Predicting Academic Performance: A Systematic Literature Review

Arto Hellas*
University of Helsinki
Helsinki, Finland
arto.hellas@cs.helsinki.fi

Petri Ihantola*
University of Helsinki
Helsinki, Finland
petri.ihantola@helsinki.fi

Andrew Petersen*
University of Toronto Mississauga
Mississauga, Canada
petersen@cs.toronto.edu

Vangel V. Ajanovski
Saints Cyril and Methodius University
Skopje, Macedonia
vangel.ajanovski@finki.ukim.mk

Mirela Gutica
British Columbia Institute of Technology
Burnaby, Canada
mirela_gutica@bcit.ca

Timo Hynninen
South-Eastern Finland University of Applied Sciences
Mikkeli, Finland
timo.hynninen@xamk.fi

Antti Knutas
LUT University
Lappeenranta, Finland
antti.knutas@lut.fi

Juho Leinonen
University of Helsinki
Helsinki, Finland
juho.leinonen@helsinki.fi

Chris Messom
Monash University
Melbourne, Australia
christopher.messom@monash.edu

Soohyun Nam Liao
University of California San Diego
San Diego, USA
snam@eng.ucsd.edu

ABSTRACT

The ability to predict student performance in a course or program creates opportunities to improve educational outcomes. With effective performance prediction approaches, instructors can allocate resources and instruction more accurately. Research in this area seeks to identify features that can be used to make predictions, to identify algorithms that can improve predictions, and to quantify aspects of student performance. Moreover, research in predicting student performance seeks to determine interrelated features and to identify the underlying reasons why certain features work better than others. This working group report presents a systematic literature review of work in the area of predicting student performance. Our analysis shows a clearly increasing amount of research in this area, as well as an increasing variety of techniques used. At the same time, the review uncovered a number of issues with research quality that drives a need for the community to provide more detailed reporting of methods and results and to increase efforts to validate and replicate work.

*Co-leaders

ITiCSE '18 Companion, July 2–4, 2018, Larnaca, Cyprus

ACM ISBN 978-1-4503-6223-8/18/07.
https://doi.org/10.1145/3293881.3295783

CCS CONCEPTS

• **Social and professional topics** → *Computer science education.*

KEYWORDS

educational data mining, analytics, learning analytics, prediction, performance, literature review, mapping study

ACM Reference Format:
Arto Hellas, Petri Ihantola, Andrew Petersen, Vangel V. Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, Chris Messom, and Soohyun Nam Liao. 2018. Predicting Academic Performance: A Systematic Literature Review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '18 Companion), July 2–4, 2018, Larnaca, Cyprus.* ACM, New York, NY, USA, 25 pages. https://doi.org/10.1145/3293881.3295783

1 INTRODUCTION

In the psychology and education-related scholarly literature, work on determining factors that contribute to academic performance has existed at least for a century. For example, in the late 1910s, a wide range of tests such as verbal memory tests were conducted with freshmen in an attempt to tease out factors that correlate with academic performance [45, 246, 397]. While the early work was conducted within psychology, interest in identifying individuals with particular performance characteristics such as the ability to program emerged soon thereafter [236].

While the earlier studies on identifying programming ability were mostly focused on attempting to find individuals who could perform as programmers [236], subsequent work spanned to trying to identify factors that predict students' computer aptitude [112] and performance in programming courses [340]. The reasons for predicting aptitude are numerous of which Evans and Simkin outline several [112]: identifying potential majors, discriminating among applicants, advising students, identifying productive individuals, identifying those who may best profit from additional guidance, improving classes, determining importance of existing predictors, and exploring the relationship between programming abilities and other cognitive reasoning processes.

Parallel to this, there has been a stream of research that studies why students struggle with learning to program. A rather famous example of this work is the research on the rainfall problem, where students are expected to write a program that reads in numbers and then calculates their average [347]. While the problem has been studied rather extensively [333], one of the interesting approaches taken to study the problem was the instrumentation of working environments. For example, when studying students working on the rainfall problem, Davies employed a system that recorded key presses for further analysis [88].

Since then, systems that record students' working process have been further adopted into computing education research [168]. This adoption of instrumented tools, combined with modern hardware, has created an opportunity to directly observe and react to student data, which has invigorated the research in models that can be used to predict academic performance.

This report is the outcome of an ITiCSE working group that is seeking to connect various communities – including those outside of computing education – that are supporting the work of predicting academic performance in computing courses. The working group is composed of internationally-diverse members with a variety of academic interests, and it worked for a period of three months, including an intensive five-day meeting at ITiCSE in July 2018.

We outline the results of a systematic literature review containing 357 articles that describes the breadth of work being done on the prediction of student performance in computing courses. In addition to the review itself, which summarizes the types of performance being predicted and the factors and methods used to perform the predictions, we identify trends in feature and method use over time, and offer insights obtained as we read. We believe this work is a mapping that will help to connect researchers in this area by identifying clusters of related work being published in different venues and highlighting opportunities for collaboration, integration, and broader dissemination.

1.1 Research Questions and Scope

The group initially sought to find literature related to identifying students who are "academically at-risk" [148], but initial forays into the literature highlighted that the term *at-risk* is often used to identify youth in disadvantaged circumstances. Several other terms, including *performance*, were used to focus the work on students at academic risk. As a result, the *prediction of student performance* became the focus of the group. We explored the following research questions:

(1) What is the current state of the art in predicting student performance?
   (a) How is performance defined? What types of metrics are used for describing student performance?
   (b) What are the features used for predicting performance?
   (c) What methods are used for predicting performance?
   (d) Which feature and method combinations are used to predict which types of student performance?
(2) What is the quality of the work on predicting student performance?

Both the metrics describing students performance and features used to predict them can be considered as variables. In the following, we will call input variables as features and output variables (i.e., performance metrics) as predicted (or output) variables. The term method refers to how output variables have been derived (possibly from the inputs).

The search terms used to identify the articles for this study are described in Section 3, but we include a discussion of the term *performance* here to define the scope of this work. For the purposes of this literature review, we defined the term broadly, including performance on a single assessment, performance in a course, retention in or continued progress through a program, and successful matriculation from a program. However, we only considered prediction of quantifiable metrics that are directly related to the course or the program such as grades, pass/fail probability, or retention in a program. We do not include articles that predict proxies, such as depression or team cohesion, that are not quantifiable or directly related to academic performance, even if they are likely to affect it. Similarly, we exclude articles that deal with predictions that may be a product of academic performance, such as employability, or articles that do not directly predict performance, such as those that primarily evaluate pedagogical interventions.

As a result of this broad focus on performance, our review covers a wide range of factors and contexts. As shown in Figure 1, students face challenges at many points in their academic career, and the factors contributing to those challenges vary widely, from family and economic stress to issues with academic preparation and a lack of study skills. We did not exclude any factors or methods that were identified in the articles that were included, though we attempted to cluster them into higher-level tags, such as "wealth" or "social factors" due to the diversity of individual factors examined.

1.2 Report Outline

This report is organized as follows. In the subsequent section, Section 2, we provide an overview of the existing review and survey articles on predicting student performance. The section discusses the exclusion and inclusion criterion, time spans, and suggested terminology and taxonomies in the existing literature. We draw upon this work when formulating our search and inclusion criteria, which are described in Section 3. Sections 4 and 5 present the results of the review, with Section 4 focusing on descriptive statistics and the results of a textual analysis and Section 5 presenting issues that emerged as reviewers completed the process. We provide several calls to the community and highlight possible directions for future work in Section 6. The list the articles reviewed is provided in Table 13, and we provide the form used to extract data from each reviewed article in an appendix. The form is provided so that readers can review our method, identify sources of bias, and if they wish, modify or extend our form for their use in similar projects.

2 SYNTHESIZING PREVIOUS REVIEWS ON PREDICTING STUDENT PERFORMANCE

This section provides a synthesis of previous reviews and surveys on predicting student performance. Our goal was to identify categories and trends already identified in existing reviews, so that our review uses common terminology, where possible, while also contributing a new understanding of the literature in the field. The synthesis described in this section informed the instrument we used in our study. A description of the data extraction instrument is provided in Section 3.1.3, and details are provided in Appendix A.

The corpus for the synthesis was created through metadata searches on the article indexes Scopus, IEEE Xplore, and the ACM Digital Library. For the search, the search string "`student AND predict* AND performance AND (review OR survey)`" – or the equivalent variant expected by each article engine – was used. The search was conducted in June 2018, and resulted in a total of 147 hits, including duplicates that were indexed in multiple sites. Manual inspection of article titles and abstracts and exclusion of one-page extended abstracts reduced the count to 13 reviews and surveys of the area. Because of the focus on *performance prediction*, broad reviews of educational data mining and learning analytics, such as [39, 102, 168, 269, 318], were omitted.

The quality of the reviews varied. In particular, the bounds of the search performed in the review or survey were not always well defined. Of the 13 articles, eight [160, 185, 198, 204, 252, 253, 269, 335] listed at least some of

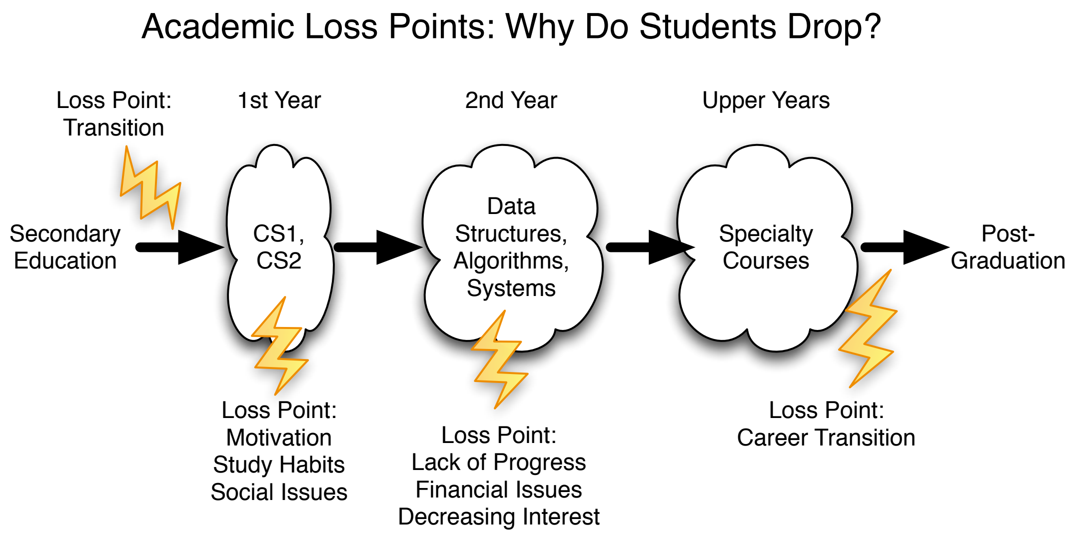

Figure 1: Points in a student's academic life, using CS as an example, where data for performance prediction can be gathered and performance prediction can be done. Both the first and second year loss factors can continue, to a lesser extent, through to graduation.

the sources, venues or fields that they had searched. Only six [198, 204, 252, 253, 269, 335] listed the keywords or search terms used. Eight [99, 160, 185, 198, 252, 253, 269, 335] described a year range or the absence thereof. All surveyed a limited number or range of articles.

In this review of surveys, we collected and synthesized what these existing reviews and surveys reported about factors for predicting student performance, methods used for predicting student performance, and how the articles described performance. Of the reviewed articles, five [160, 198, 204, 253, 335] summarized factors, two [252, 335] summarized methodologies, and four [160, 198, 204, 252] offered insights into the meaning of performance or what was being predicted. None of the surveys discussed all three of these issues, but five [160, 198, 204, 252, 335] discuss two of them.

2.1 Factors

Table 1 contains a listing of the high-level factors used for predicting student performance. The category sizes range widely, from 3 to 10, indicating that the review authors saw value in various levels of abstraction when categorizing factors. We drew inspiration from this list when constructing the form we used to extract data. Our aim was to create a broad categorization that enables identification of the source of the data being used, so we synthesized the categories used in the reviews into Family Background, Demographic Data, Working Conditions, Educational Background, Course Data (current or parallel), Student Motivation, and Psychological / Affective / Learning Scales. We split motivation and other scales into two categories because motivation is a complex construct that has been extensively used and studied for predicting performance, so we expected to see a variety of types of motivation (and corresponding instruments) used in the articles we reviewed.

2.2 Methods

Muthukrishnan et al. [252] categorized methods used for predicting performance into four high level categories: Decision trees, Regression, Clustering, Dimensionality reduction / other. Overall, other articles – if they included methods for predicting performance – did not explicitly attempt to provide a categorization, but provided lists of methods that have been used for prediction. In our form, we used the high-level categories Classification (supervised learning), Clustering (unsupervised learning), Mining (finding frequent patterns and/or feature extraction), and Statistical (correlation, regression, t-testing, etc), and we encouraged reviewers to provide details on the specific techniques used underneath these high-level categories.

A few of the articles compared the performance of methods through metrics such as accuracy that were extracted from the surveyed literature. While this approach has appeal, and we did extract such performance metrics from the articles we reviewed, we found that the contexts of the studies varied so significantly and the quality of reporting varied so widely that a meta-review would not be meaningful. A comparison of reported accuracy ratings would, in our opinion, lead the reader to misconceptions about the performance of various methods.

Table 1: High-level categorization of factors that can be used for predicting performance from the articles included in the synthesis.

Study	Factors
[160]	Activity and course features, Demographic features, Learning behavior features, Self-reported features, Student history record and performance, Student record and performance in current course, Others / unclear features
[198]	Academic, Family, Institutional, Personal, Social
[204]	Academic performance, Socio-economic, Personal information
[253]	Learning behaviour data, Learning network data, Learning level data, Learning emotional data, Other
[335]	Cumulative Grade Point Average, Engage time, External assessments, Extra-curricular activities, Family support, High-school background, Internal Assessment, Social interaction network, Study behavior, Student demographic, Student interest

2.3 Definitions of Performance

Most of the articles included in the synthesis did not explicitly define performance. This may be a consequence of the articles being reviews of other articles, effectively creating a situation where other articles are included as long as they predict performance – no matter if they defined what it means. The closest to a definition of performance was provided in [160], where the authors suggest that, "studies tended to predict course performance (successful/unsuccessful), course grades and student retention/dropout in online/blended learning contexts."

We found the lack of definition problematic, as it led some surveys to include articles with significantly different goals. Therefore, we reviewed the articles for trends to see what was most included. Many agreed that assessments, defined broadly, are the key metric. For example, one review notes that:

> ... students performance can be obtained by measuring the learning assessment and co-curriculum. However, most of the studies mentioned about graduation being the measure of students success. Generally, most of higher learning institutions in Malaysia used the final grades to evaluate students performance. Final grades are based on course structure, assessment mark, final exam score and also extracurricular activities. [335]

Another article, from 2017, agrees that student performance can be observed using internal assessment metrics:

> Most of the Indian institution and universities using *[sic]* final examination grade of the student as the student academic performance criteria. The final grades of any student depend on different attributes like internal assessment, external assessment, laboratory file work and viva-voce, sessional test. The performance of the student depends upon how many grades a student score in the final examination. [198]

However, some articles define performance more broadly. For example, [204] defines performance as, "... a measure of a student's competence for future courses." They note that:

> In addition to passing or failing a course, the grade obtained is also of interest to course instructors. Furthermore, academic advisors would be interested in the time required for a student to complete a degree and his/her ability of enrolling in multiple programs. [204]

We agree with the inclusion of these wider factors. As a result, our data extraction form initially included a range of academic performance factors, including various forms of internal course assessment, but also including program retention and likelihood of graduation as possible values to be predicted.

Another survey, which provided an overview of the values being predicted in the articles they reviewed, agrees that the literature defines performance broadly. They noted that the majority of articles they saw predicted final grades in courses, but they saw other predictions being made when the context is outside of the traditional classroom [252]. This led us to include several questions about the population being studied in our data extraction form, including the topic of the course in which the prediction is being employed, the format of the course (e.g., traditional, online, hybrid, MOOC, etc.), and the type of students in the study (e.g., K-12, non-majors, graduate students, etc.). All of these questions were to be left blank if they were not germane to the article being reviewed.

2.4 Summary

Overall, we saw a broad range of surveys in terms of quality, area and amount of the literature covered, and focus. We see an opportunity to provide a higher-level view of the methods being used and to survey the literature over a longer period of time than the existing reviews. We also believe there is an unfilled need for analysis that relates the methods and features used to the context in which they were applied.

3 SYSTEMATIC LITERATURE REVIEW

In this section, we will first describe the how the literature review was conducted (Section 3.1) and then relate statistics to describe the included articles (Section 3.2). We consider our wok to be a Systematic Literature Review (SLR). However, the methodological distinction between systematic mapping studies and SLR studies is often blurred [190, 283]. Our work can be argued to be somewhere between these two approaches. In particular, we aim to create a high level synthesis but we will not describe each identified work separately as some SLR studies do.

3.1 Methodology

3.1.1 Identification of Relevant Literature. We began by collecting articles about predicting student performance that were already known to the experts in the working group. Based on this set of known articles, we tested multiple search terms by looking at the following three indexes: (1) Scopus, (2) IEEE, and (3) ACM. We started with the search terms used in the previous surveys (see Section 2). After multiple iterations, we decided to use the following search string:

> (at-risk OR retention OR persistence OR attrition OR performance) AND
> (prediction OR modelling OR modeling OR detection OR predict OR "machine learning") AND
> ("computer science" OR informatics OR engineering OR programming OR cs)

The searches were conducted in June 2018. The syntax of the search strings was adjusted for each index. After combining the results, filtering out articles not written in English, and using an automated process to remove duplicates based on article title, publication year, and DOI, a corpus of ~4,200 articles was created. The working group manually reviewed the article titles and abstracts to flag potentially relevant articles (see Section 3.1.2 for details). A total of 743 articles were flagged. We removed all articles published prior to 2010 from this list, resulting in a set of 565 articles for close analysis. (2010 was selected as a starting point because we detected a marked increase in articles published at this point; it also provides a focus on recent work.) At this point, only the articles clearly out of our scope or published prior to our cutoff had been removed.

Review of the included articles identified that several known to the group were not found in the set to be reviewed. In most cases, it was determined that the articles were not included because they did not focus on *prediction of performance* – as defined in the previous section. Instead, they focused on modeling learners' skills or on the discussion of factors that might be related to performance, but without clear predictive goal. After consultation, we decided to omit those articles. In other cases, we found that the article's publication venue was not indexed in any of our sources. In particular, articles from the International Conference of Educational Data Mining were not included. We manually added the relevant articles from this source to our list, leading to a final total of 589 articles for analysis.

3.1.2 Inclusion Criteria. The inclusion criteria was *the article must discuss predicting student academic performance*. This necessitated a working definition of *academic performance*, which we developed in the previous section. In particular, as we flagged articles for review, we looked for a value to be predicted that was related to a single assessment, a course, or progress through a program. We only considered quantifiable metrics that are directly related to a course or program that students are enrolled in, such as course activities, exercise points, etc. We did not include articles that predict proxies not directly related to academic performance of an *individual*. More explicitly, work was not included if its focus was:

- Predicting team performance and dynamics
- Work placement
- Affective states (e.g., happiness, anxiety, depression)
- Intent (e.g., intent to enter a university)

- Automatic assessment tools (e.g., automatically assessing programming problems, automatically assessing essays, detecting plagiarism, teaching interventions, and recommender algorithms), if the article did not clearly and separately discuss performance prediction in an included context

In rare cases, authors stated that a journal publication was constructed from a conference publication and included all the relevant information from the previous version of the article. In these cases, only the journal version was included. Similarly, theses and dissertations were removed, as related journal articles were included. This resulted in 497 articles remaining for a more detailed review.

