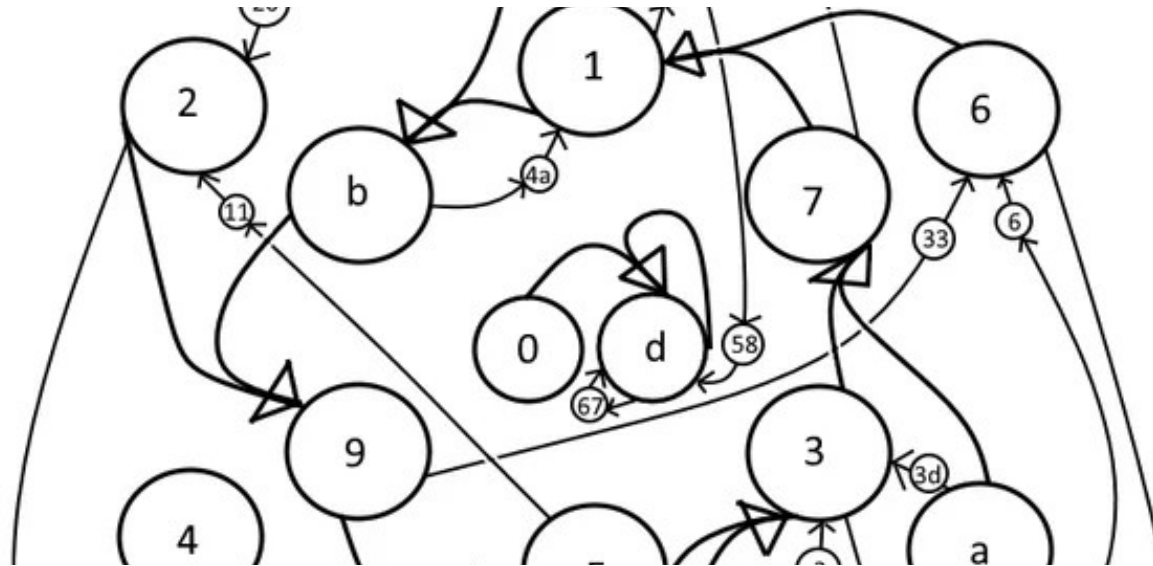# ML Exercise Preliminary Plan

Matthias Komm & Jörn Bach

CMS DAS 13.10.-17.10.2025

DESY, Hamburg

Conceptual Idea

- 1. Introduction, Motivation → 10min presentation by facilitators

- 2. All: Intro Exercise on data preparation via uproot, numpy, matplotlib

  → goal: exploratory data analysis, dataset: ttbar semileptonic

  → some feature reconstruction: e.g. W mass peak reconstruction

- 3. Students choose 1 (or 2) of 4 exercises:

  – Binary classification

  – Multiclass classification

  – Regression

  – Convulational DNN for regression

- Available from earlier schools: ttH, ttW, DY sample or tt semileptonic sample (few 10k events)

- Based on jupyter notebooks

- Datasets from MC, saved in root

- Potential bottleneck:

  - We will need some sort of GPU nodes

  - maybe if there's temporary DESY accounts give them GPU quota for these exercises

  - Experimentation is encouraged so training times should be kept fast!