Working Group 4

Hipernuclei identification with neural networks

Maria Paula M. Palhares[1]

Alexandre A. P. Suaide[1]

October 04, 2023

1 - University of São Paulo, Brazil

ex • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Index • Inc

Working Group: 4

Description: Application of machine learning techniques in high energy physics.

Index

Topics Covered

Start

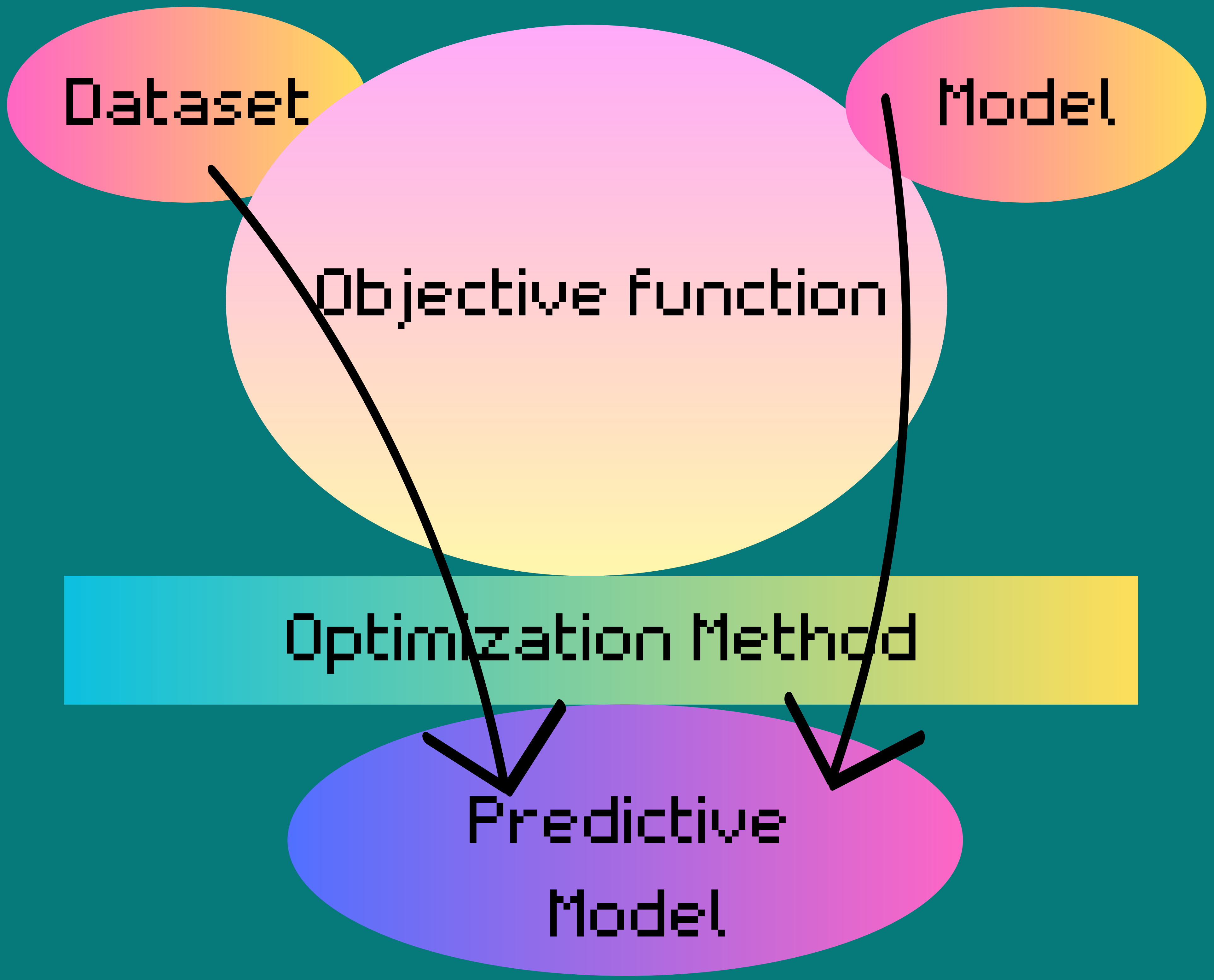
- ML in HEP
- NNs
- ALICE
- PhD project

MACHINE LEARNING

Terminology

- The use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.

