Using Machine Learning to Improve our Understanding of the Jet Background in A+A Collisions

Tanner Mengel

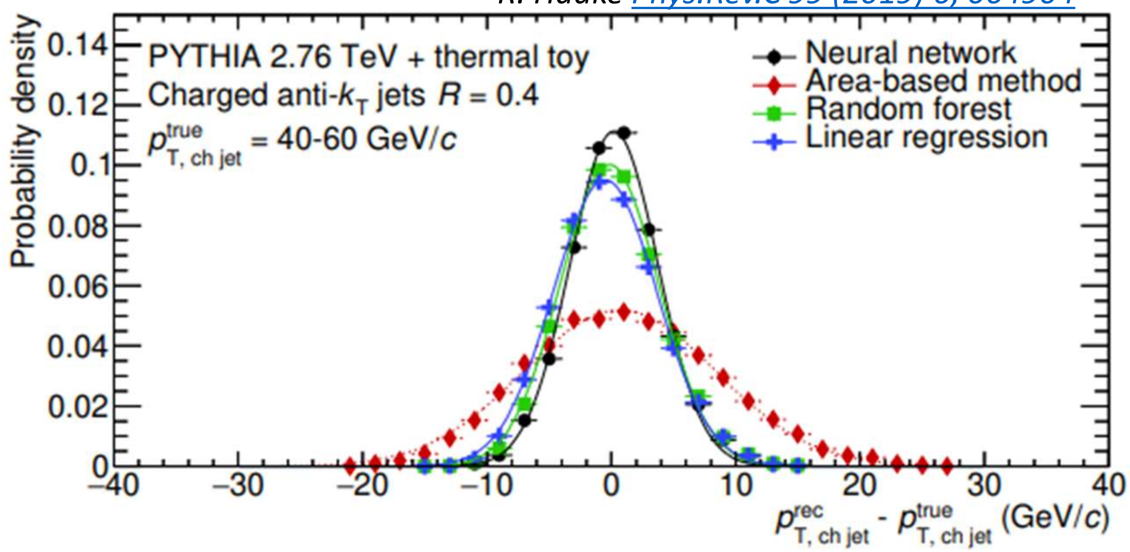
University of Tennessee

Special Thanks to Charles Hughes, Antonio Silva and Patrick Steffanic!

Motivations

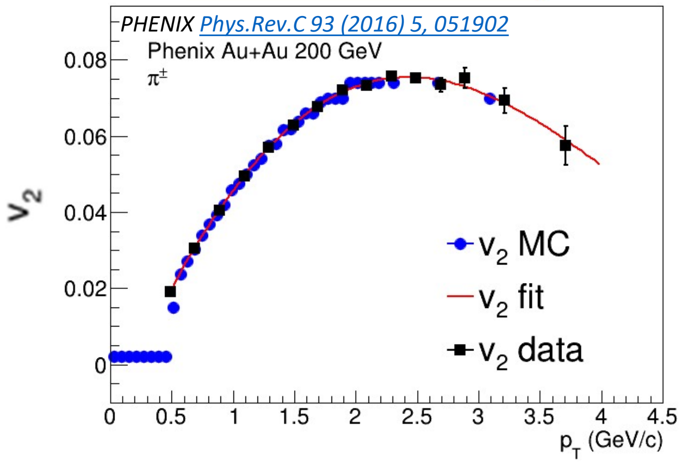
- Using a deep neural network for jet background subtraction increases momentum resolution for LHC energies
  - *Develop similar method for RHIC energies!*
  - *What is the 'Special Sauce' of the DNN?*

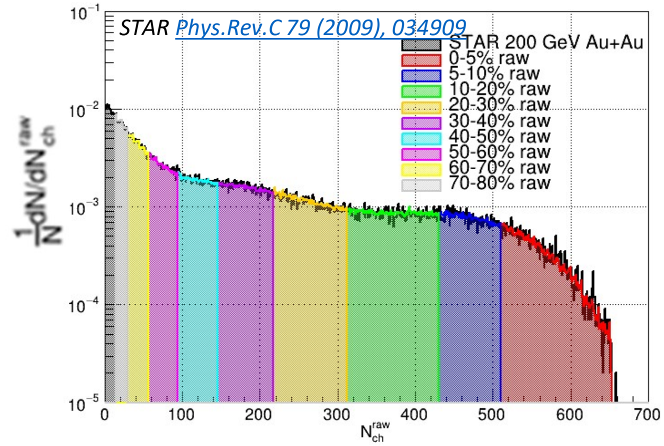
R. Haake Phys.Rev.C 99 (2019) 6, 064904



Event Simulation

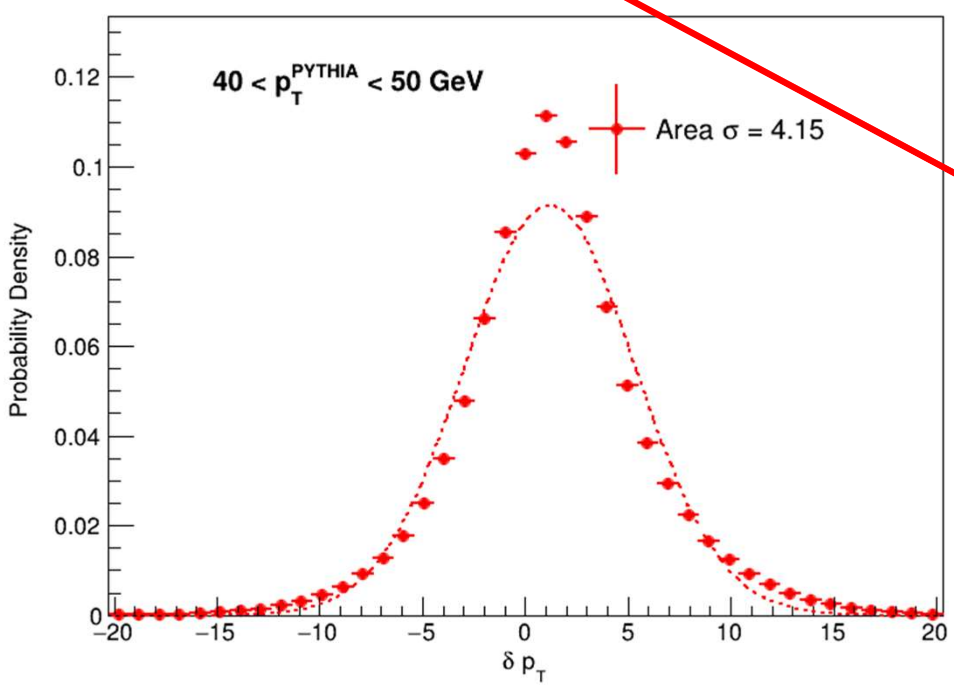
- PYTHIA8 (Signal):
  - 25 Million (1 million per p_T hard bin) p+p events at 200 GeV, Tune 14
- TennGen (Background): *C. Hughes et al Phys. Rev. C 106 (2022), 044915*
  - **Multiplicity**: Sampled from corrected N_{ch} distribution *STAR Phys.Rev.C 79 (2009), 034909*
  - $\boldsymbol{p_T}$: Identified particle p_T spectrum fit with Boltzmann-Gibbs Blast wave *PHENIX Phys.Rev.C 88 (2013) 2, 024906*
  - $\boldsymbol{\phi}$: Identified particle flow harmonics (v_2, v_3, v_4) *PHENIX Phys.Rev.C 93 (2016) 5, 051902*
  - $\boldsymbol{\eta}$: Uniform distribution |η| < 1.1
- Merge PYTHIA8 charged particles with TennGen Au+Au 200 GeV background
- Find anti-k_T jets
  - Only save jets with p_T^{PYTHIA} > 5.0 GeV
  - ~30 Million jets per dataset
- Take p_T^{PYTHIA} to be truth value
  - Train-Test split: 20/80%



