Development of a flavor tagging algorithm by using deep learning

Summer camp on ILC, 2021

T. Onoe[A], T. Suehara[A], K. Kawagoe[A], T. Yoshioka[A]
(Kyushu Univ.[A])

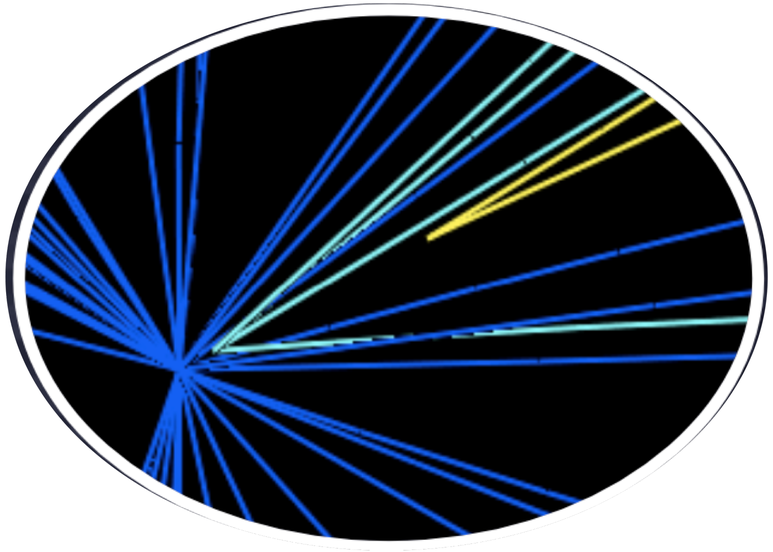
KYUSHU UNIVERSITY

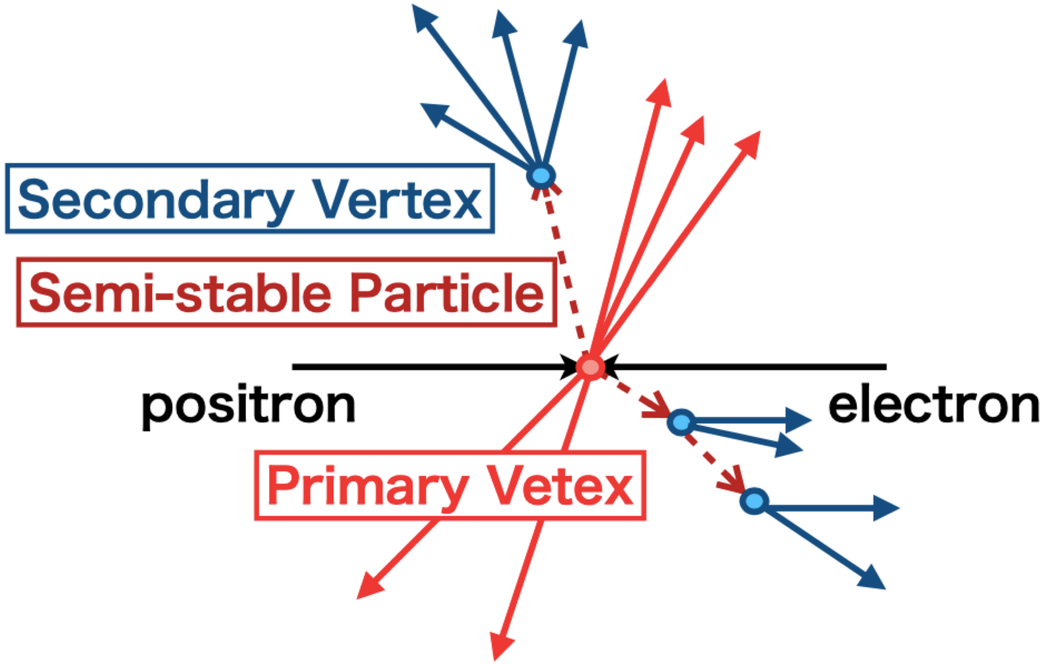
Contents

Introduction

Flavor Tagging

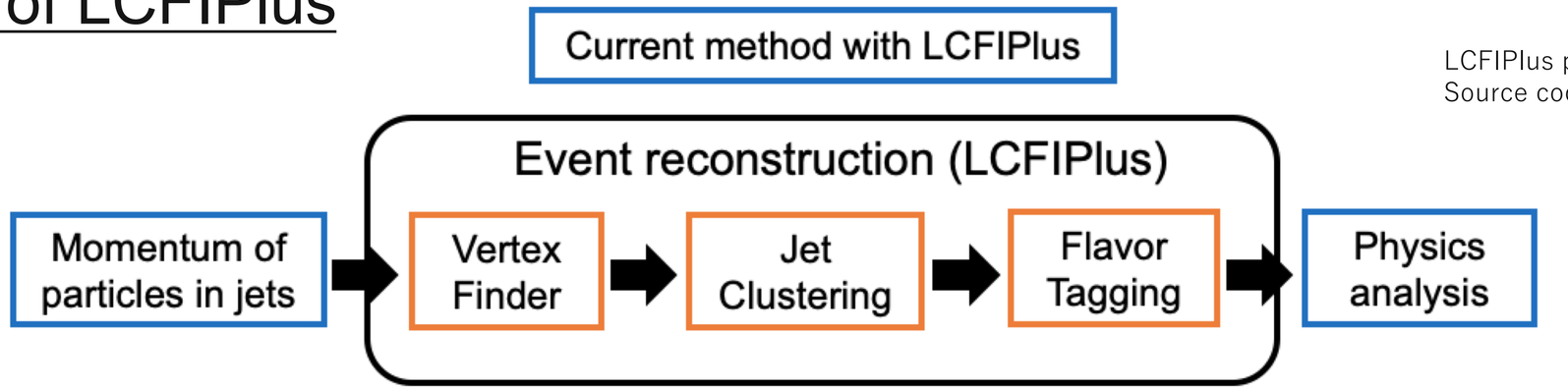
- Jets are bandles of hadrons originated by quarks and gluons.
- It is important to identify quarks/gluons (b/c/g/uds) of the origins of the jets.
  e.g. separation of H → bb, cc, gg
- Flavor tagging is the algorithm which classify the quarks.
- b/c hadrons can fly before their decay, because of their finite lifetimes
  → Vertex finding is important for the flavor tagging





Introduction

Structure of LCFIPlus



LCFIPlus paper: NIM A 808 (2016) 109-116
Source codes: https://github.com/lcfiplus/LCFIPlus

- Vertex Finder : Find primary and secondary vertices
- Jet Clustering : Reconstruct jets by clustering particles
- Flavor Tagging : Classify jets as b/c/others

・In the LCFIPlus, the flavor tagging algorithm is based on the Boosted Decision Trees, which is traditional ML.

