

Characterisation of Neutrino-Induced Particles at DUNE

2107794

Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

Particles from simulated Deep Underground Neutrino Experiment (DUNE) data were classified as ‘tracks’ and ‘showers’, and events as CC ν_μ , CC ν_e , and NC ν_x . This was to support DUNE’s use of neutrino oscillations to investigate Charge-Parity (CP) symmetry, aiming to explain the universe’s matter-antimatter asymmetry and accurately measure δ_{CP} . A projective likelihood estimator and Boosted Decision Tree (BDT) managed particle classification with a BDT accuracy of 97%. A Convolutional Neural Network (CNN) was employed for image-based event classification, achieving 85% accuracy, excelling in muon neutrino event identification at 93% accuracy.

I. Introduction

The Standard Model predicts that the Big Bang would produce equal amounts of matter and antimatter [1]. However, the universe shows a clear matter dominance, indicating early universe processes favoured matter over antimatter, a detail not fully explained by existing Standard Model physics. Investigating Charge-Parity violation (CPV), where physical laws differ between matter and antimatter, is key to understanding this imbalance. Neutrino oscillations (neutrinos changing flavours as they propagate space) offer a way to study CPV by comparing neutrino

and antineutrino behaviours [2, 3, 4].

A Liquid Argon Time Projection Chamber (LArTPC) is a kind of neutrino detector that can reconstruct particle paths from neutrino interactions with high precision and will be crucial in studying CPV [5]. The Deep Underground Neutrino Experiment (DUNE) will use these detectors on an unprecedented scale, enabling precise reconstruction of neutrino interactions with image-like clarity. With the most intense neutrino beam, a deep underground site, and large detectors, DUNE is set to significantly advance understanding of CPV and neutrinos [6].

Machine learning is key to automating

particle and event characterisation in these experiments, handling voluminous, high-dimensional data that would otherwise be unmanageable [6]. This project developed methods including a projective likelihood estimator, Boosted Decision Trees and a Convolutional Neural Network to characterise neutrino-induced particles at DUNE.

II. Theory

A. Neutrinos

Neutrinos are fundamental, electrically neutral particles that interact via the weak nuclear force and ignore electromagnetic and strong forces. They exist in three types, or flavours: electron (ν_e), muon (ν_μ), and tau (ν_τ) neutrinos, each associated with a specific lepton. Neutrinos' antimatter counterparts are called anti-neutrinos [7].

Neutrinos are detected indirectly through their weak force interactions with other particles, mediated by three bosons: the W^+ , the W^- and the Z^0 bosons. The W^\pm bosons are associated with charged-current (CC) interactions involving an exchange of charged particles and changes in flavour. The Z^0 bosons are associated with neutral-current (NC) interactions involving no change in flavour and no change in the electric charge of the participating particles [7, 8].

CC interactions produce a leading lepton indicating the neutrino's flavour, alongside other particles from neutrino-nucleus interactions. The leading lepton must be identified to infer the neutrino flavour. NC interactions generate particles without a leading lepton, leaving the neutrino's flavour unidentified [7].

Neutrinos can change flavours as they propagate through space, a phenomenon known as *neutrino oscillation* [9]. The three flavours are termed '*weak eigenstates*.' Additionally, neutrinos have three distinct and definite *mass eigenstates*, which follow a specific time evolution. Weak eigenstates are coherent linear combinations (quantum superpositions) of the mass eigenstates. For example,

$$|\nu_e\rangle = u_{e1} |\nu_1\rangle + u_{e2} |\nu_2\rangle + u_{e3} |\nu_3\rangle, \quad (1)$$

where $|\nu_e\rangle$ is the weak eigenstate produced alongside an electron in a CC interaction, $|\nu_i\rangle$ are the mass eigenstates and u_{ei} are coefficients that describe the probability amplitudes of each mass eigenstate in the ν_e state [7, 8]. Neutrinos are produced and detected as weak eigenstates and travel as superpositions of the mass eigenstates.

The Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix describes how the weak and the mass eigenstates are 'mixed'. The PMNS matrix contains six fundamental parameters that need to be determined ex-

perimentally. Most parameters have been measured to a high degree of accuracy and precision, however, the complex phase factor δ_{CP} related to Charge-Parity symmetry has not been precisely measured to date [7].

Phase differences accumulate between the mass eigenstates as they propagate through space, thus a neutrino may oscillate into a different weak eigenstate (e.g. $\nu_\mu \rightarrow \nu_e$, this is ν_e appearance and ν_μ disappearance). The likelihood of a neutrino maintaining its state is its ‘survival probability’ (e.g. $\mathbb{P}(\nu_\mu \rightarrow \nu_\mu)$), while the chance it changes state is its ‘transition probability’ (e.g. $\mathbb{P}(\nu_\mu \rightarrow \nu_e)$) [7].

B. In Search of Charge-Parity Symmetry Violation

Charge-Parity (CP) symmetry means that the laws of physics should remain unchanged if a particle is replaced by its antimatter counterpart (charge conjugation, C) and its spatial position is inverted (parity, P), changing \vec{x} into $-\vec{x}$. CP violation (CPV) implies that this symmetry is broken and does not hold in the universe. Meaning processes involving matter and antimatter are not mirror images in terms of their behaviour and are governed by different laws of physics [2, 10, 11]. In the PMNS matrix, CP symmetry is captured by δ_{CP} , the CP phase factor [7].

CPV can be detected and δ_{CP} measured

by examining if neutrinos and antineutrinos oscillate differently. A probe used is to check whether $\mathbb{P}(\nu_\mu \rightarrow \nu_e) = \mathbb{P}(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$. If $\sin(\delta_{CP}) \neq 0$, then the oscillation probabilities of neutrinos and their antimatter counterparts, antineutrinos, will be asymmetric [12]. Differences in these probabilities would indicate CPV.

To measure the difference in neutrino and antineutrino appearance and disappearance rates, identifying the neutrino flavour initiating the interaction is essential. Measuring the energy of the particles involved is necessary to estimate the energy of the initial neutrino. This process involves classifying events and analysing particle topologies. Along with characterising the leading lepton to identify the neutrino flavour, distinguishing particles as ‘tracks’ or ‘showers’—the two primary event topologies—is necessary to estimate particle energies [13].

C. Detector Technologies and DUNE

Historically, neutrino detectors mainly employed liquid scintillators and Cherenkov radiation [9, 15, 16, 17]. Recently, Liquid Argon Time Projection Chambers (LArTPCs) have gained traction due to their advantages [14, 18]. LArTPCs offer high-resolution 3D imaging, precise energy measurement, and scalability. Additionally, the

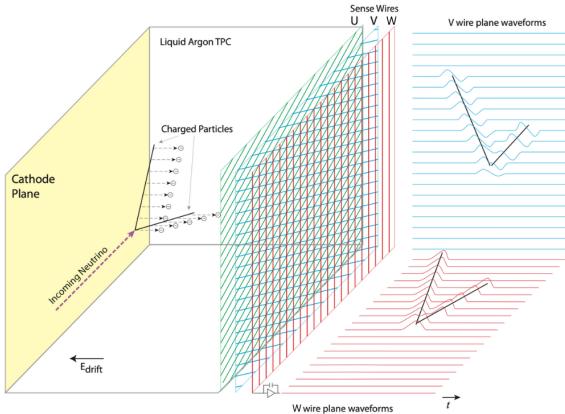


Figure 1: Schematic of a typical LArTPC detector, the particles are detected by three wire planes (u, v, w) [14].

high density of liquid argon increases the probability of neutrino interactions within the detector [14].

Shown in Figure 1, LArTPCs use liquid argon (Ar) to capture neutrino interactions by detecting charged particles via their ionisation trails. An electric field directs the ionisation electrons to three wire planes (u, v, w), with u and v as induction planes allowing electrons through, and w as the collection plane. The drift time provides a position coordinate for the drifting electrons. The position of the charge on each of the three wire planes is combined with the drift time to create three views, u, v and w , of each event. The waveforms recorded by each wire serve as a measure of the energy deposited by the particles. Thus, energy depositions are reconstructed as *hits* with associated energy values and position coordinates [14]. LArTPCs' high resolution in both space and energy enables precise par-

ticle identification, crucial for neutrino oscillation and CPV studies [5, 19]. Though LArTPCs allow 3D imaging, neutrino interaction events are most naturally visualised as 2D images in the wire position and the drift time parameter space populated by hits.

