Original Article

IJEEE

International Journal of Electrical Engineering Education
2017, Vol. 54(2) 105–118

DOI: 10.1177/0020720916688484
journals.sagepub.com/home/ije

SAGE

Predicting performance of electrical engineering students using cognitive and non-cognitive features for identification of potential dropouts

Sara Sultana[1], Sharifullah Khan[1] and Muhammad A Abbas[2]

Abstract

The student dropout rate in universities is fascinating, especially among the students of Electrical Engineering. Even the most developed European countries face 40% to 50% dropout rate of engineering students during their first year, and the rate can be as high as 80% for some engineering disciplines. This problem calls attention of educators and university administration to take measures which can help in the reduction of the dropout rate and assist students in successfully completing their degree. Among many other solutions to control the student dropout rate, one is the adoption of a prediction mechanism whereby students can be warned about their potentially poor performance so that they can improve their performance resulting in better grades. Most of the existing prediction mechanisms apply various machine learning techniques on student cognitive features. In addition, non-cognitive features also have significant impact on students' performance; however, they have been sparsely applied for prediction. This research aims at improving the existing prediction mechanism by exploiting both cognitive and non-cognitive features of students for predicting their results. It has been found in the result analysis that addition of cognitive features increases prediction accuracies of decision tree; however, the addition does not play a significant role in other techniques. The study also identified the individual cognitive features that should be considered by students and universities to cater for drop outs.

[1]School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan
[2]Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan

Corresponding author:
Sharifullah Khan, School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan.
Email: sharifullah.khan@seecs.edu.pk

Keywords
Educational data mining, decision tree, logistic regression, Naïve Bayes, neural networks, student performance prediction, non-cognitive data

Introduction

Dropout or the termination of studies at a premature level is a problem faced by higher education institutions.[1,2] This dropout not only effects students and their institution, rather it has a broader negative impact on the society as well. Students in developing countries compete very hard to get admission in universities especially in the engineering universities. But any dropout case will lose an opportunity for oneself and the others who were not able to get admission at his place. Moreover, there is a negative impact to the increase in higher education literacy rate for which governments are subsidising and making financial and administrative arrangements. The identified reasons for the dropout problem include financial limitations, poor attendance, parental pressure, job opportunity, early marriage, demographic attributes, and poor academic performance.[3] An analysis of the literature reveals that more or less these reasons are similar among advanced and developing countries in higher education. But in developing countries, a negligible number of studies have been conducted to address this dropout problem.

Luckily, student dropout is not an uncontrollable problem. Early prediction of the "risk group" combined with counselling is considered to be an effective solution for the said problem. Historically, teachers and parents have used subjective and objective approaches to monitor and support a candidate dropout. This monitoring was based upon the academic grades and results of the student (i.e., cognitive features) achieved so far. Several studies also included the feedback from teachers of the early semesters and student mentors. Nearly all higher education institutions have academic information systems which keep record of the academic performance of every student during and before his induction.

A relatively newer field of educational data mining (EDM) explores how educational data, resides in academic information systems, can be used meaningfully to generate information that would help in making decisions to improve educational process.[4,5] In the context of dropout problem, EDM is used for prediction to timely cater any potential dropout. Several EDM techniques for student performance prediction have been reported such as regression analysis, decision tree, Naïve Bayes, and neural networks. A comparison of these techniques concludes that decision tree technique predicts student performance more accurately than all other techniques. However, existing work has only accounted previous performance data (i.e., cognitive features) of students to do the prediction. The reason is, the previous academic performance data are easily and readily available from academic information systems.

On the other hand, several studies have shown that other features like behaviour, attitude, and environment of a student, which are classified as non-cognitive features, also have significant impact on student performance.[6] The testing and evaluation organizations in the US are now using non-cognitive features like self-image, leadership, motivation, creativity, and extroversion to measure student performance in General Record Examinations (GRE) tests.[6] Non-cognitive features are important for performance prediction; however, the existing literature have sparsely considered these features for dropout predication. The present work will explore the effectiveness of the non-cognitive features in improving accuracy of the decision tress technique for the dropout prediction.

