

BACHELORTHESIS

Klassifikation atmosphärischer Neutrino-Signale in JUNO mittels Neuronaler-Netze

Kjell Simon Rosenow

Fakultät: Mathematik, Informatik und Naturwissenschaften Studiengang: Physik B.Sc. Matrikelnummer: 7404143

Erstgutachter: Prof. Dr. Caren Hagner Zweitgutachter: Dr. Daniel Bick

vorgelegt am 16.09.024

Zusammenfassung

Diese Arbeit erforscht den Einsatz des Jiangmen Underground Neutrino Observatory (JU-NO)-Detektors zur Untersuchung von atmosphärischen Neutrinos und deren Beitrag zur Bestimmung der Neutrinomassenhierarchie. Als einer der weltweit größten Flüssigszintillator-Detektoren bietet JUNO mit seiner hohen Präzision und Empfindlichkeit einzigartige Möglichkeiten zur Entschlüsselung kritischer Oszillationsparameter. Der Fokus liegt auf der Implementierung von Klassifikations-Neuronalen Netzen zur Datenanalyse, einer Methode, die sich durch ihre Fähigkeit auszeichnet, komplexe Muster in Neutrinoereignissen zu erkennen und zu klassifizieren. Trotz der noch ausstehenden Datenerfassung durch JU-NO wird in dieser Arbeit detailliert beschrieben, wie simulierte Daten zur Optimierung der neuronalen Netzwerke genutzt werden können. Durch die Anwendung von Bayes'scher Optimierung wurden F1-Scores für die 3 entwickelten Modelle bestimmt, mit M1: 45%, M2: 66% und M3: 69%, sie unterscheiden sich an verwendeten Trainingsdaten, welches zeigt, wie relevant die Rekonstruktion der Michel-Elektronen und Neutronen im JUNO Detektor.

Abstract

This work explores the use of the Jiangmen Underground Neutrino Observatory (JUNO) detector to study atmospheric neutrinos and their contribution to determining the neutrino mass hierarchy. As one of the world's largest liquid scintillator detectors, JUNO offers unique opportunities to unravel critical oscillation parameters with its high precision and sensitivity. The focus is on the implementation of classification neural networks for data analysis, a method distinguished by its ability to recognize and classify complex patterns in neutrino events. Although JUNO has yet to collect actual data, this work describes in detail how simulated data can be used to optimize the neural networks. Through the application of Bayesian optimization, F1-scores for the three developed models were determined as follows: M1: 45

Inhaltsverzeichnis

1 Einleitung						
2	Net	ıtrinophysik	4			
	2.1	Neutrinos im Standardmodell	4			
		2.1.1 Lücken im Standardmodell	4			
	2.2	Neutrinooszillationen	5			
		2.2.1 Im Vakuum	5			
		2.2.2 In Materie	6			
	2.3	Neutrinointeraktionen	8			
		2.3.1 Neutrale und geladene Streuung am Elektron	8			
		2.3.2 Neutrale und geladene Streuung am Nukleon	9			
	2.4	CP-Verletzung	11			
	2.5	Neutrino Quellen	11			
		2.5.1 Atmosphärische Neutrinos	11			
		2.5.2 Reaktorneutrinos	12			
	2.6	Offene Fragen in der Neutrinophysik	13			
		2.6.1 Massenordnung	13			
3	Jiar	agmen Underground Neutrino Observatory (JUNO)	15			
0	3.1	Aufbau des Detektors	15			
	0.1	3.1.1 Myon-Veto-System	16			
		3.1.2 Flüssigszintillator (FS)	16			
	3.2	Szintillation	16			
	0	3.2.1 Ziele des Detektors	17			
		3.2.2 Bestimmung der Massenhierarchie mit atmosphärischen Neutrinos .	18			
1	Kiiı	astliche neuronale Netze	10			
т	4 1	Prinzin künstlicher neuronaler Netze	10			
	1.1	4.1.1 Künstliche Neuronen	19			
		4.1.2 Architekturen künstlicher neuronaler Netze	20			
	42	Training	$\frac{20}{21}$			
	1.2	4.2.1 Trainingsalgorithmen	21			
		4.2.7 Überannassung und Regularisierung	$\frac{21}{22}$			
	43	Evaluierungsmetriken	$\frac{22}{24}$			
	1.0	4.3.1 Konfusionsmatrix (KMatrix)	$\frac{24}{24}$			
		4.3.2 Beceiver Operating Characteristic (BOC) Kurve	24 26			
	$\Delta \Lambda$	Bavesianische Ontimierung	26			
	1.7	4.4.1 Schritte der Bayesianische Optimierung	26			
			20			

5	Sim	ulierte Daten und aggregierte Features	28							
	5.1	Wahre und Ereignis Datensätze	30							
		5.1.1 Simulierte Ereignisparameter	30							
		5.1.2 Simulierte Messdaten	30							
	5.2	Wahl der aggregierten Features	30							
		5.2.1 "Number of Photoelectrons" Features	30							
		5.2.2 FHT Features	34							
		5.2.3 Features auf der Ereignis-Ebene	34							
	5.3	Klassifikationsfähigkeit der Features	36							
6	Aufbau und Auswertung der Netze									
	6.1	Netz Architektur	39							
		6.1.1 Aktivierungsfunktion und Dropout	39							
	6.2	Training und Optimierung	40							
		6.2.1 Optimierte Modelle	41							
	6.3	Evaluierung	41							
7	Zusa	ammenfassung	44							
\mathbf{A}	Anh	ang	47							
	A.1	Danksagung	50							
	A.2	Eidesstattliche Erklärung	51							

Kapitel 1 Einleitung

Neutrinos sind fundamentale, elektrisch neutrale Elementarteilchen, die eine zentrale Rolle im Standardmodell der Teilchenphysik einnehmen. Da sie nur über die schwache Kernkraft und die Gravitation mit Materie wechselwirken, sind sie schwer nachweisbar und bleiben weitgehend unerforscht. Besonders atmosphärische Neutrinos, die beim Aufprall von kosmischer Strahlung auf die Erdatmosphäre entstehen, bieten einzigartige Möglichkeiten zur Untersuchung dieser exotischen Teilchen.

Die Entdeckung der Neutrinooszillation im 20. Jahrhundert war ein entscheidender Durchbruch, der die Existenz von Neutrinomassen implizierte und das Standardmodell der Teilchenphysik erweiterte. Neutrinooszillation beschreibt das Phänomen, bei dem Neutrinos während ihrer Ausbreitung in verschiedene Geschmackszustände oszillieren. Dies ermöglicht es, wichtige Informationen über die Differenzen der Neutrinomassenquadrate und die Mischungswinkel zu extrahieren, jedoch bleibt die absolute Massenordnung der Neutrinos eine der größten offenen Fragen der modernen Physik. Auf die physikalischen Grundlagen der Neutrinophysik wird in Kapitel 2 genauer eingegangen.

Der Jiangmen Underground Neutrino Observatory (JUNO)-Detektor in China bietet aufgrund seiner hohen Präzision und Empfindlichkeit eine hervorragende Gelegenheit zur Untersuchung von Neutrinooszillationen. Als einer der weltweit größten Flüssigszintillator-Detektoren ist JUNO dazu in der Lage, Neutrinos mit einer beispiellosen Genauigkeit zu messen. Sein primäres Ziel ist die Entschlüsselung der Massenordnung der Neutrinos durch die präzise Bestimmung der Oszillationsparameter.

In dieser Arbeit werden wir uns darauf konzentrieren, wie der JUNO-Detektor zur Detektion von atmosphärischen Neutrinos eingesetzt werden kann und wie diese Messungen zur Bestimmung der Neutrinomassenhierarchie beitragen können, wird in Kapitel 3 weiter ausgeführt.

Hinsichtlich der Datenanalyse wird der Einsatz von Klassifikations-Neuronalen Netzen untersucht, die aufgrund ihrer Fähigkeit komplexe Muster zu erkennen und zu klassifizieren, eine vielversprechende Methode bieten, um die Vielfalt der Neutrinoereignisse präzise zu identifizieren und zu analysieren, was ausschlaggebend für die genaue Bestimmung der Oszillationsparameter ist. Wie neuronale Netze fähig sind, diese Zusammenhänge zu analysieren, wird in Kapitel 4 dargelegt.

Diese Datenanalyse ist bereits seit einigen Jahren im Gange, obwohl JUNO noch nicht am erfassen von Daten ist, wie diese simuliert und für die Verwendung in neuronalen Netzen verarbeitet werden wird in 5 gezeigt. Worauf die Optimierung und Bewertung der neuronalen Netze in 6 dargestellt werden.

Kapitel 2 Neutrinophysik

2.1 Neutrinos im Standardmodell

Das Standardmodell der Teilchenphysik (SM) [10] beschreibt die Elementarteilchen und unterscheidet zwischen Fermionen, die durch einen halbzahligen Spin charakterisiert sind und als Bausteine der Materie dienen, sowie Bosonen, die einen ganzzahligen Spin besitzen und die Fundamentalkräfte vermitteln. Die Fermionen werden weiter in Quarks und Leptonen unterteilt. Quarks interagieren miteinander durch ihre Farbladung und Gluonaustausch und bilden so Hadronen, die aus mehreren Quarks zusammengesetzt sind. Zu den Leptonen gehören die elektrisch neutralen Neutrinos sowie ihre geladenen Partnerleptonen. Im Gegensatz zu diesen Partnerleptonen und Quarks besitzen Neutrinos weder eine elektrische Ladung noch Farbladung und weisen zudem eine deutlich geringere Masse als alle anderen Fermionen auf. Aufgrund des Fehlens dieser Ladungen können Neutrinos ausschließlich über die schwache Wechselwirkung (SWW), vermittelt durch die W^{\pm} - und Z^0 -Bosonen, interagieren.

2.1.1 Lücken im Standardmodell

Obwohl das SM viele Prozesse korrekt beschreibt und Vorhersagen machte, die später experimentell bestätigt wurden, wie etwa die Existenz des Higgs-Bosons [6], ist es nicht ohne Schwächen. Ein bemerkenswertes Beispiel für die Unzulänglichkeiten des SM sind Neutrinos, da das Modell sie als masselos beschreibt. Diese Annahme wurde jedoch durch Experimente, wie dasjenige am Super-Kamiokande Detektor [9], widerlegt.

2.2 Neutrinooszillationen

Basierend auf Zuber [23] folgt: Leptonen werden als Fermionen in drei verschiedene Flavors eingeteilt, die Elektronen, die Myonen und die Tauonen. Die Neutrino-Flavors erben diese Namen mit dem Elektron-Neutrino, Myon-Neutrino und dem Tau-Neutrino. Neutrinooszillation bezeichnet die periodische Änderung der Neutrino-Flavor während der Ausbreitung durch den Raum. Die Idee der Neutrinooszillation wurde das erste Mal durch Pontecorvo [15] vorgeschlagen, um die Diskrepanz des gemessenen solaren Neutrino Stroms (siehe Abschnitt: 2.5) zu dem theoretischen zu vergleichen. Die Messungen, welche den Neutrinostrom der Sonne ausreichend gut bestimmten, um Neutrinooszillation zu beweisen, fanden am Super-Kamiokande Detektor statt und wurden 1988 von Fukuda u. a. [9] veröffentlicht.