ML Concept


Dataset
Model
Objective function
Optimization Method
Predictive
Model

MACHINE LEARNING

Automated monitoring of detector performance and data quality

Higher accuracy particle tagging

Event selection, reconstruction and classification

ML on HEP

Particle Physics

Special role

- Large amounts of experimental data + realistic MC simulations.
- Structured data at multiple scales.
- Details of systematic uncertainties..

ML categories

Supervised Learning

- Classification and Regression.
- Labelled data: desired output as supervisionary signal and input as vector.
- Correct result is known beforehand.

Unsupervised Learning

- Clustering and Association.
- Arbitrary model: obtain a deeper understanding of the data by recognising its basic structure or pattern of distribution.
- Input dataset devoid of any labelled outputs.

Semi Supervised Learning

- Classification and Clustering.
- Real-world learning issues: huge quantity input data, some of which are labelled.
- Hybrid model: combine supervised (or “discriminative”) and unsupervised (or “generative”) .

Reinforcement Learning

- Classification and Control.
- Interaction with enviroment of the problem.
- It chooses a current course of action based on previous encounters (exploitation) and fresh options (exploration) based on trial and error.

ML categories

Supervised Learning

- Classification and Regression.
- Labelled data: desired output as supervisionary signal and input as vector.
- Correct result is known beforehand.

Ok

Unsupervised Learning

- Clustering and Association.
- Arbitrary model: obtain a deeper understanding of the data by recognising its basic structure or pattern of distribution.
- Input dataset devoid of any labelled outputs.

Semi Supervised Learning

- Classification and Clustering.
- Real-world learning issues: huge quantity input data, some of which are labelled.
- Hybrid model: combine supervised (or “discriminative”) and unsupervised (or “generative”) .

Reinforcement Learning

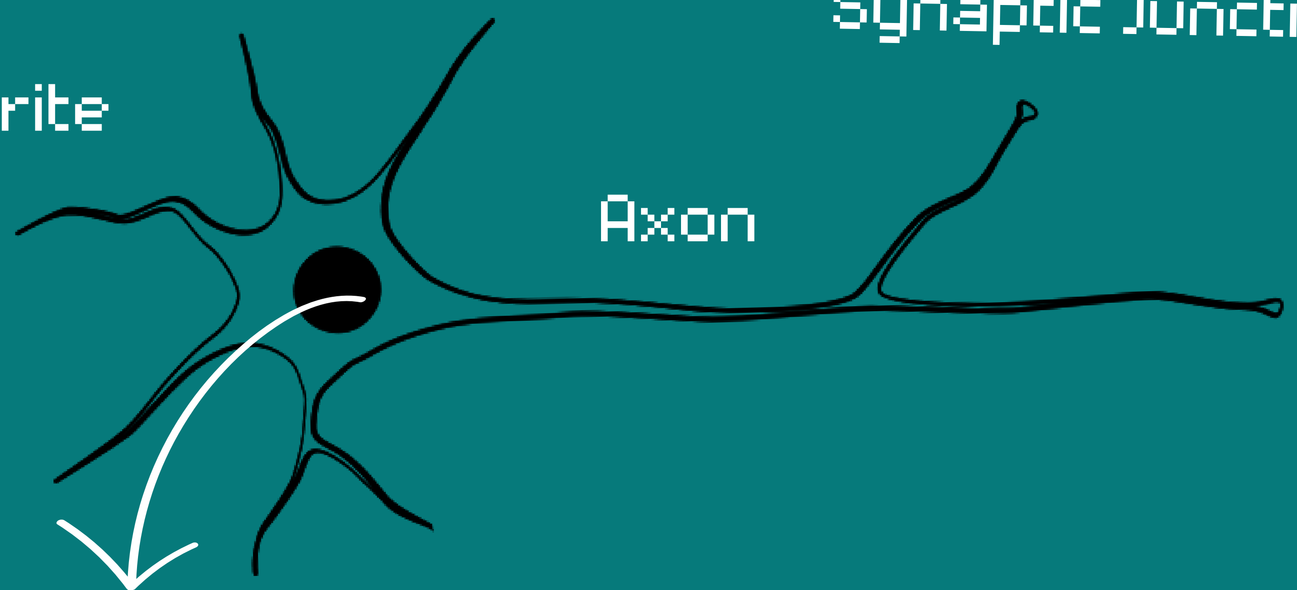
- Classification and Control.
- Interaction with enviroment of the problem.
- It chooses a current course of action based on previous encounters (exploitation) and fresh options (exploration) based on trial and error.

