OpenML Sharing Deep Learning Models

A.Danila, B.Enache, G.Goncheva, L.A.Hijl, A.Mares, V.Minev, T.M.Nguyen, A.L.L.Nijhuis, C.T. Nohai, D.M.Serban, T.Zahariev, S.Zarkova

What is machine learning?

Nowadays, Artificial Intelligence (AI) is becoming extremely popular. It comes to mind when we hear about self-driving cars, speech recognition or many of the smart algorithms that surround us. Without any doubt, it is an exciting time to be alive and to be part of all these technological improvements. We can "teach" computers by feeding them information in the form of observations. This is called Machine Learning and is apart of AI. Every child needs to be shown a few different dogs, before they can recognize the next dog they see and distinguish it from any other animal. The same holds for the computer. There are different patterns and similarities in data and the way of helping the computer "observe" these can be quite different. This is why researchers in the field of Machine Learning are eager to collaborate and share their insights.

How can OpenML help you?

OpenML is the name of an online platform which facilitates the sharing of machine learning tasks, models and benchmarks between users. Anyone can create its own machine learning model and train it with the available tasks on the platform. If a certain task is not available, the user can create and upload one. Afterwards, he can upload the model together with its benchmark (i.e. accuracy) on the platform. In this way, other users can see how good that model is and decide whether to use it or not. If there is no good model available for a certain task, the user can try to create a better one and upload it on the platform. This site is aimed to help the machine learning community by giving the users the choice to pick the best model available for a certain task.
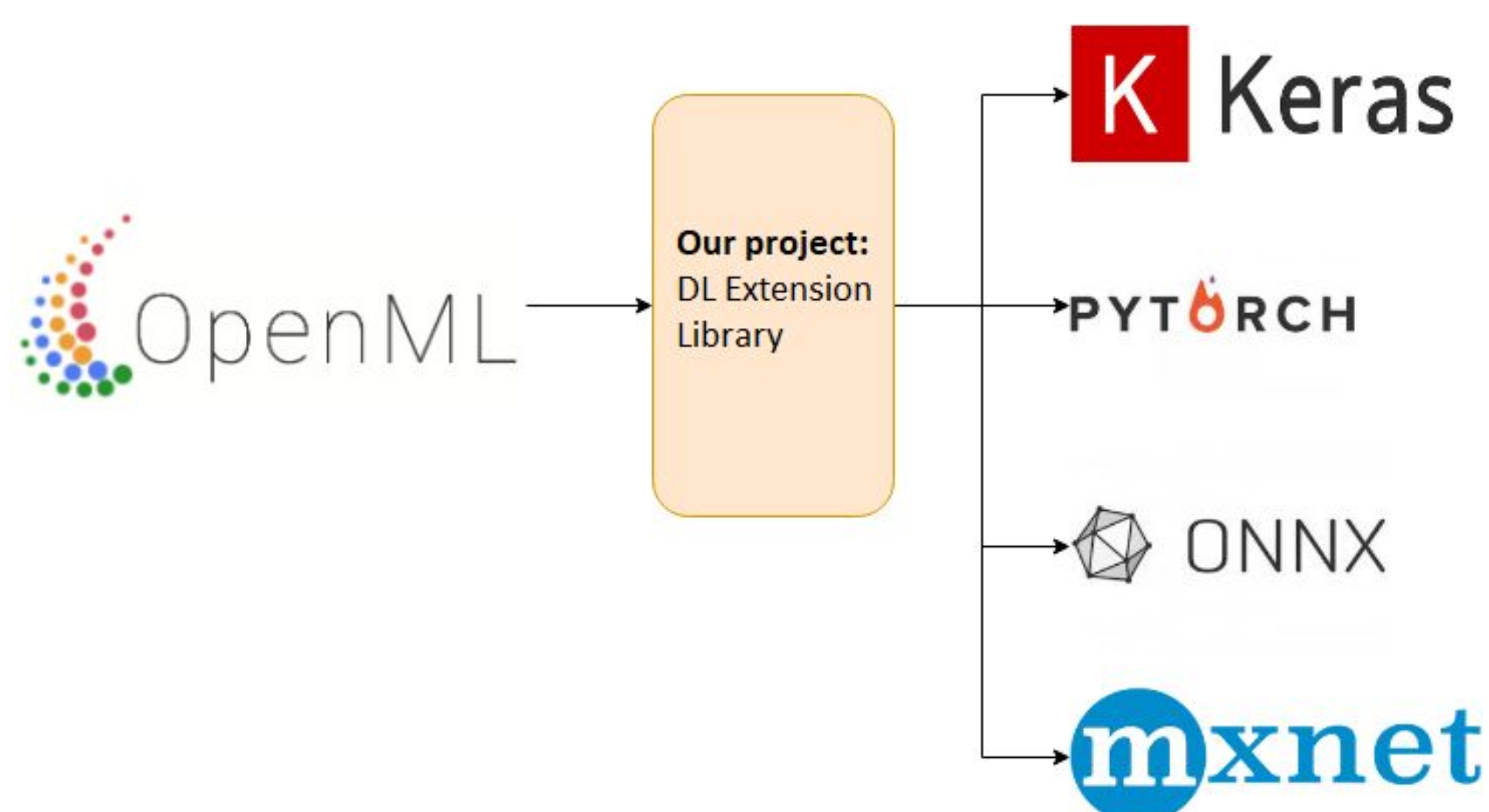

Figure 1: The objective of the project

Problems to be solved

Deep learning libraries

Currently, the OpenML-Python API (the API used to interact with OpenML) supports only machine learning models created with the Scikit-learn framework. The aim is to extend the functionality of the OpenML-Python API to support more frameworks such as Keras, PyTorch, MXNet, ONNX. This is depicted in Figure 1 as well.

