

In-Detail Analysis on Custom Teaching and Learning Framework

Subhabrata Sengupta
Institute of Engineering and
Management
Kolkata, India

Anish Banerjee
Institute of Engineering and
Management
Kolkata, India

Satyajit Chakrabarti
Institute of Engineering and
Management
Kolkata, India

ABSTRACT

In the field of Information recovery, the primary goal is to discover importance just as the most significant data as for certain inquiries. In any case, the primary issue with respect to recovery has consistently been, that the pursuit region is tremendous to such an extent that it has gotten hard to recover relevant data productively. It has been seen that the conventional ontological authoritative data causes superfluous additional CPU cost, while the client inquiries, for the most part, focus on a particular space. What's more, another difficult issue in such manner is the key expression extraction from the question which has likewise a significant job for pinpointing looking to a particular recovery space. By centering, these restrictions and difficulties, we have focused on our data recovery framework, especially towards assessment question recovery so as to take into account the interest of different assessment-related inquiries. The inquiry data has been composed according to the ontological relationship among different classes and a characteristic language parser will be utilized during key-phrase extraction for proficient recovery of inquiries most ideally requested as for the level of importance to the questions. Open learning analytics (OLA) is a moderately new part of learning analytics (LA) which rose because of the developing interest for self-sorted out, organized, and long-lasting learning opportunities. In this paper, we present the goal - question - indicator (GQI) approach for PLA and give the applied, structure, usage and assessment subtleties of the pointer motor segment of the open learning analytics platform (OpenLAP) that draws in end-clients in the pointer age process by supporting them in defining objectives, offering conversation starters, and self-characterizing pointers..

General Terms

Your general terms must be any term which can be used for general classification of the submitted material such as Pattern Recognition, Security, and Algorithms et.al.

Keywords

Learning Analytics, Open Learning Analytics, Personalized Learning Analytics, OpenLAP, Semantic web, Search Engine, Personalization, NLP, Web-Link Categorization, Parse Tree

1. INTRODUCTION

Till now whatever work has been done in the related field it has been accomplished either for prescient content age or for the field of textual similarity. We will utilize both these work progress to build up our framework that will procedure an inquiry given as an information message and will be thinking of labels at that point will contrast its semantic worth and all the inquiries with the comparative labels and afterwards will be relegated the weight-age of the inquiry its nearest to. This

will be a progressive work of computerization in the field of E-Learning. Learning Analytics (LA) is a rising examination field that manages the turn of events and improvement of techniques to help the learning procedure by misusing the instructive information. As of late, another examination region has risen called Open Learning Analytics (OLA) which centres around increasingly open, self-sorted out and arranged learning situations. So as to recover information from an information base (KB), information labourers, for example, doctors or money related investigators, frequently face the test of learning explicit inquiry dialects (e.g., SQL and SPARQL1). Notwithstanding, the quick pace of changing inquiry dialects to various sorts of KBs (e.g., Relational Databases, Triple Stores, NoSQL stores, and so on.) makes it hard for clients to stay aware of the most recent improvements of such question dialects that permit them to get to the information they require for their work. This circumstance forestalls clients without broad PC preparing from successfully using the accessible data in the KB. Creating easy to use characteristic language interfaces will make it simpler for non-specialized clients to get to the data in the KB in an instinctive way. In this paper, we are proposing a system of a Natural Language Interface that permits clients to inquire the basic KBs with normal language questions. Rather than requesting that the clients give the whole inquiry all alone, our framework will make recommendations to assist the clients with completing their inquiries. Given a total inquiry, our framework will be parsed to its First Order Logic (FOL) portrayal utilizing a sentence structure got from interlinked datasets; at long last, our framework produces dynamic examination for the outcome sets so as to assist clients with gaining a superior comprehension of the subject just as the particular questions.