3.1.3 Data Extraction. Based on the research questions and the meta-survey presented in Section 2, a preliminary version of the data extraction form was constructed. The working group members were divided into pairs, and each pair evaluated five articles using the form. Afterwards, the form was adjusted based on their responses to allow reviewers to provide more precise details about factors and methods. The high-level taxonomy of the resulting instrument is as follows:

- Prediction
  - What is being predicted
  - Details on what is being predicted
- Context
  - Number of subjects
  - Population
  - Course topic
  - Training mode
  - Education type
- Data features used to predict performance
  - Data set information
  - Family background (parental information, socioeconomic status, etc.)
  - Demographic data
  - Working Conditions
  - Education (background)
  - Course data (current or parallel)
  - Student motivation
  - Psychological / affective / learning scales
  - Used surveys / standardized questionnaires
- Methods and techniques
  - Method type
  - Classification (supervised learning)
  - Clustering (unsupervised learning)
  - Mining (finding frequent patterns / feature extraction)
  - Statistical techniques
  - Results details
- Quality factors of the article
  - Is there a clearly defined research question?
  - Is the research process clearly described?
  - Are the results presented with sufficient detail?
  - Does the article discuss threats to validity?
  - Are there separate training and prediction data sets (where relevant)?
  - Has the work been verified in a second population?
  - Are the data collection instruments linked or included?
  - Were all the features described in sufficient detail to identify them?
  - Additional notes on quality

To avoid ambiguity in the data extraction, all categories were implemented as check-boxes with an additional open-text field at the end for listing features not in the predefined list of options. Details of the form are provided in Appendix A, where all of the options provided are listed. We provide these details partially to allow review of our data collection instrument but also in the hope that others may find the instrument useful.

Despite our attempts to list all of the common answers, the open text boxes were frequently used. In some cases, it even had to be used to document multiple items, so we established a protocol of separating items with a semicolon. Our parsing scripts extracted all of the items identified, resulting in lists of items for each high-level item (context, data features, methods, and quality factors). Pairs of reviewers reviewed each list to partition related items. For example, all statistical tests of variance were collected into a single category, and this category was used in the reports in Section 4.

3.2 Descriptive Results

During data extraction, reviewers continued to apply the inclusion criteria and excluded articles that did not explicitly discuss prediction of academic performance. In total, data was obtained from 357 articles. Table 2 presents the number of publications per year. Data from 2018 should be omitted from an analysis of long-term trends, since the work was completed in July, leaving half a year for additional work to be published. Focusing on 2010 through 2017, then, we see an unmistakable increase in work published each year.

Table 2: Number of papers reviewed per year of publication.

Years	Count
2010	7
2011	20
2012	22
2013	36
2014	43
2015	53
2016	70
2017	75
2018	31
Total	357

Table 3 lists the disciplines in which prediction was performed. Recall that our search terms focused on computer science, engineering, and informatics, so it's unsurprising that most of the work we reviewed performed prediction in a CS or, more generally, STEM context. Mathematics consists of one third of STEM. The "Multi disciplinary" category refers to work that explicitly predicted performance in two or more contexts in different disciplines. Most of these were within CS or STEM as well. The rationale of this preliminary analysis is to illustrate the scope of this survey: the focus truly is in engineering and mostly in computing.

Table 3: The discipline in which the prediction was being performed, if a specific discipline is named.

CS	126	34.9%
STEM	98	27.1%
Other	39	10.8%
Multi disciplinary	30	8.3%
Unclear	14	3.9%

Table 4 presents the venues in which work has been published. We only list venues with three or more reviewed articles due to the large number of specialized venues which contributed one or two articles. In the venues with multiple articles, we saw computing education (SIGCSE, ITiCSE, ICER, etc.), engineering education (FIE, ICEED, EDUCON, etc.), STEM education (ISEC), and learning analytics and data mining (LAK, EDM, L@S, etc.).

About half of the articles we reviewed (171) were the only work from a particular venue, suggesting that there is broad interest in the topic. The single list of single-article venues includes conferences focusing on software engineering, machine learning, general education, and psychology. The topic is of interest to researchers and practitioners in a wide variety of fields, so we find work tuned to specific disciplines and regions in venues catering to those communities. Unfortunately, this result also suggests that the community is highly dispersed, which makes dissemination, collaboration, and validation challenging.

3.2.1 Topic modeling and discovered themes . As a part of our exploratory mapping, we sorted the documents into topics using the Latent Dirichlet Allocation (LDA) algorithm [51]. We used LDA as a statistical text mining method for assigning documents into topics, which are detected using word association and distributions [50]. The underlying mechanism in LDA is a probabilistic Bayesian network model, in which each document is characterized by certain topics, and each topic is defined by a specific set of words, which co-occur with a certain probability. To summarize, the topics of each document are defined by a set of words that often appear together, and the topics often characterize themes found in the literature.

For the analysis, we used a modified version of the *NAILS* script [191], which utilizes the *topicmodels* R package [155] and visualized with the *LDAvis* library [338]. Semantic coherence, a quality value for deciding the number of topic models [243], was calculated using the *R stm* library [313]. Additionally, the *LDAvis* library was also used to calculate the distance between topics on a scatter-plot, which approximates the semantic relationships between the topics with multidimensional scaling. It is a method similar to factor analysis and allows the level of similarity between objects to be visualized. The inter-topic distance was calculated using Jensen-Shannon divergence [338]. LDA-based topic modeling is a commonly used method for text analysis and equivalent methods have been used to statistically analyze scientific texts in previous studies [84, 180, 282, 387].

After examining the local maximums in the semantic coherence results, we proceeded with a topic model with 11 topics. The topic modeling results and the themes we discovered are summarized in Table 5. Topics whose relative size was less than two percent were excluded from analysis to avoid unjustified generalizations. The sum of the sizes adds up to just over 100% due to rounding issues.

All of the themes identified are connected due to their focus on students, prediction, and performance. In order to avoid repetition, these focuses are not explicitly mentioned in each row of the table. Furthermore, it should be noted that this initial exploratory categorization is based on probabilistic modeling and document corpora. The inter-topic distances and the prevalence of each topic are presented in Figure 2. The results are presented on two axes via multidimensional scaling, with prevalence of each topic denoted by size of the circle and distances by distances between the circles.

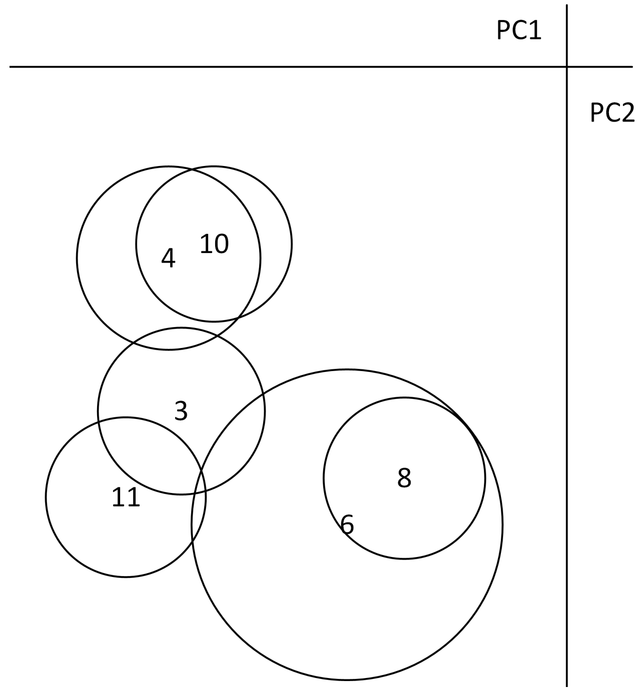

Figure 2: Distance and relative size of modeled topics

4 THEMATIC ANALYSIS

4.1 Predicted Values (Outputs)

To answer research question (1a), about the types of metrics are used for describing student performance, we categorized the types of performance that the studies aim to predict. Table 6 describes these values in the articles that were reviewed. Some articles predicted more than one category of value, so the percentages reflect the relative number of articles that showed interest in that method for describing student performance. The most popular indicator for performance is course grade, which is used as a predictor in a quarter of the studies examined, as well as an additional 13.6% that predicted a grade range. Overall, we can see that summative performance metrics, such as GPA, are the preferred prediction target. Other approaches, such as post-course outcomes or assessment of learning outcomes, are represented but are in the minority. Unfortunately, in 12.2% of the studies, it was not clear what the actual metric to predict student performance was, and therefore these metrics were classified as "Unspecified or Vague Performance."

Table 7 lists the frequency at which various indicators of performance have been predicted over the years. As earlier, data from 2018 should not be considered when evaluating long-term trends. Most factors have experienced comparable growth over the entire period we observed, but prediction of "knowledge gain" has only appeared recently.

4.2 Features Used (Inputs)

Table 8 lists the frequency at which various features have been used to predict student performance. 29 input features were identified, with performance in the course of interest, engagement in the course of interest, and performance in previous courses standing out as the most common data being used in predictions.

General demographic data has been less frequently used recently, but some aspects of demographic data, like gender and health, have become increasingly important. There also appears to have been a surge in interest in some psychometric factors, like self-regulation and self-efficacy, and in data extracted from logs, such as time on task.

We also examined how many factors were used to perform a prediction, on average, and were disappointed to find that most articles used a small number of factors. While studies may appear to use multiple factors, in many cases, these factors are measuring similar things. There are few studies that combine different data sources, such as performance in previous courses, demographic data, and psychometric data. The variance was actually larger in earlier years of the study, suggesting that there may have been more exploratory articles in the early 2010's, with current efforts focusing on the impact of a small number of factors.

4.3 Methods Used

Table 9 lists the frequency at which various statistical and machine learning methods have been used to predict student performance.

Table 4: Publication venues with at least three included papers.

Venue	Count
International Conference on Learning Analytics and Knowledge (LAK)	21
International Conference on Educational Data Mining	17
Frontiers in Education Conference (FIE)	13
ACM Technical Symposium on Computer Science Education (SIGCSE)	10
ASEE Annual Conference and Exposition	8
International Journal of Engineering Education	8
International Conference on Advanced Learning Technologies (ICALT)	7
ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE)	6
International Computing Education Research Conference (ICER)	5
International Conference on User Modeling, Adaptation, and Personalization	5
International Conference on Inventive Computation Technologies (ICICT)	4
IEEE International Conference on Engineering Education (ICEED)	4
IEEE Global Engineering Education Conference (EDUCON)	4
Integrated STEM Education Conference (ISEC)	4
Computers & Education	4
IEEE Transactions on Learning Technologies	3
Annual Conference on Information Technology Education (SIGITE)	3
IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT)	3
Journal of College Student Retention: Research, Theory & Practice	3
ACM Conference on Learning @ Scale	3

Table 5: Summary of discovered topic modeling-based themes

Topic	Theme	Relative size	Papers in the theme
Topic 1	(excluded because of small size)	0.29%	
Topic 2	(excluded because of small size)	0.58%	
Topic 3	STEM education, self-efficacy, persistence factors, motivation and gender.	9.9%	[44, 47, 79, 107, 119, 149, 153, 154, 158, 166, 173, 196, 206, 218, 219, 262, 265, 266, 279, 281, 306, 307, 332, 356, 382, 384, 386, 405, 407, 413, 415, 416, 423, 424]
Topic 4	Behavior modeling and grade prediction. Scores, exams, and assignments.	20.8%	[8–10, 21, 33–35, 54, 60, 61, 67, 75, 77, 85, 96, 105, 106, 108, 109, 115, 139, 143, 146, 150, 156, 159, 174, 182, 187, 188, 192, 193, 203, 205, 207, 209, 212–214, 223, 237, 238, 250, 257, 287, 293, 294, 302, 303, 309, 311, 312, 349–354, 360, 368, 378, 380, 381, 388–390, 393, 398, 399, 409, 420]
Topic 5	(excluded because of small size)	0.9%	
Topic 6	Data modeling, computation approaches, algorithms, and training.	13.7%	[2, 3, 13, 19, 20, 24, 26, 52, 57, 65, 66, 78, 86, 90, 97, 98, 135, 151, 164, 165, 179, 184, 197, 210, 211, 229, 248, 256, 260, 261, 264, 271, 280, 308, 314, 319, 344, 345, 365, 372, 395, 408, 414, 417, 418, 421, 422]
Topic 7	(excluded because of small size)	0.58%	
Topic 8	Prediction, classification, and educational data mining. Classification and accuracy.	21.6%	[1, 11, 14, 15, 17, 18, 23, 25, 27, 36–38, 42, 46, 48, 55, 70, 73, 87, 91, 93, 103, 116, 120, 130, 133, 136, 140, 142, 145, 147, 176, 181, 186, 189, 194, 199, 215, 221, 222, 224, 226–228, 234, 241, 242, 245, 258, 268, 278, 288, 291, 295, 298, 304, 317, 320–322, 324, 328, 336, 339, 342, 343, 358, 359, 363, 371, 379, 391, 400]
Topic 9	(excluded because of small size)	3.2%	
Topic 10	Online activity, time, and performance. Social factors and motivation.	12.3%	[5, 12, 40, 41, 43, 49, 53, 58, 59, 69, 80, 81, 101, 110, 111, 123, 126, 127, 129, 132, 134, 138, 141, 144, 175, 195, 201, 247, 251, 259, 270, 272–274, 286, 305, 310, 357, 383, 410, 412, 419]
Topic 11	Predicting grades, scores, and success. Retention and at-risk students.	16.7%	[7, 16, 22, 28–32, 63, 64, 71, 72, 76, 83, 89, 94, 95, 100, 104, 114, 118, 137, 152, 157, 161, 163, 169, 170, 172, 177, 178, 232, 240, 249, 267, 285, 296, 297, 299, 315, 325–327, 330, 341, 346, 348, 355, 361, 366, 373, 377, 385, 396, 402, 404, 411]

Statistical methods such as linear modeling (31.3%) including linear regression and ANOVA are the most common methods we observed, with various graph models and decision trees close behind. In general, classification techniques are used more frequently than clustering techniques (and our reviewers note that clustering is occasionally used as a preparatory step before applying a model). The variety of unique methods used has increased in the period we observed.

4.4 Cross-tabulation of Features (Inputs) to Performance Values (Outputs)

Finally, we explored the relationship between the value being predicted and the inputs used in the prediction. Figure 3 is an association map that illustrates the most frequently observed combinations of inputs and outputs. For example, course grades are often predicted by other grades and

Table 6: Values being predicted in the reviewed papers. Some papers predicted more than one value.

Tag	Count	%
Course Grade or Score	88	24.4%
Exam / Post-test Grade or Score	53	14.7%
Course Grade Range (e.g., A-B/C-F, Pass/Fail)	49	13.6%
Program or Module Graduation / Retention	48	13.4%
Unspecified or Vague Performance	44	12.2%
GPA or GPA Range (including CGPA, SGPA)	44	12.2%
Assignment Performance (e.g., grade, time to completion)	41	11.4%
Course Retention / Dropout	20	5.5%
Knowledge Gain	8	2.2%
Number of Courses Passed or Failed	4	1.1%

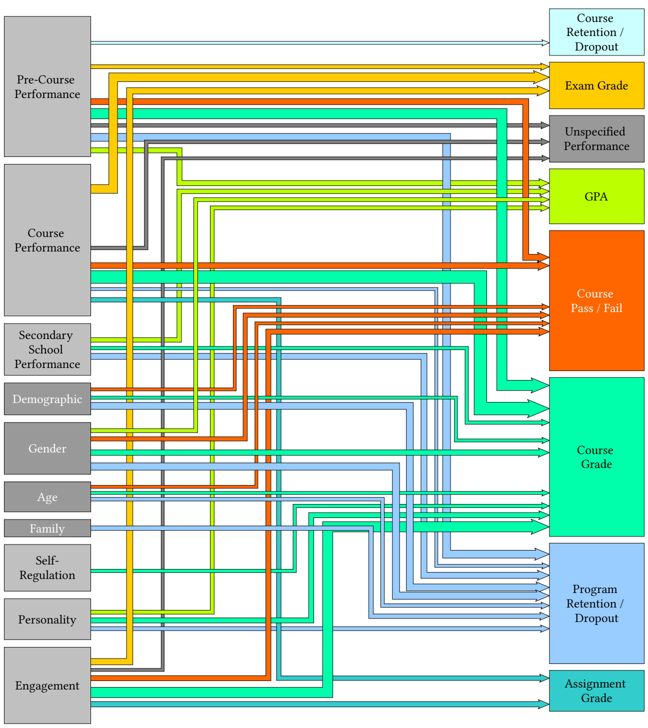

Figure 3: Most frequently researched features as predictors (left side) for predicted values (right side). The graph only includes links that were explored by at least 10 articles. The thickness of the links denotes the number of articles that explored such a predictor.

engagement in the course. Program retention uses multiple demographic factors, as well as performance in earlier academic settings.

Table 10 provides more detail on these relationships by relating particular input features with predicted values. Performance in prior courses or in secondary education is one of the most widely used inputs. Gender is also widely used as an input, though as noted in Section 5.7, gender data is generally not collected in an inclusive manner.

5 DISCUSSION

In the previous section, we presented a description of the data we obtained from reviewing the articles identified by our review. In this section, we offer insights into the data that may not be explicitly reflected in the data but which we experienced as we completed the review.

5.1 Publication Trends

Tables 8 and 9 contain the number of times that factors and methods, respectively, were used in articles in various years. While the unique number of methods and factors grew in the first years we observed, it has not obviously increased over the past four years. However, we did see many recent articles experimenting with variants of previously used methods. We also saw increasing numbers of articles using machine learning techniques, such as support vector machines (SVMs) and probabilistic graph models (such as Hidden Markov Models). Similarly, while the number of unique high-level factor categories has not increased in recent years, we saw a shift in the particular factors being used, with submissions, log data, and similar artifacts being much more heavily utilized in recent years.

As we noted earlier, we did not see as many demographic factors being used to predict course grades, and that may be an area for future investigation. We also saw a general absence of work that used motivation, which we had expected to see. Instead, many articles seemed to be using engagement, perhaps as a proxy. More generally, psychometric data is less frequently utilized, and the increase in usage of self-efficacy and self-regulation data might be signalling growth in that area.

5.2 Contexts

Most of the work we reviewed was performed in a post-secondary context, and the figures in the previous section should generally be interpreted as applying to that educational level. However, we did observe some work in the K-12 and secondary environments. GPA, single course, and program retention predictions are not (or are less) relevant in these environments, and much of the work we observed was predicting interest (which we excluded from this study) or performance in modules or individual exercises.

We also briefly considered the relationship between the discipline being studied and the value being predicted. This analysis is presented in Table 11. It appears that the disciplines are being investigated in a fairly uniform manner. However, as discussed earlier, most of the work identified here is in the context of engineering. Findings in other disciplines might be different.

5.3 Quality

During our review, we evaluated articles on several aspects of quality. Table 12 displays the results of this effort. Several of the results are disheartening. For example, in almost one out of ten articles, we had trouble identifying what was being predicted (e.g., "student performance" or "student success" without qualification).

5.3.1 Reporting Results. Other attributes are also directly related to our ability to interpret the results. Only a third (33%) of the articles we examined included a directly stated research question. Another 40% stated their goals without presenting them as research questions. In some cases, we also had trouble identifying how the prediction was being made and whether the data was reliable. In several articles, it was difficult to determine which data was being used to perform the prediction. For example, data presented in the methodology section might have been used for prediction or simply presented as demographic data to help describe the population. More seriously, in about a third of the articles we reviewed, the features being used in prediction were not described in sufficient detail for us to identify them with confidence. For example, some articles indicated that "interest," "personality traits," "motivation," or "stress" were features, but these terms can reflect different quantities. Standard scales for measuring some of these features

Table 7: The aspects of student performance predicted over the years

Performance metric being predicted	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	Percentage
Course Grade or Score	2	3	3	4	14	12	15	24	7	84	22.28%
Exam / Post–test Grade or Score	1	3	2	4	6	10	8	10	7	51	13.53%
Course Grade Range (e.g. A–B/C–F, Pass/Fail)	1	3	5	7	6	6	10	7	2	47	12.47%
Program or Module Graduation / Retention / Dropout	1	3	3	5	3	4	6	12	7	44	11.67%
GPA or GPA Range (incl. CGPA, SGPA)	2	2	2	7	8	6	7	5	2	41	10.88%
Unspecified or Vague Performance		1	4	4	8	5	13	2	3	40	10.61%
Assignment Performance (e.g. grade, time to completion)		3	1	3	3	8	11	9	2	40	10.61%
Early Detection of Course Retention / Dropout			1	2	1	4	3	5	2	18	4.77%
Knowledge Gain						2	2	2	2	8	2.12%
Number of Courses Passed or Failed		1		1			1		1	4	1.06%

Year

Table 8: Use of factors to predict student performance by year

Features used	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	Percentage
Course Performance	4	3	6	10	23	23	26	34	12	141	13.09%
Pre–Course Performance	2	8	8	13	16	23	21	33	15	139	12.91%
Engagement	3	3	7	9	11	17	31	27	5	113	10.49%
Gender	1	5	2	7	19	14	12	16	10	86	7.99%
Personality	1	2	3	9	7	9	14	14	6	65	6.04%
Demographic	1	7	3	4	14	13	6	8	9	65	6.04%
Secondary School Performance	1	6	2	5	11	12	5	11	5	58	5.39%
Age	1	4	2	5	5	8	10	11	7	53	4.92%
Family		6	2	3	9	8	8	9	7	52	4.83%
Task Time		1	2	4	3	6	11	9	5	41	3.81%
Motivation			4	5	3	5	8	5	3	33	3.06%
Self–Regulation				1	1	5	6	9	6	28	2.60%
Log Data		5	2	3	1	2	7	6	2	28	2.60%
Location	1	1		1	6	7		5	4	25	2.32%
Discipline		2		3	3	2	2	6	2	20	1.86%
Wealth	1	3	2	1	1	2	2	4	3	19	1.76%
Social		1	3	4	1	1	5	4		19	1.76%
Organization				1	1	4	7	4	2	19	1.76%
High School Quality		2		1	5	5		5	1	19	1.76%
Self–Efficacy			1		1	1	4	4	2	13	1.21%
Post–Course Performance			2		2		2	3	4	13	1.21%
Language					2	2	1	4	2	11	1.02%
Health						2	2	3	1	8	0.74%
Concept Assessment		1	2	1		1		1		6	0.56%
Browsing History				1		1	1			3	0.28%

Year

Table 9: Use of methods to predict student performance by year. Only methods identified at least five times are included.

