

Received February 6, 2020, accepted February 29, 2020, date of publication March 6, 2020, date of current version March 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2978979

Supporting Teachers to Monitor Student's Learning Progress in an Educational Environment With Robotics Activities

SAMANTHA ORLANDO¹, ELENA GAUDIOSO¹, AND FELIX DE LA PAZ¹

Artificial Intelligence Department, School of Computer Science, UNED, 28015 Madrid, Spain

Corresponding author: Samantha Orlando (sorlando1@alumno.uned.es)

This work was supported by the Italian Ministry Ministero dell'Istruzione, dell'Università e della Ricerca, MIUR, under Grant 488/2001.

ABSTRACT Educational robotics has proven its positive impact on the performances and attitudes of students. However, the educational environments that employ them rarely provide teachers with relevant information that can be used to make an effective monitoring of the student learning progress. To overcome these limitations, in this paper we present IDEE (Integrated Didactic Educational Environment), an educational environment for physics, that uses EV3 LEGO Mindstorms® educational kit as robotic component. To provide support to teachers, IDEE includes a dashboard that provides them with information about the students' learning process. This analysis is done by means of an Additive Factor Model (AFM). That is a well-known technique in the educational data mining research area. However, it has been usually employed to carry out analysis about students' performance data outside the system. This can be a burden for the teacher who, in most cases, is not an expert in data analysis. Our goal in this paper is to show how the coefficients of AFM provide valuable information to the teacher without requiring any deep expertise in data analysis. In addition, we show an improved version of the AFM that provides a deeper understanding about the students' learning process.

INDEX TERMS Educational technology, computer aided instruction, computer applications, adaptive systems, STEM.

I. INTRODUCTION

New developments in educational robotics are making possible for teachers to bring technology into the classroom more than ever before, in particular in the fields of Science, Technology, Engineering, and Mathematics (STEM) [1], [2]. In fact, nowadays there is an ever-increasing interest in educational robotics. Experiences like the one reported in [3] show the great interest that robotics raises in students. Educational robotics offers interesting teaching and learning opportunities at various levels and school subjects [2], [4]–[7] mainly to acquire skills on programming [8], robotics or electronic engineering. For example, in [9] the authors present a robot-visual programming environment interface to let K-12 students construct their own programs to manipulate certain robot components. In [10], the authors present how a small autonomous robot is used to motivate students in

an electronic engineering laboratory. In [11], the authors compare a robotics curriculum by using both commercial and open software and hardware. In addition, educational robotics has also an important role as a motivational component [12], [13]. Constructivism and constructionism are two main theories behind educational robotics. Following the Papert paradigm of constructionism [14], students should construct their knowledge under the guidance of the teacher. Papert believed that to encourage students' active learning, besides the theoretical content, it is required to provide practical exercises and up-to-date tools in laboratories. This is especially true in scientific subjects, such as physics.

Although several studies report the positive results obtained from experimental courses [15], [16], educational robotics still remains an extracurricular activity usually isolated from the official curriculum [17]. This is in part due to the belief that students try to solve the task in robotics classes with a guess-and-check technique by using the feedback of the digital technology [18]. To employ robotics tools in

The associate editor coordinating the review of this manuscript and approving it for publication was Alba Amato¹.

learning environments for educational use, it is needed to support teachers in the students' learning assessment by providing relevant information about the students' learning outcomes [19], [20]. The assessment of the students should not be limited to grading them as it should be more informative by identifying the subjects where they excel or find difficulties.

To support the learning assessment, educational data mining research area [21] offers several methods to analyze data collected from learning environments to better understand the learning outcomes or the engagement of the students. Nevertheless, since educational robotics is not yet embedded in learning environments, teachers do not get an analysis of the learning process of their students while using the robotics system.

In this paper we present IDEE (Integrated Didactic Educational Environment), a learning environment for physics focused on the concepts that students must acquire by following the official regulations.¹ IDEE [22], [23] can be used to carry out laboratory activities with a robotic kit and to work on the mandatory concepts that students have to acquire in a physics class in secondary education. To support teachers, IDEE tracks students' interaction data and provides a dashboard that allows the teachers to analyze and assess the progresses and difficulties of the class group. Specifically, our aim is to identify those students who have difficulties on certain concepts or to detect which concepts are too difficult or trivial to learn. To this end, we have used Additive Factor Model (AFM) [24]. AFM is an effective model of the student performance. The right interpretation of the model coefficients helps in identifying the students that are actively learning and the skills that result more difficult to learn. Our goal in this paper is to show how the coefficients of the AFM along with the learning curves and the interactive graphs may offer valuable information to the teachers about the learning performance of the students.

The rest of the paper is organized as follows. In the next section, we describe some related work on the use of educational robotics emphasizing the need to support teachers while using in class educational environments that use robotics. In section III, we make an overview of IDEE, and how it allows to develop and present the contents of a regular curricular programme. In section IV more details on the AFM model are given and how teachers are supported in IDEE. Finally, in section V we present our conclusions and lines for future work.

II. RELATED WORK

With the rapid development of computer technologies, the importance of acquiring computational skills has been widely recognized in recent education [25].

In fact, a major trend, in the research area on design of educational applications, is the development of learning programming environments with the goal of making the

¹In 2006 the European Parliament recommended a teaching method based on competences that structures the didactic programming according to learning units.

programming tasks more friendly and pleasing to beginner students [26]. In this context, Scratch [27] was released with the goal of providing to novice students an environment to learn programming. In order to help students in programming with Scratch, some learning environments track students' interactions to provide hints when needed. This is the case of iSnap [28]. iSnap is an educational environment that tracks students' interactions to provide hints to students when programming with Scratch.

With the proliferation of robotics tools, the new version of Scratch, Scratch3, includes interfaces to program different robotic tools, such as, Lego EV3 Mindstorm®. The increasing interest to encourage children in programming tasks has led to several initiatives. A major example is the Open Roberta Fraunhofer IAIS lab [29], which looks to engage children by the use of different robots.

Besides programming teaching, educational robotics is being used also in other subjects. For example, in [30] educational robotics is used to support teaching of native language. The authors propose two activities with Lego Minstorms EV3 that can be performed by the teacher. In [17] the authors, propose the use of Bee-Bots in a course unit on road safety.

Educational robotics has been used also in sciences education. For example, an educational social robot has been used to motivate students in reading science texts [31]. In [32] the authors propose the use of a MARRtino robot to teach Newtonian concepts. In [33] students are asked to build and use a rescuer robot to apply knowledge on electricity. All these approaches evaluate the impact of using educational robotics by comparing the level of knowledge of the students before and after the robotics experience.

Nevertheless, due to the unpredictable nature of a class that uses robots where students may interact in different ways [34], there are only few proposals to support teachers [35]. Effective use of educational robotics in class requires learning environments that take into account the characteristics of the students, the specific learning objectives and the assessment procedures that will best measure and guide learning.

Educational data can be analyzed in order to get useful insights to assess the learning process and to support both teachers and students. Educational data mining research area has employed techniques mainly from machine learning area in order to analyze educational data [36]. The goal of many of these applications is to predict students behavior [35]. For example, in [37] the authors show how machine learning techniques can be used to early predict the student success in a block-based programming environment. In [38] the authors propose an approach to early predict students' difficulties in open-ended programming tasks. In [39] the authors analyze the students' interactions with massive open online courses with the aim to predict students' motivation.

