A Systematic Mapping Study of Educational Technologies based on Educational Data Mining and Learning Analytics

Edona Doko, Lejla Abazi Bexheti
Faculty of Contemporary Sciences and Technologies
SEE University, Tetovo, Macedonia
ed15197@seeu.edu.mk; l.abazi@seeu.edu.mk

Abstract—**The traditional flipped classroom is defined based on the video instructions of flipped class at home. In addition, students have time in class to work with their teacher on key learning activities. With introduction of Flipped Classroom model and rapid development of data sciences for taking the advantages of the learners' data to optimize learning process there has been an increasing consideration of exploiting Educational Data Mining and Learning Analysis. Large amount of data taken from performed actions of students and teachers in FC has made data-driven learning and self-paced learning possible to perform different algorithms of Learning Analytics and Educational Data Mining. The goal in this paper is to devise a systematic mapping study, in order to explore and analyze existing researches about implementations of learning strategies and learning algorithms. We selected 122 papers and assorted according to field of interest, video field of interests and technology used in FC.**

Keywords-flipped classroom (FC); learning analytics (LA); educational data mining (EDM)

I. INTRODUCTION

The late progress in technology and ideology has opened a completely new direction for education research. FC has been gaining popularity and as the biggest advantage mentioned by students are the video learning interactivities [3]. That is because it offers a new approach to teaching and learning in education. When it comes to increased use of interactive learning via videos in FC there are many discussions, misunderstanding or challenges to identify the topics that students have more difficulties with. Based on this premise, we are focusing our research on analyzing EDM and LA articles to come up with better learning strategies in the future. Educational technologies have big influence in the realization of better learning strategies in FC. In this paper, the methodology used and the defined research questions is introduced first. Secondly, the classification scheme of field of interest, video field of interest and technology used in FC is presented. Afterwards five research questions are answered and two are proposed as future aim research. On the discussion part the time series of papers according to video field of interest is included. Data-driven learning system promoting self-paced learning via videos are proposed for future research areas.

II. METHODOLOGY

The main goal of this mapping study is to define and answer the research questions based on analyzed articles. After the conducted search the relevant papers are selected and classification scheme is defined. We have answered the research questions as result of the systematic map and outcome of the process. It often provides a visual summary, a map, of its results [7]. First, are collected all publications needed for the interested field. At the same time an overview of this research area is provided, the quantity, type of research and available results are identified. The second step is the conduct search for primary studies, and excluded are the studies that are not relevant to answer. Here the research questions to drive the structure are provided. The third step is ensuring that the scheme takes the existing studies into account and providing better results. Answering the research questions provides data extraction and complete mapping of studies, by identifying, analyzing and interpreting the relevant evidence. The classification scheme drives the field of interests, video field of interest and technology used in FC.

A. Research Questions and Search Strategy

The aim of our study is to analyze publications of some educational technologies based on some research questions:

- What are the main fields of interest investigated in the papers?
- How have analyzed publications in this research field changed over time?
- What kind of educational technologies are used for learning?
- What are the contributions of video fields of interests in FC?
- What technologies are used in FC?

Most of the explored research publications are extracted from digital libraries as IEEE-Xplore, ACM, and very small number of articles from Springer and IJIRCCE. The search strings in Table 1 are used to search in digital libraries mentioned above.

978-1-5386-5683-9/18/$31.00 ©2018 IEEE

TABLE I. SEARCH STRINGS

No.	*Search String*	*No. of papers*
SS1	((("Educational data mining") OR "learning analytics") AND video)	151
SS2	(((((("Data-driven learning") OR "self-paced learning") AND video) AND MOOC) AND Flipped Classroom)	93
SS3	((((("Educational data mining") OR "learning analytics") AND Video) AND Flipped Classroom)	41

Large number of articles showed on different search strings. There are selected just the ones that we saw reasonable to include as more appropriate and help achieve our goal. Most of papers are published in recent years. Fig. 1 shows the number of published papers by year. The papers that are published last year are from the first half of 2017. From the selected papers further analysis is conducted and in this study papers related to EDM and LA, self-paced learning and data-driven learning, FC learning and other online learning videos are included. As a result, after removing duplicates and irrelevant papers, only 122 articles remained.

III. CLASSIFICATION SCHEME

The classification scheme is presented in three columns where we have included the main fields of interests related to the research (Fig. 2). The educational technologies that are researched are the main fields on which we will focus in order to propose better learning strategies in the future. The field of interests are defined in the first column to come up with video field of interests. They can be used to create better student models and to be investigated in order to achieve better results in the video field on interest of the second column. The second column is the part where we think improvements can be made in order to have better results in interactivity and learning strategies using videos. Based on the analysis from the collected papers we found the gap in the 'video field of interest', in which can be contributed further. Technology used in FC that needs to be fulfilled in order the future goals be verified is shown in the third column. More details are presented in research questions in results part. Additionally, we classified papers according to the field of interest.

