Detector monitoring with artificial neural networks at the CMS experiment at the CERN Large Hadron Collider

Adrian Alan Pol[1,2] Gianluca Cerminara[2] Cecile Germain[1] Maurizio Pierini[2]

[1]Université Paris-Saclay, Orsay, France [2]CERN, Meyrin, Switzerland

The CMS Data Quality Monitoring (DQM) system

- **Guarantees high-quality data** for physics analyses:
  - *online monitoring*: **live** feedback during data acquisition;
  - *offline monitoring*: certify the data quality using offline processing.
- Online DQM assesses detector behavior and **identifies emerging problems**:
  - comparison to reference distributions;
  - comparison supported by **predefined tests**;
  - tests designed to identify **known failure modes**.
- Challenges of online w.r.t. offline, relevant to machine learning emerge:
  - the **latency** of the evaluation process;
  - absolute **normalization** of the histograms not possible;
  - higher **granularity** of the problems to spot;
  - no availability of the ground truth (**labels**).

Why machine learning for Online DQM?

Two-layer monitoring protocol was adopted by the CMS Collaboration for LHC Run I (2010-2012) and in Run II (2015-2018). The challenges include:

- **latency**: human intervention and thresholds require sufficient statistics;
- **volume budget**: amount of data a human can process in a finite time;
- **scalability of static thresholds**: assumptions on potential failure scenarios;
- **human driven decision process**: alarms based on shifter judgment;
- **changing running conditions**: reference samples change over time;
- **manpower**: the effort to train a shifter and maintain instructions.

Details on the infrastructure used for this Data Quality Monitoring (DQM) are given in [1].

Image-like processing of geographically organized data

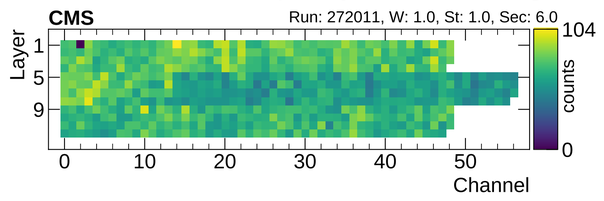
Test case: Drift Tube (DT) muon chambers hit occupancy

Data recorded by the DT chambers of the muon spectrometer during LHC Run II:

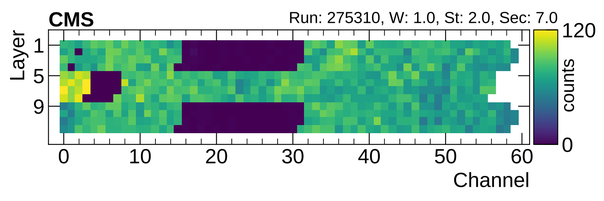
- hit occupancy C^k contains the total **number of electronic hits at each readout channel**. It is a 2-dimensional array organized along **layer** (row) and channel (column) indices

$$C^k = \{x^k_{i,j}; 1 \leq i \leq l, 0 \leq j < n_i\},$$

  $l = 12$ is the number of layers, n_i is the number of channels in layer i
- **expected**: small variance of hit occupancy between neighboring channels; example

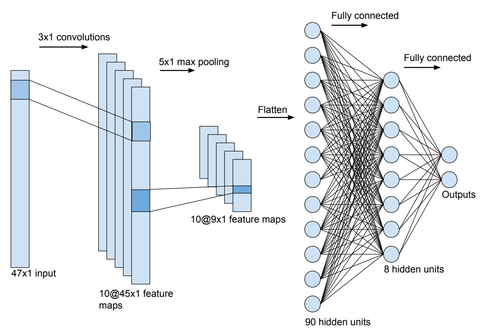


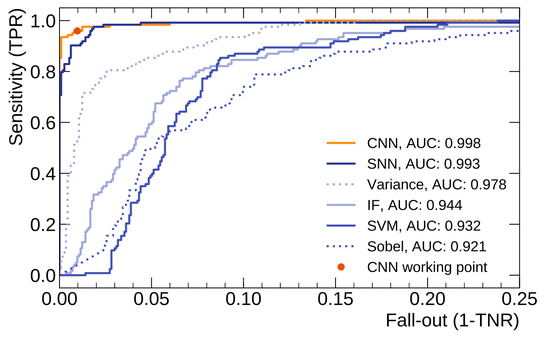
- **anomalous**: noisy or inefficient area; example, low occupancy across the 12 layers



