MILKI-PSY Cloud: MLOps-based Multimodal Sensor Stream Processing Pipeline for Learning Analytics in Psychomotor Education

Michal Slupczynski[1], Ralf Klamma[1]

[1]*RWTH Aachen University, Aachen, Germany*

Abstract
Psychomotor learning develops our bodies in organized patterns with the help of environmental signals. With modern sensor arrays, we can acquire multimodal data to compare activities with stored reference models of body motions. To do this on a large scale in an efficient way, we need cloud-based infrastructures for the storage, processing and visualization of psychomotor learning analytics data. In this paper, we propose a conceptual sensor stream processing pipeline for this purpose. The solution is based on the so-called MLOps approach, a variation of the successful DevOps model for open source software engineering for large-scale machine learning solutions based on standard components. This processing pipeline will facilitate the multimodal analysis of many training scenarios collaboratively.

Keywords
multimodal learning analytics, explainable AI, cloud infrastructuring, machine learning as a service, psychomotor learning, big data, MILKI-PSY

1. Introduction

Psychomotor learning is the development of muscular activities as organized patterns informed by signals taken from the environment. We research sports activities like running and engineering activities like human-robot collaboration. Streaming data sensors like video, motion capture suits, and IMU sensors are used to acquire data from real-world training scenarios to compare them to stored reference models of motion in the training situation. New streaming data sensors demand big data solutions for the storage, processing and visualization of multimodal learning analytics data. This conceptual paper presents the MILKI-PSY Cloud (MPC), an application of existing Machine Learning for IT Operations (MLOps) approaches to provide a collaborative Machine learning (ML)-based cloud solution to process multimodal sensor streams and provide direct and long-term feedback in psychomotor learning scenarios. We start by describing the background and related work of our contribution in Section 2. In Section 3, we present our data processing pipeline as an application of the MLOps methodology. We conclude this contribution and give an outlook on further research in Section 4.

---

2. Related Work

Multimodal stream data ingestion   Multimodal sensor applications generate vast amounts of heterogeneous and multidimensional data with unique properties [1] that needs to be transmitted, stored and analyzed in a scalable manner. Distributed communication systems such as Apache Kafka [1] provide a foundation optimized for efficiency and low latency ingestion of stream data [2]. The processing of different modalities is then done via data fusion, which can be subdivided into early fusion (merging before the ML process), late fusion (learning each modality independently, then fusing) and cross-modality fusion (learning across modalities) [3].

Time series databases (TSDB)   Time series databases (TSDB) are database systems that are designed for storage, processing, and visualization of time series data (TSD) such as sensor readings [4, 5, 6], which can manage big quantities of information and enable stream data versioning, time-related searches [7] and analysis mechanisms such as trend detection, or anomaly/outlier sensing [1]. One of the most notable examples of TSDB is InfluxDB [2].

Annotation of multimodal data   Raw sensor data requires additional annotation to provide labels for the ML systems to infer the prediction outcomes correctly. Usually, different data modalities require specialized applications to support the annotation process. Several offline and online tools have been proposed for data annotation such as Mova [8], Microsoft PSI [9] or the Visual Inspection Tool [10].

Multimodal data processing pipelines   The usage of processing pipelines for multimodal ML data has been discussed and compared in several publications [3, 11, 12]. Commonly, the pipeline phases include feature extraction, data fusion, model training and exploitation of the trained ML applications. Implementing a user-oriented approach during the development of multimodal learning environments is important to better align to learner needs [12]. A relatively novel addition to multimodal ML pipelines is the application of MLOps principles [13, 14] to build a collaborative environment and provide immersive learning feedback in psychomotor education.

ML artifact management   A versioning system for ML models such as DVC [3] provides crucial tracking and managing of changes to the ML pipeline. A system like TorchServe [4] can then be used to publish and serve the different versions of the ML models. While different stages of the MLOps pipeline can be supported by separate tools, an integrated platform like OpenML [5] can provide all the benefits in a central place.