Neural Networks

Ok

al Networks • Neural Networks • Neural Networks • Neural Networks • Neural Networks • Neural Networks • Neural Netwo

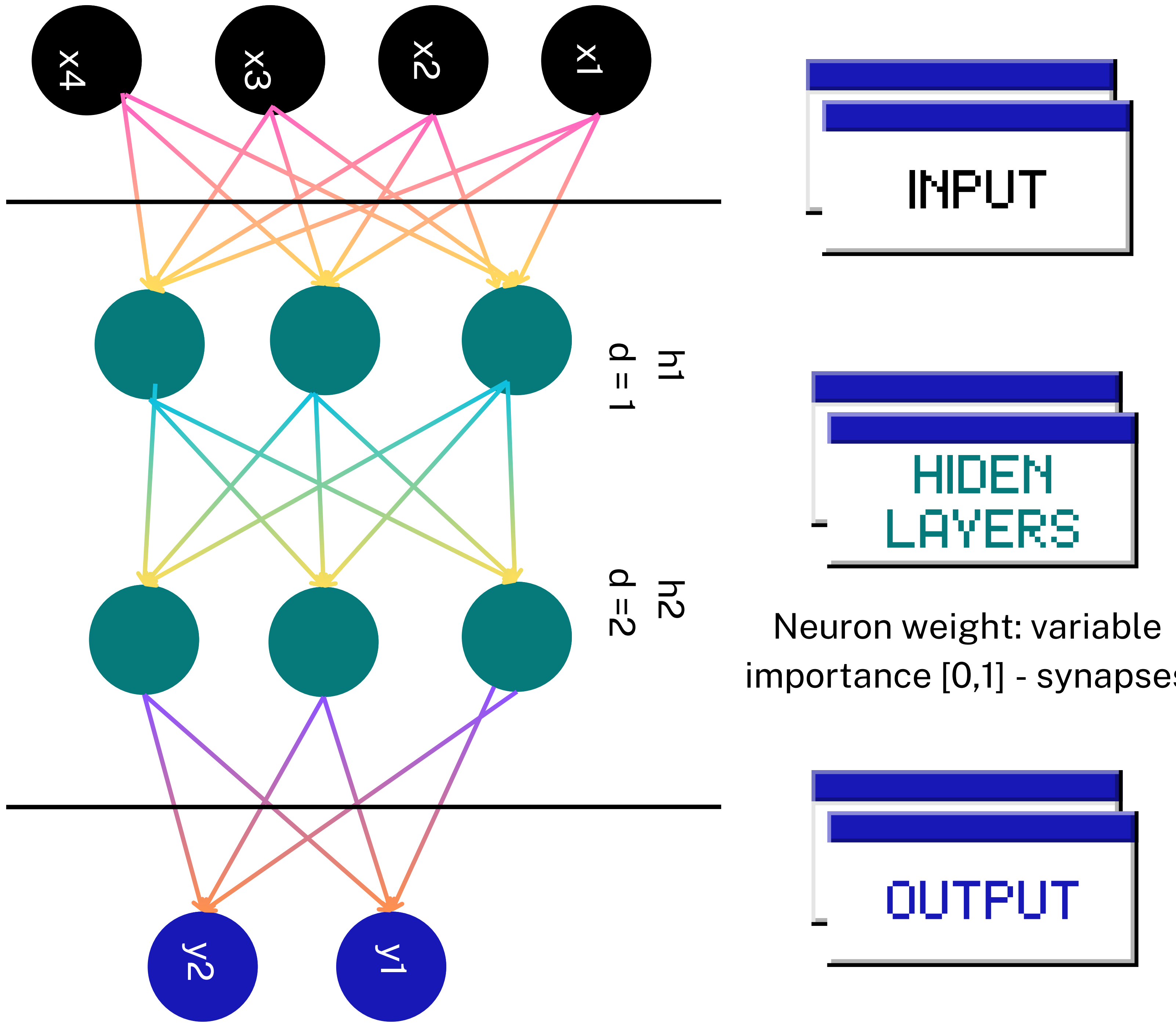
Biological Neural Networks



Definition

- Human brain: 10^12 neurons.
- Dentrites - Signal reception by the cell body [negatively charged] - travels down the axon to the synaptic junction where neurotransmitters travel through the synaptic cleft.

