Final Projects

PHYS591000 2022.03.30

Outline

- Format and grading policy for final projects
- Project 1: Neutrino Oscillation
- Project 2: FlyCircuit (Brains of fruit flies)
- Project 3: Dijet Anomaly Detection

Final Project format

- (13 in total) 5 groups with 2 people + 1 group with 3
- 2 groups will ‘compete’ in the same projects → 3 projects
- Competition on Kaggle; Will specify the metric to compare the performances of your models.
- Find your own teammate(s)* and think of a cool name of your team!

*If you have trouble finding your teammate, let us know.

Final Project assignment

- We will assign projects on a ‘first come first served’ basis.
    - The team with 3 members should work on Project 3 (Dijet Anomaly with GAN).
- Please email the TA’s about members of your group, the name of your team, and (for groups of 2) projects of your choice (put them in order of preference) **no later than 5PM Wed. April 13.**

Final Project Grading Policy

- Final Project - 30%
  - Oral presentation: 15% (All have to speak)
  - Final project results: 10% (whether your model and/or fit results make sense)
  - Kaggle competition: 5% (1st or 2nd place)

- **Deadline for Kaggle submission: 5PM Tuesday May 31**

Final Project Oral Presentations

- There are two parts of oral presentations for each group:
    - On Week 16 (June 1): A 10-min ‘intro’ of your task and your model from each group before we reveal the Kaggle competition results LIVE in class.
    - The speaker will stay for another 5 mins to discuss the competition results.

Final Project Oral Presentations

- There are two parts of oral presentations for each group:
    - On Week 17 (June 8): A 15-min presentation on the whole story, including domain knowlege, data structure and visualization, model architecture, final results, and what you've learned from comparison to the other group on the same project.

Final Project Oral Presentations

- There are two parts of oral presentations for each group:
    - One member for each part.
    - For the group with 3 members, there will be an extra 15-min talk (on June 8) on whether your model can detect the ‘fake’ distribution generated by your competitor.

GAN is network with two parts: one generate (fake) data and the other tries to distinguish them from real data, and thus they can train each other.

Final Project Attendence Requirment

- Note: 10 points (10% of final grades) off for absence without permission for the following events

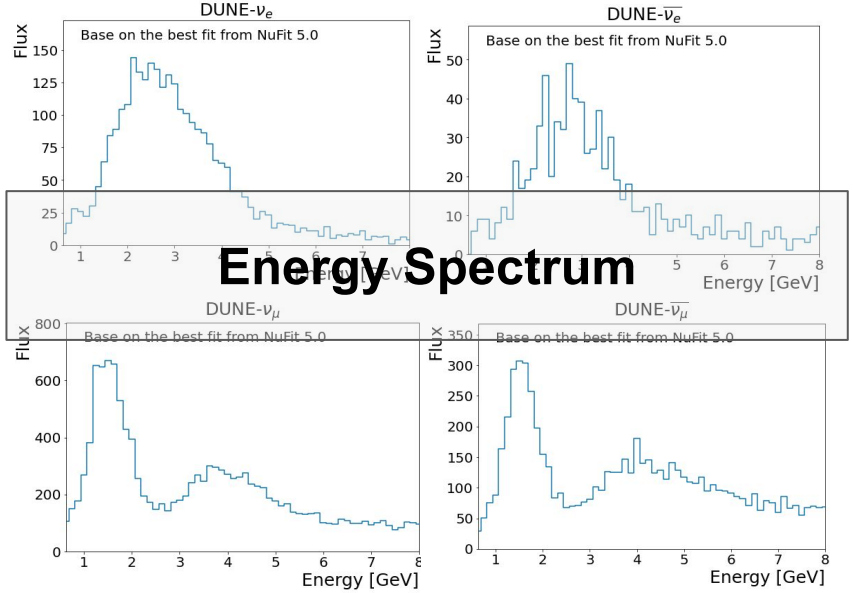
    - Guest lecture (Week 15 05/25; TBC)
    - Final project discussions (Week 16 06/01)
    - Final project oral presentations (Week 17 06/08)

Project 1: Neutrino Oscillation

- Estimate θ_{23} and δ_{cp} from spectrum
- Task: Regression of parameters
- Competition metric: RMSE

$$\begin{bmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{bmatrix} = \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & c_{23} & s_{23}\\ 0 & -s_{23} & c_{23}\end{bmatrix}\begin{bmatrix}c_{13} & 0 & s_{13}e^{-i\delta_{CP}}\\ 0 & 1 & 0\\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13}\end{bmatrix}\begin{bmatrix}c_{12} & s_{12} & 0\\ -s_{12} & c_{12} & 0\\ 0 & 0 & 1\end{bmatrix}}_{\text{PMNS Matrix}}\begin{bmatrix}\nu_1\\ \nu_2\\ \nu_3\end{bmatrix}$$

