International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

Multimodal Learning Analytics And Its Challenges- A Systematic Review

Dr P S S Akilashri[1], Ms S Thulasi Bharathi[2]

[1]*Head, Associate Professor, Department of Computer Science, National College, (Autonomous), Trichy, TamilNadu, India.*
[2]*Research Scholar, Department of Computer Science, National College (Autonomous), Trichy, TamilNadu, India.*

Email: [1]akilas27@gmail.com,[2]srimas88@yahoo.co.in

Abstract: In recent years data has been growing enormously at the faster rate originating from different sources, with different rate. It is essential to learn what that data mean. Data generated from multiple modes termed as "multimodal". In this paper, a systematic review has done with Multimodal Data (MMD) and Learning Analytics (LA) for better understanding of human learning. This study focusses on what and how MMD has been used to understand the learning. The study consisted of 34 papers for analysis collected from MMLA research community journals (ACM, LAK, IEEE, and BJET). This paper discusses about the State of art of MMLA and its implications using Machine Learning Techniques. It includes a Multimodal Learning Analytics Model for better understanding of the learning process. It addresses the role of MMLA in various domains and the challenges in collecting, processing and analysing the MMD.

Keywords: Multimodal Data, Learning Analytics, MMLA, Machine Learning, Deep Learning, Supervised models

Abbreviation: MMD- MultiModal Data; MMLA – MultiModal Learning Analytics; ACM – Association of Computer Machineries; EEG- ElectroEncepheloGraphy; ECG- ElectroCardioGraphy; LA - Learning Analytics. BJET - British Journal of Educational Technology; LAK – Learning Analytics and knowledge.

1. INTRODUCTION

Analytics is used to explore the moments of learning such as learner's experiences and behaviour while learning with the help of technologies. Learning analytics (LA) is a thriving area where data about the learners and the learning context are collected and analysed to understand and optimize learning outcomes [1]. Today the most of the research done on LA relies on the data collected from digital platforms. There is a shift in learning analytics from the unidimensional data collection methods[2], whereas in the current scenario, there is a change in analysing the online learning process, because the methods of collecting, processing and analysing the traditional learning systems were not sufficient to analyse the learning behaviour of today's modern learning systems [3] (E.g.: Learning Management System). To comprehend the modern learning process, the system collects data from multiple sources and applies multiple techniques and different measurements to a dataset [4]]. This is mentioned as multimodality of data [5] it plays a role to do learning analysis with more accuracy. The role of multimodal data in Learning Analytics involves measuring and collecting facial expression data, and physiological data (EEG, Heart rate,) including processing and reporting of behaviour data from heterogeneous sources to perform analytics

2. WHAT IS MULTIMODAL DATA? – NEED FOR MULTIMODALITY - A REVIEW

The multimodal data collected from multiple sources such as wearable sensors, Eye-tracking devices, EEG etc. was explored with the help of technology to find the relationships between them using Machine Learning Techniques. The advancement in the analytical part makes us to understand the complex learning occurrences [7]. By using the existing MMLA studies the different types of multimodal data classification are tabulated in Table 1.

Types of data	**Depiction about the Data**	**Paper**
Eye tracking data	About the students attention patterns and pupillary response	[6],[12],[13]
Video data	About their engagement in the class	[7],[8],[15]
Audio data	About the Learning performance	[24]
EEG	About the deep mental processes such as long/short-term memory (LSTM) load and cognitive workload	[12],[13],[16]
Face Data	About the Emotional responses like happiness, boredom, sadness and engagement	[22],[24]
Wrist Band	About to learn the normal routines of an individual	[16],[19],[20]
Click Stream Data	About to predict the movement-motor learning performance	[8],[25]

Table 1: Types of Multimodal Data and its depictions

The perceptions from multimodal data was enabled to get more information rather than collected the individual sources. The prediction of skill acquisition made out of MMD (E.g.: Facial video, eye tracking, EEG) was far better than using individual stream of data [16]. The need for analysing multimodal data becomes vital as it is almost present in all the domains of the science and technology. The application domains of multimodal data has been discussed in Section 3.1.2.