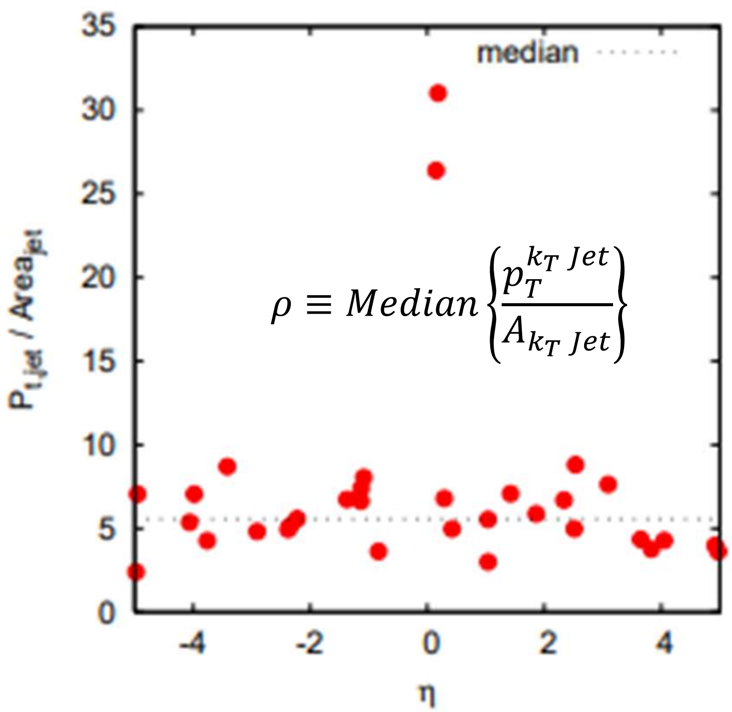


Jet Background Subtraction

- Area Based Background Subtraction:
  - $p_T^{Corr.} = p_T^{Uncorr.} - \rho A_{jet}$



G. Soyez https://arxiv.org/pdf/0905.2851.pdf

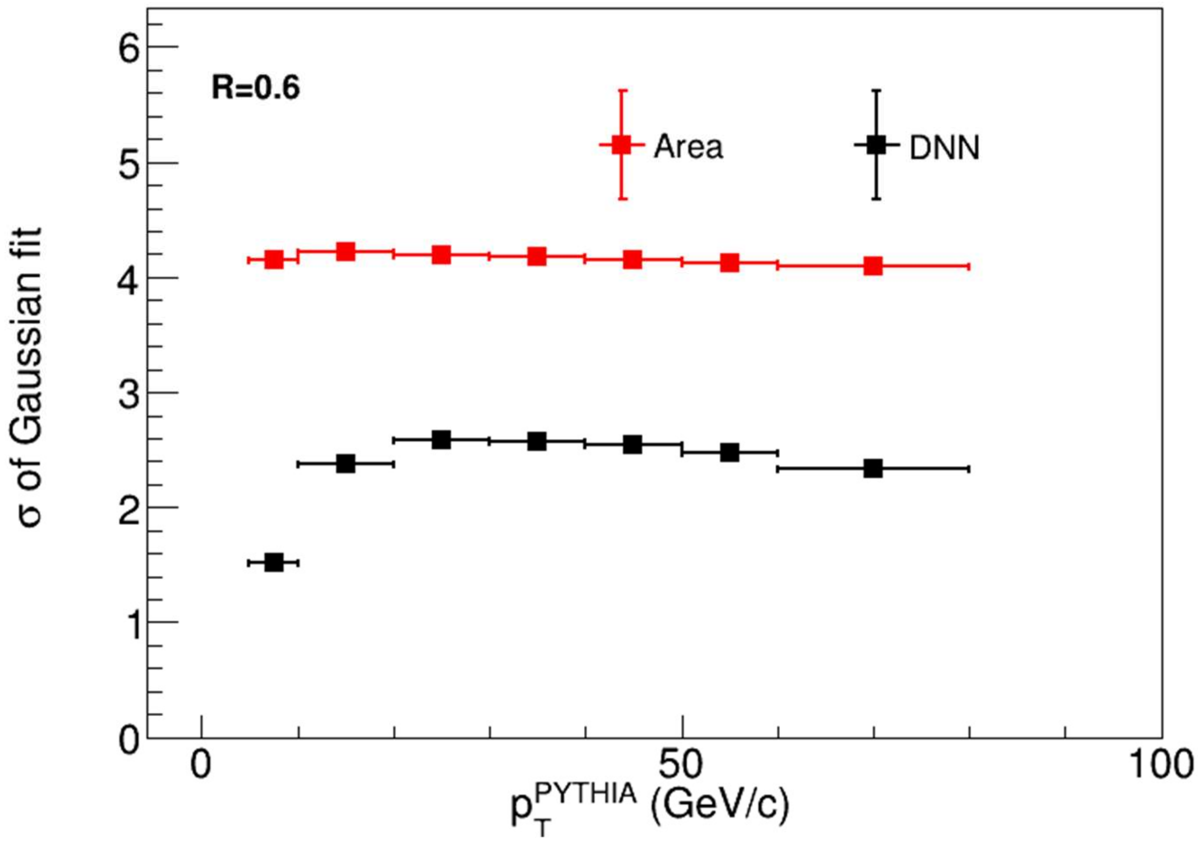


$$\rho \equiv Median\left\{\frac{p_T^{k_T\ Jet}}{A_{k_T\ Jet}}\right\}$$

Deep Neural Network

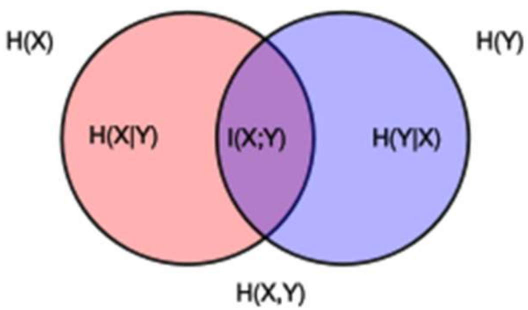
- Deep Neural Network (DNN):
  - Architecture: [N, 100,100, 50, 1]
  - Features: $p_T^{Jet,Uncorr.}$, N_{tracks} , λ_{Jet}, $p_T^{Track\ 0} - p_T^{Track\ 7}$
- Similar architecture to neural network laid out in paper

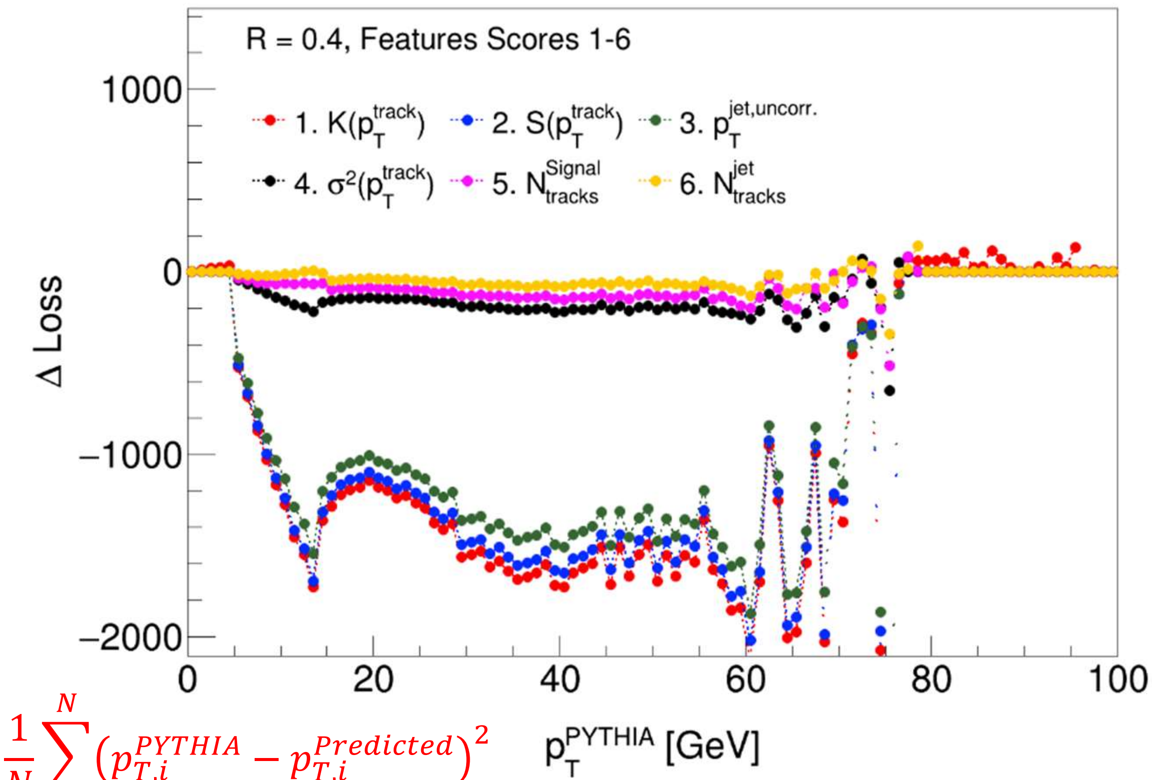
R. Haake [Phys.Rev.C 99 (2019) 6, 064904]



