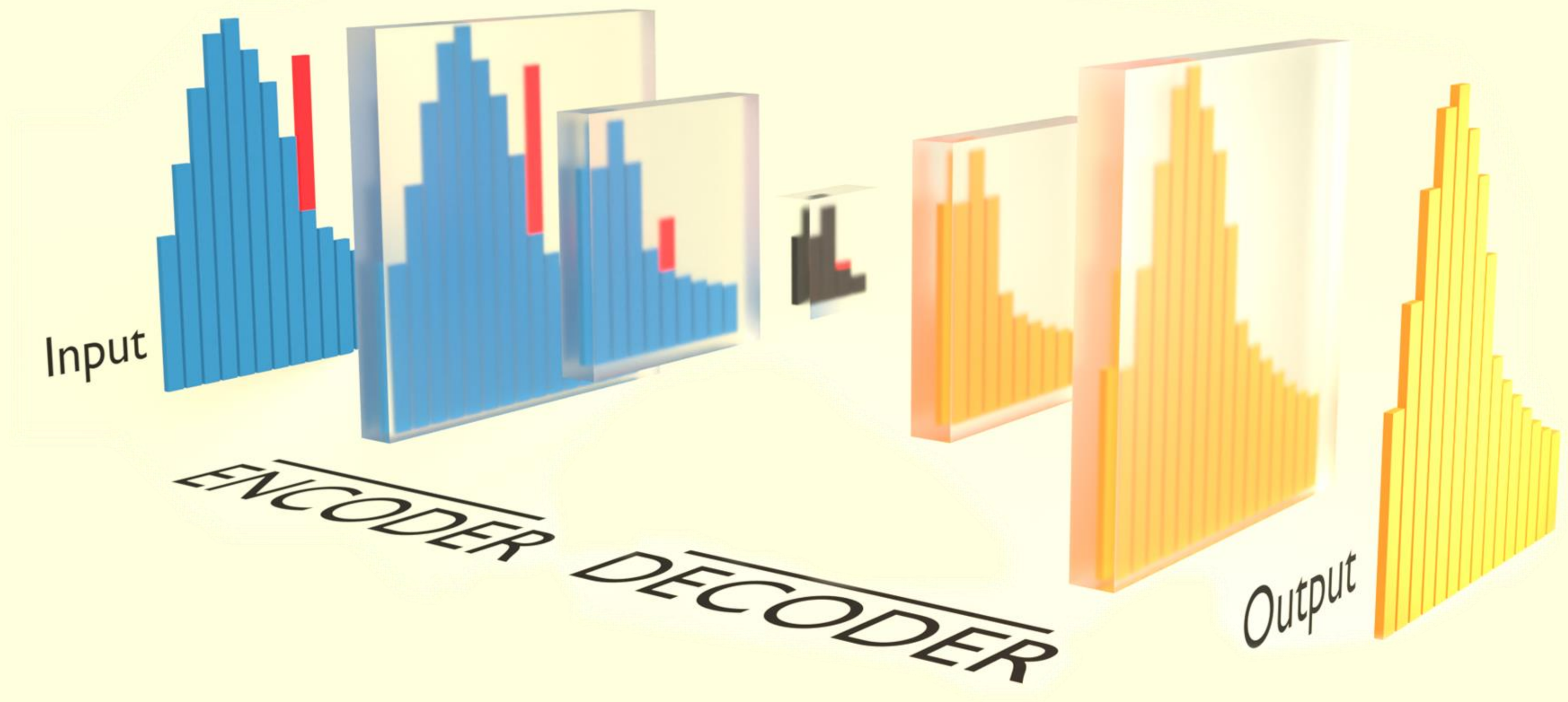
ML-based anomaly detection in jets and missing momentum for data quality monitoring in Run 3 in CMS

Alkis Papanastassiou[1,2,3] *

On behalf of the CMS Collaboration

[1]Istituto Nazionale di Fisica Nucleare, [2]Università degli studi di Firenze, [3]Baker Hughes

* alkis.papanastassiou@cern.ch



Motivation

- CMS data are gathered in time intervals (*lumisections* / LSs) ~$23.31s$.
- Data certification works on Run level → $\mathcal{O}(1000)$ LSs.
- Experts monitor several reconstructed distributions → Monitor Elements (MEs).
- Misbehaving subsystems in one LS would cause **the rejection of a whole run**.
- Certifying every LS is unfeasible without an automated approach → **AutoEncoder** (AE) [1]