Purpose on this study

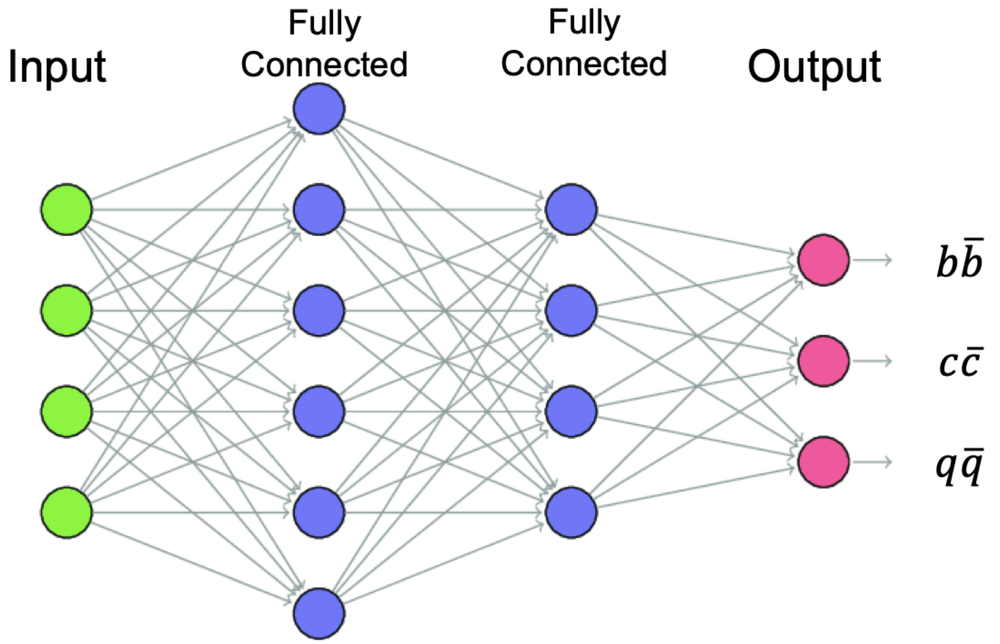
- ・Improve the performance of the flavor tagging by introducing deep-learning techniques.
- ・Combine vertex finding and flavor tagging in single DNN structure. (not done yet)

Introduction

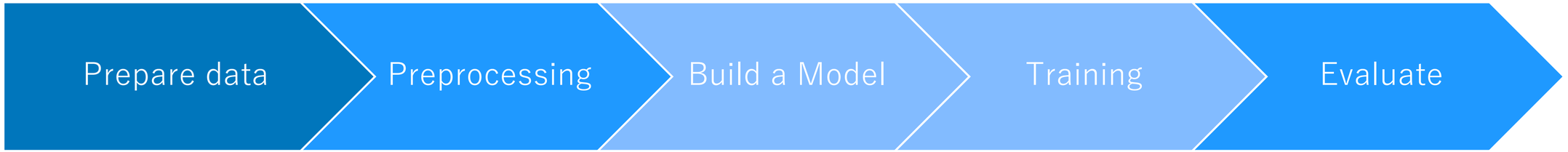
Deep Learning

- One of the Machine Learning technologies.
- Machines learn by using deep neural networks (DNNs) and extract features from big data.
- It is Supervised Learning

- There are various practical networks.



- To use deep learning, it is necessary to follow the flow.



