INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

Identifying Sign of Grit among Self-Paced MOOC Students Based on Clickstream Data

Nadirah Mohamad, Nor Bahiah Ahmad, Norsham Idris
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
Email: nadirah23@live.utm.my, bahiah@utm.my, norsham@utm.my

Submitted: 12/01/2018. Revised edition: 16/05/2018. Accepted: 21/05/2018. Published online: 31/05/2018

***Abstract*—Previous research that investigate and predict student's performance share various results concerning the factors that drive the students to retain and success. The factor ranging from intrinsic to extrinsic. However, few MOOC research attempt to scrutinize grit factor. Previous researches results are from survey. However, self-report survey is insufficient because the students may have different interpretation about the questions and measurement for themselves. Therefore, this is a case study that attempts to investigate the sign of grit among self-paced MOOC students using proposed features based on clickstream data. Also, the study observes whether the grit traits have relation with student's performance. This study found that many students have high grade despite having low grit traits. However, statistical analysis shows that there is still significant correlation between grit traits and student's grade. Other factor like unrecorded effort outside online learning may contribute to this result which can be investigated by future study. There is potential for better result by combining proposed method with collected survey.**

***Keywords*—Student's Performance, Grit, MOOC, Self-paced, Clickstream**

I. INTRODUCTION

Student's retention and performance has become the subject of research in online learning over the past decades. In these recent years, with advance of technology, MOOC becomes phenomenon and the issue of low retention and poor performance has become constant debate. Many researches have been done to investigate the student's online interaction and their feedback. The investigation is crucial in delving deeper into the root of the problem and to assist the students towards their learning goal whether to complete the course, achieve high grades or grant the students with certificate.

Previous research that investigate and predict student's performance share various results concerning the factors that drive the students to retain and success. The factor ranging from intrinsic to extrinsic factors like internal factor (satisfaction, enjoyment and interest) [1] and rewards (badges, grades and certificate) [2]. However, not many MOOC research attempt to scrutinize grit factor. Grit is known as "perseverance and passion for long-term goals" which "corresponds to the e-learners´ persistent capacity to pursue their learning goals" [3]. Especially in self-pace MOOC, the environment required the grit traits among students to encourage and motivate them achieving their goal. The reason is, MOOC do not have specific end date that most of the students who took self-paced MOOC devote several months or years to complete or achieve their goals. Apart from perseverance, the students need to have passion or interest to retain in online learning.

In a study [4], one of the principle emerges from the analysis of student's comment is interest group which shows interest is important for them to maintain online interaction and avoiding dropout. Respondents in a study [5] stated that MOOC voluntary engagement allow them to dropout anytime where they pick topics only to their interest. The importance of grit lead to many research on this topic in recent years and as previous studies recommended [6,7], investigating into different population such as this MOOC environment may support diverse measurements and reveal interesting result in order to measure online learning success. Next section discusses on the role of grit, related studies and grit traits that can be identified among MOOC students.

II. LITERATURE REVIEW

Grit has gains focus since Duckworth discussed the theory which initiated from a question on the reason that some individuals can achieve more than their peers with same intelligence [3]. Grit is known with two main traits which are perseverance of effort (perseverance) and consistency of interest (passion). Regardless of the domain, the traits can be found among successful people which distinguish them among the others. In previous studies, grit is found to be associated with level of education and age which means the students with higher grit score, have higher level of education or age [8]. This finding can be explained with the student's changing behaviour which become more matured as they grow up. The students also possess the traits of self-control, self-conscious and self-directed which can be related with grit.

Meanwhile, several studies found the association between grit and student's performance. For example, [9] found the association of current achievement with perseverance of effort while [8] found previous grade related with consistency of interest. A study [10] found grade related through engagement that produces productivity. The relation between grit and academic performance is also proved scientifically by [11], where fractional amplitude of low-frequency fluctuations (fALFF) in right dorsomedial prefrontal cortex (DMPFC) are associated with student's academic performance in which right DMPFC brain area is known for self-regulated and goal settings that shapes the traits of grit. However, there is also a study that found contra result like [6]. Even though the research that links grit with academic performance is still low and with diverse result, investigation in this MOOC context may reveal potential benefits.