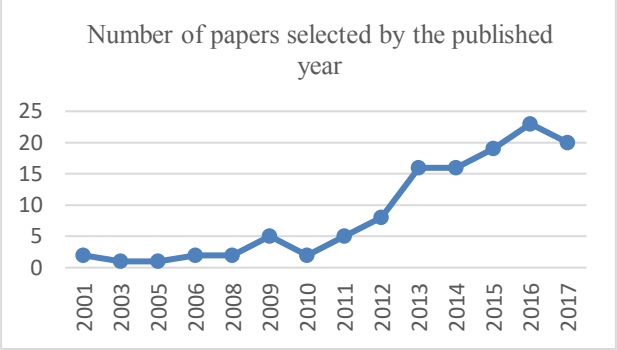

Figure 1. Number of published papers by year until mid-2017

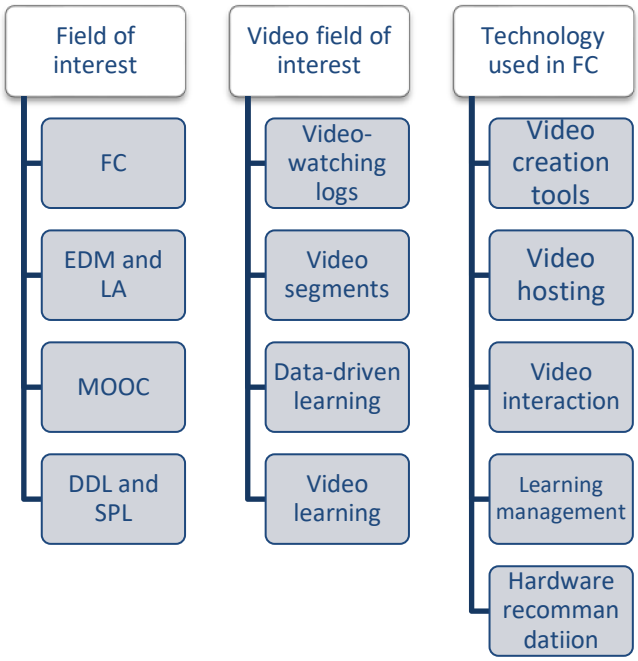

Figure 2. Classification scheme

IV. RESULTS

A. *RQ1: What are the main fields of interest investigated in the paper*

The main fields to investigate for papers are introduced as the first question in order to provide the number of papers for every field of interest and its percentage. The papers are chosen in order to provide additional benefits that can be used in FC. In Table 2 are shown the percentages of research papers for each field of interest. Analyzing and comparing the fields of interests presented in Table 2 we concluded that FC environment. more research and experiences are needed in order to improve the learning strategies and generate better student models. We found important to focus on utilizing a data-driven approach to make better decisions and use specific data mining techniques to enhance student learning experience in FC.

TABLE II. NUMBER OF PAPERS CHOSEN BY FIELD OF INTEREST

Field of interest	*Number of papers*	*Percentage*
Flipped classroom	31	25.4%
Educational Data Mining and Learning analytics	45	36.9%
MOOC	15	12.29%
Data-driven learning and self-paced learning	12	9.84%
Other	19	15.57%

B. *RQ2: How have analyzed publications in this research field changed over time?*

The interest in research on the presented field is increased by years. According to the selected literature the proposed topic has become increasingly popular during the years. We noticed that from 2012 there is an increased interest and it is growing through years. The increased interest through years that is reflected in the number of published papers during last years in Fig. 1 is shown. It is important to mention that we haven't included papers published during the second half of the 2017. Only some relevant papers of the first half of 2017 are included.

C. *RQ3: What kind of educational technologies are used for learning?*

First learning analytics summer institute was held in 2013. According to Baker & Inventado until 2022 all education research will involve analytics and data mining [15]. It is an emerging and promising field in education. The concept of learning analytics is wide and may have different meanings for different people. In addition, it shares many attributes with other closely related fields, such as EDM research and IT-supported learning processes [11]. LA and EDM help discover the hidden patterns from row data and respond to educational questions and problems [13], [12]. Improving analysis of data is the main goal of supporting research and practice in educations. LA has a relatively greater focus on human interpretation of data and visualization. Most of the selected papers for this study proved that experiments should use more data mining techniques [14]. Further, data mining algorithms could be districted into e-learning system. In this way we can use and benefit from the data mining techniques in different methods.