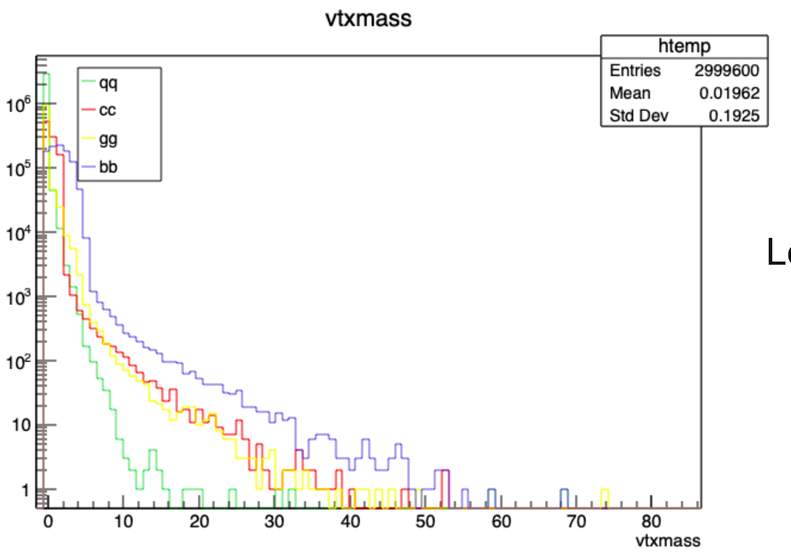
Data information

Prepare data

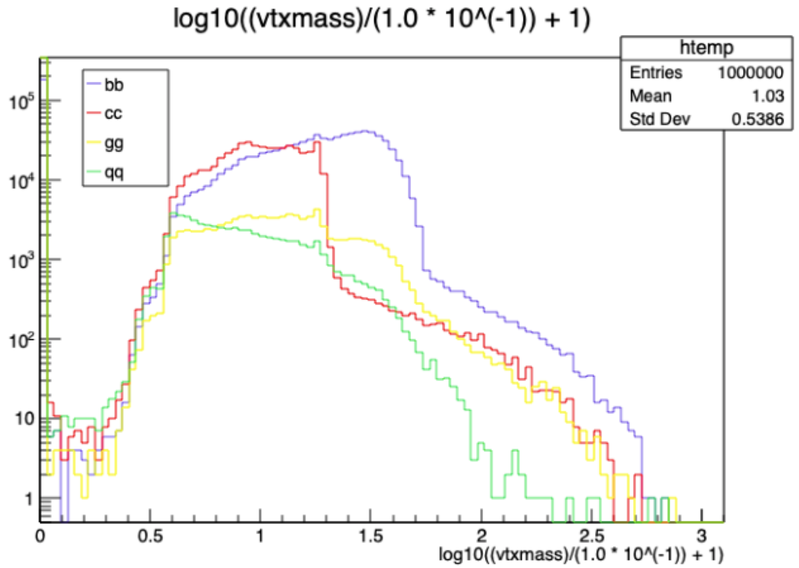
- 4 million events data from ILD simulation
- 124 variables from vertex finder
  (e.g. number of vertices, position/mass/probability/number-of-tracks of each vertex,
  displacement of tracks from the interaction point etc.)

Preprocessing

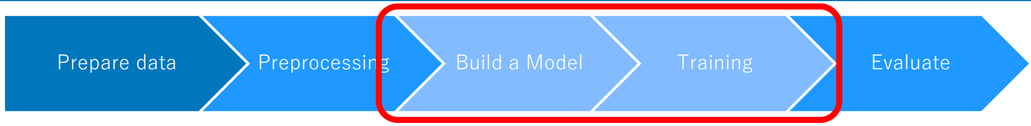
Distribution of the input variables should be flatten and scaled
before feeding to the input of the network by transformation of the variables.



Logarithmic transformation



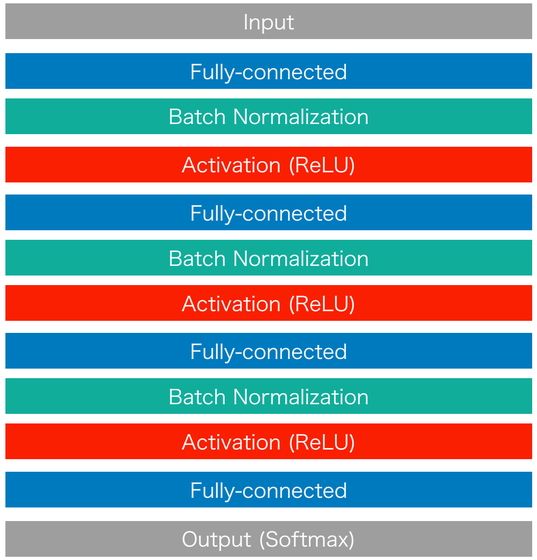
Model & Training



Build a Model

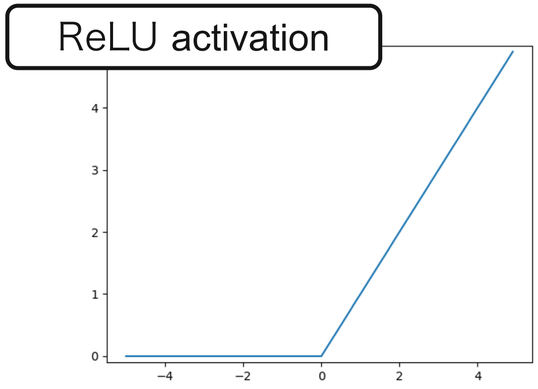
・Simplest DNN model as first trial and basis for comparison to modern networks.