Data mining techniques can be also used to describe how the learning process develops [21]. The goal is to give teachers information to improve materials or to reinforce certain concepts in the course. For example, in [40] the authors apply

clustering techniques to group students with similar interaction patterns. The clusters are shown to the teacher who can provide some recommendations about additional material only to certain groups. Supervised learning techniques have also been employed to predict student's outcomes and to provide teachers with a dashboard to inspect those predictive models [41]. In [39] several machine learning techniques are used to help teachers to detect unmotivated students at their early stages in order to timely provide support to them.

However, educational environments that use robotics do not usually track student interactions. When used in the classroom, the teacher usually can only assess the final work done by the student without any information on his learning process, but he cannot trace the causes of the results. An exception may be the work described in [42]. There, the authors describe a learning environment that helps the teacher to supervise coding activities and to detect the right moments for intervention to help the students. This is done by computing a set of rules derived by using decision trees applied to interaction data.

Thus, it is necessary to include new tools to track and model students' progress in learning environments that use robotics.

IDEE is a learning environment created to be used in class by the teacher to explain physics content by employing robotics for certain activities. All the student's interactions with the environment are tracked and stored in a database. To analyze this data we use AFM [24]. AFM has been widely employed to support the assessment of the level of assimilation of the contents of the course by the students [43]. AFM is generally applied within the framework available at PSLC-Datashop [44] and it requires a certain amount of background knowledge about the techniques involved. A great majority of teachers have not such background knowledge and thus we have implemented AFM as a component in IDEE to allow teachers to perform their own analysis of the activities using robotics. In addition, to ease the use of the dashboard, in IDEE we also provide learning curves [45] and interactive graphs that can be helpful for the teacher to understand the AFM parameters.

III. IDEE: INTEGRATED DIDACTIC EDUCATIONAL ENVIRONMENT

A. OVERVIEW

IDEE is a learning environment that provides a framework to study learning units in physics through the use of robotics [23].

Each experience in IDEE analyzes a physical phenomenon. The students study the phenomenon as true scientists by following the steps of Galileo Galilei's scientific method: first they observe the phenomenon, then they collect experimental data that will lead them to formulate hypotheses that will finally be validated by a mathematical law [23].

In IDEE, students have to manipulate robotics as laboratory artifacts in order to test their hypothesis about the solution of each learning activity. Students can visualize their

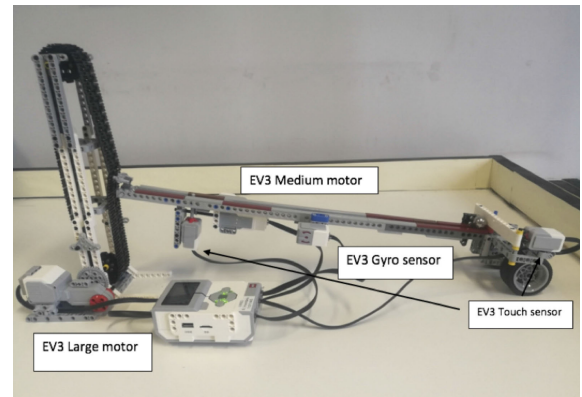


FIGURE 1. The robot designed for the inclined plane experiment in IDEE.

proposed solutions as often as they wish while observing the behavior of the robot.

Currently, IDEE contains 4 experiences: Uniform rectilinear motion, uniformly accelerated motion, inclined plane and free fall.

For each experience different types of activities are defined:

- *Introduction to the robot*: the objective of this activity is to allow students to know how the robot works. To do this, students have to connect the different sensors and motors to the robot. As an example, in Figure 1 we show the robot used in the experience of the inclined plane.
- *Laboratory*: in this activity, students have to use a set of blocks to program an action of the robot, for example, drop a ball on an inclined plane or make a car travel a distance in the uniformly accelerated motion experiment. In these activities the robot returns certain data such as the time spent by the ball to roll down the inclined plane. Students gather and enter this data in IDEE.
- *Test*: a test is defined to make students reflect on the physics concepts involved in the experiment.
- *Final problem*: it is a problem that students may solve without the use of the robot. This final problem allows the theoretical analysis of the phenomenon.

While solving a problem, IDEE automatically evaluates the solution provided by the student. Figure 2 depicts a screenshot of IDEE learning environment and an evaluation of a provided solution, given by IDEE.

B. IDEE ARCHITECTURE

The architecture of IDEE is shown in Figure 3 and it is based on:

- Django²: it is a framework that allows the creation and management of Python web applications. It handles the user management and the interaction with the database.

²<https://www.djangoproject.com>

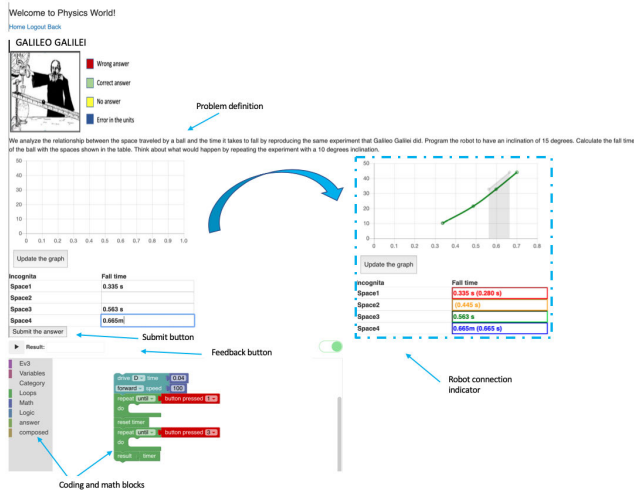


FIGURE 2. General interface of IDEE. Each learning activity in IDEE is composed by the following elements: the statement of the activity, the blocks that will be necessary to build the solution and the feedback button. In the overlay image it is shown the solution given to the student (the correct and incorrect answers and the graph that results from the solutions).

navigate through the different activities (problems or tests), the interface provides the activity statement and a working area. In IDEE students work on the activity and manipulate the robot by using the programming editor that is based on drag-and-drop blocks and developed with Blockly (some blocks for communicating with the robot have been implemented specially for IDEE). The program in Blockly is internally converted into JavaScript and interpreted by IDEE. When students finish the problem, they submit the solution and the interaction data is stored in the database. All the data collected during the activities, as well as the answers to tests and problems related with them and with the involved skills are stored in the PostgreSQL database.

The dashboard developed with Dash allows teachers to actively monitor the collected data (see section IV-B). Dash allows the creation of interactive views of the data together with statistical graphs and tests.

C. STUDENTS' KNOWLEDGE TRACING IN IDEE

The learning activities offered in IDEE work on a series of skills that embody the knowledge that students must acquire. Student actions are coded as correct or incorrect and categorized in terms of the skills needed to perform the contents of the learning unit.

In Table 1 we show some examples of the skills defined for the learning unit of the kinematic study of movement.

Activities in IDEE work on different skills, but not all of the involved skills have the same relevance. Thus, when creating an activity we have to define a *score* parameter for each skill that represents the level of knowledge that the skill contributes to (W_i).⁵ Table 2 shows some of the activities defined in IDEE in two different experiences (inclined plane and uniform rectilinear movement), the skills involved and their corresponding scores.

IDEE tracks the students' interaction data by storing it in the database. In particular, IDEE stores, in a table called *profile table*, the relationship between the activity (the field *element_id*), the skill (the field *skill_id*) and the knowledge level that the student acquires in that skill. We call this last field *accumulated score*. The *accumulated score* determines the degree in which each skill is achieved by the students when they do the activity and it is obtained for each skill and for each student taking into account the *score* parameter introduced before.