The paper is organized into different parts. In the first part, the cognitive and non-cognitive features are discussed that affect students' performance. This section covers literature review and critically analyses the data mining techniques that are often used to predict student performance. After that, the method of research is discussed where the participants, data sets, instrument for data collection, procedure for data collection, and analysis are discussed. Then, the prediction results are analysed and discussed followed by a section on conclusions.

Cognitive and non-cognitive features

The prediction of a dropout has been traditionally associated with the academic performance only (i.e., cognitive features). Features for prediction in the existing studies[7] included student's gender, age, marital status, number of children, occupation, computer literacy, marks in face-to-face meetings, and marks in examination. In another study,[8] the demographic and cognitive features included were gender, birth year, birth place, living place and country, type of previous, education, previous educational institute, previous education marks, university admittance year, admittance exam, achieved score, university majors, present semester, and total university score. Gender and previous aptitude results were used as predictors of student performance in Golding and McNamarah,[9] but these factors were found not very effective in student academic performance. Zafra et al.[10] used marks of assignments and quizzes as predictors for student performance.

On the other hand, few contemporary studies in the domain of educational psychology and higher education have shown evidence of associations between increased academic performance and non-cognitive features.[2,4,11] A research on impact of self-concept of the students revealed that it helps students in performing better and also helps teachers in imparting education in a better manner.[12] Another investigation shows that non-cognitive features like student behaviour and attitude towards assignments affect their performance and determine how much time they will allocate to a subject at home.[13] In a thesis, Flynt[14] stated that student hostility, extroversion, self-image and self-esteem, and leadership abilities help in determining their performance.

However, there has not been significant research done on the student dropout prediction using non-cognitive features.[2,4,11] This study incorporates several non-cognitive features for improving accuracy of the data mining techniques. These non-cognitive features are time management, self-concept, self-appraisal, leadership, and community support. Moreover, the present study hypothesized that with the consideration of non-cognitive features in addition to cognitive ones, the accuracy of the dropout predication can be increased. The idea is that different non-cognitive features helped in improving student performance[2,4,11] which in-turn will help in controlling dropout rate. For example, time management helps students keep track of hours that they spend in meaningfully increasing their academic performance. The student self-concept reveals how she views herself and whether she has a high or low image of herself. The realistic self-appraisal helps students in rationally evaluating the challenges and opportunities in their lives that can affect their academic performance. The leadership feature measures the governance skills of a student and the community support feature shows how much a student is or is not supported by the family, university, and surroundings in his personal matters.

EDM techniques

EDM is the process of extracting knowledge from data so that effective educational decisions can be made. This process involves selection of desired data set, pre-processing of the data, transforming data into right format, applying data mining methods and techniques on this data, and finally interpreting and evaluating results.[15] Student performance prediction can be carried out using many EDM techniques such as regression analysis and decision tree methods. There are many researches that have reported comparison of different EDM techniques to find which one is more effective in predicting student grades accurately. For example, the decision tree method has been concluded to be a more accurate method of prediction as compared to regression analysis.[15]

Four data mining prediction methods used in this research are decision tree algorithm, logistic regression, Naïve Bayes function, and neural networks. The first technique is decision trees algorithm.[16] This algorithm helps in making decision in a tree-like graph pattern where decisions and their possible outcomes are displayed in IF-THEN format, for instance, IF weather = dry THEN play = YES. According to Han et al.,[16] a decision tree is a flow-chart-like tree structure in which there are different nodes; each one representing a check or test on an attribute value, and where the leaves of the tree depict different classes. In a decision tree, each node (reflecting an instance space) is divided into two or more sub-nodes based on a certain discrete criteria of the input attribute values. The second EDM technique is logistic regression which helps in predicting outcomes by explaining relationships between dependent and independent variables.[17,18] Therefore, if there is positive relationship between age and experience of 50 employees, then with the help of logistic regression, a researcher can prove that 51st employee will have less experience since he is younger in age. The third method for prediction through EDM is Naïve Bayes classifier which predicts outcomes in a simple probabilistic manner. In this method, labels are assigned to the problems and the classifier assumes that each label contributes in the outcome generation independently.[16,19] Hence, dry weather is associated with outcome "play"

Table 1. Summary of EDM techniques used in research.