---

[1]https://kafka.apache.org
[2]https://www.influxdata.com
[3]https://dvc.org/
[4]https://github.com/pytorch/serve
[5]https://openml.org

Machine Learning for IT Operations (MLOps)   MLOps [13, 14] is a DevOps-based [15] practice that aims to reliably [16] and sustainably [17] design, deploy, test and monitor ML models in cloud and edge systems [18, 19]. MLOps extends DevOps as follows: continuous integration requires data and model versioning and validation, continuous delivery considers automatic and steady delivery of the ML models, while continuous testing involves continuous retraining of the ML models to adapt to a steady stream of input data [20, 21, 22].

Explainable AI (XAI)   Explanations are meta-information in ML-based systems that describe the meaning or significance of an input instance for a certain output categorization [23]. A variety of ML mechanisms can be used on sensor-based TSD [24], while the usage of XAI methods facilitate trust and confidence in the models applied to the TSD [25].

3. MILKI-PSY Cloud (MPC): ML-based Multimodal Sensor Stream Processing Pipeline for Direct Learner Feedback

A multimodal sensor network produces continuous data streams with various sampling frequencies, which should be processed in near real-time to offer learning analytics and learner feedback [26]. Figure 1 gives an overview of the proposed MPC infrastructure, which is structured in three phases: data ingestion *(phase 1)*, machine learning *(phase 2)* and data exploitation *(phase 3)*. This contribution is based on the MPC architecture [27].

Feature pipeline *(phase 1)* - Multimodal stream data ingestion   First, a data ingestion system *(1.1)* should be used to provide scalable communication between the sensor devices and the MPC. Annotating the raw sensor data is done in the next step of the pipeline *(1.2)* to provide the labels necessary for the ML feature detection and decision-making. To allow repeatable ML training, the data from multiple concurrent source streams need to be aggregated and stored with some versioning mechanism. A TSDB provides mechanisms for versioned data storage and analysis *(1.3)* and should thus be used for data ingestion and leverage.

ML pipeline, continuous training *(phase 2)*   Next, we use a Federated ML approach for model building and evaluation *(2.1)* [28]. Trained and evaluated ML models need to be versioned *(2.2)*. Once a new version of a model is checked into the storage system, it needs to be deployed and made available *(2.3)*. After training and serving the ML model, the deployed pipeline needs to be monitored and continuously evaluated for the model performance *(2.4)*. The measures used for this evaluation need to be defined by domain experts and calculated automatically in the context of the project depending on the specific ML use case. Alternatively, a MLOps solution *(2.X)* can provide an integrated solution for model building, versioning and serving.

If a drop in accuracy or a sufficiently large amount of new input data (e.g. a new expert recording) is detected through pipeline monitoring *(2.4)*, a retraining of the ML pipeline *(phase 2)* can be triggered as described by continuous training. For example, an ML algorithm could be trained to differentiate whether the feedback given by the system correlates with positive learner progress and change the feedback mechanism if it does not.

