



AI Enabled Data Quality Monitoring with Hydra

Thomas Britton

David Lawrence

Kishansingh Rajput

The Challenge

- **Every run** produces an initial **22 plots**. More thorough monitoring is performed offline and produces **109 plots**. With a run lasting **~3 hours** every day there are **between ~175 and 875** plots to look at.
 - To preserve sanity I looked at closer to 175 plots, but there is no reason a machine couldn't aid in looking at all of them...
- Often times a single plot being “off” is not an indication of problems. Need to look at all the plots to determine cause and severity
 - Trigger studies: Often look like big problems but are not. Can be hard to catch when shift logs have scant details

Introducing Hydra

- Hydra aims to be an extensible framework for training and managing A.I. for near real time monitoring
 - If you need it to tell a dog from cat I can have hydra do that, without system modification, **now**
- Most importantly, Hydra allows me to embrace my inner sloth:

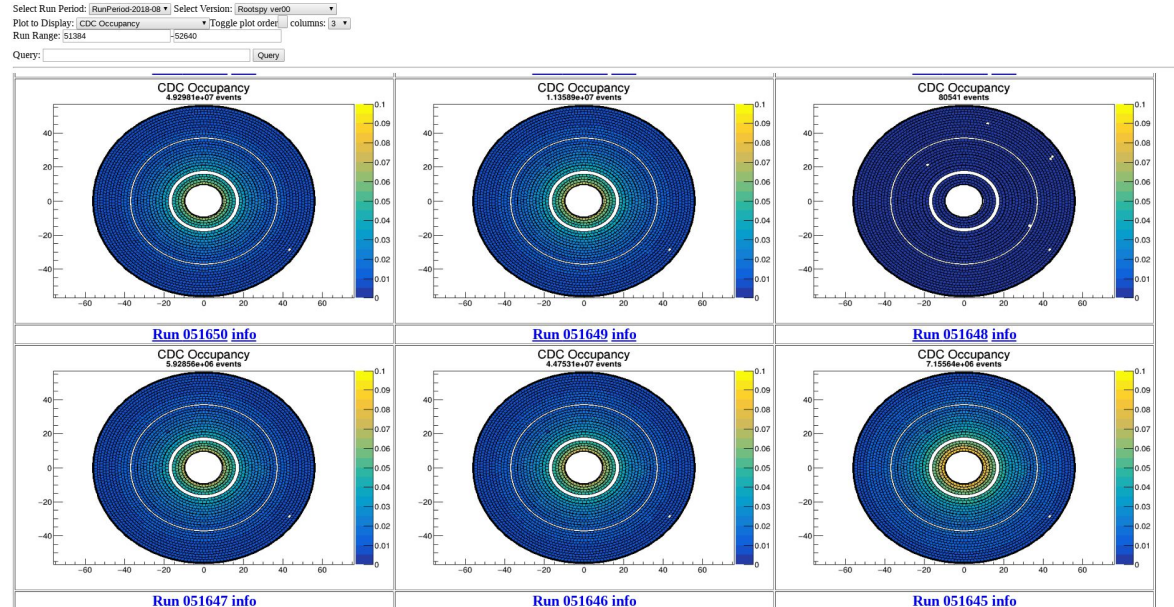


Koboldpress.com

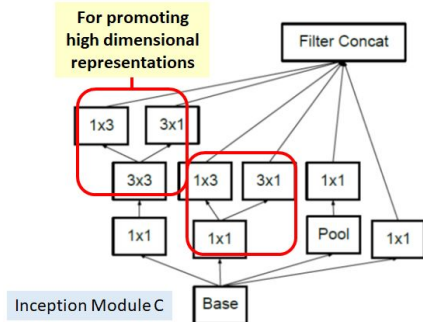
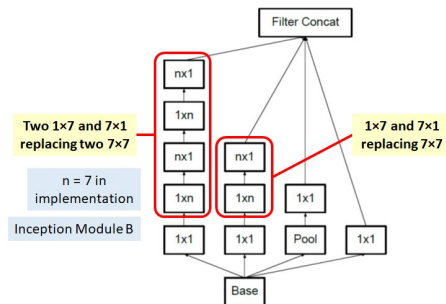
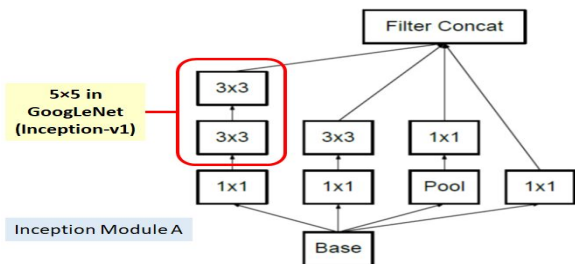
Data