Methods used \ Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	Percentage
Statistical: Linear Modeling	3	6	4	9	9	14	27	26	12	110	17.71%
Classification: Probabilistic Graphical Model	1	4	7	7	12	11	13	18	7	80	12.88%
Classification: Decision Trees	1	2	4	5	12	11	16	16	7	74	11.92%
Statistical: Correlation	2	5	2	8	4	9	10	13	4	57	9.18%
Classification: Neural Network	1	1	5	6	7	7	11	10	3	51	8.21%
Classification: SVM	1	2	5	4	5	5	7	11	5	45	7.25%
Classification: Classification				3	5	12	10	9	3	42	6.76%
Statistical: Latent Variable Models			1	2	2	2	10	5	5	27	4.35%
Classification: Random Forest		3	2	1	2	3	8	3	3	25	4.03%
Clustering: Partitioning–Based			3	1	3	1	3	4	4	19	3.06%
Classification: Nearest Neighbor		2	1	2	1	2	4	3	2	17	2.74%
Statistical: Variance Analysis					1	7	3		2	13	2.09%
Classification: Smote				1	3	2		4		10	1.61%
Mining: Text					2	3	2	2		9	1.45%
Classification: Logistic Regression	1		2				3	2	1	9	1.45%
Statistical: PCA				2		2	1	1	1	7	1.13%
Mining: Time					2	2		2		6	0.97%
Mining: Mining					1	2		1	1	5	0.81%
Mining: Association Rule						3	1	1		5	0.81%
Clustering: Hierarchical Clustering		1	1	1				1	1	5	0.81%
Clustering: Density–Based				1		1	1	1	1	5	0.81%

Table 10: Cross-tabulation of features (inputs) and performance values (outputs).

Performance metrics \ Features	Course Performance	Pre–Course Performance	Engagement	Gender	Demographic	Secondary School Performance	Personality	Family	Age	Self–Regulation	Motivation	Task Time	Log Data	Location	Discipline	Wealth	High School Quality	Social	Post–Course Performance	Self–Efficacy	Organization	ncept Assessment	Language	Health	Browsing History	Total	Percentage
Course Grade or Score	42	35	35	19	10	12	18	8	10	11	10	8	4	3	5	3	3	5	4	4	4	3	2	2	1	261	21%
Program or Module Graduation / Retention / Dropout	11	28	5	25	24	20	14	15	13	4	3	2		9	8	8	7	6	1	3	1		1	1		209	17%
Course Grade Range (e.g. A–B/C–F, Pass/Fail)	20	22	19	15	11	6	7	7	11	4	2	10	5	4	4	2	1	4	2	1	4	1		2	1	165	13%
Exam / Post–test Grade or Score	32	17	23	7	8	9	9	5	5	4	4	8	7	2	3		3	1	2	1	1		1	1		153	12%
Unspecified or Vague Performance	15	18	11	13	10	7	7	8	5	4	3	4	2	4	1	3	3	2	4	1	4		1		1	131	10%
GPA or GPA Range (incl. CGPA, SGPA)	9	18	3	13	8	15	13	7	7	3	3	1	1	5	3		2	2		3	4	1	3	1	1	126	10%
Assignment Performance (e.g. grade, time to completion)	20	7	21	2	1	3	4		1	2	5	7	9			1					4	1	1			89	7%
Early Detection of Course Retention / Dropout	7	10	9	7	7		6	2	5		4	5	1	2	1	2		1		2			1			72	6%
Knowledge Gain	1	3	2	4	2		3	2	2	1	2	1	1	2	1	1			2	2	1		1			34	3%
Number of Courses Passed or Failed	1	2		1	1	4	1	1	1					1	1	1	1									16	1%

Table 11: Comparison of disciplinary context and predicted values

	CS	STEM	Other	Unclear	Multi disciplinary
Course Grade or Score	42	22	7	2	6
Exam / Post-test Grade or Score	29	12	3	0	1
Course Grade Range (e.g. A-B/C-F, Pass/Fail)	18	17	4	1	6
Assignment Performance (e.g. grade, time to completion)	17	12	1	2	0
Unspecified or Vague Performance	14	6	3	7	5
GPA or GPA Range (incl. CGPA, SGPA)	6	11	12	1	5
Program or Module Graduation / Retention / Dropout	8	13	5	1	5

Table 12: Indicators of quality collected during the literature review

Question	Yes	No	Vague / Unclear	N/A
Is there a clearly defined research question?	118	97	143	3
Was the value being predicted clearly defined?	326	34	1	0
Were all the features described in sufficient detail to identify them?	197	110	52	2
Are the data collection instruments linked or included?	91	148	38	84
Is the research process clearly described?	235	66	60	0
Does the article discuss threats to validity?	64	262	30	5
Are the results presented with sufficient detail?	223	87	47	4
Has the work been verified in a second population?	28	316	11	6

(such as "interest" [235, 394] or "motivation" [131, 230, 375]) exist, but without a definition and without naming or providing the instruments used to collect the data, it's uncertain how the data was collected and whether the instruments used have been properly evaluated.

These omissions have a significant, detrimental effect on the ability of scholars to evaluate and replicate published work. We believe it is critically important that, as a community, we push for rigour and additional detail in how work is introduced and evaluated. We respectfully call on the community of authors and reviewers to normalize the explicit presentation of research questions; the use of a threats to validity section; a detailed description of what is predicted and of the features used for prediction including a justification of why these features have been selected; and the inclusion of the scales used, either as citations or appendices.

5.3.2 Quality of Results. Other aspects of quality we tracked speak to our ability to trust the results. Unfortunately, 24% of the articles we reviewed did not provide the minimum necessary statistical details to evaluate the results. Less than one in five explicitly discussed threats to validity using those terms, and only a quarter mentioned potential bias at all. Both of these issues make it difficult for us to trust that the authors have carefully considered their analyses.

We also saw a number of methodological flaws. For example, it's standard practice to separate the training sets (data used for fitting the model) and the test data set used for empirical validation. However, we found that many of the articles we evaluated where a model was constructed did not have a separate training set or, at least, did not mention having such a set.

Finally, we saw few efforts to generalize results. Very few (7.8%) of the studies we examined evaluated the predictions in more than one context (with a second population). A previous ITiCSE working group on education data mining [168] has already identified that much of data mining and learning analytics work in computing education is based on data from a single institution and even from a single course. They called for replication of work in multiple contexts, and we reiterate this call.

5.4 State of the Art

The field of predicting student performance is large and so far the work has not converged sufficiently for us to be able to determine the state of the art. Instead, based on observations from the survey process, we highlight a set of dimensions in which research is being conducted; methodology, data, usability, and application.

The *methodological dimension* is a time-wise (and consequently partially hype-wise) continuum of methods used for data analysis. It ranges from basic statistical analysis and correlations to the use of state of the art machine learning methods. For example, as deep learning has recently received increased attention in the machine learning research community, it has also been applied in predicting student performance.

The *data dimension* consists of three parts: data collection, quantity and granularity, each of which form their own dimensions. First, there is a shift from using data collection methods that need to be manually extracted, such as surveys filled using pen and paper, to using data collection methods that are automatic, such as working environment instrumentation and online surveys. Second, there is a movement from using single-semester single-course data, typically with post-hoc analysis, towards data from multiple semesters or multiple courses. Third, there is a movement from using coarse-grained data such as course grades towards more fine-grained data such as instrumented working environment data and physiological measurements. These dimensions are also related to starting to take a more holistic view of a student; recently, for example, student health, self-regulation, and study skills have gained attention.

The *usability dimension* is related to the data quantity dimension: given the movement from single-course single-semester data towards data from multiple courses or semesters, we are seeing predictive models that are evaluated on separate data sets. This increase of variability in the data can lead to models that generalize better. This usability dimension is related to the final dimension, which is the *application dimension.* Here, we are moving from research that shows that something can be done to doing something with the data, for example within the context of intelligent tutoring systems.

These dimensions highlight the wide range of approaches in this area, and research in even one of the dimensions may lead to improved performance prediction models. At the same time, while researchers typically push for novel contributions, we must highlight the importance of replication studies – even seemingly simple change of context may lead to different results – and reattempting existing work where one of the dimensions is adjusted. For example, work that may have looked like a dead end some time ago could currently lead to new findings as research in the dimensions is evolving.

5.5 Comparing Approaches

Drawing on the above observations and multiple dimensions, we briefly discuss the meaningfulness of comparing approaches for predicting students' performance. Researchers who work on predicting performance typically consider at least the methodological and data dimensions. Here, the data comes from the context that is observed and in which future academic performance is predicted.

In our survey, we found almost no datasets that have been published for wider use, and furthermore, only approximately 25% of the reviewed articles linked or included the used data collection instruments. This reflects on the findings of the ITiCSE 2015 working group on educational data mining, who point out the scarceness of open data sources [168].

Researchers can compare methodological approaches on the data at their disposal. Similarly, if they have access to multiple datasets, they can compare methodological approaches on all of those data sets. At the same time, as long as the data that is used to make predictions is not made open or publicly available, or even shared to a smaller group of researchers, making justified comparisons across studies is not straightforward.

5.6 The Risk of Models

While everyone is enthusiastic about the opportunities to help students that prediction can create, we should also consider the potential risk of misuse of the prediction mechanisms that are being created. Here, we first propose a deliberately extreme view on the factors that have been used to predict student performance by describing an "ideal" student:

The blood type of an ideal computer science student is O. He is a male of Asian descent and comes from a wealthy, well-respected family. His parents are well-educated, and he has no spouse or children. The student works on campus for eight hours each week. He reviews his work for at least three hours and no more than four hours each day after classes and does not visit [website] or play [popular game]. He has high self-efficacy, is highly motivated, and has a good self-regulation score. He has done well in all previous classes at our institution and was admitted with high marks from top quality primary and secondary schools.

We can, of course, see that this model is faulty. It is likely biased by the students currently in our program (not those who could be in our program) and is based on a number of immutable characteristics that reflect a background of privilege. Using this model to predict success has a high risk of causing harm. For example, if such a model were to be used for deciding who should enter a program, then the population would become increasingly privileged.

We have not, in this study, explicitly explored issues with bias or examined if the articles we reviewed considered the risk of using certain factors or publishing particular models. However, we saw little evidence that these issues are described in the literature we examined. We call on the community to be cautious in proposing and using models and ask that they consider the role that researchers and educators can play in public discourse.

5.7 Ethical Considerations

Finally, we raise several ethical issues resulting from our review: lack of consent, lack of inclusion, potentially unethical practices when collecting data, and issues of anonymity.

While we did not collect data on consent during the review, we recall seeing very few explicit indications that the data used in the articles we reviewed had been gathered (a) from public sources, (b) with consent from members of the population being studied, or (c) with the explicit approval of administrators responsible for institutional data. Textual analysis of the results confirms this impression. We searched the articles for indicators such as "consent," "authorized,""permission," "IRB," "review board," or "ethics," and found little evidence that approval for data collection is explicitly discussed. While not all institutions and societies require consent, we argue that it should at least be normal practice to acknowledge that ethical issues like the use of personal data, have been considered. We call for the community to normalize the reporting of the source of data and any consent or authorization processes (or the lack of a requirement for such processes) used in the collection of the data.

Very few of the articles we evaluated were explicitly inclusive with respect to gender. As far as we can tell, gender data was frequently presented with only two binary choices and without options like "another," "transgender," or "prefer not to disclose." It is also possible that a third option may have been present but was unused or unreported, but we could not, in many cases, evaluate this possibility since, in another example of reporting issues, the data collection instruments were not provided.

Data used in prediction studies is often secondary data as it is often gathered as normal practice during a course or program of study. Very few prediction studies recruit participants outside of an existing educational context or run formal experiments. Using secondary data can be problematic, as it may not have been anonymized appropriately and consent may not have been sought for use in a different context. For example, in one article, web traffic data from dormitories was monitored for normal network purposes, but the data was then stored and used in a study without notification.

Finally, we saw a few cases where participants were not adequately anonymized. This was, fortunately, rare, but in one notable example, actual student names appeared in a table alongside criteria used in student counselling.

We call on the reviewer community to screen for issues with anonymity and data collection and to request at the least the acknowledgement that ethical issues like consent have been considered and, where required, have been reviewed by an appropriate institutional review board.

5.8 Threats to Validity

As we were aware that systematic reviews have a number of limitations, we worked diligently to identify potential risks and to mitigate them.

A known risk of external validity is not having reviewed enough (or the appropriate) articles and including irrelevant material. The corpus for synthesis was created through metadata searches on the article indexes Scopus, IEEE Xplore, and ACM Digital Library. However, we found that some relevant articles are not indexed in these libraries. For example we discovered that articles published in the Journal of Educational Data Mining (JEDM) and LearnTechLib were not included in our corpus. We manually added articles from the Educational Data Mining community to our list, but it's likely that other un-indexed venues were missed. Similarly, it's possible that other search terms, such as the keyword "forecast," would uncover more material. However, the literature we are aware of is included, and we believe we have included a large enough sample to be representative. With respect to the inclusion of irrelevant material, we were careful to define explicit inclusion criteria and practiced the application of these criteria in pairs before beginning data collection. We are also confident that we eliminated inappropriate entries, such as poster abstracts, during the manual inspection of article titles and abstracts.

The study design and consistent application of the review template are both potential internal validity issues. The risk related to study design is around the question: Did we miss anything? It is important to mention here that we extracted the data based on our research question. Looking to our data with a different research focus would result in other data being extracted. We have provided our extraction form in Section 3.1.3 for transparency and review. Another risk when using several reviewers to extract data is the lack of consistency. To mitigate this risk, we began by reviewing articles in pairs; only after this initial session we proceeded to individual extractions. Furthermore, the first two-thirds of the data collected was collected when we were working in the same physical location, where questions were raised as they arose. The last third was collected asynchronously, but we feel that the reviewers were, by this point, comfortable with the extraction process.

6 CONCLUSION AND FUTURE WORK

The goal of this ITiCSE working group was to determine the current state of the research on predicting student academic performance. The work was conducted as a systematic literature review, reviewing relevant articles indexed by Scopus, IEEE Xplore and ACM indexes by June 2018. Works from the International Conference on Educational Data Mining were also examined, as the reviewers were aware of their existence and found that that the venue was not indexed in the databases searched. The final dataset that was analyzed contained 357 articles. These articles are listed in Table 13, categorized by the values they predicted.

6.1 Summarizing Research Question Results

Our main research questions of this work were: (1) What is the current state-of-the-art in predicting student performance? and (2) What is the quality of that work?

To summarize, during the recent years, there has been a clear increase in the amount of published research in the area, which can also be seen in the emergence of venues relevant for the topic. The majority of the work is looking at predicting easily attainable metrics such as individual course grade (38%), individual exam grade (14.7%), program retention or dropout (13.4%), GPA or cumulative GPA (12.2%), and assignment performance (11.4%). A small number of recent articles also examine measures that seek to better quantify learning performance, such as knowledge gain or speed in which the student will complete an assignment.

The features that have been used to predict student performance can be broadly split into five categories: demographic (e.g., age, gender), personality (e.g., self-efficacy, self-regulation), academic (e.g., high-school performance, course performance), behavioral (e.g., log data) and institutional (e.g., high-school quality, teaching approach). The majority of the articles used academic data for prediction (e.g., predicting course performance based on high-school performance). The use of data describing student behavior in a course (log data), while becoming more popular within the computing education research domain, is still relatively rare.

The methodologies that are used can be split into Classification (supervised learning, e.g., Naive Bayes, Decision Trees), Clustering (unsupervised learning, e.g., partitioning data), Statistical (e.g., correlation, regression), Data mining (identifying features and trends) and other methods. We found that (linear) regression models and classification methods are among the most frequently used tools, where the former is typically a method for the prediction, while for the latter the classification algorithms are often compared, leading to multiple prediction results.

When considering the quality of the existing work, there is room for improvement. The best articles that we read utilized data from multiple contexts and compared multiple methods to investigate a feature or variable of interest using multiple methods. However, we saw little re-use and sharing of data, which would allow us to compare methods or features, and we saw weaknesses in research methods and reporting. From the included articles, 33% included a clear research question, 18% discussed validity issues, and 8% verified the work in a second population. The last result echoes the finding of an ITiCSE working group on Educational Data mining [168], where the majority of the studies focused on a single course or a single institution, having no separate population with which the work would have been replicated with. On a positive note, 90% of the articles clearly defined what the target variable – or the predicted value – was, but this may partially be due to our inclusion criteria.

While there are no strong observable trends in emerging techniques, we highlighted a set of dimensions in which research is being conducted: methodology, data, usability, and application. Contributions to the body of predicting student performance can come in all of these dimensions – a researcher can, for example, study whether novel machine learning methods improve prediction when compared to other methods. There are areas with increasing interest such as the use of fine-grained log data, biometric data, and data describing students' tendencies such as self-regulation questionnaires. So far, no silver bullet has emerged.

6.2 Calls to the Community

Based on our literature review (see Table 2), interest in this area is growing. However, to make future studies more meaningful and to enable researchers to build upon published results, our first call to the community is to improve reporting standards. As an example, we provide a checklist for a minimum viable article/paper (MVP) that is reporting on work predicting student performance. Table 14, provides a checklist that each article focusing on predicting student performance should include. The list is a minimum requirements list, which essentially outlines the need for explicit and clear methodology and results.

Our second call to the community echoes that of the ITiCSE 2015 Educational Data Mining working group. We call for open data sets for developing approaches for predicting student performance. Having open data sets or *a baseline standard* would help researchers compare their data and methods with the results of others, helping them direct their efforts to more accurate models.

Our third call to the community, as revealed by the number of venues discovered during the literature survey, is related to the highly distributed community. We call for explicitly seeking collaborators in other communities, perhaps as part of a replication or comparison effort – such an effort would also help disseminate advancements. Submitting to and attending venues where work from a particular project has not yet been published would also be welcome and effective for creating connections.

Our fourth call to the community is related to the lack of comparing and replicating existing work. So far, relatively little work replicates previous efforts and more, but still relatively few, articles explicitly compare published methods for predicting performance. Doing this work would help identify more effective methods and would also provide an opportunity for broader collaboration. Changes to the previous algorithms can be iterative by nature. For example, work that measures the impact of adding an additional data source to an existing approach or that uses underutilized data sources, such as psychometrics, would be interesting.

Our final call to the community is related to reporting outcomes. While there is a push towards developing new methods, data, and so on, publishing information on approaches that did not work is important. Some of this work is published implicitly, for example in research where feature selection methods rule out a set of variables. We hope, however, that researchers would highlight methods and features that did not work with a similar zeal that is used to report what worked. Only through understanding both what works and what does not can we form a holistic understanding of the topic.