3. RELATED WORKs

3.1 Multimodal learning analytics Research [MMLA]

The unification of MMD with computational analyses was stated as MMLA. The research in learning analytics was focused on the computer mediated and structured learning [13], [14]. Here the learning is not analysed only with unimodal it has multiple modes (cognitive mode, behavioural mode) needs to be analysed to perform prediction.

Exactly, this paper is going to focus on the following questions from the existing literature

RQ1: What is the state of the art of research in MMLA?

RQ2: What are the Application Domains in MMLA?
RQ3: What are the challenges in MMLA?

3.1.1 The State-of-art of MMLA and Machine Learning

The broad search of peer-reviewed articles was done in August 2020. The term "Multimodal Learning Analytics", "MMD" were used to collect articles which covers wide range of articles published in MMLA community. The motivation of this paper is to identify the developments in the region of MMLA. The proponents of MMLA suggested the multimodal data can create complex models of learning [10]. For instance, in [7] the MMLA enhanced the understanding in detecting and modelling student learning behaviour.

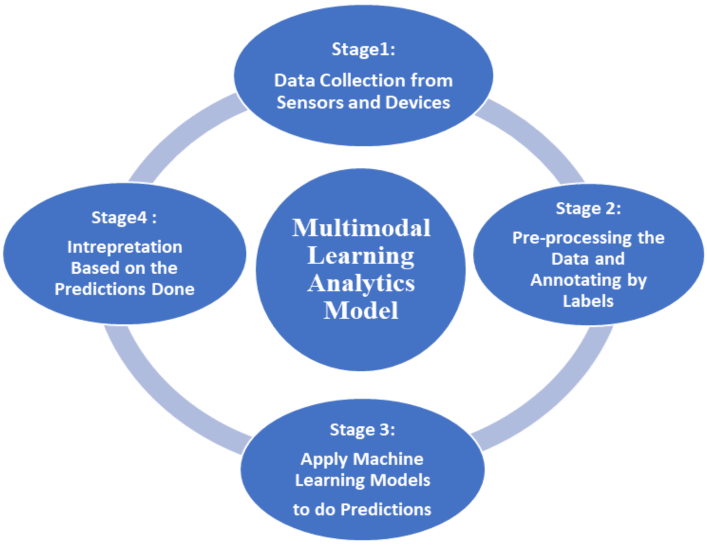

Figure 1: Multimodal Learning Analytics model

For better understanding the MMLA Model was depicted in the *figure 1*. It includes *i)* various modalities of data collected from physical spaces(Sensors and devices) *ii)* Collected data pre-processed and stored in digital spaces (Pre-processing and annotation) *iii)* Apply Machine Learning techniques like classification, regression and neural networks to predict the data, as a result feedback and interpretation was given [20],[22]. The usage of multimodal data especially clickstream mode, explored the expected benefits that employ physiological sensing for designing learning technologies [19]. The improved accuracy of prediction was proven against the unimodal data's and hence generalized methodology was created for MMLA pipelines [17], [19]. As the technology was growing the granularity of the data and modality of the data has been analysed by training the algorithms using the emerging Machine Learning algorithms. The multimodal data was been approached using Neural Networks, Regression Models, and Classification under supervised learning models [22].

The below figure 2 depicts that different types or modes of data can be utilized for analysing the learning pattern of a human being in the view of behavioural and contextual pattern.