When a student gives a correct answer, IDEE determines the value of the *accumulated score* for each skill in the following manner:

- if the activity is a test, the accumulated score is equal to the score of the skill in the activity.
- If the activity is a problem or a laboratory exercise, the *accumulated score* (AS_{ij}) for skill i and student j is

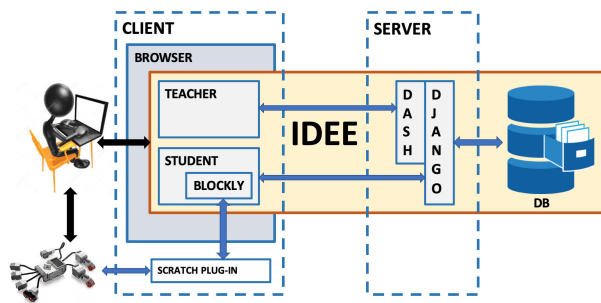


FIGURE 3. The architecture of IDEE.

- Blockly: it is an open source visual programming environment³ developed by Google. It is used to allow the students to implement programs during the learning activities. It has been chosen because its ease of use.
- Scratch Plug-in: it allows the communication between the Blockly program and the Lego EV3 robot through a bluetooth connection.
- Dash⁴: it is used to develop the teacher dashboard.
- PostgreSQL database: it is used to store and retrieve the data recorded during the activities.

IDEE uses the EV3 LEGO Mindstorms® educational kit as the robotic component. This kit has been widely employed in educational settings [46] due to its versatility, popularity among students and its diverse features including sensors, motors and a programmable device that controls the movement and reaction of the robot [47].

Students and teachers interact with IDEE through a web-based interface. Besides a menu that allows them to

³<https://developers.google.com/blockly/>

⁴<http://www.dash-project.org>

⁵These vales are defined by the teacher according to the didactic programming of the learning unit.

TABLE 1. Skills included in IDEE to achieve the contents of the didactical programming of the learning unit of kinematic study of movement.

LEARNING UNIT	KINEMATIC STUDY OF MOVEMENT		
COMPETENCE	Ability to describe the rectilinear motion of a body using the equations that connect space, speed and time.		
CONTENTS	Uniform rectilinear motion	Uniformly accelerated movement	
	Velocity	Free fall	
	Distance and displacement	Acceleration	
SKILLS	Graphics interpretation (skill_1)	Reference system (skill_2)	Relate magnitude (skill_3)
	Operation maths (skill_7)	Uniformly accelerated movement (skill_8)	Velocity (skill_10)
	Time (skill_11)	Knowledge in robotics (skill_12)	Uniform rectilinear motion (skill_13)

TABLE 2. Example of skills involved in several activities in IDEE. The degree of completion (or score parameter) of each skill when performing certain activity is shown in parenthesis. (URM :Uniform Rectilinear Movement).

Learning activity	skill_1	skill_2	skill_7	skill_8	skill_10	skill_11	skill_12
Robot inclined plane							✓(1)
Laboratory inclined plane		✓(0.8)		✓(0.5)			✓(1)
Problem inclined plane			✓(0.8)	✓(0.8)		✓(0.8)	✓(0.8)
Robot URM							✓(1)
Laboratory space/time URM		✓(0.8)			✓(0.8)		✓(0.8)
Problem URM		✓(0.8)				✓(0.8)	
Question2_test2	✓(1)	✓(0.5)					
Question7_test4	✓(1)	✓(1)					

calculated as in equation (1):

$$AS_{ij} = \frac{n_j}{N_j} * W_i \tag{1}$$

where n_j is the number of correct answers given by the student j in the activity, N_j is the number of total answers in the activity and W_i is the score parameter of skill i in the activity.

For example, let’s consider a student working in the laboratory of the inclined plane; the student has to calculate the time of fall of a ball on an inclined plane considering different lengths. The laboratory of the inclined plane is related to different skills: skill_12 with score 10 ($W_{12} = 10$); skill_3 with score 8 ($W_3 = 8$) and skill_8 with score 5 ($W_8 = 5$). In this activity, the students have to answer four different questions (one for each value of the length of the inclined plane). In the student’s response shown in Figure 4 there are: two exact answers (shown in green), a wrong one (shown in red) and one with incorrect units (shown in blue); in this case ($n_j = 2$ and $N_j = 4$), the *accumulated score* should take 2/4 as a factor to be multiplied by the scores of the student on the different skills in the activity. In Figure 4, we show the values for the *accumulated score* that are stored in the *profile* table for this particular example.

In addition, to support student’s learning, IDEE presents different versions of the same final activity problem. Thus,

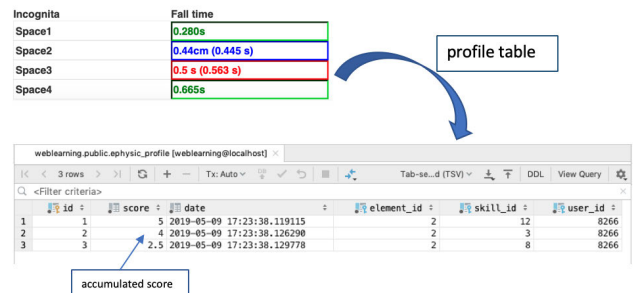


FIGURE 4. An example of a student’s answers and its tracking in the database of IDEE.

each version depends on the level of knowledge of each student. This level is obtained taking into account the average of accumulated scores ($AvgScore$) of the students in the activities on which he worked.

Thus, a student j with a high level of knowledge ($AvgScore_j > 0.8$) will not be offered any additional support. A student with a medium level of knowledge ($0.5 < AvgScore_j \leq 0.8$) will be guided in the logical steps necessary to solve the problem (logical help). Finally a student with a low level of knowledge ($AvgScore_j < 0.5$) will also be supported on the mathematical formula necessary to solve the problem (mathematical help). These thresholds have been defined by taking into account the intervals of grades usually considered

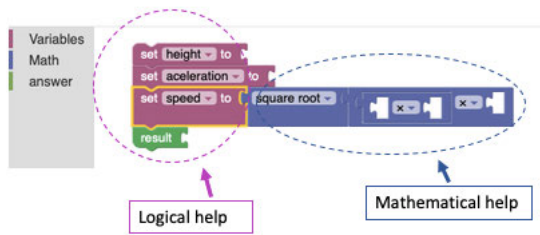


FIGURE 5. The working area for a student with a low knowledge level and the hints that are provided.

at schools (below a grade of 5 it is considered that the student needs to improve, between 5 and 8 it is considered that the student has a good level of knowledge and above 8 the student is considered to excel).

As an example, Figure 5 depicts the working area in IDEE for a student with a low level of knowledge.

IV. STUDENTS' INTERACTIONS ANALYSIS

In IDEE students' interactions analysis is done by means of AFM (Additive Factor Model) [24]. Our goal is to show how the coefficients obtained in AFM can provide useful information to a teacher. In this section we do an overview of AFM and describe the experiences carried out with IDEE and the results obtained.

A. OVERVIEW OF ADDITIVE FACTOR MODEL

The goal of AFM is to represent the probability that a student correctly answers on an activity by taking into account those in which the student has been working on before. The AFM is a statistical algorithm for modeling learning and performance that uses multiple logistic regression with student-success (0 or 1) as the dependent variable and with an independent variable for each opportunity interaction (i.e. the number of attempts of each student on each skill) [48]. The AFM is based on a binary Q-matrix that represents the iterations of each student with each skill while performing the activity. Below we show a simplified example of the Q-matrix used by AFM for one student:

$$\begin{pmatrix} & \text{skill}_1 & \text{skill}_2 \\ \text{activity}_1 & 1 & 1 \\ \text{activity}_2 & 1 & 0 \\ \text{activity}_3 & 1 & 1 \end{pmatrix}$$

We can see that the student has worked on skill_1 in all the activities and on skill_2 in activity_1 and activity_3.

