



Contribution ID: 138

Type: **not specified**

b-hive: a modern start-to-finish flavor tagging framework

Deep-learning-based algorithms have become essential for a large fraction of tasks in high-energy-physics (HEP) research, like the identification of jets originating from heavy quarks at the Compact Muon Solenoid (CMS) Experiment at the Large Hadron Collider (LHC) at CERN. Existing frameworks are often limited in capability and based on older software libraries by today's industry standards.

With our newly developed framework "b-hive", we aim to build a full package capable of handling a flavor tagging model with modern tools. From reading HEP-specific file formats, such as ROOT, to constructing datasets for a model, training, and evaluating a model, b-hive bundles all in one cohesive framework. It uses common Python packages provided by a conda environment in which "coffea", "law" and "PyTorch" are the key elements. Using a conda environment makes deployment on different computing sites straightforward. Finally, versioning of tasks by "law" improves reproducibility.

Modularity being a key aspect of b-hive, it is easy to adjust the framework depending on user needs: whole modules can be exchanged and new ones can be added and integrated into the pipeline. For smaller changes and more common distinctions between the needs of different working groups within a large collaboration, a global configuration is used to prevent divergence in the code base, making it suitable for parallel developments.

Author: JUNG, Alexander (RWTH Aachen University)

Presenter: JUNG, Alexander (RWTH Aachen University)

Session Classification: Poster Session

Track Classification: Performance and Upgrade Tools