A. Grit and MOOC Environment

MOOC is known as another medium of learning in this global era. What distinguish MOOC with other medium of learning is in term of its openness for connection with global participants and access to the top universities resources freely. Another characteristic of MOOC that rarely discussed is its self-paced mode. Most of MOOC developed today is self-paced which opposed with session based. As cited in [12], self-paced MOOC has growth exponentially since 2013. The report indicated that most of students nowadays are following self-paced MOOC which the learning occurs according to their own time and their own will. The environment which encourages independent, allocates goal with longer term. To achieve that long term goal, the students required to depend on themselves to exert more effort in interacting online like viewing resources and answering problem solving questions. Other than that, they need to have high interest to initiate their effort. When there is interest, curiosity is stimulated which promotes student's effort to get to the answer during learning process [13]. The characteristics discussed can be translated into grit traits.

Identifying grit traits among students give advantage to instructor in understanding their student more such as whether they need to provide higher difficulty of materials for grittier student [14]. Revealing student's grit traits also gives potential in revealing person's ability to self-direct and perseverance towards achieving goal. As stated previously, grit is known to be associated with long term goal. However, when the result shows that student is not gritty, there is possibility that he is not familiar with long term learning mode. The result does not mean the instructor needs to change the student's goal, instead the instructor can advices on better learning strategy or, giving priority in providing feedback for this group of students. Therefore, exploring and revealing sign of grit in self-paced MOOC students are important.

B. Sign of Grit Through Clickstream Data

In term of methodology, Duckworth survey scale is widely known to be used in measuring grit [15]. Duckworth grit survey consists of six questions under consistency of interest (COI) category and six questions under perseverance of effort (POE) category. The respondents of the study answered the survey and reported how they feel about each questions. However, self-report survey is insufficient because the students may have different interpretation about the questions and measurement for themselves. For example, a student may think that she is not diligent enough compare to her peers who is very diligent. A study [10] also shares the same concern regarding this issue while [16] suggests to use performance measure instead of self-evaluation measure which can shows the students real performance that reveals grit traits. Moreover, in MOOC environment with student's various background, the method may consume time and creates student's apathy [17]. Thus, this study would like to explore other possible approach which is mining through clickstream data.

Clickstream data is result from user's interaction with online learning system that consists of user's navigation path. Analysis from this recorded data can reveal student's behaviour including understanding about their interest [18]. Also, the data can present their sequential activities from registration until the final session. User's navigation history can show whether the students visit MOOC every week diligently or have visited for only few times. Apart from that, the data can show whether the students active doing all the activities or have only read earlier topics. These advantages become clickstream potential in revealing sign of grit. Therefore, this study attempts to investigate the sign of grit among self-paced MOOC students with proposed features based on the clickstream data. Also, the study observed if the grit traits have relation with student's performance.

III. METHODOLOGY

This study used a MOOC data from Stanford University based on edX platform) consist of about 2.4 million records. After filtered (only students who have grade records with no missing valued is used), only 582 students are analysed. As

suggested in previous studies, actual grade which resulted from student's online exercise and exam is used for this study. Grade is associated with students' learning goal in this course as in most of other courses based on edX platform (standard of achievement is given based on the student's grade).

As for the measurement, apart from limitation discussed in previous section, privacy is also the main concern when collecting and analysing MOOC data. However, this study put the greatest effort in using available data to provide useful insight. Moreover, Duckworth's theory [3] is still being used for features construction. In order to construct features from clickstream data based on the grit theory, this study refers to [19]. The method involves the process which construct features based on literature review. Fig. 1 shows the process which is important to be discussed in this study to reveal and explains the grit traits among students.

