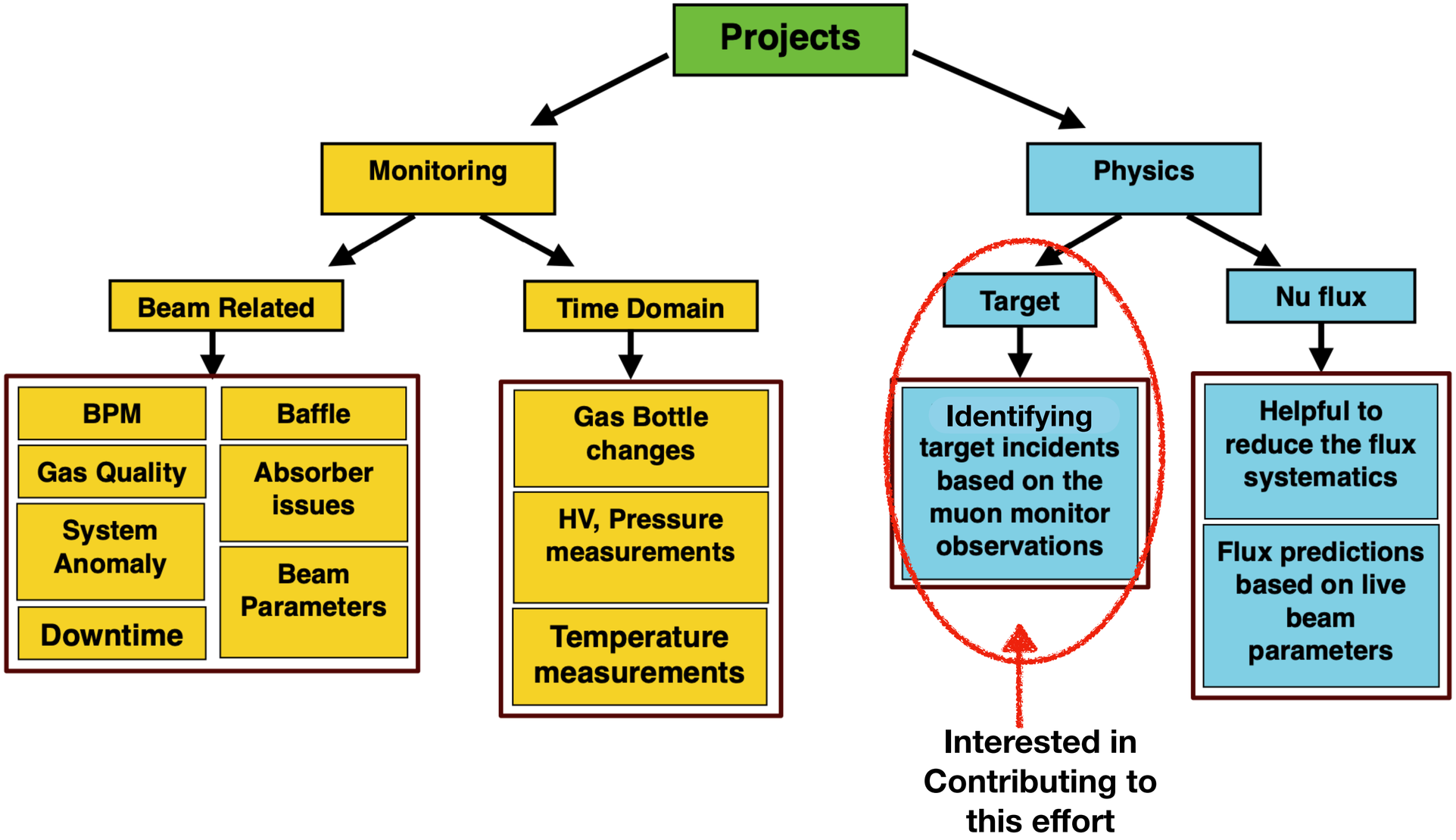
AI/ML Project Plans for NuMI beamline Data

Sudeshna Ganguly

01/15/2021

Project Plan Proposed By Athula



Identify Target Incidents from MM Observations

Impact on neutrino flux

- Important properties of target to monitor : target density, target position etc
- Need to know how stable target is

