



# Event Selection and Categorization



# Permanent Decisions

- CMS takes data at 40MHz
- Hardware triggers identify 100k of potential interesting events
  - Transverse energy, electrons or muons, jets
- In 100ms the higher level trigger further reduces this to 1k events to collect
  - Most information about the 99% we don't select is lost



# Event classification

- Trained systems for event classification are now being used in a variety of fields
  - **Classifying stars in astrophysics**
    - 40 row confusion matrix where correct assignment rates are 85% and higher
    - New star categories can be discovered by allowing unclassified things to be marked as unknown
  - **Star catalogs can now be automatically produced**



# Putting Some Ideas Together

- Is there information we could get out of the 99k of events we currently throw out?
  - Cross sections, calibration samples, cross checks on trigger efficiency?
- We have a model that explains most of what we see?
  - The things we can't explain in the model are the most interesting, but hard to design triggers for
- We have simulation to train with



# Could this change how we trigger?

- Instead of reconstructing enough of the event to reject it, could we use event classification to identify the events?
  - How much reconstruction is needed for reasonable classification rates
  - Would unclassified events be the most interesting?
  - Could we keep statistics even on things that we will eventually not keep?



# Things to try

- We collect ~15-20 collisions every times the beams cross
  - Only one of those is a triggered event, the rest are events without a trigger bias
  - It would be interesting to try the technique of event classification on what is essentially a large zero bias sample
  - Could potentially experiment with some light quark physics



# Outlook

- This is a project that would be interesting to investigate with experts
  - Machine learning specialists
  - Industry partners
- I would like to find a way to show or eliminate feasibility with the smallest amount of work
  - Even if feasible try to identify how the resource balance would change