Programming: Factors that Influence Success Revisited and Expanded

Keith Quille

keith.quille@nuim.ie

Department of Computer Science,

Maynooth University.

Maynooth. Co. Kildare. Ireland

Aidan Mooney [S] **and Susan Bergin** [S, C]

aidan.mooney, susan.bergin {@nuim.ie}

Department of Computer Science,

Maynooth University.

Maynooth. Co. Kildare. Ireland

S = Senior Author, C = Corresponding Author

International Conference on Engaging Pedagogy (ICEP), College of Computing Technology, Dublin, Ireland, Dec. 3 & 4, 2015

Abstract

Computer science progression rates at third level institutes in Ireland from first year into second year were recently estimated at 74%. It is well acknowledged that a large contributor to the lower progression rates in computer science is that students are struggling to master fundamental concepts in their first programming module. The PreSS model, developed using multiple institutions over three years, is intended to be an early warning system and is able to predict, at a very early stage, with nearly 80% accuracy the likelihood of whether a student would be successful in an introductory programming module.

In this paper we want to replicate the accuracy of PreSS for two reasons. Firstly, the landscape, the population and the student profile has significantly changed over the last decade since PreSS was developed. Secondly, validation of any study is important as all too often studies are undertaken and never repeated potentially questioning their validity.

This paper documents two independent studies completed in two academic years (2013-2015), that will be used to validate the PreSS model. A number of factors that could potentially improve the prediction accuracy of PreSS further will be presented and their inclusion in a large international study will be discussed. This international study that commenced in September 2015 consists of nearly 1,500 students in 11 different institutions in Europe, varying from Universities and Colleges, to Community Colleges which cover the National Framework of Qualifications (NFQ) levels five up to eight.

Keywords:

Programming Performance, Predict Student Success, Validation.

1. Introduction

This paper presents a study which was undertaken to validate the model developed by Bergin (2006) over a three year period (2003-2006). Bergin successfully developed a computational model named PreSS (Predict Students Success), which predicted students' performance in an introductory to programing module based on three factors selected from a possible 25. The factors selected were: Programming self-esteem, Mathematical ability based on a high school mathematics exit examination and number of hours per week a student plays computer games (Bergin, Reilly, 2005). In general programming self-esteem was found to positively correlate to programming performance; this value was obtained using multiple methods, which are discussed in further detail in this Section and in Section 3. Mathematics ability was also found to positively correlate to programming performance, as mathematical ability increased so did the likely programming performance. This was not the case with the amount of time a student spent playing computer games which generally negatively correlated to the prediction, meaning as the amount of hours spent playing computer games increased, the likely programming performance decreased.

With this in mind it should be noted that the PreSS model is not this simple when it comes to the prediction of student success using all three factors. PreSS models the complex interactions of all three factors combined, which leads to the overall prediction. Bergin's study was undertaken in four different institutions in Ireland, consisting of one University, two Colleges and one Community College and was composed of three studies; pilot, main and epilogue. The total number of students across all of the studies was over 200, but when missing and invalid data was removed the final number of students used in the main study, replicated here, was 102.

PreSS modified the acknowledged measure of self-esteem used by Rosenberg (1965), to focus the ten questions administered towards programming self-esteem, which is presented in Section 3. Principal Component Analysis (PCA) was chosen as the data reduction algorithm to reduce these ten values to one value which numerically represents a student's self-efficacy compared to other students in the class. Machine learning classifiers attempt to predict outcomes based on training inputs and a mathematical formula. Examples of these classifiers include naïve Bayes (which was the classifier of choice, Bergin) and Support Vector Machines.

The process used in PreSS is highly sophisticated and further technical details can be found in Bergin (2006). Bergin was able to achieve an accuracy of 76.57% using 10 fold cross validation on the main study data set, which consisted of 102 students. She was able to achieve higher accuracies if the data set was examined with sub factors, for example, gender or per individual institution. The significance of this model is that it is universal, as it achieved this high accuracy across all institution types and gender.

2. Methodology

The two studies analysed in this paper, were carried out in the academic years of 2013-14 (S14) and 2014-15 (S15) respectively which were both hosted by a Community College. This is a similar institution to one of the institutions investigated in Bergin's (2006) study. The data samples were from NFQ level six courses, namely, a Computer Science course where all students completed the same "Introduction to Computer Programming" module. The language used was C# and this was the first time that this language was used with PreSS and in any study to predict programming performance. Java was the most common language used in Bergin's study and C# was not commonly used until after PreSS was developed. The grading criteria consisted of two programming assignments worth a total of 600 grade points and a written examination worth 400 grade points. Additional data was available as the two study groups also completed two other programming modules, "Windows Programming" and "Object Orientated Programming". Although the results of these other modules were not used in the immediate study, further research may be carried out with this additional data. Each student also completed an online survey that contained additional questions that are explored in Section 3. Students received consent information electronically and ethical approval was granted by Maynooth University to conduct these studies.