- Use images instead of histograms because:

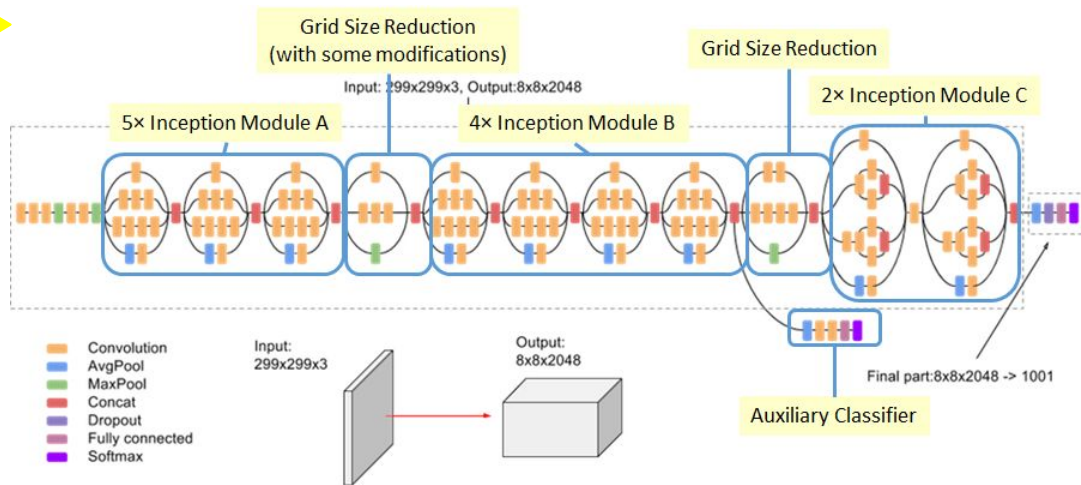
1. All images **already existed** on disk
2. It's how humans do it
 - a. **Easier to interpret**
3. Can leverage **computer vision**



