Collider searches for Dark Matter using Machine Learning

Michael Soughton
University of Sussex

Work in progress
In collaboration with Charanjit Kaur Khosa and Veronica Sanz

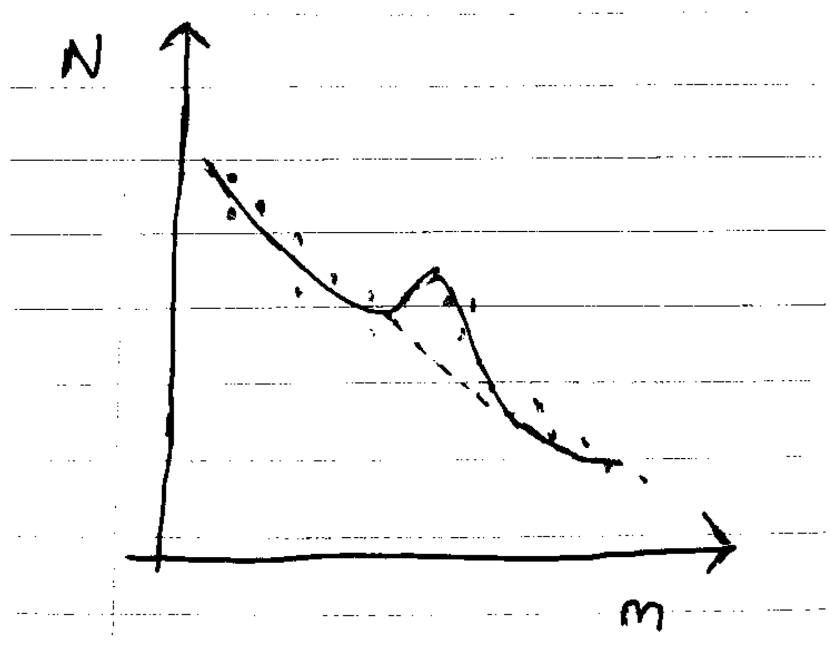
Overview

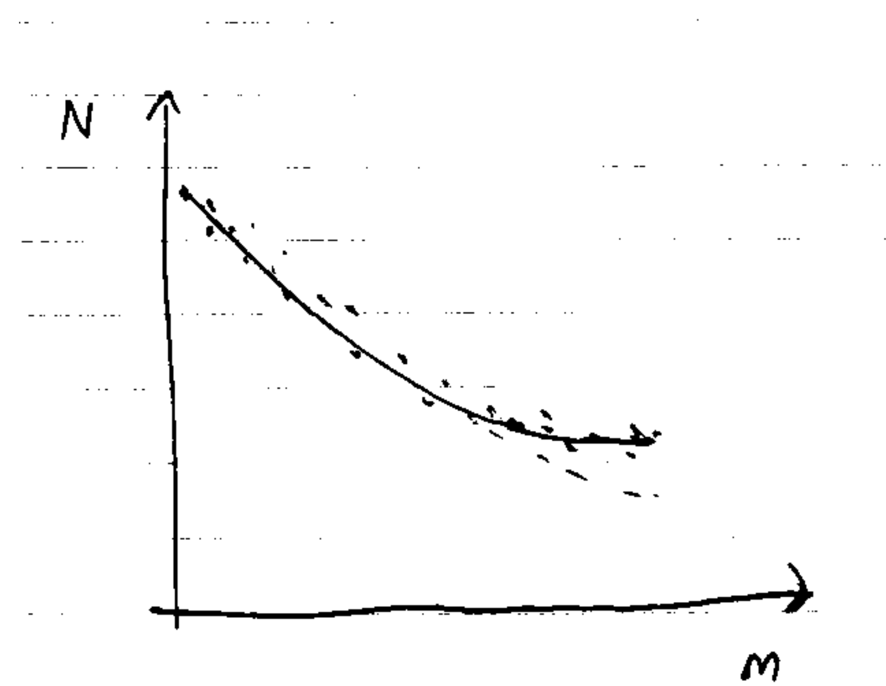
- Where is the New Physics?
- The models I will consider.
- Signatures of the models to look for.
- Machine Learning methods.
- Results.
- The road ahead.

Where is the New Physics?

- Current analysis of data is only putting deeper constraints on BSM physics

- Predictions are made across several variables, each with noisy distributions. Theoretical parameters often unconstrained.





A new approach to data analysis

- Assume that signatures of New Physics exist within the data but are well hidden in the background and lie across higher dimensional spaces.

- Train a Machine Learning algorithm to search for these signatures.

Machine Learning methods

- Machine Learning algorithms learn from and make predictions from data.

- Can learn complex non-linear patterns.

- Benefits from large datasets.