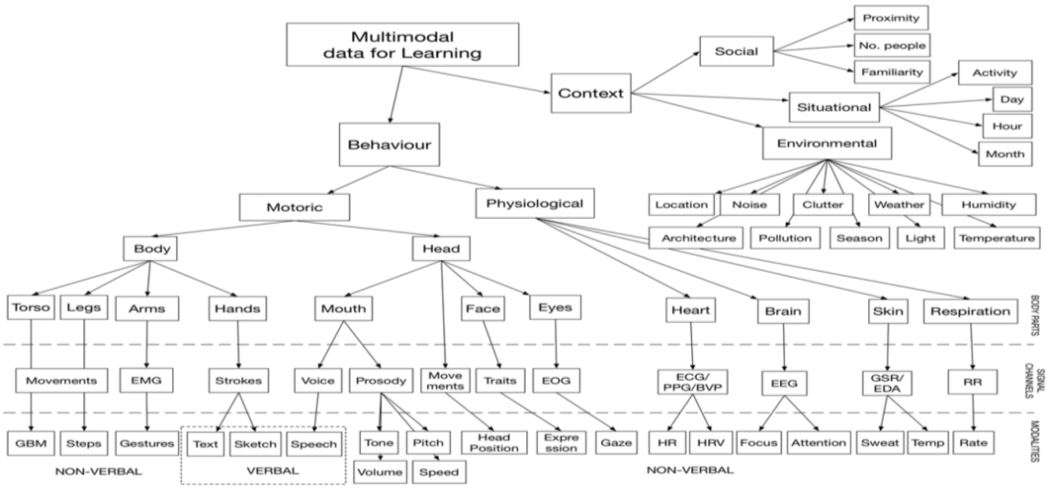

Figure 2: Multimodal data for learning [5]

The data collection, MMLA researchers have used a plethora of data streams to obtain a deeper insight about the learners' rich and "multimodal" behaviour. The fusion of multimodal data provide better prediction on the learning outcomes to interpret the complex learning process [12], [17]. In [10] the combination of system interaction, eye tracking and facial expression data predicted the intrinsic motivation and learning outcomes of students in a game based environment. In [16], [19] the combination of eye-tracking, cognition and facial measurements have been analysed to predict the attention span, cognitive pattern and affective domain respectively. And also combining these three measurements provide a comprehensive picture of learner performance, outcomes and behaviour. In [16],[19],[27] the combination of audio data, eye-tracking, system logs, video data and physiological data were used to predict the learners task performance and also these combinations have been used to define the different behavioural trajectories of learners. In [22], [28] the actions of bodily movements was detected by the video cameras and the prediction was performed using Deep Learning Techniques. To recognize the gestures and body postures Microsoft Kinect tool was employed.

In [31] the combinations of systems gaze, system log, audio and dialogue data were used to predict the learning gains in the collaborative intelligent tutoring systems and also temporal scales contributed in increasing the learning outcomes. In [22],[33],[34] the combination of interaction logs, gestures and posture predicted about the memory, conceptual understanding and cognitive workload and motivation. So the MMD can be maximum utilized by identifying the correct association between the modes to predict the ultimate goals of understanding the learner experience and behaviour. This can be achieved to a higher level of success by either the technology or the design of the process. The combinations of multimodal data and its prediction is tabulated in the table 2. But it was still a challenging factor in MMLA studies to integrate data.

Combinations of data	**Prediction about the Data**	**Paper**
System Interaction, Eye Tracking and Facial Data	Intrinsic motivation and learning outcomes of students in a game based environment	[10]
Eye-tracking, Cognition and facial	attention span ,cognitive pattern and affective domain	[16],[19]
Audio Data, Eye-tracking,	predict the learners task performance and	[16],[19],[27]

System Logs, Video Data and Physiological Data	behavioural trajectories of learners	
Gestures, postures, Eye tracking	Learners interest	[22], [28]
systems gaze, system log, audio and dialogue data	learning gains in the collaborative intelligent tutoring systems	[31]
Interaction logs, gestures, postures	Memory, conceptual understanding, cognitive workload and motivation	[22],[33],[34]

Table 2: Multimodal Data combinations and its outcome

3.1.2 Application Domains of MMLA

1*) AV modality in multisensory systems:*

It is used for a broad range of applications like *speech processing, speech recognition, speech activity, speech detection, speech enhancement, speech extraction and separation* [32]

2) Machine – Human interaction:

The ideology in MHI is to combine multiple interaction modes based on AV, and bodily movements' data (eye tracking, gesture, etc...) to learn the human physiological commands and other multisensory functions. These data are useful to predict the learners' behaviour and performance [32].