The Inception v3 Network



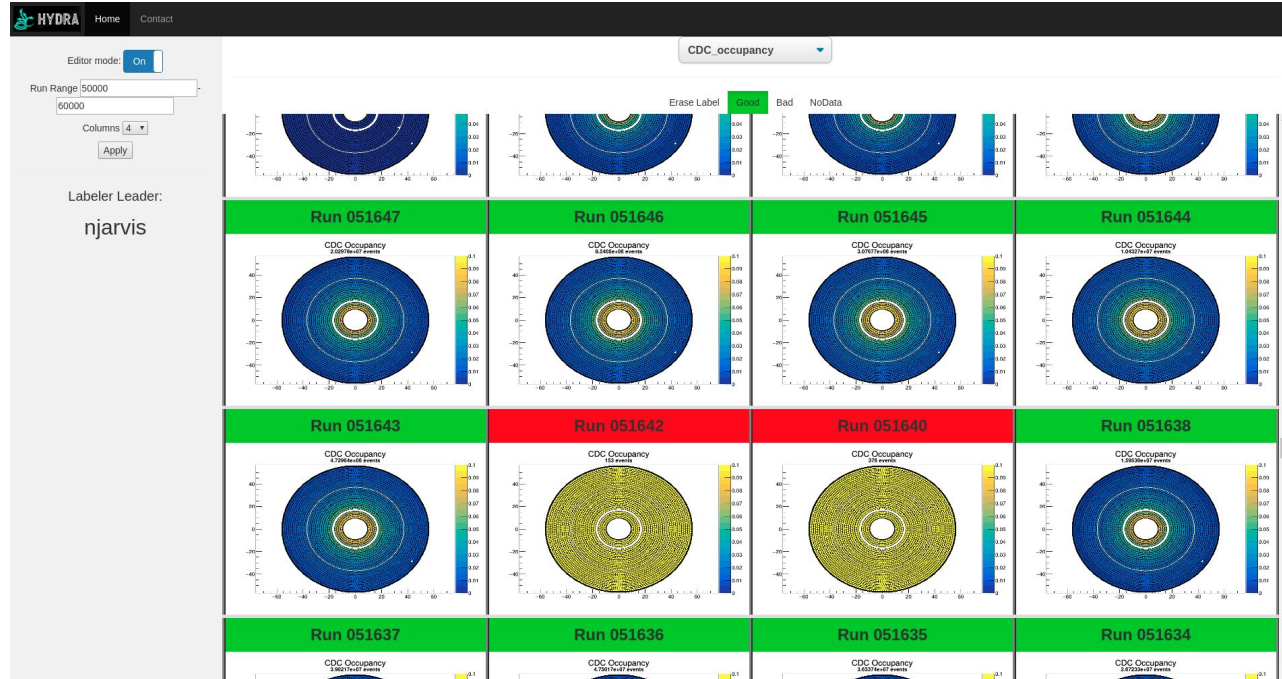
Already in
Keras and
Tensorflow



Supervised learning! aka Labeling all that Data

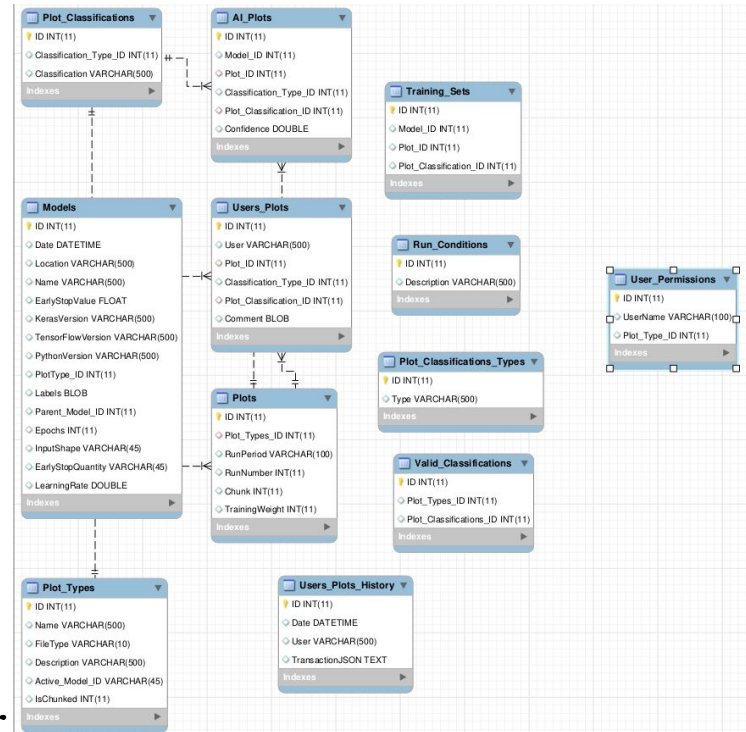
- Webpage:

~A few hours to label **all** of the plots of a given type



The Backend

- Supported by a database
 - **All** plots
 - **All** user defined labels
 - **All** models
 - **All** models' classifications with confidence*
 - *Only saved plots
- Training is virtually push button to allow for automated retraining as needed



Hydra Fast Facts

- Hydra looks at a finer time scale than any higher level monitoring the shift crew performs. **Approximately every minute**
 - Because who hits reset?
- **Operates** (conservatively) at about **3-4Hz**
 - From receiving an image to action ~**300ms**. Most of the time spent on model inference
 - Inference accounts for ~**71%** of the total processing time and is driven primarily by model size
- Currently focused on **go/no-go decisions**
 - Doctor classifying you as sick with no diagnosis as to what you are sick with. Refinement underway

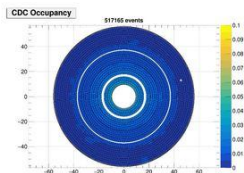


Koboldpress.com

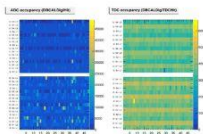
HydraRun also saw the FDC problem, which I probably would have missed inspecting it by eye.