2.2.1 Im Vakuum

Für die Beschreibung der Neutrinooszillation geht man davon aus, dass bei der Wechselwirkung sich das Neutrino in einem der drei Flavor-Eigenzuständen ν_e , ν_{μ} oder ν_{τ} befindet, diese sind Superpositionen dreier Massen-Eigenzuständen ν_1 , ν_2 und ν_3 mit den jeweiligen Massen m_1 , m_2 und m_3 . Wobei relevant ist, das die Massen der Massen-Eigenzustände nicht dieselben sind, wie die der Flavor-Eigenzustände. Der Zusammenhang folgt nach:

$$|\nu_{\alpha}\rangle = \sum_{j} U_{\alpha j} |\nu_{j}\rangle \iff |\nu_{j}\rangle = \sum_{\alpha} U_{j\alpha}^{\dagger} |\nu_{\alpha}\rangle, \qquad (2.2.1)$$

mit $|\nu_{\alpha}|$ als Zustand des Neutrinos, wobei α die Flavor- und j die Massen-Eigenzustände bezeichnet. Bei U handelt es sich um die Pontecorvo-Maki–Nakagawa–Sakata-Matrix U_{PMNS} (PMNS-Matrix), eine unitäre Matrix, was bedeutet: $U^{\dagger}U = 1$.

$$U_{\rm PMNS} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{\rm CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{\rm CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{\rm CP}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{\rm CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{\rm CP}} & c_{23}c_{13} \end{pmatrix}$$
(2.2.2)

Die Elemente der Matrix setzen sich aus s_{jk} , c_{jk} und $e^{\delta_{CP}}$ zusammen. s_{jk} , c_{jk} sind Abkürzungen für $\sin \theta_{jk}$ und $\cos \theta_{jk}$ mit den Mischwinkeln θ_{jk} und δ_{CP} ist die mögliche Verletzung der Ladungs- und Raumsymmetrie (Charge and Parity) (CP-Verletzung) in der Neutrinooszillation (siehe Abschnitt: 2.4).

Die Masse-Eigenzustände $|\nu_j\rangle$ sind stationär und entwickeln sich in Zeit und Raum nach

$$|\nu(x,t)\rangle = e^{-iE_jt}|\nu(x,0)\rangle = e^{-ipx}e^{-iE_jt}|\nu\rangle$$
(2.2.3)

mit p als Momentum, welches die Neutrinos haben, wenn sie am Ort x = 0 mit t = 0 erzeugt werden. Da es sich bei den Neutrinos um ultra-relativistische Teilchen handelt, müssen wir für die Neutrinoenergie dies berücksichtigen:

$$E_j = \sqrt{m_j^2 + p_j^2} \simeq p_j + \frac{m_j^2}{2p_j^2} \simeq E + \frac{m_j^2}{2E}$$
(2.2.4)

Dies gilt da $p \gg m_j$ und damit $E \approx p$. Für den Übergang von einem Neutrino mit Flavor $|\nu_{\alpha}\rangle$ erzeugt an der Stelle t = 0 folgt die Entwicklung des Zustands mit Gleichungen (2.2.1) und (2.2.3)

$$|\nu_{\alpha}(x,t)\rangle = \sum_{j} U_{\alpha j} e^{-ipx} e^{-iE_{j}t} |\nu_{j}\rangle = \sum_{j,\beta} U_{\alpha j} U_{j\beta}^{\dagger} e^{-ipx} e^{-iE_{j}t} |\nu_{\beta}\rangle$$
(2.2.5)

An dieser Stelle wird der Einfluss der Masse aus (2.2.4) relevant, denn aus unterschiedlichen Massen für die Neutrinos resultiert ein leicht verschiedener Phasenfaktor in (2.2.5). Dieser Unterschied führt dazu, dass ein Neutrino, welches an x = 0 erzeugt wurde, durch die Propagation des Raums periodisch den Flavor-Eigenzustand wechselt.

Die Wahrscheinlichkeit des Übergangs $P(\alpha \to \beta)(t)$ von einem Flavor-Eigenzustand $|\nu_{\alpha}(x,t)\rangle$ zu einem Flavor-Eigenzustand $|\nu_{\beta}\rangle$ ist gegeben durch

$$P(\alpha \to \beta)(t) = |A(\alpha \to \beta)(t)|^2 = |\langle \nu_\beta | \nu_\alpha(x, t) \rangle|^2$$
$$= |\sum_{j,\beta} U^{\dagger}_{k\beta} U_{\alpha j} e^{-ipx} e^{-iE_j t} \langle \nu_k | \nu_j \rangle|^2$$
(2.2.6)

mit der Übergangsamplitude $A(\alpha \to \beta)(t)$ und der orthonormalen Massen-Eigenzustände, für die $\langle \nu_i | \nu_j \rangle = \delta_{ij}$ gilt

$$P(\alpha \to \beta)(t) = |\sum_{j} U_{j\beta}^{\dagger} U_{\alpha j} e^{-ipx} e^{-iE_{j}t}|^{2} = \sum_{j} U_{j\beta}^{\dagger} U_{\alpha j} U_{l\beta}^{\dagger} U_{\alpha l} e^{-iE_{j}t}$$
(2.2.7)

und durch Einsetzen von (2.2.4) folgt

$$P(\alpha \to \beta)(L) = \sum_{j} U_{j\beta}^{\dagger} U_{\alpha j} U_{l\beta}^{\dagger} U_{\alpha l} \exp(-i\frac{\Delta m_{jl}^2}{2E}L), \qquad (2.2.8)$$

wobei $\Delta m_{ij}^2 = m_i^2 - m_j^2$ ist *L* die Distanz der Quelle und des Detektors, welches durch L = x = ct und damit unter Verwendung der natürliche Einheiten L = t folgt.

In Gleichung (2.2.8) ist zu erkennen, dass Neutrinooszillation nur existieren kann, wenn die Massendifferenz $\Delta m_{ij}^2 = 0$ ist, ansonsten folgt aus (2.2.8), dass die Übergangswahrscheinlichkeit gleich null sein muss. Damit die Massendifferenz nicht verschwindend ist und es Neutrinooszillation zwischen allen drei Flavor-Eigenzuständen gibt, müssen die drei Masse-Eigenzustände sich in ihrer Masse voneinander unterscheiden ($m_i \neq m_j \neq m_k$).

2.2.2 In Materie

Wenn Neutrinos dichte Materie durchqueren, sind weitere Überlegungen neben der Theorie im Vakuum notwendig. Aufgrund ihrer schwachen Wechselwirkung spüren Neutrinos ein Potenzial von den umgebenden Protonen, Neutronen und Elektronen. Auch beschrieben als das kohärente Vorwärtsstreuen von Neutrinos. Die Interaktionen des neutralen Stroms (siehe Abschnitt 2.3) mit der umgebenden Materie werden dabei vernachlässigt, da sie zu keinen Unterschieden zwischen den Teilchen führen. Die Wechselwirkung des geladenen Stroms von Elektronen und Neutrinos geschieht ausschließlich mit Elektronund Antielektron-Neutrinos. Es wird das Materiepotential

$$A = 2\sqrt{2}G_F N_e p \tag{2.2.9}$$

eingeführt. Hierbei wurden die Fermi-Kopplungskonstante G_F und die Elektronendichte N_e verwendet. Im Fall des Anti-Neutrinos wirkt das Materiepotential negativ. Dieser Effekt wurde als erstes von L. Wolfenstein in 1978 beschrieben [22] und durch Erweiterung der Theorie in 1986 von S. Michejew und A. Smirnow, ist er heute als der MSW-Effekt bekannt

2.3 Neutrinointeraktionen

Neutrinointeraktionen erfolgen ausschließlich über die schwache Wechselwirkung, welche durch die Vermittlung der Z^{0} - und dem W^{\pm} -Bosonen realisiert wird. Im Gegensatz zu Gluonen und Photonen, die beide masselos sind, besitzen diese Eichbosonen eine beträchtliche Masse. Die Massen betragen ungefähr $m_W \approx 80,37 \text{ GeV}$ und $m_Z \approx 91,18 \text{ GeV}$ [13]. Aufgrund dieser Massen ist die Reichweite der SWW sehr begrenzt, was sich durch folgende Abschätzung verdeutlichen lässt:

$$R_{W/Z} \approx \frac{\hbar}{m_{W/Z} \cdot c} \Rightarrow R_W \approx 2,73 \,\mathrm{pm} \text{ und } R_Z \approx 2,41 \,\mathrm{pm}$$
 (2.3.1)

Hierbei steht R für die Reichweite und m für die jeweilige Masse der Bosonen. Zum Vergleich: Der Durchmesser eines Protons beträgt etwa d = 0.8 fm, was circa 400-mal größer ist als die Reichweite der SWW.

Somit stellt die SWW einen rein subatomaren Effekt dar, der eine Vielzahl von Zerfallsprozessen vermittelt. Zu den bekanntesten Beispielen zählen der β^+ - und β^- -Zerfall:

$$\beta^{+}: p^{+} \to n + e^{+} + \nu_{e} \beta^{-}: n \to p^{+} + e^{-} + \bar{\nu}_{e}$$
(2.3.2)

In diesen Zerfällen werden Neutrinos produziert. Für die Detektion von Neutrinos werden mögliche Wechselwirkungen von Neutrinos mit Materie untersucht. Die möglichen Partner dafür sind hauptsächlich Elektronen und die Nuklei der Atome.

2.3.1 Neutrale und geladene Streuung am Elektron

Bei der Streuung eines Neutrinos an einem Elektron sind zwei verschiedene Wechselwirkungen möglich:



Abbildung 2.1: Streuung von Neutrinos ν_l an Elektronen e^- mit $l = e, \mu$ oder τ . Dabei trägt, das Z^0 -Boson den neutralen Strom und das W^{\pm} -Boson den geladenen.

Der neutrale Strom (im Englischen Neutral Current (NC)) in Abb. 2.1a mit $l = e, \mu$ oder τ ist ein elastischer Stoß der Neutrinos mit dem Elektron. Bei NC-Ereignissen hängt die an das Elektron übertragende Energie vom Streuwinkel θ ab, wodurch eine Bestimmung der Energie des Neutrinos nicht möglich ist. Im Fall, dass das Elektron stationär ist, gibt die übertragende Energie einen Mindestwert für die Energie des Neutrinos.

In dem geladenen Strom (im Englischen NC) wird in der Wechselwirkung die Ladung des Elektrons über das W^{\pm} -Boson übertragen und das Neutrino wandelt sich in ein Elektron um. Dabei behält es basierend auf Energie- und Impulserhaltung einen Großteil der Energie des Neutrinos. Die Energie des Neutrinos lässt sich durch die Vermessung des Elektrons bestimmen.