A REVIEW EXTRACTION INSTRUMENT

- Initial Vocabulary/Taxonomy for Performance
  - Assignment grade
  - Course retention / dropout
  - Course grade
  - Course grade range (e.g. A-C, D-F)
  - Course pass / fail
  - Exam grade
  - GPA
  - Graduation
  - Program retention / dropout
  - Unspecified performance
  - Not applicable
  - Other

Table 13: The 357 reviewed papers organized by the value they predict

Predicted Value	Papers
Program or Module Graduation / Retention	[7, 22, 46, 47, 62, 63, 71–73, 83, 87, 89, 94, 104, 107, 147, 149, 166, 172, 173, 177, 184, 186, 206, 208, 215, 218, 220, 231, 232, 244, 279, 296, 299, 315, 319, 327, 329, 337, 341, 358, 359, 373, 374, 407, 411, 424]
Exam / Post-test Grade or Score	[9, 10, 21, 24, 33, 35, 52, 59, 67, 68, 77, 80, 81, 85, 90, 96, 109, 114, 116, 127, 136, 152, 162, 163, 169, 193, 195, 199, 202, 205, 214, 217, 224, 233, 238, 241, 270, 274, 275, 277, 284, 287, 301, 302, 309, 314, 334, 346, 360, 368, 383, 420, 423]
Course Grade or Score	[1, 13, 14, 19, 21, 26, 27, 33, 34, 52, 57, 60, 64, 67–69, 74, 75, 78, 86, 93, 103, 105, 115, 118, 119, 122, 126–128, 133, 138, 141, 142, 144, 146, 158, 159, 171, 173–175, 183, 187, 197, 200, 203, 210, 237, 238, 240, 248, 250, 268, 272, 273, 285, 286, 292, 303, 304, 311, 312, 319, 321, 324, 326, 336, 348–351, 353–355, 357, 362, 369, 382, 391, 393, 396, 400, 409, 410, 412, 414, 417]
Number of Courses Passed or Failed	[120, 140, 176, 377]
Assignment Performance (e.g., grade, time to completion)	[52, 54, 59, 60, 65, 82, 92, 105, 106, 123, 124, 127, 143, 151, 165, 179, 187, 188, 195, 209, 216, 223, 228, 239, 242, 251, 256, 264, 286, 309, 310, 352, 364, 365, 376, 383, 389, 395, 399, 408, 415]
Unspecified or Vague Performance	[3, 6, 8, 11, 15, 18, 48, 49, 53, 55, 56, 73, 97, 100, 101, 117, 125, 132, 134, 187, 199, 229, 234, 245, 247, 271, 276, 280, 281, 288, 290, 294, 295, 298, 300, 328, 334, 342, 343, 367, 371, 392, 410]
Course Retention / Dropout	[42, 74, 91, 153, 172, 178, 192, 196, 212, 213, 218, 220, 225, 257, 261, 262, 388, 398, 403, 422]
GPA or GPA Range (including CGPA, SGPA)	[4, 12, 28–32, 36, 37, 64, 66, 73, 113, 129, 135, 137, 140, 149, 167, 181, 189, 206, 211, 218, 240, 249, 255, 263, 265, 266, 268, 291, 297, 322, 325, 339, 344, 345, 366, 385, 390, 402, 405, 413]
Course Grade Range (e.g., A-B/C-F, Pass/Fail)	[17, 20, 38, 52, 58, 61, 77, 95, 108, 110, 111, 121, 126, 130, 139, 172, 174, 182, 194, 207, 222, 226, 227, 238, 254, 257, 258, 260, 267, 278, 289, 293, 316, 317, 323, 331, 332, 361, 363, 370, 379–381, 386, 402, 403, 406, 418, 421]
Knowledge Gain	[23, 156, 164, 192, 201, 270, 401, 419]

Table 14: Checklist for a minimum viable article on predicting student performance

	Objective
☐	Define what is being predicted. If the value describing performance (e.g., course grade) consists of multiple items (e.g., course exam, course assignments), describe the contribution (weight) of each item when the performance value is calculated.
☐	Define the factors used for prediction. Describe them in such detail that a reader that is not familiar with your particular context understands them. If factors are intertwined (e.g., course assignments) with the predicted value (e.g., course exam), be explicit about the connection. Provide links to, or if not possible, include the scales and surveys that have been used when collecting data.
☐	Define the methodologies used for prediction and link the methods used by referencing appropriate articles. Unless you propose a novel method, formal proofs, etc. are not required. If you use feature selection, include details on them.
☐	Define the data. Explain where the data comes from, if it is self-reported or automatically collected and if students are compensated for participating. Moreover, if the data contains students from a course, discuss the number of students in the course, describe how many were excluded from the analysis and why, and provide descriptive statistics that outlines the data. Be specific on whether the data is from a single course, or single institution, and also discuss if a separate data set is used for validating the prediction results.
☐	Provide the results. Perform and report on the tests necessary to test required attributes of the data. Name the analyses being performed and report all the relevant statistics to allow for interpretation of the results.
☐	Discuss the reasons why specific factors, performance metrics and methods were chosen (or omitted).
☐	Reflect upon the results and consider why the methods and factors used did work or did not work. What are the particular context-specific issues that may influence the outcomes?
☐	Describe threats to validity and limitations. Note situations in which a model or approach might be applied as well as where it is not valid.

- Initial Vocabulary/Taxonomy for Population
  - K-12 (from kindergarten to high-school)
  - Minors
  - Majors
  - Non-majors
  - Professional
  - Unknown / Unclear / Vague
  - Graduate (e.g. MSc students)
  - Undergraduate (e.g. BSc students)
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Training mode
  - Blended
  - Local
  - MOOC
  - Online
  - Unclear / vague
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Education Type
  - Formal education (e.g. university education, high-school)
  - Informal education (e.g. MOOC, work-related training, ...)
  - Not applicable
- Initial Vocabulary/Taxonomy for Data set information
  - Data coming from a single institution
  - Data coming from multiple institutions
  - Data coming from a single course
  - Data coming from multiple courses
  - Data is a publicly available dataset
  - Other

- Initial Vocabulary/Taxonomy for Family background
  - Family background in General
  - Income
  - Status
  - Support
  - Parent Educational Level
  - Parent Occupation
  - Number of Siblings
  - Caretaker Role
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Demographic data
  - Demographic data in General
  - Accessibility (disability)
  - Gender
  - Age
  - Ethnicity
  - International
  - Minority
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Working Conditions
  - Working Conditions in General
  - Distance to school
  - Daily commute time
  - Access to internet
  - Basic needs (water and toilet)
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Education background
  - Education background in General
  - Accommodations (modified programs, disability)
  - Extra-curricular activities
  - Background in CS
  - Background in Math
  - Background in Physics
  - GPA (high school)
  - GPA (post-secondary)
  - GPA (unknown source)
  - High School Quality
  - Previous (hobby) experience on topic
  - Previous (work) experience on topic
  - Standardized graduate admission test (e.g. GRE)
  - Standardized undergraduate admissions test (e.g. SAT, ACT)
  - Background in other disciplines
  - Grades from previous courses (at the post-secondary institution)
  - Grades from previous courses (before post-secondary institution)
  - Education background Not applicable
  - Education background: Other
- Initial Vocabulary/Taxonomy for Course (current or parallel)
  - Course in General
  - Attendance
  - Activity (with learning resources or materials)
  - Activity (with discussion forums or chats)
  - End of term assessment (exams)
  - Time on task
  - Marks
  - Marks (tests and quizzes)
  - Marks (assignments)
  - Marks (lab work)
  - Midterm assessment (tests, quizzes, midterm exams)
  - Pre-test score / mark
  - Organizational: Pedagogical methods
  - Organizational: Teaching mode
  - Organizational: Materials Available (what types?)
  - Social: Related social data (e.g. number of friends in class)
  - Social: Unrelated social data (e.g. facebook, twitter)
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Student motivation
  - Motivation in General
  - Desired grades
  - Extrinsic
  - Intrinsic / interest / passion
  - Importance of Grades
  - Utility
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Psychological/affective/learning
  - Scales in General
  - Achievement goals
  - Emotional (self-worth, anxiety, etc)
  - Goal-orientation (performance, mastery, etc)
  - Grit
  - Self-efficacy
  - Learning Strategies (deep, surface)
  - Learning Styles (!)
  - MSLQ
  - Personality
  - Self-regulation
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for Research Method type
  - Mixed-methods
  - Qualitative
  - Quantitative
  - Not applicable
  - Other
- Initial Vocabulary/Taxonomy for classification (supervised learning)
  - Classification in General
  - Neural network
  - Adaptive boosting
  - SMOTE
  - Radial Basis
  - Naive Bayes
  - Nearest Neighbor
  - Decision Trees
  - Random Forest
  - SVM
  - Knowledge modeling
  - Bayesian network
  - Other
- Initial Vocabulary/Taxonomy for Clustering (unsupervised learning)
  - Clustering in General
  - Neural network
  - K-means
  - K-star
  - Other
- Initial Vocabulary/Taxonomy for Mining (finding frequent patterns/ feature extraction)
  - Mining in General
  - Distributed
  - Text
  - Web

  - Temporal
  - Sequential
  - Association Rule
  - GUHA Method
  - Other
- Initial Vocabulary/Taxonomy for Statistical methods used
  - Statistical methods in General
  - ANOVA
  - Correlation
  - Factor Analysis
  - Regression
  - T-Test or other test of var. between populations
  - Logistic regression
  - Structural Equation Modeling
  - Other

REFERENCES

[1] R. S. Abdulwahhab and S. S. Abdulwahab. 2017. Integrating learning analytics to predict student performance behavior. In *6th International Conference on Information and Communication Technology and Accessibility (ICTA)*. IEEE, 1–6.

[2] William H Acton, Peder J Johnson, and Timothy E Goldsmith. 1994. Structural knowledge assessment: comparison of referent structures. *Journal of Educational Psychology* 86, 2 (1994), 303–311.

[3] Seth A Adjei, Anthony F Botelho, and Neil T Heffernan. 2016. Predicting student performance on post-requisite skills using prerequisite skill data: an alternative method for refining prerequisite skill structures. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 469–473.

[4] Lilly Suriani Affendey, IHM Paris, N Mustapha, Md Nasir Sulaiman, and Z Muda. 2010. Ranking of influencing factors in predicting students' academic performance. *Information Technology Journal* 9, 4 (2010), 832–837.

[5] H.W. Aggarwal, P. Bermel, N.M. Hicks, K.A. Douglas, H.A. Diefes-Dux, and K. Madhavan. 2017. Using pre-course survey responses to predict sporadic learner behaviors in advanced STEM MOOCs work-in-progress. *Frontiers in Education Conference (FIE)* (2017), 1–4.

[6] Ángel F Agudo-Peregrina, Ángel Hernández-García, and Santiago Iglesias-Pradas. 2012. Predicting academic performance with learning analytics in virtual learning environments: A comparative study of three interaction classifications. In *Computers in Education (SIIE)*. IEEE, 1–6.

[7] Everaldo Aguiar, Nitesh V Chawla, Jay Brockman, G Alex Ambrose, and Victoria Goodrich. 2014. Engagement vs performance: using electronic portfolios to predict first semester engineering student retention. In *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge*. ACM, 103–112.

[8] Alireza Ahadi, Vahid Behbood, Arto Vihavainen, Julia Prior, and Raymond Lister. 2016. Students' syntactic mistakes in writing seven different types of SQL queries and its application to predicting students' success. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 401–406.

[9] Alireza Ahadi, Raymond Lister, Heikki Haapala, and Arto Vihavainen. 2015. Exploring machine learning methods to automatically identify students in need of assistance. In *Proceedings of the Eleventh International Computing Education Research Conference*. ACM, 121–130.

[10] Alireza Ahadi, Raymond Lister, and Arto Vihavainen. 2016. On the number of attempts students made on some online programming exercises during semester and their subsequent performance on final exam questions. In *Proceedings of the 2016 Conference on Innovation and Technology in Computer Science Education*. ACM, 218–223.

[11] Fadhilah Ahmad, NurHafieza Ismail, and Azwa Abdul Aziz. 2015. The prediction of students' academic performance using classification data mining techniques. *Applied Mathematical Sciences* 9, 129 (2015), 6415–6426.

[12] Ahmed Al-Azawei, Ali Al-Bermani, and Karsten Lundqvist. 2016. Evaluating the effect of Arabic engineering students' learning styles in blended programming courses. *Journal of Information Technology Education: Research* 15 (2016), 109–130.

[13] Mohammad Majid al Rifaie, Matthew Yee-King, and Mark d'Inverno. 2016. Investigating swarm intelligence for performance prediction. In *Proceedings of the 9th International Conference on Educational Data Mining*. 264–269.

[14] Huda Al-Shehri, Amani Al-Qarni, Leena Al-Saati, Arwa Batoaq, Haifa Badukhen, Saleh Alrashed, Jamal Alhiyafi, and Sunday O Olatunji. 2017. Student performance prediction using support vector machine and k-nearest neighbor. In *International Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 1–4.

[15] Zahyah Alharbi, James Cornford, Liam Dolder, and Beatriz De La Iglesia. 2016. Using data mining techniques to predict students at risk of poor performance. In *2016 SAI Computing Conference*. IEEE, 523–531.

[16] Ruba Alkhasawneh and Rosalyn Hobson. 2011. Modeling student retention in science and engineering disciplines using neural networks. In *Global Engineering Education Conference (EDUCON)*. IEEE, 660–663.

[17] Hind Almayan and Waheeda Al Mayyan. 2016. Improving accuracy of students' final grade prediction model using PSO. In *Information Communication and Management (ICICM)*. IEEE, 35–39.

[18] Ismail Almuniri and Aiman Moyaid Said. 2018. Predicting the performance of school: Case study in Sultanate of Oman. In *2018 International Conference on Information and Computer Technologies (ICICT)*. IEEE, DeKalb, IL, 18–21.

[19] F.M. Almutairi, N.D. Sidiropoulos, and G. Karypis. 2017. Context-aware recommendation-based learning analytics using tensor and coupled matrix factorization. *Journal on Selected Topics in Signal Processing* 11, 5 (2017), 729–741.

[20] Muzaffer Ege Alper and Zehra Cataltepe. 2012. Improving course success prediction using ABET course outcomes and grades. In *CSEDU*. 222–229.

[21] Christine Alvarado, Cynthia Bailey Lee, and Gary Gillespie. 2014. New CS1 pedagogies and curriculum, the same success factors?. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*. ACM, 379–384.

[22] Sattar Ameri, Mahtab J Fard, Ratna B Chinnam, and Chandan K Reddy. 2016. Survival analysis based framework for early prediction of student dropouts. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 903–912.

[23] Pensri Amornsinlaphachai. 2016. Efficiency of data mining models to predict academic performance and a cooperative learning model. In *International Conference on Knowledge and Smart Technology (KST)*. IEEE, 66–71.

[24] VK Anand, SK Abdul Rahiman, E Ben George, and AS Huda. 2018. Recursive clustering technique for students' performance evaluation in programming courses. In *Majan International Conference (MIC)*. IEEE, 1–5.

[25] M Anoopkumar and AMJ Md Zubair Rahman. 2016. A review on data mining techniques and factors used in educational data mining to predict student amelioration. In *International Conference on Data Mining and Advanced Computing (SAPIENCE)*. IEEE, 122–133.

[26] Adam Anthony and Mitch Raney. 2012. Bayesian network analysis of computer science grade distributions. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*. ACM, 649–654.

[27] Yojna Arora, Abhishek Singhal, and Abhay Bansal. 2014. PREDICTION & WARNING: a method to improve student's performance. *ACM SIGSOFT Software Engineering Notes* 39, 1 (2014), 1–5.

[28] Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan. 2012. Neural network model to predict electrical students' academic performance. In *International Congress on Engineering Education (ICEED)*. IEEE, 1–5.

[29] Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan. 2013. Prediction of engineering students' academic performance using artificial neural network and linear regression: a comparison. In *International Congress on Engineering Education (ICEED)*. IEEE, 43–48.

[30] Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan. 2014. Neural network and linear regression methods for prediction of students' academic achievement. In *Global Engineering Education Conference (EDUCON)*. IEEE, 916–921.

[31] Pauziah Mohd Arsad, Norlida Buniyamin, Jamalul-Lail Ab Manan, and Noraliza Hamzah. 2011. Proposed academic students' performance prediction model: A Malaysian case study. In *International Congress on Engineering Education (ICEED)*. IEEE, 90–94.

[32] Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan. 2013. A neural network students' performance prediction model (NNSPPM). In *2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*. IEEE, Kuala Lumpur, Malaysia, 1–5.

[33] Michael Mogessie Ashenafi, Giuseppe Riccardi, and Marco Ronchetti. 2015. Predicting students' final exam scores from their course activities. In *Frontiers in Education Conference (FIE)*. IEEE, Camino Real El Paso, El Paso, TX, USA, 1–9.

[34] Michael Mogessie Ashenafi, Marco Ronchetti, and Giuseppe Riccardi. 2016. Predicting Student Progress from Peer-Assessment Data.. In *Proceedings of the 9th International Conference on Educational Data Mining*. 270–275.

[35] D. Azcona and A.F. Smeaton. 2017. Targeting at-risk students using engagement and effort predictors in an introductory computer programming course. *Lecture Notes in Computer Science* 10474 LNCS (2017), 361–366.

[36] Azwa Abdul Aziz, Nur Hafieza Ismail, Fadhilah Ahmad, and Hasni Hassan. 2015. A framework for students' academic performance analysis using naïv Bayes classifier. *Jurnal Teknologi (Sciences & Engineering)* 75, 3 (2015), 13–19.

[37] Fatihah Aziz, Abd Wahab Jusoh, and Mohd Syafarudy Abu. 2015. A comparison of student academic achievement using decision trees techniques: Reflection from University Malaysia Perlis. In *AIP Conference Proceedings*, Vol. 1660. AIP Publishing, 050034.

[38] Ghada Badr, Afnan Algobail, Hanadi Almutairi, and Manal Almutery. 2016. Predicting students' performance in university courses: a case study and tool in KSU mathematics department. *Procedia Computer Science* 82 (2016), 80–89.

[39] Ryan SJD Baker and Kalina Yacef. 2009. The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining (JEDM)* 1, 1 (2009), 3–17.

[40] Gabriel Barata, Sandra Gama, Joaquim Jorge, and Daniel Gonçalves. 2016. Early prediction of student profiles based on performance and gaming preferences.