The Model

- AE: **unsupervised** neural network → no truth labels needed for training.
- **Training**:
  - on a chosen ME, performed on a *GOOD* run per-LS data.
  - performed through the minimization of the **reconstruction error** (mean squared error):

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{y_i} - \hat{y}_i)^2 \ ,$$

  where y and $\hat{y}$ are respectively the input and the output of the AE and n is the bin number.
- **Testing**:
  - on the same ME, performed on an anomalous run (*BAD*).
  - **deviations** from the learned behavior produce peaks in the reconstruction error.
- Various types of AEs, such as the *LSTM AutoEncoder* [2], can be employed.
- A ***Dense Under-complete AutoEncoder*** (*Fig. 1*) is used.

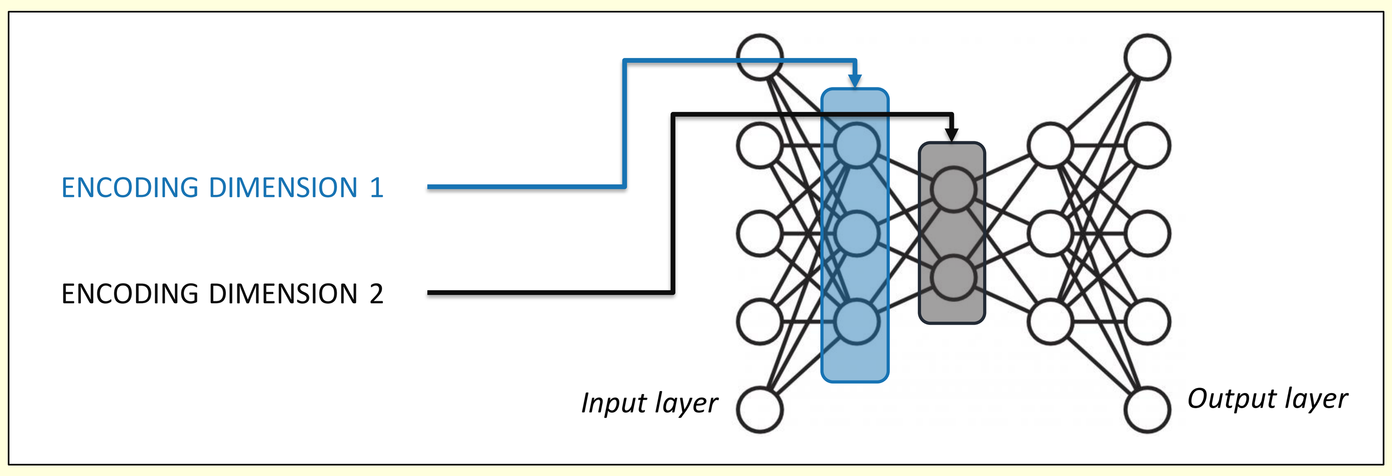


Fig.1: The Dense AutoEncoder architecture utilized is characterized by the dimensions of two layers only.

- In the following we test two *BAD* runs after training the model on a *GOOD* run using the ME *MET Significance.*