Dashboard

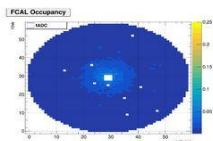
72437



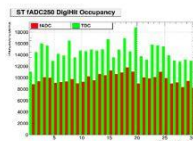
chunk 0092: Good @ 0.96700119972229



chunk 0293: Good @ 0.9977362155914307



chunk 0291: Good @ 0.9995253086090088



- **Real-time dashboard** viewable from anywhere shows the last plot analyzed as well as the class and confidence (overly precise)
- Able to detect hot channels in some detectors for later calibration. Can detect these problems early indicating hardware that may soon need replacing
- Gives a go/no-go indicator for the plots. When something seems off shift crews can see the problems and **focus on the plots that matter**
 - Notify experts, sound an alarm, take corrective action
- **Trailing 24hr window** of all potentially problematic plots seen.
 - This can give experts access to problems that may be impossible to spot in summation (intermittent short term issues)

Hydra Future Development

- **Custom, optimized models and Diagnosis**
 - Layer-wise Relevance Propagation (**LRP**) and Deep Taylor Decomposition (**DTD**)
 - Incorporating LRP/DTD to map pixels most relevant in a particular decision by the models.
 - For a bad image, those relevant pixels can then be mapped to faulty physical board(s).

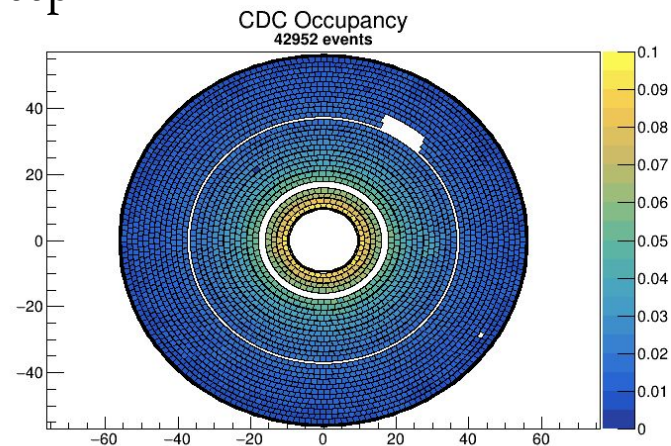


Fig: Diagnosing CDC plot (Bad) with LRP

Hydra Future Development

- **Custom, optimized models and Diagnosis**
 - Layer-wise Relevance Propagation (**LRP**) and Deep Taylor Decomposition (**DTD**)
 - Incorporating LRP/DTD to map pixels most relevant in a particular decision by the models.
 - For a bad image, those relevant pixels can then be mapped to faulty physical board(s).