- Input : 42 parameters
- 4 fully-connected layer with batch normalization and ReLU activation
- Output : 3 categories (b-, c-, uds- likeness)



Training

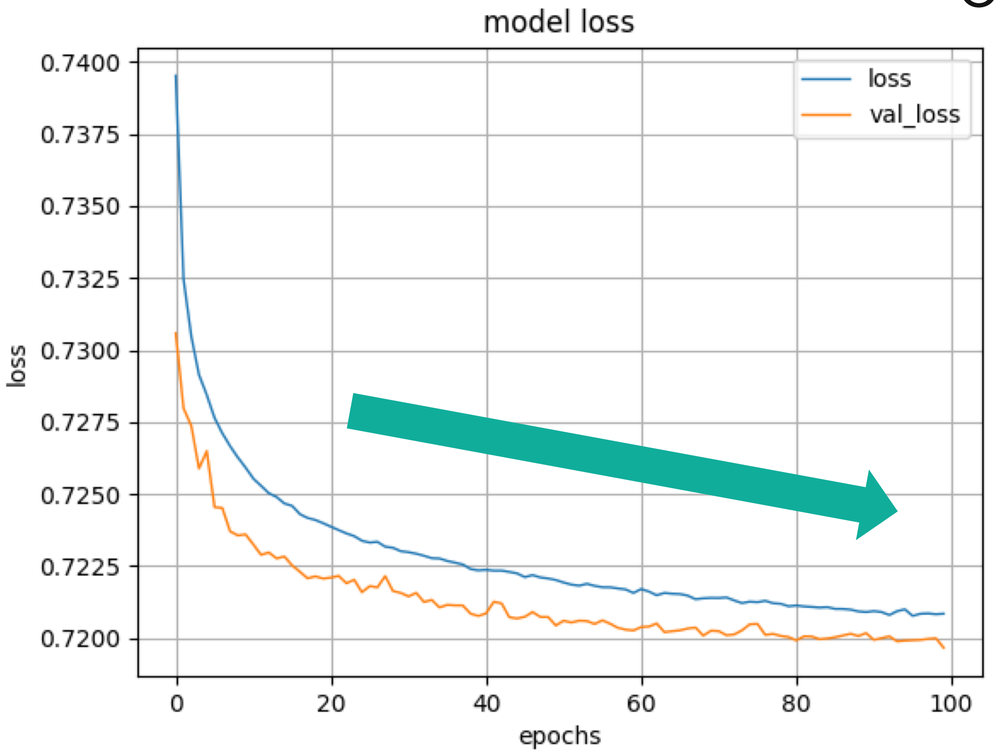
- Loss function : Categorical cross entropy
- Optimization : Adam (Learning rate : 0.01)
- The number of training : 100 epochs
- Batch size : 1024