The Deep Underground Neutrino Experiment (DUNE), under the U.S. Department of Energy and hosted by Fermilab, aims to investigate CPV in neutrino oscillations. It features a near detector at Fermilab in Illinois for beam characterisation, and a far detector 1.5 km underground at the Sanford Underground Research Facility in South Dakota, 1300 km away (see Figure 2). The far detector, a large 70 kt modular LArTPC with 5 mm wire spacing, will provide detailed images of neutrino interactions. [6]

DUNE will use ν_μ and $\bar{\nu}_\mu$ dominated beams produced by the Long-Baseline Neutrino Facility (LBNF) beamline, measured and characterised at the near detector, located 574 m away, and then at the far detector using LArTPC detectors. The oscillation probabilities of ν_μ and $\bar{\nu}_\mu$ over a range of energies will be inferred by comparing the neutrino spectra observed at the near detector and the far detector, obtaining the $\nu_e/\bar{\nu}_e$ appearance and $\nu_\mu/\bar{\nu}_\mu$ disappearance to see if neutrinos and antineutrinos oscillate differently. It aims to reach a sensitivity to CPV of better than three

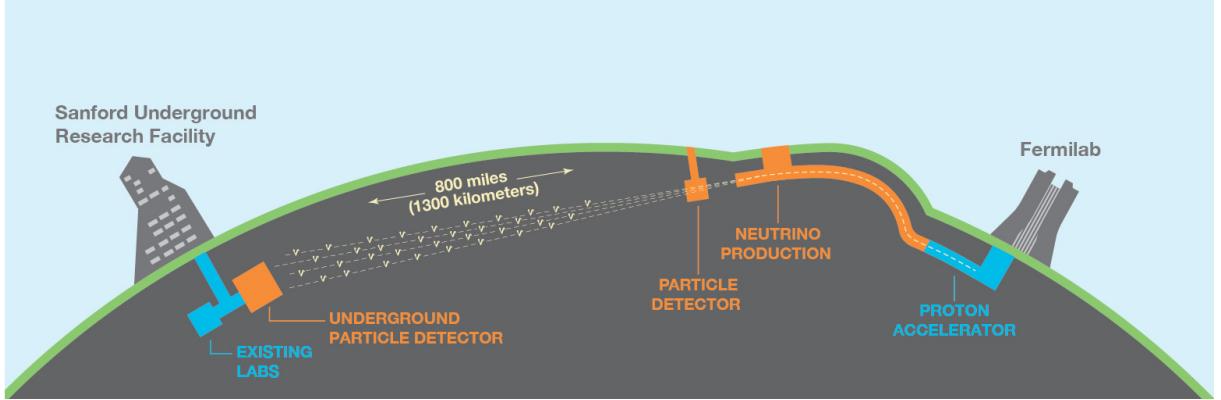


Figure 2: Illustration of the DUNE project, showing Fermilab and SURF, along with the beam production and detectors [6].

standard deviations (3σ), over more than 75% of the range of possible values of the unknown CP-violating phase δ_{CP} [6].

Simulated data that is representative of real experimental data is essential for developing classification models and showcasing the DUNE’s scientific potential. For DUNE, Monte Carlo simulations are conducted in three stages [12]: event generation with GENIE to mimic neutrino-nucleus interactions, particle tracking and interactions in liquid argon simulated by GEANT4, and finally, simulating the detector response with LArSoft [6].

This study uses two event reconstruction methods. The ‘perfect event reconstruction’ relies on the Monte Carlo truth record to accurately tag each hit with the Particle Data Group (PDG) code of the causing particle, ensuring precise identification [6]. PDG codes are numerical identifiers used to specify particles uniquely (e.g., μ^- : 13 and e^- : 11) [20]. The second method, ‘real

event reconstruction’, utilises the Pandora software package [21]. Pandora employs a multi-algorithm strategy using pattern recognition algorithms to decipher particles’ paths. Pandora’s effectiveness is measured by ‘completeness’, the portion of a particle’s original hits found in its reconstruction, and ‘purity’, the proportion of reconstructed hits belonging to the original particle. Pandora will be used to reconstruct events at DUNE.

Identifying relevant events ($CC \nu_\mu, \bar{\nu}_\mu, \nu_e$, and $\bar{\nu}_e$) while filtering out non-informative ones (NC interactions) is crucial at DUNE’s far detector. After isolating signal events and reconstructing their energies, a fit to the reconstructed neutrino energy distributions across the four event types will be conducted to deduce δ_{CP} [6, 12]. Figure 3 shows the appearance samples and how they are expected to vary with the true value of δ_{CP} using simulated data. Accurate event classification into CC

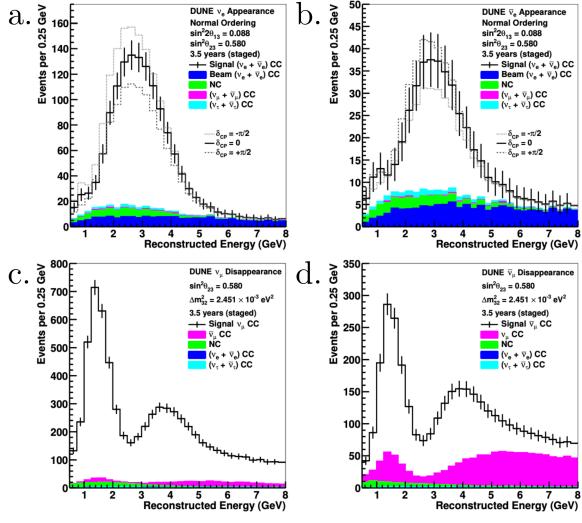


Figure 3: The simulated (a) ν_e and (b) $\bar{\nu}_e$ appearance spectra and how they vary with $\delta_{CP} = -\frac{\pi}{2}, 0$ and $\frac{\pi}{2}$ and the (c) ν_μ and (d) $\bar{\nu}_\mu$ disappearance spectra, for a 7-year beam exposure[6].

$\nu_\mu/\bar{\nu}_\mu$ or CC $\nu_e/\bar{\nu}_e$ is essential for measuring appearance/disappearance rates. Particle identification as ‘tracks’ or ‘showers’ is essential for estimating initial neutrino energies and facilitating this analysis.

D. Particle Topologies

Two major particle topologies arise at the energy levels relevant to DUNE (see Figure 4). Particles such as muons (μ^-), protons (p), charged pions (π^\pm) and kaons (K^\pm) exhibit characteristic *tracks*. Final-state muons in particular show long, straight tracks, losing little energy as they move through the liquid argon. On the other hand, final-state electrons (e^-) and photons (γ) exhibit characteristic electromagnetic (EM) *showers* [22].

Electrons initiate EM showers primarily via Bremsstrahlung leading to photon emis-

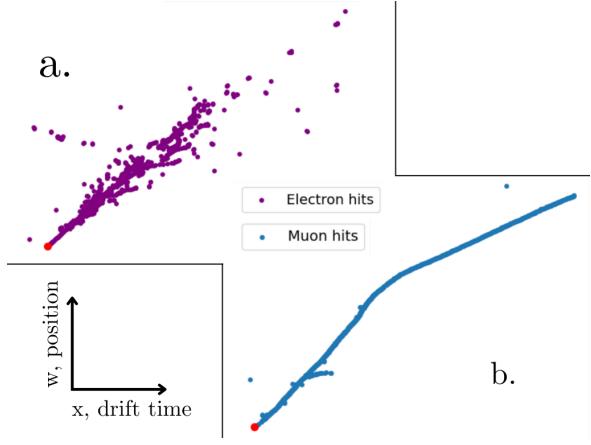


Figure 4: (a) An electron exhibiting a typical *shower* topology. (b) A muon exhibiting a typical long *track* topology. The vertices are shown as red dots.

sion. These photons can undergo pair production, forming electron-positron pairs. The process then repeats, perpetuating a cascade. There is significant energy deposition near the shower’s start [8]. In contrast, track-like particles such as muons and charged pions are characterised by a relatively uniform energy deposition ($\frac{dE}{dx}$) along their trajectory until they stop. Near the end of their path, these particles may exhibit a Bragg peak, a significant increase in energy deposition as the particle slows down and comes to a stop [19].

