

D4.8 Working Prototype and Designs (Revised)

This Deliverable D4.8 Working Prototype and Designs (Revised) outlines the current stage of development of our interactive prototype pipeline 'AMLExplorer,' which is aimed to easily train, assess, and apply AML-models in the context of expressive gesture use in music exploration. We investigate how to provide users with an easy introduction to AML and its learning process by visualising the algorithm's results, limitations, and making its results explorable. This aims to accomplish two goals: first, to contribute to general ML literacy by allowing novices to train and explore new Machine Learning (ML) approaches such as AML, and second, to provide an easy tool for ML experts to explore and evaluate AML possibilities.

This deliverable is the first step towards reaching these objectives. In order to enable more creative professionals to express themselves freely, gestures are increasingly used in Human Computer Interaction (HCI) incorporating different types of activity recognition. We use the MediaPipe Hand Pose detection model available via Tensorflow [6] to recognize full hand gesture input based on a 3D hand model to be processed and classified by AML. While in earlier deliverables we used this input to determine and represent a labelled class, this time we allow the user to train the algorithm with their own input data, making the result interactive and directly applicable. This increases the possibility of faster interaction cycles between the user and the AML model by giving it more expressiveness. We present a pipeline where we apply the AML model to recognize personal gestures and make them interactive and fluid.

In short, we introduce an interactive AML system for musical gestures, that:

1. accepts user-defined input gestures
2. trains an AML model based on a user-created data set
3. allows to explore the AML capabilities in comparison to other ML approaches
4. provides an interactive way to explore new sound patterns using AML.

We then present two use cases enabled by AMLExplorer: a creative use case aimed at creative professionals and a more educative use case aimed at computer scientists interested in AML. We describe the design of the various AMLExplorer interfaces that we created to assist users in collecting personalised data, training an AML model, understanding and effectively using AML classifications for creating new sound patterns through gestures. Finally, we describe the underlying interactive AML pipeline's implementation. We conclude with an overview of the pipeline's potential for ML literacy as well as future work on and with AML algorithms.