Machine Learning methods

- Supervised ML used for tasks where labelled data exists.
  - Linear/logistic regression
  - Neural Networks
  - Decision trees
- Unsupervised ML used for finding patterns and structure in unlabelled data.
  - Generative Adversarial Networks (GANs)
  - Variational Autoencoders (VAEs)

Dark Matter in colliders: two candidate models

- Axion Like Particles (ALPs):

$$\mathcal{L}_a = \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}M_a^2 a^2 - \frac{g_{a\gamma}}{4}\,aF_{\mu\nu}\tilde{F}^{\mu\nu} - \frac{g_{agg}}{2}\,a\,Tr\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right] + \sum_{\psi} g_a^{\psi} m_\psi a\overline{\psi}\gamma^5\psi$$

- Spin-1 mediated WIMPs:

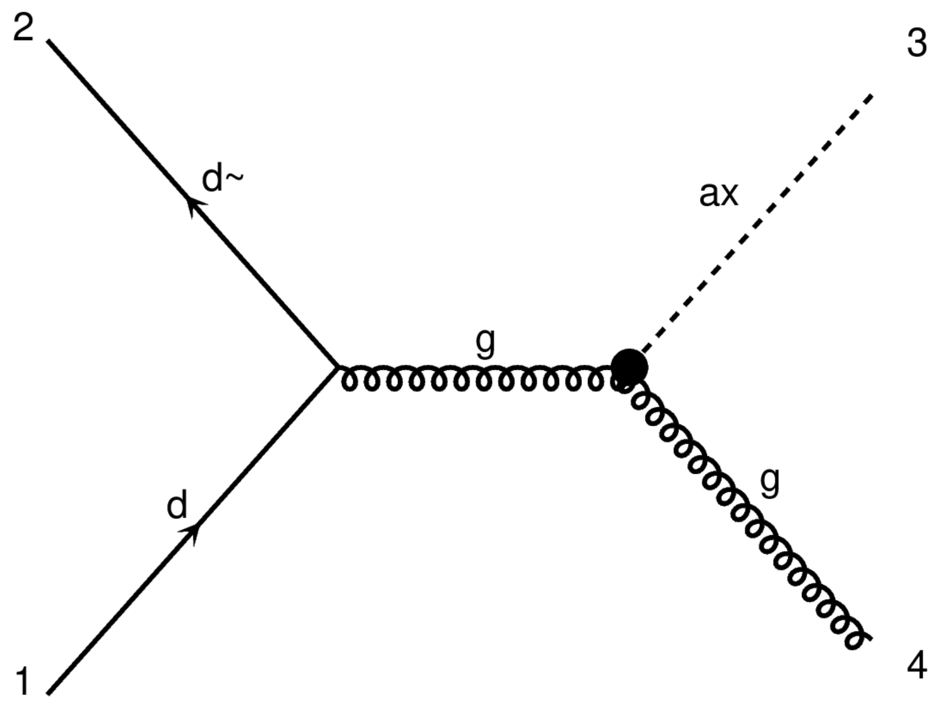
$$\mathcal{L}_{DM}^{Y_1} = \overline{\chi}\gamma_\mu\left(g_{DM}^V + g_{DM}^A\gamma^5\right)\chi Y_1^\mu$$

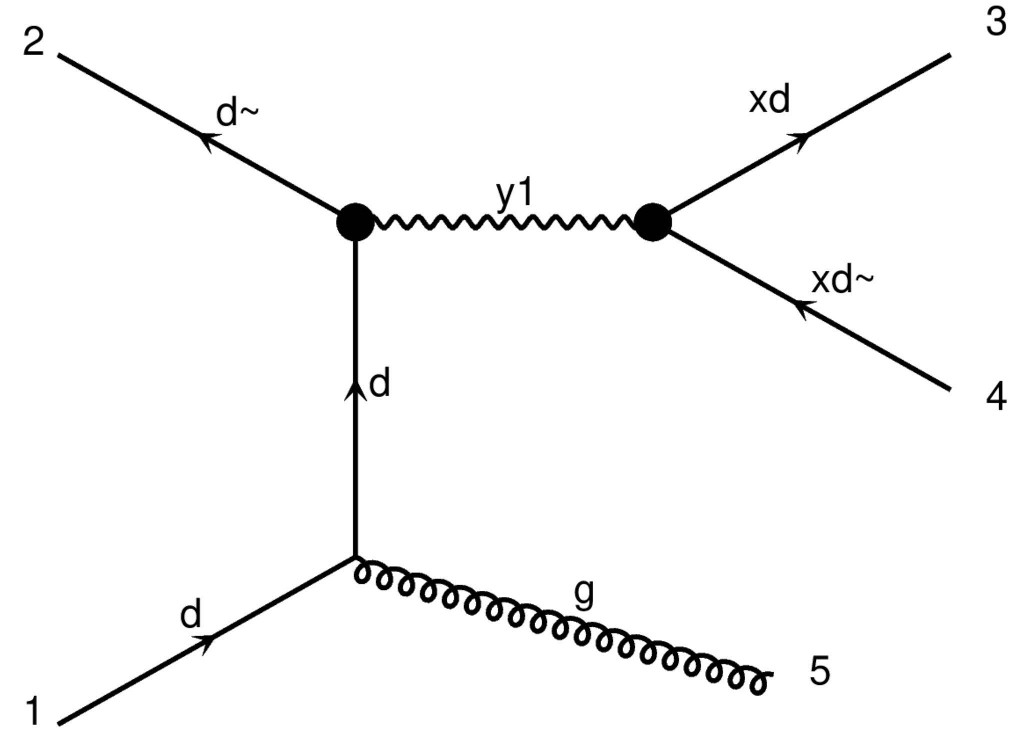
$$\mathcal{L}_{SM}^{Y_1} = \bar{t}\gamma_\mu\left(g_t^V + g_t^A\gamma^5\right)tY_1^\mu + \bar{b}\gamma_\mu\left(-g_t^A\gamma^5\right)bY_1^\mu$$

- Look into mono-jet processes, specifically the p_T^j, η_j, ϕ_j distributions.