Identify Target Incidents

1. Geometry changes(e.g. stress wave moving through target)
2. Radiation damage effects (e.g. embrittlement, swelling, corrosion etc.)
3. Changing density
4. Fractured/missing fin
5. Target position shift

How to Monitor These Target Incidents

- Typically distortion in neutrino beam spectrum tells us about changes e.g. target degradation but not at real time
- Muon Monitors could be able to see these
- Could be used to build a real-time monitor

How to Monitor These Target Incidents

- **Muon Monitors could be able to see these**

Athula has shown: possible to catch incident by looking at MM (Muon Monitor)#COR data
Neural Network with 4 inputs: Beam Intensity, Horn Current, Beam Position X and Y
2 Output: Normalized Integrated signals (MM1,2)

Incident identifying variable

$$F_{Incident} = \left[\frac{MM1COR}{MM2COR}\right]_{Pred} \cdot \left[\frac{MM2COR}{MM1COR}\right]_{Data}$$

$$\text{Noramlized Integrated Signal} = \frac{MM\#COR}{POT}$$

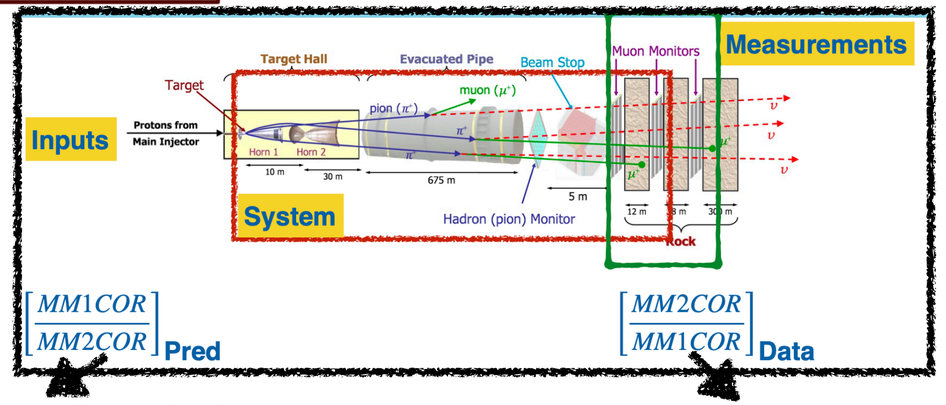
$\propto$ total muon flux passing through detector per proton

$\propto$ total muon energy lost in detector

$MM\#COR$ = Muon Monitor integarted signal with pedetal subtractioins and calibrations

$MM\#$ = Muon Monitor (1,2 or 3)

POT = Beam Intensity



Depends on the inputs

Depends on the inputs & run time

- **If $F_{incident}$ is independent of inputs, it depends on system status**
- **$F_{incident}$ could be sensitive to:**
  - **Target related incidents**
  - **Horn system related incidents**
  - **Decay pipe, absorber related incidents**
  - **Muon monitor detector failures**

From Athula
Athula's TSD topical meeting talk

How to Monitor These Target Incidents

- **Muon Monitors could be able to see these**

Athula has shown: possible to catch incident by looking at MM (Muon Monitor)#COR data
Neural Network with 4 inputs: Beam Intensity, Horn Current, Beam Position X and Y
2 Output: Normalized Integrated signals (MM1,2)

Incident identifying variable

$$F_{Incident} = \left[\frac{MM1COR}{MM2COR}\right]_{Pred} \cdot \left[\frac{MM2COR}{MM1COR}\right]_{Data}$$