D. *RQ4: What kind of video field of interests are defined?*

Students frequently face difficulties [17]. It is because of learning strategies during educational experience and have no information of the most important or difficult topics. Using the process of data collection, preparation and analysis provided by educational data mining techniques and algorithms, we propose to try to improve the learning strategies in the future to overcome some difficulties.

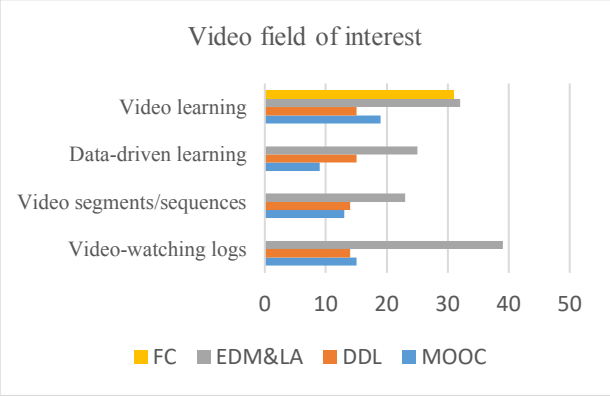

Figure 3. Number of papers analyzed based on field of interest and video field of interest

Even that the use of interactive videos and the different reports provided for the teacher attempt to overcome the gaps still there is a place for improving and resolving video learning strategies. After analyzing papers from the field of interests in classification scheme, we have defined the video field of interests showed in Fig. 3. From video watching logs, video segments, data-driven learning and video learning can be provided various ideas, techniques and algorithms to improve learning strategies. From the analyzed articles, we can use EDM to develop or refine learning strategies, predict students' performance and behavior. In addition, we can use time series data to promote self-paced learning, focus on utilizing a data-driven approach to make better decisions and use specific data mining techniques to enhance student learning experience.

E. *RQ5: What technology is used in FC?*

FC environment (Table 3) can be used in order to improve, develop and define learning strategies in the future. We have realized that better learning strategies, better understanding of learning topics and better student models can be provided. The technology has a big influence in the realization and quality of flipped classrooms. In addition to perform FC there is a need of some video tools depending on the device that can be used. Once the teacher created the video it needs to be hosted. In the video there are some interactions which reflect on the student learning level such as the questions that pops up in order to collect the analytics that watched the video, how the students have performed and the overall time watched. This is the part where we consider there is space for improvement in order to provide more data and possibilities for students in order to improve the interactivity and the learning process. It is linked with learning management solutions that help teachers to house all of a teacher's digital content. To increase the creativity of the videos, we need hardwares to create flipped class videos.

V. DISCUSSION

After analyzing 122 papers and shared experiences of flipped classroom, we believe it is so much more. There are many researches that can agree on the offered affordances and the positive effects on students' performance, attitudes, and engagement [2]. It increases student interaction "at home" and "in-class" by FC systems and tools for maximizing interactive and self-paced learning system of performed actions in video flipping classroom [1]. Data-driven learning system in Computer Science course using EDM and LA has big impact on FC.

TABLE III. TECHNOLOGY USED IN FC

Technology in FC	*Can be used*
Video Creation Tools	Screencast-o-matic…
Video Hosting	YouTube…
Video Interaction	Questions to collect the analytics and much more to be improved
Learning Management	Tracking, reporting and much more to be improved
Hardware Recommendations	Computer, tablet device…

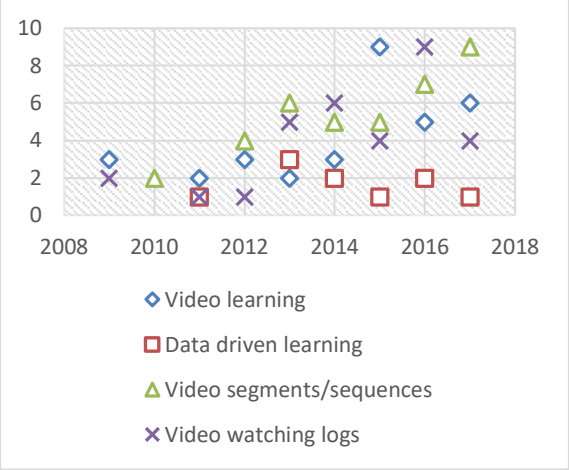

Figure 4. Time series of papers according to video field of interest

Kim et al. explored the design space of data-driven interaction techniques for educational video navigation [16]. To extract patterns through the CMS access data, we can use visualization, clustering, sequential pattern mining, classification, association rule mining and text mining. Learning systems are influenced from several areas to research and build models of EDM and LA. In education are used analytics to improve their services and to increase grades of students [6]. EDM tends to focus on developing new tools for discovering patterns in data [9]. Data mining is the process of automatically discovering useful information in large data repositories. A lot of the selected papers elaborate different video field of interest over years as showed in Fig. 4.