Feature space optimization

- Mutual Information: I(X;Y) = H(X,Y) – H(X|Y) –H(Y|X)
- Permutation Scoring: Randomly permutes feature to see change in Cost (mean squared error) evaluation

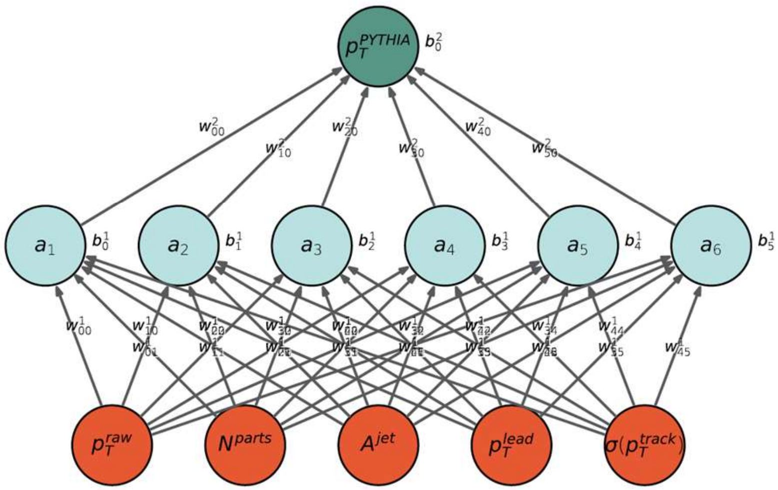


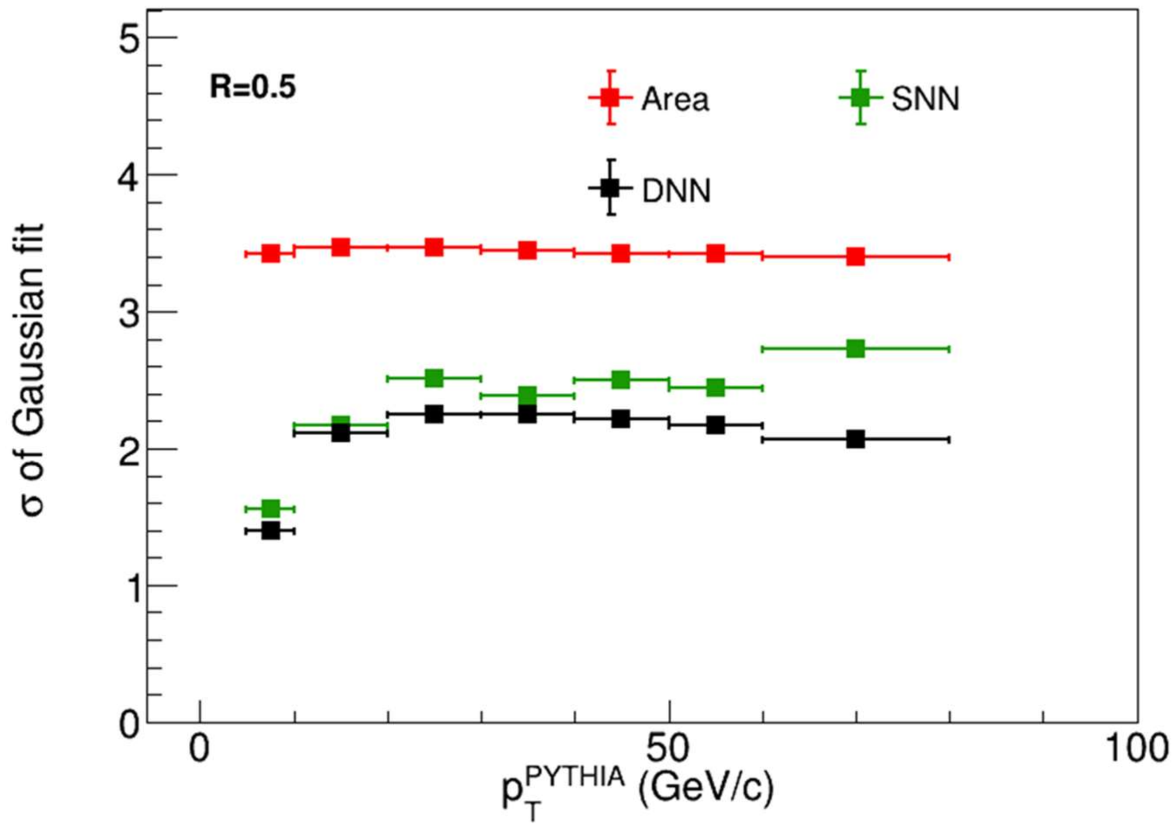


$$MSE = \frac{1}{N}\sum_{i}^{N}\left(p_{T,i}^{PYTHIA} - p_{T,i}^{Predicted}\right)^2$$

Shallow Neural Network

- Similar Performance to Deep Neural Network
- From ~16000 parameters to 36





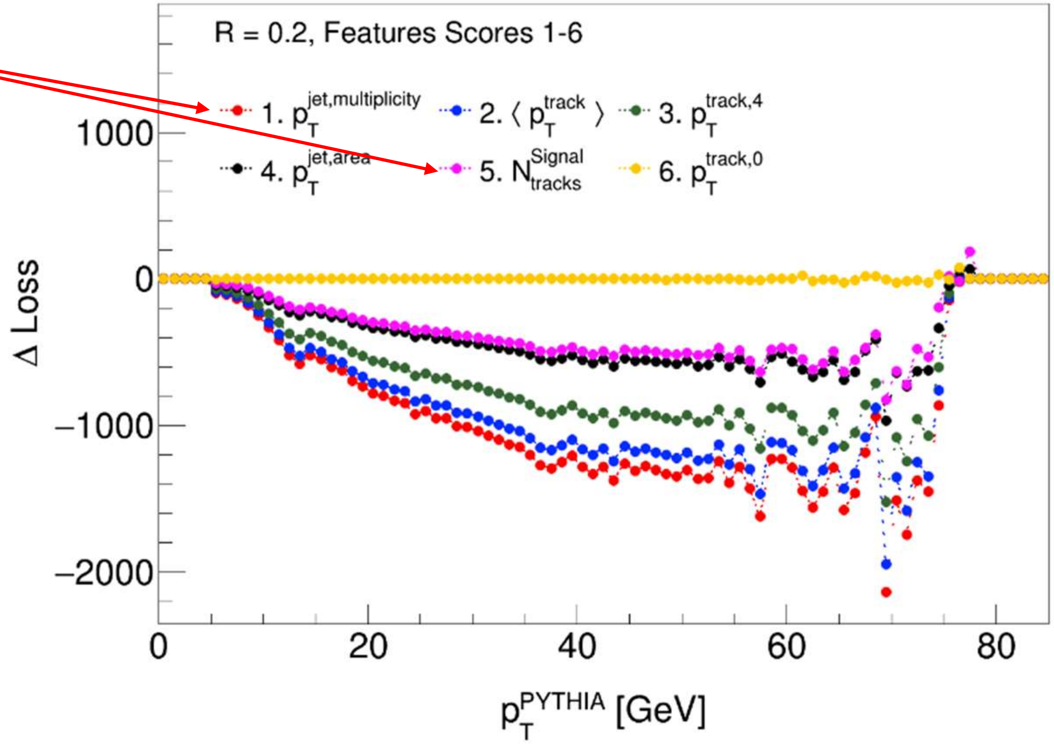
Physics Intuition + ML

- $N_{constit}^{Jet}$ appears near the top of permutation importance
- Width of residual is driven by particle number fluctuations

