**EUROPEAN AI FOR
FUNDAMENTAL PHYSICS
CONFERENCE
EuCAIFCon 2024**

Contribution ID: **99**   Type: **Flashtalk with Poster**

Quark/gluon tagging in CMS Open Data with CWoLa and TopicFlow

Wednesday, 1 May 2024 16:03 (3 minutes)

Methods for training jet taggers directly on real data are well motivated due to both the ambiguity of parton labels and the potential for mismodelled jet substructure in Monte Carlo. This talk presents a study of weakly-supervised learning applied to Z+jet and dijet events in CMS Open Data. In order to measure the discrimination power in real data, we consider three different estimates of the quark/gluon mixture fractions. These fractions are then used to train TopicFlow: a deep generative model that disentangles quark and gluon distributions from mixed datasets. We discuss the use of TopicFlow both as a generative classifier and as a way to overcome limited statistics.

Primary authors: ORE, Ayodele (Heidelberg University - ITP); GARGALIONIS, John (University of Valencia - IFIC); DOLAN, Matthew (University of Melbourne)

Presenter: ORE, Ayodele (Heidelberg University - ITP)

Session Classification: 4.1 Pattern recognition, Image analysis & Uncertainty quantification

Track Classification: Session B