Formally, given a group of N students, a set of K skills and a set of J activities or items, in AFM, the probability that a student $i \in \{1 \dots N\}$ gives a correct answer in an activity, is expressed by equation (2).

$$\begin{aligned} p &= P(Y_{ij} = 1 | \alpha_i, \beta, \gamma) \\ &= f\left(\alpha_i + \sum_{k=1}^K \beta_k * q_{jk} + \sum_{k=1}^K \gamma_k * q_{jk} * T_{ik}\right) \\ f(x) &= \frac{1}{(1 + e^{-x})} \end{aligned} \quad (2)$$

where: p is the probability of student's success on an exercise; Y_{ij} is the binary response of student i to activity j ; q_{jk} is the binary indicator that activity j uses skill k ; α_i is the student intercept that represents the student's knowledge; β_k is the skill intercept for skill k which represents the difficulty of the skill; γ_k is the skill slope that represents how quickly skill k can be assimilated and T_{ik} is the number of times the student i has worked on skill k .

Thus, in AFM model, the probability of success depends on the ability of student i (α_i), the easiness of skill k (β_k) and the learning rate for skill k (γ_k). The *students intercept* parameter is used for each student (the initial knowledge of each student is different). The *skill intercept* parameter is used for each skill (some skills are better known than others) and the *skill slope* depends on the skill but does not depend on the student (some skills are easier to learn than others).

B. ADDITIVE FACTOR MODEL IN IDEE

In educational data mining area, Q-matrices like the ones used in AFM have been widely employed to analyze data coming from intelligent tutoring systems. Nevertheless, up to our knowledge this analysis is done using PSLC-Datashop [44]. There, researchers can upload their dataset and their Q-matrices and analyze their educational data. It can be a burden for the teacher who, in most cases, is not an expert in data analysis. Our goal in IDEE is to provide a dashboard to let teacher analyzes the students' performance by means of AFM without any knowledge about the techniques involved in those analyses. To this end we have also focused on visualizations that can be useful for the teacher in interpreting what is happening in the classroom context.

To implement AFM in IDEE, we have used the machine learning Python library *scikit-learn*.⁶ In particular, IDEE applies a logistic regression classifier included in the *StatsModels* Python library [49]. To obtain the coefficients of the logistic regression, we have employed Generalized linear model (GLM) [49]. To display the interactive data visualizations we have used the library *dash_core_component* in *Dashboard*.⁷

In particular, in IDEE the teacher's module contains what follows:

- 1) *Indicators of the learning process*: a table with the AFM coefficients α , β and γ mentioned in section IV-A. These coefficients are obtained applying GLM (binomial family) to students' interaction data. The coefficient α shows the students' abilities in the skills and it is graphically represented in IDEE using a Polar graph. Accordingly, the coefficient β_k represents the easiness of skill k . The highest value for β_k implies a lower initial difficulty for skill k ; the highest value of the coefficient γ_k indicates how fast students learn

⁶<https://scikit-learn.org/>

⁷<https://dash.plot.ly/dash>

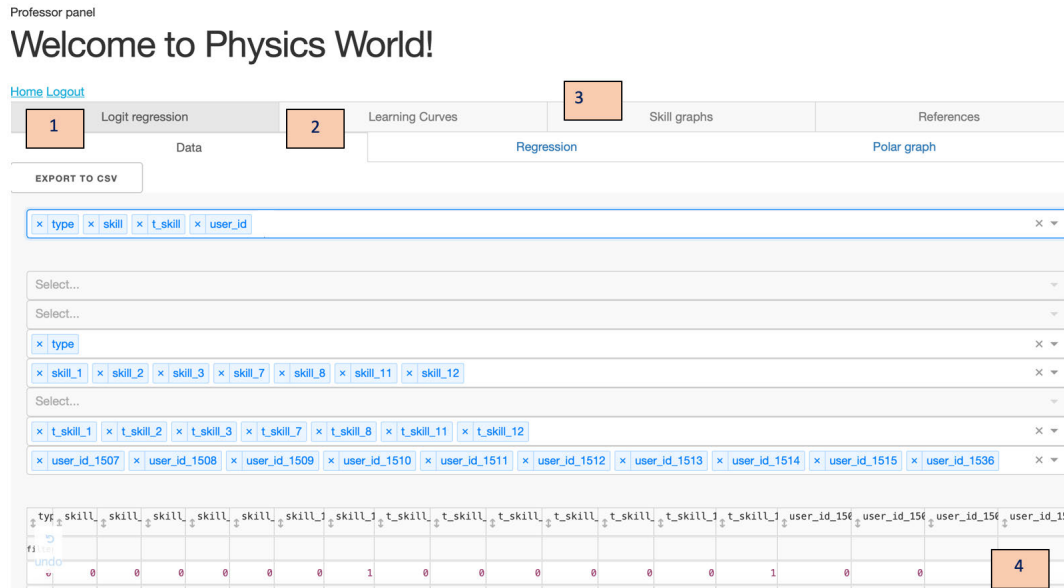


FIGURE 6. Screenshot of the teacher’s module in IDEE. It can be seen that the teacher can access different information using the available menu: indicators of learning (1); learning curves (2) or graphs with the accumulated score of students on the different skills (3) and student’s interaction table (4).

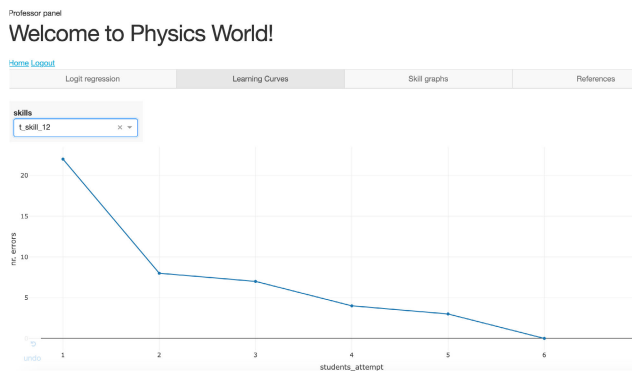


FIGURE 7. Screenshot of the teacher’s module which shows the learning curve for skill_12. In this skill, the students seem to improve their learning since the number of errors declines and the curve is smooth and decreasing.

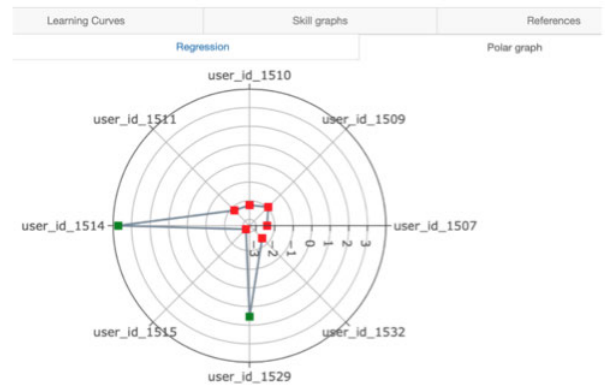


FIGURE 8. Graphical analysis of the student intercept parameter α with a Polar Graph. With red points it is shown students who presented learning difficulties on the skills analyzed, with green points the students that no present difficulties. Students are identified by their IDs but teachers has access to a list with the correspondence between IDs and students’ names.

skill k . In IDEE, for each skill k , significative values for coefficients β_k and γ_k are shown to the teacher (those with p-value < 0.05).