Local strategy

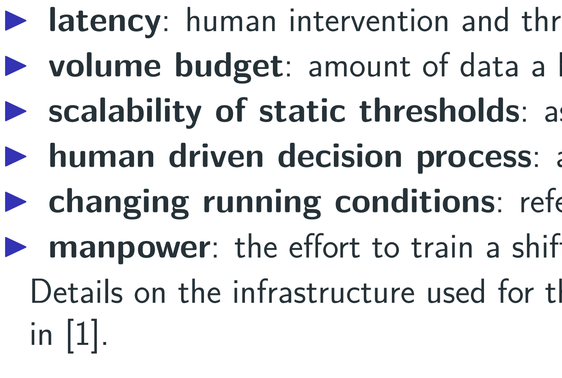
- assessing the (mis)behavior with high-granularity (**few channels**)
- data collected in each **layer** are treated **independently** from the others to detect intra-layer problems
- convolutional neural network (CNN) [2] outperforms shallow network (SNN), one-class support vector machine (μ-SVM) [3], Isolation Forest (IF) [4, 5], variance and a variation of Sobel filter [6]

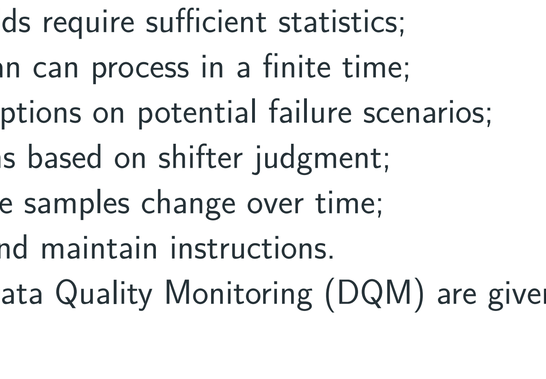




Regional strategy

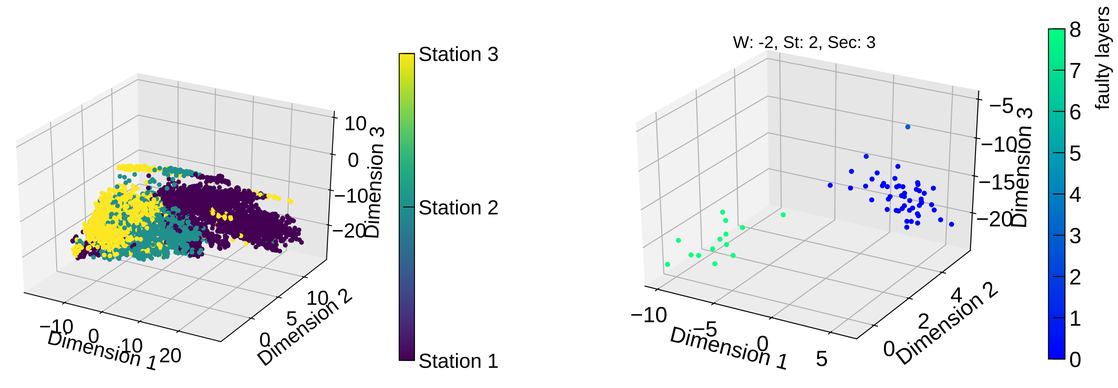
- extend local strategy to account for **intra-chamber** problems: simultaneously consider all layers in a chamber.
- semi-supervised autoencoder [7] variations trained with well-behaving chambers use **mean squared error** of input x and reconstructed $\ddot{x}$ samples: $\epsilon = \frac{1}{ij}\sum_{i,j}(x^k_{i,j} - \ddot{x}^k_{i,j})^2$
- the test is performed for intra-chamber anomalies, an example: identify layers with low efficiency (lower voltage)





Global strategy

- simultaneous use of **all the chambers** data
- the **position** impacts expected occupancy pattern
- with autoencoders, a **compressed representation** of chamber data is learned
- with **3-dimensional** bottleneck one can visually inspect those representation
- the representations **cluster** depending on their position in the detector (left: distance from the interaction point)
- the same chamber **changes** representation when problem occurs (right)



Outlook

- the **first** successfully implemented ML application in CMS DQM;
- the local approach has satisfactory performance and was successfully **implemented** in production;
- the proposed strategy is **generic** enough to be applicable to other kinds of CMS muon chambers, as well as to other sub-detectors;
- the model should be refined, e.g. integrating a mechanism of periodic retraining;
- ongoing other sub-detector efforts to apply similar strategy.

References

[1] F. De Guio, "The data quality monitoring challenge at cms: experience from first collisions and future plans," tech. rep., 2015.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[3] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[4] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pp. 413–422, IEEE, 2008.

[5] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 1, p. 3, 2012.

[6] I. Sobel, "An isotropic 3×3 image gradient operator," *Machine vision for three-dimensional scenes*, pp. 376–379, 1990.

[7] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.