Multimodal challenge: analytics beyond user-computer interaction data

Daniele Di Mitri
Open University of The Netherlands
daniele.dimitri@ou.nl

Jan Schneider
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
schneider.jan@dipf.de

Hendrik Drachsler
Open University of The Netherlands – DIPF
hendrik.drachsler@ou.nl

ABSTRACT: this contribution describes one the challenges explored in the Fourth LAK Hackathon. This challenge aims at shifting the focus from learning situations which can be easily traced through user-computer interactions data and concentrate more on user-world interactions events, typical of co-located and practice-based learning experiences. This mission, pursued by the multimodal learning analytics (MMLA) community, seeks to bridge the gap between digital and physical learning spaces. The "multimodal" approach consists in combining learners' motoric actions with physiological responses and data about the learning contexts. These data can be collected through multiple wearable sensors and Internet of Things (IoT) devices. This Hackathon table will confront with three main challenges arising from the analysis and valorisation of multimodal datasets: 1) the data collection and storing, 2) the data annotation, 3) the data processing and exploitation. Some research questions which will be considered in this Hackathon challenge are the following: how to process the raw sensor data streams and extract relevant features? which data mining and machine learning techniques can be applied? how can we compare two action recordings? How to combine sensor data with Experience API (xAPI)? what are meaningful visualisations for these data?

Keywords: multimodal learning analytics, wearables, CrossMMLA, sensor-based learning

1 BACKGROUND

The Learning Analytics & Knowledge community has acknowledged the need for extending the analysis of learning to more diverse data sources going beyond the conventional online learning systems, MOOC events or student information systems. This need stems from the necessity of taking into account physical and co-located interactions which still represent the bulkiest set of learning activities. Multimodal datasets can provide historical evidence and description of the learning process, i.e. the learner's behaviour and learning context. These data are collected automatically through wearable sensors, IoT devices and computer logs and therefore can capture only "what is visible" to some generic sensor. Such definition makes multimodal data conceptually separated by other human-driven qualitative interpretations like expert reports or teacher assessments. The latter are interpretations which describe dimensions unobservable with sensors, such as learning outcomes, cognitive aspects or affective states during learning. The analysis of multimodal data for learning has grown into a field of research called Multimodal Learning Analytics (Blikstein, 2013). These types of analytics seek to bridge complex behavioural patterns with learning theories (Worsley, 2014). In related work the multimodal approach was used in dialogic learning, during teacher-student discourses during lectures (D'mello et al., 2015); computer-supported collaborative learning during knowledge-sharing and group discussions (Martinez-maldonado et al., 2017; Schneider & Blikstein, 2015); and practice-based and open-ended learning tasks, when understanding and executing a practical learning tasks (Ochoa et al., 2013). The analysis of multimodal data in learning is a fairly new but steadily growing field of research which need support: The LAK community still lacks a programmatic approach for modelling the learning process and producing valuable feedback with multimodal data. In our understanding, this approach should clarify the collection, storage, analysis and exploitation of the multimodal data in a pragmatic and scalable manner, which can be adopted into real-life educational settings.

2 THE CHALLENGES

When describing and analysing learning with multimodal data, there still exist many open challenges (Blikstein & Worsley, 2016). For the LAK Hackathon, we identify three main challenges arising from the data-empowered feedback loop of multimodal data and learning analytics: 1) the data collection and storing; 2) the data annotation (or triangulation); 3) the data processing and exploitation.

2.1 Data collection and storing

The first step of the journey is the *data collection* and the data creation with new multimodal datasets. The sensors are most likely from different vendors and have different specifications and support, the approach used for data collection must be flexible and extensible to different sensors which collect data at different frequencies and formats. To address this challenge, we introduce the *LearningHub*, a software prototype whose purpose is to synchronise and fuse the different streams of multimodal data generated by the different sensor-applications while capturing a meaningful part of a learning task, that we call *Action Recording*. The *LearningHub* channels the data from multiple sensors and provides as output JSON files, which serialise and synchronise the values of the sensors for each sensor application. The JSON files allow to have multiple attributes with different time frequencies and formats; they provides also the logic to facilitate the action recording for storing and later retrieval.

- Research questions: 1.1) how to improve the *LearningHub*? 1.2) is the JSON serialisation of the Action Recoding the best approach for storing and retrieving? 1.3) how to link an Action Recording to an xAPI?

2.2 Data annotation

The *data annotation* challenge consists in finding a seamless and unobtrusive approach for labelling the learning process, i.e. triangulating the multimodal action recordings with evidence of the learning activities. To address this challenge, we propose another software prototype: the *Visual Inspection Tool (VIT)*. The VIT is useful for retrospectively analysing and annotating multimodal action recordings. The VIT allows to load multimodal datasets, plot them on a common time scale and triangulate them with video recordings of the learning activity. It allows to select a particular timeframe and annotate the multimodal data slice with a xAPI triplet, assigning an actor, a verb and an object. The VIT offers a human-computer interface which helps to deal with the complexity of multimodal data.

- Research questions: 2.1) how to best improve the VIT? 2.2) how to define and exchange xAPI vocabulary for multimodal activities?

2.3 Data processing and exploitation

The data processing step consists in extracting and aligning the relevant features from the "raw" multimodal data and making them suitable for exploitation, meaning by providing some personalisation benefits to the learner or the teacher. The data processing steps depend on the chosen exploitation strategy. For example, *light-weight feedback* can be generated through hardcoded rules; *historical reports* require different visualisations in that can be grouped into an analytics dashboard; *frequent patterns* or *predictions* require training either unsupervised or supervised models, store them into memory and use them to estimate the value or the class of a particular target attribute.

- Research questions: 3.1) how to process the raw sensor data streams and extract relevant features? 3.2) which data mining and machine learning techniques can be applied? 3.3) how can we compare two action recordings? 3.4) what are meaningful visualisations?

REFERENCES

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 102. http://doi.org/10.1145/2460296.2460316

Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining : using computational technologies to measure complex learning tasks, *7750*(x), 220–238. http://doi.org/http://dx.doi.org/10.18608/jla.2016.32.11

D'mello, S., Olney, A., Blanchard, N., Sun, X., Ward, B., Samei, B., & Kelly, S. (2015). Multimodal Capture of Teacher-Student Interactions for Automated Dialogic Analysis in Live Classrooms. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 557–566. http://doi.org/10.1145/2818346.2830602

Martinez-maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., & Shum, S. B. (2017). Analytics Meet Patient Manikins : Challenges in an Authentic Small-Group Healthcare Simulation Classroom, 1–5. http://doi.org/10.1145/3027385.3027401

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI '13)*, 583–590.

Schneider, B., & Blikstein, P. (2015). Unraveling Students' Interaction Around a Tangible Interface using Multimodal Learning Analytics. *JEDM - Journal of Educational Data Mining*, *7*(3), 89–116.

Worsley, M. (2014). Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors. *3rd Multimodal Learning Analytics Workshop and Grand Challenges, MLA 2014*, 1–4.