- 2) *Learning curves*: learning curves plot performance versus opportunities to practice [45]. Thus they show the number of errors in an activity for each attempt by all the students. In IDEE, learning curves are shown to the teachers for each skill.
- 3) *Graphs*: teacher can see also a graph showing the *accumulated score* for each student and each skill.
- 4) *Student’s interaction table*: a table with the activities carried out by a selected student, the time that this student has spent in the activities and the mistakes he made in the different activities.

We can see in Figure 6 a screenshot of the teacher’s module in IDEE.

C. RESULTS

IDEA has been used in a physics class in the Gredos Las Suertes School⁸ with 30 students aged between 16 and 17 and supervised by a teacher of mathematics and physics of secondary education. The goal of the experience with IDEA is to help students to understand the concepts of the learning unit of kinematic study of movement through experimentation. Prior to working with IDEA, students theoretically dealt with the topic in class with their physics teacher. In addition, students’ knowledge on the subject is assessed to ensure that they are prepared to take advantage of the experiences in the laboratory. The experience lasted two hours; each student

⁸An official permission from the Gredos Las Suertes school was obtained to carry out the experience.

Professor panel

Welcome to Physics World!

[Home](#) [Logout](#)

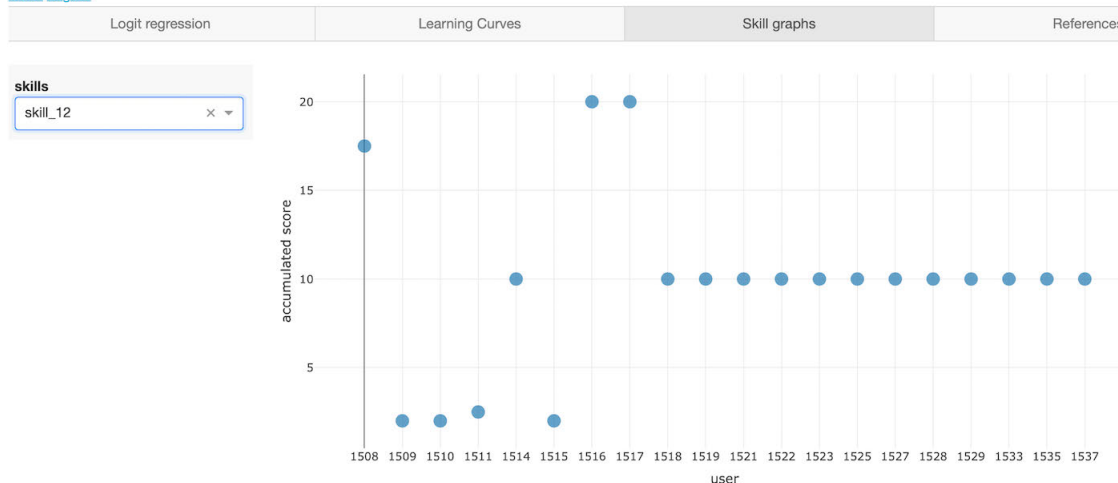


FIGURE 9. Screenshot of the teacher’s module that shows the *accumulated score* of each student for a particular skill.

TABLE 3. Indicators of the learning process.

skill_id	skill name	skill intercept (β)	Pf(> lzI)	Slope (γ)	Pr(> lzI)
skill_2	Reference system	-4.23	0.026738	3.2	0.022850
skill_12	Knowledge in robotics	-4.16	3.26e-09	0.84	0.000111

(a) The skill intercept is the initial difficulty level for each skill and the slope is the learning rate of each skill.

student	student intercept (α)	Pf(> lzI)
user_id_1507	-2.39929	0.004915
user_id_1509	-1.92893	0.024076
user_id_1510	-2.24017	0.005639
user_id_1511	-2.18629	0.006486
user_id_1514	3.71865	0.001396
user_id_1515	-3.07097	0.000200
user_id_1529	1.54727	0.037535
user_id_1532	-2.38468	0.003317

(b) Results of the student intercept parameter for significant students (in bold are shown those students with low learning).

worked with his/her computer and carried out the laboratory activities in a group of three students.

IDEE processes the data stored from this experience and fill in a derived table, made of 570 entries. Each entry is composed by the student_id, the skill he/she is working and the number of attempts in each skill he/she has previously done. This is the table used to apply AFM.

The goal in this experience was to teach the learning unit of kinematic study of movement (the competence, the contents and the nine skills covered are listed in Table 1).

The four activities created in IDEE are all aimed at understanding this learning unit.

IDEE analyzes the data with a logistic regression whose significant results are summarized in Table 3a and in Table 3b. The interpretation of the table results can be made considering the definition of the AFM described in section IV-A.

From the results in Table 3a, in the column skill intercept (β), it can be highlighted some learning deficiencies:

- Students have difficulties when they work on activities that require the physical concept of reference system; skill_2 (“Reference system”) with $\beta = -4.23$.
- Students present initial difficulties when they use robotics tools in IDEE; skill_12 (“Knowledge in robotics”) with $\beta = -4.16$.

The initial difficulties on the use of the robotic tools, highlighted by the model, appears consistent with the fact that it was the first time the students were using IDEE.

In the same Table 3a, the column slope γ allows to analyze the evolution students have when interacting with the system. Thus:

- in skill_2, the parameter $\gamma = 3.2$ shows that students learn quickly.

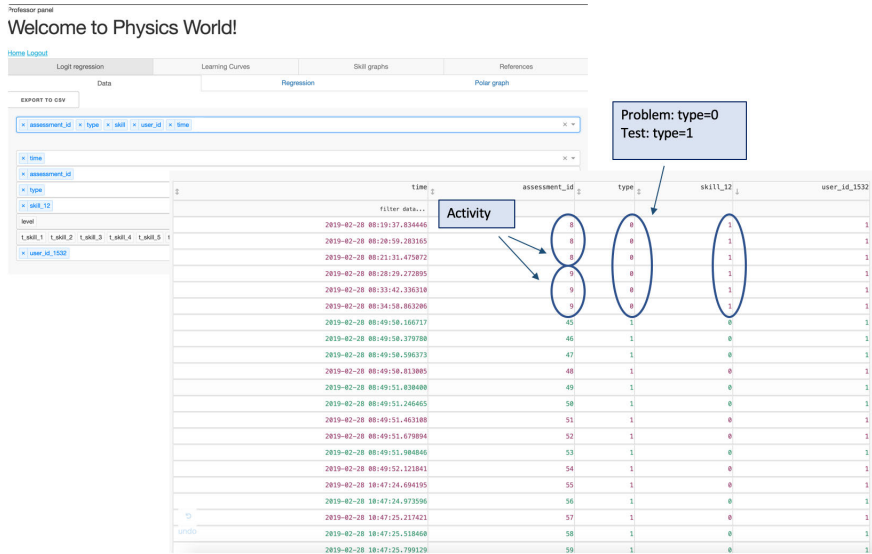


FIGURE 10. The student has worked on skill_12 six times in two problems activities and all the answers have been incorrect (they are all marked in red).

- In skill_12, the parameter $\gamma = 0.84$ shows an improvement but it is slower. The result would be better if the students were already exposed to the use of the system.

In addition, according to the results shown in Table 3b, the analysis of the parameter α evinces a group of 6 students with learning difficulties on the skills analyzed marked in bold in the table and two students who showed no difficulties (user_id_1514 and user_id_1529).