3) ***Biomedical, Health:***

The EEG data deals with understanding the elements of the brain which is essential for analysing the cognitive activities. Deep neural networks was required to analyse the EEG data. Recent medical technologies enables us to do collect data to record brain imaging [28].

4) ***Educational Research:***

The digitalization has emerged in the field of education as the online learning and distributed learning has increased significantly. As a result, to understand the learners' behaviour and to optimize their learning, LA was essential in the field of education to identify the success learning pattern [34].

5) ***Environmental studies:***

The MMLA was utilized in meteorological monitoring. Accurate measurements of vapour, smog, rain were needed for meteorological analysis and forecasting. Data in this domain was collected through Radars, Rain Gauges, Satellite borne remote sensing devices and microwave links [32]. Data collected from the above devices were integrated to make better predictions.

3.1.3 Challenges in MMLA:

From the literature review few research gaps was observed in the MMLA, here important challenges was addressed to be observe red and analysed in the future.

1) ***Lack of Data Utilization**:*

It has been explored that few modes of collected data has not been utilized efficiently. ***i)*** The EEG data was underutilized due to low signal-to-noise ratio issue. ***ii)*** Pre-processing of EEG data seems to be high and ***iii)*** Cost of the equipment (expensive) for conducting experiments. In the literature [13], [16], [19] the EEG data was used along with other modes to study the brain patterns. But utilizing these kind of data would provide more insights on deeper cognitive study.

2) Sample Size:

And also there was another issue with respect to the sample size [17]. The increase in sample size significantly increases the difficulties in collecting, processing and analysing Multimodal Data, because each data carries different set of noise sources, improving Signal-to-noise ratio (SNR) for different streams seems to be a difficult task. This problem got aggravated when there are many physiological sensors involved [13], [16], [19].

3) Heterogeneity of data measurements:

Each mode of data has its own scale of different sampling rates(for e.g.: EEG 120-550Hz, Video 20-70FPS, Audio 40.1KHz, Eye tracking 30-200Hz) which requires extra pre-processing of the data collected to ensure the temporal resolution of all the modalities[7],[25].

4) Data integration:

As MMLA is still exploring, there is still a lack of a theoretical and overall analysis in the area of multimodal data integration [29]. Meanwhile the temporal state of the data from different sources with different sampling rate and different sizes, the data integration for pre-processing and analysing was a great challenge [5],[7],[10].

5) Generalization:

MMLA studies was context-dependent and difficult to get general insights of a problem on another scenario [26]. It is also a challenge for researchers who were not aware of Noise reduction, integrating the data, feature extraction, technically.

6) Other factors:

Thus from the real time perspective, while designing a MMLA study the following parameters need to be considered such as the Targeted population(age, sample size) , objectives of the study (predicting learning outcomes) the types of modalities with knowledge of their qualities, and other collective capacities (*i.e., their ability to combine certain modalities to investigate different phenomena, such as combining video to find the moments of collaboration with eye-tracking to investigate the quality of those moments)[26]*. Another most relevant challenges of the multimodal data includes the high dimensionality, missing data, modality resolutions, fusing data techniques has to be explored [18]. The research work need to be done more in the area of feature extraction from each modality effects the educational constructs [30].

4. CONCLUSION

This paper concludes the importance of MMLA in field of science and technology. The paper focusses on the model for data collection using sensors, processing using technologies , analysing and predicting using computational methods. The MMLA research is growing still and more research needs to be done on the challenges mentioned above to reach the fullest potential in the future.

5. REFERENCES:

[1] Siemens, G.; Baker, R.S.J.d. "**Learning analytics and educational data mining: Towards communication and collaboration."** In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012), New York, NY, USA, 29 April–2 May 2012; pp. 252–254.

