

Prospects for automatic data quality monitoring at the CMS pixel detector using machine learning

Luka Lambrecht, on behalf of the CMS Collaboration e-mail: luka.lambrecht@cern.ch



Figure from [1]



- Goals: optimal usage of the LHC delivered luminosity. filter compromised data from certified data.
- Data quality monitoring: spot detector issues in real time.
- Data certification: certify data as good quality for physics analyses.
- Current manual procedure has some disadvantages:
 - Very labour intensive.
 - Sensitive to vizualization details and human errors.
 - Coarse time granularity (run based opposed to luminosity section based).



CMS Preliminary 2017 (13 TeV) CMS Preliminary CMS Preliminary 2017 (13 TeV) 2017 (13 TeV) The pixel tracker and its monitoring elements — Good - Good Good - Run 306139, LS 1112 - Run 306139, LS 1112 ----- Run 306139, LS 1112 **BPIX** FPIX Reconstruction Reconstruction Reconstruction Innermost CMS subdetector. BPIX L3 **BPIX L2** η=1.0 BPIX L4 D3 η=2.0 L4 • Crucial for charged particle tracking. r=160m Consists of 4 barrel layers and 3 endcap disks. L3 alized η=2.5 r=109mn • Monitoring elements (among others): L2 r=68mr • electric charge per track cluster. r=29mm • pixel occupancy. 20000 40000 60000 20000 40000 60000 20000 40000 Cluster charge (electrons) Cluster charge (electrons) -50.0 cm Cluster charge (electrons) Figure from [4] Figure from [2] Inputs Set of histograms per luminosity section. Cluster charge distributions split by region: BPIX L2, L3, L4, FPIX- D1, D2, D3, FPIX+ D1, D2, D3. Preprocessing Normalization. Filter on detector status. Filter on sufficient number of entries. Challenges







taking).

- Autoencoder and NMF models more accurate than baseline models.
- A few anomalies found in previously manually certified data, traced down to high voltage tests and beam dump effects.
- Remaining issues:
 - sensitivity to discrete detector condition changes in between runs.
 - how to optimally handle luminosity sections with few events.





Future developments

- Optimize choice of reference histograms for local training.
- Extend to other monitoring elements.
- Further validation and commissionning in Run-3 data.
- Implement in online DQM software for live data taking.

References

[1]: Public CMS Data Quality Information, link

- [2]: The CMS Phase-1 Pixel Detector Upgrade, CMS Note 2020/005, link
- [3]: Tracker DQM Machine Learning studies for data certification, CMS Note 2021/034, link
- [4]: Soon to be made public.



Poster presented at the 15th Pisa Meeting on Advanced Detectors, 22-28 May 2022