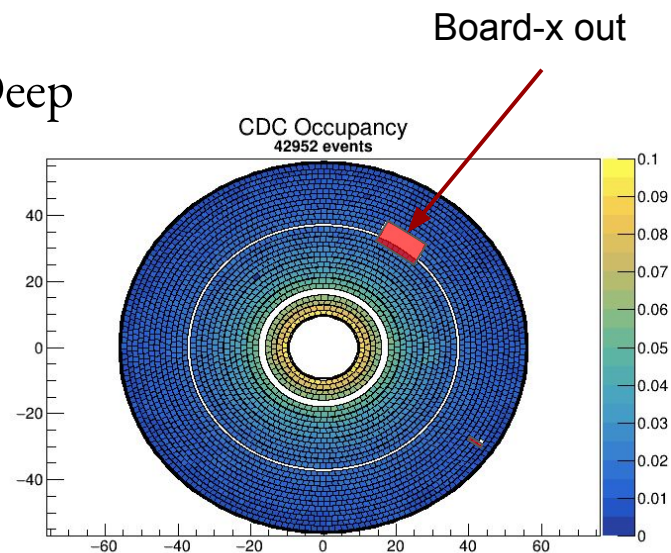


Fig: Diagnosing CDC plot (Bad) with LRP

Hydra Future Development

- **Classification splitting**
 - From a doctor saying “you are sick” to actually diagnosing a condition
- **Custom, optimized models**
 - Inference accounts for ~71% of the total processing time and is driven primarily by model size
- Ability to actually **take corrective action** as needed
 - Will require trust and more data on in situ running
- **More plot types!!**
 - Data types too