ALICE [JHEP 03 (2012), 053](#)

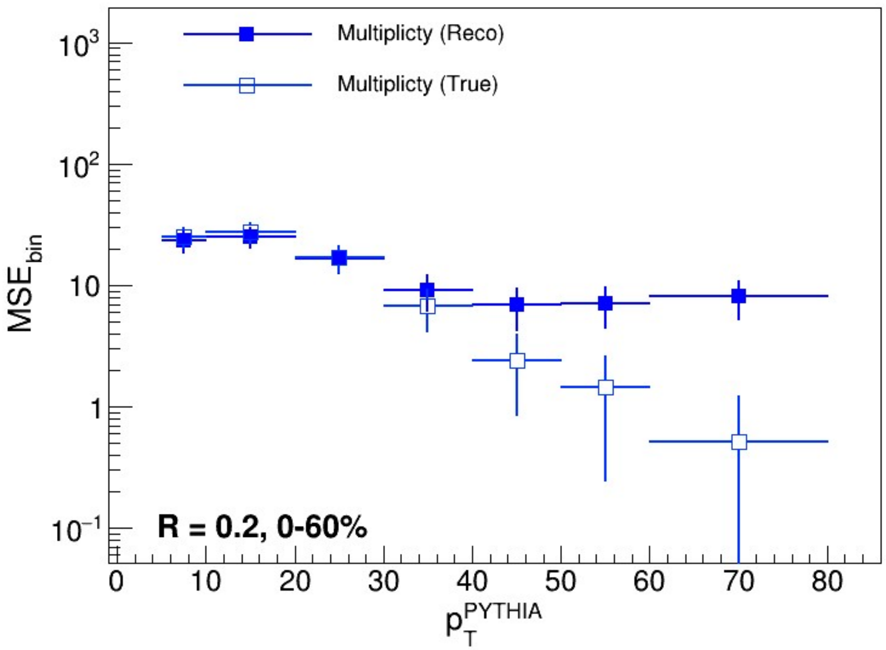
$$\sigma = \sqrt{N\sigma_{p_T}^2 + (N + 2N^2 \sum_n v_n^2 \mu_{p_T}^2}$$

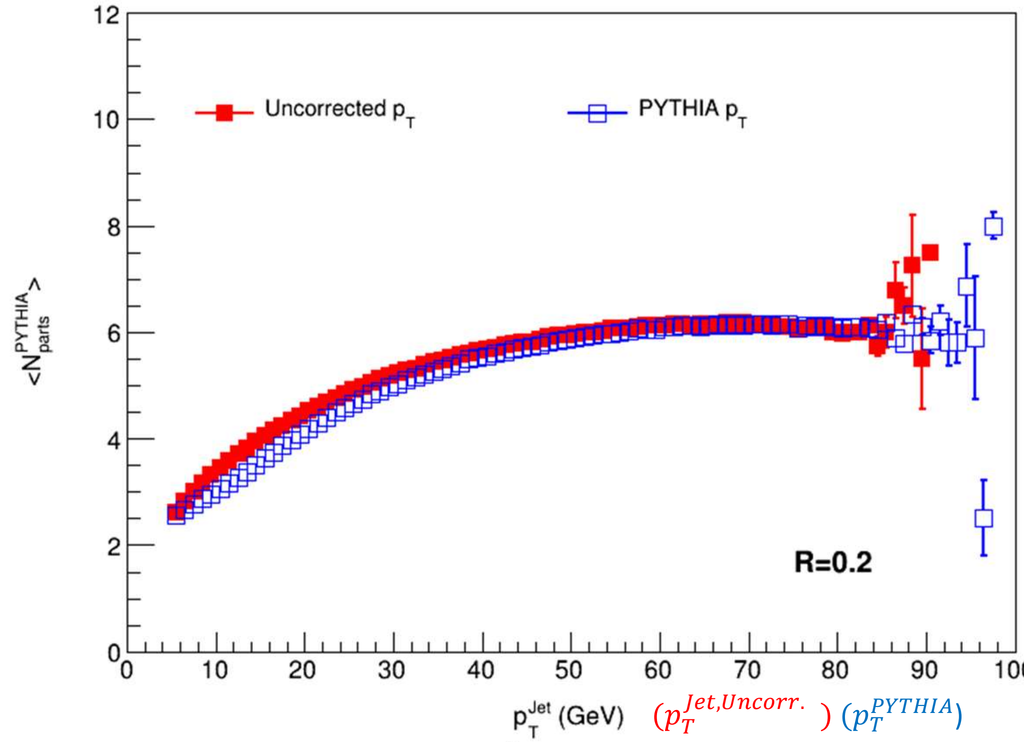
$$p_T^{Corr.} = p_T^{Uncorr.} - \rho\left(N_{constit}^{Jet}\right)$$



Multiplicity Method

- Multiplicity Based Subtraction:
  - $p_T^{Corr.} = p_T^{Uncorr.} - \rho\big(\underbrace{N_{constit}^{Jet} - \langle N_{PYTHIA}^{Jet}\rangle}_{\approx N_{Background}^{Jet}}\big)$

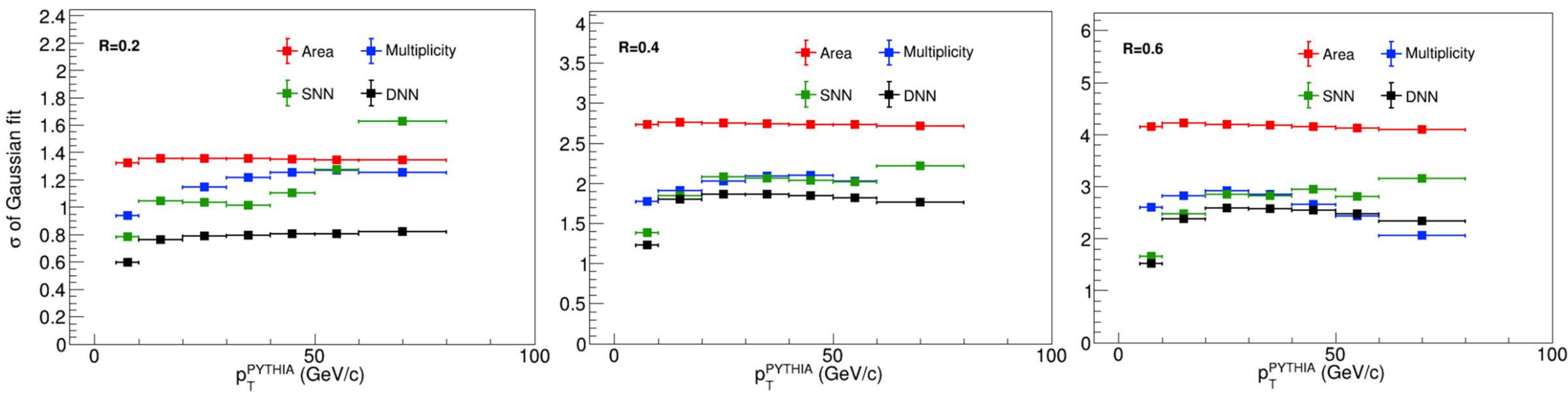




$(p_T^{Jet,Uncorr.})$ (p_T^{PYTHIA})