	EDM technique used	Data set	Findings
Kotsiantis[7]	• Decision tree • Model tree M5 • Linear regression	354	Regression method offers more accurate predictions
Kabakchieva[8]	• Decision tree J48 • Naïve Bayes • Nearest neighbour	10,300	Decision tree J8 classifier helps in better prediction of grades

independent of the other variable temperature. The final EDM technique used in this prediction research is neural networks. It is a prediction algorithm which computes outcomes on the basis of huge number of units or neural units connected to each other like neurons in human nervous system. neural networks process information through neural clusters to identify patterns and predict outcomes based on these patterns.

EDM techniques have been used in some existing researches, and we have summarized their results in Table 1. In the research by Kotsiantis,[7] prediction has been carried out on a data set of 354 instances and showed that regression method offers more accurate prediction than decision tree and model tree M5 method. Kabakchieva[8] proved on a large data set of 10,300 instances that J48 algorithm of decision tree predicts grades more accurately as compared to the prediction methods of Naïve Bayes and k-nearest neighbour. Conducting performance prediction on a smaller data set of size 96, Golding and McNamarah[9] found that student gender and school aptitude grades do not affect academic performance in universities especially when prediction is carried out using multiple regression method. Zafra et al.[10] proved that G3P-MI method is a better way of predicting student grades compared to decision stump and logistic regression particularly when a data set is small, e.g., 118 in this case. Finally, Xing et al.[20] also verified the supremacy of decision tree algorithm in accurately predicting student grades. He used a large data set without mentioning its actual size and compared prediction results of decision tree algorithm with those of logistic regression method.

Methods

Participants

In the present study, data have been collected from the first-year electrical engineering and computer science students at the School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NSUT), and Islamabad and Abasyn University, Islamabad, Pakistan. Data about non-cognitive features were collected from a sample of 128 target participants. A few of the participants were dropped, who were lacking any of the

Table 2. Features covered in data set 1 and data set 2.

	Data set 1	Data set 2
Demographic features	1) Gender 2) Student employment status 3) Mother's education	1) Gender 2) Parent's cohabitation status 3) Guardian
Cognitive features	1) Previous result 2) Sessional marks 3) Quizzes 4) Assignments 5) Projects	1) Absenteeism 2) First sessional 3) Second sessional
Non-cognitive features	1) Time management 2) Self-concept 3) Self-appraisal 4) Leadership 5) Community support	1) Study preference 2) Study time 3) Free time 4) Independence 5) Proximity to college 6) Health 7) Go out 8) School support 9) Plan for future studies

criterion features (i.e., missing data, information not verified by examination department). Hence, we reduced the original sample size from 128 to 113 (https://github.com/azeemabbas/dropoutData). Most of the extant studies related to the present work considered more or less the same sample size for their investigation.[7,9,10] However, for a more elaborated comparison, a larger data set was retrieved from an online repository. This online data set offered 650 instances. It was a student profile data set used by Student Self-Assessment Questionnaire.[21] The online data set covered 32 features, of which 15 independent and 1 dependent feature were selected for the present study, based on general acceptance in research and relevance of these features with the culture of Pakistan. The features as independent variables that were covered in the two data sets are presented in the Table 2. Variable result was treated as a dependent variable in both the data sets.

Instrument

In order to collect data from these students, a questionnaire was designed. The questions in the designed questionnaire (i.e., data collection instrument) were collected from different questionnaires. For example, the questions related to community and social support were adopted from Social Support Questionnaire,[22] leadership questions were extracted from Leadership Skills Questionnaire,[23] Student Self-Assessment Questionnaire[21] provided questions on the aspect of realistic self-appraisal and self-concept. The designed questionnaire was focused toward

"Study Habit." Moreover, the study habit questionnaire is adopted form of Virginia Gordon's University Survey: A Guidebook and Readings for New Students, and the questionnaire was further adopted under this research in order to include the critical non-cognitive factors identified in the related work section.