Observing a muon in a neutrino interaction indicates that a ν_μ initiated the process, and observing an electron shower suggests the presence of an ν_e (similarly for $\bar{\nu}_\mu$ and $\bar{\nu}_e$) [7]. Identifying these leading leptons is required when measuring $\nu_e/\bar{\nu}_e$ appearance and $\nu_\mu/\bar{\nu}_\mu$ disappearance to detect CPV. The picture gets complicated with NC interactions which leave no infor-

mation about the flavour of neutrino which caused the interaction.

At DUNE, the major background for CC ν_μ and CC $\bar{\nu}_\mu$ events will be NC interactions with charged pions (π^\pm) which can mimic the track topology exhibited by μ^\pm . For CC ν_e and CC $\bar{\nu}_e$, the background may include NC interactions with photons from π^0 decay which imitate the EM shower cascades exhibited by e^- . A small fraction of ν_e are intrinsic to the beam at DUNE and are not the result of neutrino oscillations. These also create a background noise [12]. In the final analysis, all background events need to be rejected, and signal events (CC ν_μ , $\bar{\nu}_\mu$, ν_e and $\bar{\nu}_e$) need to be identified from all the data collected at the far detector.

E. Machine Learning Methods

For DUNE to meet its objectives and precisely measure neutrinos, it's essential to classify final state particles as either track-like or shower-like. This classification is key for estimating initial neutrino energies and critical for plotting neutrino appearance and disappearance as a function of energy to determine δ_{CP} . Additionally, classifying entire events as CC ν_μ , CC ν_e , or NC ν_x is necessary for measuring ν_e appearance and ν_μ disappearance.

In this study, neutrinos and antineutrinos are analysed together due to their similar interaction topologies in LArTPC

detectors. Three machine learning algorithms for classification are explored: Projective likelihood estimators, Boosted Decision Trees (BDTs) and Convolutional Neural Networks (CNNs).

A *projective likelihood estimator* uses numerical input features to classify particles determining whether they are more likely to be a track (e.g., muons) or a shower (e.g., electrons and photons). This method employs probability density functions (PDFs) for various features that capture aspects of the particle trajectories (here relating to ‘track-ness’ or ‘shower-ness’). Termed also as a *naive Bayes estimator*, this technique involves multiplying the PDFs corresponding to each feature to compute an overall likelihood that a particle belongs to one of the two categories. This process is based on the assumption that the features are independent [23]. The likelihood of a particle being identified as a track or a shower is calculated as follows:

$$\mathcal{L}_{trk/shw} = \prod_i p_{trk/shw,i}, \quad (2)$$

where $p_{trk/shw,i}$ represents the PDF for the i th input feature, either for a track or a shower. This likelihood is normalised by the sum of the likelihoods for both categories, yielding a ratio:

$$y_{\mathcal{L},trk/shw} = \frac{\mathcal{L}_{trk/shw}}{\mathcal{L}_{trk} + \mathcal{L}_{shw}}. \quad (3)$$

By calculating this ratio for each particle, the estimator assigns it a normalised likelihood of being a track or a shower, facilitating its classification [23]. Projective likelihood estimators are a basic, fundamental machine learning technique that require manual feature engineering and present an easy and natural starting point for track-shower classification. A cut-off point needs to be decided for likelihoods where values above the cut-off point are classified as signal (here track), and the values below are classified as background (here shower).

Boosted Decision Trees (BDTs) improve particle classification by using decision trees to systematically segment data into specific categories based on binary decisions. Decision trees split the dataset into smaller subsets using input feature values until a stopping criterion is reached, with each subset eventually classified as either signal (e.g., tracks) or background (e.g., showers) based on the majority of data points it contains. Figure 5 shows a schematic of this process. This method progressively refines the classification, enhancing the performance over simpler models like the projective likelihood estimator by more accurately dividing the data into regions with distinct classifications [23].

BDTs enhance classification robustness by using a boosting technique that transforms a single decision tree into an ensemble, creating a forest. This process adjusts

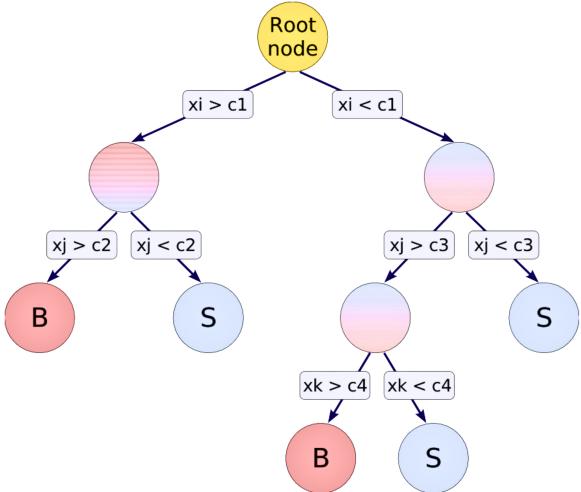


Figure 5: Schematic of a decision tree [23].

the weight of training instances, focusing subsequent trees on previously misclassified cases. Each tree's output contributes to the overall decision, with a weighted average forming the collective output. This approach improves reliability and accuracy in distinguishing between signal and background events by addressing variations in the training data [23].

The BDT method, with its sophisticated training and inference capabilities, offers more precise class differentiation within data than the projective likelihood method. It requires feature engineering and tuning of key hyperparameters, specifically the number of trees for boosting and the decision trees' maximum depth. While increasing the number of trees enhances performance, it also raises computational demands and the potential for overfitting. Similarly, deeper trees capture complex patterns but may overfit the training data.

Convolutional Neural Networks (CNNs) are a type of deep neural network optimised for visual tasks such as image classification. Central to CNNs are convolutional layers that automatically learn features from images, bypassing manual feature engineering. These layers use trainable filters to scan images, creating feature maps that highlight various aspects. CNNs, with their deep, multi-layered structure, process inputs layer by layer, evolving from simple to complex feature detection, leading to a final classification layer [12].

Customising the architecture of neural networks to fit specific classification tasks involves tuning architectural choices and hyperparameters. Key considerations include the number of convolutional layers and filters, which affect the network’s depth and its capacity to learn from data. Filter dimensions play a role, with smaller filters detecting finer details and larger ones capturing wider features. Pooling size impacts how much the input feature map is down-sampled, influencing spatial dimension reduction. Hyperparameters such as the learning rate, critical for weight updates during training, must be optimally set to avoid too rapid or too slow convergence. Batch size, the number of examples used per iteration, affects training speed and stability. The number of training epochs, or complete passes through the training dataset, is crucial to balance, preventing underfitting or overfitting.

F. Objectives

The study at hand addresses the vast data volumes expected from the DUNE experiment, focusing on classifying particles as track-like or shower-like and determining neutrino flavours in interactions by classifying events as CC ν_μ , CC ν_e or NC ν_x . These classifications are vital for examining neutrino and antineutrino oscillation patterns to detect CP violation and precisely measure the δ_{CP} phase.

A projective likelihood estimator was developed for track-shower particle classification, BDTs were employed for both track-shower particle classification and event classification, and a CNN was developed for event classification only.

III. Methodology

Simulated data representing DUNE neutrino interaction events were obtained in .root format. These files included truth records for event particles, featuring 2D hit records for u , v , and w wires, drift time, and PDG codes for particle identification. They also detailed particle parent-child relationships, energy deposition estimated via Analog-to-Digital Converter (ADC) values from the wires, neutrino interaction vertices, and reconstruction performance metrics like purity and complete-

ness. Additional data, such as counts of final state particles and 3D hit and vertex coordinates, were present but less central to this analysis.

The Python package `uproot` facilitated file access through two classes: ‘Events’ and ‘View’. ‘Events’ extracted detailed data for individual particles and events, including hit records, drift times, PDG codes, and more. ‘View’ simplified visualisation by focusing on essential data for event viewing: hit coordinates in x and z , ADC values, and true vertex locations on the appropriate wire plane. Event plots were integral for visual inspection and code debugging.

A. Particle Classification

Five features were developed to classify particles as tracks or showers. Orthogonal regression was used to derive a line of best fit for each particle, transforming $x - z$ coordinates to longitudinal (l) and transverse (t) distances. The features included: the fraction of hits close to the best-fit line, average hit separation, hits per unit length, energy profile, and variance in hit distribution along the trajectory. A correlation matrix was employed to confirm the distinct contribution of each feature to the classification process. These five features are further discussed in section IV.A.

