Pre and Post Survey-Based Prediction of Results from Student Characteristics

Jared Estad
Michael Kuehn
Jeremy Straub
Eunjin Kim

Available at: https://works.bepress.com/jeremy_straub/276/
Pre and Post Survey-Based Prediction of Results from Student Characteristics
Jared Estad, Michael Kuehn, Jeremy Straub, Eunjin Kim
Department of Computer Science, University of North Dakota

Overview
This study was created to compare a large data set of students class surveys. In this data set, there exist both pre- and post-survey data. A set of knowledge based questions exist at the end of the survey. Our intention is to use the other data with in each survey to predict based on a singular set of pre-survey data a students potential for success throughout the course. As such, we need to account for data on the pre-survey and its knowledge based questions. Then, we need to account for how this data may interact with the data on the post-survey. Questions are almost the same pre-survey to post-survey, the only difference appears in the verb tenses of the post-survey questions to account for a student completing the course. As such we largely ignore this small difference; instead we consider how the data changes and what affect this has on the student’s success in the course.

Analysis of Figure 1
Figure 1 represents the basic system function that will determine our outcome after build up. Essentially, we will use 7 groups (each groups contains similar survey questions) and calculate a percentage value based on the total number a person can answer versus what number he or she did answer. For example, someone may answer 5/7 and 4/5, so total for that group they would have 9/12. Using this information, we will pass it to the middle section of our figure the “Retrieve Percent and Set Prediction Categories”. In this area we choose a set of two groups, these two groups have the highest percentages (based on points/n). From these we create the current knowledge base score and tie it to the pair of groups. If we need to deal with these groups again, the value may be changed depending on the situation. If pre-survey groups are the same and post-survey groups are the same then the value may need to be averaged. If the case is that the groups from the pre-survey change in the post-survey, the pre-survey pair will likely decrease slightly in value while the post-survey group pair will increase slightly in value. Hopefully, this will generate a number of presets that can be used to calculate an outcome of knowledge based questions based on a pre-survey. Meaning that after running n number of pairs of surveys through our data base and program enough information will be collected to allow us to use single pre-survey and find a general accuracy range for their knowledge of data.

Purpose of Project
The purpose of this project is to generate an expert system that is capable of using data to predict the class performance of an individual student with the help of prior data sets. The system will be able to work with multiple pre- and post-surveys, from there it must examine a single pre-survey and decide what percent of knowledge based questions that person would answer correctly on the post survey, which helps us understand said students success throughout the course.

Database Usage
Though its relevance to how our system works may not be highly apparent, we do use a database to contain the survey questions and answers. In our system, if we need to access this information we will end up pulling it from the database and running our system on that information. Right now the idea is to set each survey within an object. As such a pair of objects will be created for each set of surveys (the pre- and post-survey). We will continue to send calls to the data base until all the information has been cycled through and has been used with said object.

Conclusion
We hardly expect this system to be perfect and to be honest an infinite amount of time could be spent on developing other ways to implement a solution to the problem. Ideally, we will gain results that stand to give us roughly what we would expect to see depending on what such a survey is completed. In some ways, the surveys represent people and the differences in their personalities. Like the differences in personalities, we hope to obtain different results. In the future, testing may help us to improve the system, so that it can be fairly accurate.

Group Question Samples
Below are a number of sample questions. Each sample question comes from a group (in Figure 2) and represents a kind of question that may be contained in that group.

Experimental Method
1. We obtained empty tests that contained survey opinion questions only.
2. Another type of question, knowledge based questions exist on the survey.
3. Stemming from this, we will retrieve a number of correctly answer questions for each survey and the total number of questions.
4. Using the opinion questions, we generated 7 groups based on questions that we felt were similar in nature.
5. Each group represents a certain type of question which can be answered from either 1-5 or 1-7, with a higher number representing an higher understanding or confidence.
6. Using the sums of these groups and the percent of correctly answered knowledge based questions, we draw connections between different types of answers.
7. Ideally, we will uses pairs of the two highest value groups from both pre-survey and post-survey, to calculate the usual percent of correct knowledge based questions.

Groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>Guidance Related Problem Solving</th>
<th>Basic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>21, 36, 37, 50, 51, 53, 54</td>
<td>24, 45, 46, 47, 28, 29, 32</td>
<td></td>
</tr>
<tr>
<td>31, 30, 33, 34, 38, 43, 45, 57</td>
<td>49, 52, 55, 46, 28, 29, 32</td>
<td></td>
</tr>
<tr>
<td>20, 23, 35, 39, 43, 44</td>
<td>40, 41, 45, 48</td>
<td></td>
</tr>
<tr>
<td>22, 26, 58, 27</td>
<td>Number of questions correct</td>
<td></td>
</tr>
<tr>
<td>Number of Questions</td>
<td>Number of Questions</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2— Groupings