Für alle Experimente, welche mit Neutrinos unter $\sim 11 \,\text{GeV}$ ist der geladene Strom nur für Elektron-Neutrinos möglich, sollten die Neutrinos eine höhere Energie haben ist ein inverser Myon-Zerfall möglich.

$$\nu_{\mu} + e^{-} \to \mu + \nu_{e} \tag{2.3.3}$$

Der inverse Myon-Zerfall (2.3.3) ist erst ab einer Energie von 10,9 GeV möglich [19]. Bei dem inversen Myon-Zerfall wechselwirkt ein Myon Neutrino mit einem Elektron und erzeugt ein Myon und ein Elektron Neutrino. Das entstandene Myon zerfällt nach durchschnittlichen 2,2 µs in ein Myon-Neutrino, anti Elektron-Neutrino und ein Elektron. Dieses Elektron wird Michel Elektron genannt, nach Michel [12].

2.3.2 Neutrale und geladene Streuung am Nukleon

Welche Wechselwirkung am häufigsten auftritt, hängt stark von der Energie ab.



Abbildung 2.2: Wirkungsquerschnitte für Charged Current (CC) Wechselwirkungen pro Nukleon und geteilt durch die Energie der Neutrinos. QE steht für quasi-elastische Streuung (quasi-elastic scattering), RES für Resonanz Produktion (resonance production) und DIS für tief inelastische Streuung (deep inelastic scattering). Die Grafik kommt aus [8], wo die Messdaten, welche als die verschiedenen Punkte zu erkennen sind, genauer zu ihren Experimenten zugeteilt werden.

Quasi-elastische Streuung (QE)

In der quasi-elastischen Streuung wechselwirkt das Neutrino mit einem der Neutronen oder Protonen im Fall eines anti Neutrinos. Die Produkte dieser Reaktion erzeigen dabei eine sehr deutliche Signatur für die Detektion, das entstehende Positron annihiliert mit einem anderen Elektron und das Neutron wird durch Neutron capture wieder eingefangen, wobei dies auch eine für den Prozess typische Energie freisetzt.



Abbildung 2.3: Geladene Streuung eines anti Neutrinos $\bar{\nu}_{\ell}$ and einem Proton eines Nukleons

Resonanz Produktion (RES)

Bei höheren Energien kommt es dazu, dass das Neutrino, ein Neutron, in einen höheren Energiezustand angeregt wird, welcher schnell wieder zerfällt und dabei ein Pion emittiert.



Abbildung 2.4: Streuung von Neutrinos ν_l an Neutronen e^- mit $l = e, \mu$ oder τ . Hier mit einer Resonanz Produktion.

Tief inelastische Streuung (DIS)

In der tief inelastischen Streuung ist die Energie des Neutrinos ausreichend, um in ein Neutron vorzudringen, und direkt mit den Quarks zu wechselwirken. Bei dieser Interaktion wird genügend Energie an das Quark übertragen, um einen Hadron-Schauer auszulösen.



Abbildung 2.5: Feynman-Diagramm für tief inelastisches Neutrino-Streuen.

2.4 CP-Verletzung

Ein System gilt als symmetrisch in der Physik, wenn seine Eigenschaften durch bestimmte Transformationen geändert werden können, ohne dass ein messbarer Unterschied auftritt. Unter Ladungskonjugation versteht man den Austausch der Ladungen aller Teilchen (zum Beispiel $e^-/p^+ \rightarrow e^+/p^-$, etc.), wobei das System unverändert verhält.

Die CP-Symmetrie, bekannt als "Charge-Parity"-Symmetrie, beschreibt die Invarianz unter gleichzeitiger Ladungskonjugation und räumlicher Spiegelung. Bis zu den 1960er-Jahren ging man davon aus, dass diese Symmetrie für alle physikalischen Prozesse gültig ist. In den 1960er-Jahren wurde jedoch das K-Meson entdeckt und im Zuge dessen eine Verletzung der CP-Symmetrie in der schwachen Wechselwirkung des Teilchens festgestellt [16].

Die Entdeckung der CP-Verletzung ist bedeutend, da sie hilft, die Matter-Antimatter-Diskrepanz zu erklären. CP-verletzende Prozesse verhalten sich nicht identisch für Materie und Antimaterie und können so zu dieser Diskrepanz beitragen.

In der Neutrinophysik treten CP-Verletzungen in der PMNS-Matrix-Matrix auf, in Form der CP-verletzenden Phase δ_{CP} . Da sie in der PMNS-Matrix-Matrix auftritt, spielt sie auch eine Rolle im Drei-Neutrino-Modell. In der Gleichung für die Oszillationswahrscheinlichkeit ergibt sich:

$$P(\nu_{\alpha} \to \nu_{\beta}) = \sum_{i,j} U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j} e^{i(\delta_{ij} - \delta_{ji})} \sin^2\left(\frac{\Delta m^2_{ij}L}{4E}\right), \qquad (2.4.1)$$

wobei $U_{\alpha i}$ die Elemente der PMNS-Matrix sind, Δm_{ij}^2 die Differenz der quadrierten Massen, L die Distanz und E die Energie des Neutrinos.

Sollte δ_{CP} nicht verschwindend sein, hat dies einen messbaren Einfluss auf die Neutrinooszillation. Die Phase lässt sich daher durch Experimente mit ausreichender Präzision bestimmen.

2.5 Neutrino Quellen

Wie in Abschnitt 2.3 gezeigt, können viele Prozesse Neutrinos erzeugen. Die Neutrino Quellen unterscheiden sich in der Anzahl der erzeugten Teilchen und der Energie dieser.

In Abbildung 2.6 sind die Neutrino Quellen abgebildet, welche relevante Mengen an Neutrinos in den jeweiligen Energiebereichen beitragen. Von links nach rechts erwartet man bei geringen Energien den Kosmischen Neutrino Hintergrund (CNB), ein Artefakt des Urknalls und Neutrinos welche kurz nach diesem in der primordialen Nukleosynthese (BBN) entstanden sind. Es folgen die Sonne und der diffuse Supernova Neutrino Hintergrund (DSNB), sowie die Geoneutrinos, die durch radioaktive Zerfälle in der Erde entstehen und die Reaktorneutrinos. Atmosphärische Neutrinos von 10 MeV bis 100 TeV, sowie kosmische Neutrinos bei PeV und höher.

2.5.1 Atmosphärische Neutrinos

Atmosphärische Neutrinos entstehen durch Wechselwirkungen von kosmischen Strahlen in der Erdatmosphäre, insbesondere durch den Zerfall von Sekundärmesonen wie Pionen und Kaonen.



Abbildung 2.6: Neutrino-Energiefluss $E\phi$ in Abhängigkeit von der Energie in den Einheiten cm⁻²s⁻¹. Darstellung des Grand Unified Neutrino Spectrum (GUNS) an der Erde, integriert über alle Richtungen und über die Flavors. Feste Linien sind für Neutrinos, gestrichene oder gepunktete Linien für Antineutrinos, überlagerte gestrichelte und feste Linien für Quellen sowohl von ν als auch von $\bar{\nu}$. [20]

1. Pionzerfall:

$$p + N \rightarrow \Delta^+ \rightarrow p + \pi^0, \quad \Delta^+ \rightarrow n + \pi^+$$
 (2.5.1)

Hierbei erzeugt der Zerfall der Δ -Resonanz Pionen, die dann weiter zerfallen und Neutrinos produzieren.

2. Kaonzerfall:

$$K^+ \to \mu^+ + \nu_{\mu}, \quad \mu^+ \to e^+ + \nu_e + \overline{\nu}_{\mu}$$
 (2.5.2)

Die erzeugten Neutrinos haben typischerweise Energien im Bereich von einigen GeV bis zu mehreren TeV, wobei die hier betrachtete Bandbreite von einigen GeV spezifisch für die Energie der erzeugenden Teilchen und ihre Wechselwirkungen mit der Atmosphäre ist.

2.5.2 Reaktorneutrinos

Reaktorneutrinos sind Neutrinos, die in Kernreaktoren erzeugt werden. Diese Neutrinos stammen hauptsächlich aus den Beta-Zerfällen von radioaktiven Isotopen, die während des Betriebs eines Kernreaktors entstehen. Insbesondere in thermischen Reaktoren, die Uran-235 oder Plutonium-239 als Brennstoff verwenden. Bei der Spaltung von Uran-235 oder Plutonium-239 werden große Mengen Energie freigesetzt, sowie verschiedene Spaltprodukte, darunter auch Neutronen und instabile Isotope.

Neutrinos können aus diesen Spaltprodukten durch den Beta-Zerfall erzeugt werden:

$$n \to p + e^- + \bar{\nu}_e \tag{2.5.3}$$

oder

$$\beta^-: \text{Uran-}235 \to \text{Krypton} + \nu_e + \nu e^-.$$
 (2.5.4)

Bei jedem dieser Zerfälle wird ein Elektron und ein Elektron-Neutrino ν_e emittiert. Typischerweise haben diese eine Energie von etwa 1 MeV bis 10 MeV

Reaktorneutrinos eignen sich für experimentellen Physik, da die Reaktoren Neutrinos in einer kontrollierten Menge produzieren und bei Detektion sich dadurch die durch Oszillation verschwindenden Neutrinos auffallen.

2.6 Offene Fragen in der Neutrinophysik

Da Neutrinos nur über die SWW mit anderer Materie interagieren können, gestaltet sich die Bestimmung ihrer Eigenschaften schwieriger als bei anderen Fermionen. Daher bleiben noch einige Fragen über sie unbeantwortet.

• Majorana oder Dirac

 Wenn Neutrinos Dirac-Fermionen sind, verhalten sie sich wie Elektronen oder Quarks, und ihre Antiteilchen unterscheiden sich von den Neutrinos. Falls sie Majorana-Fermionen sind, wären sie ihre eigenen Antiteilchen und würden sich nicht von ihnen unterscheiden. Dies ist grundsätzlich nur für neutrale Teilchen möglich, da die Ladung der Teilchen identisch mit der Ladung des Antiteilchens sein muss [4].

• Sterile Neutrinos

• Sterile Neutrinos sind Neutrinos ν_i , die nicht zu den drei Flavor-Eigenzuständen $\alpha = e, \mu, \tau$ gehören und daher nicht über die SWW mit anderer Materie interagieren. Diese Eigenschaften machen sterile Neutrinos zu potenziellen Kandidaten für dunkle Materie [7].

• Absolute Masse der Neutrinos

 Auch die exakte Masse der Neutrino Flavor-Eigenzustände ist noch unbekannt. Diese lässt sich nicht durch Neutrinooszillationen bestimmen, da diese nur mit den quadratischen Differenzen der Massen arbeitet. Daher werden vielfältige Experimente konzipiert, um die Masse bestimmen zu können [21].