The compiled survey tool comprises 42 questions covering demographic, cognitive, and non-cognitive aspects of the individual profile. There were six questions on student demographic background, five questions on student cognitive data, i.e., previous result, marks in quizzes, assignments, and projects, eight questions on time management (i.e., non-cognitive feature), two questions regarding self-concept (i.e., non-cognitive feature), five questions about self-appraisal (i.e., non-cognitive feature), nine questions related to leadership quality among students (i.e., non-cognitive feature), and seven questions regarding community support (i.e., non-cognitive feature). The answers of these questions helped in defining whether a student is doing great, okay or needs help in the particular non-cognitive feature.

After the compilation of questionnaire, it was shared among the students using Google documents which helped in online data collection. The time of sharing survey questionnaire with the students was a critical factor. The questionnaire was shared with the students at NUST SEECS couple of weeks after the second sessional exam. The reason is that this is the time when the students have known the marks for their both sessional as well as they have completed their projects. This allows them to have an overall picture of their internal marks and performance during the semester. Collecting data at this time helps in predicting final grades on the basis of all cognitive information that a student has. On the other hand, Abasyn University conducts one sessional exam therefore; the questionnaire was shared with their students after two weeks of sessional exam when they have both the marks of sessional and the evaluation of their semester project.

In order to find accuracy of prediction, the students were not informed about the predicted results, so that they are not intimidated and it could be found whether the prediction model forecasts grades effectively or not. However, once the results of the study are finalized, in the upcoming stages of the research and forecasting, these results might be used for counselling students before terminal exams, so that they can improve their performance. After the students responded to the questions, the students' grades were collected from the exam departments of respective universities based on the student IDs, provided in survey responses. In the second data set, collected from the online source, the final grades were also available. Therefore, the researcher was saved from conducting second stage of data collection, i.e., result data collection from exam branch.

Procedure

The data collected from both universities, during Spring 2016 semester, the primary (survey questionnaire) and the secondary source (online) were pre-processed and transformed into presentable formats. During pre-processing, the issue of missing values was handled. Since all other questions were mandatory to answer except name and roll number, the instances with both missing name and roll number were excluded from the data set, since their final grades could not be accessed from exam departments. After that data reduction took place where by the dimensions were reduced and data was transformed by converting five scale responses (e.g., ranges of GPA) into the result categories of Pass, Fail, and Probation/Warning.

Once the data were pre-processed and transformed, data mining methods were selected. These methods included decision tree j48 algorithm, logistic regression, Naïve Bayes and the neural networks. The main reason behind the selection of these methods was the frequent usage of these techniques in other-related studies. The pre-processed data in MS Excel sheets were converted in CVS format and were processed using each of the four mentioned techniques in Weka, which is a freely available data mining and machine learning software.[24] Weka is a Java-based collection of algorithms that assists in data pre-processing, classification, regression, and visualization.

Prediction

Processing data in Weka[24] give output showing how the classification of data took place, what are the given values of dependent feature (i.e., Results), predicted values of results, prediction accuracies, and confusion meatrices. From these outputs, the prediction accuracies were selected for four methods. For each method, there were two prediction accuracies; once for the combination of demographic and cognitive features and once for the prediction based on demographic, cognitive, and non-cognitive features. The prediction accuracies of four methods were compared to find whether prediction accuracy increased decreased or remained unchanged after the inclusion of non-cognitive features. The same experiment and processing was repeated on the data set 2 to find the prediction accuracies on the online data set. Then, the results were compared to find trends in results and to analyse which data mining tool offered increased or decreased prediction accuracies on each data set.

In Figure 1, the value of "Correctly Classified Instances" illustrates the accuracy of model. The accuracy of prediction using logistic regression function on data set 1 is 53.98% or 54% when prediction does not include non-cognitive features and accuracy is 51.32% or 51% when prediction is based on all three, i.e., demographic, cognitive, and non-cognitive features. These values are also compared in Table 3.