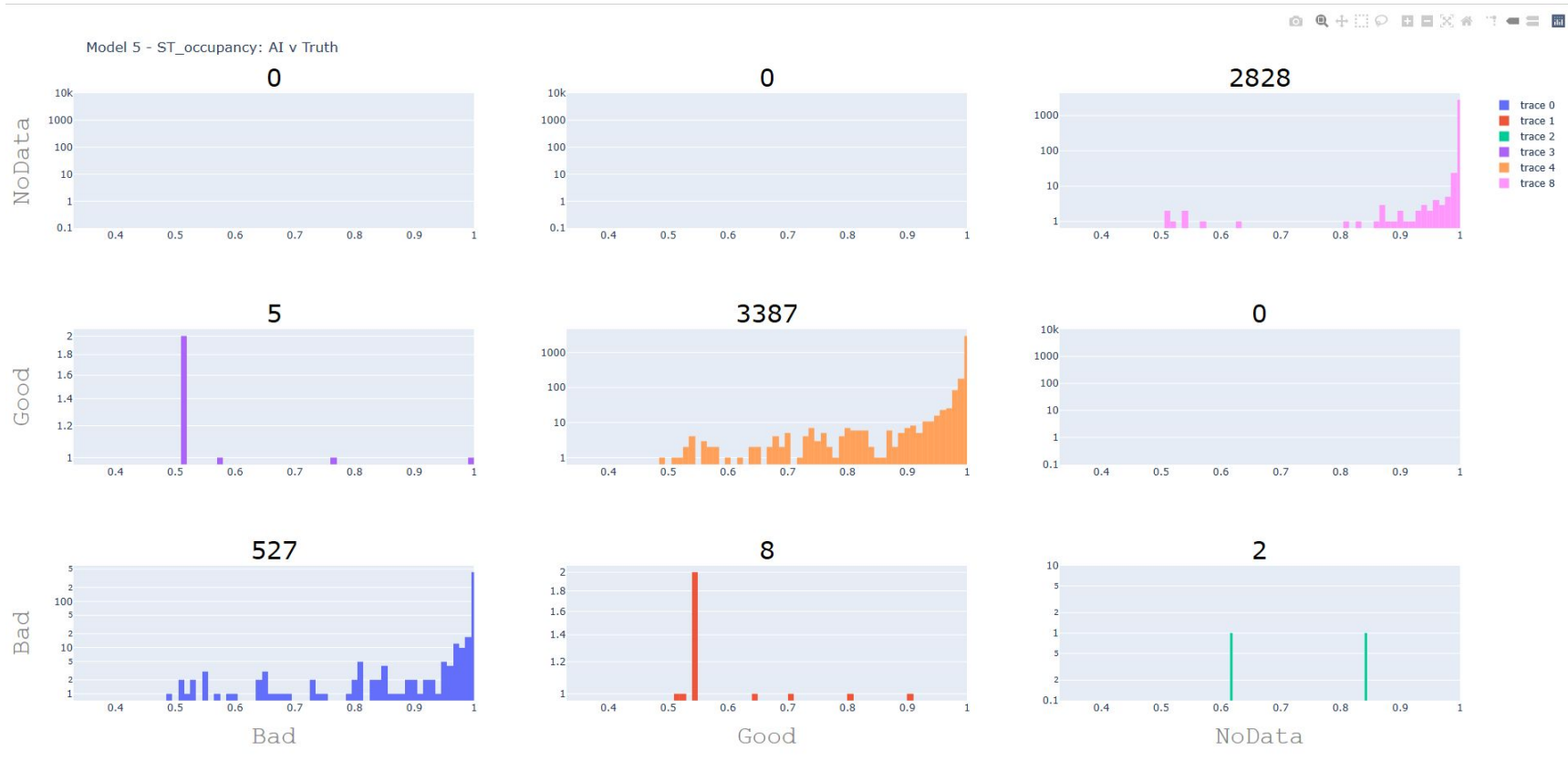


Koboldpress.com

Backup Slides



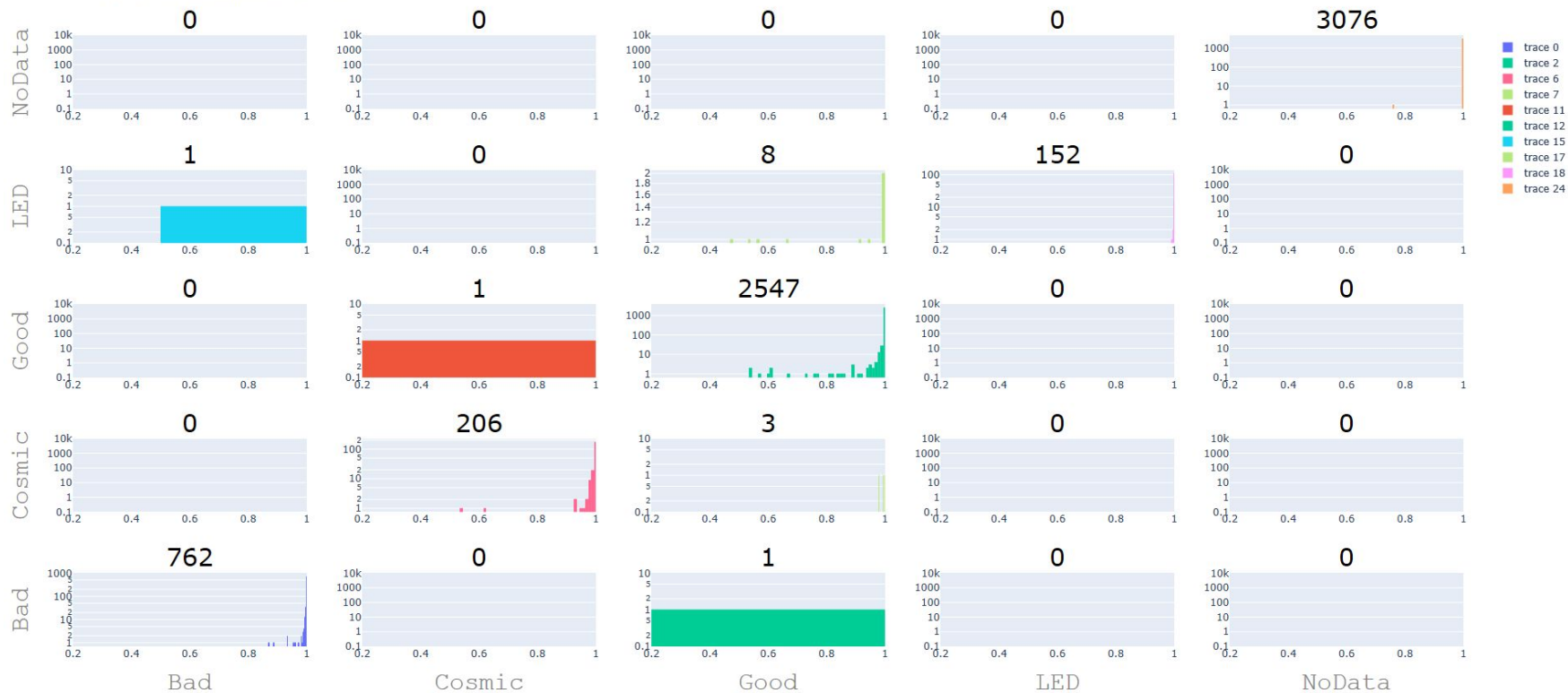
Preliminary Results (Start Counter)



