Machine learning-based particle identification of atmospheric neutrinos in JUNO

Jiaxi Liu[1], Fanrui Zeng[2], Xinhai He[1], Zhen Liu[1], Wuming Luo[1], Wing Yan Ma[2], Hongyue Duyang[2], and Teng Li[2]

[1]Institute of High Energy Physics, Chinese Academy of Science, Beijing 100049, China
[2]Shandong University, Binhai Road 72, Jimo district, Qingdao 266237, China

E-mail: liujiaxi@ihep.ac.cn

Abstract.

The Jiangmen Underground Neutrino Observatory (JUNO) is a next-generation large (20 kton) liquid-scintillator neutrino detector, which is designed to determine the neutrino mass ordering from its precise reactor neutrino spectrum measurement. Moreover, high-energy (GeV-level) atmospheric neutrino measurements could also improve its sensitivity to mass ordering via matter effects on oscillations, which depend on the capability to identify electron (anti-)neutrinos and muon (anti-)neutrinos against each other and against neutral current background, as well as to identify neutrinos against antineutrinos. However, this particle identification task is difficult in large unsegmented liquid scintillator detectors like JUNO. This paper presents a machine learning approach for the particle identification of atmospheric neutrinos in JUNO. In this method, several features relevant to event topology are extracted from PMT waveforms and used as inputs to the machine learning models. Moreover, the features from captured neutrons could also provide the capability of neutrinos versus anti-neutrinos identification. Two independent strategies are developed to utilize neutron information and to combine these two types of inputs information in different machine learning models. Preliminary results based on Monte Carlo simulations show the potential of this approach.

1 Introduction

Neutrinos have always been one of the hot topics in the field of particle physics, and the neutrino mass ordering (NMO) has been a particularly compelling problem for many physicists. For the three mass eigenstates of neutrinos, the sign of $|\Delta m_{32}^2|$ is still unknown, leaving us two possibilities: normal ordering (NO, $m_1 < m_2 < m_3$) or inverted ordering (IO, $m_3 < m_1 < m_2$).

The Jiangmen Underground Neutrino Observatory (JUNO) [1, 2] is a multipurpose neutrino experiment designed with the main physic goal of determining NMO, using a large homogeneous liquid scintillator (LS) detector to observe the oscillations of reactor electron antineutrino ($\bar{\nu}_e$) from nearby nuclear power plants. The design of JUNO's central detector is illustrated in figure 1.

Beside of reactor neutrinos, JUNO can obtain additional sensitivity to NMO by precisely measuring the matter effects in oscillations of atmospheric neutrinos. JUNO's final NMO sensitivity will be obtained through a joint analysis of both reactor and atmospheric neutrino oscillations. Figure 2 show the differences in oscillation probabilities for each atmospheric neutrino flavor, namely electron (anti)neutrinos ($\nu_e/\bar{\nu}_e$) and muon (anti-)neutrino ($\nu_\mu/\bar{\nu}_\mu$). From the figure, it can be seen that the oscillation probabilities of atmospheric neutrinos is closely related to their energy, directionality, and the neutrino type (referred to as "flavor"). Therefore the particle identification for atmospheric neutrinos is essential in

JUNO, which comprises of two major parts: The first part is distinguishing between $\nu_e/\bar{\nu}_e$ and $\nu_\mu/\bar{\nu}_\mu$ in Charged-Current (CC) interactions, as well as distinguishing them from background Neutral-Current (NC) interactions that do not provide any NMO sensitivity; The second part is the discrimination between neutrinos and antineutrinos in CC interactions, which is crucial because the matter effect on oscillation probability is opposite for neutrinos and antineutrinos.