Results

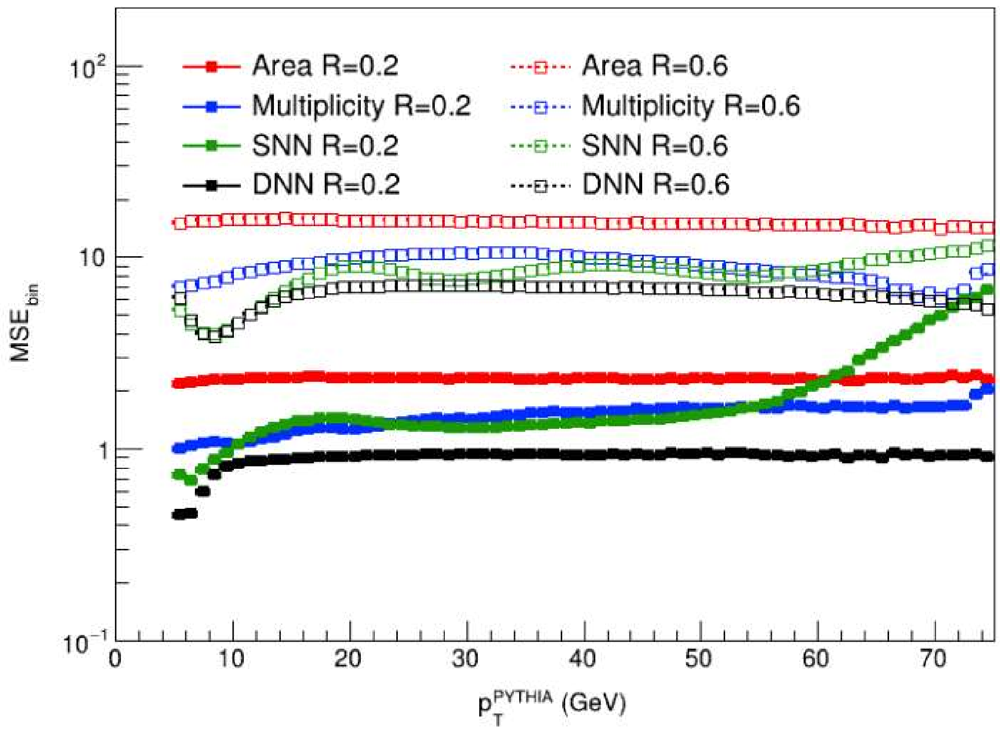
- Deep Neural Network will always perform best
- Shallow Neural Network can be engineered for similar performance
- Jet track multiplicity method comparable to ML



Results

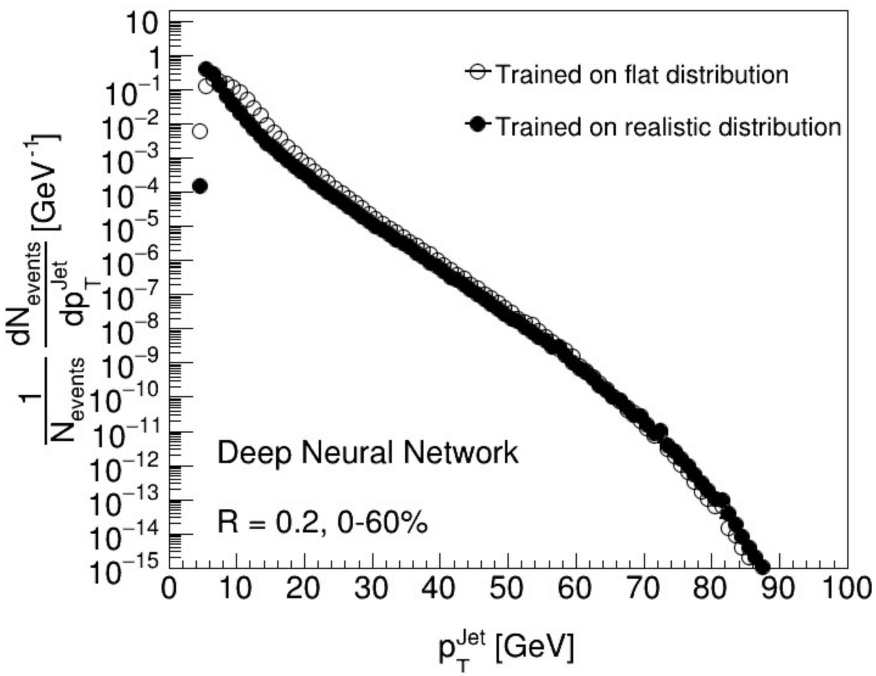
$$MSE = \frac{1}{N}\sum_{i}^{N}\left(p_{T,i}^{PYTHIA} - p_{T,i}^{Predicted}\right)^2$$

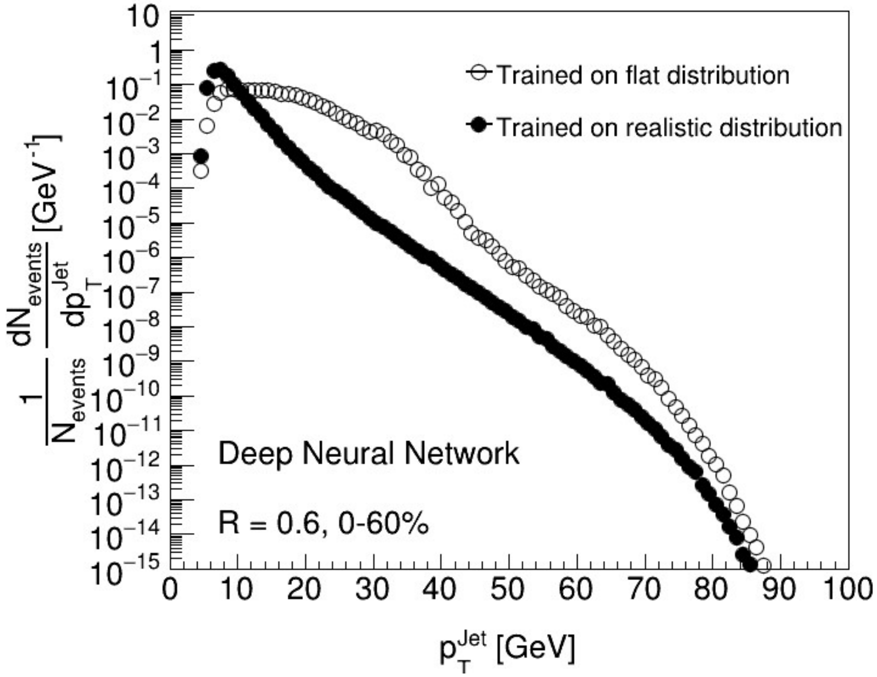
- Multiplicity and Area based methods are consistent at all momentum regions



Something to think about

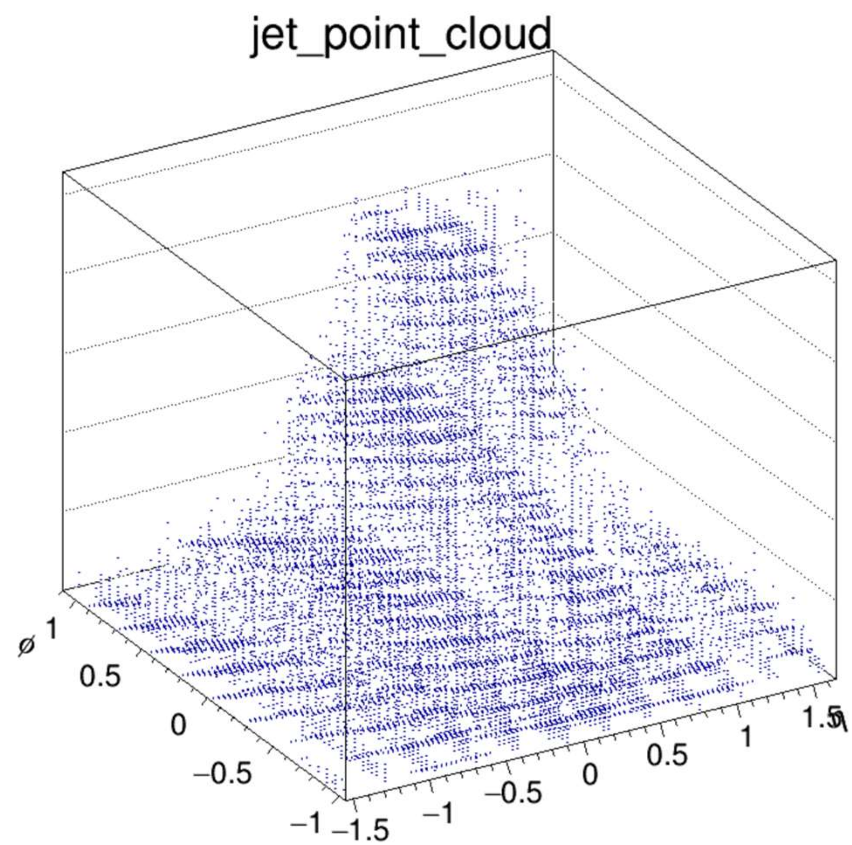
- ML is highly dependent on training set
- Need clear understanding of how your model is biased





