and multiple regression analysis

You-Shyang Chen [a,*], Ching-Hsue Cheng [b]

[a] Department of Information Management, Hwa Hsia Institute of Technology, 111, Gong Jhuan Road, Chung Ho District, New Taipei City 235, Taiwan, ROC
[b] Department of Information Management, National Yunlin University of Science and Technology, 123, Section 3, University Road, Touliu, Yunlin 640, Taiwan, ROC

ABSTRACT

Education is recognized as the key to individual success. Particularly, elementary education is vital for providing students with basic literacy and numeracy, as well as establishing foundations in mathematics, language, science, geography, history, and other social sciences. Mathematics is fundamental to numerous fields with real life applications, including natural science, engineering, medicine, and social sciences; therefore, student mathematics-learning achievement (MLA) in elementary school is valuable. This study aims to eliminate wastage of educational resources and seek suitable hybrid models for application to education. This study proposes an integrated hybrid model based on rough set classifiers and multiple regression analysis and provides a new trial of such a hybrid model to process MLA problems for elementary schools and their teachers. The proposed model is illustrated by examining a dataset from an elementary school in Taiwan. The experimental results reveal that the proposed model outperforms the listing methods in both classification accuracy and standard deviation. The rough set LEM2 (Learning from Examples Module, version 2) algorithm generates a set of comprehensible decision rules that can be applied in a knowledge-based education system designed for interested parties. Consequently, the analytical results have important implications for strategies related to mathematics teaching/learning and support to achieve teaching goals related to mathematics education.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Education aims to improve quality of life by ensuring future competence, and develops basic learning skills, including reading, writing and arithmetic, and life skills via school systems. The school system in Taiwan primarily comprises primary school, elementary school, secondary school, and university, and provides a teacher-student relationship through which teachers are supported with resources to teach the curriculum in a more effective and child friendly manner, particularly in elementary school. Elementary education involves supporting the basic thinking, learning, literacy, and mathematics skills of children aged 7–13 years old. Mathematics represents an important foundation among the courses taught in elementary education, since it is the language of technology and good mathematics skills are crucial to future life situations. Abstraction and logical reasoning complement the skills of counting, calculation and measurement, to produce the discipline known as 'maths.' Maths is an essential tool in natural science, engineering, medicine, and social sciences. Throughout the education system, students are enabled to become high achievers who are more confident and motivated in maths through regular exposure to learning situations. The goal is for education systems to maximize student learning potential and boost student motivation to learn mathematics.

Teacher–learner relationships are stronger in education systems where students have diverse cultural backgrounds. Traditionally, teachers focused on teaching and students on learning, and teachers additionally provided interactive materials to facilitate student learning achievement. The question of learning outcomes in maths may vary according to the roles of student. For teachers and students, appropriate assessment feedback provides a useful reference in correcting mistakes (such as, teaching and learning faults). Assessments involve identifying, gathering, and interpreting information about student learning outcomes. This information is then used to guide teaching, offer feedback on the progress of student learning, and control the direction of ongoing teaching/learning strategies. Restated, schools use assessment systems to provide information to correct teaching and learning in maths. Kellaghan and Greaney [25] categorized assessments of student achievements into four levels: class level assessment, school level assessment, external examinations, and national and international assessments. Hanushek and Pace [18] reported that poor education negatively affects future income and productivity for individual students, results in low learning achievement,

---

* Corresponding author. Tel.: +886 2 8941 5143; fax: +886 2 8941 5142.
E-mail address: ys_chen@cc.hwh.edu.tw (Y.-S. Chen).

limits student progress, and even harms national social and economic development. Obviously, education systems exert numerous crucial influences on schools, and mathematical assessment methods and products have varied levels and purposes [32,43,44,51]. This study focuses on school level assessment, and therefore, test or exam scores are used to assess student achievements in mathematics learning. A valuable issue is created by identifying critical variables (attributes) that influence mathematics-learning achievement (MLA), clarify the relationships among variables, and ultimately generate decision rules as knowledge-based systems for interested parties, such as schools, teachers, students, and parents or guardians. From this perspective, the influences on MLA are explored and the associated problems emphasized.