Results and analysis

Student performance prediction was carried out using two data sets, i.e., the locally collected data set and the data set available online. The prediction accuracies of different machine learning techniques including decision tree, logistic regression, Naïve Bayes, and neural networks were compared with and without the inclusion

```
Scheme:     weka.classifiers.functions.Logistic -R 1.0E-8 -M
-1 -num-decimal-places 4
Relation:    D C
Instances:   113
Attributes:  9
        Gender
        Status
        MotherEducation
        PreviousSemesterResult
        SessionalMarks
        Quizzes
        Assignments
        Projects
        Result
Test mode:   10-fold cross-validation
=== Summary ===

Correctly Classified Instances      61        53.9823 %
Incorrectly Classified Instances    52        46.0177 %

 === Confusion Matrix ===

  a  b  c   <-- classified as
 44  7  4 |  a = Pass
 10 10 11 |  b = Probation
  9 11  7 |  c = Withdraw
```

```
Scheme:     weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -
num-decimal-places 4
Relation:    D C NC
Instances:   113
Attributes:  14
        Gender
        Status
        MotherEducation
        PreviousSemesterResult
        SessionalMarks
        Quizzes
        Assignments
        Projects
        TimeManagement
        SelfConcept
        RealisticSelfAppraisal
        Leadership
        CommunitySupport
        Result
Test mode:   10-fold cross-validation
=== Summary ===

Correctly Classified Instances      58        51.3274 %
Incorrectly Classified Instances    55        48.6726 %

=== Confusion Matrix ===

  a  b  c   <-- classified as
 37 11  7 |  a = Pass
 10 11 10 |  b = Probation
  9  8 10 |  c = Withdraw
```

Figure 1. Comparison of result outputs after processing data set 1 using logistic regression function left, using demographic and cognitive features right, using demographic, cognitive, and non-cognitive features.

of non-cognitive features. Despite the difference in values of prediction accuracy, the same trend was observed in each technique, i.e., decision tree showed increase in prediction accuracy after the inclusion of non-cognitive features, Naïve Bayes showed no change and the other two methods showed reduced accuracies.

Data set 1

The data set 1 where there were 113 instances in all, the decision tree algorithm offered 61% prediction accuracy when only demographic and cognitive features were used for finding student results. The prediction accuracy increased to 65% when the non-cognitive features were also used along with the demographic and cognitive features. Result prediction using logistic regression method offered 54% prediction accuracy without the use of non-cognitive features and 51% prediction accuracy when these features were added in prediction process. In Naïve Bayes method, 61% prediction accuracy was observed both before and after the addition of non-cognitive features. The neural networks gave 54% accuracy without non-cognitive features and 52% accuracy with them. The prediction accuracies of different machine learning techniques including decision tree, logistic regression, Naïve Bayes, and neural networks were compared with and without the inclusion of non-cognitive features.

Table 3. Comparison of prediction accuracy results using four methods on two data sets.

		Decision tree	Logistic regression	Naïve Bayes	Neural networks
Data set 1	Without non-cognitive features	61% (69 instances correctly classified)	54% (61 instances correctly classified)	61% (69 instances correctly classified)	54% (61 instances correctly classified)
	With non-cognitive features	65% (73 instances correctly classified)	51% (58 instances correctly classified)	61% (69 instances correctly classified)	52% (59 instances correctly classified)
Data set 2	Without non-cognitive features	82% (535 instances correctly classified)	84% (542 instances correctly classified)	84% (546 instances correctly classified)	82% (535 instances correctly classified)
	With non-cognitive features	84% (541 instances correctly classified)	82% (530 instances correctly classified)	84% (547 instances correctly classified)	76% (490 instances correctly classified)

Table 4. Predicting student results using combinations of non-cognitive features.

Factors/methods	Decision tree (%)	Logistic regression (%)	Naïve Bayes (%)	Neural networks (%)
Demo + Cog	61	54	61	54
Demo + Cog + (TM)	60	51	59	53
Demo + Cog (SC)	**63**	51	60	52
Demo + Cog+ (RSA)	**62**	**55**	**63**	**57**
Demo + Cog + (Leadership)	**62**	52	59	**57**
Demo + Cog + (Com Sup)	58	**58**	**64**	53
Demo + Cog + (TM + SC)	63	50	59	50
Demo + Cog + (TM + RSA)	62	52	**62**	50
Demo + Cog + (Leadership + RSA)	63	52	**65**	**55**
Demo + Cog + (TM + Com Sup)	61	**57**	**62**	53
Demo + Cog + (TM + Leadership)	60	50	59	**60**
Demo + Cog + (L + Com Sup)	59	54	60	56
Demo + Cog + (Com Sup + SC)	63	53	59	51
Demo + Cog + (Com Sup + RSA)	60	**57**	**65**	54
Demo + Cog + (TM+ Com Sup + SC)	60	**57**	**65**	54

TM: time management; SC: self-concept; RSA: realistic self-appraisal; Com Sup: community support.