$$\text{Noramlized Integrated Signal} = \frac{MM\#COR}{POT}$$

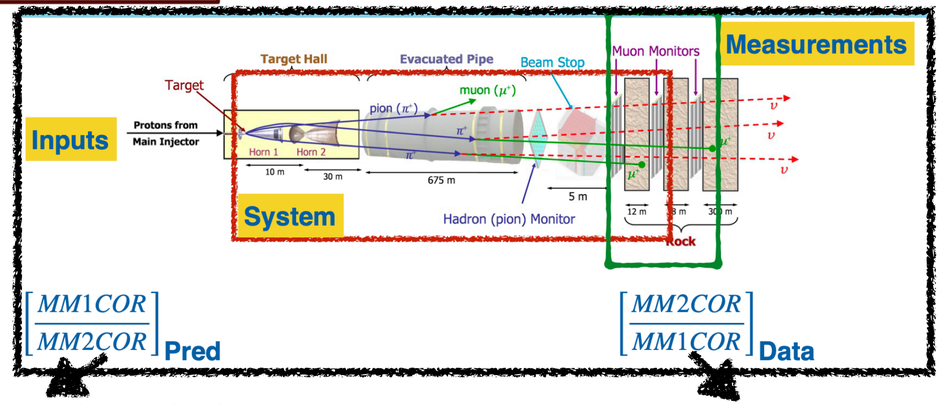
$\propto$ total muon flux passing through detector per proton

$\propto$ total muon energy lost in detector

$MM\#COR$ = Muon Monitor integarted signal with pedetal subtractioins and calibrations

$MM\#$ = Muon Monitor (1,2 or 3)

POT = Beam Intensity



Depends on the inputs

Depends on the inputs & run time

- **We can look at MM#COR**

But

- **Muon Monitor observations will depend on many other parameters**
- **Need to look at Muon Monitor pixel information and disentangle other correlations**

How to Monitor These Target Incidents

- **Muon Monitor observations will depend on many other parameters**
- **Need to look at Muon Monitor pixel information and disentangle other correlations**

- **Use ML to filter out correlations associated to target incidents**
- Split **data** into training and test sets
- Apply multivariate filter methods (removes redundant features from data)
- **For training ML model: need data!** input: 3 MM data can be used
- **Not much target data available**
- **Need MC samples with large statistics** to understand correlations b/w MM observation and target incidents
- Normalize simulated data to mimic real data (No. Of muons/ pixel -> Voltage measurement)

Important to generate large simulation sample (>300 M events each)

- **Geant4 based simulation**
- **Usually takes ~month to produce 1 simulation sample**
- **Need to scan over parameter space ~ many large samples**
- **GPU processing could be solution**
- **Geant4 does not support GPU processing**
- **Possible to write a standalone module that uses GPU processing and plug it into Geant4**
- **Easier option would be CPU processing on NERSC(16000 cores) or ALCF**

Notes from Adam Lyon

- Athula and I had a quick meeting with Adam Lyon (Associate Division Head of SCD) on how to generate large simulation sample fast on NERSC
- Based on his suggestions we could do the following:

Get answers to our simulation-specific questions

↓

Talk to Nahn Tran/Brian Nord (who are already working on using AI for accelerator physics) to figure out if they already have computing account for accelerator-related project on NERSC/something similar

↓

If “NO” is the answer to above question, apply for an account on NERSC (Nahn Tran should be the contact person)

↓

If we get computing time on NERSC, Lisa Goodenough can help us run Geant4 simulation on NERSC

Notes from Adam Lyon

- Athula and I had a quick meeting with Adam Lyon (Associate Division Head of SCD) on how to generate large simulation sample fast on NERSC
- Based on his suggestions we could do the following:

Get answers to our simulation-specific questions

1. What's the minimum size of each simulation data set (out of many configurations we would like to create)? How many million events?
2. How many different configurations do we need? What's the minimum number? 1000 or more?
3. What CPU time on Fermigrid do we take right now to run 1M events?
4. In other words, how many slots on Fermigrid do we use?
5. How much memory do we need to store output data? From all configurations?

- We want to store information such as: muon's momentum, position, particle IDs, absorber, horn information, neutrino information, tracking information, hadron production etc.
- It's a balance b/w storing large output (with all ntuples) files vs storing reduced information (histograms)
- Can we extract output files into an external hard disk? (question for the department)

Identify Target Incidents from MM Observations

Moving forward:

- Create large simulation samples for different target status to find out correlations b/w target incidents and Muon Monitor observations
- Validate simulation to compare with data
- Identify neural network inputs: Muon Monitor pixels, proton beam positions (possibly)
- Output:
  - first classify incidents
  - then define a new variable indicating target-related issues
- Perform training on simulated data