Transactions on Learning Technologies 9, 3 (2016), 272–284.
[41] L. Barba-Guamán and P. Valdiviezo-Díaz. 2017. Improve the performance of students in the mathematics learning through Bayesian model. *7th International Workshop on Computer Science and Engineering (WCSE)* (2017), 349–354.
[42] Laci Mary Barbosa Manhães, Sérgio Manuel Serra da Cruz, and Geraldo Zimbrão. 2015. Towards automatic prediction of student performance in STEM undergraduate degree programs. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*. ACM, 247–253.
[43] Dennis Barker. 1986. Developing creative problem solving in civil engineering. *Assessment & Evaluation in Higher Education* 11, 3 (Sept. 1986), 192–208.
[44] Lecia J Barker, Charlie McDowell, and Kimberly Kalahar. 2009. Exploring factors that influence computer science introductory course students to persist in the major. In *ACM SIGCSE Bulletin*, Vol. 41. ACM, 153–157.
[45] Hermine Baum, Miriam Litchfield, and MF Washburn. 1919. The results of certain standard mental tests as related to the academic records of college seniors. *The American Journal of Psychology* (1919), 307–310.
[46] Jaroslav Bayer, Hana Bydzovská, Jan Géryk, Tomás Obsivac, and Lubomir Popelinsky. 2012. Predicting drop-out from social behaviour of students. In *Proceedings of the 5th International Conference on Educational Data Mining*.
[47] Christopher T Belser, Diandra J Prescod, Andrew P Daire, Melissa A Dagley, and Cynthia Y Young. 2017. Predicting undergraduate student retention in STEM majors based on career development factors. *The Career Development Quarterly* 65, 1 (2017), 88–93.
[48] Ceasar Ian P Benablo, Evangeline T Sarte, Joe Marie D Dormido, and Thelma Palaoag. 2018. Higher education student's academic performance analysis through predictive analytics. In *Proceedings of the 2018 7th International Conference on Software and Computer Applications*. ACM, 238–242.
[49] Hoang Tieu Binh et al. 2017. Predicting students' performance based on learning style by using artificial neural networks. In *2017 9th International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 48–53.
[50] David M. Blei. 2012. Probabilistic topic models. *Commun. ACM* 55, 4 (April 2012), 77.
[51] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3, Jan (2003), 993–1022.
[52] Paulo Blikstein, Marcelo Worsley, Chris Piech, Mehran Sahami, Steven Cooper, and Daphne Koller. 2014. Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *Journal of the Learning Sciences* 23, 4 (2014), 561–599.
[53] Nigel Bosch and Sidney D'Mello. 2014. It takes two: momentary co-occurrence of affective states during computerized learning. In *International Conference on Intelligent Tutoring Systems*. Springer, 638–639.
[54] Christopher G Brinton and Mung Chiang. 2015. MOOC performance prediction via clickstream data and social learning networks. In *2015 International Conference on Computer Communications (INFOCOM)*. IEEE, 2299–2307.
[55] Norlida Buniyamin, Usamah bin Mat, and Pauziah Mohd Arshad. 2015. Educational data mining for prediction and classification of engineering students achievement. In *International Congress on Engineering Education (ICEED)*. IEEE, 49–53.
[56] Hana Bydovska and Lubomír Popelínskỳ. 2013. Predicting student performance in higher education. In *2013 24th International Workshop on Database and Expert Systems Applications (DEXA)*. IEEE, 141–145.
[57] Hana Bydžovská. 2016. A comparative analysis of techniques for predicting student performance. In *Proceedings of the 9th International Conference on Educational Data Mining*.
[58] Dino Capovilla, Peter Hubwieser, and Philipp Shah. 2016. DiCS-Index: Predicting student performance in computer science by analyzing learning behaviors. In *2016 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)*. IEEE, 136–140.
[59] R.M. Carro and V. Sanchez-Horreo. 2017. The effect of personality and learning styles on individual and collaborative learning: Obtaining criteria for adaptation. *IEEE Global Engineering Education Conference (EDUCON)* (2017), 1585–1590.
[60] Adam S Carter, Christopher D Hundhausen, and Olusola Adesope. 2015. The normalized programming state model: Predicting student performance in computing courses based on programming behavior. In *Proceedings of the Eleventh International Computing Education Research Conference*. ACM, 141–150.
[61] K. Casey and D. Azcona. 2017. Utilizing student activity patterns to predict performance. *International Journal of Educational Technology in Higher Education* 14, 1 (2017).
[62] Erin Cech, Brian Rubineau, Susan Silbey, and Caroll Seron. 2011. Professional role confidence and gendered persistence in engineering. *American Sociological Review* 76, 5 (2011), 641–666.
[63] Nihat Cengiz and Arban Uka. 2014. Prediction of student success using enrolment data. *KOS* 14, 17 (2014), 45–2.
[64] L. Chan, R. Sleezer, J.J. Swanson, M. Ahrens, and R.A. Bates. 2017. Difficulty in predicting performance in a project-based learning program. *ASEE Annual Conference & Exposition* (2017).
[65] R. Chaturvedi and C.I. Ezeife. 2017. Predicting student performance in an ITS using task-driven features. *17th IEEE International Conference on Computer and Information Technology (CIT)* (2017), 168–175.
[66] Jeng-Fung Chen and Quang Hung Do. 2014. A cooperative cuckoo search–hierarchical adaptive neuro-fuzzy inference system approach for predicting student academic performance. *Journal of Intelligent & Fuzzy Systems* 27, 5 (2014), 2551–2561.
[67] Xin Chen, Lori Breslow, and Jennifer DeBoer. 2018. Analyzing productive learning behaviors for students using immediate corrective feedback in a blended learning environment. *Computers & Education* 117 (2018), 59–74.
[68] Yujing Chen, Aditya Johri, and Huzefa Rangwala. 2018. Running out of stem: a comparative study across stem majors of college students at-risk of dropping out early. In *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. ACM, 270–279.
[69] YY Chen, Shakirah Mohd Taib, and Che Sarah Che Nordin. 2012. Determinants of student performance in advanced programming course. In *2012 International Conference for Internet Technology and Secured Transactions*. IEEE, 304–307.
[70] Fatma Chiheb, Fatima Boumahdi, Hafida Bouarfa, and Doulkifli Boukraa. 2017. Predicting students performance using decision trees: Case of an Algerian university. In *2017 International Conference on Mathematics and Information Technology (ICMIT)*. IEEE, 113–121.
[71] D.S. Choi and M.C. Loui. 2015. Grit for engineering students. *Frontiers in Education Conference (FIE)*.
[72] D.S. Choi, B. Myers, and M.C. Loui. 2017. Grit and two-year engineering retention. *Frontiers in Education Conference (FIE)* (2017), 1–3.
[73] Tjioe Marvin Christian and Mewati Ayub. 2014. Exploration of classification using NBTree for predicting students' performance. In *2014 International Conference on Data and Software Engineering (ICODSE)*. IEEE, 1–6.
[74] Mi Chunqiao, Peng Xiaoning, and Deng Qingyou. 2017. An artificial neural network approach to student study failure risk early warning prediction based on TensorFlow. In *International Conference on Advanced Hybrid Information Processing*. Springer, 326–333.
[75] Carleton Coffrin, Linda Corrin, Paula de Barba, and Gregor Kennedy. 2014. Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge*. ACM, 83–92.
[76] Michael A Collura, Shannon Ciston, and Nancy Ortins Savage. 2011. Effect of freshman chemistry on student performance in sophomore engineering courses. In *ASEE Annual Conference & Exposition*. 9.
[77] Rianne Conijn, Chris Snijders, Ad Kleingeld, and Uwe Matzat. 2017. Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *Transactions on Learning Technologies* 10, 1 (2017), 17–29.
[78] B. M. Corsatea and S. Walker. 2015. Opportunities for Moodle data and learning intelligence in virtual environments. In *2015 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*. 1–6.
[79] Jennifer D Cribbs, Cheryl Cass, Zahra Hazari, Philip M Sadler, and Gerhard Sonnert. 2016. Mathematics identity and student persistence in engineering. *International Journal of Engineering Education* 32, 1 (2016), 163–171.
[80] Scott Crossley, Ran Liu, and Danielle McNamara. 2017. Predicting math performance using natural language processing tools. In *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge*. ACM, 339–347.
[81] Diana Cukierman. 2015. Predicting success in university first year computing science courses: The role of student participation in reflective learning activities and in i-clicker activities. In *Proceedings of the 2015 Conference on Innovation and Technology in Computer Science Education*. ACM, 248–253.
[82] Ryan SJ d Baker, Zachary A Pardos, Sujith M Gowda, Bahador B Nooraei, and Neil T Heffernan. 2011. Ensembling predictions of student knowledge within intelligent tutoring systems. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 13–24.
[83] Phil Dacunto. 2016. Academic "Predestination": Does It Exist?. In *2016 ASEE Annual Conference & Exposition Proceedings*. ASEE Conferences.
[84] D. D'Amato, N. Droste, B. Allen, M. Kettunen, K. Lähtinen, J. Korhonen, P. Leskinen, B. D. Matthies, and A. Toppinen. 2017. Green, circular, bio economy: A comparative analysis of sustainability avenues. *Journal of Cleaner Production* (2017).
[85] Holger Danielsiek and Jan Vahrenhold. 2016. Stay on these roads: Potential factors indicating students' performance in a CS2 course. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 12–17.
[86] S.M. Darwish. 2017. Uncertain measurement for student performance evaluation based on selection of boosted fuzzy rules. *IET Science, Measurement and Technology* 11, 2 (2017), 213–219.
[87] Ali Daud, Naif Radi Aljohani, Rabeeh Ayaz Abbasi, Miltiadis D Lytras, Farhat Abbas, and Jalal S Alowibdi. 2017. Predicting student performance using advanced learning analytics. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 415–421.
[88] Simon P Davies. 1991. The role of notation and knowledge representation in the determination of programming strategy: a framework for integrating models of programming behavior. *Cognitive Science* 15, 4 (1991), 547–572.

[89] Saylisse Dávila, Wandaliz Torres-García, and Viviana I Cesaní. 2015. Mining the profile of successful IE students: Using historical data to drive curricular interventions. In *IISE Annual Conference*. Institute of Industrial and Systems Engineers (IISE), 2892.

[90] Rosângela Marques de Albuquerque, André Alves Bezerra, Darielson Araujo de Souza, Luís Bruno Pereira do Nascimento, Jarbas Joaci de Mesquita Sá, and José Cláudio do Nascimento. 2015. Using neural networks to predict the future performance of students. In *Computers in Education (SIIE)*. IEEE, 109–113.

[91] David de la Peña, Juan A Lara, David Lizcano, María A Martínez, Concepción Burgos, and María L Campanario. 2017. Mining activity grades to model students' performance. In *2017 International Conference on Engineering & MIS (ICEMIS)*. IEEE, 1–6.

[92] Maria De Marsico, Andrea Sterbini, and Marco Temperini. 2016. Modeling peer assessment as a personalized predictor of teacher's grades: The case of OpenAnswer. In *2016 International Conference on Information Technology Based Higher Education and Training (ITHET)*. IEEE, 1–5.

[93] Gilberto de Melo, Sanderson M Oliveira, Cintia C Ferreira, Enio P Vasconcelos Filho, Wesley P Calixto, and Geovanne P Furriel. 2017. Evaluation techniques of machine learning in task of reprovation prediction of technical high school students. In *2017 Chilean Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. IEEE, 1–7.

[94] José R Deliz, Rolando García Juan C Morales, and Gloribel Rivera. 2015. Markov chain modeling: student flow through remedial Mathematics to sophomore ISERC. In *IISE Annual Conference*. Institute of Industrial and Systems Engineers (IISE), 3171.

[95] R.M. DeMonbrun and M.G. Brown. 2017. Exploring the relationship between the use of learning technologies and student success in the engineering classroom. *ASEE Annual Conference & Exposition* (2017).

[96] Paul Denny, Andrew Luxton-Reilly, John Hamer, Dana B Dahlstrom, and Helen C Purchase. 2010. Self-predicted and actual performance in an introductory programming course. In *Proceedings of the Fifteenth Conference on Innovation and Technology in Computer Science Education*. ACM, 118–122.

[97] Jayalatchumy Dhanpal, Thambidurai Perumal, et al. 2016. Efficient graph clustering algorithm and its use in prediction of students performance. In *Proceedings of the International Conference on Informatics and Analytics*. ACM, 39.

[98] Daniele Di Mitri, Maren Scheffel, Hendrik Drachsler, Dirk Börner, Stefaan Ternier, and Marcus Specht. 2017. Learning Pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data. In *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge*. ACM, 188–197.

[99] A. Dinesh Kumar, R. Pandi Selvam, and K. Sathesh Kumar. 2018. Review on prediction algorithms in educational data mining. *International Journal of Pure and Applied Mathematics* 118, Special Issue 8 (2018), 531–536.

[100] Blazenka Divjak and Dijana Oreski. 2009. Prediction of academic performance using discriminant analysis. In *2009 31st International Conference on Information Technology Interfaces (ITI)*. IEEE, 225–230.

[101] Oana Dumitrascu and Rodica Ciudin. 2015. Modeling factors with influence on sustainable university management. *Sustainability* 7, 2 (2015), 1483–1502.

[102] Ashish Dutt, Maizatul Akmar Ismail, and Tutut Herawan. 2017. A systematic review on educational data mining. *IEEE Access* 5 (2017), 15991–16005.

[103] Omar Augusto Echegaray-Calderon and Dennis Barrios-Aranibar. 2015. Optimal selection of factors using genetic algorithms and neural networks for the prediction of students' academic performance. In *2015 Latin America Conference on Computational Intelligence (LA-CCI)*. IEEE, 1–6.

[104] Edward M. Elias and Carl A. Lindsay. 1968. *The Role of Intellective Variables in Achievement and Attrition of Associate Degree Students at the York Campus for the Years 1959 to 1963.* Technical Report Report No-PSU-68-7. Pennsylvania State University. 21 pages.

[105] Asmaa Elbadrawy, Agoritsa Polyzou, Zhiyun Ren, Mackenzie Sweeney, George Karypis, and Huzefa Rangwala. 2016. Predicting student performance using personalized analytics. *Computer* 49, 4 (2016), 61–69.

[106] Asmaa Elbadrawy, R Scott Studham, and George Karypis. 2015. Collaborative multi-regression models for predicting students' performance in course activities. In *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge*. ACM, 103–107.

[107] Lorelle Espinosa. 2011. Pipelines and pathways: Women of color in undergraduate STEM majors and the college experiences that contribute to persistence. *Harvard Educational Review* 81, 2 (2011), 209–241.

[108] Anthony Estey and Yvonne Coady. 2016. Can interaction patterns with supplemental study tools predict outcomes in CS1?. In *Proceedings of the 2016 Conference on Innovation and Technology in Computer Science Education*. ACM, 236–241.

[109] Anthony Estey and Yvonne Coady. 2017. Study habits, exam performance, and confidence: How do workflow practices and self-efficacy ratings align?. In *Proceedings of the 2017 Conference on Innovation and Technology in Computer Science Education*. ACM, 158–163.

[110] D.S. Evale. 2017. Learning management system with prediction model and course-content recommendation module. *Journal of Information Technology Education: Research* 16, 1 (2017), 437–457.

[111] Digna S Evale, Menchita F Dumlao, Shaneth Ambat, and Melvin Ballera. 2016. Prediction model for students' performance in Java programming with course-content recommendation system. In *Proceedings of 2016 Universal Technology Management Conference (UTMC)*. Minnesota, United States of America, 5.

[112] Gerald E Evans and Mark G Simkin. 1989. What best predicts computer proficiency? *Commun. ACM* 32, 11 (1989), 1322–1327.

[113] Nickolas JG Falkner and Katrina E Falkner. 2012. A fast measure for identifying at-risk students in computer science. In *Proceedings of the Ninth International Computing Education Research Conference*. ACM, 55–62.

[114] Stephen E Fancsali, Guoguo Zheng, Yanyan Tan, Steven Ritter, Susan R Berman, and April Galyardt. 2018. Using embedded formative assessment to predict state summative test scores. In *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. ACM, 161–170.

[115] Yunping Feng, Di Chen, Zihao Zhao, Haopeng Chen, and Puzhao Xi. 2015. The impact of students and TAs' participation on students' academic performance in MOOC. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 1149–1154.

[116] Manuel Fernández-Delgado, Manuel Mucientes, Borja Vázquez-Barreiros, and Manuel Lama. 2014. Learning analytics for the prediction of the educational objectives achievement. In *Frontiers in Education Conference (FIE)*. IEEE, 1–4.

[117] Ángel Fidalgo-Blanco, María Luisa Sein-Echaluce, Javier Esteban-Escaño, Francisco J García Peñalvo, and Miguel Ángel Conde. 2016. Learning analytics to identify the influence of leadership on the academic performance of work teams. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*. ACM, 377–382.

[118] Eric Fitzsimmons, Stacey Tucker-Kulesza, Xiongya Li, Whitney Jeter, and Jana Fallin. 2016. The engineering classroom is still relevant. In *2016 ASEE Annual Conference & Exposition*. ASEE Conferences.

[119] S. Ganguly, M. Kulkarni, and M. Gupta. 2017. Predictors of academic performance among Indian students. *Social Psychology of Education* 20, 1 (2017), 139–157.

[120] Ernesto Pathros Ibarra García and Pablo Medina Mora. 2011. Model prediction of academic performance for first year students. In *2011 10th Mexican International Conference on Artificial Intelligence (MICAI)*. IEEE, 169–174.

[121] Eleazar Gil-Herrera, Athanasios Tsalatsanis, Ali Yalcin, and Autar Kaw. 2011. Predicting academic performance using a rough set theory-based knowledge discovery methodology. *International Journal of Engineering Education* 27, 5 (2011), 992.

[122] Kazumasa Goda, Sachio Hirokawa, and Tsunenori Mine. 2013. Correlation of grade prediction performance and validity of self-evaluation comments. In *Proceedings of the 14th Annual ACM SIGITE Conference on Information Technology Education*. ACM, 35–42.

[123] Matthew C Gombolay, Reed Jensen, and Sung-Hyun Son. 2017. Machine Learning Techniques for Analyzing Training Behavior in Serious Gaming. *IEEE Transactions on Computational Intelligence and AI in Games* (2017).

[124] Yue Gong and Joseph E Beck. 2011. Looking beyond transfer models: finding other sources of power for student models. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 135–146.

[125] Yue Gong, Joseph E Beck, and Carolina Ruiz. 2012. Modeling multiple distributions of student performances to improve predictive accuracy. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 102–113.

[126] Andrés González-Nucamendi, Julieta Noguez, Luis Neri, and Víctor Robledo-Rella. 2015. Predictive models to enhance learning based on student profiles derived from cognitive and social constructs. In *2015 International Conference on Interactive Collaborative and Blended Learning (ICBL)*. IEEE, 5–12.

[127] Annagret Goold and Russell Rimmer. 2000. Factors affecting performance in first-year computing. *ACM SIGCSE Bulletin* 32, 2 (2000), 39–43.

[128] Lindsey Gouws, Karen Bradshaw, and Peter Wentworth. 2013. First year student performance in a test for computational thinking. In *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference*. ACM, 271–277.

[129] Geraldine Gray, Colm McGuinness, and Philip Owende. 2013. Investigating the efficacy of algorithmic student modelling in predicting students at risk of failing in tertiary education. In *Proceedings of the 6th International Conference on Educational Data Mining*.

[130] Foteini Grivokostopoulou, Isidoros Perikos, and Ioannis Hatzilygeroudis. 2014. Utilizing semantic web technologies and data mining techniques to analyze students learning and predict final performance. In *2014 International Conference on Teaching, Assessment and Learning (TALE)*. IEEE, 488–494.

[131] Frédéric Guay, Robert J Vallerand, and Céline Blanchard. 2000. On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion* 24, 3 (2000), 175–213.

[132] Jayati Gulati, Priya Bhardwaj, Bharti Suri, and Anu Singh Lather. 2016. A Study of Relationship between Performance, Temperament and Personality of a Software Programmer. *ACM SIGSOFT Software Engineering Notes* 41, 1 (2016), 1–5.

[133] Pratiyush Guleria, Niveditta Thakur, and Manu Sood. 2014. Predicting student performance using decision tree classifiers and information gain. In *2014 International Conference on Parallel, Distributed and Grid Computing (PDGC)*. IEEE, 126–129.

[134] Fransiskus Allan Gunawan et al. 2016. Fuzzy-mamdani inference system in predicting the corelation between learning method, discipline and motivation with student's achievement. In *2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. IEEE, 134–139.

[135] Necdet Güner, Abdulkadir Yaldır, Gürhan Gündüz, Emre Çomak, Sezai Tokat, and Serdar İplikçi. 2014. Predicting academically at-risk engineering students: A soft computing application. *Acta Polytechnica Hungarica* 11, 5 (2014), 199–216.

[136] Bo Guo, Rui Zhang, Guang Xu, Chuangming Shi, and Li Yang. 2015. Predicting students performance in educational data mining. In *2015 International Symposium on Educational Technology (ISET)*. IEEE, 125–128.

[137] Rami J Haddad and Youakim Kalaani. 2015. Can computational thinking predict academic performance?. In *Integrated STEM Education Conference (ISEC)*. IEEE, 225–229.

[138] P. Haden, D. Parsons, J. Gasson, and K. Wood. 2017. Student affect in CS1: Insights from an easy data collection tool. *ACM International Conference Proceeding Series* (2017), 40–49.

[139] Thomas Haig, Katrina Falkner, and Nickolas Falkner. 2013. Visualisation of learning management system usage for detecting student behaviour patterns. In *Proceedings of the Fifteenth Australasian Computing Education Conference*. Australian Computer Society, Inc., 107–115.

[140] Radhika R Halde, Arti Deshpande, and Anjali Mahajan. 2016. Psychology assisted prediction of academic performance using machine learning. In *International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 431–435.

[141] Matthew Hale, Noah Jorgenson, and Rose Gamble. 2011. Predicting individual performance in student project teams. In *2011 24th International Conference on Software Engineering Education and Training (CSEE&T)*. IEEE, 11–20.

[142] M. Han, M. Tong, M. Chen, J. Liu, and C. Liu. 2017. Application of Ensemble Algorithm in Students' Performance Prediction. *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)* (2017), 735–740.

[143] Wan Han, Ding Jun, Liu Kangxu, and Gao Xiaopeng. 2017. Investigating performance in a blended SPOC. In *2017 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE, 239–245.

[144] Qiang Hao, Ewan Wright, Brad Barnes, and Robert Maribe Branch. 2016. What are the most important predictors of computer science students' online help-seeking behaviors? *Computers in Human Behavior* 62 (2016), 467–474.

[145] Mohammed E Haque. 2012. Effect of Class Absenteeism on Grade Performance: A Probabilistic Neural Net(PNN) based GA trained model. In *ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[146] Brian Harrington, Shichong Peng, Xiaomeng Jin, and Minhaz Khan. 2018. Gender, confidence, and mark prediction in CS examinations. In *Proceedings of the 23rd Conference on Innovation and Technology in Computer Science Education*. ACM, 230–235.