Data exploitation & feedback *(phase 3)* With a properly trained and deployed ML model, a frontend application can then use it to provide feedback to the learner *(3.1)*, either directly in near real-time through interactive AR/VR elements, or long-term performance reports and other gamification elements. The usage of explainable AI principles could provide an additional layer of context to the decisions made by the ML algorithm which would benefit the domain experts and learners in gaining insights into the feedback mechanism. This would not only reduce potential bias embedded in the ML model but also enhance the trust of the end-users in the system leading to better learning results. Finally, new multimodal sensor data is generated through learner interaction with the system *(3.2)* leading to a new MPC pipeline cycle with further data ingestion.
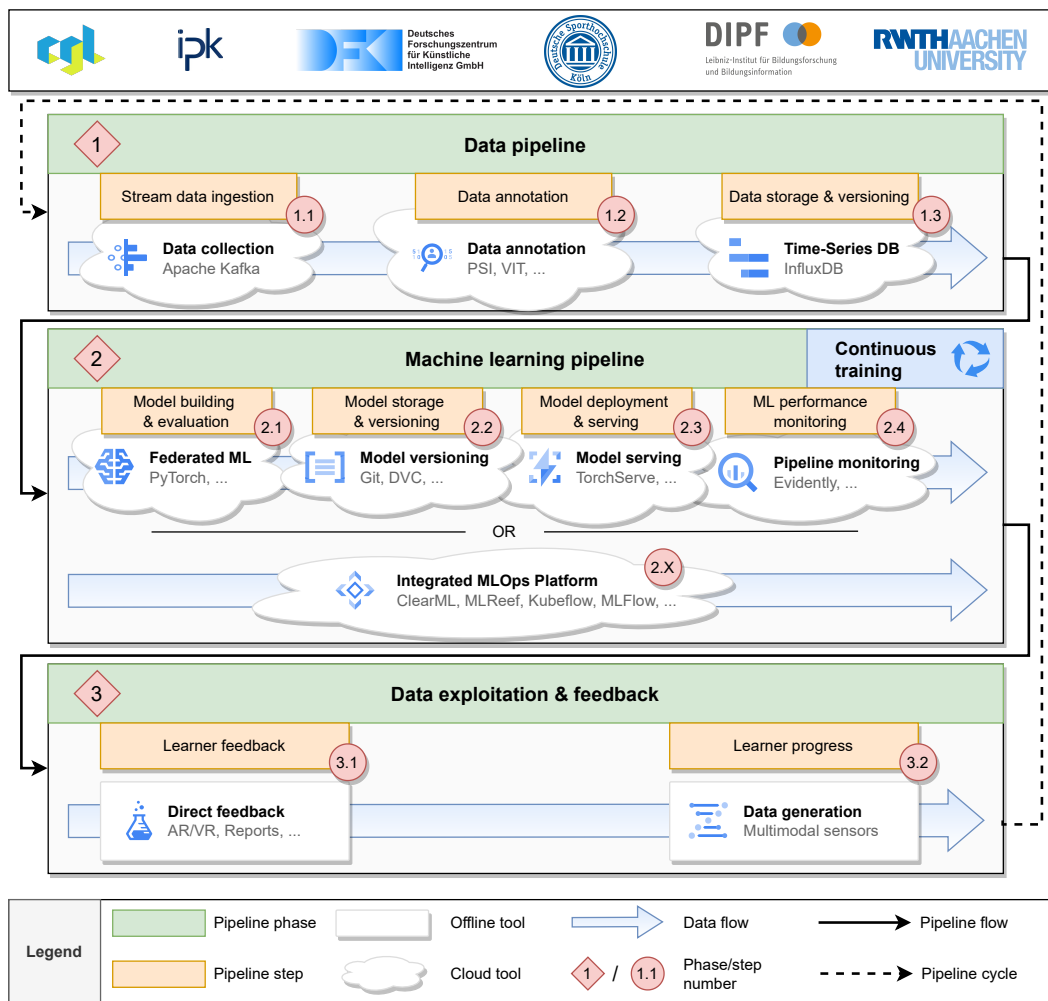
Figure 1: Proposed architecture of the MILKI-PSY Cloud

3.0.1. Stakeholders

At different stages of the proposed pipeline, different stakeholders are involved in decision-making, execution and practice. The stakeholder groups involved in the project range from domain experts to end-users, including pedagogical specialists, data science engineers, coaches, teachers and students. Recording and annotation of training data and raw sensor readings are performed by sports scientists due to their expertise in the psychomotor skills to be taught using the proposed pipeline.

The intricacies of the ML training and evaluation need to be executed by experienced ML experts, while the result of the "data exploitation & feedback" phase will mostly benefit coaches, teachers and learners. For effective interaction with the toolset used in the MPC, each stakeholder group needs to be provided with a different set of interfaces. Sports scientists need a simple way to feed their data into the system and a powerful annotation engine to add meaning to the multimodal data stream. Application developers and designers can make use of the data stored in the TSDB together with the trained and deployed ML models to create immersive learning environments and provide learner feedback. Since the proposed pipeline will be running its toolset in a cloud with powerful hardware, ML experts and researchers will have access to computing resources to process the TSD, so they can focus on model training and development. The versioning in the MLOps pipeline enables coaches, teachers and students not to require extensive ML expertise to make use of the best-performing version of the ML models. The proposed cloud pipeline enables the stakeholders to collaborate on different aspects of the immersive learning environment, from data ingestion and annotation over model training to interaction with the learners.