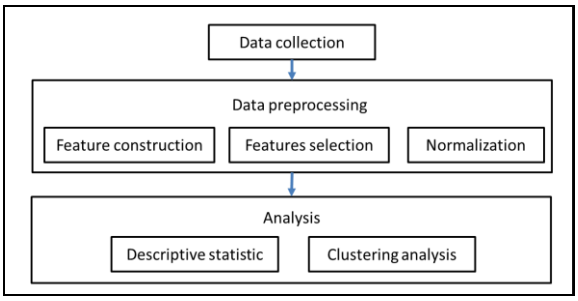

Fig. 1. Process involved in investigating sign of grit among students

After collecting data, features construction is implemented to provide a dataset based on previous studies. Then, the study has selected features to observe the correlation among the features proposed and the rank given according to the most important features. Descriptive statistic also has been done to present the distribution of data. Later, clustering analysis is done after the dataset is normalized. Next subsection explained the process with more details.

A. Features Construction

In this subsection, the method of preparing the features for grit is explained. From the definition of grit and the scale constructed by [15], the grit traits consist of consistency of interest and perseverance of effort. The features constructed are chosen based on, and closest to the grit definition and theory. Fig. 2 shows three files involved to extract information in order for all the grit features to be prepared. First, all relevant files for the feature construction is collected which include _EventXtract, _allData and _ActivityGrade.

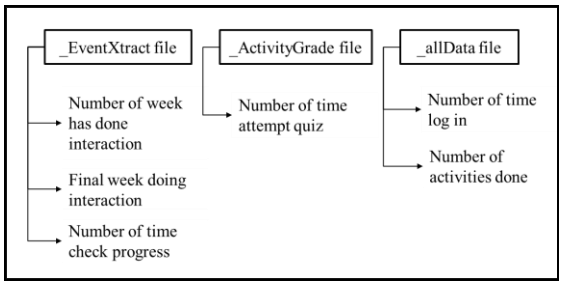

Fig. 2. Files involved to extract information for each grit features

For 'number of week' and 'final week doing interaction', this study lists and sorted all the student's url navigation with the resources information and the time at which the event initiated. The information is available in _EventXtract file. The list is divided into weeks where the time of the first event initiated, appointed as the first week. The reason for following the students' paced instead of the course paced are; for self-paced MOOC, it is unlikely to find the date of students' registration to be align with the date of course start. Also, it easier to track on the patterns of week each student takes to learn personally. For example, there is student who visits the course once every two weeks or student who persist to visit every week. Such patterns provide meaningful insight and can be observed using this feature. Meanwhile, from the same _EventXtract file, 'number of time check progress' can be constructed through the frequency of the student checking his progress with url navigation where '/courses/field/course_name/SelfPaced/ progress' is referred.

Another feature that shows student's perseverance of effort is 'number of time attempt quiz' which the information is available from _ActivityGrade file through 'number of attempts' field. The field for particular student is total up to get the score. For 'number of time logging in' and 'total activities', _allData file is analysed. The file records every session in which the student has initiated including the date, time and events in a session. The first feature is based on how many time the student initiates a session which can be observed by the total number of sessions for a particular student. For the later feature, information like number of events in a session is referred to and is total up for all sessions.

Fig. 3 shows the activities and features involved in this study to identify students with grit traits. As shown in Fig. 3, the activities selected in this study is logging in, assessment and the total activities. Meanwhile, Table I shows the features and definition for each features. According to [20], the trait of grit can be showed by the student's effort to logging in and making attempt. Therefore, representing trait 'persistent of effort', the features is collected from logging in and activities that shows student's attempt on quiz, which is assessment. All the relevant features that can be collected is listed which include 'number of weeks the student has done interaction', 'final week doing interaction', 'number of time the student log in' and 'number of time the student attempt quiz or exam'. 'Number of time the student checking progress' is included as the feature is related to both

assessment and logging in. The feature also shows student's effort in ensuring there is improvement in the student's learning. In this study, for this MOOC platform, 'number of session' which the student has is used for 'number of time log in'.

Meanwhile, 'consistency of interest' can be showed by the number of activities which the student has done. In [21] study, several student's comments show that they dropout when they could not find the topic or part that interesting. This indicates that student who is still interested in the course selected, will produces more interaction or doing more activities. Next section, each features constructed importance is discussed with features selection.