[2] Schwendimann, B.A.; Rodríguez-Triana, M.J.; Vozniuk, A.; Prieto, L.P.; Boroujeni, M.S.; Holzer, A.; Gillet, D.; Dillenbourg, P. "**Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research"**. IEEE Trans. Learn. Technol. 2017, 10, 30–41.

[3] Eradze, M.; Laanpere, M. "**Lesson Observation Data in Learning Analytics Datasets: Observata**". In Proceedings of the 12th European Conference on Technology-Enhanced Learning (EC-TEL 2017), Tallinn, Estonia, 12–15 September 2017; pp. 504–508.
[4] Rodríguez-Triana, M.J.; Prieto, L.P.; Vozniuk, A.; Boroujeni, M.S.; Schwendimann, B.A.; Holzer, A.; Gillet, D. "**Monitoring, awareness and reflection in blended technology enhanced learning: A systematic review".** Int. J. Technol. Enhanc. Learn. 2017, 9, 126–150.
[5] Di Mitri, D.; Schneider, J.; Specht, M.; Drachsler, H. "**From signals to knowledge: A conceptual model for multimodal learning analytics".** J. Comput. Assist. Learn. 2018, 34, 338–349.
[6] Alejandro Andrade, Ginette Delandshere, Joshua A. Danish, **"Using Multimodal Learning Analytics to Model Student Behaviour:A Systematic Analysis of Behavioural Framing",** Journal of Learning Analytics, 3(2),282-306.,2016. http://dx.doi.org/10.18608/jla.2016.32.14
[7] Paulo Blikstein, Marcelo Worsley, "**Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks",** Journal of Learning Analytics, 3(2), 220-238., 2015, http://dx.doi.org/10.18608/jla.2016.32.11.
[8] Quan Nguyen, Bart Rienties , Michal Huptych**, "Linking students' timing of engagement to learning design and academic performance",** LAK'18, March 7—9, 2018, Sydney, NSW, Australia.
[9] Pardo , A. , Han , F. , & Ellis , R. A. (2016). "**Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance"**. IEEE Transactions on Learning Technologies,10(1), 82 – 92 .
[10] Cukurova , M. , Giannakos , M. , & Martinez-Maldonado , R. (2020 **). "The promise and challenges of multimodal learning analytics"**, British Journal of Educational Technology.
[11] Mangaroska , K. , & Giannakos , M. N. (2018). **"Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning**". IEEE Transactions on Learning Technologies.
[12] Mangaroska , K. , Vesin , B. , & Giannakos , M. (2019). "**Cross-platform analytics: A step towards personalization and adaptation in education"**. In Proceedings of the Ninth International Conference on Learning Analytics & Knowledge (pp. 71 – 75). New York, NY : ACM .
[13] Prieto , L. P. , Sharma , K. , Dillenbourg , P. , & Jesús , M. (2016). "**Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors**". In Proceedings of the Sixth International Conference on Learning Analytics and Knowledge (pp. 148 – 157). New York, NY: ACM .
[14] Mangaroska , K. , Sharma , K. , Giannakos , M. , Træteberg , H. , & Dillenbourg , P. (2018)."**Gaze-driven design insights to amplify debugging skills: A learner-centred analysis approach**". Journal of Learning Analytics, 5 (3), 98 – 119 .
[15] Blikstein , P. (2013). **Multimodal learning analytics** . In Proceedings of the Third International Conference on Learning Analytics and Knowledge (pp. 102 – 106). New York, NY : ACM.
[16] Sharma , K. , Papamitsiou , Z. , & Giannakos , M. (2019). **Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach** . British Journal of Educational Technology , 50 (6), 3004 – 3031 .