In a recent work [3], we presented a new method of reconstructing directionality of atmospheric neutrinos in a large homogeneous LS detector. Our approach involved training several machine learning (ML) models to determine neutrino direction, utilising features extracted from the photomultiplier tube (PMT) waveforms to reflect the event topology. In this paper, we focus on the particle identification in a large homogeneous LS detector such as JUNO for atmospheric neutrinos, using a similar ML-based methodology. We leverage the input features derived from the PMT waveforms, but we also incorporate additional information from neutron capture signals to aid the flavor classification task.
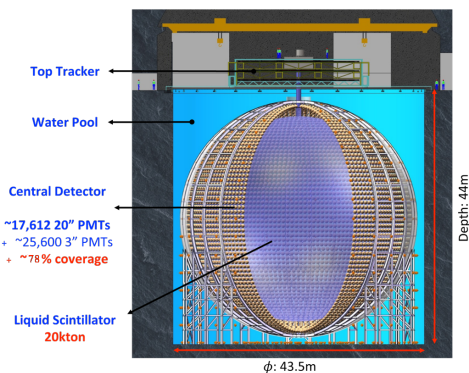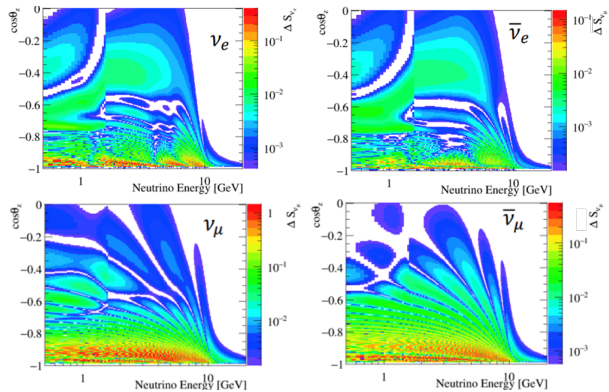

Figure 1: The JUNO detector.

Figure 2: Differences in oscillation probabilities between the two neutrino mass order (NO & IO).

2 Methodology

As detailed in the previous method on reconstructing directionality of atmospheric neutrinos [3], the light received by a PMT is the superposition of light from various points on particle tracks in the detector. The number of photo-electrons (PEs) seen by a PMT as a function of time is determined by the event topology. Therefore PMT waveforms would contain all the information about the characteristics of a given event such as energy, direction, particle type that are relevant for physics analyses. In practice, however, it can be very difficult to identify flavors directly from waveforms, given the complex relationship between the two and considering the large number of PMTs involved. Therefore instead of using the full waveforms, key features are extracted from waveforms to represent the entire shape of waveforms, and a sophisticated set of features are used for this study, including total charge (nPE), first hit time (FHT), the slope of the reconstructed waveform in the first 4 ns after first hit time, and so on.

For the 3-label identification, which is the identification of three types of atmospheric neutrino events: $\nu_\mu/\bar{\nu}_\mu$-CC, $\nu_e/\bar{\nu}_e$-CC, and NC, The key difference is from the charged lepton produced by the primary interaction. For a $\nu_\mu/\bar{\nu}_\mu$-CC event, the muon generated by the primary interaction rapidly deposits its energy and exhibits a long and straight track in the LS; For a $\nu_e/\bar{\nu}_e$-CC event, the event topology in LS resembles that of $\nu_\mu/\bar{\nu}_\mu$-CC, except that the electron from primary interaction instead creates a shower through processes like bremsstrahlung. As for an NC event, no charged lepton is produced by the primary interaction and the outgoing neutrino is invisible, leaving only the hadron part in the LS. The different event topologies for electron showers and muon tracks in the LS lead to different time distributions of PEs seen by PMTs at various angles, as shown in figure 3.