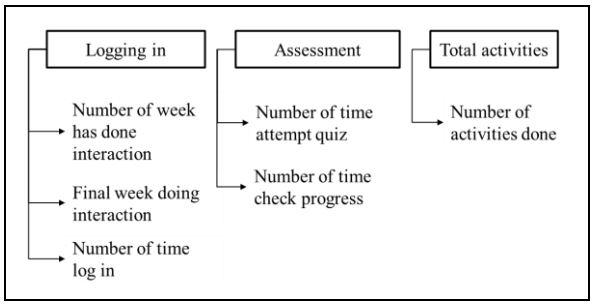

Fig. 3. Activities involved to extract information for each grit features

TABLE I. FEATURES AND DEFINITION FOR GRIT TRAITS

Features	**Definition**
Number of weeks has done interaction	Frequency of weeks which the students has interaction from the week register until the last week has done activity.
Final week doing interaction	The final or the last week in which the student has done interaction before absent.
Number of time log in.	Frequency of the students logging in the online course from register until the last activity.
Number of time attempt quiz/exam.	Frequency of the students attempt the questions of quizzes, assignments or exams.
Number of time check progress.	Frequency of the student checks the learning progress to see the status of course completion.
Number of activities done.	The total number of events of activities which the student has done from each session since register until the last activity.

B. Features selection

After feature construction, features selection is implemented using principle component analysis (PCA). PCA is known for the ability to produce a list of features priority, according to rank represents by fraction of variance [22]. Before proceed, Table II presents the correlation matrix for all grit features. From Table II, 'number of week doing activities' has the strongest correlation with 'final week doing interaction', number of time log in' and 'number of time check progress'. While for 'number of activities done' and 'number of time attempt quiz', the strongest correlation is 'number of time check progress'.

Meanwhile, result of features ranking shows that, first rank features consist of 'number of weeks has done interaction', 'final week doing interaction', 'number of time log in', 'number of time check progress', 'number of attempt' and 'number of activities done' (according to priority) with 0.63 variances. The results conclude that 'number of week has done interaction', 'final week doing interaction' and 'number of time log in' are among the strongest features proposed. However, other features like 'number of attempt' and 'number of activities done' are also important as the features exhibit association with other features and the impact towards grades is important to be investigated.

TABLE II. ACTIVITIES AND FEATURES FOR GRIT TRAITS

Features	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
NoWeek (F1)	1	0.7*	0.56	0.29	0.11	0.15
FinalWeek (F2)	0.7*	1	0.26	0.2	0.02	0.06
NoLogin (F3)	0.56*	0.26	1	0.16	0.12	0.09
CheckProgress (F4)	0.29*	0.2	0.16	1	0.14	0.16
TotActivities (F5)	0.11	0.02	0.12	0.14*	1	0.06
NoAttempt (F6)	0.15	0.06	0.09	0.16*	0.06	1

C. Analysis Method

There are two central analysis used in this study. First the study observed student's score on each grit features with descriptive statistic. Then, the correlation between each grit trait and student's grade is studied using correlation analysis. Second, the study categorized the students based on the grit features using self-organizing maps and later, compare with their grades. This analysis is done to investigate whether the student who placed in high grit category, has high grade or vice versa. Clustering is one of educational data mining technique that is able to categorize variables according to their similarity naturally [23]. The result provided insight on how the students are grouped together.

IV. RESULT AND DISCUSSION

Descriptive statistic has been produced for grit traits, Table III shows the minimum, maximum and mean value for each feature. 'Number of weeks has done interaction' means the students frequently visit the online course every week and do interaction. The longest period is the student who visited the online course for 36 weeks and the lowest is only a week. The 'final week doing interaction' is 59, which means the student diligently visited online course and the final week the student has done interaction is on the 59th week.