2. EXTENSIVE LITERATURE REVIEW

The test of customized LA and present the reasonable, structure, and usage subtleties of a standard based marker definition apparatus to help the adaptable definition and dynamic age of pointers to address the issues of various partners with different objectives and inquiries in the LA practice was tended to. [1]

The favourable circumstances, applications and difficulties of utilizing profound taking in methods for picking up bits of knowledge from the information produced from the IoE gadgets inside the instructive space for making such learning examination frameworks have been advanced. At last, an element astute correlation is given between the proposed Learning Analytics (LA) based methodology and ordinary instructing learning approach as far as execution parameters like cognizance, consideration, maintenance and accomplishment of students. [2]

At present, informative data mining will in general spotlight on growing new devices for finding designs in the information. These examples are commonly about the smaller scale ideas engaged with learning: one digit duplication, a deduction with conveys, etc. Learning examination—at any rate as it is as of now stood out from information mining—centers around applying instruments and methods at bigger scopes, for example, in courses and at schools and post-secondary organizations. Be that as it may, the two controls work with examples and expectation: If we can observe the example in the information and comprehend what's going on, we can anticipate what should come straight away and make the proper move. [3]

AI is worried about an enormous number of calculations that improve their exhibition with experience, in numerous fields of research, for example, those learning settings where understudies cooperate with learning frameworks leaving helpful tracks. Instructive Data Mining is the study of separating helpful data from the huge informational indexes or databases containing understudies communications during their learning, for instance in a virtual domain. At long last, Learning Analytics is a lot of steps for comprehension and enhancing the entire learning process, along with nature in which it happens. It is made by a few stages, where the first is carefully identified with Educational Data Mining for catching information by some AI calculations. In this paper, we talk about the convergences and connections between these three regions of research, attempting to examine their connections and steps to give a helpful outline on the taking in forms from various purposes of perspectives. Various models are presented and talked about. [4]

An algorithm, in view of representative relapse that utilizes non-exploratory information on past scores gathered by the college as information has been tried. It can foresee 60–70 per cent of variety in understudies' test scores. Applying the calculation to assess the effect of instructing techniques in a conventional differential conditions class, it was found that clickers were an increasingly compelling showing procedure when contrasted with customary written by hand schoolwork; nonetheless, online schoolwork with prompt criticism was seen as significantly more viable than clickers. [5]

The essential point of the advancement of eLAT is to process enormous informational collections in microseconds as to singular information examination premiums of educators and information protection issues, so as to push them to self-think about their innovation upgraded instructing and learning situations and to distinguish open doors for mediations and enhancements. [6]

Multimodal learning investigation (MMLA) can offer novel strategies that can create one of a kind data about what happens when understudies are occupied with community-oriented, venture-based learning exercises. Using multimodal learning investigation stage, we gathered different surges of information, handled and removed multimodal connections to address the accompanying inquiry: which highlights of MMLA are acceptable indicators of cooperative critical thinking in open-finished errands in venture-based learning? Manually entered scores of CPS were relapsed utilizing AI strategies. The response to the inquiry gives potential approaches to naturally recognize parts of a joint effort in venture-based learning. [7]

Current executions of LA principally depend on a predefined set of inquiries and pointers which isn't reasonable with regards to OLA where the markers are erratic. In this paper, we present

the objective - question - pointer (GQI) approach for PLA and give the theoretical, structure, usage and assessment subtleties of the marker motor segment of the OpenLAP that draw in end-clients in the pointer age process by supporting them in defining objectives, suggesting conversation starters, and self-characterizing pointers. [8]

New and promising methodologies, for example, neural systems, and progressively conventional relapse approaches can both be utilized to order multimodal learning examination information, and both have favourable circumstances and inconveniences relying upon the exploration questions and settings being researched. The work introduced here is a critical commitment towards creating strategies to consequently recognize the key parts of understudies' accomplishment in venture-based learning situations and to eventually assist instructors with offering proper and opportune help to understudies in these major angles. [9]