[17] Liu , R. , Stamper , J. , Davenport , J. , Crossley , S. , McNamara , D. , Nzinga , K. , & Sherin , B. (2019). **“Learning linkages: Integrating data streams of multiple modalities and timescales”**. Journal of Computer Assisted Learning , 35 (1), 99 – 109
[18] Lahat, D., Adali, T., & Jutten, C. (2015). **“Multimodal data fusion: An overview of methods, challenges, and prospects”.** Proceedings of the IEEE, 103(9), 1449–1477. http://doi.org/10.1109/JPROC.2015.2460697
[19] Giannakos , M. N. , Sharma , K. , Pappas , I. O. , Kostakos , V. , & Velloso , E. (2019). “**Multimodal data as a means to understand the learning experience”**,International Journal of Information Management , 48 , 108 – 119 .
[20] Di Mitri , D. , Scheffel , M. , Drachsler , H. , Börner , D. , Ternier , S. , & Specht , M. (2017). **“Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data.”**
[21] In Proceedings of the Seventh International Learning Analytics and Knowledge Conference (pp. 188 – 197). New York, NY : ACM.
[22] Spikol , D. , Ruffaldi , E. , Dabisias , G. , & Cukurova , M. (2018). **“Supervised machine learning in multimodal learning analytics for estimating success in project-based learning.”** Journal of Computer Assisted Learning , 34 (4), 366 – 377.
[23] Florian-Gaviria , B. , Glahn , C. , & Gesa , R. F. (2013). **“A software suite for efficient use of the European qualifications framework in online and blended courses”.** IEEE Transactions on Learning Technologies , 6 (3), 283 – 296 .
[24] Ochoa , X. , Domínguez , F. , Guamán , B. , Maya , R. , Falcones , G. , & Castells , J. (2018). **“The rap system: Automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors.”** In Proceedings of the Eighth International Conference on Learning Analytics and Knowledge (pp. 360 – 364). New York, NY : ACM.
[25] Casey , K. (2017). **“Using keystroke analytics to improve pass-fail classifiers”** . Journal of Learning Analytics , 4 (2), 189 – 211 .
[26] Kshitij Sharma , Michail Giannakos **,“Multimodal data capabilities for learning”**, British Journal of Educational Technology doi:10.1111/bjet.12993 Vol 51 No 5, 2020 1450–1484
[27] Florian-Gaviria , B. , Glahn , C. , & Gesa , R. F. (2013). “**A software suite for efficient use of the European qualifications framework in online and blended courses**”. IEEE Transactions on Learning Technologies , 6 (3), 283 – 296 .
[28] Noel , R. , Riquelme , F. , MacLean , R. , Merino , E. , Cechinel , C. , Barcelos , T. S. , … Munoz , R. (2018). “**Exploring collaborative writing of user stories with multimodal learning analytics: A case study on a software engineering course**” . IEEE Access , 6 , 67783 – 67798 .
[29] Su Mu , Meng Cui , Xiaodi Huang ,“**Multimodal Data Fusion in Learning Analytics: A Systematic Review”**, Sensors 2020, 20, 6856; doi:10.3390/s20236856. www.mdpi.com/journal/sensors
[30] Emerson, A., Azevedo, R., & Lester, J. (2020). “**Multimodal learning analytics for game-based learning”.** British Journal of Educational Technology, 51(5). https://doi.org/10.1111/bjet.12992
[31] Olsen, J., Sharma, K., Rummel, N., & Aleven, V. (2020). “**Using multimodal data to temporally analyze collaborative learning outcomes: Benefits and challenges**” , British Journal of Educational Technology, 51(5).
[32] Dana Lahat, Tülay Adalı, Christian Jutten, **“Multimodal Data Fusion: An Overview of Methods, Challenges and Prospects”**, Proceedings of the IEEE, Institute of Electrical

and Electronics Engineers, 2015, Multimodal Data Fusion, 103 (9), pp.1449-1477. ff10.1109/JPROC.2015.2460697ff.

[33] Junokas , M. J. , Lindgren , R. , Kang , J. , & Morphew , J. W. (2018). "**Enhancing multimodal learning through personalized gesture recognition"** . Journal of Computer Assisted Learning , 34 (4), 350 – 357 .

[34] Andrade , A. (2017). Understanding student learning trajectories using multimodal learning analytics within an embodied-interaction learning environment . In Proceedings of the Seventh International Learning Analytics and Knowledge Conference (pp. 70 – 79). New York, NY : ACM .