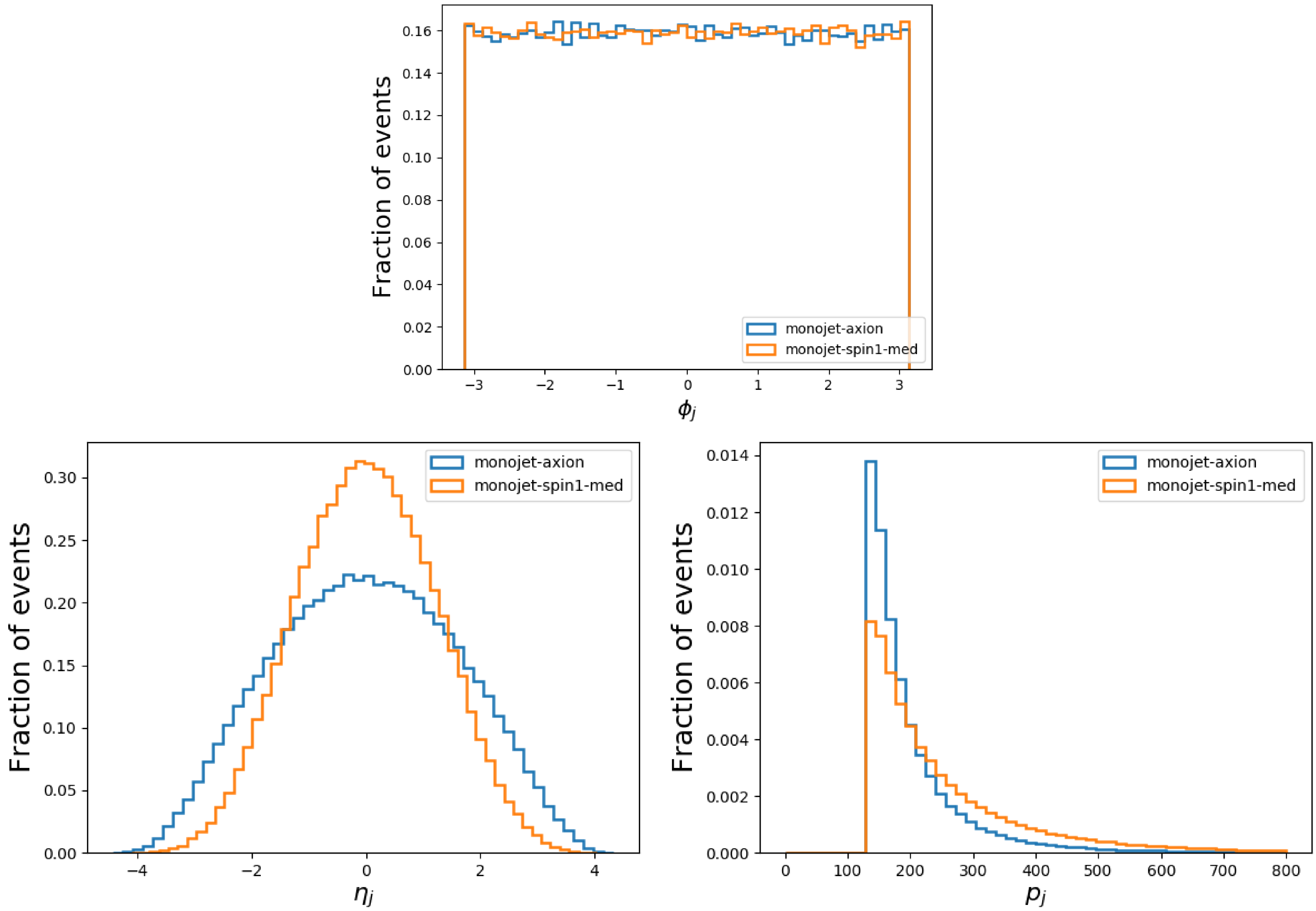
Processes considered

- Mono-jet processes are generated in MadGraph5_aMC@NLO to parton level.