Preliminary Results (FCAL)



Model 6 - FCAL_occupancy: AI v Truth



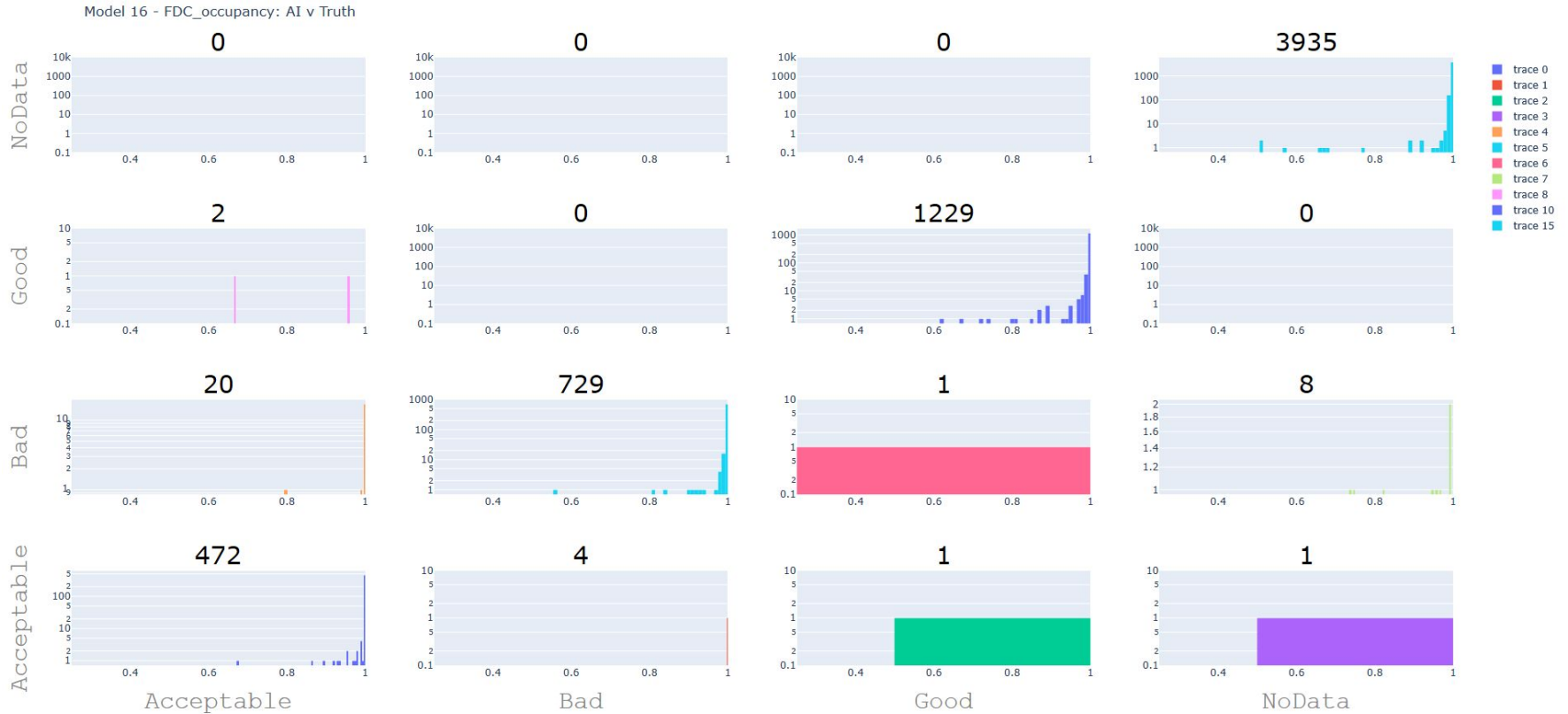
Preliminary Results (DIRC)



Preliminary Results (BCAL)



Preliminary Results (FDC)



Preliminary Results (TOF)

Model 20 - TOF_occupancy: AI v Truth