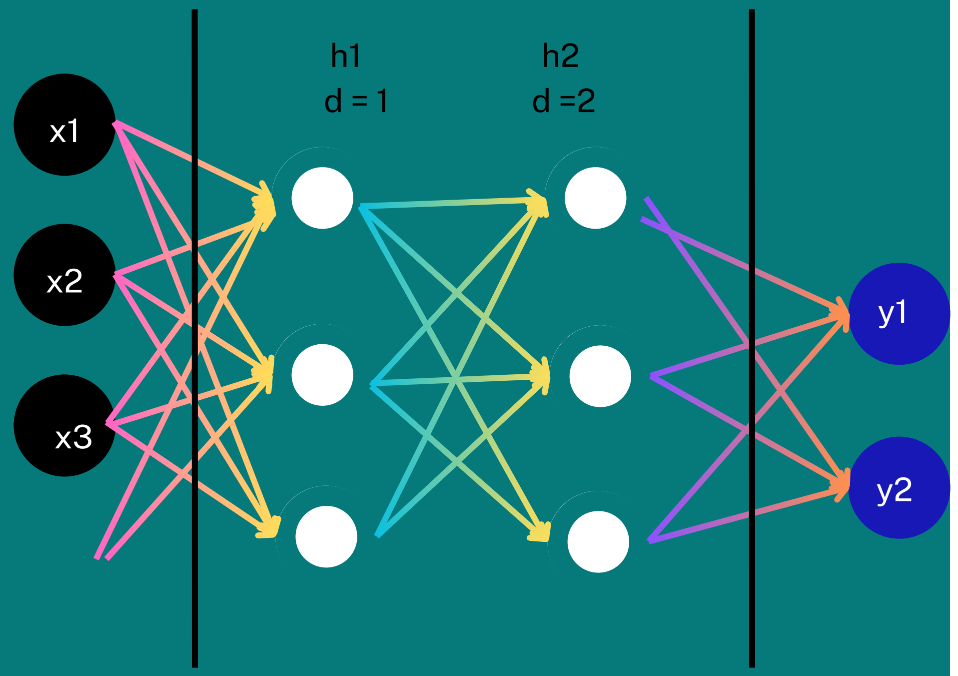
Artificial NN Overview



Neuron weight: variable importance [0,1] - synapses

- Feed-Forward network: single direction.
- Backpropagation: calculation and attribution of the error associated with each neuron, aiming for the best adjustment and fit.
- Hyperparamters: control the learning process [focusing on speed and quality].

