ID de Contribution: **17** Type: **Non spécifié**

A Data-centric AI based diagnosis method using unsupervised labeling and MLP

The quality of Input data has a profound impact in fault diagnosis. To construct high quality datasets for required diagnosis performance, it is vital to eliminate ambiguities in featured data. An example in our case study is the undesirable similarity between I-V curves in both open-circuit fault and partial shading fault. In this study, data-centric AI architecture is shaped on both I-V curve training data development and fault data maintenance. The diagnosis is based on an automated cluster-then-label algorithm. In the training data development step, all featured data are clustered to study the hidden meaning of each cluster representative. After assigning an initial label to each cluster, a data quality valuation runs and the assessment verifies representatives to detect qualified data with distinct fault labels. According to the assessment results, the iterative data maintenance process performs the data quality improvement. The objective of quality improvement is to eliminate the cluster ambiguities in a more smart and concise way. Our preliminary results of our data-centric AI-based diagnosis illustrate higher understandability than other state of art unsupervised fault diagnosis.

Auteur principal: M. ZARGARANI, mohsen (CentraleSupelec, Institute GEEPs)

Orateur: M. ZARGARANI, mohsen (CentraleSupelec, Institute GEEPs)

Classification de thématique: Data frugal Approaches, Data-centric AI