[147] Tomas Hasbun, Alexandra Araya, and Jorge Villalon. 2016. Extracurricular activities as dropout prediction factors in higher education using decision trees. In *2016 16th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 242–244.

[148] Arto Hellas, Petri Ihantola, Andrew Petersen, Vangel V. Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, Chris Messom, and Soohyun Nam Liao. 2018. Taxonomizing Features and Methods for Identifying At-risk Students in Computing Courses. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*. ACM, New York, NY, USA, 364–365.

[149] Paul R Hernandez, P Schultz, Mica Estrada, Anna Woodcock, and Randie C Chance. 2013. Sustaining optimal motivation: A longitudinal analysis of interventions to broaden participation of underrepresented students in STEM. *Journal of Educational Psychology* 105, 1 (2013), 89.

[150] J Herold, TF Stahovich, and K Rawson. 2013. Using educational data mining to identify correlations between homework effort and performance. In *ASEE Annual Conference & Exposition*.

[151] Indriana Hidayah, Adhistya Erna Permanasari, and Ning Ratwastuti. 2013. Student classification for academic performance prediction using neuro fuzzy in a conventional classroom. In *2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*. IEEE, 221–225.

[152] Jeffrey L Hieb, Keith B Lyle, Patricia AS Ralston, and Julia Chariker. 2015. Predicting performance in a first engineering calculus course: Implications for interventions. *International Journal of Mathematical Education in Science and Technology* 46, 1 (2015), 40–55.

[153] R.M. Higashi, C.D. Schunn, and J.B. Flot. 2017. Different underlying motivations and abilities predict student versus teacher persistence in an online course. *Educational Technology Research and Development* 65, 6 (2017), 1471–1493.

[154] Chia-Lin Ho and Dianne Raubenheimer. 2011. Computing-related Self-efficacy: The Roles of Computational Capabilities, Gender, and Academic Performance. In *ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[155] Kurt Hornik and Bettina Gr"un. 2011. topicmodels: An R package for fitting topic models. *Journal of Statistical Software* 40, 13 (2011), 1–30.

[156] Roya Hosseini, Peter Brusilovsky, Michael Yudelson, and Arto Hellas. 2017. Stereotype modeling for Problem-Solving performance predictions in MOOCs and traditional courses. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 76–84.

[157] Larry L Howell and Carl D Sorenson. 2014. Are Undergraduate GPA and General GRE Percentiles Valid Predictors of Student Performance in an Engineering Graduate Program? *International Journal of Engineering Education* 30, 5 (2014), 1145–1165.

[158] Pei-Hsuan Hsieh, Jeremy R Sullivan, Daniel A Sass, and Norma S Guerra. 2012. Undergraduate engineering students' beliefs, coping strategies, and academic performance: An evaluation of theoretical models. *The Journal of Experimental Education* 80, 2 (2012), 196–218.

[159] Qian Hu, Agoritsa Polyzou, George Karypis, and Huzefa Rangwala. 2017. Enriching Course-Specific Regression Models with Content Features for Grade Prediction. In *2017 International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 504–513.

[160] Xiao Hu, Christy WL Cheong, Wenwen Ding, and Michelle Woo. 2017. A systematic review of studies on predicting student learning outcomes using learning analytics. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, 528–529.

[161] Shaobo Huang and Ning Fang. 2010. Regression models of predicting student academic performance in an engineering dynamics course. In *ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[162] Shaobo Huang and Ning Fang. 2012. Work in progress: Early prediction of students' academic performance in an introductory engineering course through different mathematical modeling techniques. In *Frontiers in Education Conference (FIE)*. IEEE, 1–2.

[163] Shaobo Huang and Ning Fang. 2013. Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education* 61 (2013), 133–145.

[164] Yun Huang, Julio D Guerra-Hollstein, and Peter Brusilovsky. 2016. Modeling Skill Combination Patterns for Deeper Knowledge Tracing. In *UMAP Extended Proceedings*.

[165] Yun Huang, Yanbo Xu, and Peter Brusilovsky. 2014. Doing more with less: Student modeling and performance prediction with reduced content models. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 338–349.

[166] Bryce E Hughes. 2018. Coming out in STEM: Factors affecting retention of sexual minority STEM students. *Science Advances* 4, 3 (2018).

[167] Norhayati Ibrahim, Steven A Freeman, and Mack C Shelley. 2011. Identifying predictors of academic success for part-time students at Polytechnic Institutes in Malaysia. *International Journal of Adult Vocational Education and Technology (IJAVET)* 2, 4 (2011), 1–16.

[168] Petri Ihantola, Arto Vihavainen, Alireza Ahadi, Matthew Butler, Jürgen Börstler, Stephen H Edwards, Essi Isohanni, Ari Korhonen, Andrew Petersen, Kelly Rivers, et al. 2015. Educational data mining and learning analytics in programming: Literature review and case studies. In *Proceedings of the 2015 ITiCSE on Working Group Reports*. ACM, 41–63.

[169] Shajith Ikbal, Ashay Tamhane, Bikram Sengupta, Malolan Chetlur, S Ghosh, and James Appleton. 2015. On early prediction of risks in academic performance for students. *IBM Journal of Research and Development* 59, 6 (2015), 5–1.

[170] PK Imbrie, JJ Lin, T Oladunni, and K Reid. 2007. Use of a neural network model and noncognitive measures to predict student matriculation in engineering. In *ASEE Annual Conference & Exposition*.

[171] Anoushka Jain, Tanupriya Choudhury, Parveen Mor, and A Sai Sabitha. 2017. Intellectual performance analysis of students by comparing various data mining techniques. In *2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*. IEEE, 57–63.

[172] David James and Clair Chilvers. 2001. Academic and non-academic predictors of success on the Nottingham undergraduate medical course 1970–1995. *Medical Education* 35, 11 (2001), 1056–1064.

[173] Jamie L Jensen, Shannon Neeley, Jordan B Hatch, and Ted Piorczynski. 2017. Learning scientific reasoning skills may be key to retention in science, technology, engineering, and mathematics. *Journal of College Student Retention: Research, Theory & Practice* 19, 2 (2017), 126–144.

[174] Suhang Jiang, Adrienne Williams, Katerina Schenke, Mark Warschauer, and Diane O'dowd. 2014. Predicting MOOC performance with week 1 behavior. In *Proceedings of the 7th International Conference on Educational Data Mining*.

[175] Srećko Joksimović, Dragan Gašević, Vitomir Kovanović, Bernhard E Riecke, and Marek Hatala. 2015. Social presence in online discussions as a process predictor of academic performance. *Journal of Computer Assisted Learning* 31, 6 (2015), 638–654.

[176] E. Jove, J.A. Lopez-Vazquez, M.I. Fernandez-Ibanez, J.-L. Casteleiro-Roca, and J.L. Calvo-Rolle. 2018. Hybrid intelligent system to predict the individual academic performance of engineering students. *International Journal of Engineering Education* 34, 3 (2018), 895–904.
[177] Shimin Kai, Juan Miguel L Andres, Luc Paquette, Ryan S Baker, Kati Molnar, Harriet Watkins, and Michael Moore. 2017. Predicting student retention from behavior in an online orientation course. In *Proceedings of the 10th International Conference on Educational Data Mining*.
[178] Kyehong Kang and Sujing Wang. 2018. Analyze and Predict Student Dropout from Online Programs. In *Proceedings of the 2nd International Conference on Compute and Data Analysis*. ACM, 6–12.
[179] Tanja Käser, Nicole R Hallinen, and Daniel L Schwartz. 2017. Modeling exploration strategies to predict student performance within a learning environment and beyond. In *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge*. ACM, 31–40.
[180] Jussi Kasurinen and Antti Knutas. 2018. Publication trends in gamification: a systematic mapping study. *Computer Science Review* 27 (2018), 33–44.
[181] P. Kaur and W. Singh. 2017. Implementation of student SGPA Prediction System (SSPS) using optimal selection of classification algorithm. *Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016* 2 (2017).
[182] Gregor Kennedy, Carleton Coffrin, Paula De Barba, and Linda Corrin. 2015. Predicting success: how learners' prior knowledge, skills and activities predict MOOC performance. In *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge*. ACM, 136–140.
[183] Fulya Damla Kentli and Yusuf Sahin. 2011. An SVM approach to predict student performance in manufacturing processes course. *Energy, Education, Science, and Technology Bulletin* 3, 4 (2011), 535–544.
[184] Mohd Nor Akmal Khalid, Umi Kalsom Yusof, and Looi Guo Xiang. 2016. Model student selection using fuzzy logic reasoning approach. In *2016 International Conference on Advanced Informatics: Concepts, Theory And Application (ICAICTA)*. IEEE, 1–6.
[185] Leena Khanna, Shailendra Narayan Singh, and Mansaf Alam. 2016. Educational data mining and its role in determining factors affecting students academic performance: A systematic review. In *2016 1st India International Conference on Information Processing (IICIP)*. IEEE, 1–7.
[186] Annisa Uswatun Khasanah et al. 2017. A Comparative Study to Predict Student's Performance Using Educational Data Mining Techniques. In *IOP Conference Series: Materials Science and Engineering*, Vol. 215. IOP Publishing, 012036.
[187] Suin Kim, Jae Won Kim, Jungkook Park, and Alice Oh. 2016. Elivate: A Real-Time Assistant for Students and Lecturers as Part of an Online CS Education Platform. In *Proceedings of the Third Conference on Learning @ Scale*. ACM, 337–338.
[188] Eranki LN Kiran and Kannan M Moudgalya. 2015. Evaluation of programming competency using student error patterns. In *2015 International Conference on Learning and Teaching in Computing and Engineering*. IEEE, 34–41.
[189] KV Krishna Kishore, S Venkatramaphanikumar, and Sura Alekhya. 2014. Prediction of student academic progression: A case study on Vignan University. In *2014 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 1–6.
[190] Barbara Ann Kitchenham, David Budgen, and Pearl Brereton. 2015. *Evidence-based software engineering and systematic reviews*. Vol. 4. CRC Press.
[191] Antti Knutas, Arash Hajikhani, Juho Salminen, Jouni Ikonen, and Jari Porras. 2015. Cloud-based bibliometric analysis service for systematic mapping studies. In *Proceedings of the 16th International Conference on Computer Systems and Technologies*. ACM, 184–191.
[192] Joseph A Konstan, JD Walker, D Christopher Brooks, Keith Brown, and Michael D Ekstrand. 2015. Teaching recommender systems at large scale: evaluation and lessons learned from a hybrid MOOC. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 2 (2015), 10.
[193] Irena Koprinska, Joshua Stretton, and Kalina Yacef. 2015. Predicting student performance from multiple data sources. In *International Conference on Artificial Intelligence in Education*. Springer, 678–681.
[194] Maria Koutina and Katia Lida Kermanidis. 2011. Predicting postgraduate students' performance using machine learning techniques. In *Artificial Intelligence Applications and Innovations*. Springer, 159–168.
[195] M. Koç. 2017. Learning analytics of student participation and achievement in online distance education: A structural equation modeling. *Kuram ve Uygulamada Egitim Bilimleri* 17, 6 (2017), 1893–1910.
[196] Nicole Kronberger and Ilona Horwath. 2013. The ironic costs of performing well: Grades differentially predict male and female dropout from engineering. *Basic and Applied Social Psychology* 35, 6 (2013), 534–546.
[197] M. Kuehn, J. Estad, J. Straub, T. Stokke, and S. Kerlin. 2017. An expert system for the prediction of student performance in an initial computer science course. *IEEE International Conference on Electro Information Technology* (2017), 1–6.
[198] Mukesh Kumar, AJ Singh, and Disha Handa. 2017. Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques. *International Journal of Education and Management Engineering* (2017).
[199] S Chaitanya Kumar, E Deepak Chowdary, S Venkatramaphanikumar, and K V Krishna Kishore. 2016. M5P model tree in predicting student performance: A case study. In *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, Bangalore, India, 1103–1107.
[200] Lynn Lambert. 2015. Factors that predict success in CS1. *Journal of Computing Sciences in Colleges* 31, 2 (2015), 165–171.
[201] Charlotte Larmuseau, Jan Elen, and Fien Depaepe. 2018. The influence of students' cognitive and motivational characteristics on students' use of a 4C/ID-based online learning environment and their learning gain. In *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. ACM, 171–180.
[202] John Lee and Chak Yan Yeung. 2018. Automatic prediction of vocabulary knowledge for learners of Chinese as a foreign language. In *2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP)*. IEEE, 1–4.
[203] Un Jung Lee, Gena C Sbeglia, Minsu Ha, Stephen J Finch, and Ross H Nehm. 2015. Clicker score trajectories and concept inventory scores as predictors for early warning systems for large STEM Classes. *Journal of Science Education and Technology* 24, 6 (2015), 848–860.
[204] Cheng Lei and Kin Fun Li. 2015. Academic performance predictors. In *2015 29th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*. IEEE, 577–581.
[205] Juho Leinonen, Krista Longi, Arto Klami, and Arto Vihavainen. 2016. Automatic inference of programming performance and experience from typing patterns. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 132–137.
[206] Robert W Lent, Matthew J Miller, Paige E Smith, Bevlee A Watford, Robert H Lim, and Kayi Hui. 2016. Social cognitive predictors of academic persistence and performance in engineering: Applicability across gender and race/ethnicity. *Journal of Vocational Behavior* 94 (2016), 79–88.
[207] Leo Leppänen, Juho Leinonen, Petri Ihantola, and Arto Hellas. 2017. Predicting Academic Success Based on Learning Material Usage. In *Proceedings of the 18th Annual Conference on Information Technology Education*. ACM, New York, NY, USA, 13–18.
[208] Kin Fun Li, David Rusk, and Fred Song. 2013. Predicting student academic performance. In *2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)*. IEEE, 27–33.
[209] Xiu Li, Lulu Xie, and Huimin Wang. 2016. Grade prediction in MOOCs. In *2016 International Conference on Computational Science and Engineering (CSE) and International Conference on Embedded and Ubiquitous Computing (EUC) and 15th International Symposium on Distributed Computing and Applications for Business Engineering (DCABES)*. IEEE, 386–392.
[210] Zhenpeng Li, Changjing Shang, and Qiang Shen. 2016. Fuzzy-clustering embedded regression for predicting student academic performance. In *2016 International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 344–351.
[211] Defu Lian, Yuyang Ye, Wenya Zhu, Qi Liu, Xing Xie, and Hui Xiong. 2016. Mutual reinforcement of academic performance prediction and library book recommendation. In *2016 16th International Conference on Data Mining (ICDM)*. IEEE, 1023–1028.
[212] Jiajun Liang, Chao Li, and Li Zheng. 2016. Machine learning application in MOOCs: Dropout prediction. In *2016 11th International Conference on Computer Science & Education (ICCSE)*. IEEE, 52–57.
[213] Jiajun Liang, Jian Yang, Yongji Wu, Chao Li, and Li Zheng. 2016. Big data application in education: dropout prediction in edx MOOCs. In *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*. IEEE, 440–443.
[214] Soohyun Nam Liao, Daniel Zingaro, Michael A Laurenzano, William G Griswold, and Leo Porter. 2016. Lightweight, early identification of at-risk CS1 students. In *Proceedings of the 2016 International Computing Education Research Conference*. ACM, 123–131.
[215] Kittinan Limsathitwong, Kanda Tiwatthanont, and Tanasin Yatsungnoen. 2018. Dropout prediction system to reduce discontinue study rate of information technology students. In *2018 5th International Conference on Business and Industrial Research (ICBIR)*. IEEE, 110–114.
[216] Chen Lin, Shitian Shen, and Min Chi. 2016. Incorporating Student Response Time and Tutor Instructional Interventions into Student Modeling. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. ACM, 157–161.
[217] Che-Cheng Lin and Chiung-Hui Chiu. 2013. Correlation between course tracking variables and academic performance in blended online courses. In *2013 13th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 184–188.
[218] Lisa Linnenbrink-Garcia, Tony Perez, Michael M Barger, Stephanie V Wormington, Elizabeth Godin, Kate E Snyder, Kristy Robinson, Abdhi Sarkar, Laura S Richman, and Rochelle Schwartz-Bloom. 2018. Repairing the leaky pipeline: A motivationally supportive intervention to enhance persistence in undergraduate science pathways. *Contemporary Educational Psychology* 53 (2018), 181–195.
[219] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to program: Gender differences and interactive effects of students' motivation,

goals, and self-efficacy on performance. In *Proceedings of the 2016 International Computing Education Research Conference*. ACM, 211–220.

[220] Elizabeth Litzler and Jacob Young. 2012. Understanding the risk of attrition in undergraduate engineering: Results from the project to assess climate in engineering. *Journal of Engineering Education* 101, 2 (2012), 319–345.

[221] Ronan A Lopes, Luiz AL Rodrigues, and Jacques D Brancher. 2017. Predicting master's applicants performance using KDD techniques. In *2017 12th Iberian Information Systems and Technologies (CISTI)*. IEEE, 1–6.

[222] Manuel Ignacio Lopez, JM Luna, C Romero, and S Ventura. 2012. Classification via clustering for predicting final marks based on student participation in forums. (2012).

[223] Jingyi Luo, Shaymaa E Sorour, Kazumasa Goda, and Tsunenori Mine. 2015. Predicting Student Grade Based on Free-Style Comments Using Word2Vec and ANN by Considering Prediction Results Obtained in Consecutive Lessons.. In *Proceedings of the 8th International Conference on Educational Data Mining*.

[224] Ling Luo, Irena Koprinska, and Wei Liu. 2015. Discrimination-Aware Classifiers for Student Performance Prediction.. In *Proceedings of the 8th International Conference on Educational Data Mining*.

[225] Cheng Ma, Baofeng Yao, Fang Ge, Yurong Pan, and Youqiang Guo. 2017. Improving Prediction of Student Performance based on Multiple Feature Selection Approaches. In *Proceedings of the 2017 International Conference on E-Education, E-Business and E-Technology*. ACM, 36–41.

[226] Xiaofeng Ma and Zhurong Zhou. 2018. Student pass rates prediction using optimized support vector machine and decision tree. In *Computing and Communication Workshop and Conference (CCWC), 2018 IEEE 8th Annual*. IEEE, 209–215.

[227] Kartika Maharani, Teguh Bharata Adji, Noor Akhmad Setiawan, and Indriana Hidayah. 2015. Comparison analysis of data mining methodology and student performance improvement influence factors in small data set. In *2015 International Conference on Science in Information Technology (ICSITech)*. IEEE, 169–174.

[228] Sapan H Mankad. 2016. Predicting learning behaviour of students: Strategies for making the course journey interesting. In *2016 10th International Conference on Intelligent Systems and Control (ISCO)*. IEEE, 1–6.

[229] J James Manoharan, Dr S Hari Ganesh, and M Lovelin Ponn Felcia. 2014. Discovering Student's Academic Performance Based on GPA using k-Means Clustering Algorithm. In *IEEE World Congress on Computing and Communication Technology*.

[230] Andrew J Martin. 2001. The Student Motivation Scale: A tool for measuring and enhancing motivation. *Journal of Psychologists and Counsellors in Schools* 11 (2001), 1–20.

[231] Lebogang Mashiloane and Mike Mchunu. 2013. Mining for marks: a comparison of classification algorithms when predicting academic performance to identify "students at risk". In *Mining Intelligence and Knowledge Exploration*. Springer, 541–552.

[232] Cindi Mason, Janet Twomey, David Wright, and Lawrence Whitman. 2018. Predicting engineering student attrition risk using a probabilistic neural network and comparing results with a backpropagation neural network and logistic regression. *Research in Higher Education* 59, 3 (2018), 382–400.

[233] John M Mativo and Shaobo Huang. 2014. Prediction of students' academic performance: Adapt a methodology of predictive modeling for a small sample size. In *Frontiers in Education Conference (FIE)*. IEEE, 1–3.

[234] M Mayilvaganan and D Kalpanadevi. 2014. Comparison of classification techniques for predicting the performance of students academic environment. In *Communication and Network Technologies (ICCNT)*. IEEE, 113–118.

[235] Joseph P Mazer. 2013. Validity of the student interest and engagement scales: Associations with student learning outcomes. *Communication Studies* 64, 2 (2013), 125–140.

[236] WJ McNamara and JL Hughes. 1961. A review of research on the selection of computer programmers. *Personnel Psychology* 14, 1 (1961), 39–51.

[237] Yannick Meier, Jie Xu, Onur Atan, and Mihaela Van Der Schaar. 2015. Personalized grade prediction: A data mining approach. In *2015 International Conference on Data Mining (ICDM)*. IEEE, 907–912.