Therefore, for 3-label identification, the methodology is simply extracting the features from the primary trigger time window which reflect the topology of charged leptons produced by the primary interaction, and use them as inputs to the ML models. As for $\nu/\bar{\nu}$ identification, it is a rather challenging task in a large homogeneous LS detector without a magnetic field, given that LS detectors have neither Cerenkov rings nor tracks. However, additional information beyond features from primary trigger can help us in this work. Because of the V-A structure of $\nu/\bar{\nu}$ interactions, ν-CC events would have statistically larger hadronic energy fraction than $\bar{\nu}$-CC, which means that more secondary neutrons will be produced from hadrons for ν-CC events. On the other hand, as the neutrino energy increases, $\bar{\nu}$-CC
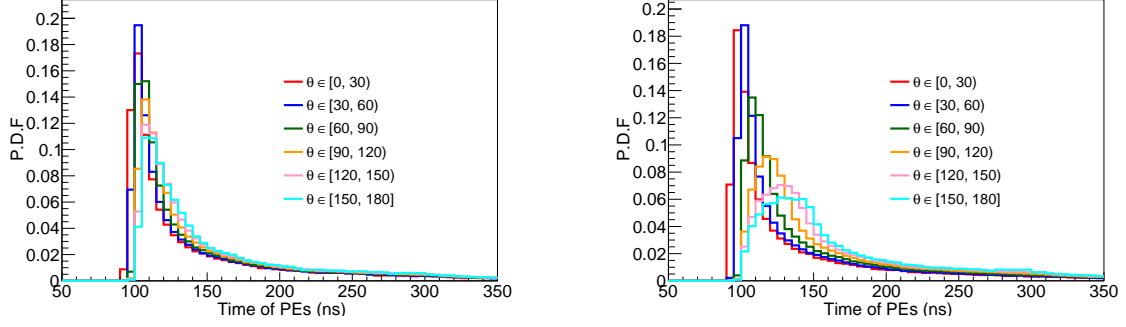
Figure 3: The time distributions of PEs for PMTs at different angles in an electron (left) and a muon (right) event, with the same kinetic energy and vertex position. This reflects the difference in event topology for the two type of events.

would tend to produce more neutrons in primary interactions than ν-CC. For the overall topology of final state neutrons, the distinction between ν and $\bar{\nu}$ arises from these two factors, which dominate in different neutrino energy ranges. Moreover, these final state neutrons can be selected with a high-efficiency cut on energy and time after primary trigger in a LS detector. As a result, the information from the captured neutrons, such as neutron multiplicity and their spatial distribution, is able to provide extra capability in the ν and $\bar{\nu}$ discrimination task.

Figure 4 shows the schematic workflow in this study, containing the two steps detailed earlier: the 3-label classification and the $\nu/\bar{\nu}$ discrimination. For each of the three identification tasks shown in the figure, an individual ML model is trained only for this task. This approach allows each model to focus on a specific classification task with different input features. In addition, we have independently developed two strategies in utilising all the information described above to perform the flavor-identification tasks.
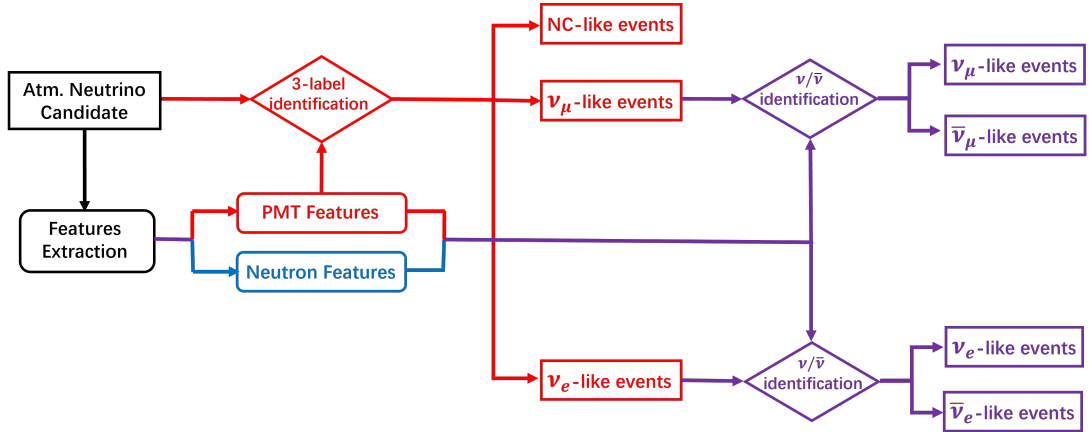

Figure 4: The schematic workflow of atmospheric neutrino particle identification.

2.1 Strategy 1

In this strategy, we are selecting neutron candidates by applying a time window ranging from 10 μs to 1 ms after the prompt signal , with an energy range of 2 MeV to 2.7 MeV. The vertices of these neutron capture candidates are reconstructed using the method from Ref. [4].