By using the dashboard implemented in IDEE, the teacher is able to understand all these results without being an expert in educational data mining. Notice that our goal is not to accurately predict the students' success but to show the teachers which skills seem too difficult or too easy and which students seem to be experiencing problems in the activities. As it has been described in section IV-B, in IDEE the statistical results of the analysis are graphically displayed to the teacher by means of learning curves and interactive graphs.

The main goal of the learning curves in IDEE is to let the teacher focus on the progress or the difficulties of the class group in those skills that are considered relevant. As an example, we give a possible set of steps that the teacher may follow to understand the situation of his/her class. Looking at Figure 7, teacher can see the learning curve displaying the first 6 attempts for all the students in the skill_12 (this skill is significant both considering β_{12} and γ_{12} as shown in Table 3a). The students show a high initial difficulty (high number of errors). During the activities the number of errors declines, so the curve is smooth and decreasing. Thus in this skill students improve in their learning.

From the graph in Figure 8, the teacher can see with a graphical representation of Table 3b that a group of 6 students presented learning difficulties (red points) on the significant

$$\begin{pmatrix}
 & skill_1 & skill_2 & answer & student_1 & student_2 \\
 activity_1 & 1 & 1 & 1 & 1 & 0 \\
 activity_2 & 1 & 0 & 0 & 1 & 0 \\
 activity_3 & 1 & 1 & 0 & 1 & 0 \\
 activity_1 & 1 & 1 & 0 & 0 & 1 \\
 activity_2 & 1 & 0 & 0 & 0 & 1 \\
 activity_3 & 1 & 1 & 1 & 0 & 1
 \end{pmatrix}$$

FIGURE 11. Example of the Q-matrix used by AFM.

skills and two students who presented no difficulties (green points). Both the table and the graph are available in the dashboard. In addition, looking at Figure 9, the teacher can inspect accumulated scores for skill_12 of those students with difficulties. Students with IDs 1509, 1510, 1511 and 1515 have a low value for accumulated score while students with IDs 1507 and 1532 do not have value for accumulated score. By selecting these students in the student's interaction table, in Figure 10, the teacher can see that the student with ID 1507 has not worked but the student with ID 1532 has worked but he has never given correct answers. With this information, teachers may decide to offer additional help or supervise more closely certain students.

This type of analysis turns out to be more effective for the teacher than a test to verify the skills acquired during the activities, indeed a test gives no information to the teacher about the learning process of the students and therefore about their difficulties [16].

D. DEEPER UNDERSTANDING OF STUDENTS' LEARNING PROCESS

As we have seen in the subsections above and in equation (2), AFM only considers the student's interactions on different skills to understand student's learning process. In this section we make a deeper analysis of the recorded data and we try

$$\begin{pmatrix} & skill_1 & skill_2 & answer & student_1 & student_2 & type \\ activity_1 & 1 & 0.6 & 1 & 1 & 0 & 1 \\ activity_2 & 0.4 & 0 & 0 & 1 & 0 & 0 \\ activity_3 & 0.3 & 0.7 & 0 & 1 & 0 & 1 \\ activity_1 & 1 & 0.6 & 0 & 0 & 1 & 1 \\ activity_2 & 0.4 & 0 & 0 & 0 & 1 & 0 \\ activity_3 & 0.3 & 0.7 & 1 & 0 & 1 & 1 \end{pmatrix}$$

FIGURE 12. Example of the Q-matrix used by eAFM (the column *type* can take two different values: 0 and 1 that indicates if the activity is a problem or a test respectively).

TABLE 4. Example of interactions of two different students with three activities. The columns *skill_1* and *skill_2* show the value of the score of each skill in the activities. The columns *student_1* and *student_2* show if each particular student has worked or not with each activity.

	skill_1	skill_2	student_1	student_2
activity_1	score=10	score=6	true	false
activity_2	score=4	-	false	false
activity_3	score=3	score=7	false	true

to understand if other components, such as the type of activities that students were interacting with (test and problems), or the type of contents they were exposed to (using the score parameter of the skills), can add more insights on the student learning process.

To this end, in the AFM equation (3) we have included two new parameters: A is the type of activity the student is working on (namely *type*) and W_j is the *score* parameter (i.e. the weight of the different skills on the activity).

So, given a group of N students, a set of K skills and a set of J activities or items, in this enhanced AFM (eAFM), the probability that a student $i \in \{1 \dots N\}$ gives a correct answer in an exercise, is expressed by equation (3):

$$\begin{aligned} p &= P(Y_{ij} = 1 | \alpha_i, \beta, \gamma) \\ &= f(\alpha_i + \gamma * \mathbb{1}_A(j) + \sum_{k=1}^K \beta_k * q_{jk} * W_k \\ &\quad + \sum_{k=1}^K \gamma_k * q_{jk} * T_{ik}) \\ f(x) &= \frac{1}{(1 + e^{-x})} \end{aligned} \quad (3)$$

where $A \subset J$ is the set of tests defined for the activities in IDEE and W_k is the score of skill k . Thus $\mathbb{1}_A(j)$ is 1 if $j \in A$ (i.e. if j is a test) and 0 otherwise. With this new model, the parameter γ denotes the difficulty of the type of exercise (problem or test).

Table 4 shows an example of how two different students have worked on three different activities.

With the traditional AFM we should have the same Q-matrix for *student_1* and *student_2*, therefore the model will determine the same learning for the two students (see Figure 11). With eAFM we have the Q-matrix shown in Figure 12. In this case, *student_1*, gave a correct answer in *activity_1*. Notice that the value of score parameter of *skill_1* for *activity_1* is 1. In this same activity, *student_2* did

TABLE 5. Comparison in terms of BIC, AIC and RMSE between AFM and eAFM.

MODEL	BIC	AIC	RMSE
AFM	875.11	679.55	0.4158
AFM + Type	869.90	670.00	0.4113
AFM + Type + Score (eAFM)	864.92	665.02	0.4105

not answer correctly, but he was right in *activity_3*. Nevertheless the value of score parameter for *skill_1* in *activity_3* is 0.3. So on *skill_1* the two students have a different learning.

The eAFM takes into account that *student_1* has a different amount of learning from *student_2* (something that was omitted with the previous model).

To check if this eAFM could be used to provide learning indicators as done in section IV-C, we have employed GLM using this new Q-matrix.

The *type* parameter is significant with $\gamma = 1.95$, this value indicates that the tests are on average easier than the problems. In fact, with the eAFM the significance of the user with ID 1529 is lost. This new analysis is in agreement with the data since this student does not show difficulties.

The eAFM model has been compared with the previous AFM in terms of the AIC (Akaike Information Criterion), the BIC (Bayesian Information Criterion) and the RMSE (Root Mean-Squared Error) with cross validation [50], [51]. In the Table 5 we show the results obtained. A decrease by 10 points (or more) in AIC is deemed to be a reasonable indicator of a better more suitable model according to prior practice [52]. However, further experimentation is needed to get a more solid evidence.

V. CONCLUSION AND FUTURE WORK

Usually, educational robotics has been employed in extracurricular activities isolated from classrooms. Providing support to teachers in educational robotics environments is crucial to promote the use of robotics in classrooms. In this paper we present IDEE, an educational environment that provides learning activities that uses robotics. In addition, we show how teachers can be supported in IDEE through the use of AFM, a well-known technique to represent the learning process in learning environments.

We show that the application of AFM may be useful for a teacher who uses data from experiences in physics classes that include theory, problems and laboratories. The results display valuable information about learning process. In our

work, the reliability of the estimation of the coefficients is a crucial aspect for model inference.

