

# MLJAR - Platform for building machine learning models

**Piotr Płoński**

Warsaw University of Technology



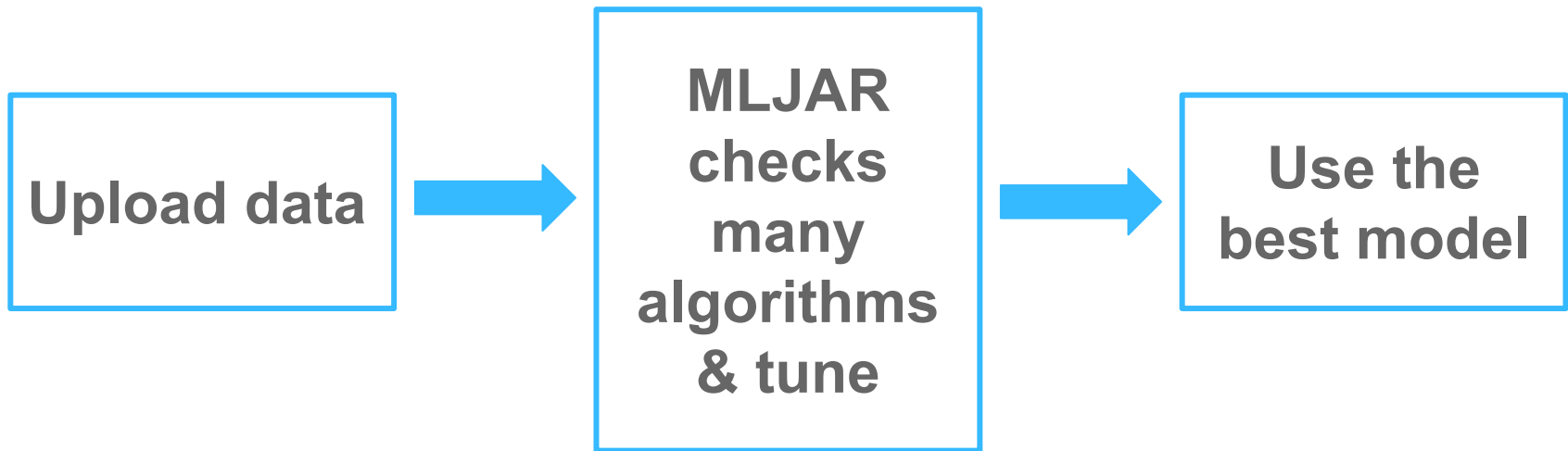
# The problem

- Building good ML model requires checking a lot of learning algorithms and tuning
- People don't have time or patience to do this
- There is a need for proper tools ...



## The Tool - MLJAR

- MLJAR is a platform for data scientists for building ML models
- It works with tabular data and images (can be extended)



**Demo time!**

# SUSY Benchmark

AUC on test set	Low level	High level	Complete
BDT from P.Baldi et al, TMVA implementation	0.850	0.835	0.863
BDT, from mljar.com, xgboost implementation	0.858	0.867	0.877

# The needs

- Web part is on AWS servers (cost is on me)
- Computations can be done on AWS (\$\$\$)
- Or on local machines (with Nvidia GPUs)



# Data from neutrino experiments

- Data is sparse
- Efficient algorithms are needed

btgraham / SparseConvNet

Watch 25 Star 169 Fork 71

Code Issues 3 Pull requests 0 Projects 0 Wiki Pulse Graphs

Spatially-sparse convolutional networks. Allows processing of sparse 2, 3 and 4 dimensional data. Build CNNs on the square/cubic/hypercubic or triangular/tetrahedral/hyper-tetrahedral lattices.

<https://github.com/btgraham/SparseConvNet/wiki>

## keras2cpp

This is a bunch of code to port Keras neural network model into pure C++.

★ 17 ● C++

## keras-sparse-check

Compare training on dense vs sparse data on data from neutrino experiments

★ 0 ● Python

**We need proper tools for sparse data!**