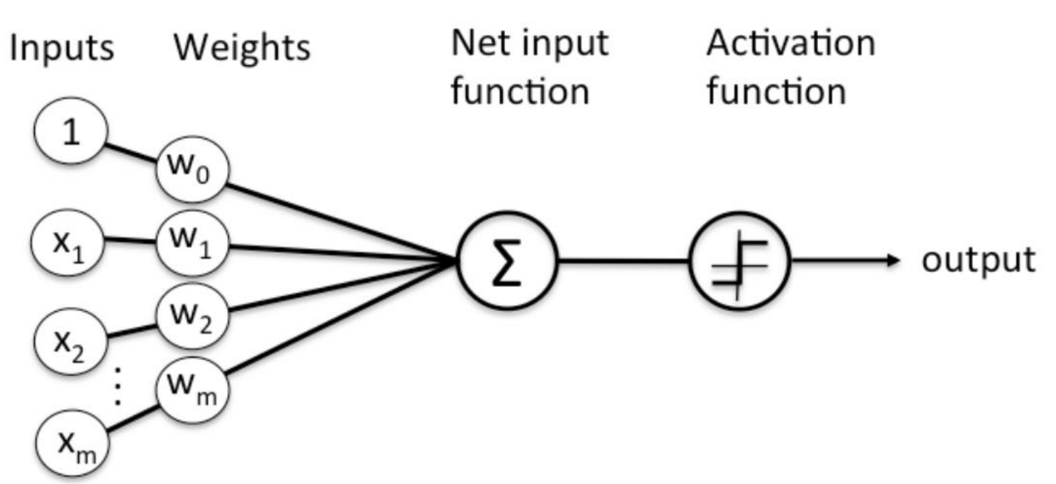


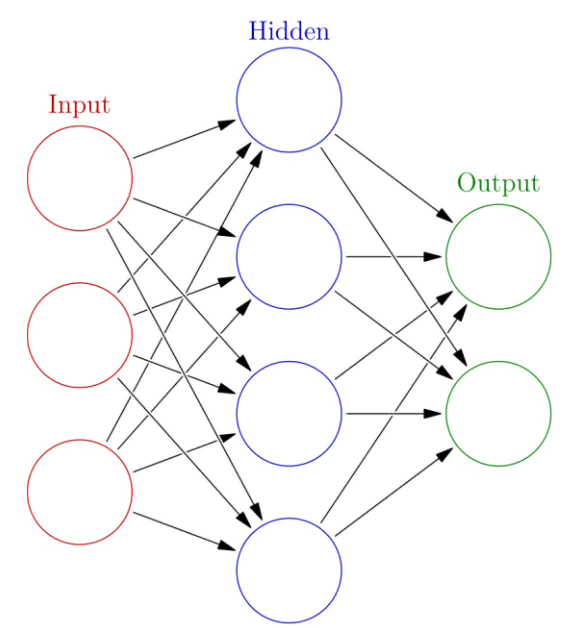
1D distributions



Finding patterns with ML

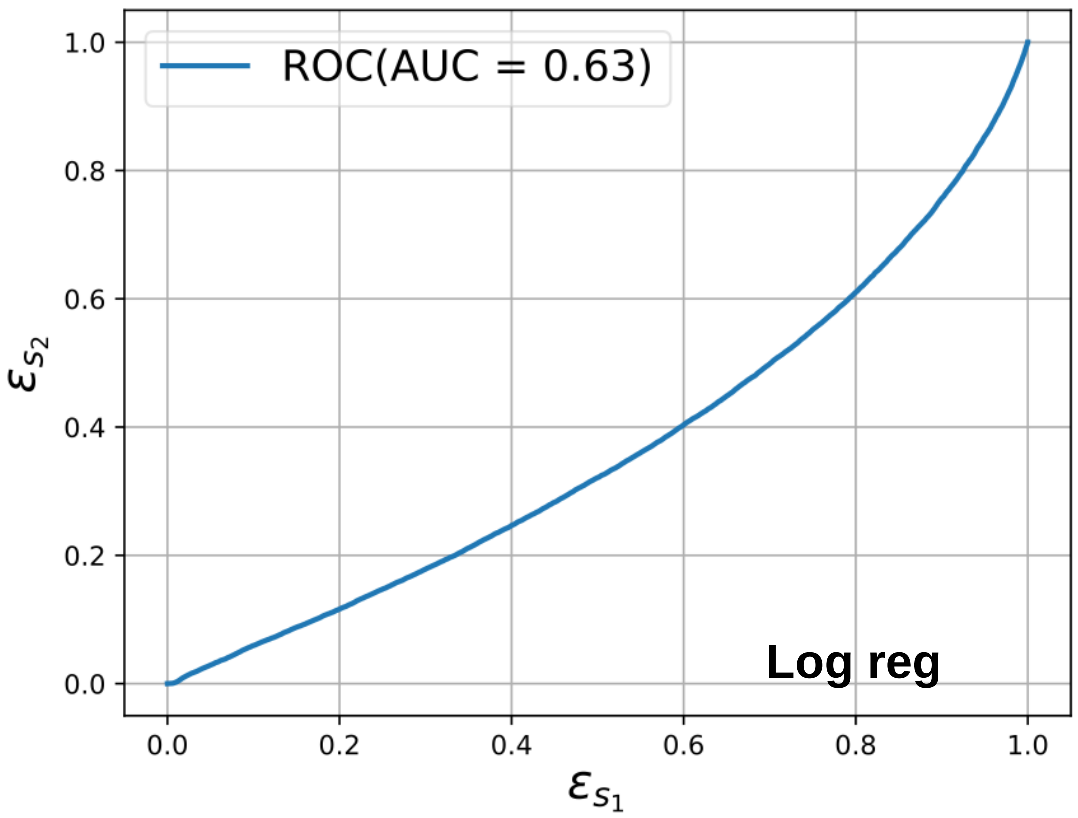
- Aim to predict output *y* from input data ***x***.
- Parameters of model learnt in training phase.
- Logistic regression finds parameters through minimising a ‘cost function’.
- Neural Networks consist of nodes that perform a linear transform then a non-linear function.