TABLE III. MINIMUM AND MAXIMUM NUMBER FOR EACH FEATURES FOR GRIT

Features	**Minimum**	**Maximum**	**Mean**
Number of weeks has done interaction	1	36	5.04
Final week doing interaction	1	59	9.99
Number of time log in	1	580	20.96
Number of time check progress.	0	91	17.23
Number of activities done	2	37452	787.90
Number of attempt quiz	0	115	75.67

This does not necessarily mean the student keeps visiting the course every week but could be several weeks or few weeks. The lowest reported is the student who visits only on the first week. In term of frequency of logging in, the maximum number is 384 times compare to the student who only login once which the result shows a wide gap. Then, the 'number of time check progress' shows 91 for maximum and there is a student who did not even checked his progress. Meanwhile, the maximum 'number of activities done' is 37452 for the whole period which shows the student's interest in completing the activities given.

For mean value, the study concluded that on average the student retain in the course for five weeks, in which the student last visit is on his 10th week. The study assumes that the students return late because of other commitment as the student's level of education are predominantly master or doctorate. Most of the students have log in the course 21 times and checked their progress 17 times. On average, the students have done 788 activities like watch video, view content or answer quiz, and attempt all quizzes given for 76 times. The study observed that there is wide gap between minimum and maximum for all features. The skewness results also show high skewness for all features (1.2 to 11.6) exclude 'number of attempt quiz' which is moderately skew (-0.9). However, the study intends to include all the data to investigate and the reason which later is discussed in section IV.

Meanwhile, contrary to most of previous research which found low performance among MOOC students, the result from this study shows that 571 students (98%) score more than 50% while 11 students (2%) score less than 50%. However, it would be interesting to investigate the result of grit traits for both groups. In order to find out whether each feature is correlate with student's grade, correlation analysis is done. The highest correlation is the 'number of weeks which student has done interaction' which is 0.20. Followed by the 'number of time student check progress', 0.198. Then, 'number of time logging in', 0.184 and 'final week student has done interaction', 0.142. The lowest, 'total number of activities' is 0.108. From the result, the study concluded that even though weak, all the features have significant correlation with student's grade.

Before the second analysis, which clustering analysis is used, dataset is normalized to allow normal distribution across features which have various measurements. Then, dataset is analysed with self-organizing maps technique using Weka tool. Self-organizing maps is known to support high-dimensional data and provides clear visualization on cluster produced [23]. From the analysis, most of students belong to *cluster1* (49%), followed by *cluster2* (22%), cluster0 (18%) and *cluster3* (11%). Meanwhile, Fig. 4 shows the average score of grit traits for each clusters. Further investigation found that cluster3 consist of students with high score for grit traits, followed by *cluster2*, *cluster1*, and *cluster0*.

Next, the association between clusters formed and grades are discussed. Fig. 5. shows the result of the clustering analysis where there are four clusters which this study expects to categorize the student according to student's level of grit. Most students with low grades belong to *cluser0* and few more in *cluster1* and *cluster3*. Because most of the students achieve high grades, the students can be seen clustered in all groups. Through cluster interpretation, the study concluded that most of the students who belong in *cluster3* have high score for all features, while most students who belong to *cluster0* has low score on all features. However, the study also found that the clusters contain mix result of features. Meaning, some students have low score on grade but belong to *cluster3* that has many students with high grit. While there are also students with high score on grade but belong to *cluster0* that has many students with low grit, as shown in Table IV.

When compared with the student's grade, most students result depicted their clusters. However, as explained previously, noticed that student with ID *033* has unexpected result. After investigation, the student did log in to MOOC (low score can be seen before normalization) but the student really focusses on completing all the given quiz or exam which means the student has low interest on the course but determine to achieve high result, thus, produced low score for grit. Meanwhile, student with ID *324* showed high effort with high score for the first two features and belong to cluster with highest grit. However, the student's grade is very low. After deep observation, the student did put a lot of effort but almost half of the student's week are not filled by activities which means the student just read course info or login for few minutes. Also, this result may have been contributed by other factor that need further investigation such as the student's understanding. Moreover, based on previous studies, there are many factors that contribute to student's online interaction. This means the factors may include these grit traits or may highly depend on the student's intelligence.