2D images from the w -wire, chosen for its high signal-to-noise ratio and role in

collecting drifting electrons, were used for particle classification due to their adequacy and ease of use. Data were filtered to exclude trajectories with fewer than 10 hits, unidentified particles (PDG code = 0), and those with purity and completeness below 75%, ensuring the removal of noisy data.

A projective likelihood estimator, using these five features, was used for classifying particles as tracks or showers. Accuracy was optimised using a Receiver Operating Characteristic (ROC) curve, which plots *efficiency* against *purity* at various likelihood cut-off points. Efficiency captures the fraction of true tracks/showers that are correctly classified as tracks/showers, and purity captures the fraction of correctly classified tracks/showers in all that are reported as tracks/showers. The optimal cut-off point is chosen to maximise both metrics.

A BDT, using the same five features, was deployed to improve track-shower classification. A grid search across tree depths from 1 to 21 identified the optimal depth for the highest accuracy, with the number of boosting trees capped at 100 to balance computational efficiency and reduce overfitting risks. The optimal cut-off point for classification was determined using a ROC curve.

Performance evaluation involved metrics like accuracy and AUC (Area Under the ROC Curve) values, alongside tools like

confusion matrices, ROC curves, and efficiency plots. AUC values range from 0.5 (random guessing) to 1 (perfect classification), indicating the model’s ability to distinguish between the target classes. Confusion matrices compare actual versus predicted classifications, highlighting accuracy and misclassifications. Binomial errors were used to measure uncertainty in the confusion matrices’ elements. Accuracy itself is the proportion of correct predictions over all observations. Efficiency, the rate of correctly identified tracks/showers, was analysed against the number of hits in a particle to assess model limitations. Visual inspection facilitated the investigation of misclassified cases. Beyond track-shower particle classification, event classification to identify signal neutrino events ($\text{CC } \nu_\mu/\bar{\nu}_\mu$ and $\text{CC } \nu_e/\bar{\nu}_e$) and reject background events ($\text{NC } \nu_x$) is vital to achieving DUNE’s aims. This is the main classification task in measuring $\nu_e/\bar{\nu}_e$ appearance and $\nu_\mu/\bar{\nu}_\mu$ disappearance - which is fundamental to constraining and precisely measuring the parameters of the PMNS matrix, including δ_{CP} . This study treated neutrino and antineutrino events as the same as they display mostly similar topologies.

B. Event Classification

To facilitate event classification using BDTs, a new logic flow was developed.

The biggest particle, defined by the largest number of hits, from each event, was obtained and classified as track or shower using the existing BDT. If it was a track, a BDT was used to classify it as muon or background, this new BDT used two new manually engineered features. If it was a shower, a different BDT was employed to classify it as an electron or background, this BDT also used two new features. If the particle was a muon, the event was classified as CC ν_μ , if it was an electron the event was classified as CC ν_e , else it was classified as NC ν_x .

For distinguishing muons from background track-like particles in neutrino interactions, two features were developed: the **total track length**, derived from longitudinal coordinates in a transformed coordinate system, and the **total charge per track length**, estimated using ADC values. These features differentiate muons from charged pions and protons, enhancing particle classification accuracy. These are further discussed in section IV.D.

To classify shower-like particles as electrons or background photons in neutrino interactions, two features were used: the **vertex gap**, measuring the distance from the neutrino vertex to the track start, and the **initial track charge**, based on energy deposited in the first 10% of the track length, estimated using ADC values. These criteria distinguish electrons from photons, which

may originate from π^0 decay. These features are further discussed in section IV.D.

A truth record was created for each event by obtaining all the primary parent particles in the event, and checking if any were a muon or an electron. Events with primary muons or electrons were tagged as CC ν_μ or CC ν_e , respectively; others were marked as NC ν_x .

To improve event classification, CNNs were considered. Event images were created using the ‘View’ class, and a basic CNN architecture was implemented to begin with. The CNN did not use individual particles for event classification, thus particle reconstruction was no longer relevant for particle identification. Perfect reconstruction data was used to create the images. The CNN classified events as CC ν_μ , CC ν_e or NC ν_x . The truth record used for the BDTs was transformed into a one-hot encoding format suitable and necessary for CNNs.

Performance evaluation for the CNN included visual inspection and accuracy was used as a metric for overall performance. A 3×3 confusion matrix was used to evaluate performance for each target class. Loss metrics were used to find the optimal training times and epoch numbers to avoid overfitting.

These models classified individual particles and whole events, tasks that are crucial in the DUNE data analysis pipeline,

directly incumbent for measuring neutrino energies, appearance and disappearance rates, and thus pin down δ_{CP} .

IV. Results and Discussion

To begin particle track-shower classification, probability density functions were calculated and evaluated for each of the five features.

A. The Features

Energy: This feature exploited the energy profiles of showers (e.g., electrons) and tracks (e.g., muons) described in section II.D, quantified using:

$$\frac{\text{Energy in the last } 20\%}{\text{Total energy}}. \quad (4)$$

Higher values were expected for tracks due to the Bragg peaks at the end of the particle paths. Figure 6 shows the normalised probability density functions (PDFs) obtained for the feature. Tracks and showers are differentiated and, as expected, the ratio is generally greater for tracks. This feature is limited in cases where particles do not stop within the detector and thus show a consistent energy deposition till the end of the path, no Bragg peak is observed. Such cases are few and sit at the lower end of the PDF, overlapping with the showers. Overall, this feature provided powerful differentiation be-

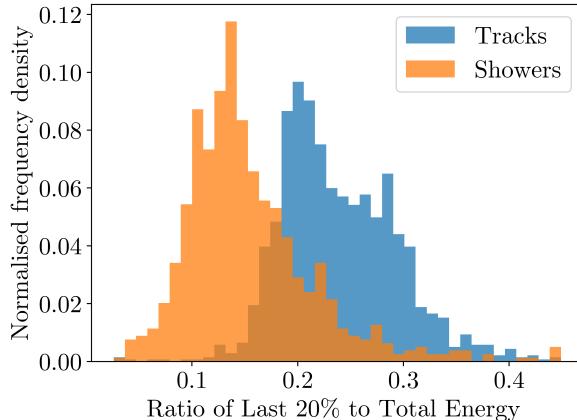


Figure 6: Histogram showing the energy feature’s probability density for 2000 DUNE data particles, using perfect reconstruction.

tween tracks and showers.

Two peaks are observed for tracks. Upon further investigation, these capture the difference in the size of the Bragg peak for muons and protons, as protons deposit a larger amount of energy towards the end of their tracks.

Standard deviation in hits: This feature differentiates showers and tracks by assessing hit dispersion along the orthogonally regressed best-fit line. Tracks generally follow a relatively straight path due to consistent energy loss and minimal scattering in the detector medium. Showers, on the other hand, form a branching cascade of secondary particles, leading to significant deviation from linearity and increased scattering, thereby widening the transverse distances from the best-fit line. The feature quantifies this distinction using the standard deviation of transverse distances ($l-t$ coordinates) for the last 30% of hits.

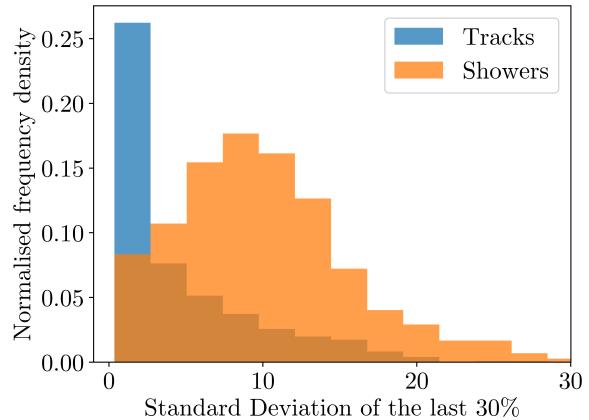


Figure 7: Histogram showing the standard deviation feature’s probability density for 2000 DUNE data particles, using perfect reconstruction.

This feature’s PDF is shown in Figure 7.

Tracks exhibit a uniform, narrow hit distribution around the best-fit line, corresponding to a lower standard deviation. In contrast, showers, due to their branching structure, display a broader hit distribution. Notably, showers overlap with tracks near lower (at 0-2) standard deviations. This occurs in low-energy showers from electrons and photons, which cannot develop into full showers and thus resemble short tracks, leading to standard deviations near zero. While this feature struggles to distinguish between tracks and showers in such scenarios, it generally enhances classification accuracy.