[238] Yannick Meier, Jie Xu, Onur Atan, and Mihaela Van der Schaar. 2016. Predicting grades. *IEEE Transactions on Signal Processing* 64, 4 (2016), 959–972.

[239] Juan A Méndez and Evelio J González. 2013. A control system proposal for engineering education. *Computers & Education* 68 (2013), 266–274.

[240] Vilma Mesa, Ozan Jaquette, and Cynthia J Finelli. 2009. Measuring the impact of an individual course on students' success. *Journal of Engineering Education* 98, 4 (2009), 349–359.

[241] Mvurya Mgala and Audrey Mbogho. 2015. Data-driven intervention-level prediction modeling for academic performance. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ACM, 2.

[242] V. Mhetre and M. Nagar. 2018. Classification based data mining algorithms to predict slow, average and fast learners in educational system using WEKA. *Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017* 2018-January (2018), 475–479.

[243] David Mimno, Hanna M. Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 262–272.

[244] Youngkyoung Min, Guili Zhang, Russell A Long, Timothy J Anderson, and Matthew W Ohland. 2011. Nonparametric survival analysis of the loss rate of undergraduate engineering students. *Journal of Engineering Education* 100, 2 (2011), 349–373.

[245] Tripti Mishra, Dharminder Kumar, and Sangeeta Gupta. 2014. Mining students' data for prediction performance. In *Advanced Computing & Communication Technologies (ACCT)*. IEEE, 255–262.

[246] Margaret Montague, MM Reynolds, and MF Washburn. 1918. A Further Study of Freshmen. *The American Journal of Psychology* 29, 3 (1918), 327–330.

[247] F. Moradi and P. Amiripour. 2017. The prediction of the students' academic underachievement in mathematics using the DEA model: A developing country case study. *European Journal of Contemporary Education* 6, 3 (2017), 432–447.

[248] S. Morsy and G. Karypis. 2017. Cumulative knowledge-based regression models for next-term grade prediction. *Proceedings of the 17th SIAM International Conference on Data Mining (SDM)* (2017), 552–560.

[249] Rachel Mosier, John Reck, Heather Yates, and Carisa Ramming. 2017. Standardized Tests as a Predictor for Success in Construction, Architecture, and Architectural Engineering Programs. In *2017 ASEE Annual Conference & Exposition Proceedings*. ASEE Conferences, Columbus, Ohio.

[250] Jonathan P Munson and Joshua P Zitovsky. 2018. Models for Early Identification of Struggling Novice Programmers. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. ACM, 699–704.

[251] Juliet Mutahi, Andrew Kinai, Nelson Bore, Abdigani Diriye, and Komminist Weldemariam. 2017. Studying engagement and performance with learning technology in an African classroom. In *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge*. ACM, 148–152.

[252] SM Muthukrishnan, MK Govindasamy, and MN Mustapha. 2017. Systematic mapping review on student's performance analysis using big data predictive model. *Journal of Fundamental and Applied Sciences* 9, 4S (2017), 730–758.

[253] Kew Si Na and Zaidatun Tasir. 2017. Identifying at-risk students in online learning by analysing learning behaviour: A systematic review. In *2017 International Conference on Big Data and Analytics (ICBDA)*. IEEE, 118–123.

[254] Àngela Nebot, Francisco Mugica, and Félix Castro. 2010. Fuzzy predictive models to help teachers in e-learning courses. In *2010 International Joint Conference on Neural Networks (IJCNN)*. IEEE.

[255] Àngela Nebot, Francisco Mugica, Félix Castro, and Jesús Acosta. 2010. Genetic fuzzy system for predictive and decision support modelling in e-learning. In *2010 International Conference on Fuzzy Systems (FUZZ)*. IEEE, 1–8.

[256] Prema Nedungadi and M. S. Remya. 2014. Predicting students' performance on intelligent tutoring system - Personalized clustered BKT (PC-BKT) model. In *Frontiers in Education Conference (FIE)*. IEEE, Madrid, Spain, 1–6.

[257] Celia Gonzalez Nespereira, Ana Fernández Vilas, and Rebeca P Díaz Redondo. 2015. Am I failing this course?: risk prediction using e-learning data. In *Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality*. ACM, 271–276.

[258] U. Ninrutsirikun, B. Watanapa, C. Arpnikanondt, and N. Phothikit. 2017. Effect of the Multiple Intelligences in multiclass predictive model of computer programming course achievement. *IEEE Region 10 Annual International Conference, Proceedings/TENCON* (2017), 297–300.

[259] Julieta Noguez, Luis Neri, Andres González-Nucamendi, and Víctor Robledo-Rella. 2016. Characteristics of self-regulation of engineering students to predict and improve their academic performance. In *Frontiers in Education Conference (FIE)*. IEEE, 1–8.

[260] Xavier Ochoa. 2016. Adaptive multilevel clustering model for the prediction of academic risk. In *Latin American Conference on Learning Objects and Technology (LACLO)*. IEEE, 1–8.

[261] S. Oeda and G. Hashimoto. 2017. Log-Data Clustering Analysis for Dropout Prediction in Beginner Programming Classes. *Procedia Computer Science* 112 (2017), 614–621.

[262] Viola Osborn and Philip Turner. 2002. Identifying at-risk students in LIS distributed learning courses. *Journal of Education for Library and Information Science* (2002), 205–213.

[263] Rozita Jamili Oskouei, Mohsen Askari, and Phani Rajendra Prasad Sajja. 2013. Perceived Internet Usage Behaviours as Predictor to Outlier Detection in Students' Communities in Academic Environments. *Indian Journal of Science and Technology* 6, 7 (2013), 4923–4935.

[264] Korinn Ostrow, Christopher Donnelly, Seth Adjei, and Neil Heffernan. 2015. Improving student modeling through partial credit and problem difficulty. In *Proceedings of the Second Conference on Learning @ Scale*. ACM, 11–20.

[265] Aini Nazura Paimin, Maizam Alias, RG Hadgraft, Julianna Kaya Prpic, et al. 2013. Factors affecting study performance of engineering undergraduates: Case studies of malaysia and australia. In *Research in Engineering Education Symposium, REES 2013*. 180–186.

[266] Aini Nazura Paimin, Roger G Hadgraft, J Kaya Prpic, and Maizam Alias. 2016. An application of the theory of reasoned action: Assessing success factors of engineering students. *International Journal of Engineering Education* (2016).

[267] Stuart Palmer. 2013. Modelling engineering student academic performance using academic analytics. *International Journal of Engineering Education* 29, 1 (2013), 132–138.

[268] Mrinal Pandey and S Taruna. 2018. An Ensemble-Based Decision Support System for the Students' Academic Performance Prediction. In *ICT Based Innovations*. Springer, 163–169.

[269] Zacharoula Papamitsiou and Anastasios A Economides. 2014. Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society* 17, 4 (2014).

[270] Zacharoula Papamitsiou, Anastasios A Economides, Ilias O Pappas, and Michail N Giannakos. 2018. Explaining learning performance using response-time, self-regulation and satisfaction from content: an fsQCA approach. In *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. ACM, 181–190.

[271] Zacharoula Papamitsiou, Eirini Karapistoli, and Anastasios A Economides. 2016. Applying classification techniques on temporal trace data for shaping student behavior models. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 299–303.

[272] A. Pardo, F. Han, and R.A. Ellis. 2017. Combining University student self-regulated learning indicators and engagement with online learning events to Predict Academic Performance. *IEEE Transactions on Learning Technologies* 10, 1 (2017), 82–92.

[273] Abelardo Pardo, Feifei Han, and Robert A Ellis. 2016. Exploring the relation between self-regulation, online activities, and academic performance: A case study. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 422–429.

[274] Abelardo Pardo, Negin Mirriahi, Roberto Martinez-Maldonado, Jelena Jovanovic, Shane Dawson, and Dragan Gašević. 2016. Generating actionable predictive models of academic performance. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 474–478.

[275] Zachary A Pardos, Ryan SJD Baker, Maria OCZ San Pedro, Sujith M Gowda, and Supreeth M Gowda. 2013. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International Conference on Learning Analytics & Knowledge*. ACM, 117–124.

[276] Zachary A Pardos, Sujith M Gowda, Ryan SJd Baker, and Neil T Heffernan. 2012. The sum is greater than the parts: ensembling models of student knowledge in educational software. *ACM SIGKDD Explorations Newsletter* 13, 2 (2012), 37–44.

[277] Zachary A Pardos, Sujith M Gowda, Ryan Shaun Joazeiro de Baker, and Neil T Heffernan. 2011. Ensembling Predictions of Student Post-Test Scores for an Intelligent Tutoring System.. In *Proceedings of the 4th International Conference on Educational Data Mining*. 189–198.

[278] Priyanka Anandrao Patil and RV Mane. 2014. Prediction of Students Performance Using Frequent Pattern Tree. In *2014 International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 1078–1082.

[279] A Borrego Patrick. 2018. Predicting Persistence in Engineering through an Engineering Identity Scale. *International Journal of Engineering Education* 34, 2A (2018).

[280] Dimple V Paul, Chitra Nayagam, and Jyoti D Pawar. 2016. Modeling Academic Performance using Subspace Clustering Algorithm. In *2016 Eighth International Conference on Technology for Education (T4E)*. IEEE, 254–255.

[281] Wei Peng, Rabindra A Ratan, and Laeeq Khan. 2015. Ebook uses and class performance in a college course. In *2015 48th Hawaii International Conference on System Sciences (HICSS)*. IEEE, 63–71.

[282] Birgit Penzenstadler, Ankita Raturi, Debra Richardson, Coral Calero, Henning Femmer, and Xavier Franch. 2014. Systematic mapping study on software engineering for sustainability (SE4S). In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. ACM, 14.

[283] Kai Petersen, Sairam Vakkalanka, and Ludwik Kuzniarz. 2015. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology* 64 (2015), 1–18.

[284] Chris Piech, Mehran Sahami, Daphne Koller, Steve Cooper, and Paulo Blikstein. 2012. Modeling How Students Learn to Program. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*. ACM, New York, NY, USA, 153–160.

[285] Norman Poh and Ian Smythe. 2014. To what extend can we predict students' performance? A case study in colleges in South Africa. In *2014 Symposium on Computational Intelligence and Data Mining (CIDM)*. IEEE, 416–421.

[286] Elvira Popescu, Mihai Dascalu, Alex Becheru, Scott Crossley, and Stefan Trausan-Matu. 2016. Predicting Student Performance and Differences in Learning Styles Based on Textual Complexity Indices Applied on Blog and Microblog Posts: A Preliminary Study. In *2016 16th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 184–188.

[287] Leo Porter, Daniel Zingaro, and Raymond Lister. 2014. Predicting student success using fine grain clicker data. In *Proceedings of the Tenth International Computing Education Research Conference*. ACM, 51–58.

[288] Wanthanee Prachuabsupakij and Nuanwan Soonthornphisaj. 2014. Hybrid sampling for multiclass imbalanced problem: Case study of students' performance prediction. In *2014 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. IEEE, 321–326.

[289] Anjana Pradeep, Smija Das, and Jubilant J Kizhekkethottam. 2015. Students dropout factor prediction using EDM techniques. In *2015 International Conference on Soft-Computing and Networks Security (ICSNS)*. IEEE, 1–7.

[290] Raymond Ptucha and Andreas Savakis. 2012. How connections matter: factors affecting student performance in stem disciplines. In *Integrated STEM Education Conference (ISEC)*. IEEE, 1–5.

[291] Utomo Pujianto, Erwina Nurul Azizah, and Ayuningtyas Suci Damayanti. 2017. Naive Bayes using to predict students' academic performance at faculty of literature. In *2017 5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*. IEEE, 163–169.

[292] S.K. Pushpa, T.N. Manjunath, T.V. Mrunal, A. Singh, and C. Suhas. 2018. Class result prediction using machine learning. *Proceedings of the 2017 International Conference On Smart Technology for Smart Nation, SmartTechCon 2017* (2018), 1208–1212.

[293] Jiezhong Qiu, Jie Tang, Tracy Xiao Liu, Jie Gong, Chenhui Zhang, Qian Zhang, and Yufei Xue. 2016. Modeling and predicting learning behavior in MOOCs. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. ACM, 93–102.

[294] Keith Quille and Susan Bergin. 2018. Programming: predicting student success early in CS1. a re-validation and replication study. In *Proceedings of the 23rd Conference on Innovation and Technology in Computer Science Education*. ACM, 15–20.

[295] Nachirat Rachburee and Wattana Punlumjeak. 2015. A comparison of feature selection approach between greedy, IG-ratio, Chi-square, and mRMR in educational mining. In *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)*. IEEE, 420–424.

[296] J. Raigoza. 2017. A study of students' progress through introductory Computer Science programming courses. *Frontiers in Education Conference (FIE)* (2017), 1–7.

[297] D.R. Raman and A.L. Kaleita. 2017. Enhancing student success by combining pre-enrollment risk prediction with academic analytics data. *ASEE Annual Conference & Exposition* (2017).

[298] L. Ramanathan, S. Dhanda, and S. Kumar D. 2013. Predicting Students' Performance using Modified ID3 Algorithm. *International Journal of Engineering and Technology (IJET)* 5, 3 (June 2013), 2491–2497.

[299] Nichole Ramirez, Joyce Main, and Matthew Ohland. 2015. Academic Outcomes of Cooperative Education Participation. In *2015 ASEE Annual Conference and Exposition Proceedings*. ASEE Conferences, Seattle, Washington, 26.140.1–26.140.13.

[300] Shiwani Rana and Roopali Garg. 2016. Application of Hierarchical Clustering Algorithm to Evaluate Students Performance of an Institute. In *2016 2nd International Conference on Computational Intelligence & Communication Technology (CICT)*. IEEE, 692–697.

[301] S. Rana and R. Garg. 2017. Prediction of students performance of an institute using ClassificationViaClustering and ClassificationViaRegression. *Advances in Intelligent Systems and Computing* 508 (2017), 333–343.

[302] A Ravishankar Rao. 2017. A novel STEAM approach: Using cinematic meditation exercises to motivate students and predict performance in an engineering class. In *Integrated STEM Education Conference (ISEC)*. IEEE, 64–70.

[303] A Ravishankar Rao. 2018. Simultaneously educating students about the impact of cell phone usage while creating a metric to predict their performance. In *Integrated STEM Education Conference (ISEC)*. IEEE, 143–148.

[304] Raisul Islam Rashu, Naheena Haq, and Rashedur M Rahman. 2014. Data mining approaches to predict final grade by overcoming class imbalance problem. In *17th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 14–19.

[305] G. Raura, F. Efraín, A. Ponce, and O. Dieste. 2017. Experience does not predict performance: The case of the students-academic levels. *2017 Ibero-American Conference on Software Engineering* (2017), 57–70.

[306] Kenneth Reid and PK Imbrie. 2008. Noncognitive characteristics of incoming engineering students compared to incoming engineering technology students: A preliminary examination. In *ASEE Annual Conference & Exposition*.

[307] Rachelle Reisberg, Joseph A Raelin, Margaret B Bailey, Jerry Carl Hamann, David L Whitman, and Leslie K Pendleton. 2011. The Effect of Contextual Support in the First Year on Self-Efficacy in Undergraduate Engineering Programs. *ASEE Annual Conference & Exposition* (2011), 14.

[308] Zhiyun Ren, Xia Ning, and Huzefa Rangwala. 2017. Grade Prediction with Temporal Course-wise Influence. (2017).

[309] Zhiyun Ren, Huzefa Rangwala, and Aditya Johri. 2016. Predicting performance on MOOC assessments using multi-regression models. In *Proceedings of the 9th International Conference on Educational Data Mining*.

[310] Alexander Repenning and Ashok Basawapatna. 2016. Drops and Kinks: Modeling the Retention of Flow for Hour of Code Style Tutorials. In *Proceedings of the 11th Workshop in Primary and Secondary Computing Education*. ACM, 76–79.

[311] Nicholas Rhodes, Matthew Ung, Alexander Zundel, Jim Herold, and Thomas Stahovich. 2013. Using a Lexical Analysis of Student's Self-Explanation to Predict Course Performance. In *Proceedings of the 6th International Conference on*