Past frameworks for characteristic language inquiries over complex connected datasets require the client to enter a total and very much shaped inquiry and present the appropriate responses as crude arrangements of elements. Utilizing a component-based language with a full conventional semantics, we have built up a framework that can bolster rich autosuggest and to convey progressively produced investigation for each outcome that it returns. [10]

As understudies cooperate with the different learning advances in their course of study, they make computerized follows that can be caught and examined. These computerized follow the structure of the new sort of information that is as often as possible utilized in learning investigation to create noteworthy suggestions that can bolster understudy learning. This paper investigates the utilization of such examination to address the moves obstructing the limit of teachers to give customized input at scale. The contextual investigation announced in the paper demonstrated how the methodology was related to a positive effect on understudy impression of input quality and on scholarly accomplishment. The investigation was led with first-year undergrad designing understudies took a crack at PC frameworks. [11]

3. PROBLEM STATEMENT

3.1 LA Model

In this area, fundamental traits of LA (Learning Analytics) model will be presented. Essential thought of LA model is to improve the learning and showing process by utilizing diverse instrumental investigation. The Society for Learning Analytics Research [12] defines LA as Learning Analytics is the estimation, assortment, examination, and revealing of information about students and their specific circumstances, for the motivations behind comprehension and upgrading learning and the earth in which it happens.

3.1.1 The 5-steps Oblinger LA Model:

Proposed in [14] in 2007, this model is the first LA model and has been represented the starting point for planning LA real systems. It is characterized by a set of questions, grouped in the following five steps, as shown in Fig. 1:

3.1.1.1 The Capture step

This step is the one where ML is involved to gather useful data. In fact, in this step, one has to answer some questions concerning what data is being collected, how frequently is the data collected, where is the data going to be stored, which format is going to be used to represent all events and if the observations securely stored;

3.1.1.2 The Report step

This step addresses the information flow collected in the first step with the aim of producing some reports. So the questions concern who and how frequently will receive the reports, what kind of information needs to be reported and how the reports will be accessed;

3.1.1.3 The Predict step

This step concerns the prediction algorithms: their accuracy, their particular prediction, their input data and finally what predictions are reported to stakeholders.

3.1.1.4 The Act step

This step concerns the actions to be deployed into the learning environment to improve the overall learning process;

3.1.1.5 The Refine step

In this step, the questions concern a reasoning about the results obtained in the previous steps, included their usefulness.

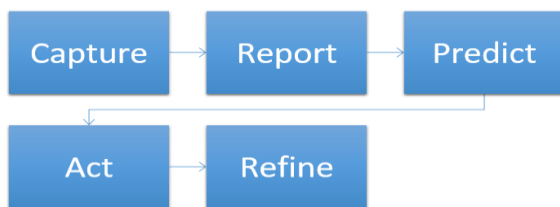


Fig. 1. The 5-steps LA model [14]

3.1.2 The Learning Analytics Cycle LA Model

Proposed in 2012 by D.Clow [15], this model changes the past model structure a hypothetically grounded Learning Analytics Cycle, where a significant pretends the last criticism to the framework. The model is made by four stages, as appeared in Fig. 2. Students are the members to a course, casual students in a MOOC, etc. The subsequent stage Data, i.e., is the age and catches of information which can be created during the learning procedure. After there is the Metrics step, which gives a few experiences into the learning procedure. This progression requires a product framework ready to envision data for the administration of the course, i.e., a successful dashboard so as to comprehend the learning procedure pattern. The last advance is the Intervention step where an input to improve the learning procedure is expected.