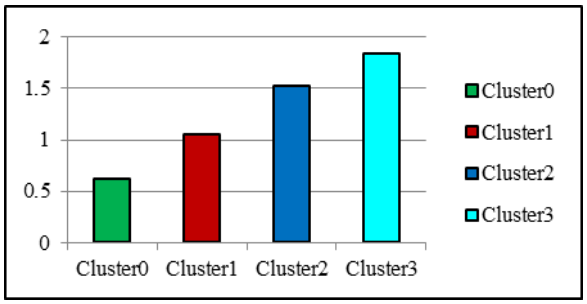

Fig. 4. Students according to four clusters based on average score of grit

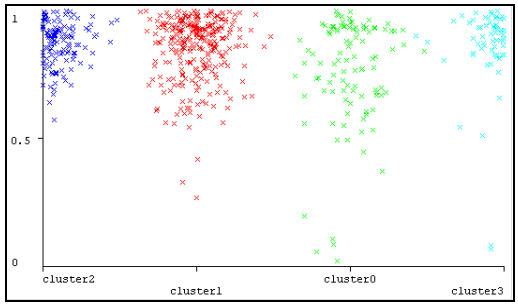

Fig. 5. Students according to four cluster based on grit traits (X) axis and grade (Y) axis

TABLE IV. SAMPLE OF STUDENTS WITH CLUSTER ASSIGNED BASED ON FEATURES CONSTRUCTED

ID	NoWeek	FinalWeek	NoLogin	ChckProgress	NoAttempts	TotActivities	Cluster	Grade
352	1	0.62	0.66	0.1	0.82	0.06	cluster3	0.95
324	0.6	1	0.02	0.04	0.09	0.02	cluster3	0.03
273	0.54	0.62	0.17	0.32	0.72	0.1	cluster3	0.82
097	0.31	0.29	0.07	0.25	0.73	0.07	cluster2	0.84
624	0.34	0.31	0.07	0.52	0.8	0.07	cluster2	0.83
314	0.11	0.09	0.02	0.3	0.64	0.01	cluster2	0.81
538	0.03	0.02	0.01	0.09	0.82	0.01	cluster1	0.74
688	0.06	0.05	0.01	0.02	0.77	0	cluster1	0.74
628	0.29	0.28	0.07	0.08	0.57	1.3	cluster1	0.73
033	0.06	0.03	0.02	0.19	0	0	cluster0	1
024	0.03	0.03	0	0.03	0.03	0.00	cluster0	0.14
143	0	0	0	0.02	0.18	0.00	cluster0	0.12

V. CONCLUSION

This study has investigated the sign of grit among self-paced MOOC students using proposed features based on clickstream data. Also, the study has observed whether the grit traits has relation with student's online performance. The study found that based on clickstream data, the analysis produced mix result. Most students have high grade despite having low grit traits. This result may have been contributed by other factor like unrecorded effort outside online learning which can be investigated by future study. Even though weak, correlation between all features with grades is still a significant. If collecting data is not the constraint, other studies can combine both the survey and proposed method with more investigation optimizing other features based on specific activities. For example, collecting data on number of time watch video which represent student's interest on given topics. Similar study can also be conducted on more datasets with various online learning environment to observe the implication. Also, further studies are needed to investigate the association between grit traits and student's retention in completing the course in online learning. Through this study, we expect to contribute to grit-related study and research that investigate factor of student's performance or retention.

ACKNOWLEDGMENT

The authors would like to express our deepest gratitude to Research Management Centre (RMC), Universiti Teknologi Malaysia (UTM) for the support in R&D (Tier 1, 17H75) and UTM Big Data Centre, Soft Computing Research Group (SCRG) for the inspiration in making this study success. Also, CAROL, Stanford University, thank you for the opportunity and provide access to explore the data.

REFERENCES

[1] Fırat, M., Kılınç, H., & Yüzer, T. V. (2018). Level of Intrinsic Motivation of Distance Education Students in E-Learning Environments. *Journal of Computer Assisted Learning*, 34(1), 63-70.

[2] Khalil, M., & Ebner, M. (2017). Clustering Patterns of Engagement in Massive Open Online Courses (MOOCs): The Use of Learning Analytics to Reveal Student Categories. *Journal of Computing in Higher Education*, 29(1), 114-132.

[3] Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-term Goals. *Journal of Personality and Social Psychology*, 92(6), 1087.