Educational Data Mining.
[312] Jeff Ringenberg, Marcial Lapp, Apoorva Bansal, and Parth Shah. 2011. The Programming Performance Prophecies: Predicting Student Achievement in a First-Year Introductory Programming Course. In *ASEE Annual Conference & Exposition.* American Society for Engineering Education.
[313] Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley. 2017. *stm: R Package for Structural Topic Models.*
[314] Tim Rogers, Cassandra Colvin, and Belinda Chiera. 2014. Modest analytics: using the index method to identify students at risk of failure. In *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge.* ACM, 118–122.
[315] Samuel L Rohr. 2012. How well does the SAT and GPA predict the retention of science, technology, engineering, mathematics, and business students. *Journal of College Student Retention: Research, Theory & Practice* 14, 2 (2012), 195–208.
[316] Cristobal Romero, Pedro G Espejo, Amelia Zafra, Jose Raul Romero, and Sebastian Ventura. 2013. Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education* 21, 1 (2013), 135–146.
[317] Cristóbal Romero, Manuel-Ignacio López, Jose-María Luna, and Sebastián Ventura. 2013. Predicting students' final performance from participation in on-line discussion forums. *Computers & Education* 68 (2013), 458–472.
[318] Cristóbal Romero and Sebastián Ventura. 2010. Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40, 6 (2010), 601–618.
[319] S. Rovira, E. Puertas, and L. Igual. 2017. Data-driven system to predict academic grades and dropout. *PLoS ONE* 12, 2 (2017).
[320] Sagardeep Roy and Anchal Garg. 2017. Analyzing performance of students by using data mining techniques a literature survey. In *2017 4th Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON).* IEEE, 130–133.
[321] Sagardeep Roy and Anchal Garg. 2017. Predicting academic performance of student using classification techniques. In *2017 4th Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON).* IEEE, 568–572.
[322] Sandra Milena Merchan Rubiano and Jorge Alberto Duarte Garcia. 2015. Formulation of a predictive model for academic performance based on students' academic and demographic data. In *Frontiers in Education Conference (FIE).* IEEE, 1–7.
[323] Reynold A Rustia, Ma Melanie A Cruz, Michael Angelo P Burac, and Thelma D Palaoag. 2018. Predicting Student's Board Examination Performance using Classification Algorithms. In *Proceedings of the 2018 7th International Conference on Software and Computer Applications.* ACM, 233–237.
[324] Chew Li Sa, Emmy Dahliana Hossain, Mohammad bin Hossin, et al. 2014. Student performance analysis system (SPAS). In *2014 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M).* IEEE, 1–6.
[325] Syafawati Ab. Saad, Nor Hizamiyani Abdul Azziz, Siti Aisyah Zakaria, and Nornadia Mohd Yazid. 2015. Performance of engineering undergraduate students in Mathematics: A Case Study In UniMAP. In *American Institute of Physics*, Vol. 1691.
[326] S. Sadati and N.A. Libre. 2017. Development of an early alert system to predict students at risk of failing based on their early course activities. *ASEE Annual Conference & Exposition.*
[327] William E Sadler. 1997. Factors Affecting Retention Behavior: A Model To Predict At-Risk Students. In *37th Annual Forum of the Association for Institutional Research.* AIR, Orlando, FL, 23.
[328] Medha Sagar, Arushi Gupta, and Rishabh Kaushal. 2016. Performance prediction and behavioral analysis of student programming ability. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI).* IEEE, 1039–1045.
[329] Farhana Sarker, Thanassis Tiropanis, and Hugh C Davis. 2013. Exploring student predictive model that relies on institutional databases and open data instead of traditional questionnaires. In *Proceedings of the 22nd International Conference on World Wide Web.* ACM, 413–418.
[330] Farhana Sarker, Thanassis Tiropanis, and Hugh C Davis. 2014. Linked data, data mining and external open data for better prediction of at-risk students. In *2014 International Conference on Control, Decision and Information Technologies (CoDIT).* IEEE, 652–657.
[331] Patrick D Schalk, David P Wick, Peter R Turner, and Michael W Ramsdell. 2011. Predictive assessment of student performance for early strategic guidance. In *Frontiers in Education Conference (FIE).* IEEE, S2H–1.
[332] Mark Schar. 2016. Connecting for Success: The Impact of Student-to-Other Closeness on Performance in Large Scale Engineering Classes. In *ASEE Annual Conference & Exposition.*
[333] Otto Seppälä, Petri Ihantola, Essi Isohanni, Juha Sorva, and Arto Vihavainen. 2015. Do we know how difficult the rainfall problem is?. In *Proceedings of the 15th Koli Calling Conference on Computing Education Research.* ACM, 87–96.
[334] Sami Shaban and Michelle McLean. 2011. Predicting performance at medical school: can we identify at-risk students? *Advances in Medical Education and Practice* 2 (2011), 139.
[335] Amirah Mohamed Shahiri, Wahidah Husain, et al. 2015. A review on predicting student's performance using data mining techniques. *Procedia Computer Science* 72 (2015), 414–422.
[336] Ashkan Sharabiani, Fazle Karim, Anooshiravan Sharabiani, Mariya Atanasov, and Houshang Darabi. 2014. An enhanced bayesian network model for prediction of students' academic performance in engineering programs. In *Global Engineering Education Conference (EDUCON).* IEEE, 832–837.
[337] Duane F Shell, Leen-Kiat Soh, Abraham E Flanigan, and Markeya S Peteranetz. 2016. Students' initial course motivation and their achievement and retention in college CS1 courses. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education.* ACM, 639–644.
[338] Carson Sievert and Kenneth E. Shirley. 2014. LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces.* 63–70.
[339] Md Fahim Sikder, Md Jamal Uddin, and Sajal Halder. 2016. Predicting students yearly performance using neural network: A case study of BSMRSTU. In *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV).* IEEE, 524–529.
[340] Simon, Sally Fincher, Anthony Robins, Bob Baker, Ilona Box, Quintin Cutts, Michael de Raadt, Patricia Haden, John Hamer, Margaret Hamilton, Raymond Lister, Marian Petre, Ken Sutton, Denise Tolhurst, and Jodi Tutty. 2006. Predictors of Success in a First Programming Course. In *Proceedings of the 8th Australasian Conference on Computing Education.* Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 189–196.
[341] Larry D Singell and Glen R Waddell. 2010. Modeling retention at a large public university: Can at-risk students be identified early enough to treat? *Research in Higher Education* 51, 6 (2010), 546–572.
[342] Mamta Singh, Jyoti Singh, and Arpana Rawal. 2014. Feature extraction model to identify at–risk level of students in academia. In *2014 International Conference on Information Technology (ICIT).* IEEE, 221–227.
[343] M. Sivasakthi. 2018. Classification and prediction based data mining algorithms to predict students' introductory programming performance. *Proceedings of the International Conference on Inventive Computing and Informatics (ICICI)* (2018), 346–350.
[344] Ahmad Slim, Gregory L Heileman, Jarred Kozlick, and Chaouki T Abdallah. 2014. Employing markov networks on curriculum graphs to predict student performance. In *2014 13th International Conference on Machine Learning and Applications (ICMLA).* IEEE, 415–418.
[345] Ahmad Slim, Gregory L Heileman, Jarred Kozlick, and Chaouki T Abdallah. 2014. Predicting student success based on prior performance. In *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM).* IEEE, 410–415.
[346] Marisol Solis-Foronda. 2017. Predictors of Licensure Examination for Teachers (LET) Performance: A Mediation Analysis. In *Proceedings of the International Conference on Digital Technology in Education.* ACM, 74–78.
[347] E. Soloway. 1986. Learning to Program = Learning to Construct Mechanisms and Explanations. *Commun. ACM* 29, 9 (Sept. 1986), 850–858.
[348] Sheryl Sorby, Edmund Nevin, Eileen Mageean, Sarah Sheridan, and Avril Behan. 2014. Initial Investigation into Spatial Skills as Predictors of Success in First-year STEM Programmes. In *2014 42nd Conference European Society for Engineering Education (SEFI).* Birmingham, UK, 9.
[349] Shaymaa E Sorour, Kazumasa Goda, and Tsunenori Mine. 2015. Estimation of student performance by considering consecutive lessons. In *2015 IIAI 4th International Congress on Advanced Applied Informatics (IIAI-AAI).* IEEE, 121–126.
[350] Shaymaa E Sorour, Jingyi Luo, Kazumasa Goda, and Tsunenori Mine. 2015. Correlation of grade prediction performance with characteristics of lesson subject. In *2015 15th International Conference on Advanced Learning Technologies (ICALT).* IEEE, 247–249.
[351] Shaymaa E Sorour and Tsunenori Mine. 2016. Building an Interpretable Model of Predicting Student Performance Using Comment Data Mining. In *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI).* IEEE, 285–291.
[352] Shaymaa E Sorour and Tsunenori Mine. 2016. Exploring students' learning attributes in consecutive lessons to improve prediction performance. In *Proceedings of the Australasian Computer Science Week Multiconference.* ACM, 2.
[353] Shaymaa E. Sorour, Tsunenori Mine, Kazumasa Goda, and Sachio Hirokawa. 2014. Predicting students' grades based on free style comments data by artificial neural network. In *Frontiers in Education Conference (FIE).* IEEE, Madrid, Spain, 1–9.
[354] Shaymaa E Sorour, Tsunenori Mine, Kazumasa Godaz, and Sachio Hirokawax. 2014. Comments data mining for evaluating student's performance. In *2014 IIAI 3rd International Conference on Advanced Applied Informatics (IIAI-AAI).* IEEE, 25–30.
[355] T Stanko, O Zhirosh, D Johnston, and S Gartsev. 2017. On possibility of prediction of academic performance and potential improvements of admission campaign at IT university. In *Global Engineering Education Conference (EDUCON).* IEEE, 862–866.

[356] Lesley Strawderman, Bill Elmore, and Arash Salehi. 2009. Exploring the impact of first-year engineering student perceptions on student efficacy. In *ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[357] Chung-Ho Su. 2016. The effects of students' motivation, cognitive load and learning anxiety in gamification software engineering education: a structural equation modeling study. *Multimedia Tools and Applications* 75, 16 (2016), 10013–10036.

[358] E. Sugiharti, S. Firmansyah, and F.R. Devi. 2017. Predictive evaluation of performance of computer science students of unnes using data mining based on naÏve bayes classifier (NBC) algorithm. *Journal of Theoretical and Applied Information Technology* 95, 4 (2017), 902–911.

[359] S. Sultana, S. Khan, and M.A. Abbas. 2017. Predicting performance of electrical engineering students using cognitive and non-cognitive features for identification of potential dropouts. *International Journal of Electrical Engineering Education* 54, 2 (2017), 105–118.

[360] Emily S Tabanao, Ma Mercedes T Rodrigo, and Matthew C Jadud. 2011. Predicting at-risk novice Java programmers through the analysis of online protocols. In *Proceedings of the Seventh International Workshop on Computing Education Research*. ACM, 85–92.

[361] Ashay Tamhane, Shajith Ikbal, Bikram Sengupta, Mayuri Duggirala, and James Appleton. 2014. Predicting student risks through longitudinal analysis. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1544–1552.

[362] William T Tarimo, Fatima Abu Deeb, and Timothy J Hickey. 2016. Early detection of at-risk students in CS1 using teachback/spinoza. *Journal of Computing Sciences in Colleges* 31, 6 (2016), 105–111.

[363] S Taruna and Mrinal Pandey. 2014. An empirical analysis of classification techniques for predicting academic performance. In *2014 International Advance Computing Conference (IACC)*. IEEE, 523–528.

[364] Nguyen Thai-Nghe, Tomas Horv, Lars Schmidt-Thieme, et al. 2011. Personalized forecasting student performance. In *2011 11th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 412–414.

[365] Nguyen Thai-Nghe and Lars Schmidt-Thieme. 2015. Multi-relational factorization models for student modeling in intelligent tutoring systems. In *2015 Seventh International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 61–66.

[366] Siu-Man Raymond Ting and R Man. 2001. Predicting academic success of first-year engineering students from standardized test scores and psychosocial variables. *International Journal of Engineering Education* 17, 1 (2001), 75–80.

[367] Amit Kumar Tiwari, Divya Rohatgi, Akhilesh Pandey, and Anil Kumar Singh. 2013. Result prediction system for Multi-Tenant database of University. In *2013 International Advance Computing Conference (IACC)*. IEEE, 1340–1344.

[368] Sabina Tomkins, Arti Ramesh, and Lise Getoor. 2016. Predicting post-test performance from online student behavior: a high school MOOC case study. In *Proceedings of the 9th International Conference on Educational Data Mining*.

[369] Edmundo Tovar and Óliver Soto. 2010. The use of competences assessment to predict the performance of first year students. In *Frontiers in Education Conference (FIE)*. IEEE.

[370] Evis Trandafili, Alban Allkoçi, Elinda Kajo, and Aleksandër Xhuvani. 2012. Discovery and evaluation of student's profiles with machine learning. In *Proceedings of the Fifth Balkan Conference in Informatics*. ACM, 174–179.

[371] Bruno Trstenjak and Dzenana Donko. 2014. Determining the impact of demographic features in predicting student success in Croatia. In *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 1222–1227.

[372] Athanasios Tsalatsanis, Ali Yalcin, and Autar Kaw. 2009. Application of emerging knowledge discovery methods in engineering education. *ASEE Annual Conference & Exposition* (2009).

[373] Jaan Ubi, Innar Liiv, Evald Ubi, and Leo Vohandu. 2013. Predicting student retention by comparing histograms of bootstrapping for Charnes-Cooper transformationlinear programming discriminant analysis. In *2013 Second International Conference on e-Learning and e-Technologies in Education (ICEEE)*. IEEE, 110–114.

[374] M.F. Uddin and J. Lee. 2017. Proposing stochastic probability-based math model and algorithms utilizing social networking and academic data for good fit students prediction. *Social Network Analysis and Mining* 7, 1 (2017).

[375] Robert J Vallerand, Luc G Pelletier, Marc R Blais, Nathalie M Briere, Caroline Senecal, and Evelyne F Vallieres. 1992. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement* 52, 4 (1992), 1003–1017.

[376] Eric Van Inwegen, Seth Adjei, Yan Wang, and Neil Heffernan. 2015. An analysis of the impact of action order on future performance: the fine-grain action model. In *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge*. ACM, 320–324.

[377] Barend Van Wyk, Cecilia Louw, and Wiecher Hofman. 2013. Mathematics: A powerful pre-and post-admission variable to predict success in Engineering programmes at a University of Technology. *Perspectives in Education* 31, 4 (2013), 114–128.

[378] L. Vea and M.M. Rodrigo. 2017. Modeling negative affect detector of novice programming students using keyboard dynamics and mouse behavior. *Lecture Notes in Computer Science* (2017), 127–138.

[379] S.K. Verma, R.S. Thakur, and S. Jaloree. 2017. Fuzzy association rule mining based model to predict students' performance. *International Journal of Electrical and Computer Engineering* 7, 4 (2017), 2223–2231.

[380] Arto Vihavainen. 2013. Predicting Students' Performance in an Introductory Programming Course Using Data from Students' Own Programming Process. In *2013 13th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 498–499.

[381] Arto Vihavainen, Matti Luukkainen, and Jaakko Kurhila. 2013. Using students' programming behavior to predict success in an introductory mathematics course. In *Proceedings of the 6th International Conference on Educational Data Mining*.

[382] F Ruric Vogel and Salomé Human-Vogel. 2016. Academic commitment and self-efficacy as predictors of academic achievement in additional materials science. *Higher Education Research & Development* 35, 6 (2016), 1298–1310.

[383] Birgit Vogel-Heuser, Martin Obermeier, Steven Braun, Kerstin Sommer, Fabian Jobst, and Karin Schweizer. 2013. Evaluation of a UML-based versus an IEC 61131-3-based software engineering approach for teaching PLC programming. *IEEE Transactions on Education* 56, 3 (2013), 329–335.

[384] Christina Vogt. 2006. The crucial role of engineering faculty on student performance. In *ASEE Annual Conference & Exposition*. IEEE.

[385] Pattaramon Vuttipittayamongkol. 2016. Predicting factors of academic performance. In *2016 Second Asian Conference on Defence Technology (ACDT)*. IEEE, 161–166.

[386] Isabel Wagner. 2016. Gender and performance in computer science. *ACM Transactions on Computing Education (TOCE)* 16, 3 (2016), 11.

[387] Chong Wang and David M Blei. 2011. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 448–456.

[388] Feng Wang and Li Chen. 2016. A Nonlinear State Space Model for Identifying At-Risk Students in Open Online Courses.. In *Proceedings of the 9th International Conference on Educational Data Mining*. 527–532.

[389] Lisa Wang, Angela Sy, Larry Liu, and Chris Piech. 2017. Deep knowledge tracing on programming exercises. In *Proceedings of the Fourth Conference on Learning @ Scale*. ACM, 201–204.

[390] Rui Wang, Gabriella Harari, Peilin Hao, Xia Zhou, and Andrew T Campbell. 2015. SmartGPA: how smartphones can assess and predict academic performance of college students. In *Proceedings of the 2015 ACM IInternational Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 295–306.

[391] Masna Wati, Wahyu Indrawan, Joan Angelina Widians, and Novianti Puspitasari. 2017. Data mining for predicting students' learning result. In *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*. IEEE, 1–4.

[392] Christopher Watson, Frederick WB Li, and Jamie L Godwin. 2013. Predicting performance in an introductory programming course by logging and analyzing student programming behavior. In *2013 IEEE 13th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 319–323.

[393] Christopher Watson, Frederick WB Li, and Jamie L Godwin. 2014. No tests required: comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*. ACM, 469–474.

[394] Sam C Webb. 1951. A generalized scale for measuring interest in science subjects. *Educational and Psychological Measurement* 11, 3 (1951), 456–469.

[395] Christian Weber and Réka Vas. 2016. Applying connectivism? Does the connectivity of concepts make a difference for learning and assessment?. In *2016 International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 002079–002084.

[396] D. Wham. 2017. Forecasting student outcomes at university-wide scale using machine learning. *ACM International Conference Proceeding Series* (2017), 576–577.

[397] Sophie D White, Sybil May, and MF Washburn. 1917. A Study of Freshmen. *The American Journal of Psychology* 28, 1 (1917), 151–155.

[398] Jacob Whitehill, Joseph Williams, Glenn Lopez, Cody Coleman, and Justin Reich. 2015. Beyond prediction: First steps toward automatic intervention in MOOC student stopout. In *Proceedings of the 8th International Conference on Educational Data Mining*.

[399] Febrianti Widyahastuti, Yasir Riady, and Wanlei Zhou. 2017. Prediction model students' performance in online discussion forum. In *Proceedings of the 5th International Conference on Information and Education Technology*. ACM, 6–10.

[400] Febrianti Widyahastuti and Viany Utami Tjhin. 2017. Predicting students performance in final examination using linear regression and multilayer perceptron. In *2017 10th International Conference on Human System Interactions (HSI)*. IEEE, 188–192.

[401] Joseph B Wiggins, Joseph F Grafsgaard, Kristy Elizabeth Boyer, Eric N Wiebe, and James C Lester. 2017. Do You Think You Can? The Influence of Student Self-Efficacy on the Effectiveness of Tutorial Dialogue for Computer Science. *International Journal of Artificial Intelligence in Education* 27, 1 (2017), 130–153.

[402] Irmgard U Willcockson, Craig W Johnson, William Hersh, and Elmer V Bernstam. 2009. Predictors of student success in graduate biomedical informatics training: introductory course and program success. *Journal of the American Medical Informatics Association* 16, 6 (2009), 837–846.

[403] Annika Wolff, Zdenek Zdrahal, Andriy Nikolov, and Michal Pantucek. 2013. Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the Third International Conference on Learning Analytics & Knowledge*. ACM, 145–149.

[404] Thomas F Wolff, Steven M Cramer, and Barbara A Masi. 2011. Retention: Quantifying the Apples and Oranges. In *ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[405] Anna Woodcock, William G Graziano, Sara E Branch, Meara M Habashi, Ida Ngambeki, and Demetra Evangelou. 2013. Person and thing orientations: Psychological correlates and predictive utility. *Social Psychological and Personality Science* 4, 1 (2013), 116–123.

[406] Xinhui Wu, Junping Yang, and Changhai Qin. 2012. Student Achievement Databases Assist Teaching Improvement. In *Advances in Electronic Commerce, Web Application and Communication*. Springer, 209–214.

[407] Yonghong Jade Xu. 2018. The Experience and Persistence of College Students in STEM Majors. *Journal of College Student Retention: Research, Theory & Practice* 19, 4 (2018), 413–432.

[408] Haiqin Yang and Lap Pong Cheung. 2018. Implicit Heterogeneous Features Embedding in Deep Knowledge Tracing. *Cognitive Computation* 10, 1 (2018), 3–14.

[409] T.-Y. Yang, C.G. Brinton, C. Joe-Wong, and M. Chiang. 2017. Behavior-Based Grade Prediction for MOOCs Via Time Series Neural Networks. *Journal on Selected Topics in Signal Processing* 11, 5 (2017), 716–728.

[410] Yu Yang, Hanqing Wu, and Jiannong Cao. 2016. Smartlearn: Predicting learning performance and discovering smart learning strategies in flipped classroom. In *2016 International Conference on Orange Technologies (ICOT)*. IEEE, 92–95.

[411] Nong Ye, Ting Yan Fok, Xin Wang, James Collofello, and Nancy Dickson. 2018. The PVAD Algorithm to Learn Partial-Value Variable Associations with Application to Modelling for Engineering Retention. *IFAC-PapersOnLine* 51, 2 (2018), 505–510.

[412] Florence Yean Yng Ling, Poh Khai Ng, and Mei-yung Leung. 2010. Predicting the academic performance of construction engineering students by teaching and learning approaches: Case study. *Journal of Professional Issues in Engineering Education & Practice* 137, 4 (2010), 277–284.

[413] Her-Tyan Yeh, Wei-Sheng Lin, and Chaoyun Liang. 2014. The effects of imagination between psychological factors and academic performance: The differences between science and engineering majors. *International Journal of Engineering Education* 30, 3 (2014), 746–755.

[414] Osman Yildiz, Abdullah Bal, and Sevinc Gulsecen. 2013. Improved fuzzy modelling to predict the academic performance of distance education students. *The International Review of Research in Open and Distributed Learning* 14, 5 (2013).

[415] Chong Ho Yu. 2012. Examining the relationships among academic self-concept, instrumental motivation, and TIMSS 2007 science scores: A cross-cultural comparison of five East Asian countries/regions and the United States. *Educational Research and Evaluation* 18, 8 (2012), 713–731.

[416] Hsiu-Ping Yueh, Chi-Cheng Chang, and Chaoyun Liang. 2013. Are there differences between science and engineering majors regarding the imagination-mediated model? *Thinking Skills and Creativity* 10 (2013), 79–90.

[417] Amelia Zafra, Cristóbal Romero, and Sebatián Ventura. 2009. Predicting academic achievement using multiple instance genetic programming. In *Ninth International Conference on Intelligent Systems Design and Applications (ISDA)*. IEEE, 1120–1125.

[418] Amelia Zafra and SebastiáN Ventura. 2012. Multi-instance genetic programming for predicting student performance in web based educational environments. *Applied Soft Computing* 12, 8 (2012), 2693–2706.

[419] Leping Zeng, Dan Chen, Kun Xiong, Aihua Pang, Jufang Huang, and Lianping Zeng. 2015. Medical University Students' Personality and Learning Performance: Learning Burnout as a Mediator. In *2015 7th International Conference on Information Technology in Medicine and Education (ITME)*. IEEE, 492–495.

[420] Wei Zhang, Xujun Huang, Shengming Wang, Jiangbo Shu, Hai Liu, and Hao Chen. 2017. Student performance prediction via online learning behavior analytics. In *2017 International Symposium on Educational Technology (ISET)*. IEEE, 153–157.

[421] Qing Zhou, Youjie Zheng, and Chao Mou. 2015. Predicting students' performance of an offline course from their online behaviors. In *International Conference on Digital Information and Communication Technology and Its Applications (DICTAP)*. 70–73.

[422] Ke Zhu. 2014. Research based on data mining of an early warning technology for predicting engineering students' performance. *World Transactions on Engineering and Technology Education* 12, 3 (2014), 572–575.

[423] Daniel Zingaro, Michelle Craig, Leo Porter, Brett A Becker, Yingjun Cao, Phill Conrad, Diana Cukierman, Arto Hellas, Dastyni Loksa, and Neena Thota. 2018. Achievement Goals in CS1: Replication and Extension. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. ACM, 687–692.

[424] J.P. Zwolak, R. Dou, E.A. Williams, and E. Brewe. 2017. Students' network integration as a predictor of persistence in introductory physics courses. *Physical Review Physics Education Research* 13, 1 (2017).