Visualization

OpenML does not have any functionality to display the neural network structure of a deep learning model. It also does not offer the user any information related with the learning curve graph of the model after training.
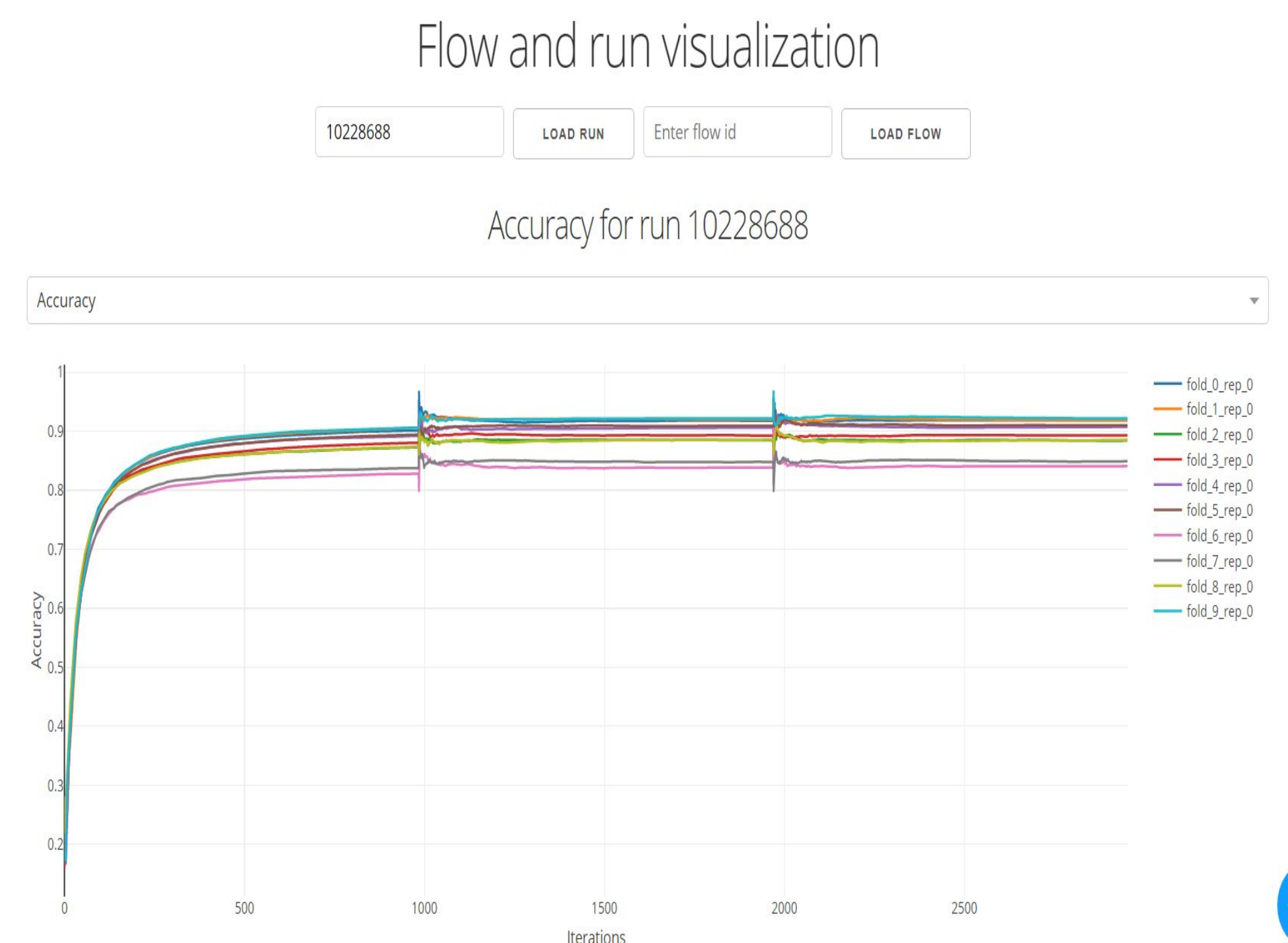

Figure 2: The learning curve of a model

Our solution

OpenML-Python API has its own format to interact with the machine learning models. The solution that the team came up with is to convert the format for each of the deep learning frameworks (Keras, PyTorch, MXNet and ONNX) into the format supported by the OpenML-Python API. In this way all functionalities that OpenML-Python API supports with Scikit-learn framework will be supported by the new frameworks.

For visualization, the Dash framework is being used to generate all the requested graphs.