The overall architecture of the model for strategy 1 is shown in figure 5. The input PMT features extracted from the waveforms are transformed into a point cloud, with each feature sequentially added after the 3-dimensional (x,y,z) coordinates of the PMTs. The input is fed into a point cloud-based model PointNet++ [5]. For the captured neutrons, the only useful information to reflect event topology is their reconstructed vertices, therefore the point cloud for neutrons only contains the absolute positions in the

detector. By adopting this data format, the loss of information is minimized, ensuring that not only the quantity of neutrons but also their spatial distribution is retained.

Because the point cloud of neutron is much sparser than the one of PMTs, PointNet++ is too complex for neutrons and does not perform well. Furthermore, if both point clouds are inputted into the model together, the neutron information will be ignored by the model due to the overwhelming difference in quantity. Therefore a separate DGCNN-based model [6] with more capability to handle sparse point clouds is applied, in order to extract features from the reconstructed neutron information. The two parts of inputs are fed into the PointNet++-based model and the DGCNN-based model, and finally concatenated at the fully-connected Layer. The outline of the strategy is shown on figure 6.
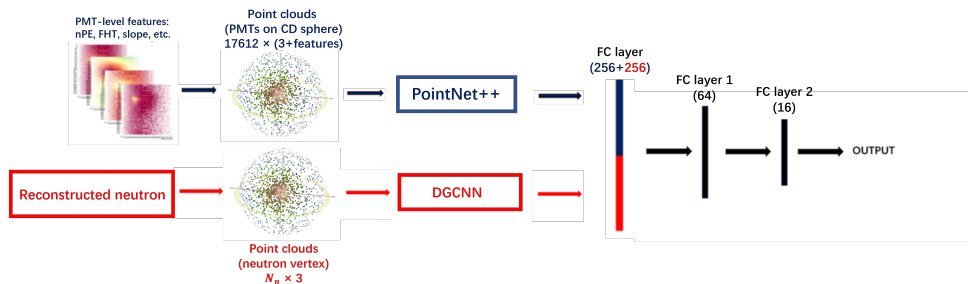

Figure 5: Schematic diagram of PointNet++/DGCNN model architecture used for strategy 1.
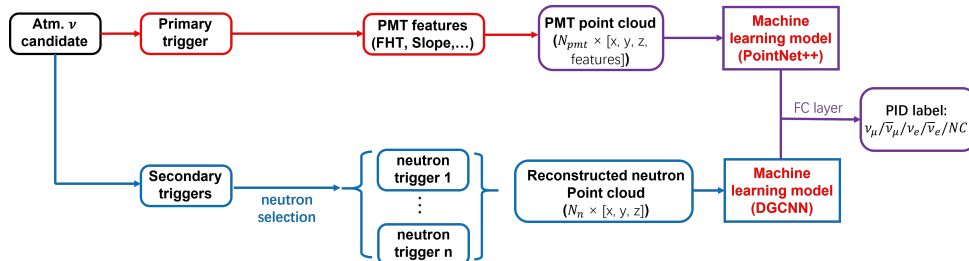

Figure 6: Outline of strategy 1.

2.2 Strategy 2

In this strategy, the multiple neutron candidate triggers are merged into one, and the features for neutron, namely nPE and FHT, are extracted from PMT waveforms just like the primary trigger. This treatment for neutrons ensures the uniformity of the input format, and enables the machine learning model to process the input data more efficiently. The architecture of the model for this strategy is illustrated in figure 7.

The DeepSphere model [7] is used in this strategy, which processes the PMT features by treating them as spherical images, since the PMTs are originally distributed on the surface of the detector sphere. This approach maintains rotation covariance and avoids distortions caused by projecting the data onto a planar surface. The outline of the strategy is shown on figure 8.