Based on learner behavioral patterns in Shabandar et al. papers, two sets of features and compared in terms of their suitability for predicting the course outcome of learners participating in MOOCs [8] are presented. Bell has described a system for capturing time-varying sequences of behavioral patterns of agents or robots from a video clips [10]. The detailed time slices of action are 'coarsened' to provide gross, molecular units of behavior using Rough sets methods [4]. The traditional algorithm for mining sequential pattern from web log data is used by Swati Singh Lodhi proposing a new algorithm "Sequential ID3" [5]. Since we wanted to find the gap of where and how we can help improving data-driven learning systems, in the same time promoting self-paced learning, the papers will not be enough to come up with further detailed algorithms. We will consider it for future goals and analysis.

VI. FUTURE RESEARCH AREA

By having all information available online, students must become self-learners and seek access to information on their own. Some future research areas can be to promote better self-paced learning, data-driven learning system using EDM and LA. In addition, one can propose a model and strategies for better learning environment to identify the most important or difficult topics that a student has on each video, teacher knows exactly which parts of the contents in each topic are most difficult for learners to be further improved, improve tracking and reporting or provide case study in FC using data-driven learning. Two research questions proposed as future research are:

- What kind of questions need to be answered for learning analytics and what algorithms can be used for answering them?
- What kind of questions need to be answered for educational data mining and what algorithms can be used for answering them?

VII. CONCLUSION

According to the analyzed literature, FC have become increasingly popular in higher education since 2012. A systematic study in order to improve learning interactivity via videos in educational technologies is introduced in this study. In addition we have introduced some research questions and some future research questions which have to be further elaborated. We were looking on time series of papers according to the gap of video field of interest. Most of the papers were supported by different analysis and models of EDM and LA

REFERENCES

[1] Doods, M. (2015). Evidence for the Flipped Classroom in STEM.

[2] Amresh A., Adam R. C., & John F. (2013). Evaluating the Effectiveness of Flipped Classrooms for Teaching CS1. *IEEE*.

[3] Heng N. M. (2014). Teaching Tip: The Flipped Classroom. *Journal of Information Systems Education.*

[4] Jorge R. & Peter M. (2016). Learning to Surf: Explaining the Flipped Classroom (FC) to Science Students Using an Analogy, Science and Education Publishing. *American Journal of Educational Research.*

[5] Qian, Gong, Yuan & Zhang (2016). Extracting Enterprises Collaborative Network from Massive Online Documents. *IEEE*.

[6] Liu & Zhang (2016). Emotion and associated topic detection for course comments in a MOOC platform. *IEEE*.

[7] Krithivasan, Gupta, Shandilya & Lala (2016). Auto-Tagging for Massive Online Selection Tests: Machine Learning to the Rescue. *IEEE*.

[8] Shabandar, Hussain, Laws, Keight, Lunn & Radi (2017). Machine Learning Approaches to Predict Learning Outcomes in Massive Open Online Courses. *IEEE*.

[9] Yueshun & Ping (2009). A study of Distance Learning Pattern Analysis and Evaluation Based on Data Mining. *IEEE*.

[10] Bell, Center, Beck, Miller & Herrera (2007). Video Mining - Learning Patterns of Behaviour via an Intelligent Image Analysis System. *IEEE*.

[11] Medeiros & Carvalho (2005). Applying Text Mining and Machine Learning Techniques to Gene Clusters Analysis. *IEEE*.

[12] Chang & Yang (2016). Semisupervised Feature Analysis by Mining Correlations Among Multiple Tasks. *IEEE*.

[13] Saskia R. (2013). Datafication of Education Enhancing Teaching and Learning Through Data Mining and Learning Analyses. *IEEE*.

[14] Ratnapala & Deegalla (2014), Students Behavioural Analysis in an Online Learning Environment Using Data Mining. *IEEE*.

[15] Baker & Inventado (2014). Educational Data Mining and Learning Analytics. *SPRINGER*.

[16] Kim, Guo, Cai, Li, Gajos & Miller (2014). Data-driven interaction techniques for improving navigation of educational videos. *ACM*.

[17] Gregory M., Teodora R. S. & Kathleen E. (2013). Inverting (Flipping) Classrooms – Advantages and Challenges. *ASEE*.