The exploration of factors related to MLA can be modeled as classification analysis and is formatted in statistical methods and artificial intelligence (AI) techniques. Earlier studies have detailed the statistical methods of addressing learning achievement problems [10,22,38,54], which rely on restrictive assumptions, such as linear separability, multivariate normality, and independence of the predictive variables [41]. Unfortunately, multivariate normality assumptions regarding independent variables are frequently violated, and models become increasingly complex given nonlinear relationships in the input/output dataset [53]. Recently, commentators thus have recommended employing more efficient methods for processing such classification problems [2,3,57] based on AI techniques that have emerged as alternatives to statistical methods for solving real-world problems. The common tools used in AI techniques for classification problems, such as rough set theory (RST), have become significant research trends and have attracted attention from both practitioners and researchers [42,59]. Particularly, intelligent hybrid systems comprise several models or procedures. Moreover, such ensemble classifiers have been proven to outperform stand-alone models [41]. Given the limitations of AI techniques when applied to stand-alone models, it is worth proposing an intelligent hybrid system to amplify the advantages of the individual models and minimize their limitations. Interest in designing and employing various intelligent hybrid systems thus has increased considerably over the past decade [41]. Furthermore, based on review of the literature [4,35,42], this study notes that a significant trend in knowledge discovery is to build different rule-based models to explain the used dataset, an approach that has been reasonably effective. Numerous previous studies have found the performances of various classifiers to depend heavily on data and context [24]; therefore, designing reliable hybrid models that can be applied to the education industry to address classification problems is an interesting issue. Research on integrating AI techniques (for example, RST method) and statistical methods (for example, multiple regression analysis, MRA) to tackle classification problems involved in assessing MLA is lacking, and little experience of applying such techniques exists in the education industry. Given the above reasons, this study thus focuses on filling these knowledge gaps and conducting an intelligent hybrid model to solve a real life application problem involving MLA.

As stated above, this study focuses on the 'application' aspect of the hybrid rough sets-based model, and aims to construct an intelligent hybrid model primarily based on rough set classifiers and using multiple regression analysis to implement to classify MLA problems. This study focuses on the following three key topics: (1) Construct a hybrid model to assess the performance of the rough set classifiers in effectively classifying MLA in the education industry; (2) Identify the influences on MLA to support the needs that provide methods to increase the effectiveness of mathematics learning, and the needs that provide solutions to increase the teaching efficiency; and (3) Generate comprehensible decision rules applicable in a knowledge-based education system for elementary schools, teachers, students, and parents or guardians, to improve teacher–student relationships and communications.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature on mathematics learning achievement; Section 3 then introduces basic concepts of rough set theory and the rule extraction-LEM2 algorithm. Next, Section 4 describes the proposed model, including its algorithms, presents verification, and makes comparisons; discussion and findings are described in the next section. Conclusions are finally drawn in Section 6, and future research directions are discussed.

2. Related work

This section first reviews the issues related to assessing mathematics learning achievement.

A Chinese proverb holds that learning is a treasure that remains with its owner forever. Human learning is a mediator of the acquisition of new knowledge, behaviors, skills, values, or preferences, and may occur as part of education, interaction with others, or specific training. Learning can be classified as informal or formal. Formal learning occurs within a constructed teacher–student structure such as a school, namely an educational system, while informal learning is organized outside the formal learning system or during real life, and results from experience of day-to-day situations. Formal learning involves providing students with a solid grounding in basic thinking, learning and literacy, as well as offering various courses, such as language, science, mathematics, music, and art. Teachers offer students this formal learning in school, and mathematics is particularly important to formal learning. Mathematics is an essential requirement in real-world fields, such as natural science and social science; therefore, mathematics is closely associated with real life and is crucial for everybody.

From the perspective of schools, assessing the quality of mathematics learning is important, because understanding the factors related to student MLA is a priority. Mizala and Romaguera [31] showed that students with more educated parents and higher family income obtained higher mathematics scores than students with less educated parents and lower family income. Moreover, the effective homework has been shown to facilitate student achievement. Cooper et al. [6] proved that students with more positive attitudes to homework achieved higher mathematics grades than those with negative attitudes. Additionally, Cooper and Valentine [5] found that students who spent more time on homework exhibited higher academic achievement. Trautwein et al. [55] indicated that more frequent homework was associated with higher mathematics achievement. Additionally, House [21] identified that students exposed to more frequent explanations of mathematical rules and definitions achieved higher mathematics test scores. Finally, Whicker et al. [58] noted that cooperative learning was associated with higher mathematics achievement scores.

3. Theoretical background

This section introduces previous studies related to the technical foundation used to address mathematics learning achievement, including rough set theory and the rule extraction-LEM2 algorithm.

3.1. Rough set theory

RST, first proposed by Pawlak [36], used mathematical modeling to solve class data classification problems, and proved very useful for decision support systems, especially when the decision process involved hybrid data, vague concepts, and uncertain data. The rough set process involves starting with a relational database, a table of objects with attributes, and attribute values for each object.