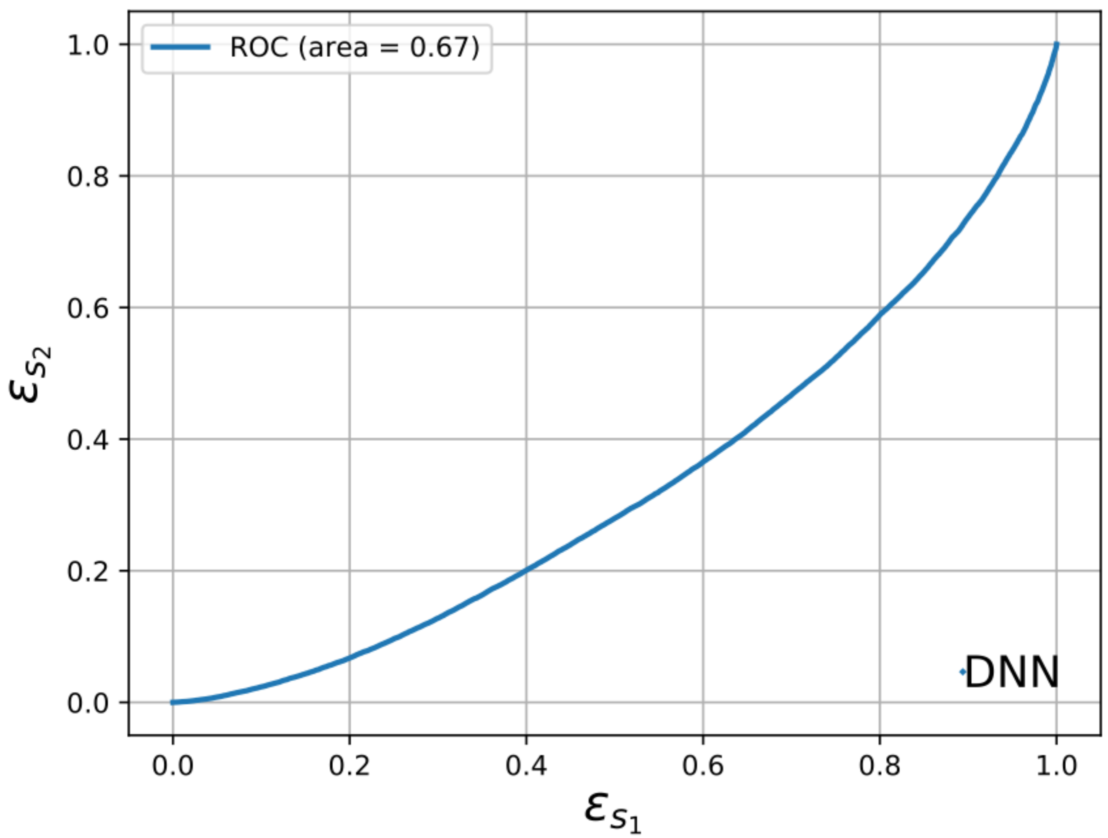




Logistic regression and Deep Neural Network (DNN)