Hits per unit length: This feature distinguished tracks and showers based on their hits per unit length. Tracks show uniform hits due to consistent energy loss ($\frac{dE}{dx}$) and ionisation patterns. In contrast, showers, from a cascade of secondary particles,

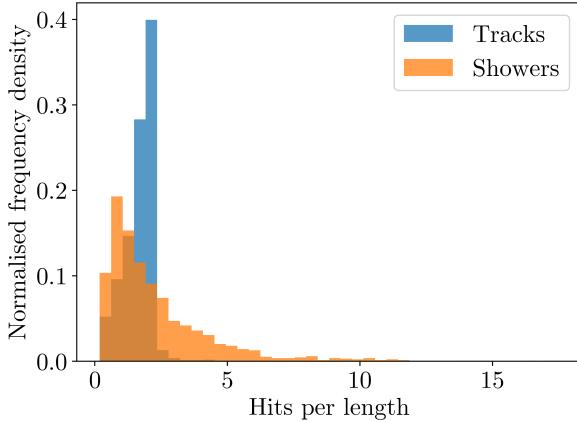


Figure 8: Histogram showing the probability density function for hits per length for 4000 DUNE data particles, using perfect reconstruction.

result in a sparse and wide-ranging distribution of hits per unit length. The feature was computed by dividing the total hits by the trajectory length, estimated from the range of longitudinal coordinates in $l - t$ coordinates. Showers were expected to register lower values than tracks. The PDF, shown in Figure 8, exhibited expected distributions. Tracks and showers were well differentiated.

Mean separation between hits: This feature differentiates between showers and tracks based on hit separation: hits in tracks are closer, whereas in showers, they’re more spread out. It calculates and averages the distances between hits, expecting larger values for showers and smaller for tracks. Figure 9 illustrates that while the feature generally distinguishes between the two, some showers overlap with tracks. Again, low-energy shower-like particles (electrons and muons) that re-

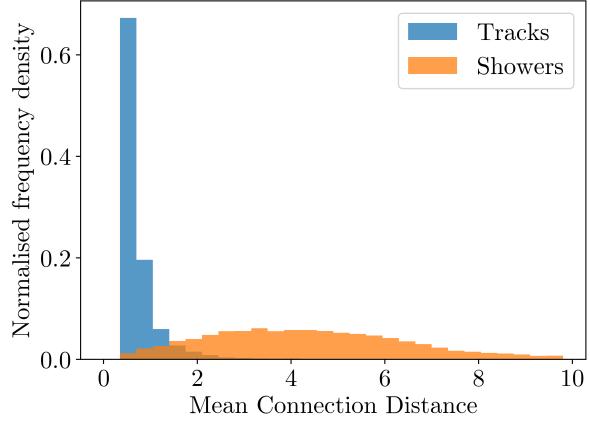


Figure 9: Histogram showing the probability density function for mean separation between hits for particles from 9000 DUNE events, using perfect reconstruction.

semble small tracks rather than developing full showers pose problems and reduce the feature’s effectiveness in low-energy cases. However, overall this feature captures the differences between tracks and showers excellently.

Fraction of proximal hits to the line of best fit:

This variable distinguishes showers from tracks based on the proximity of hits to the line of best fit: in showers, hits are typically further from the line than in tracks. This feature was calculated by obtaining the ratio of the total number of hits and the hits outside 5% of the coordinate range in the transverse (t) direction to the line best fit, with small ratios expected for tracks and large ratios expected for showers. Figure 10 shows expected behaviour, this feature captured differences between tracks and showers very well.

The correlation matrix in Figure 11 for the five features indicates most are largely

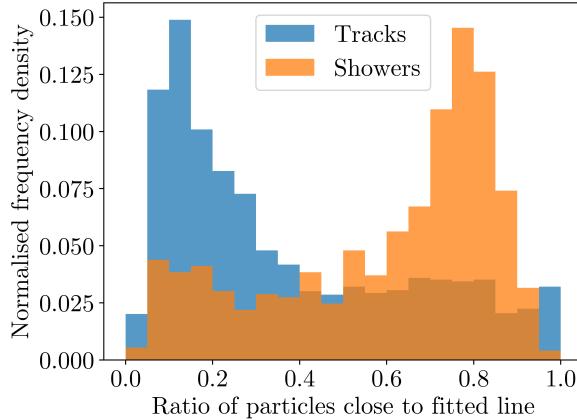


Figure 10: Histogram showing the probability density function for fraction of proximal hits to the line of best fit for particles from 1000 DUNE events, using perfect reconstruction.

uncorrelated, demonstrating their independence and distinct contribution to the classification task. This independence is crucial for the projective likelihood method, which presupposes input variable independence. Notably, the fraction of proximal hits and the standard deviation of the last 30% of hits show a correlation above 0.5, reflecting their shared focus on hit dispersion from the best-fit line. Mean separation and standard deviation of the last 30% of hits also correlate above 0.5 as they show similar distributions. Despite these correlations, the effects are not significant enough to impact the classification effectiveness of the projective likelihood approach.

B. Projective Likelihood Estimator

The data was divided into training and testing sets with a 70:30 ratio. The model was trained using both perfect reconstruction and real reconstruction. The training

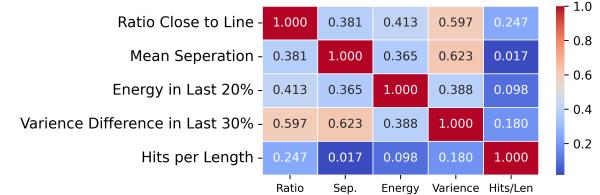


Figure 11: The correlation matrix for the variables used in the projective likelihood estimator and the BDT.

phase involved computing the features for each particle to create normalised PDFs for both track-like and shower-like particles. Equations 2 and 3 were then used to create the projective likelihood estimator. During testing, particle features were evaluated against these PDFs to determine their classification. A ROC curve was used to identify the best cut.

The AUC values for perfect and real reconstruction were 0.986 and 0.976, respectively, indicating the model's strong ability to distinguish between tracks and showers. The performance difference likely stems from Pandora's imperfect particle reconstructions, where particles might be incorrectly merged or split, a problem absent in perfect reconstruction data.

The confusion matrices in Figure 12 assess the classifier's performance, showing high reliability in distinguishing track-like events with a 99% accuracy for tracks and 87% for showers in real reconstruction data, the reconstruction that will be used at DUNE. This demonstrates the classifier's effectiveness in detecting ν_μ and,

True Track -	0.987 ± 0.002	0.01 ± 0.02
True Shower -	0.13 ± 0.03	0.87 ± 0.01
a.	as Track	as Shower
True Track -	0.964 ± 0.002	0.04 ± 0.01
True Shower -	0.08 ± 0.01	0.921 ± 0.004
b.	as Track	as Shower

Figure 12: The confusion matrices for (a) real and (b) perfect particle reconstruction.

to a slightly lesser extent, ν_e appearances, essential for the experiment’s goals. The disparity in detection rates between tracks and showers highlights challenges in classifying shower-like events. With an overall real reconstruction accuracy of 93%, there was potential for further optimisation.

Visual inspection of misclassified particles helped understand this model’s limitations and the errors indicated in the confusion matrix (Figure 12), shedding light on the factors behind the 7% of incorrectly classified events. This analysis revealed several consistent features among the misclassifications.

Misclassified Showers: Low energy electrons and photons, which lose energy rapidly and fail to form extensive shower-like cascades, leave minimal trails resembling tracks, leading to their misclassification by the model. Figure 13 provides examples. The input features were ineffective in distinguishing between showers and tracks in these low-energy scenarios, as

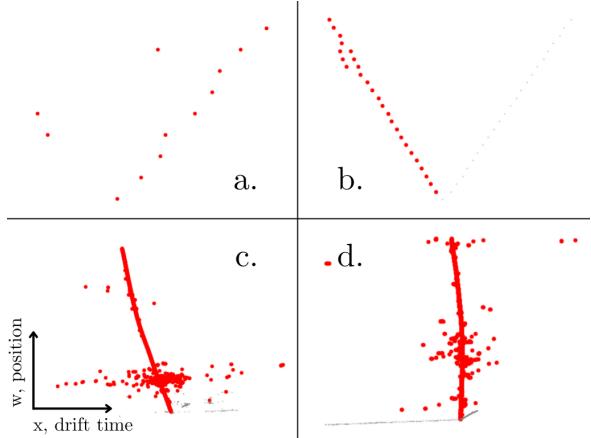


Figure 13: (a) and (b): These showers, likely low energy electrons or photons, were misclassified as tracks. (c) and (d): These tracks, showing some delta rays (electrons) emitting from the muon trajectories, were misclassified as showers.

mentioned before.