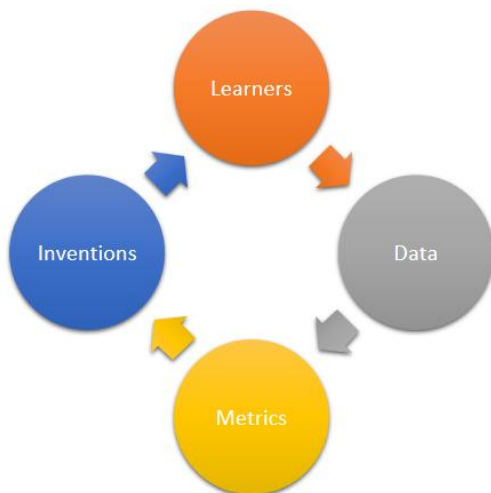


Fig. 2. The Learning Analytic Cycle [15]

3.2 The Teacher-Inquiry Model

As of late, another field of research has been proposed as supplementing cooperative energy with LA, called Teaching Analytics (TA) [16] [17]: Instructing Analytics alludes to the techniques and advanced devices to assist instructors with breaking down and improve the instructive plans before the conveyance. Besides, later advancements on Teaching Analytics additionally bolster examination of how the educator conveys the instructive structures. Considering this perspective too, Sergis and Sampson in 2017 [18], proposed another model, which speaks to incorporation among LA and TA, appeared in Fig. 3, and increasingly appropriate to help instructor request. This model is spoken to by the accompanying 6-steps cycle.

In the first step, Problem Identification, the educator distinguishes a few issues to improve the learning procedure.

In the second step, Develop Inquiry Questions, the instructor creates request questions; what is to explore, which instructive information is to be accumulated.

The third step, Educational Design, the instructor plans the instructive structure to be conveyed to begin the request.

In step four, Deliver Educational Design and gather information, the educator conveys the instructive plan to the students and gathers the instructive information.

In sync five, Analyze instructive information, the educator examinations instructive information to evoke bits of knowledge to address the request characterized question.

At long last, in sync six, Reflect on Data, the instructor overhauls the learning procedure to respond to the characterized request question.

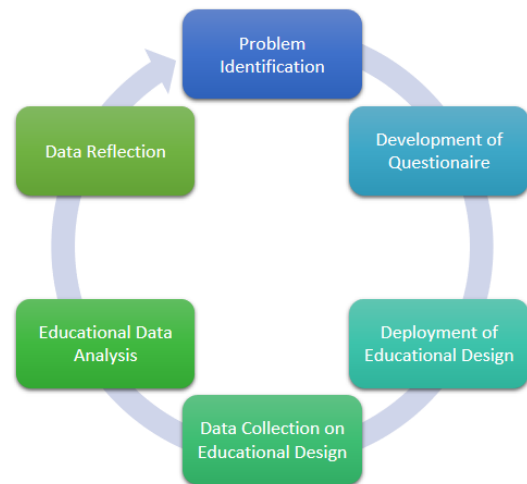


Fig 3. The Teacher Inquiry model [18]

3.3 Extracting Knowledge from Data

This segment delivers the methods used to remove information from information, following a great base up approach. Here two research zones are engaged: Machine Learning and Data Mining.

3.4 Machine Learning

Tom M. Mitchell gave the most referred to the meaning of Machine Learning in his book Machine Learning [19]. The writing proposes numerous calculations that fall under this definition: choice trees, fake neural systems, Bayesian systems, Instance-based calculations, grouping and hereditary

calculations just to make reference to the most utilized. A typical trait of all ML calculations is that they run on and gain from information, to discover rules, concealed examples and to anticipate future practices. For instance, Artificial Neural Networks attempt to emulate neurons for acknowledgement and order. ML algorithms have been utilized in instructive zones, for example, Technology Enhanced Learning (TEL) to foresee understudies' practices. In crafted by Zafra and Ventura [20], the creators utilize the Multiple Instance-Based ML calculations to foresee understudies' presentation dependent on work they have done on a Virtual Learning Platform. In [21], information separated from the Moodle LSM are utilized to deduce the learning styles of the students, utilizing grouping strategies. In [22] and in [23], Bayesian Networks K-NN Instance-Based calculations are utilized to construe the evaluations in a friend appraisal condition. Along these lines, ML strategies and calculations are predominantly utilized in instructive settings to foresee future practices.