2.6.1 Massenordnung

Die Neutrino Flavor-Eigenzustände setzen sich aus den Massen-Eigenzuständen m_1, m_2 und m_3 zusammen. Im Folgenden wird die Konvention genutzt, dass $m_1 < m_2$ die kleinere Differenz der quadrierten Massen (DQM) beschreibt, dies resultiert, mit nur zwei unabhängigen DQMs, in zwei möglichen Massenordnungen der Massen-Eigenzustände. Die normale Massenordnung (NO) definiert durch m_1, m_2, m_3 in ansteigender Reihenfolge und damit $m_1 \leq m_2 \ll m_3$ und die invertierte Massenordnung (IO) mit $m_3 \ll m_1 \leq m_2$. Die beiden Massenordnungen unterscheiden sich demnach durch das Vorzeichen, der DQM in der m_3 involviert ist.

Hier entspricht $\Delta m_{sol}^2 = \Delta m_{12}^2$ und $\Delta m_{atm}^2 = \Delta m_{31}^2$ (NO) = Δm_{23}^2 (IO). Die Bezeichnungen der DQMs als Solar und Atmosphärisch kommen aus dem historischen Kontext ihrer Bestimmung. Durch die Auswertung früherer Experimente konnten bereits Grenzen für die Massenquadratdifferenzen ermittelt werden:



Abbildung 2.7: Neutrino Massen-Eigenzustände für die NO und die IO (nicht Maßstabsgetreu) [18].

$$\Delta m_{sol}^2 = m_2^2 - m_1^2 = (7,42 \pm 0,21) \cdot 10^{-5} \,\mathrm{eV}^2,$$

$$\Delta m_{atm}^2 = m_3^2 - m_i^2 = \pm (2,510 \pm 0,027) \cdot 10^{-3} \,\mathrm{eV}^2,$$
(2.6.1)

mit i = 1 und dem positiven Vorzeichen bei Δm_{31}^2 für die NO und i = 2 mit dem negativen für die IO. Da die DQMs maßgeblich die Neutrinooszillation bestimmen, lassen sie sich auch über diese messen. Ein ausreichend sensibler Detektor wie JUNO sollte durch Messungen der Neutrinooszillation fähig sein, die Massenordnung zu bestimmen.

Kapitel 3

Jiangmen Underground Neutrino Observatory (JUNO)

Das Jiangmen Underground Neutrino Observatory (JUNO) wurde 2008 vorgeschlagen, um die Massenhierarchie der Neutrinomassen zu bestimmen. Für dieses Experiment werden Reaktor Antineutrinos von den Yangjiang und Taishan Atomkraftwerken (AKWs) [5] verwendet. Der Standort des Detektors befindet sich 53 km entfernt von den beiden AKWs in Kaiping um Jiangmen in der Provinz Guangdong. An diesem Ort befindet sich der Detektor bei 700 m unter der Erdoberfläche. Nach der Zulassung des Projektes in 2013 durch die Chinesische Akademie der Wissenschaften begann die Konstruktion des Detektors in 2015.

3.1 Aufbau des Detektors

JUNO besteht aus einem zentralen Flüssigszintillator-Detektor, ein Cherenkov-Detektor und ein Myon-Tracker. Abbildung 3.1 zeigt den Aufbau der Detektoren.



Abbildung 3.1: Aufbau vom JUNO Detektor

Die Acryl-Sphäre (hier in Grün) hält, mit einem Radius von 17,7 m, 20 kt an Flüssigszintillator (FS). Umgeben ist, der FS von einem Gitter mit einem Radius von 20,05 m. An das Gitter werden die Photoelektronenvervielfachers (PMTs) befestigt. Die PMTs bedecken dabei 80% der Kugeloberfläche. Die gesamte Struktur wird in einem ausgegrabenen Zylinder mit Radius von 21,75 m platziert und der restliche freie Raum wird mit Ultra reinem Wasser gefüllt.

3.1.1 Myon-Veto-System

In Abbildung 3.1 befinden sich im Wasser neben dem zentralen Detektor weitere 2400 PMTs. Das Wasser fungiert damit als Cherenkov-Detektor. Das Ziel des Cherenkov-Detektors ist es, kosmische Myonen zu detektieren. Zusammen mit einem Myon-Tracker, welcher an der Wasseroberfläche montiert ist, bilden sie das Myon-Veto-System. Dieses System wird verwendet, um Neutrino Ereignisse, die Myonen erzeugen, von kosmischen Myonen zu unterscheiden und damit diese Ereignisse aus dem Hintergrund zu entfernen.

3.1.2 Flüssigszintillator (FS)

Beim Flüssigszintillator von JUNO wird Linearalkylbenzol (LAB) verwendet, dies eignet sich aufgrund dessen niedrigen chemischen Reaktivität, Stabilität bei höheren Temperaturen, gute Transparenz und Lichtausbeute. Zu diesem werden in JUNO 2,5 gperL2.5 g/L 2,5-Diphenyloxazole (PPO) als Fluoreszenzmittel und 3 mgperL3 mg/L p-Bis- (o-Methylstyryl)-benzol (bis-MSB) als Wellenlängenumwandler, welche dafür sorgt, dass das emittierte Szintillationslicht die für die PMTs optimale Wellenlänge haben [5].

3.2 Szintillation

Ein Szintillator ist ein Material, welches, wenn geladene Teilchen oder Photonen es durchdringt, es Photonen ausstrahlt. Ein FS-Detektor misst die Energie eines Teilchens durch Photonen, welche am Rand des Detektors von PMTs detektiert werden. Da die PMTs nur Photonen detektieren, muss im Idealen fall die gesamte Energie des Teilchens zu Photonen konvertiert werden. Ein guter FS darf keine Energie permanent absorbieren. Dafür ist das Material des Szintillators auf die Art konzipiert, dass sollte es ionisiert werden oder elastisch gestoßen werden, es die absorbierte Energie als Photon abgibt. Wie schnell der Prozess von Absorption und Immission ist bestimmt, wie lange die Photonen vom Interaktions-Vertex¹zu den PMTs brauchen. Wie die Energie vom Teilchen auf den Szintillators übertragen wird, hängt von dessen Energie und Masse ab. Hier dargestellt durch den relativistischen Geschwindigkeitsfaktor $\beta\gamma$ mit $\beta = v/c$ und dem Lorentz-Faktor γ . Für den Bereich 0, $1 \leq \beta\gamma \leq 1000$ lässt sich der mittlere Energieverlust des Teilchens mit der Bethe-Gleichung

$$\left\langle -\frac{dE}{dx}\right\rangle = K \frac{z^2 Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln\left(\frac{2m_e c^2 \beta^2 \gamma^2 W_{\text{max}}}{I^2}\right) - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right]$$
(3.2.1)

bestimmen [13, S.576]. Dabei sind die Einheiten in Tabelle 3.1 und für W_{max} gilt

$$W_{\rm max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}.$$
(3.2.2)

 $^{^1\}mathrm{Der}$ Ort, an dem die Wechselwirkung mit dem Neutrino stattgefunden hat.

Symbol	Definition	Symbol	Definition
Z	Ladungszahl des Materials	K	$\frac{1}{4\pi N_A r_e^2 m_e c^2}$
z	Ladungszahl des Teilchens	m_e	Elektronenmasse
A	Atomare Masse des Materials	r_e	klassischer Elektronenradius
M	Teilchenmasse	N_A	Avogadro-Konstante
Ι	Mittlere Energie für Ionisation	$\delta(eta\gamma)$	Korrektur des Dichte-Effekts

Tabelle 3.1: Definitionen für die Einheiten der Bethe-Gleichung

Für den in dieser Arbeit betrachteten Energiebereich (1 GeV bis 20 GeV) beschreibt die Bethe-Gleichung den Energieverlust der Myonen. Bei der Bewegung der Myonen durch den Szintillator geben die Myonen ihre Energie hauptsächlich durch elastische Stöße an Moleküle oder Atome des Szintillators ab. Diese werden durch die absorbierte Energie angeregt und emittieren nach kurzer Zeit die Energie in Form eines Photons.

Elektronen oder Positronen werden im Gegensatz zu den Myonen bei diesem Energiebereich nicht korrekt von der Bethe-Gleichung beschrieben. Sie verlieren einen Großteil ihrer Energie durch Bremsstrahlung. Bremsstrahlung ist der Effekt, in dem elektrisch geladene Teilchen durch elektrische Felder abgelenkt werden, und dabei ein Photon abgeben. In der FS geschieht dies durch die Coulomb-Felder der Atome.

Das entstehende Photon führt zu bei ausreichender Energie ($E_{\text{Grenze}} > 1,022 \text{ MeV}$) zu Paarerzeugung ($\gamma \rightarrow e^- + e^+$). Die entstandenen Paare werden ebenfalls gebremst und es entsteht ein Elektroschauer. Er endet, sobald die Energie des Photons geringer ist als E_{Grenze} und nicht mehr ausreicht, um weitere Paare zu erzeugen.

Photoelektronenvervielfacher (PMT)

Photoelektronenvervielfachers (PMTs) sind hochempfindliche Detektoren, die in der Lage sind, einzelne Photonen zu detektieren. Der grundlegende Mechanismus eines PMT beruht auf dem photoelektrischen Effekt, bei dem Photonen auf eine photoempfindliche Oberfläche treffen und Elektronen emittieren. Diese Emission findet typischerweise an der Kathode des Geräts statt, die aus einem Material besteht, das für die Wellenlängen des zu erfassenden Lichtes empfindlich ist.

Sobald ein Photon ein Elektron aus der Kathode freisetzt, wird dieses Elektron durch ein elektrisches Feld beschleunigt. Das beschleunigte Elektron trifft auf ein sekundäres Emissionsmaterial, üblicherweise in Form von Fotokathoden und setzt dort weitere Elektronen frei. Dieser Prozess der Sekundäremission wiederholt sich für die neu losgelösten Elektronen. Die Anzahl der erzeugten Elektronen steigt exponentiell, was zu einem erheblichen Signalverstärkungsfaktor führt.

Das designierte Signal wird schließlich von einem Anodenblech detektiert, wo die gesammelten Elektronen, typischerweise $10 \cdot 10^7$ an dieser Stelle, in eine messbare elektrische Ladung umgewandelt werden.

3.2.1 Ziele des Detektors

JUNOs Hauptziel ist die Bestimmung der Neutrino Massenordnung, für die Reaktorantineutrinos verwendet werden. Die Bestimmung der Massenordnung ist auch über andere Ereignisse möglich. Im Folgenden betrachten wir die Bestimmung der Massenordnung über atmosphärische Neutrinos.

3.2.2 Bestimmung der Massenhierarchie mit atmosphärischen Neutrinos

Folgender Abschnitt beruht auf Quelle [3, S.111ff]. Atmosphärische Neutrinos entstehen durch Zerfallsprozesse in der Atmosphäre (siehe Unterabschnitt 2.5.1). Für Neutrinos, die nicht direkt über dem Detektor entstehen, bedeutet dies bis zu ~ 13.000 km an Materie, die durchdrungen werden muss. Durch die hohe Entfernung und Dichte der Erde ist für diese Neutrinos der MSW-Effekt (siehe Unterabschnitt 2.2.2) relevant für die Übergangswahrscheinlichkeiten.