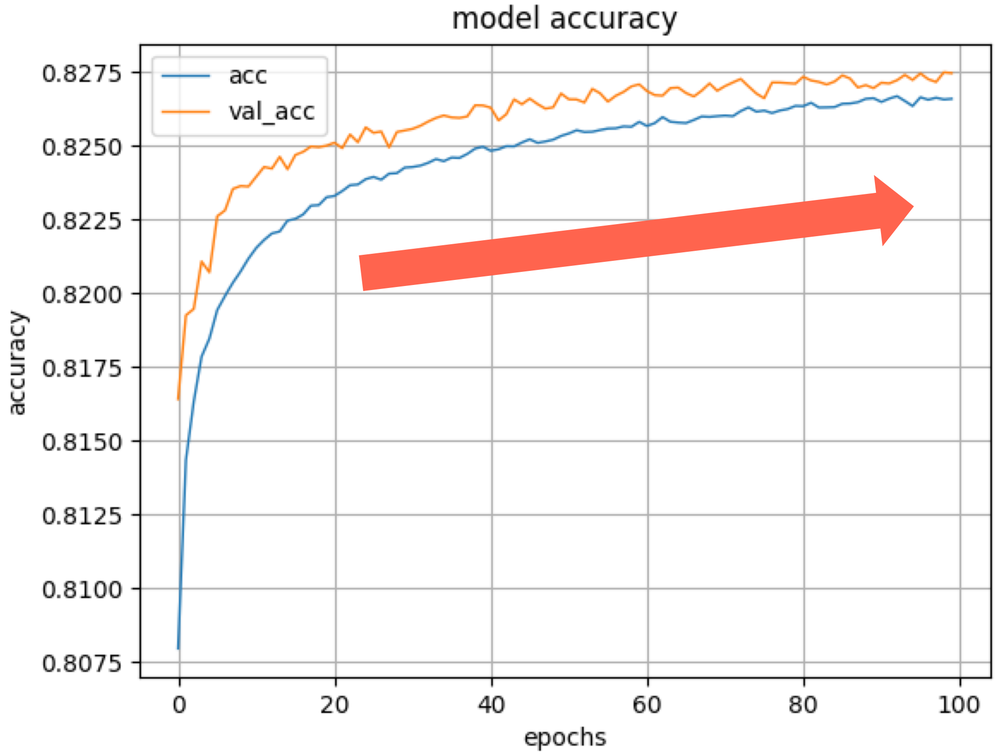


Result

Evaluate

Overall result



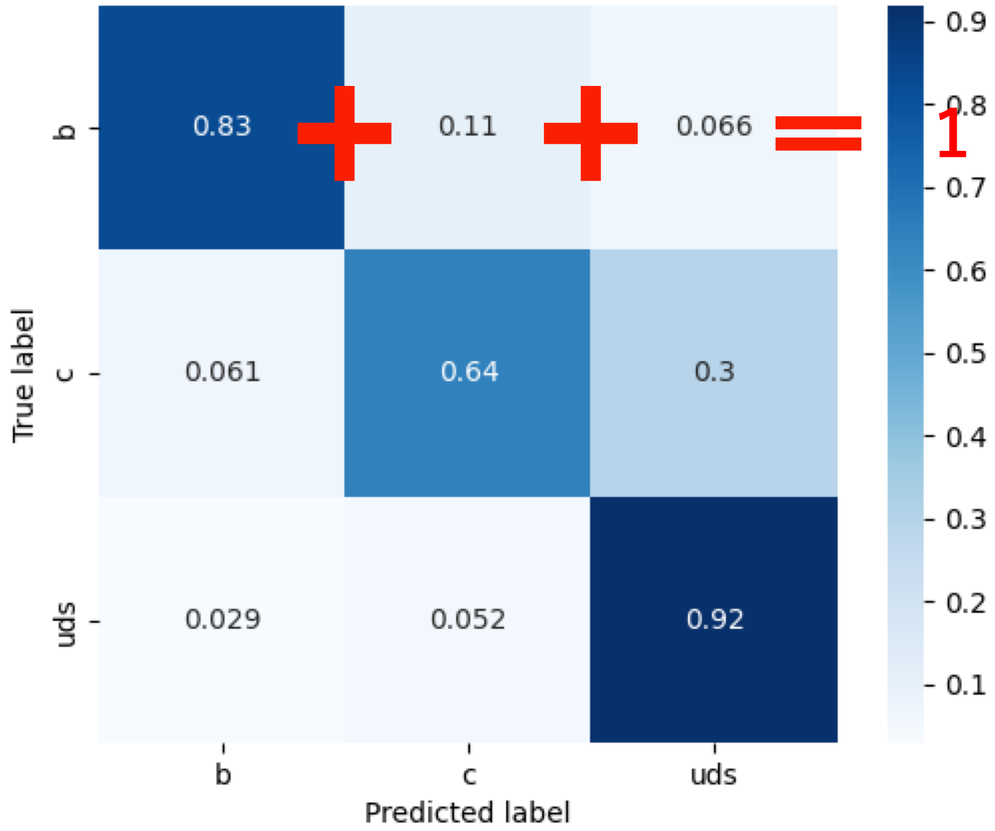


From plots, we can see that the model has comparable performance on both train and validation datasets (labeled test).