Definition

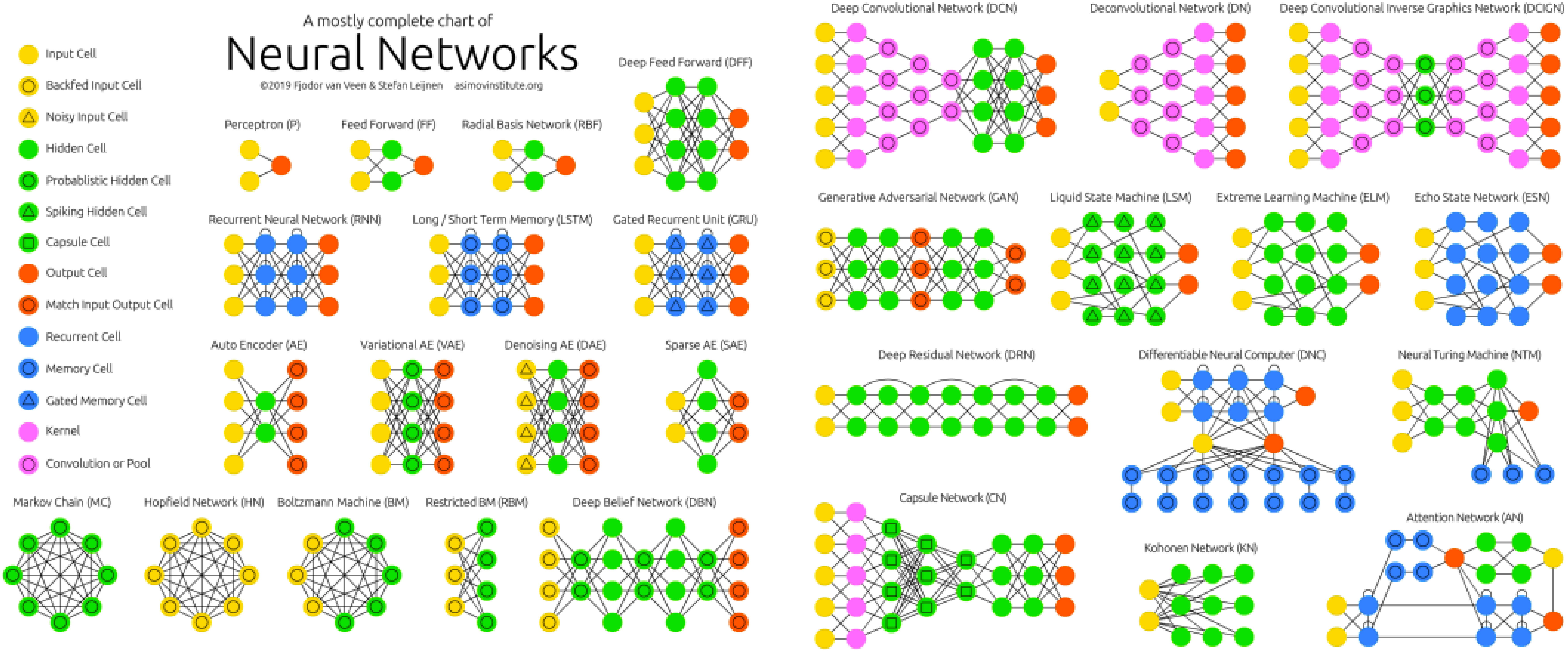


Mathematical description: output of a node generally

$$z = a_j^d = h(\sum_j W_{ij}^d \cdot a_i^{d-1} + b_j^d)$$

- Mimic: linear additivity for the input and strong non-linearity for the resulting output.
- Activation function [h] is used to apply transformations to the output vector for each element: [node]: nonlinear function like sigmoid, ReLU, and tanh.
- Weight tensor between node i in layder d-1 and node j in layder d: excitatory and inhibitory.
- Bias term. Neuron fires when the integrated input signal >> bias.

NN chart


A mostly complete chart of
Neural Networks
©2019 Fjodor van Veen & Stefan Leijnen asimovinstitute.org
Input Cell
Backfed Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Capsule Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Gated Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Differentiable Neural Computer (DNC)
Neural Turing Machine (NTM)
Capsule Network (CN)
Kohonen Network (KN)
Attention Network (AN)

ALICE

Current NN applications

PID	TPC detector	MFT-MCH track matching
Combine information from different detectors	Energy-loss calibration	Classification giving the score for a correct match

ALICE

Current NN applications

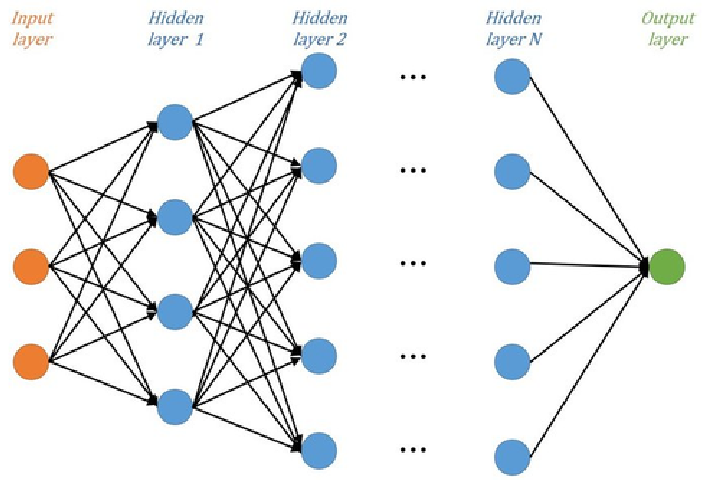
PID	TPC detector	MFT-MCH track matching
Combine information from different detectors	Energy-loss calibration	Classification giving the score for a correct match