Abbildung 3.2: Überlebenswahrscheinlichkeit der Elektron-Neutrinos in der NO. Dargestellt in Abhängigkeit von der Energie der Neutrinos E_{ν} und $\cos \theta$ mit dem Einfallswinkel θ und logarithmischen Achsen. Entnommen aus der Quelle [3, S.113]

Abbildung 3.2 zeigt den Einfluss des MSW-Effektes auf die Neutrinos abhängig von ihrem Einfallswinkel θ und ihrer Energie E_{ν} . Für diese Verteilungen wurde angenommen, dass die NO korrekt ist. Für die IO wären die Graphen vertauscht, aufgrund des des Vorzeichenwechsels der Massendifferenz.

Basierend auf diesen Unterschied der beiden Massenordnungen ist die Bestimmung der korrekten Massenordnung möglich. Dafür ist es jedoch nötig, die Energie sowie den Einfallwinkel der Neutrinos zu rekonstruieren und Neutrinos von Anti-Neutrinos zu unterscheiden.

Kapitel 4 Künstliche neuronale Netze

Künstliche Neuronale Netze (KNN) erben ihren Namen von den Nervenzellen des Gehirns und so wie das Hirn lernt, soll auch das neuronale Netz lernen und Zusammenhänge erkennen.

4.1 Prinzip künstlicher neuronaler Netze

KNNs werden dafür verwendet, Muster in Daten zu finden. Der Vorteil der KNNs im Gegensatz zu einem Menschen ist die Fähigkeit gelernte Muster auf tausende Daten zu übertragen oder Muster zu finden, die sich in den Daten verstecken.

4.1.1 Künstliche Neuronen

Viele Neuronen zusammen bilden ein KNN und wie die Nervenzellen des Gehirns bekommen sie Informationen und verarbeiten diese zu einer Ausgabe.



Abbildung 4.1: Caption

Abbildung 4.1 zeigt den Aufbau des *j*-ten Neurons. Das Neuron erhält einen Vektor \vec{x} , welcher mit den Gewichten $\vec{w_j}$ skalar multipliziert wird, nach:

$$net_j = \sum_{i=0}^n x_{ij} \cdot w_i \quad \Leftrightarrow \quad net_j = \sum_{i=1}^n x_i \cdot w_{ij} + \theta_j$$
 (4.1.1)

Die Konvention $x_0 = 1$ gilt hier. net_j stellt die Netzeingabe des *j*-ten Neurons dar, wie in Abbildung 4.1. θ_j ist der Schwellwert oder Bias, ein konstanter Term, der zur gewichteten Summe addiert wird.

Vor der Ausgabe o_j wird die Netzeingabe häufig durch eine Aktivierungsfunktion φ transformiert, um die nicht lineare Beziehung zwischen Eingaben und Ausgaben abzubilden.



Abbildung 4.2: Haupttitel für das gesamte Bild.

Mit den Funktionen aus Abbildung 4.2 definiert durch:

$$\varphi^{step} = \begin{cases} 0 & x < 0\\ 1 & \text{sonst} \end{cases}$$
(4.1.2)
$$\varphi^{slf} = \begin{cases} 0 & x < -0, 5\\ x + 0, 5 & -0, 5 \le x \le \\ 1 & 0, 5 < x \end{cases}$$
(4.1.4)
$$\varphi^{sig}_{a} = \frac{1}{1 + e^{-ax}} , a \in \mathbb{R}$$
(4.1.3)
$$\varphi^{ReLU}_{a} = \max(0, x)$$
(4.1.5)

 $\varphi^{ReLO} = max(0,x)$ (4.1.5) Die Ausgabe o_j wird entweder an das nächste Neuron weitergeleitet, als Teil eines neuen Vektors oder als Teil der Ausgabe des Netzes.

4.1.2 Architekturen künstlicher neuronaler Netze

Die Architektur eines KNNs bezeichnet die Struktur der Verbindungen zwischen Neuronen. Die Wahl der Architektur richtet sich nach dem Anwendungsbereich des Netzes.

In dieser Arbeit werden ausschließlich "feedforward"-Netze betrachtet, bei denen die Neuronen ihre Ausgaben lediglich an die jeweils nächste Schicht weiterleiten.

Abbildung 4.3 zeigt ein vollständig verbundenes KNN. Die erste Schicht ist die Eingabeschicht, die Neuronen dieser Schicht erhalten den Vektor \vec{x} und geben ihre Ausgaben an die erste versteckte Schicht weiter. Die Neuronen der versteckten Schichten empfangen die Ausgaben aller Neuronen der vorherigen Schicht und leiten ihre Ausgaben an die nächste Schicht weiter, bis die Ausgabeschicht erreicht wird. Die letzte Schicht des KNN



Abbildung 4.3: Ein komplett verbundenes KNN mit \vec{x} für den Eingabevektor, i_j für die Eingabe-, h_j^i für die Versteckte- und o_j für die Ausgabeschicht. Das KNN hat insgesamt 6 Schichten.

ist die Ausgabeschicht, welche einen Vektor mit Dimensionen entsprechend der Anzahl der Neuronen der letzten Schicht erzeugt.

4.2 Training

Wenn ein KNN initialisiert wird, werden alle Gewichte \vec{w}^j zufällig generiert, die Bias werden auf null gesetzt. Im Fall von PyTorch, welches für die gezeigten Modelle genutzt wird, verfolgt die Erstellung der Gewichte durch eine Normalverteilung mit dem Mittelwert $\mu = 0$ und der Standardabweichung, $\sigma = \sqrt{n}$ wobei n die Anzahl der Neuronen in der jeweiligen Schicht ist [14].

Ein nicht trainiertes KNN gibt rein zufällige Ausgaben aus. Um die Ausgabe zu verbessern, ohne manuell an Gewichte zu verändern, muss das KNN in trainiert werden.

4.2.1 Trainingsalgorithmen

Beim Trainieren werden, anders als bei nicht maschinellen Lernen, dem KNN keine Regeln direkt beigebracht. Das Netz erhält große Mengen an Eingabe-Daten und zu jedem Vektor der Eingabe-Daten erhält es eine richtige Lösung, die das Netz ausgeben soll. Um zu bewerten, wie falsch das KNN geraten hat und herauszufinden, welche Gewichte zu ändern sind, wir eine Verlustfunktion definiert. Basierend auf dieser Funktion führt man eine Fehlerrückführung durch und optimiert die Gewichte des Netzes:

Verlustfunktion (Loss-Function)

Die Fehlerfunktion vergleicht die Ausgabe des KNN mit den wahren Werten. Wie der Fehler L berechnet wird, hängt vom Ziel des Netzes ab. Für ein Regressionsproblem nutzt man den mittleren Quadratfehler (Mean Squared Error, MSE):

$$L(t,o) = \frac{1}{2} \sum_{i=1}^{n} (t_i - o_i)^2$$
(4.2.1)

mit o als Ausgabe des Netzes, t für die wahren Werte und n für die Anzahl der Ausgabeneuronen.

Für Klassifikationsprobleme mit mehr als zwei Klassen wird die kategorische Kreuzentropie (Categorical Cross-Entropy) verwendet:

$$L(t, o) = -\sum_{i=1}^{n} t_i \cdot \log o_i$$
(4.2.2)

Mit den Variablen aus Gleichung 4.2.1, wobei die wahren Werte in Gleichung 4.2.2 ein Vektor an binären Werten ist. In beiden Fällen ist die Verlustfunktion 0, wenn die Ausgabe dem wahren Wert entspricht. Das Ziel des Trainings ist, diesen Wert zu reduzieren.

Fehlerrückführung (Backpropagation)

Die Verlustfunktion gibt ein Maß zur Bewertung des KNN, um das KNN zu verbessern ist es nötig herauszufinden, welche Neuronen für den Fehler sorgen. Fehlerrückführung leitet die Verlustfunktion nach den Gewichten der Neuronen ab. Um die Gradienten auch für Neuronen, die nicht Teil der Ausgabeschicht sind, werden sie iterativ von der Ausgabeschicht zur Eingabeschicht berechnet.

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}}$$
(4.2.3)

Mit w_{ij} als Gewicht des *j*-ten Neurons der *i*-ten Schicht und der Rest aus Abbildung 4.1.

Optimierung (Optimizer)

In der Optimierung werden die Gewichte angepasst, um die Verlustfunktion zu minimieren. Die Gradienten der Verlustfunktion, berechnet durch die Fehlerrückführung, sind abhängig von den jeweiligen Gewichten. Die Gewichte werden mithilfe des Gradientenabstiegsverfahrens angepasst. Da es sich um ein Minimierungsproblem handelt, zielt das Verfahren darauf ab, für alle Gewichte ein Minimum zu erreichen. Der Gradientenabstieg erreicht dieses Minimum durch:

$$w_{ij} = w_i j - \alpha \cdot \frac{\partial L}{\partial w_{ij}} \tag{4.2.4}$$

Das Gewicht w_{ij} wird in die Richtung angepasst, die die Verlustfunktion am stärksten reduziert. Hierbei bestimmt die Lernrate α , wie stark sich die Gewichte in einer Iteration ändern. Kleine Werte für die Lernrate ($\approx 10 \cdot 10^{-5}$) führen zu einer langsamen Annäherung an das Minimum, wobei die Gefahr besteht, in lokalen Minima stecken zu bleiben. Hohe Werte hingegen können dazu beitragen, lokale Minima zu überwinden, führen jedoch zu einem erratischen Lernverhalten, da sie zu großen Schritten in jeder Iteration führen und damit das Risiko erhöhen, das Minimum zu verfehlen.

4.2.2 Überanpassung und Regularisierung

Beim Testen der Fähigkeiten des neuronalen Netzes während des Trainings nutzt man einen Teil der Daten, mit denen das KNN nicht trainiert wird. Dieser nicht genutzte Datensatz sind die Validierungsdaten. Mit ihnen wird überprüft, ob das KNN in der Lage ist, Daten korrekt zu bewerten, die ihm nicht bereits bekannt sind.

Überanpassung ist der Prozess, bei dem das KNN beginnt, die Trainingsdaten zu genau zu erkennen. Auch wenn das Ziel darin besteht, die Daten möglichst perfekt vorhersagen zu können, beeinträchtigt dies die Leistungsfähigkeit auf Daten außerhalb des Trainingsdatensatzes. Ob ein KNN beginnt, zu überanpassen, lässt sich an dem Verlust der Trainingsdaten im Vergleich zum Verlust der Validierungsdaten erkennen. Sollte der Verlust der Validierungsdaten in jeder Iteration des Training-Loops (Epoche) höher sein als der Verlust der Trainingsdaten, ist es wahrscheinlich, dass das KNN überanpasst.

Es ist möglich, länger zu trainieren und dabei Overfitting zu vermeiden. Mögliche Methoden dafür sind:

- Auslassen (Dropout)
 - Dropout wird als ein Wert von 0 bis 1 definiert. Er gibt die Wahrscheinlichkeit an, dass in einer Iteration über das KNN eine Verbindung zwischen zwei Neuronen nicht genutzt wird. Dies verhindert, dass sich das KNN zu stark auf einzelne Verbindungen zwischen den Schichten verlässt.
- L1/L2-Regularisierung
 - Beide Arten fügen der Verlustfunktion einen Strafterm hinzu.