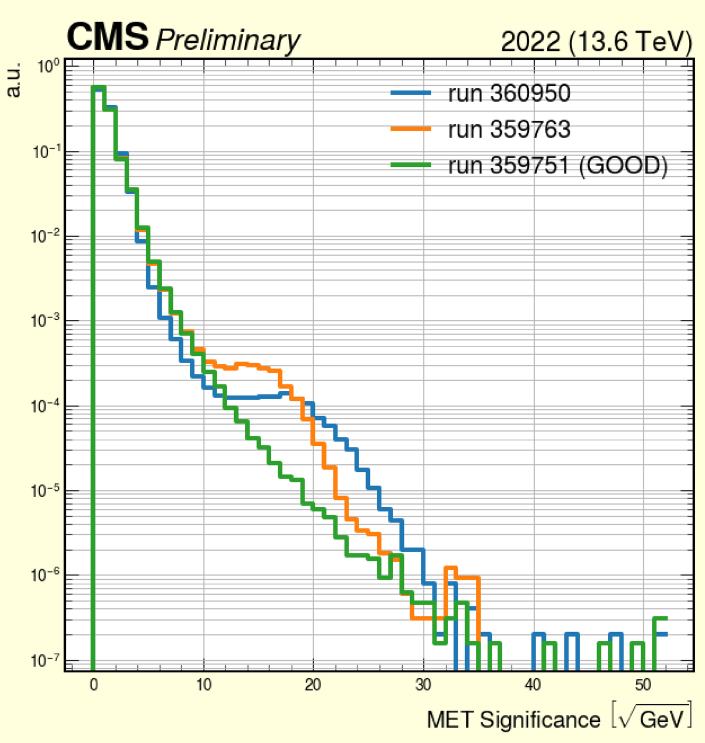


Fig.2: The histograms of MET Significance for the three runs. The anomaly is visible in the histogram center for the blue and orange runs.

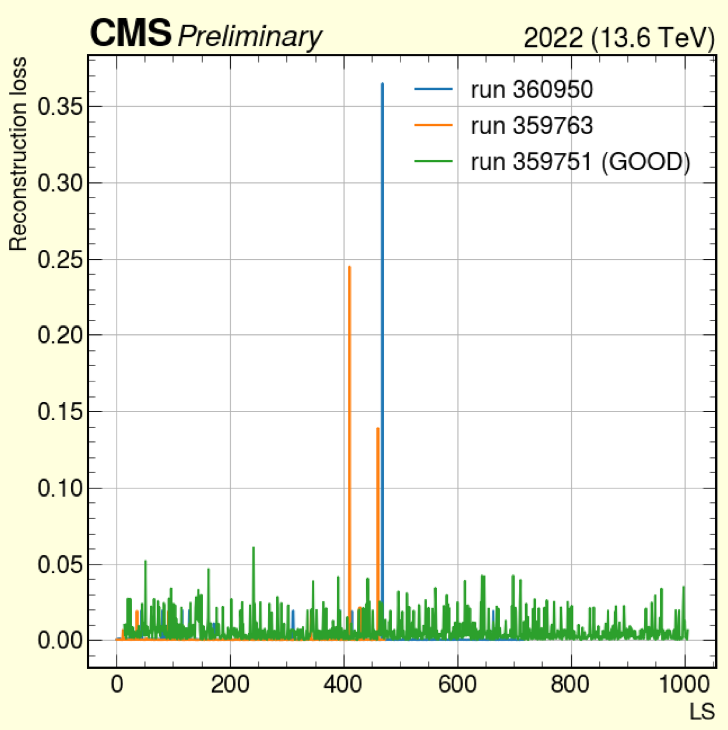


Fig.3: The reconstruction error of the AE as a function of the LS number for the three runs. Some peaks are visible for the two anomalous runs.

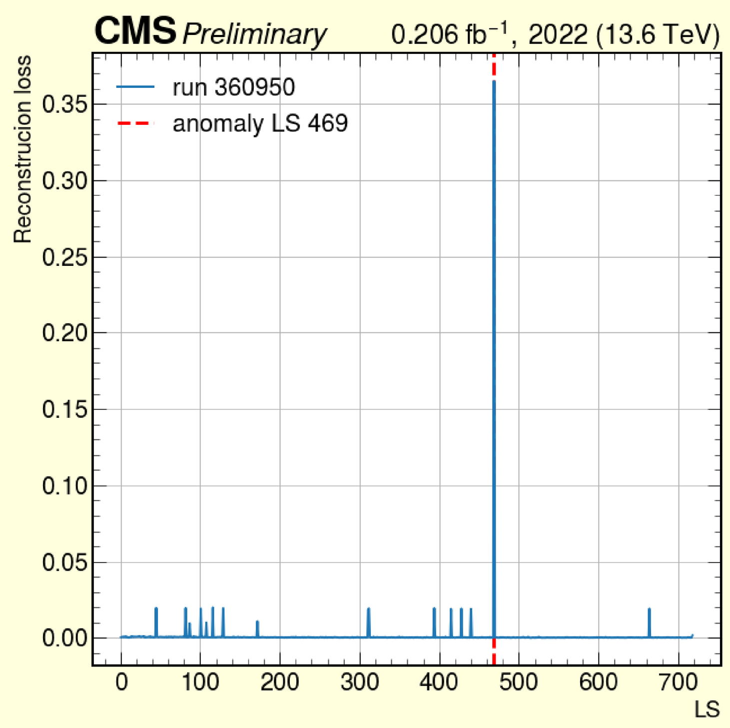


Fig.4: The reconstruction error of the AE as a function of the LS number for run 360950.

