

Overview

- The physics
- Data preparation
- ML algorithms used
 - XGBoost
 - LightGBM
 - MLPClassifier
 - PyTorch
 - Tensorflow
 - Projection and clustering
- Conclusion and improvements

Heavy Neutral Leptons

- Explanation of Neutrino mass
 - "Seesaw mechanism"

$$\mathcal{L}_{K+D+M} = \sum_{\alpha=e,\mu,\tau} \sum_{I}^{N} \left(i\nu_{\alpha L}^{\dagger} \overline{\sigma}^{\mu} \partial_{\mu} \nu_{\alpha L} \right)$$

$$+ i N_{IR}^{\dagger} \sigma^{\mu} \partial_{\mu} N_{IR}$$
$$- \left(m_{\alpha I} \nu_{\alpha L}^{\dagger} N_{IR} - \frac{i M_I}{2} N_{IR}^{\dagger} \sigma_2 N_{IR}^* + h.c. \right) \right)$$

three generations of matter interactions / force carriers (fermions) (bosons) Ш Ш ≃1.28 GeV/c2 ≃173.1 GeV/c2 mass ≃2.2 MeV/c² 0 ≈124.97 GeV/c² charge 0 2/3 2/3 2/3 Η С t g u 1⁄2 1⁄2 spin 1⁄2 1 0 charm gluon higgs top up SCALAR BOSONS ≃4.7 MeV/c² ≈96 MeV/c² ≃4.18 GeV/c² **ഗ** 0 DUARK -1/3 $-\frac{1}{3}$ -1/3 0 S b a V 1/2 1/2 1/2 1 strange down bottom photon ≃0.511 MeV/c² ~105.66 MeV/c2 ≃1.7768 GeV/c² ≈91.19 GeV/c2 E BOSONS BOSONS -1 -1 -1 0 Ζ е Ш τ 1/2 1/2 1/2 Z boson electron muon tau S EPTON <0.17 MeV/c² <18.2 MeV/c² <1.0 eV/c² ≈80.39 GeV/c² 0 0 0 ±1 **D GAUG** VECTOR I Ve Vμ Vτ W 1/2 1/2 electron tau muon W boson neutrino neutrino neutrino

Standard Model of Elementary Particles

SHiP Detector





Decay or random muon noise?

Data preparation

- HNL decay simulation provided by Mads Hyttel, Edis Tireli and Oleg Ruchayskiy (1e5 data points for 21 different HNL masses)
- Muon noise simulation provided by us :) (1e6 data points)
 - Details can be provided, but are more physics than machine learning
- Variables: four-momenta of outgoing particles

["E_mu_plus", "E_mu_minus", "p_mu_plus_x", "p_mu_minus_x", "p_mu_plus_y", "p_mu_minus_y", "p_mu_plus_z", "p_mu_minus_z"]

• Target: "Truth" (1 or 0), thus accuracy is easy to measure

XGBoost

• Bayesian Optimization on N = 1e5 sample

The best hyperparameters are :

{'colsample_bytree': 0.7560154270067307, 'gamma': 8.81647216347955, 'max_depth': 9.0, 'min_child_weight': 9.0, 'reg_alpha': 101.0, 'reg_lambda': 0.6158 033470421029}



XGBoost

• SHAP Values



LightGBM

- RandomSearch with 10 iterations
- No correlation between the HNL mass and accuracy The best parameters are: {'max_depth': 8, 'n_estimators': 8, 'num_leaves': 67}

MLPClassifier

- RandomSearch with 10 iterations
- No correlation between the HNL mass and accuracy

The best parameters are: {'hidden_layer_sizes': 88, 'activation': 'tanh', 'solver': 'sgd', 'alpha': 0.05, 'learning _rate': 0.001}



Tensorflow

- Bayesian optimization
 - Learning rate = 0.01

Model: "sequential"

Layer (type)	Output Shape	Param #
input_layer (Dense)	(32, 12)	108
dense (Dense)	(32, 25)	325
dense_1 (Dense)	(32, 10)	260
output (Dense)	(32, 1)	11
Total params: 704		
Trainable params: 704		
Non-trainable params: 0		



Projection





UMAP



10

0

20



-i0

•

•

-30

-60

Background Noise

HNL Decay

-ż0







0

10

20



TSNE

Gaussian Mixture Modelling



Conclusion and Improvements

- In general very good results
 - Tree based > NN > Unsupervised
- Improvements
 - More sophisticated simulations ("It would take a small group 1 year to do" Oleg)
 - Less shortcuts on computational power



Thank you for your time!

Back-up slide: Momentum histograms (Reminder: The noise/signal ratio is 10)

х 50000 х 80000 У v Z 70000 40000 60000 30000 50000 40000 20000 30000 20000 10000 10000 0 -20 20 60 40 -400 -1000 -500 -1500500 0 1000 1500 le6 1.0 х У y 800000 z 0.8 600000 0.6 400000 0.4 200000 0.2 0 0.0 1000 2000 500 1000 1500 -2000 -10003000 0 -1500-1000-500 0

HNL Decay