3.5 Data Mining

Data Mining can be characterized likewise as the procedure engaged with removing fascinating, interpretative, valuable and novel data from information [24]. It has been utilized for a long time by organizations, researchers and governments to filter through volumes of information like carrier traveller records, registration information and the grocery store scanner information that produces statistical surveying reports. So the primary attribute of this region of research is that it needs a lot of information to be viable. A significant model is the use of DM to web-based business. These days online business sites store countless tracks speaking to guests' activities: from this information, organizations attempt to demonstrate clients so to be increasingly viable in focused publicizing to build their business. To this point, Google proposes Google investigation which takes into consideration examination about clients. Facebook and Amazon are other two instances of the utilization of DM to huge arrangements of unstructured information, put away in log documents.

4. PROPOSED MODEL

The main aim in modifying or in changing the way E Learning model is perceived in the modern world. Machine Learning, Deep learning and various other native and complex integrated Machine Learning models made it possible to automate nearly everything. The aim of ours is to automate the examination and the Question analysis part of the E Learning system. The various parts of it consist of:

1. Question Similarity Determination
2. Question Analytics
3. Question Weightage determination

4.1 Question Similarity Determination

The auto-suggestion is based on the idea of left corner parsing. Given a query segment as it finds all grammar rules whose left corner i.e. on the right side matches the left side of the lexical entry of questions. Then all leaf nodes will be traced in the grammar that can be reached by using the adjacent element of. For all reachable leaf nodes (i.e., lexical entries in our grammar), if a lexical entry also satisfies all the linguistic constraints, then it will be treated as a valid suggestion.

4.2 Question Analytics

Instead of only retrieving a list of entities, our system provides several different types of analytics for different result sets. In many situations, the result is a set of records rather

than one single entry. This provides us the opportunity to perform and provide further analyses of the result set for the users. Our system provides several types of analytics. Descriptive analytics summarize the facts in the result set. For instance, for the question "What is Pointer?", our system will describe the distribution of all concepts used for such question in the result set. We also compare the keywords in the result set on different dimensions (e.g., Programming Language, Chapters, and Modules etc). Moreover, we compute trends via exponential smoothing for entities that have a temporal dimension. We might adopt the Stanford CoreNLP toolkit for recognizing person, organization, and location from the news corpus. Given an entity, we show its frequency count and how its sentiment may change over time. This information may provide further insights to users in order to support their own analysis. This Analytical overview and understanding of a question's sentimental representation will help out in getting its emotional value and an understanding of its sentimental comparison with other questions.

4.3 Question Weightage Determination

Work-Flow of this process is as follows:

4.3.1 Step1: Text Pre-processing

Here the Text or question given goes through basic pre-processing like removal of stop-words, special characters, concerting the uppercase characters to lower case.

4.3.2 Tagging and Tag-Finding

Here we search for tags in the processed text from a group of pre-defined list of tags for different sets of fields of question. When found then tags are associated to the text.

4.3.3 Searching for the Similar Question

The processed text is compared to all the questions in the dataset by taking the questions from there individually with respect to similar tagging and their semantic value is compared and the question with maximum resemblance is stored. Weight-age of the stored question is assigned to the original question given as input and is appended to the section.

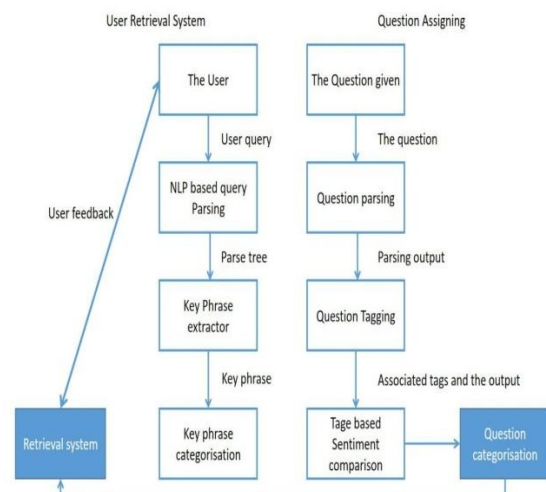


Fig. 4. Proposed System Architecture

In the proposed system the question is analyzed using various NLP tools and the is judged and compared on various parameters to get the essence of the question and the complexity. The testing system build to analyse the questions in the IT and Computer Science domain when fed with the