Data set 2

By processing the data set 2 using decision tree method in Weka, the demographic and cognitive features predicted results with 82% accuracy, and when the non-cognitive features were added in the processing too, the accuracy increased to 84%. Using logistic regression method, 84% accuracy was achieved with the help of demographic and cognitive features and 82% accuracy when third category of features, i.e., non-cognitive factors were added. Naïve Bayes offered 84% prediction accuracy with and without non-cognitive factors while neural networks produced 82% accurate results without non-cognitive features and 76% accuracy with them.

Prediction accuracies using combination of non-cognitive features

Although the use of five non-cognitive features altogether with the other demographic and cognitive features helped in increasing prediction accuracy using decision tree method, when selected one or two, not all five, features were used prediction of results, other three methods logistic regression, Naïve Bayes, and neural networks also showed some increase in the prediction accuracy. The following Table 4 shows how adding different non-cognitive features of data set 1 altered prediction accuracies. The bold values in the Table 4 show the increase in accuracy experienced after inclusion of certain non-cognitive features.

Besides showing increased prediction accuracy with overall five non-cognitive features, the decision tree method also increases accuracy when demographic and cognitive features are processed along with self-concept (SC), realistic self-appraisal (RSA) or leadership individually. The accuracy of logistic regression-based prediction increases when the method includes specific non-cognitive features like realistic self-appraisal, community support (Com Sup), combination of Time management (TM) and Community support, community support and RSA, and Time management, community support and self-concept.

The prediction accuracy of Naïve Bayes method increased when the demographic and cognitive features were processed for the prediction along with RSA and community support individually, or with combinations of TM and RSA, leadership and RSA, TM and community support, community support and RSA and TM, community support, and self-concept. The prediction accuracy of neural networks increased when the demographic and cognitive features were supported by RSA or leadership individually or with the combination of leadership and RSA, and TM and leadership.

Conclusion

The primary objective of this research was to use EDM tools in such a way that students can be informed about their potential poor performance so that they can improve their performance and reduce dropout from the university. The data analysis and results proved that certain non-cognitive features help in improving the prediction accuracy of results while others do not. For example, time management feature was not found to improve prediction accuracy prominently showing that at university level, the number of hours spent in studying is less important and quality of time matters more. The second non-cognitive feature self-concept was not found to be a good performance predictor so the academicians might appreciate to focus more on the use of realistic self-appraisal which increases prediction accuracy individually as well as in combination with other non-cognitive features. Universities and colleges imparting engineering degrees will also notice in this study that leadership is a more relevant performance predictor for students of social sciences. In this research conducted on engineering students, leadership was a poor predictor of performance. Finally, the universities should try to improve student community support by offering counselling on regular basis because results show that community support helps in increasing performance more often.

It is observed that when a few non-cognitive features are coupled with cognitive features, they can result in better prediction accuracy. Leadership, for example, either does not help in increasing prediction accuracy or does so by a very little percentage. Community support helps in increasing accuracy more often, both individually as well as in combination with other non-cognitive features. Therefore, the non-cognitive features increase the prediction accuracy but the researcher's critical job is to find which non-cognitive features to include in prediction and which to ignore. Leadership might be a more relevant performance predictor for students of the social sciences, but the students of electrical engineering might not benefit a lot from it.