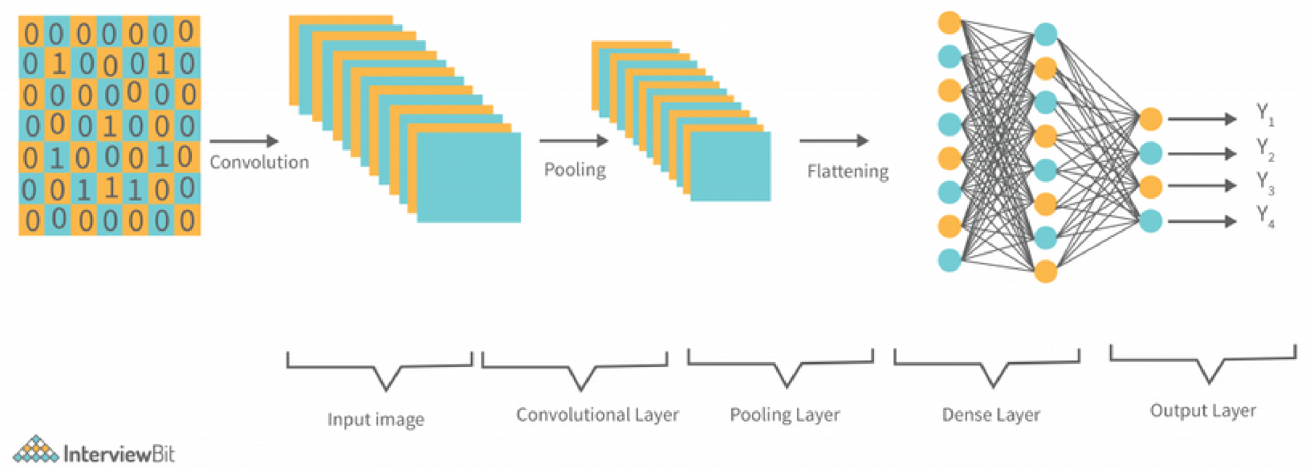
Base article

Identifying Ds mesons from radiative W decays at the LHC

- **Propost:** Development and validation of three neural networks to identify Ds mesons [particles with strangeness] through radioactive decays of W bosons. Refinement in RUN 3 data accuracy.

- **Focus:** Deep Neural Network & Convolutional Neural Network & Combined model





- **Challenge**: Differentiate between jets originating from radiatively produced mesons and background jets composed by quarks and gluons.

Reference

arXiv:2207.13587

Authors

E. Bakis, N. de Groot, N Vranjes

Identifying D mesons from radiative W decays at the LHC

Sample:

- Collisions pp at sqrt(13) TeV, analysing **160k events** of W in Ds + γ.
- **Jets**: Reconstruction via pFlow with anti-kt and clustering ΔR = 0.4 to satisfy pt > 20 GeV and |η| < 5.0. WIll be considered as Ds mesons if the angular distance of the true Ds particle is ΔR < 0.2.
- **Background**: Containing pp ⟶ gg and pp ⟶ qq process, both containing 80k events.
- Dataset was divised for Training [70%] and Testing [30%].

Variables

Input variables

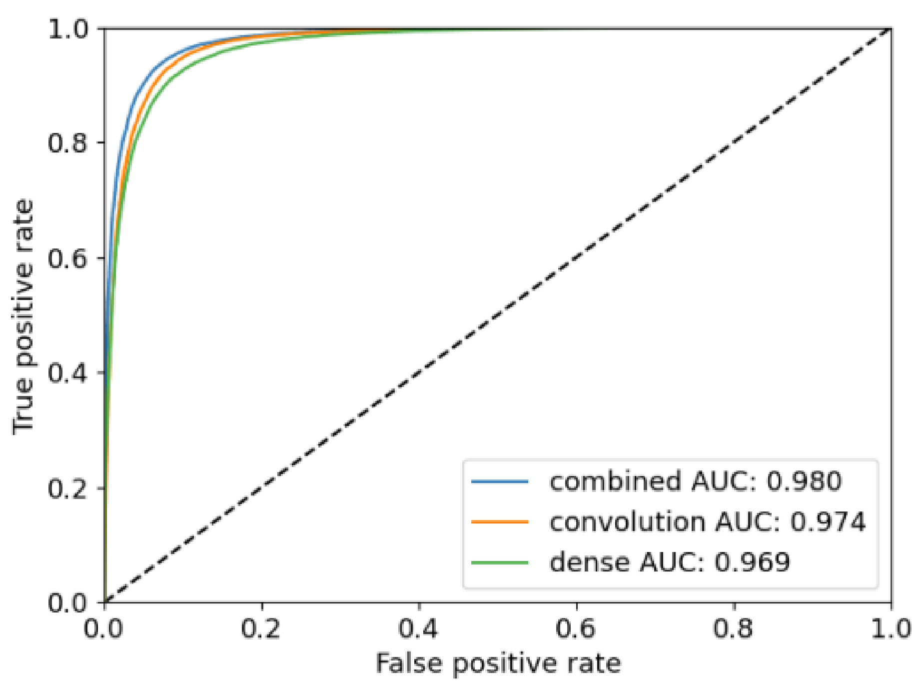
Name	Description
$\Delta\eta$	width of the jet in η
$\Delta\phi$	width of the jet in ϕ
m_{tr}	invariant mass of all charged tracks in the jet
m_j	invariant mass of all constituents of the jet
n_{ch}	charged particle multiplicity
n_0	neutral particle multiplicity
$\lvert Q \rvert$	absolute value of the total charge
$\lvert q_j \rvert$	jet charge (p_T weighted charge sum, $\Sigma_i\, q_i \cdot p_{Ti}^{1/2} / \Sigma_i\, p_{Ti}^{1/2}$)
b-tag	output of the b-tagging algorithm
R_{em}	average ΔR with respect to the jet axis weighted by electromagnetic energy
R_{track}	p_T weighted average ΔR for tracks
f_{em}	fraction of EM energy over total neutral energy of the jet
p_{core1}	ratio of sum p_T in a cone of ΔR <0.1 and the jet p_T
p_{core2}	ratio of sum p_T in a cone of ΔR <0.2 and the jet p_T
f_{core1}	ratio of sum ET in a cone of $\Delta R < 0.1$ and the jet total ET
f_{core2}	ratio of sum ET in a cone of $\Delta R < 0.2$ and the jet total ET
f_{core3}	ratio of sum ET in a cone of $\Delta R < 0.3$ and the jet total ET
$(p_T^D)^2$	λ_0^2
LHA	Les Houches Angularity; $\lambda_{0.5}^1$
Width	λ_1^1
Mass	λ_2^1
E_{had}/E_{em}	ratio of the hadronic versus electromagnetic energy deposited in the calorimeter
τ_0, τ_1, τ_2	N-Subjettiness