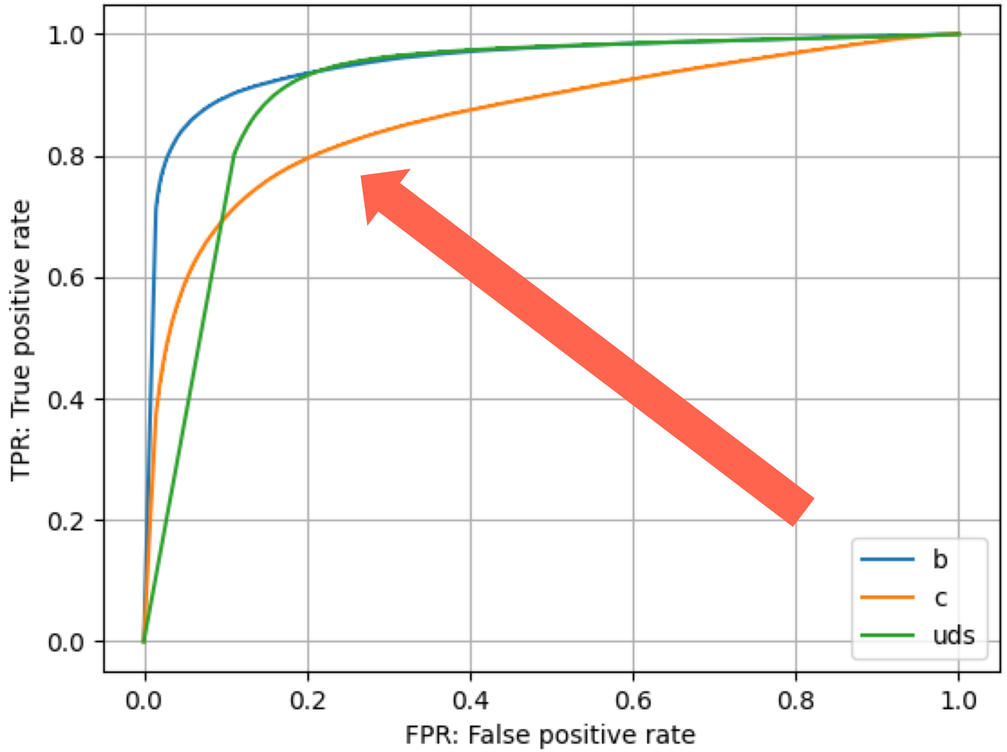
Result

Prepare data | Preprocessing | Build a Model | Training | Evaluate

Evaluate



uds events are classified well,
but c events are classified at 65% accuracy.