Acknowledgements

A special mention is given here to Dr Asad Anwar Butt, Ms Farzan Ahmad, and Mr Fahad Ahmed Satti for their guidance in this research work. Thanks to Ms Komal Malik who tuned English of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. European Commission. Dropout and completion in higher education in Europe Annex 3: country case studies Europe policy briefings Australia, U.S.A., http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2015/annex-3-country-case-studies_en.pdf (2015, accessed 26 December 2016).
2. Sittichai R. Why are there dropouts among university students? Experiences in a Thai University. *Int J Educ Dev* 2012; 32: 283–289.
3. Sheard M. Hardiness commitment, gender, and age differentiate university academic performance. *Br J Educ Psychol* 2009; 79: 189–204.
4. Aulck L, Velagapudi N, Blumenstock J, et al. Predicting student dropout in higher education. In: *2016 ICML Workshop on #Data4Good: Machine Learning in Social Good Applications*, New York, NY, June 2016, pp. 16–20, ArXiv eprints https://arxiv.org/abs/1607.02450.
5. Cortez P and Silva AMG. *Using data mining to predict secondary school student performance*. London: EUROSIS, 2008.
6. Kyllonen P. *The case for noncognitive assessments*. R&D Connections, 2005, https://www.ets.org/Media/Research/pdf/RD_Connections3.pdf (accessed 26 December 2016).
7. Kotsiantis SB. Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. *Artif Intell Rev* 2012; 37: 331–344.
8. Kabakchieva D. Predicting student performance by using data mining methods for classification. *Cybern Inf Technol* 2013; 13: 61–72.
9. Golding P and McNamarah S. Predicting academic performance in the School of Computing & Information Technology (SCIT). In: *Proceedings 35th IEEE Annual Conference Frontiers in Education, FIE '05*, Indianapolis, Indiana, USA, October 2005, p. S2H.
10. Zafra A, Romero C and Ventura S. Predicting academic achievement using multiple instance genetic programming. In: *2009 Ninth international conference on intelligent systems design and applications*, Pisa, Italy, September 2009, pp. 1120–1125.

11. Dekker G, Pechenizkiy M and Vleeshouwers J. Predicting students drop out: a case study. In: *Proceedings of the 2nd International Conference on Educational Data Mining*, Cordoba, Spain, July 2009, pp. 41–50.
12. Pesu L, Viljaranta J and Aunola K. The role of parents' and teachers' beliefs in children's self-concept development. *J Appl Dev Psychol* 2016; 44: 63–71.
13. Janssen S and O'Brien M. Disentangling the effects of student attitudes and behaviors on academic performance. *Int J Scholarsh Teach Learn* 2014; 8: 7.
14. Flynt CJ. *Predicting academic achievement from classroom behaviors*. Doctoral Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, August 2008.
15. Sachin RB and Vijay MS. A survey and future vision of data mining in educational field. In: *2012 Second International Conference on Advanced Computing & Communication Technologies*, Rohtak, India, January 2012, pp. 96–100.
16. Han J, Pei J and Kamber M. *Data mining: concepts and techniques*. Amsterdam: Elsevier, 2011.
17. Tan P-N, Steinbach M and Kumar V. *Introduction to data mining*. Harlow: Pearson Education India, 2006.
18. Witten IH and Frank E. *Data mining: practical machine learning tools and techniques*. San Francisco, CA: Morgan Kaufmann, 2005.
19. Hand DJ, Mannila H and Smyth P. *Principles of data mining*. Cambridge, MA: MIT Press, 2001.
20. Xing W, Guo R, Petakovic E, et al. Participation-based student final performance prediction model through interpretable Genetic Programming: integrating learning analytics, educational data mining and theory. *Comput Human Behav* 2015; 47: 168–181.
21. Student Self-Assessment Questionnaire. Ryerson University, http://www.ryerson.ca/content/dam/arts/sec/documents/Arts_questionnaire.pdf (accessed 26 December 2016).
22. Sarason IG, Levine HM, Basham RB, et al. Assessing social support: the social support questionnaire. *J Pers Soc Psychol* 1983; 44: 127.
23. Leadership Skills Questionnaire, http://www.csueu.org/DesktopModules/Bring2mind/DMX/Download.aspx?EntryId=189&Command=Core_Download&language=en-US&PortalId=0&TabId=260 (accessed 26 December 2016).
24. Hall M, Frank E, Holmes G, et al. The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl* 2009; 11: 10–18.