Table 1: DNN input variables.

Results

Identifying Ds mesons from radiative W decays at the LHC

ROC curves



igure 3: ROC curves for the different network type

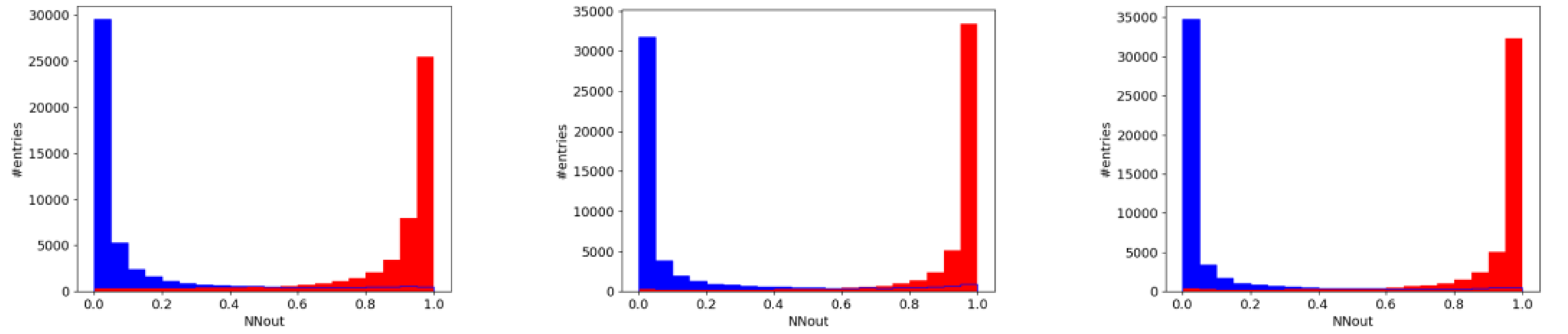
Overview

Network type	Test sample	Training sample	AuC
DNN	D_s vs mixed	D_s vs mixed	0.969
CNN	D_s vs mixed	D_s vs mixed	0.974
Combined	D_s vs mixed	D_s vs mixed	0.980
	D_s vs gluon	D_s vs mixed	0.994
	D_s vs quark	D_s vs mixed	0.964
	D_s vs gluon	D_s vs gluon	0.994
	D_s vs quark	D_s vs quark	0.965

ROC curve: Illustrates the performance of a binary classification system as the discrimination threshold varies, represented by the ratio RPV [True Positives / Total Positives] versus RPF [False Positives / Total Negatives].