Flavor Eigenstates — Mass Eigenstates



ML Regressor

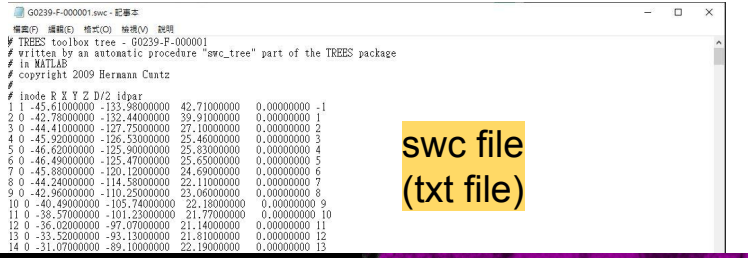
parameter estimate

Parameter	Normal ordering	Inverted ordering
θ_{12} [°]	$33.44^{+0.78}_{-0.75}$	$33.45^{+0.78}_{-0.75}$
θ_{13} [°]	$8.57^{+0.13}_{-0.12}$	$8.61^{+0.12}_{-0.12}$
θ_{23} [°]	$49.0^{+1.1}_{-1.4}$	$49.3^{+1.0}_{-1.2}$
Δm^2_{21} [$\times 10^{-5}$ eV^2]	$7.42^{+0.21}_{-0.2}$	$7.42^{+0.21}_{-0.2}$
Δm^2_{3l} [$\times 10^{-3}$ eV^2]	$+2.514^{+0.028}_{-0.027}$	$-2.497^{+0.021}_{-0.020}$
δ [°]	195^{+51}_{-25}	286^{+27}_{-32}

NuFit 5.0 (2019)

Project 2: FlyCircuit (Brain neurons of fruit flies)

- Cluster the local neuron and projection neuron.
- Task: train a classification model
- Competition metric: AUC



classification model:
DNN
CNN
…

Label Table
(0 for local;
1 for projection)

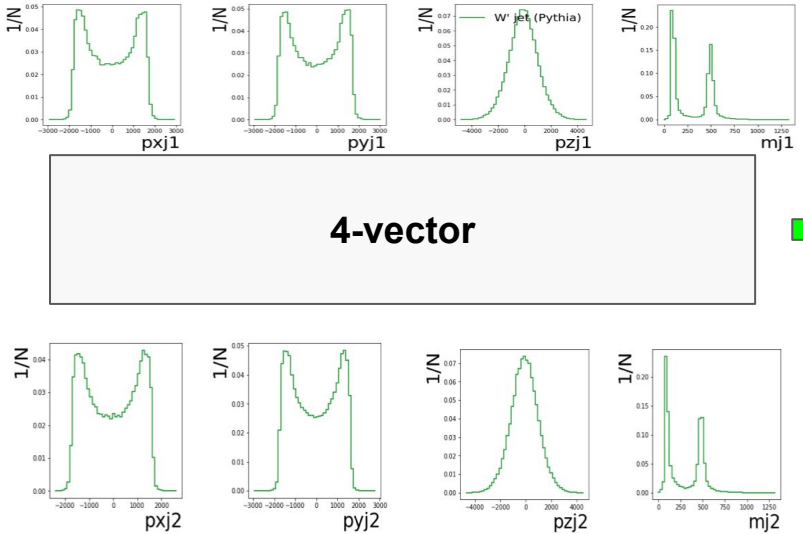
ID	label
5-HT1B-F-500007	0
5-HT1B-M-000000	1
5-HT1B-M-700003	1
5HT1A-F-100032	1
5HT1A-F-200016	1
5HT1A-F-300013	1
5HT1A-F-300026	1
5HT1A-F-800014	1
5HT1A-F-900001	1
…	

SPIN(2014)
NPIN(2021)

https://lhco2020.github.io/homepage/

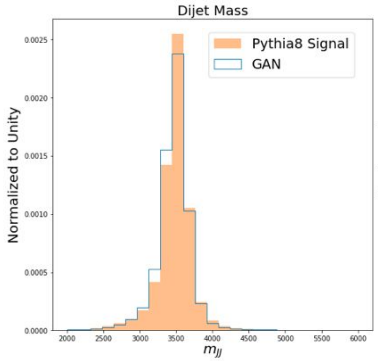
Project 3: Dijet Anomaly Detection

- Use generative model to build a fast simulator
- Task: train a generative model
- Competition metric: KL divergence



Generative Models

Density Estimators:
GAN



Check Kinematics