L1: Verlust = Originaler Verlust –
$$\lambda \sum_{i=1}^{n} |w_i|$$
 (4.2.5)

L2: Verlust = Originaler Verlust –
$$\lambda \sum_{i=1}^{n} w_i^2$$
 (4.2.6)

Beide Regularisierungsarten bestrafen das KNN für hohe Gewichte. Der Parameter λ bestimmt, wie stark die Strafe den Verlust beeinflusst. Die L1-Regularisierung führt dazu, dass viele Gewichte auf null reduziert werden, wodurch sie sich besonders für einfache Probleme eignet. Die L2-Regularisierung hingegen fördert keine verschwindenden Gewichte, sondern sorgt dafür, dass das KNN versucht, die Gewichte möglichst kleinzuhalten, wobei einzelne große Gewichte stark bestraft werden.

- Früher Stopp
 - Eine Methode, die zwar das Überanpassen nicht direkt reduziert, aber das KNN daran hindert, sich zu überanpassen, besteht darin, eine Bedingung für einen frühen Stopp zu implementieren. Dies geschieht, wenn der Verlust der Validierungsdaten beginnt, für längere Zeit größer als der Verlust der Trainingsdaten zu bleiben.
- Batch-Normalisierung
 - Batch-Normalisierung ist eine Form von Regularisierung, welche Eingaben einer Schicht normalisiert. Es wird verwendet um instabile Prozesse einzuschrenken und die Werte der Vektoren in Schach zu halten. Dafür werden in jeder Schicht der Mittelwert μ_B und die Varianz σ_B^2 berechnet, mit

$$\mu_B = \frac{1}{n} \sum_{i=1}^n x_i \text{ und } \sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_B)^2$$
(4.2.7)

Hier sind x_i die Elemente des zu normalisierender Vektors und m die Anzahl der Neuronen in der vorherigen Schicht. Mit dem Mittelwert und der Varianz wird der Vektor normalisiert.

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}},\tag{4.2.8}$$

wobei ϵ eine kleine Konstante ist, um Division durch null zu vermeiden. Es resultiert der normalisierte Vektor \vec{x} welcher mit

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}},\tag{4.2.9}$$

skaliert und verschoben wird. γ und β sind lernbare Parameter, die dem Netz mehr Variation erlauben.

4.3 Evaluierungsmetriken

Die Evaluierung eines neuronalen Netzes erfolgt anhand verschiedener Metriken, die die Vorhersagequalität quantifizieren. Die Wahl der Metrik hängt dabei von der Art des Problems ab, sei es Klassifikation oder Regression. Im Folgenden werden ausschließlich Klassifikationsmetriken betrachtet.

4.3.1 Konfusionsmatrix (KMatrix)

In der Statistik und maschinellem Lernen ist eine Konfusionsmatrix ein spezifischer Tabellenaufbau, der es ermöglicht, die Leistung eines Klassifikationsmodells zu visualisieren. Sie stellt die tatsächlichen gegenüber den vorhergesagten Klassifikationen dar und verfolgt die Anzahl der richtigen und falschen Vorhersagen, jeweils aufgeschlüsselt nach den Klassen. Eine typische Konfusionsmatrix für ein binäres Klassifikationsproblem sieht folgendermaßen aus:



Abbildung 4.4: Konfusionsmatrix mit oben den vorhergesagten Werten und links mit den realen Werten.

Die vier möglichen Ausgänge eines binären Klassifikationsproblems sind:

• Wahre Positive (TP):

 $\circ~$ Die tatsächlich Positiven werden korrekt als positiv
 klassifiziert.

- Wahre Negative (TN):
 - Die tatsächlich Negativen werden korrekt als negativ klassifiziert.
- Falsche Positive (FP):
 - Die tatsächlich Negativen werden fälschlicherweise als positiv klassifiziert.
- Falsche Positive (FN):
 - Die tatsächlich Positiven werden fälschlicherweise als negativ klassifiziert.

Diese Matrix bietet die Grundlage für verschiedene Leistungsmetriken, wie Genauigkeit, Präzision, Sensitivität und F1-Score, welche helfen, die Leistung eines Modells tiefgreifender zu verstehen.

• **Genauigkeit**: Sie gibt den Anteil der korrekt klassifizierten Instanzen an. Formal lässt sich die Genauigkeit definieren als:

$$Genauigkeit = \frac{Anzahl korrekter Vorhersagen}{Gesamtanzahl der Vorhersagen}$$
(4.3.1)

• **Präzision**: Diese Metrik betrachtet den Anteil der korrekt vorhergesagten positiven Klassen unter allen als positiv klassifizierten Ereignisse. Für die positive Klasse gilt:

$$Pr\ddot{a}zision = \frac{Die \text{ wahren Positiven}}{Die \text{ wahren Positiven} + Die falschen Positiven}$$
(4.3.2)

• Sensitivität: Die Sensitivität gibt den Anteil der korrekt erkannten positiven Ereignissen an, bezogen auf alle tatsächlich positiven Ereignisse:

$$Sensitivität = \frac{Die \text{ wahren Positiven}}{Die \text{ wahren Positiven} + Die falschen Negativen}$$
(4.3.3)

• **F1-Score**: Der F1-Score ist eine Kombination von Präzision und Sensitivität dar und wird verwendet, um ein Gleichgewicht zwischen diesen beiden Metriken herzustellen:

$$F1 = \frac{\text{Pr}\ddot{a}\text{z}\text{ision} \cdot \text{Sensitivit}\ddot{a}t}{\text{Pr}\ddot{a}\text{z}\text{ision} + \text{Sensitivit}\ddot{a}t}$$
(4.3.4)

Die Wahl der Evaluationsmetrik muss auf die jeweilige Problemstellung abgestimmt werden, damit sie aussagekräftig ist.

4.3.2 Receiver Operating Characteristic (ROC) Kurve

Die "Receiver Operating Characteristic" (ROC) Kurve ist ein grafisches Werkzeug zur Evaluierung der Leistungsfähigkeit einer binären Klassifikation. Sie stellt die True Positive Rate (TPR) auf der y-Achse gegen die False Positive Rate (FPR) auf der x-Achse für verschiedene Schwellenwerte dar. Ein Klassifikator, der zufällig rät, erzeugt eine diagonale Linie von der unteren linken Ecke zur oberen rechten Ecke der ROC-Kurve, während ein perfekter Klassifikator einen Punkt in der oberen linken Ecke erreichen würde, wo die TPR gleich 1 und die FPR gleich 0 ist.

$$TPR = \frac{\text{Die wahren Positiven}}{\text{Die echten Positiven}}$$
(4.3.5)

$$FPR = \frac{\text{Die falschen Positiven}}{\text{Die echten Negativen}}$$
(4.3.6)

Aus der ROC-Kurve lässt sich die "Area Under the Curve" (AUC) bestimmen. Die AUC bietet eine aggregierte Leistungsmetrik über alle möglichen Klassifikationsschwellenwerte. Ein AUC-Wert von 0,5 deutet auf einen Klassifikator hin, der nicht besser als zufälliges Raten ist, während ein Wert von 1,0 einen perfekten Klassifikator anzeigt. Ein Wert unter 0,5 ist aktiv schlechter als einfaches Raten. Dies macht es jedoch nicht zu einem schlechten Klassifikator. Die Fähigkeit, im Extremfall von einem AUC-Wert von 0,0, alle falschen Positiven ohne ein wahres Positives zu klassifizieren ist ebenfalls eine perfekte Klassifikation. Der Unterschied ist, dass der Klassifikator invers mit der Klasse zusammenhängt, eine Eigenschaft, die ein KNN lernen kann.

4.4 Bayesianische Optimierung

Basierend auf Quelle [17]. In der Initialisierung eines KNN ist es erforderlich, viele Parameter manuell auszuwählen, wie zum Beispiel die Lernrate, die Anzahl der Schichten oder die Größe des Dropouts. Diese manuellen Parameter werden als Hyperparameter des KNN bezeichnet. Im Gegensatz zu den Gewichten oder Bias werden sie nicht während des Trainings gelernt. Da es jedoch auch für die Hyperparameter Werte gibt, die ideal für das KNN sind, ist deren Bestimmung entscheidend für optimale Vorhersagen.

Die Bayesianische Optimierung (BO) versucht, die ideale Kombination von Hyperparametern zu bestimmen. Dies erfolgt durch die Erstellung mehrerer Netze mit zufälligen Hyperparametern, die von einer Zielfunktion, wie der in Abschnitt 4.3, bewertet werden. Basierend auf diesen Bewertungen wird ein weiteres Netz erstellt. Dieses Netz verwendet keine zufälligen Hyperparameter mehr, sondern erhält die Parameter, die am vielversprechendsten für die Optimierung der Zielfunktion sind.

4.4.1 Schritte der Bayesianische Optimierung

- 1. **Definition der Zielfunktion**: Die Zielfunktion $f(\mathbf{x})$ ist eine aus Abschnitt 4.3, wobei \mathbf{x} den Vektor der Hyperparameter darstellt. Die Zielfunktion ist oft teuer zu berechnen, da sie mehrere Trainingsdurchläufe erfordert.
- 2. Wahl eines probabilistischen Modells: Ein Gaussian Process (GP) ist häufig das bevorzugte Modell, um die Zielfunktion f zu approximieren. Der GP wird durch eine Prior-Verteilung definiert, die die Struktur der Zielfunktion beschreibt.

- 3. Akquisitionsfunktion: Die Akquisitionsfunktion bewertet, wie nützlich es wäre, die Zielfunktion an einem bestimmten Punkt \mathbf{x} zu evaluieren. Typische Akquisitionsfunktionen sind:
 - *Expected Improvement (EI)*: Erwartete Verbesserung der Zielfunktion gegenüber dem besten bisher bekannten Wert.
 - *Probability of Improvement (PI)*: Wahrscheinlichkeit, dass der Funktionswert an einem bestimmten Punkt besser ist als der beste bisherige Wert.
 - Upper Confidence Bound (UCB): Ein Kriterium, das sowohl den Mittelwert als auch die Unsicherheit des GP berücksichtigt.
- 4. Optimierung der Akquisitionsfunktion: Der Punkt \mathbf{x}_{n+1} wird bestimmt, welcher die Akquisitionsfunktion maximiert. Dieser Punkt stellt den nächsten Kandidaten für die Evaluierung dar.
- 5. Evaluierung und Update: Die Zielfunktion $f(\mathbf{x}_{n+1})$ wird am ausgewählten Punkt \mathbf{x}_{n+1} durch Training des neuronalen Netzes und ausgewertet. Der GP wird mit der neuen Beobachtung aktualisiert, um das Modell zu verbessern.