Conclusions

- Machine Learning can be used as a tool to improve understanding of a complex problem (Jet background subtraction)
- Jet track multiplicity can be used to improve background subtraction over current methods
- ML methods need refinement and clear understanding of what it's doing

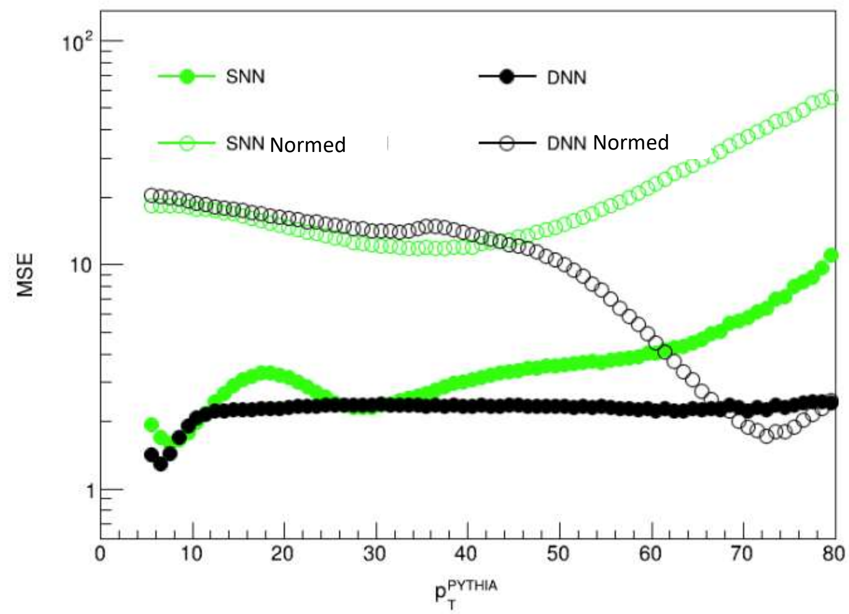


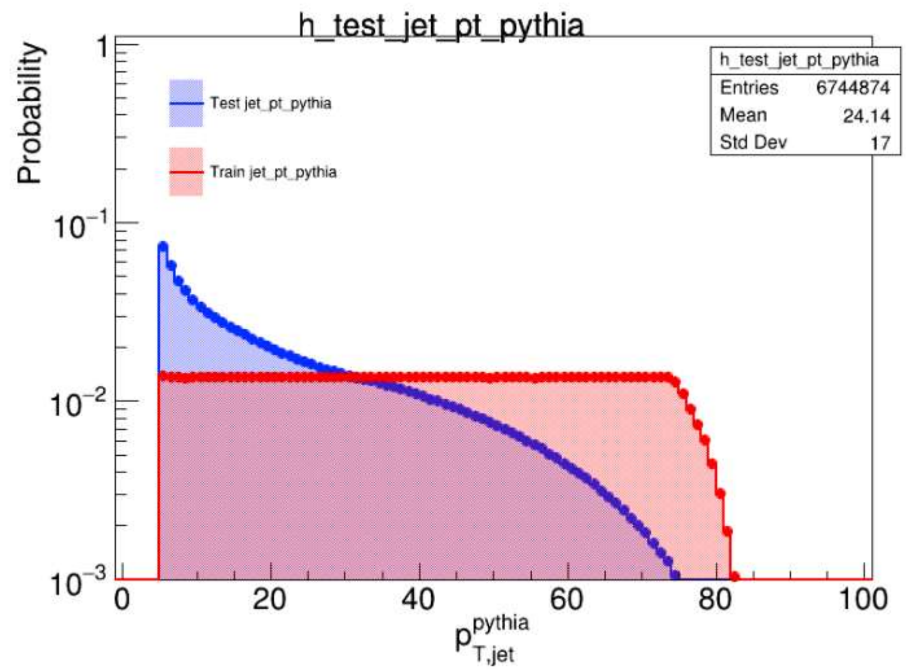
Back-up

The *'s

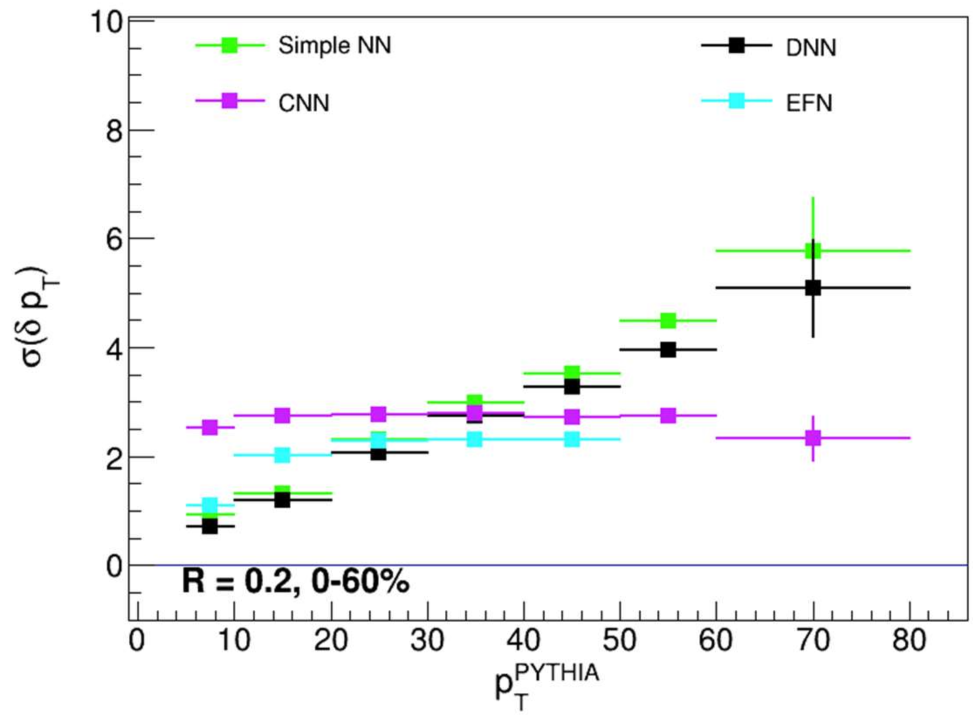
$$MSE = \frac{1}{N}\sum_{i}^{N}\left(p_{T,i}^{PYTHIA} - p_{T,i}^{Predicted}\right)^2$$

- The relationship between input features and corrected jet momentum is dependent on … the jet momentum!