3. Instruments

This paper re-used the instruments/valid questions from the original study by Bergin, ensuring that the revisited analysis would be un-biased. PreSS uses a total 12 questions, resulting in three inputs to the model. The first input as mentioned in Section 1, was the modified self-esteem questionnaire, which accounted for 10 of the 12 questions.

The Rosenberg (1965) Self-Esteem (RSE) questionnaire was adapted to apply to programming self-esteem. The RSE scale is perhaps the most widely used self-esteem measure in social science research. The scale consists of 10 questions and has been shown to have generally high inter-item and test-retest reliability. Each of the questions was modified to relate to programming self-esteem and not to self-esteem directly, for example the first question was changed from "On the whole, I am satisfied with myself" to "On the whole, I am satisfied with my programming progress". To insure that the instrument continued to maintain high-reliability after modification Cronbach's alpha for the modified instrument were derived. The original instrument reported reliability in the range of 0.82 to 0.88, Rosenberg (1965). The modified instrument maintained high reliability with an alpha value of 0.91, Bergin (2006). The full set of modified RSE questions is presented in Table 1.

Q1) On a whole I am satisfied with my programming progress?
Q2) At times I think that I am no good at all at programming?
Q3) I feel that I have a number of good programming qualities?
Q4) I am able to complete programming items as well as most other students in my class?
Q5) I feel that I do not have much programming ability to be proud of?
Q6) I certainly feel useless at programming at times?
Q7) I feel that I am a person of worth, at least on a plane with other programmers in my class?
Q8) I wish I could have more respect for my programming ability?
Q9) All in all, I am inclined to feel that I am a failure at programming?
Q10) I take a positive attitude towards my programming ability?
Each question had four possible answers: 1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree

Table 1: The modified Rosenberg self-esteem questions with possible answers, as used in PreSS and PreSS#.

Principle Component Analysis (PCA) as mentioned in section 1, essentially performs an orthogonal transformation (rotation of data in multi-dimensions) to find the covariance eigenvectors with the largest eigenvalues, which represents the largest distribution or effect of the data set, hence selecting the principal component. This was implemented in PreSS, selecting eigenvalues > 1, to reduce the 10 questions of the programming self-esteem attribute to one principal component, thus obtaining a single value to accurately represent self-esteem. The two remaining questions consisted of: how many hours a student plays computer games per day and the student's Leaving Certificate Mathematics examination result (comparable to a high school exit examination). Both these values were normalised and used in the model.

A web based system named PreSS# (Quille, Bergin & Mooney, 2015), was developed in the academic year 2013-2014 to replace Bergin's paper based collection method with a web based version. PreSS# is a scalable cloud based system, with security and validation at its core thus it was the ideal tool to be used as the data collection method for the two independent studies used in this paper. This instrument is the primary data collection tool, used in the large international study which started in September 2015.

The final instrument used in this paper was a pre-survey, administered before the first study commenced with the S14 group of students. The aim of this pre-survey was to capture additional factors, which may have value for inclusion in the large study. These questions were selected based on hypothesis that some factors, which may have not existed or significantly impacted the model during Bergin's study, may be of value today and could contribute to the model's prediction ability. The pre-study questionnaire is presented in Table 2 and is discussed further in Section 7.

1.	Age Bracket?
2.	Gender?
3.	How many hours per day would you play computer games on a mobile device?
4.	If you play games on a mobile device, what genre of games do you play the most?
5.	How many hours per day would you play computer games on a Console, PC or laptop?
6.	If you play games on a console, PC or laptop, what genre of games do you play the most?
7.	How many hours per day would you use the internet (not including social media or messaging services)?
8.	What would your primary use of the internet consist of (not including social media or messaging services)?
9.	How many hours per day would you use a social networking service?
10.	If you do use social networking, what particular service do you use the most?
11.	How many hours per day would you use a messaging service?
12.	If you do use messaging service, what particular service do you use the most?

Table 2: The questions used in the pre-study questionnaire.

4. Background

Initial Study (S14)

The first study consisted of 34 students. There was no absent data entries so no student was excluded from the final sample which consisted of five females and 29 males. The overall final results showed the divide of weak to strong students was 16:18 respectively and the ratio between mature students (students under the age of 23) and non-mature students (students 23 years of age or older) was 12:22 respectively.