Kapitel 5 Simulierte Daten und aggregierte Features

Das Trainieren von neuronalen Netzen erfordert große Datenmengen, um ein effektives Modell zu kreieren. Dazu werden Informationen benötigt über die Art des Ereignisses¹. Informationen, die bei realen Daten nicht bekannt sind. Bzw. ist das KNN dafür gedacht, diese Informationen aus den realen Daten zu bestimmen.

Ein weiterer Faktor ist, dass Algorithmen zur Datenanalyse zum Start der Datenerfassung funktionsfähig sein müssen, damit die Daten ausgewertet werden können. Die Arbeit an Modellen für die Datenverarbeitung beginnt damit meist zeitgleich mit Beginn des Projektes.

Simulationen sind nur nützlich, wenn sie die Realität korrekt nachbilden. Da jeder Detektor andere Umgebungsparameter hat, anders aufgebaut ist und andere Ziele hat, wird eine Simulation für das Projekt maßgeschneidert. Die JUNO Kollaboration simuliert die Ereignisse mittels einer Monte-Carlo (MC) Simulation. Dabei handelt es sich um eine von Daya-Bay [2] wiederverwendete und an JUNO angepasste Simulation.

Parameter	Wert	Parameter	Wert
Target Masse	$20\mathrm{kt}$	1. Birks Factor	$6.5 \cdot 10^{-3} \mathrm{g/cm^2/MeV}$
Target Radius	$17,7\mathrm{m}$	2. Birks Factor	$1.5 \cdot 10^{-6} \left({ m g/cm^2/MeV} \right)^2$
Target Dichte	$20\mathrm{gcm^{-3}}$	Dämpfungslänge	$20\mathrm{m}$
Massenanteil von C	$87,\!92\%$	Absorptionslänge	$20\mathrm{m}$
Massenanteil von H	$12{,}01\%$	Rayleigh-Streuungslänge	$20\mathrm{m}$
Emissionszeit τ_1	$4{,}93\mathrm{ns}$	$ au_1$ -Gewicht	0,799
Emissionszeit τ_2	$20,\!6\mathrm{ns}$	$ au_2$ -Gewicht	0,171
Emissionszeit τ_3	$190\mathrm{ns}$	$ au_3$ -Gewicht	0,03
Optische Deckung	75%	Lichtausbeute	$10.400 { m MeV}^{-1}$
Quantum Effizienz	35%		

Tabelle 5.1: Parameter für die MC Simulation in JUNO.

In Tabelle 5.1 sind die Parameter von JUNOs MC Simulation aufgelistet aus Quelle

¹In dieser Arbeit sind dies die Typen an zu klassifizierenden Neutrinos: NC, ν_e , $\bar{\nu}_e$, ν_{μ} , und $\bar{\nu}_{\mu}$

[3, S.194].

• Target

• Dimensionen des Flüssigszintillators, mit Masse, dem Kugelradius und der Dichte des FS.

• Massenanteile

• Die Massenanteile sind die Prozenteile der Masse des FS, welche durch die jeweiligen Atome kommen.

+ τ Emissionszeit und Gewicht

 $\circ\,$ Die τ Parameter beschreiben die Emission des Lichtes über die Zeit als eine Superposition dreier exponentieller Zerfalle,

$$I(t) = I_0 \cdot \sum_{i=1}^3 w_i e^{-t/\tau_i}$$
(5.0.1)

mit den Emissionszeiten τ_i und den τ -Gewichten w_i .

• Optische Deckung und Quantum Effizienz

- Die optische Deckung bezieht sich auf die Kugeloberfläche des Detektors, die mit PMTs abgedeckt ist.
- Die Quantum Effizienz ist der Anteil an Photonen, die bei optimaler Wellenlänge vom PMT detektiert werden.

• Birks Faktoren

• Im Weg von Interaktions-Vertex zu den PMTs geht Energie verloren, dieser Prozess wird Quenching genannt und wird durch Birks' Gesetz beschrieben,

Quenchte Energie =
$$\frac{\text{Im FS deponierte Energie}}{1 + C_1 \frac{dE}{dx} + C_2 \left(\frac{dE}{dx}\right)^2}$$
, (5.0.2)

hier ist $\frac{dE}{dx}$ die Energieverlustrate und C_1 und C_2 sind die für JUNO spezifischen Birks Faktoren.

• Ausbreitungsparameter

- $\circ\,$ Dämpfungslänge: Entfernung, nachdem die Chance, dass ein Teilchen nicht im FS absorbiert wurde, Detektor auf 1/e gesunken ist.
- \circ Absorptionslänge: Entfernung, nachdem die abgestrahlte Energie auf 1/egesunken ist.
- **Rayleigh-Streuungslänge**: Mittlere Entfernung, nach der ein Photon gestreut wird.
- Lichtausbeute
 - $\circ\,$ Anzahl der erzeugten Photonen pro MeV im FS.

5.1 Wahre und Ereignis Datensätze

Die in der Arbeit verwendeten Daten wurden mit der JUNO-Simulationssoftware [11] generiert. Bei dieser Simulation wird die Interaktion der atmosphärischen Neutrinos mit dem FS, unter Betracht der Wirkungsquerschnitte und den möglichen Wechselwirkungen der entstandenen Sekundärteilchen, simuliert. Der erzeugte Datensatz enthält 19787 Ereignisse und beschränkt sich nur auf atmosphärische Neutrinos im Energiebereich von 1 GeV bis 20 GeV.

5.1.1 Simulierte Ereignisparameter

Die Ereignisparameter sind die wahren Daten, welche zum Erstellen der Labels für das KNN genutzt werden. Dafür werden die Informationen über NC oder CC und die Teilchen ID²genutzt. Die Labels unterscheiden zwischen allen NC Ereignissen und den jeweiligen Neutrino-Typen. Neben den Daten, die für die Labels verwendet werden, enthalten die Ereignisparameter alle Informationen, die das Ereignis beschreiben. Diese werden jedoch im folgenden nicht weiter verwendet. Ein anderer Parameter wurde bereits im Vorhinein zur Begrenzung des Datensatzes verwendet. Bei der Simulation werden nur "komplett enthalten" Ereignisse betrachtet. Sollte ein Ereignis nicht "komplett enthalten" sein im Detektor, ist eine Rekonstruktion der Energie nicht mehr möglich, da ein unbekannter Teil der Energie außerhalb des Detektors abgegeben wurde. Dieser Schnitt der Daten basierend auf einem Wahrheitswert, welcher nicht verfügbar ist bei der Datenerfassung, lässt sich jedoch rechtfertigen, da das ausgetretene Teilchen im äußeren Cherenkov-Detektor erfasst werden.

5.1.2 Simulierte Messdaten

Neben den Ereignisparametern werden die Messdaten des simulierten Detektors bestimmt. Es wird sich auf die Messdaten der 17612 großen Photonenvervielfachers (LPMTs) begrenzt. Diese geben für ein Ereignis jeweils zwei Werte aus. Die im Messzeitraum gemessenen "Number of Photoelectrons" (NPE) und den Moment, in dem das erste Photon das jeweilige LPMT trifft. Die im Weiteren "First Hit Time" (FHT) genannten Werte bestimmt. Zusätzlich zu diesen Daten enthält jedes Ereignis ein.

5.2 Wahl der aggregierten Features

Für das Training des KNN werden die Messdaten nicht direkt verwendet. Der Datensatz ist dafür zu groß³und mit 35224 Werten und damit auch Eingabe-Neuronen pro Ereignis ist die nötige Rechenleistung zu hoch. Um die Daten zu komprimieren und die Menge an nötigen Eingabe-Neuronen zu reduzieren, werden manuell aggregierte Features (Eigenschaften) bestimmt. Das Ziel hier ist, die Daten auf eine Art zusammenzufassen, die dem KNN das Lernen der Muster erleichtert.

5.2.1 "Number of Photoelectrons" Features

Aus den NPE-Daten werden 4 Features erstellt



Abbildung 5.1: Enter Caption

Summierte NPE

Die NPE wird über alle LPMTs summiert. Die Summierte NPE ist die Gesamtmenge an Photoelektronen, die in einem Ereignis detektiert wurden.

Sumierte NPE =
$$\sum_{i}$$
 NPE (5.2.1)

Dieses Feature ist im Fall eines CC-Ereignisses, direkt von der Gesamtenergie abhängig. Dies lässt sich gut in 5.1 erkennen, in dem die summierte Menge der Photoelektronen nach der wahren Energie des Ereignisses aufgetragen sind.

Mittelpunkt der Ladung (ML)

Die Bestimmung des Mittelpunkts der Ladung in einem Neutrinodetektor ist von zentraler Bedeutung für die präzise Rekonstruktion der Energie des Ereignisses.

Die Mitte der Ladung bezieht sich dabei auf den gewichteten Mittelpunkt aller detektierten Photoelektronen (NPE). Damit gibt dieser Punkt eine Näherung an den Vertex der Wechselwirkung. Eine exakte Bestimmung dieses Punktes verbessert die Rekonstruktion der Energie. Wissen über den Vertex ermöglicht es, quenchende Effekte in der Rekonstruktion in Betracht zu ziehen.

Abbildung 5.3 zeigt die Entfernung des Mittelpunkts der Ladung zum Koordinatenursprung für die verschiedenen Klassen. In diesem Fall ist dies der Koordinatenursprung, die der Mittelpunkt JUNOs. Berechnet durch

$$\vec{r}_{\rm ML} = \frac{\sum_{i} \vec{r}_{i} \cdot \text{NPE}_{i}}{\sum_{i} \text{NPE}_{j}},\tag{5.2.2}$$

mit $\vec{r_i}$ für die Position des *i*-ten LPMTs und NPE_i als die Anzahl der Photoelektronen, die das *i*-te LPMT gemessen hat. In Abbildung 5.3 sind die Normen der Vektoren abgebildet.



Abbildung 5.2: Enter Caption



Abbildung 5.3: Beschriftung der Grafik

KAPITEL 5. SIMULIERTE DATEN UND AGGREGIERTE FEATURES

Es ist anzumerken, dass die bestimmte Näherung bereits durch zwei Vorannahmen verfälscht wird. In Abbildung 5.3 ist die maximale Entfernung des Mittelpunkt der Ladung (ML)s zum Mittelpunkt ~ 12 m. Die LPMT befinden sich jedoch ~ 12 m vom Mittelpunkt entfernt. Dies folgt aus dem Ausschluss aller nicht "komplett enthalten" Ereignisse. Interaktionen, die ein ML in dem Bereich von 12 m bis zu den LPMTs stattfinden, werden demnach in diesen Daten nicht auftauchen. Die zweite Vorannahme resultiert aus Gleichung 5.2.2. Für die Berechnung des ML werden die NPE verwendet, welche jedoch von quenchende Effekten beeinflusst werden. Die Gewichtung von LPMTs die weit vom ML entfernt sind, werden durch den höheren Energieverlust beim Weg durch den FS weiter reduziert. Die bestimmten ML befinden sich demnach näher am Mittelpunkt als es in der Abbildung 5.3 scheint.