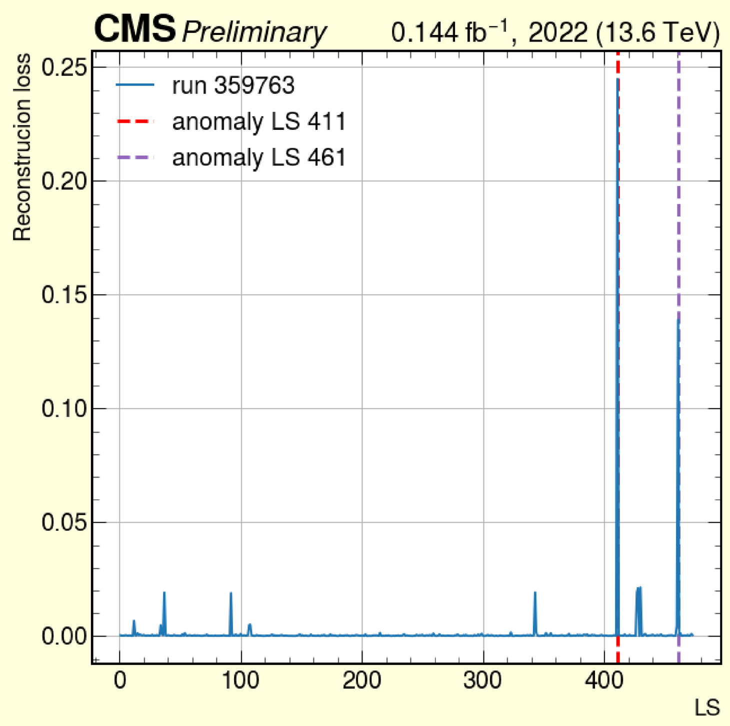


Fig.5: The reconstruction error of the AE as a function of the LS number for run 359763.

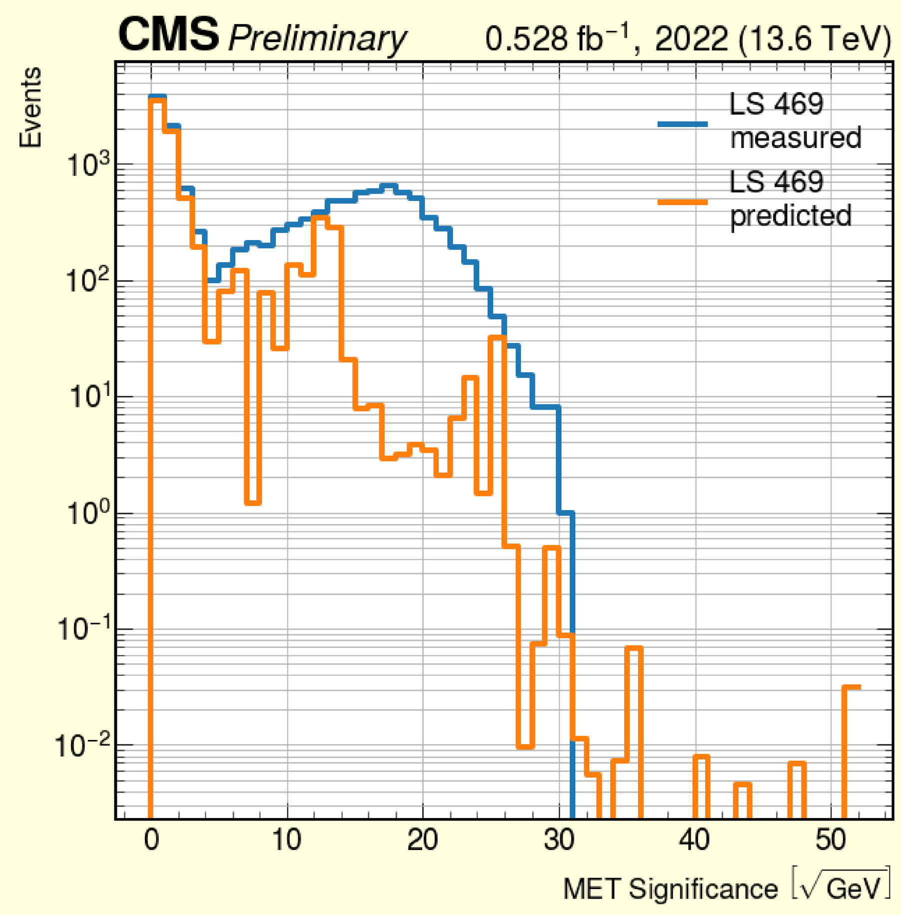


Fig.6: The input histogram (blue) and the output (orange) of the AE for the LS corresponding to peak of the reconstruction loss for run 360950.