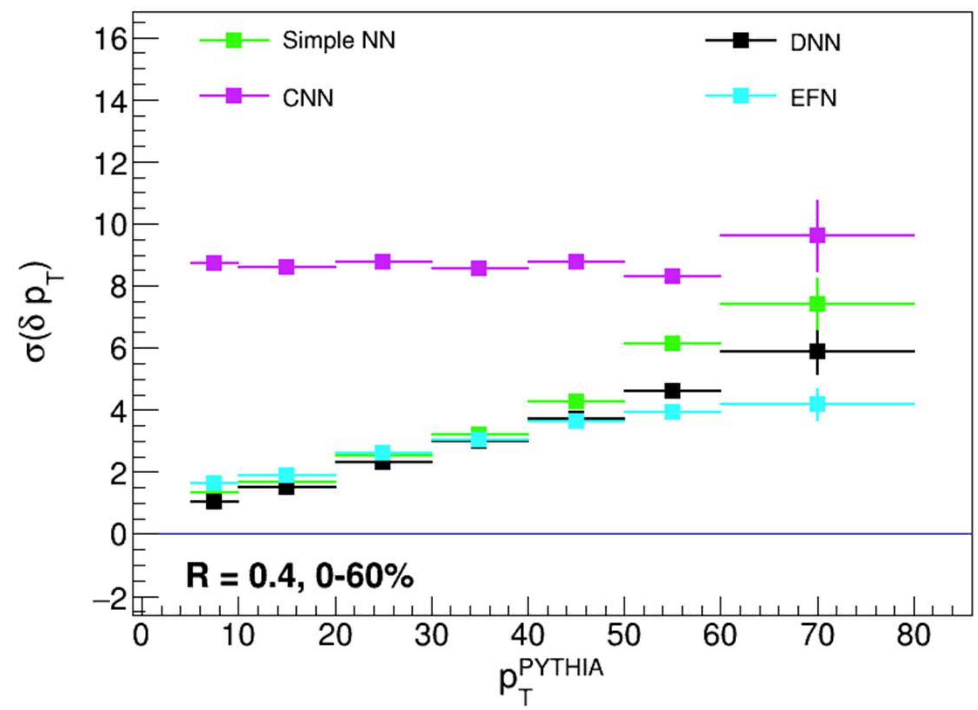




Coming soon



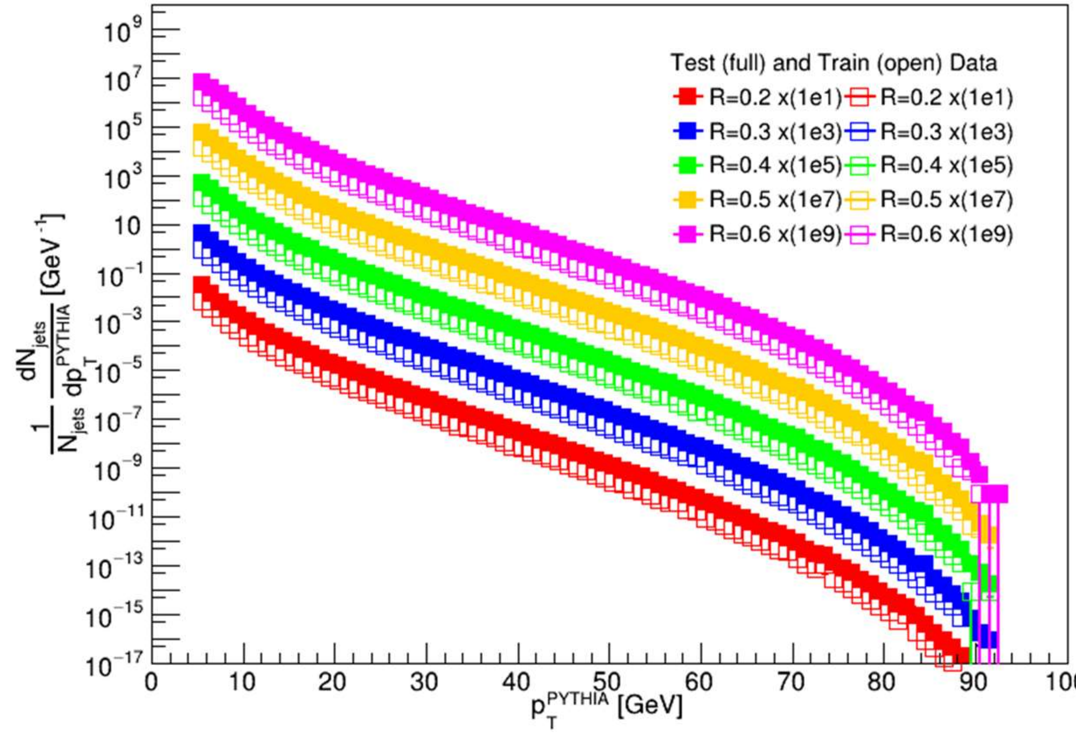
Simple NN
DNN
CNN
EFN
σ(δ p_T)
R = 0.2, 0-60%
p_T^PYTHIA



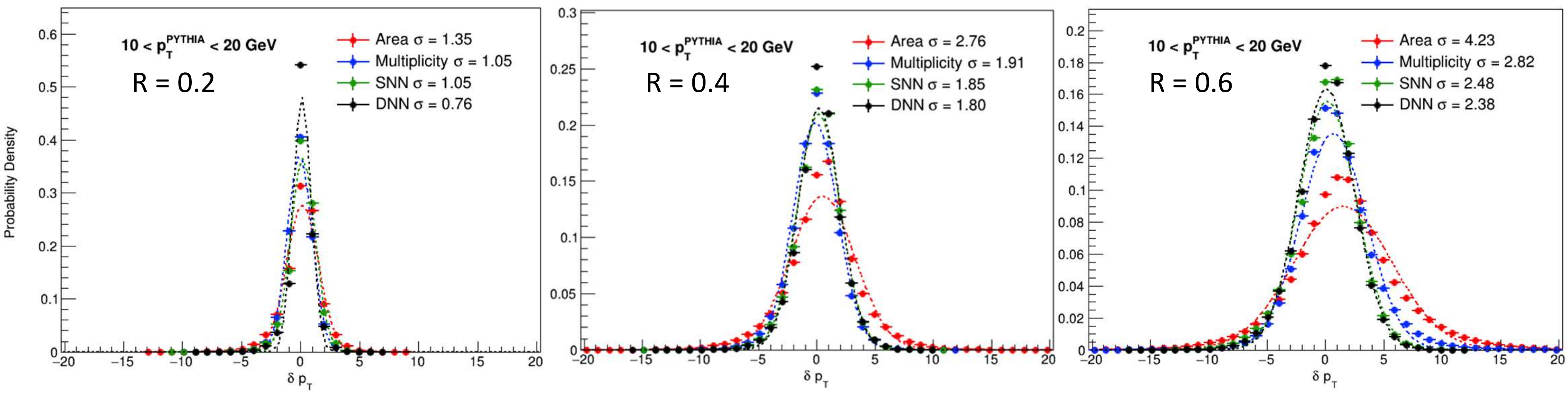
Simple NN
DNN
CNN
EFN
σ(δ p_T)
R = 0.4, 0-60%
p_T^PYTHIA

Jet Simulation Routine

- Merge p+p PYTHIA events with TennGen Au+Au background
- Find anti-k_T jets
  - Only save jets with p_T^{PYTHIA}> 5.0 GeV
- Take p_T^{PYTHIA} to be truth value
  - Train-Test split: 20/80