4. Conclusions and Outlook

This contribution presents a ML-based sensor data processing pipeline to facilitate direct and long-term learner feedback in psychomotor education. Pipeline automation through MLOps is critical in ensuring reproducibility, helping to reduce manual infrastructuring inherent with ML development. The usage of XAI to support the ML process is beneficial in facilitating trust and providing individualized feedback for multimodal psychomotor learning scenarios. Future work includes further implementation and evaluation of the proposed framework in a practical environment. We will focus on an end-user-oriented MLOps execution and evaluation.

Acknowledgments

---

[6]https://milki-psy.de/

References

[1] F. A. Tsvetanov, Storing data from sensors networks, IOP Conference Series: Materials Science and Engineering 1032 (2021) 012012. doi:10.1088/1757-899x/1032/1/012012.

[2] B. R. Hiraman, C. Viresh M., K. Abhijeet C., A study of apache kafka in big data stream processing, in: 2018 International Conference on Information , Communication, Engineering and Technology (ICICET), IEEE, 2018, pp. 1–3. doi:10.1109/ICICET.2018.8533771.

[3] W. C. S. Iv, R. Kapoor, P. Ghosh, Multimodal classification: Current landscape, taxonomy and future directions, ACM Computing Surveys (2022). doi:10.1145/3543848.

[4] T. Dunning, B. E. Friedman, Time series databases: New ways to store and access data / Ted Dunning and Ellen Friedman, first edition ed., O'Reilly, Beijing, 2014.

[5] Google Cloud Architecture Center, Processing streaming time series data: overview, 2021. URL: https://cloud.google.com/architecture/processing-streaming-time-series-data-overview.

[6] S. K. Jensen, T. B. Pedersen, C. Thomsen, Time series management systems: A survey, IEEE Transactions on Knowledge and Data Engineering 29 (2017) 2581–2600. doi:10.1109/TKDE.2017.2740932.

[7] H. Parviainen, The complete guide to time series data, 2021. URL: https://www.clarify.io/learn/time-series-data.

[8] O. Alemi, P. Pasquier, C. Shaw, Mova, Proceedings of the 2014 International Workshop on Movement and Computing (2014) 37–42. doi:10.1145/2617995.2618002.

[9] D. Bohus, S. Andrist, A. Feniello, N. Saw, M. Jalobeanu, P. Sweeney, A. L. Thompson, E. Horvitz, Platform for situated intelligence, 2021. doi:10.48550/arXiv.2103.15975.

[10] D. Di Mitri, J. Schneider, R. Klemke, M. Specht, H. Drachsler, Read Between the Lines, in: unknown (Ed.), Proceedings of the 9th International Conference on Learning Analytics & Knowledge - LAK19, ACM Press, New York, New York, USA, 2019, pp. 51–60. doi:10.1145/3303772.3303776.

[11] T. Baltrusaitis, C. Ahuja, L.-P. Morency, Multimodal machine learning: A survey and taxonomy, IEEE transactions on pattern analysis and machine intelligence 41 (2019) 423–443. doi:10.1109/TPAMI.2018.2798607.

[12] K. Sharma, M. Giannakos, Multimodal data capabilities for learning: What can multimodal data tell us about learning?, British Journal of Educational Technology 51 (2020) 1450–1484. doi:10.1111/bjet.12993.

[13] D. Kreuzberger, N. Kühl, S. Hirschl, Machine learning operations (mlops): Overview, definition, and architecture, 2022. doi:10.48550/arXiv.2205.02302.

[14] N. Talagala, Why mlops (and not just ml) is your business' new competitive frontier, 2018. URL: https://www.aitrends.com/machine-learning/mlops-not-just-ml-business-new-competitive-frontier/.

[15] A. Dyck, R. Penners, H. Lichter, Towards definitions for release engineering and devops, in: 2015 IEEE/ACM 3rd International Workshop on Release Engineering, IEEE, 2015, p. 3. doi:10.1109/RELENG.2015.10.