3 Performance

Two neutrino samples with different neutrino spectra are used in this work: one with a flat neutrino energy spectrum to mitigate the influence of neutrino energy dependence to the model, and the other one has flux calculated from Honda *et al.* [8] which could reflect the real atmospheric neutrino distribution. For the purpose of training and testing the ML models, a sub-sample is selected from the flat sample for testing; on the other hand, the full Honda sample is not used for training and is used solely for testing purposes.
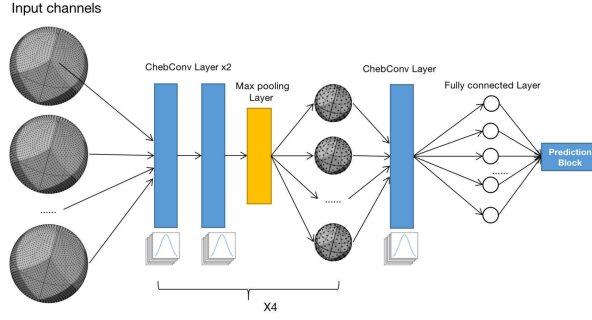
Figure 7: Schematic diagram of DeepSphere model architecture used for strategy 2.
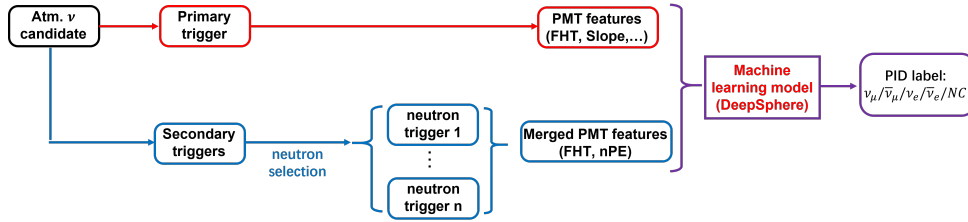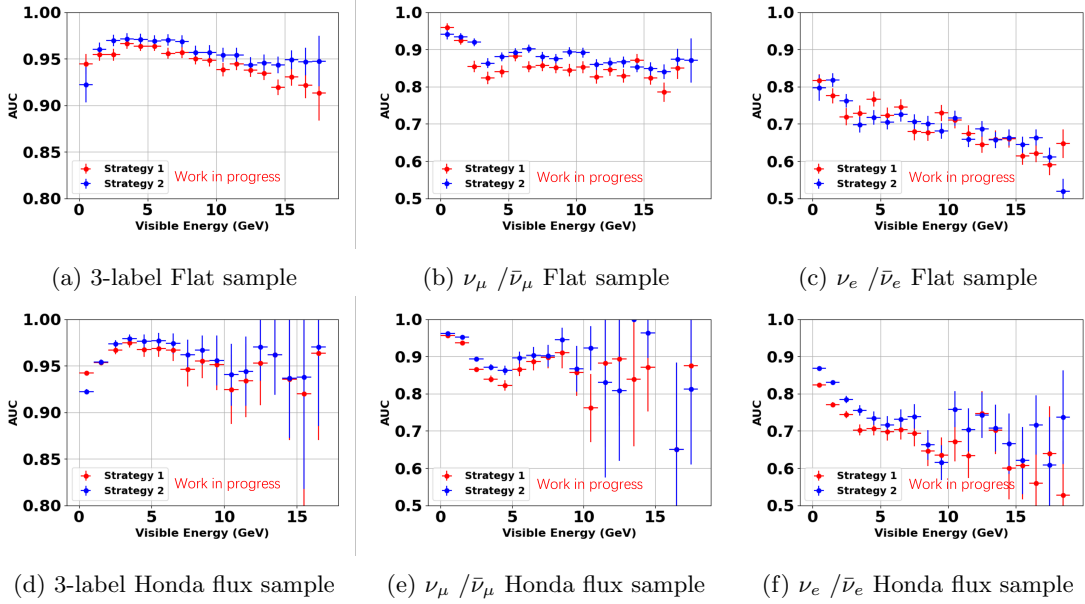

Figure 8: Outline of strategy 2.

(a) 3-label Flat sample

(b) ν_μ /$\bar{\nu}_\mu$ Flat sample

(c) ν_e /$\bar{\nu}_e$ Flat sample

(d) 3-label Honda flux sample

(e) ν_μ /$\bar{\nu}_\mu$ Honda flux sample

(f) ν_e /$\bar{\nu}_e$ Honda flux sample

Figure 9: Comparison of AUC scores as a function of visible energy for 3-label classification (9a and 9d), ν_μ /$\bar{\nu}_\mu$ (9b and 9e) and ν_e /$\bar{\nu}_e$ (9c and 9f). The large error bars at higher energies of the Honda flux sample are due to the small statistics in this region.