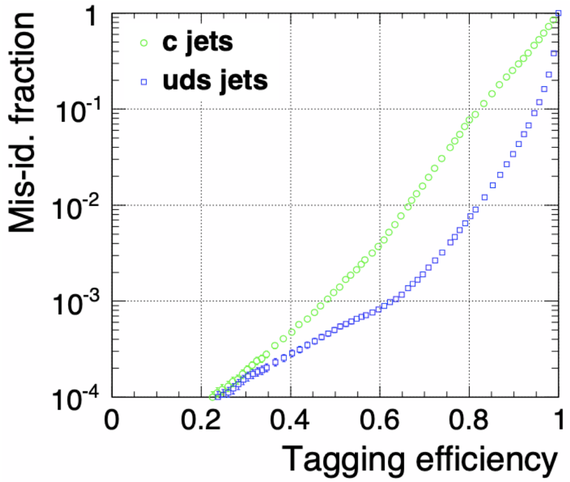


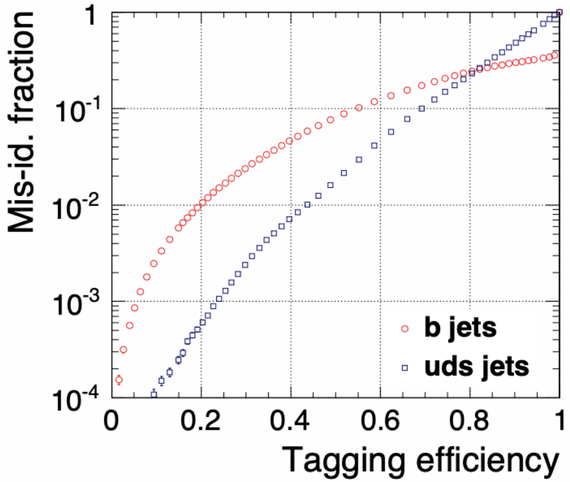
The more the curve is to the upper left of the graph, The better the classification is.

Result

Evaluate

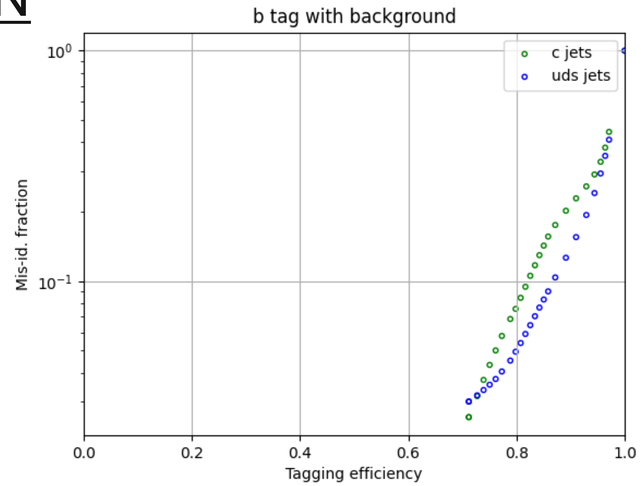
LCFIPlus

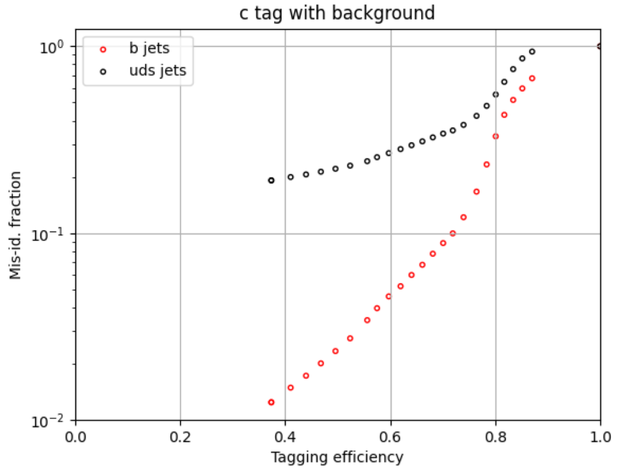




LCFIPlus is better identification now.

DNN





Summary & Next step

Summary

- We are in the process of constructing the network for flavor tagging.
- The accuracy of DNN was about 82%.

Next Step

- Detailed comparison with LCFIPlus results
- Hyperparameter tuning
- Trying more intelligent networks such as graph neural networks

Thank you for listening.