question “How to find duplicate number on Integer array in Java” yields the following results.

```

... Enter The Question: How to find duplicate number on Integer array in Java

Post Evaluation Weightage determined: 6

Data Structure and Algorithm : 1
Programming Languages : 1
Code/Error Analysis : 0
    
```

Fig. 5. Weightage Given to the Question

The parameter the the question covers of the topics it depict is also shown using a horizontal bar chart.

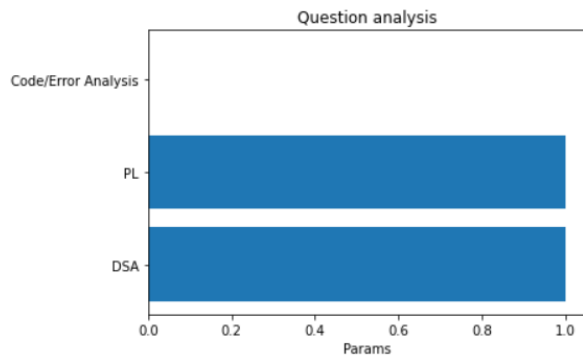


Fig. 6. Parameter Tagging on Input Question

The word clouds of the question help us to analyse the essence of the input question and also proves the accuracy of the model.



Fig. 7. Word Cloud without word restriction



Fig. 8. Word Cloud with restriction on Tagged words

5. CONCLUSIONS AND FUTURE DIRECTIONS

Open learning examination (OLA) is a moderately new part of learning investigation (LA) which rose because of the developing interest for self-composed, organized, and deep-rooted learning openings. OLA manages to take in

information gathered from different learning situations and settings, investigated with a scope of examination techniques, and for various partners with assorted premiums and destinations. This assorted variety in various components of OLA is a test which should be tended to by embracing a customized learning investigation (PLA) model. Current executions of LA depend on a predefined set of inquiries and pointers which isn't appropriate with regards to OLA where the markers are erratic. This paper presents the goal - question - indicator (GQI) approach for PLA and give the reason, structure, and execution and assessment subtleties of the pointer motor segment

6. ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Satyajit Chakrabarti, Director, Institute of Engineering & Management, for his valuable and constructive suggestions and overall help in conducting this research.

I would also like to thank Mr. Anish Banerjee, Student, Institute of Engineering & Management, for his great support during the planning and development of this research work.

7. REFERENCES

- [1] Arham Muslim, et al. ,”A Rule-Based Indicator Definition Tool for Personalized Learning Analytics”, Ahad et al. Smart Learning Environments (2018)
- [2] Marie Bienkowski et. Al, “Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief”,U.S. Department of Education Office of Educational Technology,2012
- [3] Filippo Sciarrone et. Al. “Machine Learning and Learning Analytics: Integrating Data with Learning”, IEEE, 2015
- [4] Fedor Duzhin et. Al. ,”Machine Learning-Based App for Self-Evaluation of Teacher-Specific Instructional Style and Tools”, Educ. Sci. 2018
- [5] Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, “Design and Implementation of a Learning Analytics Toolkit for Teachers.” Educational Technology & Society,2015.
- [6] Daniel Spikol et. Al. “Using Multimodal Learning Analytics to Identify Aspects of Collaboration in Project-Based Learning”, CSCL 2017 Proceedings
- [7] Muslim, A., Chatti, M., Mughal, M. and Schroeder, U. “The Goal - Question - Indicator Approach for Personalized Learning Analytics”, In Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017)
- [8] Daniel Spikol et. Al. “Supervised machine learning in multimodal learning analytics for estimating success in project-based learning”, J Comput Assist Learn. 2018.
- [9] Dezhao Song, “Natural Language Question Answering and Analytics for Diverse and Interlinked Datasets”, Proceedings of NAACL-HLT 2015
- [10] Pardo et. Al. “Using learning analytics to scale the provision of personalised feedback”, British Educational Research Association, 2017
- [11] A. Zafra and S. Ventura, “Predicting student grades in learning management systems with multiple instance learning genetic programming,” in Educational Data Mining - EDM 2009, Cordoba, Spain, July 1-3, 2009.