We also propose the eAFM whose results, presented in section IV-D, are encouraging. The score and the type of activity seem valid variables to allow a more accurate analysis of the students' learning process. As future line of work, we plan to keep working on the eAFM. We would like to test if a better assessment of the learning progress of the students can be obtained by taking into account the learning path of the students.

ACKNOWLEDGMENT

The authors want to acknowledge the Scuola Italiana in Madrid, the Santa Francisca Cabrini School, and the Gredos Las Suertes School for their support in the experiments with IDEE.

REFERENCES

- N. A. Bascou and M. Menekse, "Robotics in k-12 formal and informal learning environments: A review of literature," in *Proc. ASEE Annu. Conf.*, Berlin, Germany: Springer, 2016, p. 26119.
- A. Khanlari, "Effects of educational robots on learning STEM and on students' attitude toward STEM," in *Proc. IEEE 5th Conf. Eng. Edu. (ICEED)*, Dec. 2013, pp. 62–66.
- A. Eguchi, "RoboCupJunior for promoting STEM education, 21st century skills, and technological advancement through robotics competition," *Robot. Auto. Syst.*, vol. 75, pp. 692–699, Jan. 2016.
- E. Menegatti and M. Moro, "Educational robotics from high-school to master of science," in *Proc. Conf. Simulation, Modeling Program. Auto. Robots*, 2010, pp. 639–648.
- N. W. W. Hong, E. Chew, and J. W. Sze-Meng, "The review of educational robotics research and the need for real-world interaction analysis," in *Proc. 14th Int. Conf. Control, Autom., Robot. Vis. (ICARCV)*, Nov. 2016, pp. 1–6.
- N. Spolaôr and F. B. V. Benitti, "Robotics applications grounded in learning theories on tertiary education: A systematic review," *Comput. Edu.*, vol. 112, pp. 97–107, Sep. 2017.
- Q. Wu, S. Wang, J. Cao, B. He, C. Yu, and J. Zheng, "Object recognition-based second language learning educational robot system for chinese preschool children," *IEEE Access*, vol. 7, pp. 7301–7312, 2019.
- A. B. Williams, "The qualitative impact of using LEGO MINDSTORMS robots to teach computer engineering," *IEEE Trans. Educ.*, vol. 46, no. 1, p. 206, Feb. 2003.
- S. P. Krishnamoorthy and V. Kapila, "Using a visual programming environment and custom robots to learn c programming and K-12 STEM concepts," in *Proc. 6th Annu. Conf. Creativity Fabr. Edu.*, New York, NY, USA, 2016, pp. 41–48.
- K. E. Newman, J. O. Hamblen, and T. S. Hall, "An introductory digital design course using a low-cost autonomous robot," *IEEE Trans. Educ.*, vol. 45, no. 3, pp. 289–296, Aug. 2002.
- C.-C. Chang and Y. Chen, "The impact of an integrated robotics STEM course with a sailboat topic on high school Students' perceptions of integrative STEM, interest, and career orientation," *EURASIA J. Math., Sci. Technol. Edu.*, vol. 14, no. 12, p. 1614, Aug. 2018.
- L. Narbutaitė, R. Damasevicius, E. Kazanavicius, and S. Misra, "Using collaborative robotics as a way to engage students," in *Towards Extensible Adaptable Methods Comput.*, S. Chakraverty, A. Goel, and S. Misra, Eds. New York, NY, USA: Springer, 2018, pp. 385–397.
- I. Ahmed, N. Lubold, and E. Walker, "ROBIN: Using a programmable robot to provide feedback and encouragement on programming tasks," in *Artificial Intelligence in Education (Lecture Notes in Computer Science)*. New York, NY, USA: Springer-Verlag, 2018, pp. 9–13.
- S. Papert, *Mindstorms. Children, Computers and Powerful Ideas*. New York, NY, USA: Basic books, 1980.
- F. B. V. Benitti, "Exploring the educational potential of robotics in schools: A systematic review," *Comput. Edu.*, vol. 58, no. 3, pp. 978–988, Apr. 2012.
- L. Xia and B. Zhong, "A systematic review on teaching and learning robotics content knowledge in K-12," *Comput. Edu.*, vol. 127, pp. 267–282, Dec. 2018.
- A. Ioannou, C. Socratous, and E. Nikolaedou, "Expanding the curricular space with educational robotics: A creative course on road safety," in *Lifelong Technology-Enhanced Learning (Lecture Notes in Computer Science)*, V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachler, R. Elferink, and M. Scheffel, Eds. New York, NY, USA: Springer, 2018, pp. 537–547.
- M. Venturini, "How teachers think about the role of digital technologies in student assessment in mathematics," Ph.D. dissertation, Faculty Educ. Simon Fraser Univ. (Canada) and Faculty Math. Univ. Bologna, Italy, 2015. Accessed: Feb. 5, 2020. [Online]. Available: <http://summit.sfu.ca/item/15703>
- D. Alimisis, "Educational robotics: Open questions and new challenges," *Themes Sci., Technol. Educ.*, vol. 6, no. 1, pp. 63–71, 2013.
- I. Jormanainen and E. Sutinen, "Role blending in a learning environment supports facilitation in a robotics class," *Educ. Technol. Soc.*, vol. 17, no. 1, pp. 294–306, 2014.
- A. Bogarín, R. Cerezo, and C. Romero, "A survey on educational process mining," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 8, no. 1, p. e1230, Jan. 2018.
- S. Orlando, E. Gaudio, and F. de la Paz, "IDEE: A visual programming environment to teach physics through robotics in secondary schools," in *Robotics in Education*. New York, NY, USA: Springer, 2018, pp. 241–246.
- S. Orlando, F. de la Paz, and E. Gaudio, "Design and implementation of a robotics learning environment to teach physics in secondary schools," in *Proc. 8th Int. Work-Confer. Interplay Between Natural Artif. Comput., (IWINAC)*, Almería, Spain, Jun. 2019, pp. 69–76.
- H. Cen, K. Koedinger, and B. Junker, "Learning factors analysis: a general method for cognitive model evaluation and improvement," in *Intelligent Tutoring Systems*, M. Ikeda, K. D. Ashley, and T. Chan, Eds. Berlin, Germany: Springer, 2006, pp. 164–175.
- J. Nouri, L. Zhang, L. Mannila, and E. Norén, "Development of computational thinking, digital competence and 21st century skills when learning programming in K-9," *Edu. Inquiry*, vol. 11, pp. 1–17, Jun. 2019.
- F. J. García-Peñalvo, A. M. Rees, J. Hughes, I. Jormanainen, T. Toivonen, and J. Vermeersch, "A survey of resources for introducing coding into schools," in *Proc. 4th Int. Conf. Technol. Ecosystems Enhancing Multiculturalism*, New York, NY, USA, 2016, pp. 19–26.
- M. Armoni, S. O. Meerbaum, and A. M. Ben, "From scratch to 'real' programming," *ACM Trans. Comput. Educ.*, vol. 14, no. 4, pp. 1–15, 2015.
- D. Y. T. W. Price and L. D., "ISnap: Towards intelligent tutoring in novice programming environments," in *Proc. ACM SIGCSE Tech. Symp. Comput. Sci. Educ.* New York, NY, USA: Springer, 2017, pp. 483–488.
- M. Ketterl, B. Jost, T. Leimbach, and R. Budde, "Tema 2: Open Roberta - a Web based approach to visually program real educational robots," *Tidskriften Lärning Medier*, vol. 8, no. 14, pp. 1–22, Jan. 2015.
- A. Milašinić, B. Anelić, L. Pushkar, and A. S. Kržić, "Using robots as an educational tool in native language lesson," in *Robot. Edu.*, M. Merdan, W. Lepuschitz, G. Koppensteiner, R. Balogh, and D. Obdržálek, Eds. Springer, 2019, pp. 296–301.
- J. E. Michaelis and B. Mutlu, "Supporting interest in science learning with a social robot," in *Proc. Interact. Design Children (IDC)*, New York, NY, USA, 2019, pp. 71–82.
- P. Ferrarelli, W. Villa, M. Attolini, D. Cesareni, F. Micale, N. Sansone, L. Pantaleone, and L. Iocchi, "Improving students' concepts about newtonian mechanics using mobile robots," in *Robot. Edu.*, W. Lepuschitz, M. Merdan, G. Koppensteiner, R. Balogh, and D. Obdržálek, Eds. Cham, Switzerland: Springer, 2019, pp. 113–124.
- A. Badeleh, "The effects of robotics training on students' creativity and learning in physics," *Edu. Inf. Technol.*, to be published, doi: [10.1007/s10639-019-09972-6](https://doi.org/10.1007/s10639-019-09972-6).
- A. Savard and V. Freiman, "Investigating complexity to assess student learning from a robotics-based task," *Digit. Experiences Math. Edu.*, vol. 2, no. 2, pp. 93–114, Sep. 2016.
- O. Zughoul, F. Momani, O. H. Almasri, A. A. Zaidan, B. B. Zaidan, M. A. Alsalem, O. S. Albahri, A. S. Albahri, and M. Hashim, "Comprehensive insights into the criteria of student performance in various educational domains," *IEEE Access*, vol. 6, pp. 73245–73264, 2018.
- A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," *IEEE Access*, vol. 5, pp. 15991–16005, 2017.