[16] M. Borg, Agility in software 2.0 – notebook interfaces and mlops with buttresses and rebars, in: A. Przybylek, A. Jarzebowicz, I. Lukovic, Y. Y. Ng (Eds.), Lean and Agile Software Development, volume 438 of *Lecture Notes in Business Information Processing*, Springer

International Publishing, Cham, 2022, pp. 3–16. doi:10.1007/978-3-030-94238-0_1.

[17] D. A. Tamburri, Sustainable mlops: Trends and challenges, in: 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), IEEE, 2020, pp. 17–23. doi:10.1109/SYNASC51798.2020.00015.

[18] C. Min, A. Mathur, U. G. Acer, A. Montanari, F. Kawsar, Sensix++: Bringing mlops and multi-tenant model serving to sensory edge devices, 2021. URL: http://arxiv.org/pdf/2109.03947v1.

[19] E. Raj, D. Buffoni, M. Westerlund, K. Ahola, Edge mlops: An automation framework for aiot applications, in: 2021 IEEE International Conference on Cloud Engineering (IC2E), IEEE, 2021, pp. 191–200. doi:10.1109/IC2E52221.2021.00034.

[20] R. Dawson, Why is DevOps for Machine Learning so Different?, 2019. URL: https://hackernoon.com/why-is-devops-for-machine-learning-so-different-384z32f1.

[21] A. Tripathi, Mlops: A complete guide to machine learning operations: Mlops vs devops, 2021. URL: https://ashutoshtripathi.com/2021/08/18/mlops-a-complete-guide-to-machine-learning-operations-mlops-vs-devops/.

[22] Y. Zhou, Y. Yu, B. Ding, Towards mlops: A case study of ml pipeline platform, in: 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE), IEEE, 2020, pp. 494–500. doi:10.1109/ICAICE51518.2020.00102.

[23] A. Das, P. Rad, Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey, 2020. URL: http://arxiv.org/pdf/2006.11371v2.

[24] L.-R. Jácome-Galarza, M.-A. Realpe-Robalino, J. Paillacho-Corredores, J.-L. Benavides-Maldonado, Time series in sensor data using state-of-the-art deep learning approaches: A systematic literature review, in: Á. Rocha, P. C. López-López, J. P. Salgado-Guerrero (Eds.), Communication, Smart Technologies and Innovation for Society, volume 252 of *Smart innovation, systems and technologies*, Springer Singapore, Singapore, 2022, pp. 503–514. doi:10.1007/978-981-16-4126-8_45.

[25] T. Rojat, R. Puget, D. Filliat, J. Del Ser, R. Gelin, N. Díaz-Rodríguez, Explainable artificial intelligence (xai) on timeseries data: A survey, 2021. doi:10.48550/arXiv.2104.00950.

[26] M. K. Garg, D.-J. Kim, D. S. Turaga, B. Prabhakaran, Multimodal analysis of body sensor network data streams for real-time healthcare, in: J. Z. Wang, N. Boujemaa, N. O. Ramirez, A. Natsev (Eds.), Proceedings of the international conference on Multimedia information retrieval - MIR '10, ACM Press, New York, New York, USA, 2010, p. 469. doi:10.1145/1743384.1743467.

[27] M. Slupczynski, R. Klamma, Milki-psy cloud: Facilitating multimodal learning analytics by explainable ai and blockchain, in: R. Klemke, K. Sanusi, D. Majonica, A. Richert, V. Varney, T. Keller, J. Schneider, D. Di Mitri, ..., N. Riedl (Eds.), First International Workshop on Multimodal Immersive Learning Systems (MILeS 2021), 2021, pp. 22–28. URL: http://ceur-ws.org/Vol-2979/paper3.pdf.

[28] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. Nitin Bhagoji, K. Bonawitz, Z. Charles, ..., S. Zhao, Advances and open problems in federated learning, Foundations and Trends in Machine Learning 14 (2021) 1–210. doi:10.1561/2200000083.