Features : p_T^j, η_j, ϕ_j



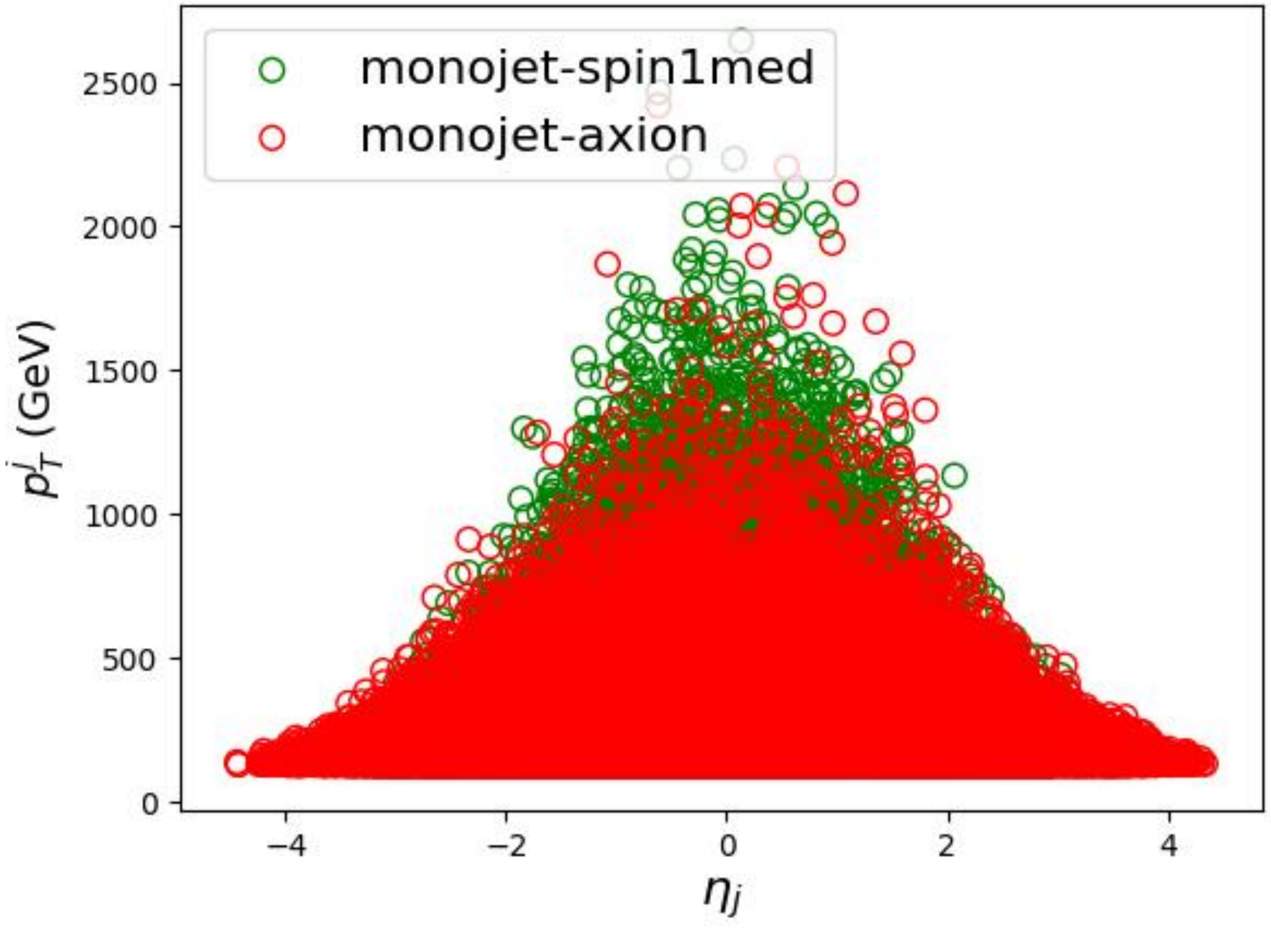


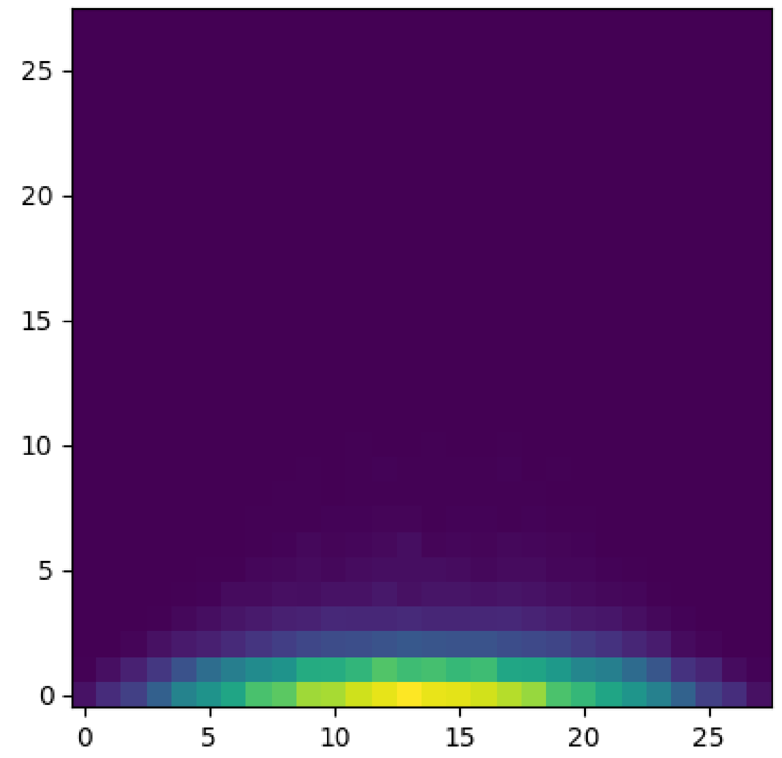
2D distributions

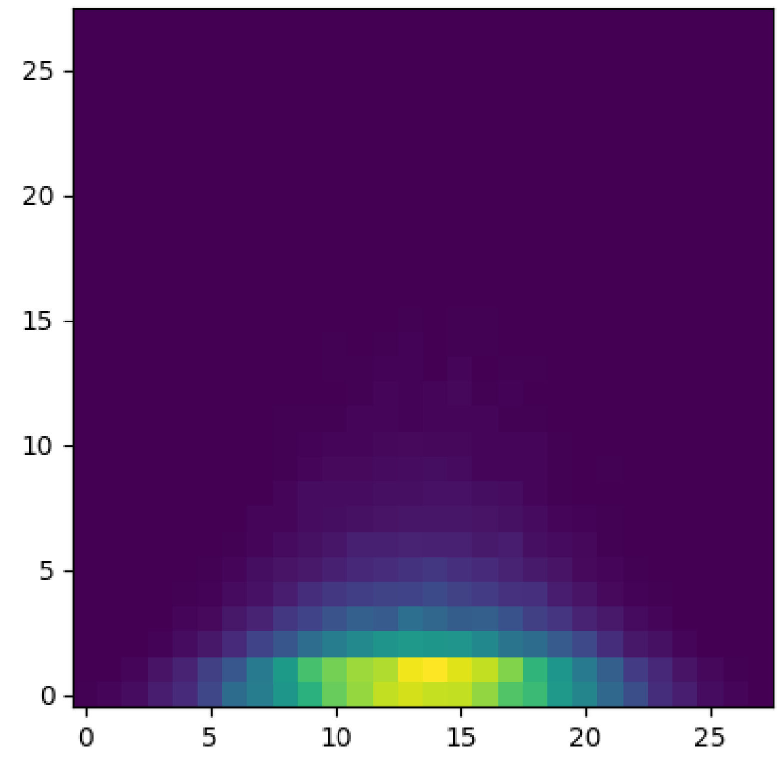
Binning: 28x28 (50K events per image)