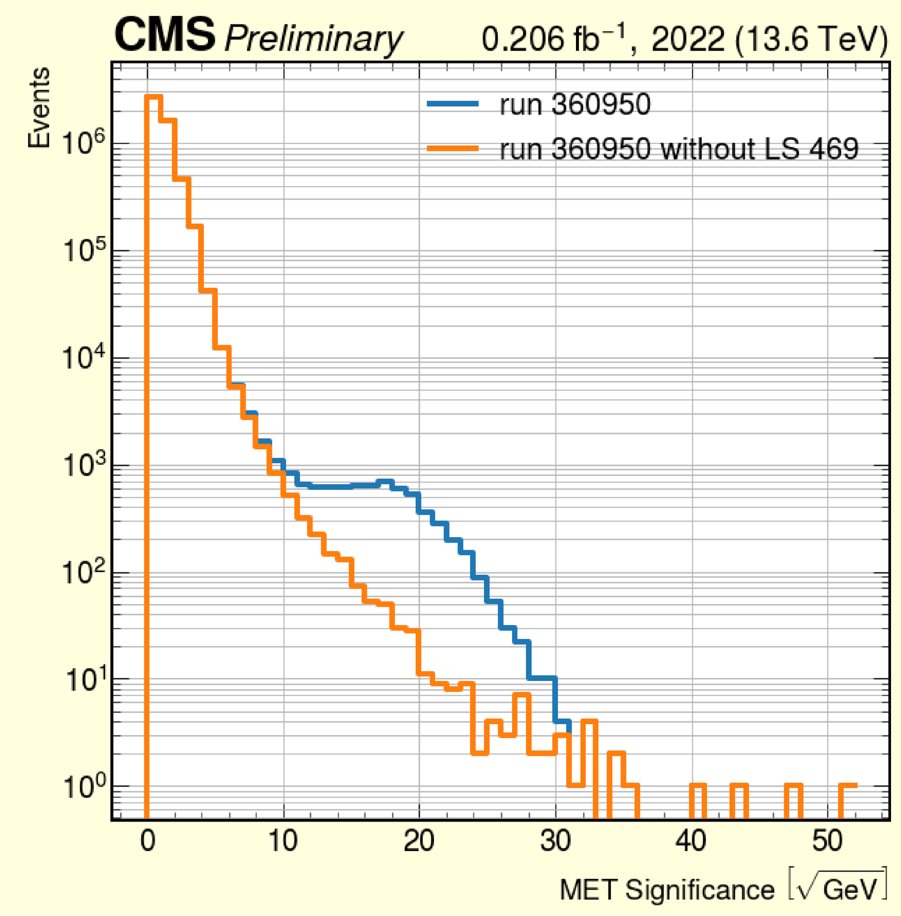


Fig.7: The elimination of LS 469 *for run* 360950 *completely removes the anomaly from the run.*

Conclusions

- We developed an **AutoEncoder-based Anomaly Detection Tool** capable of detecting anomalies in DQM MEs with a per-LS granularity [3].
- We tested the tool on several runs flagged *BAD* by JME DQM and identified the source of the anomalous behavior in a limited set of LSs.
- In particular, in the examples presented here, we removed one LS from each anomalous run and verified that the remainder was no longer anomalous.
- The equivalent luminosity recovered from the two runs is $350\ pb^{-1}$, or around 1% of the entire 2016 CMS dataset.
- Exploiting the per-LS granularity in DQM and systematically employing the tool we presented will enable an **increase in the efficiency of the DC procedure**, ultimately resulting in a larger dataset available for physics analyses.

References

[1] Ian Goodfellow, et al. Deep Learning. MIT Press, 2016.

[2] Yuanyuan Wei, et al. "LSTM-Autoencoder based Anomaly Detection for Indoor Air Quality Time Series Data." (2022).

[3] **CMS DP note**: DP-2023-010, https://cds.cern.ch/record/2854697?ln=pt

CMS DP note