Misclassified Tracks: Muon tracks emitting delta rays—secondary electrons triggering electromagnetic cascades—can mimic showers. Perfect reconstruction often captures these as part of the muon track, leading to their misclassification as showers by the model. Figure 13 illustrates such instances. The input features did not effectively address these complexities in track-like topologies involving secondary interactions.

Similar conclusions were derived using Figure 14, which shows how classification efficiency varied with the number of hits in a particle. Tracks were consistently well classified in both perfect and real reconstructions, whereas showers with few hits were classified less efficiently. The decreased efficiency for showers in real reconstruction could stem from Pandora’s imper-

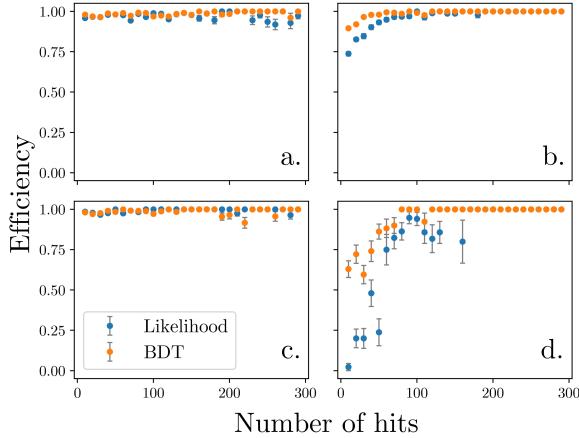


Figure 14: Efficiency for perfect reconstruction (a) tracks and (b) showers as a function of number of hits in a particle. Efficiency for real reconstruction (c) tracks and (d) showers. Likelihood results are shown in blue and BDT in orange.

fect particle reconstructions, such as the incorrect merging of particles.

In the DUNE experiment, even a small percentage of misclassified events could notably affect sensitivity and precision, impacting neutrino property measurements and CP-violating phase determination. Enhancing the classifier by integrating new and refining existing features could mitigate these limitations. Nonetheless, employing more sophisticated machine learning approaches like Boosted Decision Trees (BDTs) could yield substantial improvements using the same features.

C. Boosted Decision Tree

Using a BDT with the same features as the projective likelihood method improved real reconstruction classification accuracy to 97%, a 4% increase. The BDT, trained with a 70/30 data split, utilised an ROC

True Track	0.987 ± 0.002	0.01 ± 0.02
True Shower	0.05 ± 0.03	0.947 ± 0.007
a.	as Track	as Shower
True Track	0.980 ± 0.002	0.02 ± 0.01
True Shower	0.03 ± 0.01	0.972 ± 0.002
b.	as Track	as Shower

Figure 15: The confusion matrices for (a) real and (b) perfect particle reconstruction for the BDT.

curve to determine the optimal cut-off for peak classification efficacy. A grid search identified optimal hyperparameters, finding a tree depth of 12 to be most effective.

The BDT demonstrated a significant improvement over the likelihood approach, as reflected in AUC values of 0.991 for perfect and 0.988 for real reconstruction from ROC curve analysis. This marks a considerable improvement, particularly noticeable in shower identification, where it achieved an 8% increase in accuracy using the real reconstruction data. The performance gap between perfect and real reconstructions narrowed greatly. These improvements would greatly impact DUNE’s sensitivity and precision in measuring neutrino energies, appearance and disappearance rates.

Figure 15 presents confusion matrices for evaluating the BDT. Results for real reconstruction show a 99% correct detection rate for tracks and a 95% rate for showers. This demonstrates the BDT’s high reliability in identifying track-like events and a signifi-

cant 8% improvement in shower detection over the likelihood method, enhancing the ability to identify electron neutrinos crucial for ν_e appearance measurements. The BDT effectively addresses the limitations encountered by the likelihood method, particularly in distinguishing showers, even in low-energy scenarios where the likelihood struggled, while maintaining the same high performance for track classification.

Figure 14 tells the same story, showing that classification efficiencies for showers with a low number of hits, and therefore generally low energies, improved significantly. This improvement is solely attributed to the BDT’s more advanced inference method.

The next key task for the DUNE experiment involves distinguishing signal events (CC ν_μ and CC ν_e) from non-contributory background events (NC ν_x), necessitating whole event classification. A strategy was crafted employing the existing track-shower classifying BDT along with two newly developed BDTs specifically for DUNE event classification, tested solely on perfect reconstruction data due to time constraints.

D. Particle Based Event Classification

The identity of the biggest particle, defined by the largest number of hits, in an event, was used to identify the event as

a whole. This is a coarse heuristic which served as a good starting point for particle-based event classification. The largest particle was first categorised as a track or a shower using the pre-existing BDT. Subsequently, tracks were further differentiated into muons or background (e.g., charged pions/protons), and showers into electrons or background (e.g., photons), using two new BDTs equipped with features tailored for these distinctions. Events were then classified based on the largest particle: muons indicated CC ν_μ events, electrons suggested CC ν_e events, and other identifications implied NC ν_x events.

To classify tracks, the **total length** of the track, derived from the difference between the first and last longitudinal coordinates in the $l - t$ coordinate system, was used. This leverages the observation that protons, with higher energy loss rates, tend to have shorter tracks than muons and charged pions. Charged pions, capable of generating secondary particles through interactions, may show reduced primary track lengths and can exhibit kinks or sudden direction changes from pion-nucleus interactions.

The feature **total charge divided by track length** (using ADC values) served to differentiate between muons and charged pions/protons, providing a rough estimate of energy deposition rate per unit length ($\frac{dE}{dx}$) along the particle’s 3D trajectory.

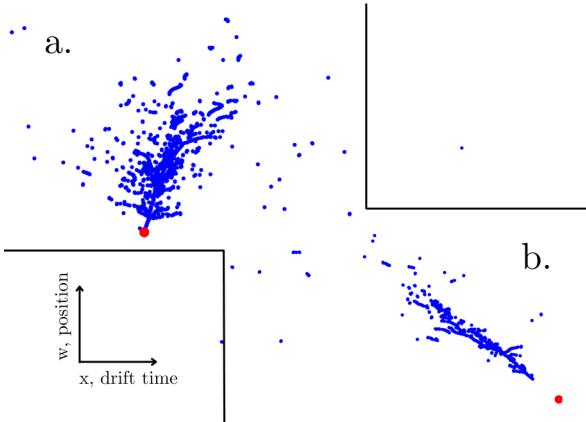


Figure 16: (a) An electron interacts with the medium almost instantly, while a gap is observed for (b) the photon.

Each particle type exhibits distinct $\frac{dE}{dx}$ characteristics. Muons, with their smaller interaction cross-section, lose energy consistently and minimally across various energies. Charged pions, while similar to muons at lower energies, can undergo strong interactions with nuclei, leading to inelastic collisions and potential abrupt stops or secondary particle production at higher energies. Protons, being significantly heavier, have a higher energy loss rate, marking a faster energy deposition than muons and pions.

To classify showers, the **vertex gap**—the distance between the neutrino vertex and the start of the particle track—was utilised. Electrons from neutrino interactions begin interacting with the medium almost instantly, resulting in minimal distance between the neutrino interaction vertex and the electron track’s onset due to immediate energy loss via ion-

isation and bremsstrahlung. Conversely, photons, being neutral, can traverse some distance before their interactions cause an electron-positron pair production, initiating an electromagnetic shower. This leads to a detectable conversion gap between the neutrino interaction point and the shower’s start. Figure 16 shows example events demonstrating these behaviours.

The feature **charge at the beginning of the particle track**, derived from the total energy (estimated using ADC values) in the initial 10% of the track divided by this segment’s length, was employed to reflect the conversion gap through energy deposition rates. Electrons immediately deposit energy upon creation, while photons begin energy deposition only after converting to an electron-positron pair, which then deposits energy at a doubled rate. This distinction leads to markedly different energy deposition patterns at track starts, effectively captured by this feature.