[4] Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC Design Principles: A Pedagogical Approach from the Learner's Perspective. *eLearning Papers*, 33.

[5] Liyanagunawardena, T. R., Parslow, P., & Williams, S. A. (2014). Dropout: MOOC Participants' Perspective. *Proceedings of the European MOOC Stakeholder Summit 95.*

[6] Bazelais, Paul, David John Lemay, and Tenzin Doleck. (2016). How Does Grit Impact College Students' Academic Achievement in Science? *European Journal of Science and Mathematics Education.* 4(1), 33-43.

[7] Aparicio, Manuela, Fernando Bacao, and Tiago Oliveira. (2017). Grit in the Path to e-Learning Success. *Computers in Human Behavior*, 66, 388-399.

[8] Brooks, Nita G., and Scott J. Seipel. (2017). Grit and the Information Systems Student: A Discipline-Specific Examination of Perseverance and Passion for Long Term Goals. *Proceedings of the EDSIG Conference*, 2473, 3857.

[9] Wolters, Christopher A., and Maryam Hussain. (2015). Investigating Grit and Its Relations with College Students' Self-regulated Learning and Academic Achievement. *Metacognition and Learning*, 10(3), 293-311.

[10] Hodge, Brad, Brad Wright, and Pauleen Bennett. (2017). The Role of Grit in Determining Engagement and Academic Outcomes for University Students. *Research in Higher Education*, 1-13.

[11] Wang, Song, Ming Zhou, Taolin Chen, Xun Yang, Guangxiang Chen, Meiyun Wang, and Qiyong Gong. (2017). Grit and the Brain: Spontaneous Activity of the Dorsomedial Prefrontal Cortex Mediates the Relationship between the Trait Grit and Academic Performance. *Social Cognitive and Affective Neuroscience.* 12(3), 452-460.

[12] Cunningham, J. A., Bitter, G., Barber, R., & Douglas, I. (2017). Using Traces of Self-Regulated Learning in a Self-Paced Mathematics MOOC to Predict Student Success.

[13] Ainley, M., & Hidi, S. (2014). Interest and Enjoyment. *International Handbook of Emotions in Education*, 205-227.

[14] Fancsali, Stephen E., Steven Ritter, John Stamper, and Tristan Nixon. (2013). Toward "hyper-personalized" Cognitive Tutors. *AIED 2013 Workshops Proceedings.* 7, 71-79.

[15] Duckworth, A. L., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (GRIT–S). *Journal of Personality Assessment*, 91(2), 166-174.

[16] Willingham, D. T. (2016). Ask the Cognitive Scientist: Grit Is Trendy, but Can It Be Taught? *American Educator*, 40(2), 28.

[17] Yukselturk, Erman, Serhat Ozekes, and Yalın Kılıç Türel. (2014). Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program. European *Journal of Open, Distance and E-learning*, 17(1), 118-133.

[18] Aguiar, E., Nagrecha, S., & Chawla, N. V. (2014). Predicting Online Video Engagement Using Clickstreams. arXiv preprint arXiv:1405.5147.

[19] Romero, C., Ventura, S., Pechenizkiy, M., and Baker, R. S. (Eds.). (2011). *Handbook of Educational Data Mining*. CRC Press.

[20] Agnihotri, Lalitha, Ani Aghababyan, Shirin Mojarad, Mark Riedesel, and Alfred Essa. (2015). Mining Login Data for Actionable Student Insight. *International Educational Data Mining Society.*

[21] El Said, G. R. (2017). Understanding How Learners Use Massive Open Online Courses and Why They Drop Out: Thematic Analysis of an Interview Study in a Developing Country. *Journal of Educational Computing Research*, 55(5), 724-752.

[22] Bro, Rasmus, and Age K. Smilde. (2014). Principal Component Analysis. *Analytical Methods*, 6(9), 2812-2831.

[23] Gallén, R. C., & Caro, E. T. (2018). A Benchmarking Study of K-Means and Kohonen Self-Organizing Maps Applied to Features of Mooc Participants. *European Journal of Open, Distance and E-learning*, 21(1).