Identifying Ds mesons from radiative W decays at the LHC

Output distributions

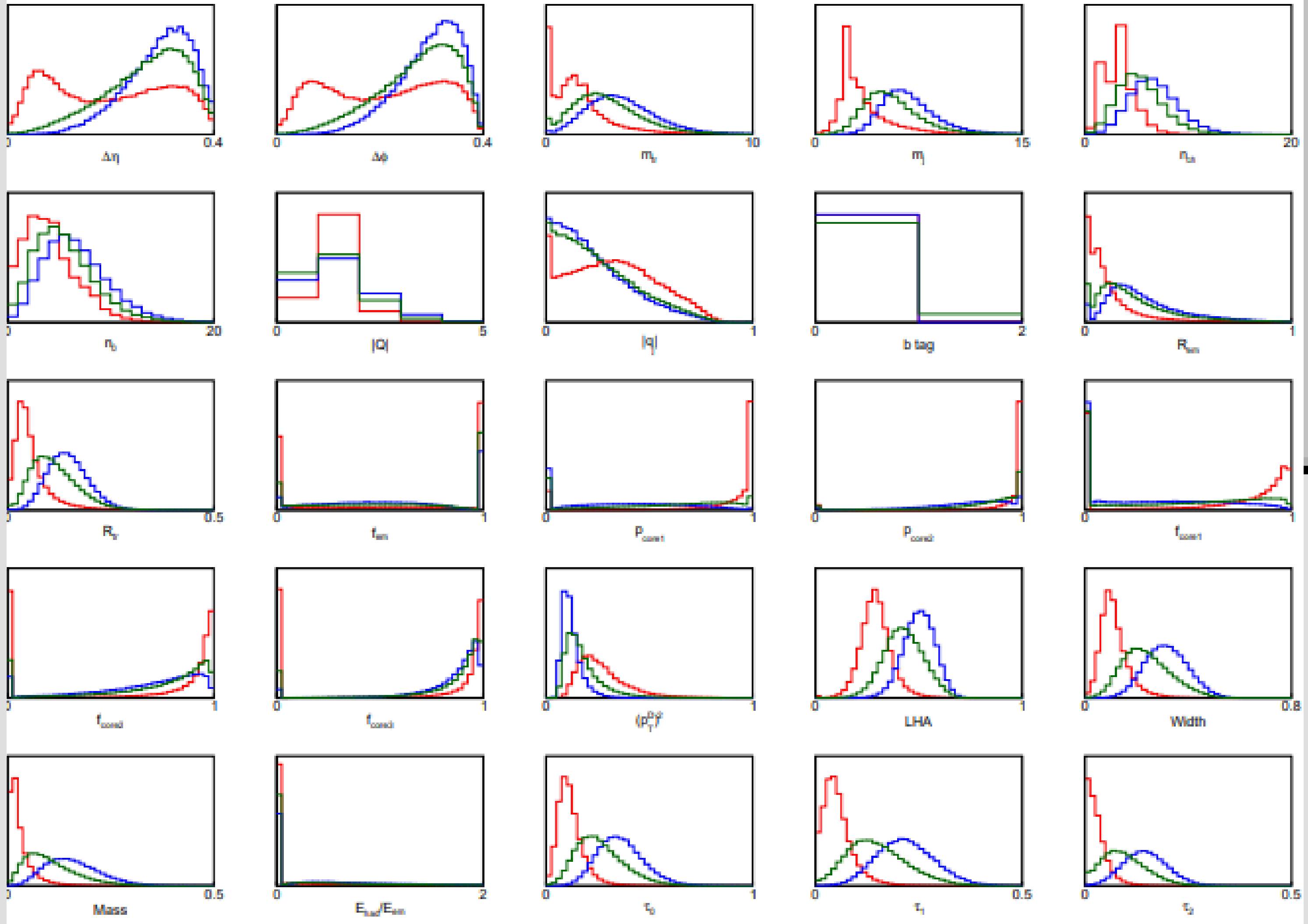


(a) DNN (b) CNN (c) Combined network

- **Signal efficiency: 67%** to combined model and **52%** to DNN and CNN models.
- **Background rejection: 25%** (Combined model), **21%** (DNN model) and 14% (CNN model).

Identifying Ds mesons from radiative W decays at the LHC

DNN variables distributions



Colours

- Signal presented (red),
- gg (blue) and qq (green) background

Propouse

Ok

Λnn bound-state investigation

Breef description

- Hot-topic on hipernuclear physics area.

- Main objetive: experimental carachterization of Λnn bound-state nature throught femtoscopy analysis and ML application.

- **ML application**: Characterization and validation of neural networks to identify hyper nuclei, using hypertritium as a test. [arXiv:2107.10627] [arXiv:1907.06906] [arXiv:2209.07360]

Λnn bound-state investigation

Status, challengs and perspectives

- Simulation via **Pythia** of the **decay of hypertritium** into **3He + pion** in **p-p collisions at 5.02 TeV** in order to obtain the behavior of the kinetic variables [transverse moment of the primary vortex of 3He and pion, DCA between particles, ŋ].

- **Challenge**: Adaptation of neural networks to the ALICE framework for particle identification and obtaining real p-p collision data from Run3.

- **Perspectives**: Future collaboration with WG3 to implement machine learning to identify hypernuclei in high-energy collisions.

Thank you!

maria.palhares@usp.br