These features were then calculated for the biggest particles in all the events and used as input features for the two BDTs. The 46,424 events large dataset went through 70/30 training and testing splits at each step. Unseen events were classified through this workflow, achieving a 76% accuracy rate, with 24% of events misclassified. Each step of the logic flow must be investigated to evaluate and discuss this classification performance.

True ν_μ CC	0.77 ± 0.02	0.033 ± 0.007	0.20 ± 0.01
True ν_e CC	0.033 ± 0.007	0.78 ± 0.02	0.18 ± 0.01
True ν_x NC	0.16 ± 0.02	0.13 ± 0.02	0.72 ± 0.02

as ν_μ CC as ν_e CC as ν_x NC

Figure 17: The confusion matrix for event classification using BDTs.

Figure 17 presents a confusion matrix assessing the performance of this method. This approach yields 77% accuracy for CC ν_μ events, 78% for CC ν_e events, and 72% for rejecting background events. Accurate event classification is crucial for achieving DUNE’s objectives; these relatively low accuracies would adversely affect the analysis of ν_e appearance and ν_μ disappearance rates, hindering CPV detection and precise δ_{CP} measurement. While this method serves as a useful starting point, its effectiveness is limited.

The low accuracy observed is significantly influenced by the performance of two BDTs. The track-classifying BDT achieved 80% accuracy, while the shower-classifying BDT reached 84%. Enhancing the sensitivity of existing features to particle type differences and adding new features could improve these accuracies. Conversely, the classification of the largest particles in events as tracks or showers, which achieved a 99% accuracy rate, did not significantly affect the overall method’s performance.

The primary limitation of this approach is its reliance on the largest particle to

represent the entire event, which may not always correspond to the relevant leading lepton, leading to significant potential for errors. This necessitates the development and implementation of more sophisticated algorithms for accurately identifying the leading lepton in each event. Additionally, considering only the largest particle overlooks the broader event context, which could offer crucial insights for classifying whole events. Future methodologies should incorporate analysis of all interaction products for more effective particle-based event classification.

Convolutional Neural Networks (CNNs), which use an image-based classification technique address some of these issues and promise better performance due to their more advanced inference methods. Using event images, they take into account the whole of the event and not just the biggest particle. CNNs were implemented to classify simulated DUNE neutrino events as CC ν_μ , CC ν_e and NC ν_x .

E. Convolutional Neural Network

Event truth records were obtained in a one-hot encoded format as outlined in section III. Using an image-based approach with CNNs renders the particle reconstruction method irrelevant. Images for the CNN were generated using perfect reconstruction data, which provide a com-

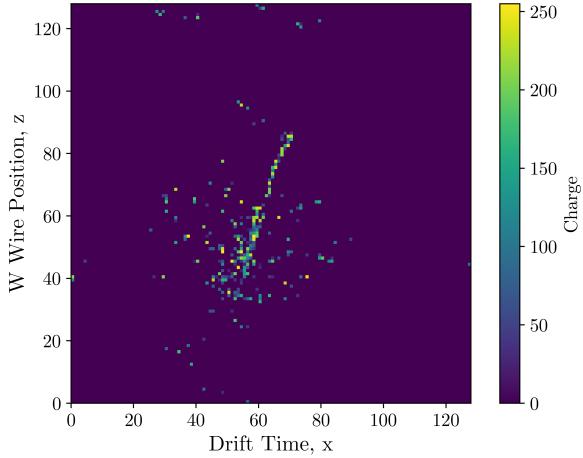


Figure 18: Example 128×128 pixel image. Initially, images were scaled to fit the whole event in the set image size, and colour (using ADC values) was used to indicate charge.

plete record of each event. No data cuts and manually engineered features were required.

Image creation involved the 'View' class for the w -wire plane, capturing essential data (drift time (x), wire position (z), and ADC values) for generating 2D event images. Initially, 128×128 pixel images were created, with pixel colour indicating hit charge based on ADC values; dark pixels represented no hit. Images were scaled according to event size, defined by the minimum and maximum $x - z$ coordinates. A total of 9310 event images were produced, divided into a 70:30 training and testing split. Figure 18 shows an example image.

These 128×128 images were used to train a CNN with a basic image classification architecture. The CNN processed these images, applying a single convolutional layer with 32 (3×3) filters and the ReLU ac-

tivation function to extract features like edges and textures. A subsequent Max-Pooling2D layer halved the feature maps' spatial dimensions to streamline data and lessen computational demands. A Dropout layer randomly deactivated 25% of neuron activations during training to prevent overfitting by avoiding reliance on certain neuron paths. The architecture then flattened the feature maps into a 1D vector, leading to a dense output layer with neurons equal to the class count, using softmax to produce a class probability distribution (CC ν_μ , CC ν_e , and NC ν_x), reflecting the network's class predictions.

Using the initial images, the CNN attained an overall accuracy of 74.9%, a marginal improvement over the previous method using BDTs. Before refining the CNN architecture, efforts were focused on enhancing the image quality to further boost performance.

Scaling images to a uniform size based on the size of the event distorted the spatial relationships and details within the data, with smaller event features seeming the same size as larger features across the images. Events in particle physics experiments can vary greatly in size and energy. Scaling images introduces inconsistencies in the physical scale represented by the pixel across different events. This inconsistency makes it challenging for the CNN to learn and generalise from the spatial patterns of

the interactions, as the same physical distance could correspond to a different number of pixels in scaled images. For example, a 5 cm distance would look the same as a 30 cm distance in these images as they are scaled to fit the event in the whole image, losing crucial information about the physical dimensions of the interaction products. Therefore instead of scaling, new images were created with the events cropped.

Image resolution was enhanced to 256×256 pixels, and a cropping interval centred around the mean x and z positions were employed, significantly enhancing image quality while maintaining the physical dimensions of events. Figure 19, depicting the same event as Figure 18, illustrates that the updated images accurately preserve event features without distortion. Additionally, removing the colour representation of the charge simplified the images, further improving CNN performance.

Using these refined images with the same CNN architecture, and using the same events for training, improved event classification accuracy to 83.6%. This represents an 8% improvement over the prior image creation method and a 10% enhancement compared to the original particle-based event classification approach. These optimised images were utilised in the final CNN.

The CNN architecture was enhanced to improve its efficacy. The final structure,

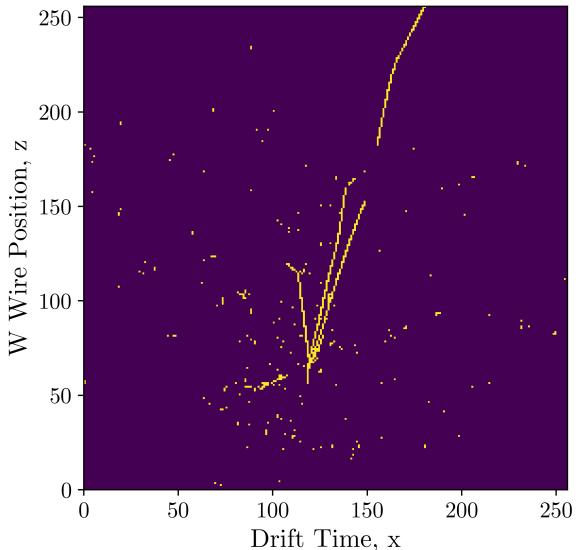


Figure 19: Example 256×256 pixel image used in the final model, showing the same event as Figure 18. The images were cropped instead of scaled, thus original physical dimensions were preserved.