Second Study (S15)

The second study consisted of 26 students. There was no absent data entries so no student was excluded from the final sample. Four females and 22 males took part in the study. The overall final results showed the divide of weak to strong students was 9:17 respectively and the ratio between mature students and non-mature students was 3:23 respectively.

5. Results

The two studies were investigated to examine if they could replicate the results of Bergin's study, using the PreSS model. They were trained using all of Bergin's 102 student data samples from the main study. The process used in this investigation was identical to the process used by PreSS, using the same machine learning toolbox and methods to compute prediction. Both prediction accuracies using S14 and S15 were independently compared to Bergin's accuracy using a Welch's T Test and a binomial distribution to calculate the variance. The accuracy achieved in the S14 study was 76.47% producing a P value of 0.8592 and a T value of 0.1778 thus showing that there was no statistical difference between the two accuracies. The accuracy achieved in the S15 study was 76.92% producing a P value of 0.5669 and a T value of 0.5748 thus showing that there was no statistical difference between the two accuracies. This means that even though the PreSS model was developed nearly a decade ago, similar levels accuracy are achievable in a considerably changed landscape today.

6. Discussion

Both studies, S14 and S15, achieved accuracies statistically similar to that of Bergin's (2006) study using PreSS. This is a significant result as both studies consisted of very different student profiles with age, results and gender ratios differing significantly. Furthermore the student profiles from these studies grouped together differed significantly to Bergin's student profiles, due to the time lapse between studies, institution representation (Community Colleges represented less than 7% of institutions in Bergin's study) and programming language used. This considered, PreSS still produced statistically similar accuracies using S14 and S15. This implies that even though Bergin's PreSS model was developed over ten years ago, it is still valid, universal and not biased towards an institution, student profile, programming language, era or any other factor, thus always achieving the same high level of accuracy.

7. Further Investigations

During the two studies S14 and S15, 13 additional factors were included that the authors had selected based on proposed research questions and the pre-survey administered before S14 commenced (Table 2, Section 3), with three factors showing potential value: hours spent on social media, gender and age. Bergin (2006) suggested that gender may be of value to the model and based on her hypothesis, gender was incorporated into the additional factors recorded in S14 and S15.

It was also hypothesised that age may positively affect the accuracy of PreSS. This was based on prior experience that mature students may rate their self-esteem lower than its true value, resulting in an incorrect prediction. Research has shown that there is a positive relationship between age and programming ability/ attainment (Morrison, Murphy-Hill, 2013).

The use of social media was hypothesised as a new factor for programming prediction. This was based on the significant average time spent on social media as identified in the pre-survey which was almost two hours per day per student. Interestingly a bi modal distribution was observed due to the variance between mature and non-mature students social media habits. As this factor was not prominent ten years ago, the authors felt this factor had some merit for inclusion in the new studies.

8. Conclusions and Future Work

Bergin (2006) developed a semi-automated computational model named PreSS that could predict a student's academic performance in programming with an accuracy of about 80% after only 4-6 teaching hours. It is important to validate studies such as PreSS as all too often studies are validated once and never repeated and in our case almost a decade had passed since the models initial validation. In this paper Bergin's model for predicting student programming success has been validated using two new separate independent studies. This shows that the PreSS model is still fit for purpose and valid with accuracies of about 80% even with the change in landscape and technology.

The authors are currently conducting an international study in 11 institutions, with the aim of improving upon the PreSS model, with hypothesised increases in the model's accuracy with the introduction of the new factors. This is important as the selection of new factors to be included in the international study was limited, as it may discourage institutions from taking part if the time required to complete the survey is to too large or disruptive.

References:

Quille, K., Bergin, S., & Mooney (2015), A. PreSS#, A Web-Based Educational System to Predict Programming Performance.

Bergin, S., (2006), A computational model to predict programming performance, Maynooth University PhD thesis.

Bergin, S., & Reilly, R. (2005, February). Programming: factors that influence success. In ACM SIGCSE Bulletin (Vol. 37, No. 1, pp. 411-415). ACM.

Rosenberg, M., (1965), Society and the adolescent self-image, Princeton University Press.

Morrison, P., & Murphy-Hill, E. (2013, May). Is programming knowledge related to age? an exploration of stack overflow. In Mining Software Repositories (MSR), 2013 10th IEEE Working Conference on (pp. 69-72). IEEE.

Mooney, O., Patterson, V., O' Connor, M., Chantler, A. (2010), A study of progression in Irish higher education, Higher Education Authority, http://www.hea.ie/content/2010.

Watson, C., (2014), Failure Rates in Introductory Programming Revisited, 2014 conference on Innovation & technology in computer science education. ITiCSE'14, June 21–25, 2014, Uppsala, Sweden.