- [37] E. Emerson, A. Smith, C. Smith, F. Rodríguez, W. Min, E. Wiebe, B. Mott, K. Boyer, and J. Lester, "Predicting early and often: Predictive student modeling for block-based programming environments," in *Proc. 12th Int. Conf. Educ. Data Mining*, Montreal, QC, Canada, 2019, p. 48.
- [38] Y. Mao, R. Zhi, F. Khoshnevisan, T. Price, T. Barnes, and M. Chi, "One minute is enough: Early prediction of student success and event-level difficulty during novice programming tasks," in *Proc. 12th Int. Conf. Educ. Data Mining*, Montreal, QC, Canada, 2019, pp. 119–128.
- [39] R. Al-Shabandar, A. J. Hussain, P. Liatsis, and R. Keight, "Analyzing learners behavior in MOOCs: An examination of performance and motivation using a data-driven approach," *IEEE Access*, vol. 6, pp. 73669–73685, 2018.
- [40] E. Gaudio, M. Montero, L. Talavera, and F. Hernandez-del-Olmo, "Supporting teachers in collaborative student modeling: A framework and an implementation," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2260–2265, Mar. 2009.
- [41] N. Diana, M. Eagle, J. Stamper, S. Grover, M. Bienkowski, and S. Basu, "An instructor dashboard for real-time analytics in interactive programming assignments," in *Proc. 7th Int. Learn. Anal. Knowl. Conf. (LAK)*, New York, NY, USA, 2017, pp. 272–279.
- [42] I. Jormanainen, M. Beynon, and E. Sutinen, "An Abductive environment enables Teacher's intervention in a robotics class," in *Proc. 17th Int. Symp. Artif. Life Robot.*, 2015, p. 121.
- [43] G. G. C. Durand and N. Y. S. Belacel Bouslimani Léger, "Review, computation and application of the additive factor model (AFM)," Nat. Res. Council Canada, Winnipeg, MB, Canada, Tech. Rep. 23002483, 2017.
- [44] K. Koedinger, R. Baker, K. Cunningham, A. Skogsholm, B. Leber, and J. Stamper, "A data repository for the edm community: The PSLC datashop," in *Handbook Education Data Mining*, C. Romero, S. Ventura, M. Pechenizkiy, and R. Baker, Eds. Boca Raton, FL, USA: CRC Press, 2010, pp. 43–53.
- [45] B. Martin, A. Mitrovic, K. R. Koedinger, and S. Mathan, "Evaluating and improving adaptive educational systems with learning curves," *User Model. User-Adapted Interact.*, vol. 21, no. 3, pp. 249–283, Aug. 2011.
- [46] T. Toivonen, I. Jormanainen, and M. Tukiainen, "An open robotics environment motivates students to learn the key concepts of artificial neural networks and reinforcement learning," in *Advances in Intelligent Systems and Computing*. New York, NY, USA: Springer, 2017.
- [47] E. Afari, U. the Petroleum Institute Abu Dhabi, and M. S. Khine, "Robotics as an educational tool: Impact of Lego mindstorms," *Int. J. Inf. Edu. Technol.*, vol. 7, no. 6, pp. 437–442, 2017.
- [48] J. Stamper and K. R. Koedinger, "Human-machine student model discovery and improvement using data," in *Proc. 15th Artif. Intell. Educ.* Berlin, Germany: Springer-Verlag, 2011, pp. 353–360.
- [49] S. Seabold and J. Perktold, "StatsModels: Econometric and statistical modeling with Python," in *Proc. 9th Python Sci. Conf.*, S. van der Walt and J. Millman, Eds., Austin, TX, USA, 2010, pp. 57–61.
- [50] J. Stamper, K. R. Koedinger, and E. McLaughlin, "A comparison of model selection metrics in datashop," in *Proc. Educ. Data Mining*, 2013, pp. 284–287.
- [51] R. Liu and K. Koedinger, "Variations in learning rate: Student classification based on systematic residual error patterns across practice opportunities," in *Proc. Educ. Data Mining Soc.*, Jun. 2015, pp. 420–423.
- [52] K. Burnham and D. Anderson, "Multimodel inference: Understanding AIC and BIC in model selection," in *Sociol. Methods Res.*, vol. 33, pp. 261–304, Nov. 2004.



SAMANTHA ORLANDO received the B.Sc. degree in mathematics from the University of Salerno, Italy, and the double master's degree for teaching mathematics and physics in secondary schools. She is currently pursuing the Ph.D. degree in intelligent systems with the Department of Artificial Intelligence, National Distance University of Spain (UNED), Madrid, Spain. She is a secondary school Teacher in Italy, and recognized in Spain, with a long teaching experience. Her research focuses on adaptive educational systems and the use of robotics in learning environments. Her main aim is to improve the learning process of the students by introducing in class new technologies that can guide and support the teacher.



ELENA GAUDIO received the B.Sc. degree in mathematics from the Universidad Complutense de Madrid, Madrid, Spain, in 1996, and the Ph.D. degree in sciences from the National Distance University of Spain (UNED), Madrid, in 2003. She is currently an Associate Professor with the Department of Artificial Intelligence, UNED. She has authored or coauthored more than 30 articles about educational and adaptive systems in International journals and conferences. Her research interests include educational systems, recommender systems, adaptive systems, and machine learning.



FELIX DE LA PAZ received the B.Sc. and M.Sc. degrees in physics from the Universidad Complutense de Madrid, Spain, in 1995, and the Ph.D. degree in physics from the Universidad Nacional de Educacion a Distancia, Madrid, Spain, in 2003. He is currently an Associate Professor with the Artificial Intelligence Department, Universidad Nacional de Educacion a Distancia. His main research interests include navigation and representation tasks in mobile robots, human-robot interaction, and social robotics.

...