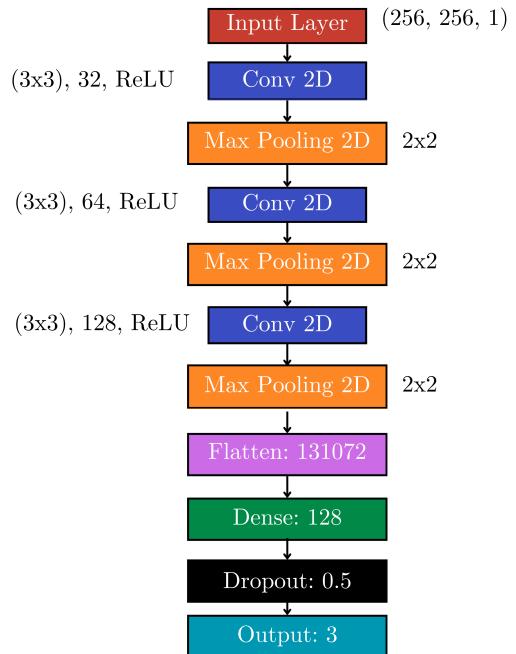


Figure 20: The final CNN architecture included three convolutional layers for 256×256 pixel images.

outlined in Figure 20, included strategic modifications for better feature extraction, overfitting reduction, and classification performance. It featured three convolutional

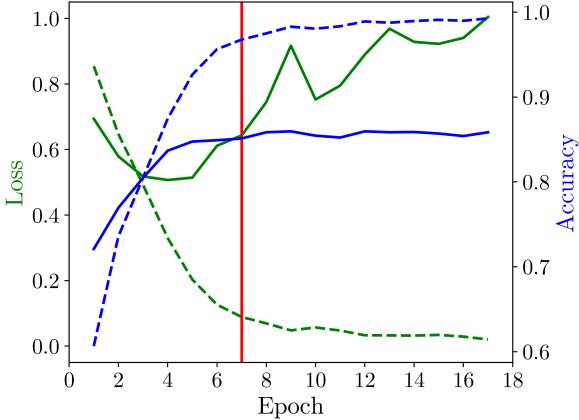


Figure 21: Loss and accuracy for training (dashed lines) and validation (solid lines) data during training. The red vertical line highlights the network results at epoch 7, where the accuracy plateaus.

layers, succeeded by max pooling, with filter counts of 32, 64, and 128 of size (3,3) and ReLU activation. These layers increase in depth to detect a broad spectrum of features, from simple to intricate, maintaining spatial feature map dimensions through 'same' padding. Following each convolutional layer, max pooling reduced representation size, cutting down parameters and computational demand, thus mitigating overfitting.

After processing through the convolutional layers, the output is flattened and passed to a dense layer with 128 neurons and ReLU activation, integrating the features into higher-level patterns for classification. A subsequent dropout layer with a 0.5 rate effectively reduces overfitting by randomly disabling half of the activations during training, preventing dependency on specific features. The concluding dense layer, using softmax activation,

True ν_μ CC	0.930 ± 0.006	0.031 ± 0.003	0.040 ± 0.003
True ν_e CC	0.071 ± 0.004	0.837 ± 0.007	0.092 ± 0.005
True ν_x NC	0.125 ± 0.007	0.126 ± 0.007	0.749 ± 0.01

as ν_μ CC as ν_e CC as ν_x NC

Figure 22: The confusion matrix for event classification using a CNN.

outputs classification probabilities for the multi-class scenario. This architecture is tailored to effectively manage event classification complexity, ensuring computational efficiency and model generalisation.

The CNN was trained using Python 3.9.18 and Tensorflow 2.14.0. The model was compiled with the Adam optimiser and a categorical cross-entropy loss function, selections that balanced learning efficiency and accuracy optimisation for categorical outcomes. With a batch size of 128 and a 0.001 learning rate, the sparse testing data doubled as a validation set, aiding in epoch optimisation and overfitting prevention. After 7 epochs, where test accuracy stabilised and loss began to rise, training was stopped, ensuring the model's generalisability and overfitting avoidance as depicted in Figure 21.

The CNN achieved an 85% accuracy in event classification, with Figure 22 displaying the performance per event type. Notably, CC ν_μ event classification improved to 93% accuracy, marking a 21% increase over the particle-based BDT method. CC ν_e classification also improved by 5%, reaching 84% accuracy. Background event

identification improved slightly, achieving 75% accuracy.

The advancements highlight CNNs as the most effective method for event classification within DUNE’s objectives, indicating the potential for even better results with more computational resources and complex CNN architectures to extract finer details from images. While the CNN demonstrated promise, especially with a 93% accuracy in identifying CC ν_μ events critical for measuring ν_μ disappearance, it also showed limitations in background event (NC ν_e) rejection, underscoring areas for improvement. This performance is noteworthy when compared to DUNE’s state-of-the-art convolutional visual network (CVN), which achieves 95% accuracy for CC ν_μ and 90% for ν_e events [12], reflecting the inherent challenge in ν_e event identification. Overall, the CNN approach significantly outperforms the BDT, particle-based method.

As ‘black boxes’, CNNs present a unique problem of interpretability in physics applications. The projective likelihood estimator and the Boosted Decision Trees were all based on features derived from observed physical features of the particles involved, while for the CNN, the decision-making process is not clear and well understood. Further investigation and interpretability analysis of the CNN is beyond the scope of this paper, however, tools such as fea-

ture maps, and changing the testing data to see how the CNN reacts, may be useful.

The use of simulated data also raises concerns about the difference in performance between simulation and real data, though CNNs have been successfully used on real neutrino data [13]. The models will fail to perform on real data if the simulations are not representative of the true physical phenomena. Testing on real events is crucial to validate simulation accuracy and model applicability.

The findings here are crucial for DUNE’s data analysis strategies, highlighting the importance of accurate automated event classification for signal detection and background rejection amidst the experiment’s extensive data volume. Such classifications are essential for CP violation detection and refined δ_{CP} measurement, the key to exploring why the universe exhibits matter dominance despite its symmetrical matter and antimatter origins. As machine learning evolves, its relevance in physics grows, becoming increasingly central to managing the high-dimensional data characteristic of contemporary large-scale experiments.

V. Conclusion

Original features were developed to classify track-like and shower-like particles in neutrino-induced interactions. The projective likelihood estimator provided a start-

ing track-shower classification accuracy of 93% with showers showing lower correct classification. The Boosted Decision Tree provided excellent improvements in classifying track-like and shower-like particles reaching an accuracy of 97%. This is most relevant in particle energy reconstruction, crucial in measuring neutrino energies at DUNE.

The particle-based event classification logic flow using BDTs classified events with a 76% accuracy, demanding serious performance improvements. The CNN addressed this greatly, classifying events with an accuracy of 85%, especially excelling at identifying muon neutrino events with a 93% accuracy. However, the CNN showed decreased background rejection capabilities, only correctly identifying NC ν_x events with 75% accuracy and found it generally harder to identify electron neutrino events with an 84% accuracy. These results are most crucial when identifying signal events and rejection events at DUNE to measure ν_e appearance and ν_μ disappearance rates which give insights into CPV in neutrino oscillations and can constrain δ_{CP} .

References

- [1] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics* (Oxford University Press, 2007).
- [2] I. I. Bigi and A. I. Sanda, *CP Violation* (Cambridge University Press, 2009).
- [3] J. Dodd and B. Gripaios, *The Ideas of Particle Physics* (Cambridge University Press, 2020).
- [4] D. J. Griffiths, *Introduction to Elementary Particles* (Wiley-Vch, 2011).
- [5] B. Baller *et al.*, Journal of Instrumentation **9**, T05005 (2014).
- [6] B. Abi *et al.*, arXiv:2002.03005 (2020).
- [7] M. Thomson and C. U. Press, *Modern Particle Physics* (Cambridge University Press, 2018).
- [8] D. H. Perkins, *Introduction to High Energy Physics* (Cambridge University Press, 2011).
- [9] Y. Fukuda *et al.*, Physical Review Letters **81**, 1562 (1998).
- [10] J. R. Fry, Reports on Progress in Physics **63**, 117 (2000).
- [11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Physical Review Letters **13**, 138 (1964).
- [12] B. Abi *et al.*, Physical Review D **102** (2020).
- [13] A. Abed Abud *et al.*, The European Physical Journal C **82** (2022).
- [14] R. Acciarri *et al.*, Journal of Instrumentation **12**, P02017 (2017).

- [15] A. Bellerive, J. Klein, A. McDonald, A. Noble, and A. Poon, Nuclear Physics B **908**, 30 (2016), Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.
- [16] M. Messier, Nuclear Physics B **908**, 151 (2016), Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.
- [17] T. Katori and J. M. Conrad, arXiv:1404.7759 (2014).
- [18] P. Jones, arxiv:1110.1678 (2011).
- [19] D. Kalra, Journal of Physics: Conference Series **2374**, 012163 (2022).
- [20] L. Garren, I. G. Knowles, T. Sjöstrand, and T. Trippe, The European Physical Journal C **15**, 205 (2000).
- [21] R. Acciari *et al.*, The European Physical Journal C **78** (2018).
- [22] J. L. Marshall and M. A. Thomson, The European Physical Journal C **75** (2015).
- [23] A. Hoecker *et al.*, *TMVA 4 Toolkit for Multivariate Data Analysis*, 2018.