March 4, 2016

An Expert System for the Prediction of Student Performance in an Initial Computer Science Course

Michael Kuehn
Jared Estad
Jeremy Straub
Tom Stokke
Scott Kerlin

Available at: https://works.bepress.com/jeremy_straub/319/
An Expert System for the Prediction of Student Performance in an Initial Computer Science Course

Michael Kuehn, Jared Estad, Jeremy Straub, Tom Stokke, Scott Kerlin
Department of Computer Science, University of North Dakota

Overview

There are many factors that play a part in how a student performs during a course. Some of these factors can be looked at in hindsight to help suggest why a student may or may not have done well in class. This poster assesses the potential of predicting student performance in an introductory computer science class using information about students' preparation, attitudes and study habits. An expert system has been utilized for this purpose. The expert system accepts data related to seven different categories of preparation, belief and attitude and, through the partial activation of multiple rules, predicts an outcome for each student on the post-test (which should correlate with and is used as a surrogate for the student’s final course grade). This poster presents our findings and correlation between certain factors relevant to students’ success. It shows how these factors can be used to predict a student’s grade and discusses the significance of being able to do so. Thus far, the results from the first sample of data appear promising. The work has shown that certain groups of self-assessment questions have clearly out performed other groups in predicting final test grades. The use of expert system techniques to predict student grades can provide an insight on different factors that affect student performance.

Knowledge of characteristics that may lead to poor performance allows instructors to potentially identify students who may be at risk of a low grade and need supplemental instruction, tutoring or other support at the outset of the semester.

Walkthrough

Upon receiving the opinion based questions that would be asked to the students. The first step was to find unique ways of categorizing them into manageable groups of data. An example would be to take question number 1 and question number 14 and find that they may imply that the student already knows something about the course. Or in other words they have pre-knowledge of what they are going to learn. In this example, question #1 may imply that the student has experience in advanced programming. For each student according to how they answer each question effects the percentage of confidence for the category specified. Then our second part of the expert system takes the largest two categories and asks the third part of the system for the results pertaining to those two categories. It will then make the prediction based on that result. While the third part of the system will have to change the values of the results each category gives based on whether or not the expert system has made a correct decision. For example if the system gives a percentage that is too high, then it will need to lower its percent to better reflect what students are actually achieving for a percent. The percent it gives and changes based off its true percent. While the second part of the system will adjust a 5% range above and below the true percent to predict.

Significance

Educational assessment is often a controversial subject. Disagreement exists over what specific results are desirable, how to achieve these most-desirable results for students, and this even extends as far as disagreement as to what results should be generated and assessed [1-3]. Some may have a (perhaps well-founded) fear that determining an evaluative criteria may allow an administration to ‘clean house’ of those not subscribing to an approach (or not performing under a particular assessment technique which does not lend itself to their approach or style of teaching). While there may be many different factors that affect a student’s grade positively and negatively, there is still not a suitable system for determining what factors play the most important role in determining student performance. This project explores how a well-defined assessment approach, evaluated with expert systems techniques can be used to predict student performance and potentially allow at-risk students to be identified for additional aid, prior to reaching a point of ‘no return’ for the class, where they cannot recover. This assessment not only has an evaluative/summative role, but can play a formative one as well.

Acknowledgement

This paper is based on, revised and updated from [1].

Tools

Within this project there are a couple reasons to store our questions into a database. With the values from each question stored within the database, the expert system can call each question and how many points it is worth. The use of the database allows for our data to be easy to manage and sort for specific questions. The database also allows for questions to be added and removed with ease to the expert system in case the test questions change. The next important use of the database is to store the results of the two higher categories, with it’s predictions, and the actual results. This may be useful to better understand which categories are being called, and whether or not specific categories are more correct in prediction then others. Allowing for an inference on how one could go about changing the expert system to help it make better decision for future work. An example would be if the top two categories where almost always the same two categories , then one could see that, that was the case. Then think of ways of changing the categories to better represent their answers.

Structure

There are four main parts that go into implementing our expert system. The first main part is the organization of the pre-evaluation self assessment questions into reliable categories. The second part is created to determine which categories should be used on the current students grade prediction. The third part is to create a final grade range that will be used as it’s predicted test grade range. This part needs to be allowed to change it’s values given based on the results it receives. It also has to have a combination of all the questions categories grouped by two. Then the fourth part is the creation of the database. This data base will be used to supply the questions to the expert system. It will also need to store the data that the expert system used to make it’s decisions, along with the actual results that were achieved.

Example Questions

Question #1: How experienced are you in creating computer programs that display 3D objects?
Circle 1-3: 1 2 3
(1 being not confident 3 being confident)

Question #2: How often do you comment your code?
Circle 1-3: 1 2 3
(1 not at all, 2 sometimes, 3 always)

References