$\eta_j : [-4,4]$

$p_T^j : [130,2000]GeV$

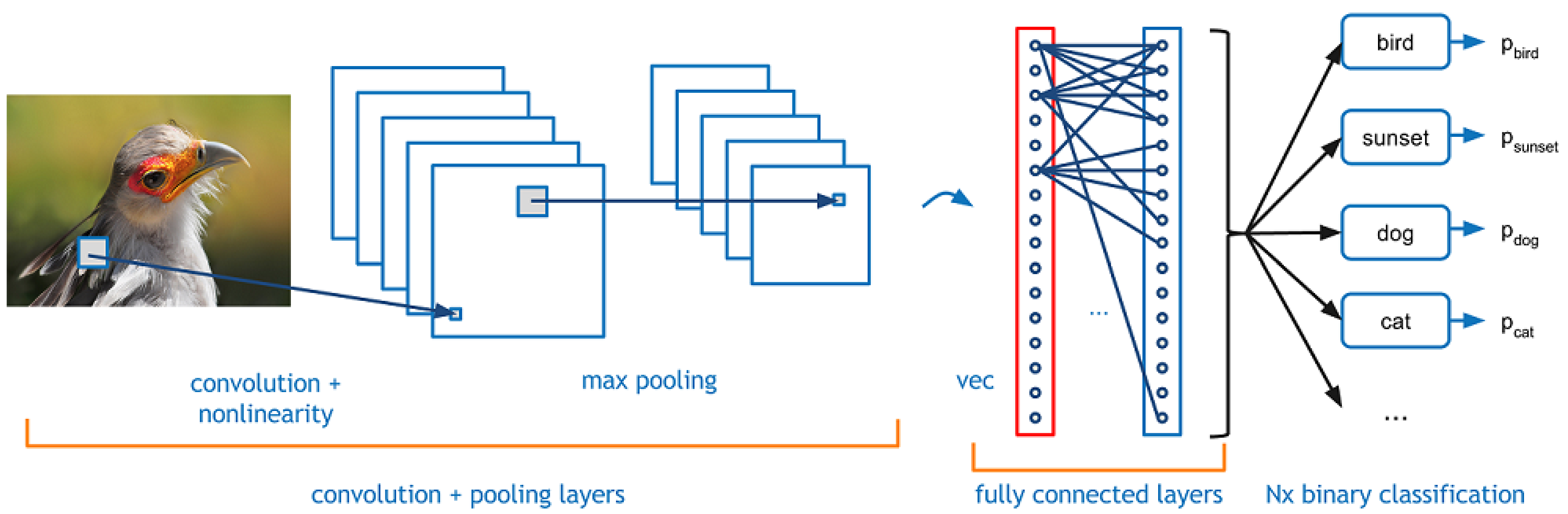






Convolutional Neural Networks (CNNs)

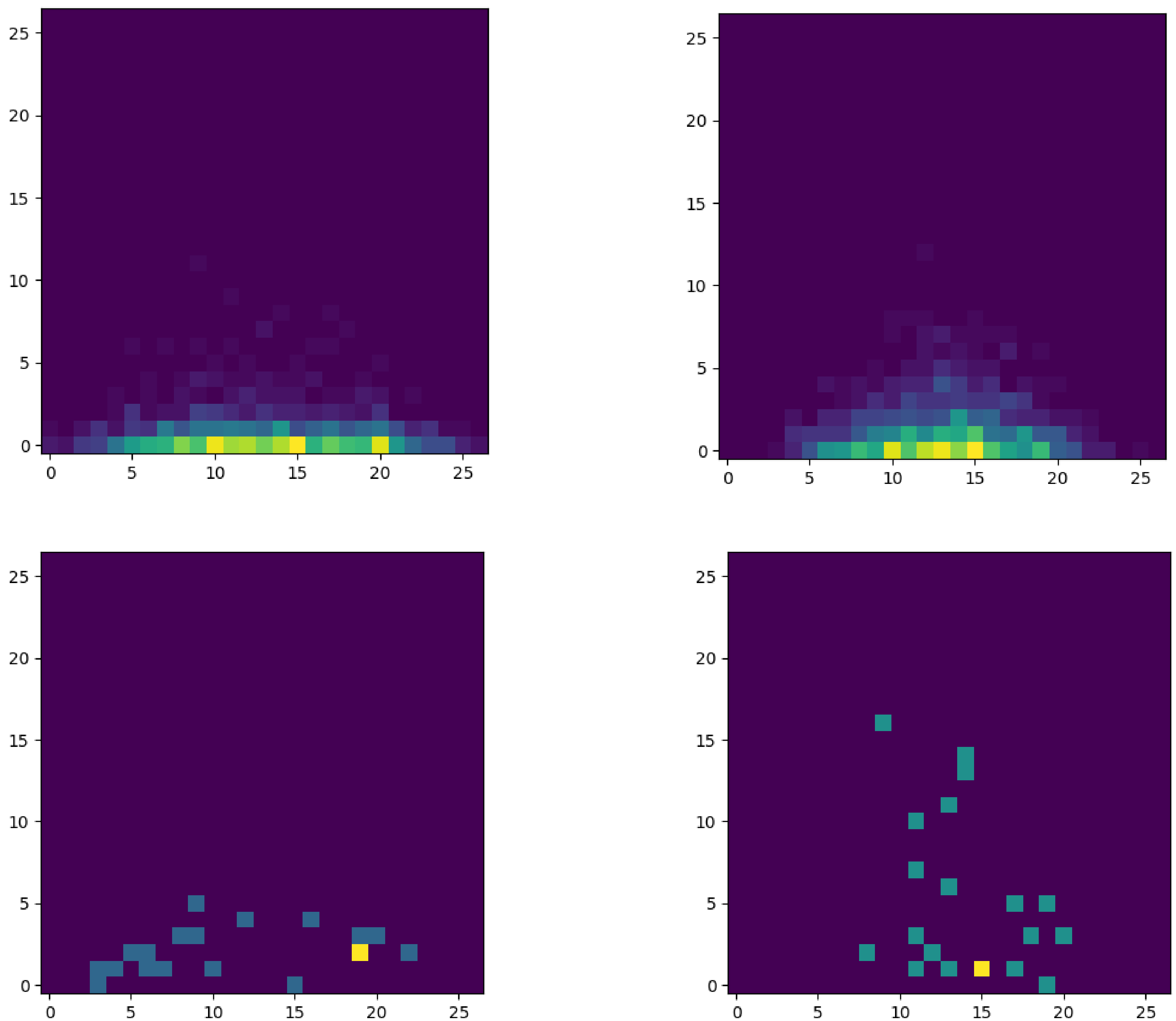
- Finds patterns from spatial features of data – correlations between variables.