Results

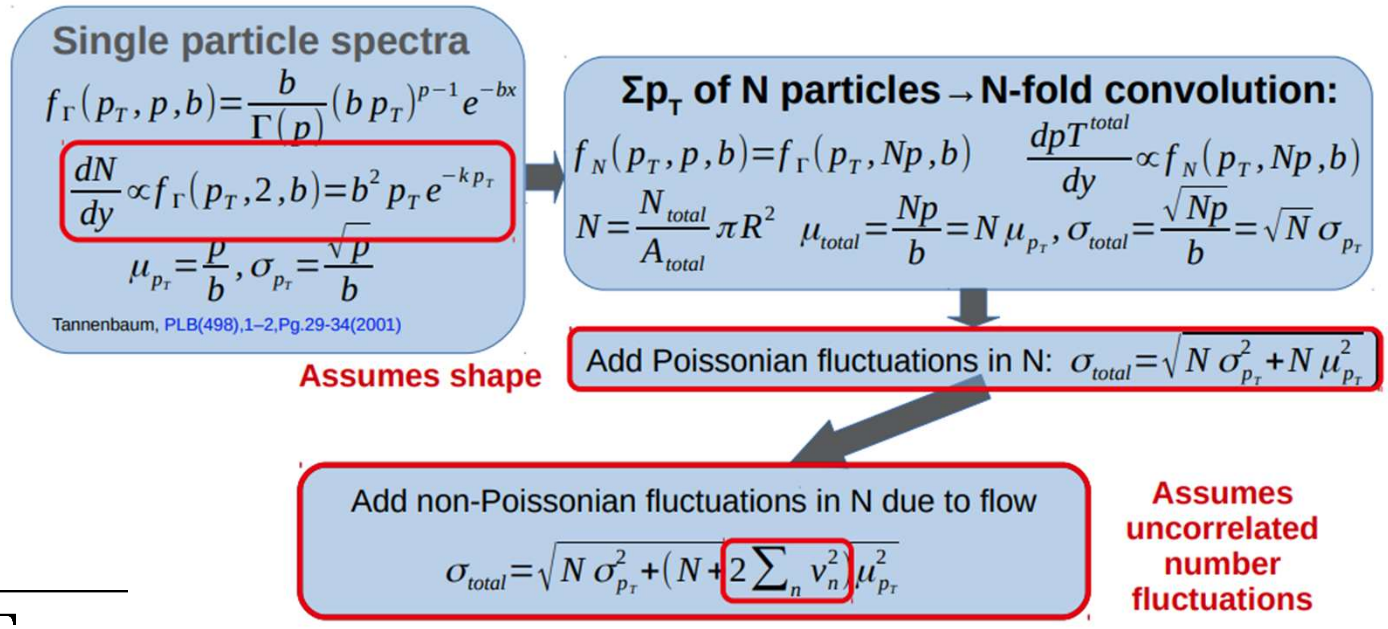


$$\delta p_T = \left(p_{T,i}^{PYTHIA} - p_{T,i}^{Predicted}\right)$$

Physics Intuition + ML

- $N_{constit}^{Jet}$ appears near the top of permutation importance
- Width of residual is driven by particle number fluctuations

ALICE [JHEP 03 (2012), 053]

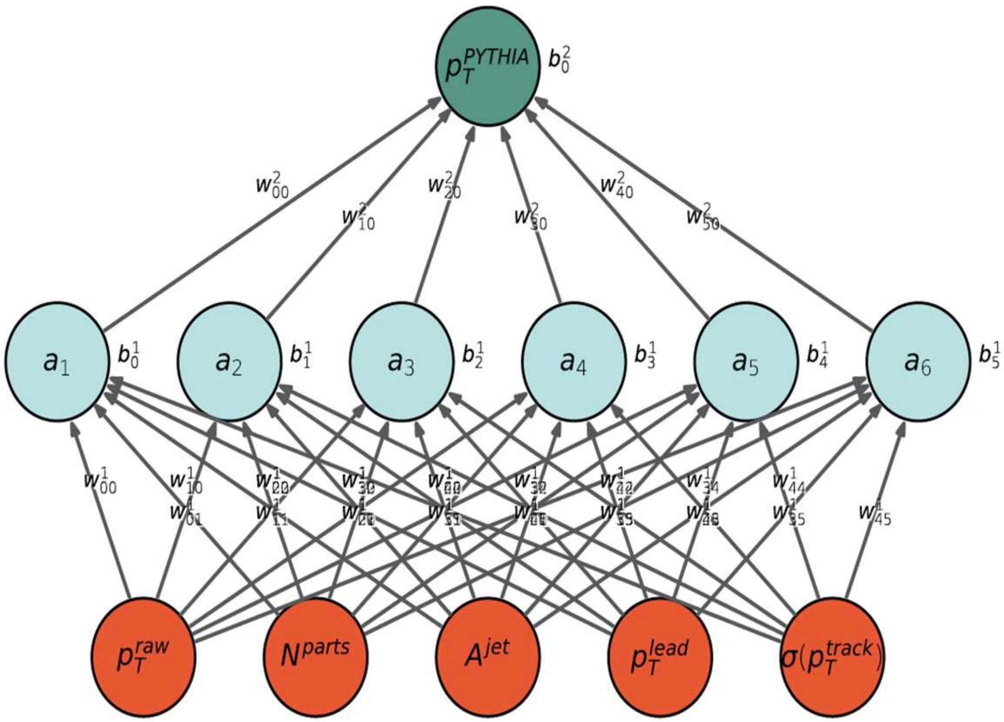


$$\sigma = \sqrt{N\sigma_{p_T}^2 + (N + 2N^2\sum_n v_n^2\mu_{p_T}^2}$$

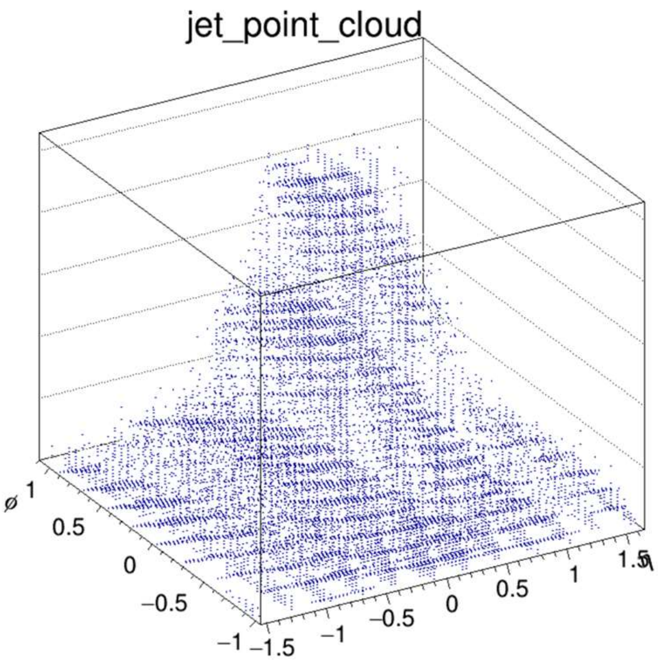
$$p_T^{Corr.} = p_T^{Uncorr.} - \rho\, N_{constit}^{Jet}$$

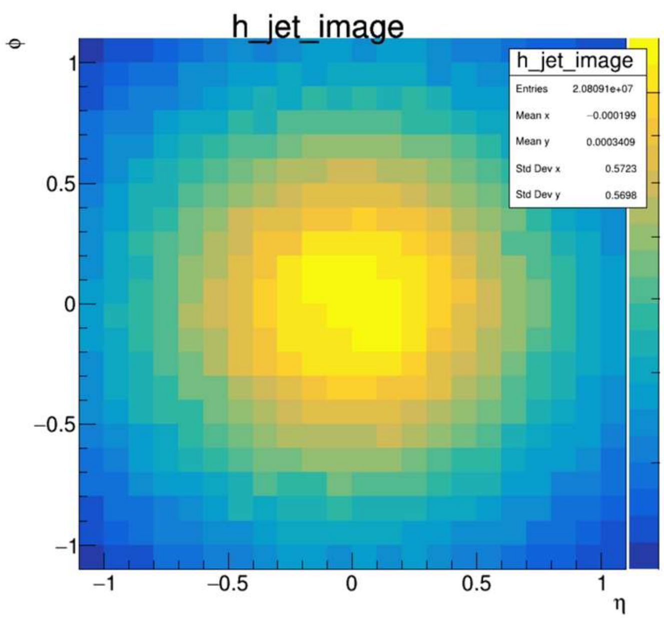
Machine Learning Models

- TensorFlow Keras Sequential Models:
  - All Models:
    - ADAM Optimizer
    - Standard Scaling
    - 100 Epochs (unless early stopping is triggered)
    - Cost Function/Metrics = Mean Sq. Error
    - Batch size = 1000



Jet Representations





["jet_pt_raw","jet_nparts"
,"jet_angularity",
"jet_track_pt_0","jet_trac
k_pt_1","jet_track_pt_2",
"jet_track_pt_3","jet_trac
k_pt_4","jet_track_pt_5",
"jet_track_pt_6",
"jet_track_pt_7]

Energy Flow Network

- IRC safe network
  - Proof via deep sets Theorem: *https://arxiv.org/abs/1810.05165*