The actual output of the ML models for classification task is not just the predicted label for each event, but a vector of scores represents the possibility for each predicted label. By adjusting the thresholds on those scores, The efficiency and purity for each label can be optimised therefore could not fully reflect the model's capability of classification. Therefore, the Area Under the Receiver operating characteristic (ROC) Curve (AUC) is used to evaluate the classification performances, as it is independent of the score cut selection and is not affected by class imbalances in the data as well.

Figure 9 shows a comparison of the ROC AUC scores for the two strategies for all atmospheric neutrino

events with the energy range of 0.5 to 20 GeV for both the flat sample and Honda flux sample. The results indicate that the two strategies are consistent with each other. Figure 10 shows the true composition for each of the four CC categories as functions of L/E in Honda flux sample, which will be used as the inputs for NMO sensitivity study. In this upcoming NMO study, the efficiency and purity of each label will be tuned to obtain the best sensitivity.
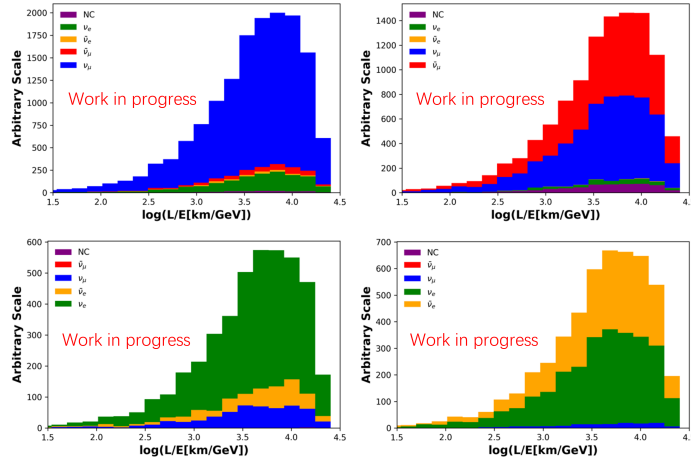

Figure 10: L/E plots for each CC category.

4   Summary

In this paper, we presented a general machine learning approach of atmospheric neutrino particle identification, using the features extracted from the PMT waveforms in the primary trigger as well as the information of the captured neutron candidates. Two individual particle-identification strategies with different types of machine learning models were developed to cross validate the methods. The two strategies show consistent performances, and will be both applied in JUNO's future NMO sensitivity analyses for cross check. Preliminary results using Monte Carlo simulations demonstrate the great potential of this approach, as well as JUNO's capability for atmospheric neutrino oscillation measurements.

References

[1] Fengpeng An et al. Neutrino physics with JUNO. *J. Phys. G*, 43(3):030401, feb 2016.

[2] Angel Abusleme et al. JUNO physics and detector. *Progress in Particle and Nuclear Physics*, 123:103927, 2022.

[3] Zekun Yang et al. First attempt of directionality reconstruction for atmospheric neutrinos in a large homogeneous liquid scintillator detector. *Phys. Rev. D*, 109(5):052005, 2024.

[4] Gui-hong Huang, Wei Jiang, Liang-jian Wen, Yi-fang Wang, and Wu-Ming Luo. Data-driven simultaneous vertex and energy reconstruction for large liquid scintillator detectors. *Nucl. Sci. Tech.*, 34(6):83, 2023.

[5] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[6] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5):1–12, 2019.

[7] N. Perraudin, M. Defferrard, T. Kacprzak, and R. Sgier. Deepsphere: Efficient spherical convolutional neural network with healpix sampling for cosmological applications. *Astronomy and Computing*, 27:130–146, 2019.

[8] M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa. Atmospheric neutrino flux calculation using the nrlmsise-00 atmospheric model. *Phys. Rev. D*, 92:023004, Jul 2015.