- Proceedings of the 2nd International Conference on Educational Data Mining., 2009, pp. 309–318. [Online]. Available: <http://www.educationaldatamining.org/EDM2009/uploads/proceedings/zafra.pdf>
- [12] R. Ferguson, “The state of learning analytics in 2012: a review and future challenges,” Tech. Rep. KMI-12-01, vol. 4, 2012.
- [13] J. P. Campbel, P. B. DeBlois, and D. G. Oblinger, “Academic analytics: A new tool for a new era,” *EDUCAUSE Review*, vol. 42, pp. 40–57, 1996.
- [14] D. Clow, “The learning analytics cycle: Closing the loop effectively,” in *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, ser. LAK '12. New York, NY, USA: ACM, 2012, pp. 134–138. [Online]. Available: <http://doi.acm.org/10.1145/2330601.2330636>
- [15] G. Gauthier, “Using teaching analytic to inform assessment practices in technology mediated problem solving tasks,” in *Proceeding of Workshop on Teaching Analytics at the 3rd Conference on Learning Analytics and Knowledge LAK 2013*, April 2013, pp. 1–8.
- [16] L. P. Prieto, S. Villagra, V. I. M. J. Abelln, A. Martnez-Mons, and Y. Dimitriadis, “Recurrent routines: Analyzing and supporting orchestration in technology-enhanced primary classrooms,” *Computers & Education*, vol. 57, pp. 1214–1227, 08 2011.
- [17] S. Sergis and D. Sampson, “Teaching and learning analytics to support teacher inquiry: a systematic literature review,” in *Learning Analytics. From Research to Practice*, A. Pea-Ayala, Ed. Berlin: Springer, 2017, pp. 25–63.
- [18] T. Mitchell, *Machine Learning*. Mc Graw-Hill International Editions, 1997.
- [19] A. Zafra and S. Ventura, “Predicting student grades in learning management systems with multiple instance learning genetic programming,” in *Educational Data Mining - EDM 2009*, Cordoba, Spain, July 1-3, 2009. Proceedings of the 2nd International Conference on Educational Data Mining., 2009, pp. 309–318. [Online]. Available: <http://www.educationaldatamining.org/EDM2009/uploads/proceedings/zafra.pdf>
- [20] V. Efrati, C. Limongelli, and F. Sciarrone, “A data mining approach to the analysis of students’ learning styles in an e-learning community: A case study,” in *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge*, C. Stephanidis and M. Antona, Eds. Springer International Publishing, 2014, pp. 289–300.
- [21] M. De Marsico, F. Sciarrone, A. Sterbini, and M. Temperini, “Supporting mediated peer-evaluation to grade answers to open-ended questions,” *EURASIA Journal of Mathematics Science and Technology Education*, vol. 13, no. 4, 2017.
- [22] A. Sterbini and M. Temperini, “Correcting open-answer questionnaires through a bayesian-network model of peer-based assessment,” in *Proceedings of the International Conference on Information Technology Based Higher Education and Training, ITHET 2012*, 2012.
- [23] U. Fayyad, G. Piatetsky-shapiro, and P. Smyth, “From data mining to knowledge discovery in databases,” *American Association for Artificial Intelligence*, vol. 17, pp. 37–54, 1996.
- [24] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2006.