* Dataset of 250000 events, with an ID column, 30 feature columns, a weight column and a label column.

Some details to get started:

* all variables are floating point, except PRI\_jet\_num which is integer
* variables prefixed with PRI (for PRImitives) are “raw” quantities about the bunch collision as measured by the detector.
* variables prefixed with DER (for DERived) are quantities computed from the primitive features, which were selected by  the physicists of ATLAS
* it can happen that for some entries some variables are meaningless or cannot be computed; in this case, their value is −999.0, which is outside the normal range of all variables

For detailed information on the semantics of the features, labels, and weights, see the [technical documentation](http://higgsml.lal.in2p3.fr/documentation) from the [LAL website](http://higgsml.lal.in2p3.fr/) on the task.