Running the networks

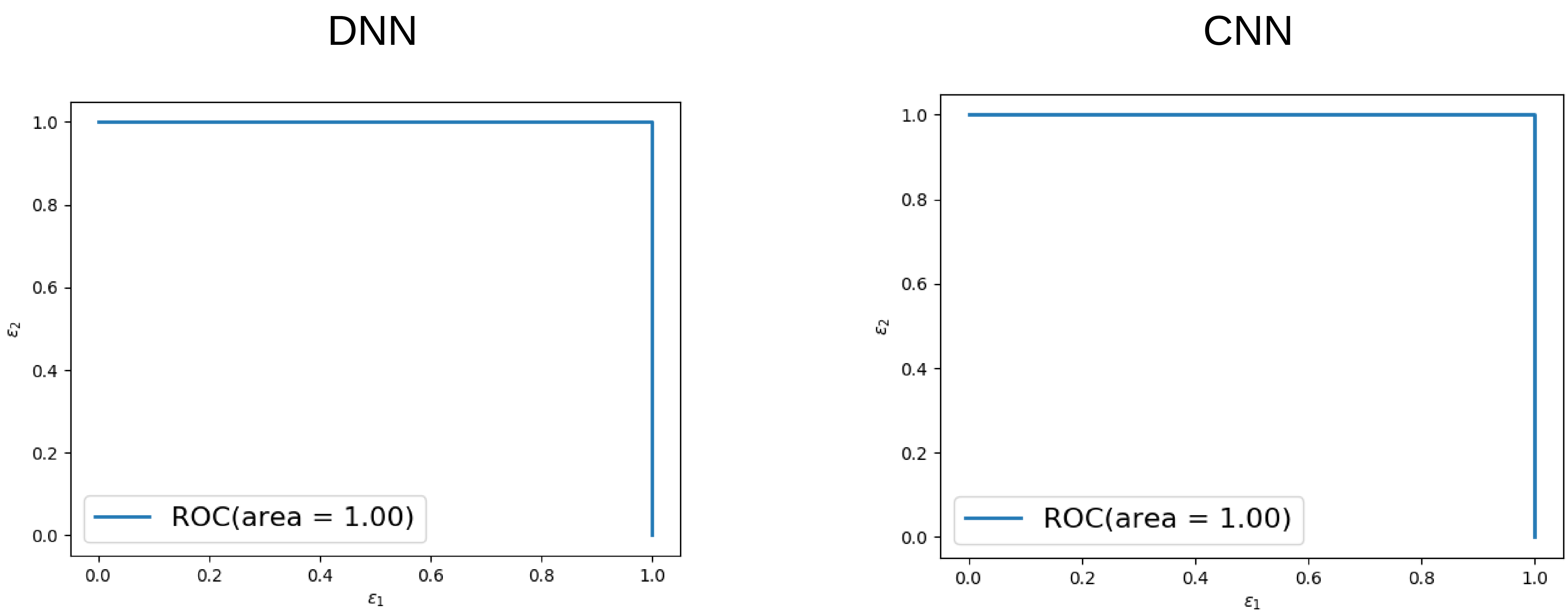
- **Information of correlation between variables exists in 2D histograms.**
- **Train on images with 1000 events per image.**
- **Training : Validation : Test Sample = 320:40:40.**
- **Hidden layers: 2**
- **Loss function: Binary cross-section**
- **Activation functions : ReLU and sigmoid/softmax.**

Testing against events with 20 events/image



Comparison of scores for density plots

- DNN accuracies:
  - Training: ~100%
  - Test: ~100%
  - Against data of 20 events/image: ~89.6%
- DNN mean absolute errors:
  - Training: ~2.11×10^{-5}
  - Test: ~1.36×10^{-5}
  - Against data of 20 events/image: ~0.165
- CNN accuracies:
  - Training: ~100%
  - Test: ~100%
  - Against data of 20 events/image: ~83.7%
- CNN mean absolute errors:
  - Training: ~2.14×10^{-5}
  - Test: ~3.29×10^{-5}
  - Against data of 20 events/image: ~0.332



DNN
CNN
ROC(area = 1.00)
ROC(area = 1.00)
0.0
0.2
0.4
0.6
0.8
1.0
ε_1
ε_2

Summary and the next steps

- Continue LO parton level analysis to distinguish between signals with background.
- Explore (algorithm) model accuracy with different (particle physics) model parameters.
- Perform NLO parton level analysis.
- Perform detector level analysis (jet distributions).
- Outlier exposure (unsupervised learning).

Thanks for listening!