Neben dem ML sind noch wertere physikalische Eigenschaften interessant, die zu betrachten sind. Unter anderem sollte das Finden eines Maßes für die Ausbreitung des MLs, Informationen über den Typ der Sekundärteilchen geben, bei den verschiedenen Interaktionskanäle Sekundärteilchen entstehen, die Ihre Energie an den FS unterschiedlich übertragen. Es wird der mittlere Abstand, der PMTs zum ML berechnet, indem die Entfernung der PMTs zum ML mit den NPE gewichtet werden.

$$\bar{R} = \frac{\sum_{i} |\vec{R}_{ML,i} - \vec{R}_{i}| \cdot \text{NPE}_{i}}{\sum_{i} \text{NPE}_{j}},$$
(5.2.3)

Mit den bestimmten Mittelwerten lässt sich die gewichtete Standardabweichung des ML bestimmen, welche sich als Ausbreitung des ML interpretieren lässt.



Abbildung 5.4: Beschriftung der Grafik

Die Abbildung 5.4 zeigt die Verteilungen der Standardabweichungen, klassifiziert nach den möglichen Wechselwirkungen. Die es ist zu erkennen, dass die Myon Neutrino Interaktionen eine kleinere Standardabweichung im Verhältnis zu den Elektron-Neutrinos haben.

5.2.2 FHT Features

die FHT-Daten enthalten Informationen über die Geschwindigkeit der Übergabe der Energie an den FS und danach die PMTs.

Die Maximale FHT ist ein Indikator, wie gut der Detektor ausgeleuchtet wurde, da eine große Menge an Photonen die Wahrscheinlichkeit reduzieren, dass ein LPMT nicht getroffen wird.



Abbildung 5.5: Enter Caption

Ähnlich verhält sich auch die minimale FHT

Mit den Daten lässt sich zudem ein Maß für die Ausbreitungsgeschwindigkeit definieren, indem betrachtet wird, wie viele PMTs nach einer bestimmten Zeit ein Photon gemessen haben. Dies zeigt, ähnlich wie bei der FHT_{MAX} , wie viele Photonen im Zeitraum von 0 ns bis 120 ns produziert werden.

5.2.3 Features auf der Ereignis-Ebene

Die bis hier diskutierten aggregierten Features sollen dem KNN die Rohdaten auf eine Art geben, die die Verarbeitung beim Training erleichtern soll. Ereignis-Ebene Features hingegen sind Rekonstruktionen der physikalischen Parameter des Ereignisses.

Wie in Abschnitt 2.3 bereits gezeigt gibt es Wechselwirkungen, die nur für die jeweiligen Teilchen typisch sind (z.B. das Michel-Elektron des Myons). Damit können Ereignis-Features die Klassifikationsfähigkeit der KNN verbessern. In der betrachtend dieser Features ist zu bedenken, dass jedes dieser Ereignis-Features rekonstruiert werden muss, was einen hohen Rechenaufwand benötigt. Es ist also von Interesse herauszufinden, welche der features die KNN am meisten verbesser. und wie realistisch es ist, das Ereignis-Feature genügen präzise zu rekonstruieren.

Die Daten in den folgenden Grafen wurde von der JUNO-Kollaboration simuliert. Für die Simulation wurden Auflösungen von 1 m für, den Vertex und 10 cm für die Neutron/Michel-Elektron Position angenommen. Mit den Definitionen



Abbildung 5.6: Enter Caption



Abbildung 5.7: Enter Caption

Mittelwer
$$(\vec{x}) = \frac{1}{n} \sum_{i=1}^{n} x_i$$
, Max $(\vec{x}) = max(\{x_1, x_2, \dots, x_n\})$,

quadratische Mittel (QMW)
$$(\vec{x}) = \sqrt{\sum_{i=1}^{n} x_i^2}, \quad \text{Summe}(\vec{x}) = \sum_{i=1}^{n} x_i, \quad (5.2.5)$$

wurden insgesamt 26 Ereignis-Features rekonstruiert. Die ersten 6 sind:

- Rekonstruktionen der gequenchten deponierten Energie.
- Die Anzahl an Neutronen.
- Die Anzahl an Michel-Elektronen.
- Die Entfernung des Mittelpunkts der deponierten Energie zu Mittelpunkt von JUNO
- die aufsummierte Entfernung der Neutronen und Michel-Elektronen zum Interaktions-Vertex
- das Verhältnis der Länge des Interaktion-Vertex zu der aufsummierten Entfernung

Die anderen 20 Features sind jeweils die Funktionen aus Gleichung 5.2.5, mit

$$\vec{x} = \begin{cases} |\vec{R}_i - \vec{R}_{Z/q/v}| \\ \theta_i/\varphi_i \end{cases} , \qquad (5.2.6)$$

wobe
i $\vec{R_i}$ die Vektoren zu den Neutronen und Michel-Elektron
en sind. Die übrigen Variablen sind:

• Das Neutronen und Michel-Elektronen Positionszentrum \vec{R}_Z

$$\vec{R}_Z = \sum_{i=1}^n \frac{\vec{R}_i}{n}.$$
(5.2.7)

- Das Zentrum der deponierten Energie \vec{R}_q .
- Der Interaktions-Vertex \vec{R}_v .
- θ und φ sind die Koordinaten der Neutronen und Michel-Elektronen im sphärischen Koordinatensystem mit Mittelpunkt von JUNO als Ursprung.

5.3 Klassifikationsfähigkeit der Features

Ein KNN mit Features arbeiten zu lassen, die keine Klassifikationsfähigkeiten haben, wäre eine Verschwendung an Rechenleistung. Um die Klassifikationsfähigkeit eines Features zu bestimmen, wird die ROC-Kurve und deren AUC-Werte berechnet (siehe Unterabschnitt 4.3.2). ROC-Kurven sind am besten für binäre Klassifikationsprobleme geeignet. Um zu bewerten, ob sich ein Feature für das KNN lohnt, werden die AUC-Werte im quadratischen Mittel und im besten Einzelfall bewertet.

Feature	FHT_{120}	$\mathrm{FHT}_{\mathrm{MIN}}$	FHT _{MAX}	$_{\rm K} {\rm FHT}_{\rm ML}$	$\sum \text{NPE}$	$R_{\rm ML}$	\bar{R}	$\sigma(R_{ m ML})$
QMW AUC	0.5301	0.5256	0.5328	0.5289	0.5616	0.5461	0.5472	0.5494
Größte AUC	0.5596	0.5394	0.5825	0.5317	0.6135	0.5649	0.5660	0.5683
Klasse	NC	$ar{ u}_{\mu}$	NC	$ u_{\mu}$	NC	$ar{ u}_{\mu}$	$ar{ u}_{\mu}$	$ar{ u}_{\mu}$

Tabelle 5.2: Quadratisches Mittel und höchstes AUC. Für das Höchste gibt die Klasse den Typ der Interaktion an.

Feature	QE_{dep}	$n_{ m Neutron}$	$n_{\rm Elektron}$	$d_{QE_{dep}}$, Vert	$d_{ m sum}$	Verhältnis
QMW AUC	0.5588	0.5625	0.7081	0.7944	0.7231	0.5100
Größte AUC	0.6083	0.6699	0.8221	1.0000	0.9901	0.5212
Klasse	NC	NC	$\bar{ u}_e$	NC	NC	$ u_e$

Tabelle 5.3: Caption

Die AUC-Werte der Features werden für jede der 5 Klassen berechnet. alle AUC-Werte welche kleiner als 0.5 folgt

$$AUC' = \begin{cases} AUC & AUC \le 0, 5\\ 1 - AUC & \text{sonst} \end{cases},$$
(5.3.1)

dies ist nötig, da AUC-Werte, die größer und kleiner als 0,5 sind, sich für Klassifikationen eignen. Wird dies nicht getan, werden AUC-Werte, die in invers in einer Klasse gut sind und normal in einer anderen auf 0,5 gemittelt. Von den berechneten AUC-Werte wird der mit der stärksten Ausprägung, zusammen mit der Klasse in der er exzelliert notiert. Das quadratische Mittel wird gebildet und ebenfalls in Tabelle 5.2, 5.3 und 5.4 notiert.

In Tabelle 5.2 ist zu erkennen, dass von den aggregierten Features keins die Klassen signifikant bestimmen kann, mit dem größten Ausreißer von 0,6135 von den, \sum NPE wobei der QMW hier signifikant kleiner ist. Tabelle 5.3 hingegen hat, neben den Neutronen und Michel-Elektronen von den hohen AUC-Werten erwartet waren, Features, die alle NC-Ereignisse perfekt klassifizieren können.

In Abbildung 5.8 wird das Problem klar. Nach dem Histogramm sind die Rekonstruktion Entfernungen der deponierten Energie der NC-Ereignisse zum Interaktions-Vertex im Durchschnitt 70 m und für manche sogar 100 m. Da der JUNO einen weitaus geringeren Durchmesser hat, kann dies keine realistische Messung sein. Weitere Vergleiche mit der JUNO-Kollaboration weißen keine Features dieser Art auf, es scheint ein Fehler in der Auslese der Daten geschehen zu sein. In Tabelle 5.4 finden sich ähnliche Anomalien. Die Rekonstruktion des Vertex R_v scheint fehlerhaft. Alle Ereignis-Features werden im folgenden nicht in den KNN verwendet



Abbildung 5.8: Histogramm der Entfernung des Zentrums der deponierten Energie zu dem Interaktions-Vertex. Die NC-Ereignisse sind dabei unrealistisch weit weg.

Feature			θ			ς	ρ	
Funktion	QMW	Max	Mittel	Summe	QMW	Max	Mittel	Summe
QMW AUC	0.5097	0.5118	0.5799	0.6578	0.5117	0.5212	0.5859	0.6637
Größte AUC	0.5132	0.5175	0.6209	0.7217	0.5201	0.5309	0.6307	0.7384
Klasse	$\bar{\nu}_e$	$ u_e$	$\bar{ u}_e$	$ u_e$	ν_e	$\bar{ u}_e$	$\bar{ u}_e$	$\bar{\nu}_e$
Feature	R_Z				R_q			
RMS AUC	0.7213	0.7223	0.7061	0.6110	0.6300	0.6256	0.6246	0.6122
Größte AUC	0.7722	0.7738	0.7498	0.6553	0.7201	0.7089	0.6686	0.6696
Klasse	$\bar{ u}_{\mu}$	$ar{ u}_{\mu}$	$ar{ u}_{\mu}$	$ u_e$	$\bar{ u}_{\mu}$	$ar{ u}_{\mu}$	$ar{ u}_{\mu}$	$ u_{\mu}$
Feature				F	v_v			
Funktion	QN	IW	Max		Mittel		Summe	
QMW AUC	0.7	072	0.7074		0.7101		0.6949	
Größte AUC	1.0	000	1.0000		0.9999		0.7393	
Klasse	N	C	N N	ſC	NC		ν_e	

Tabelle 5.4: Caption

Kapitel 6 Aufbau und Auswertung der Netze

In diesem Kapitel geht es um die explizite Architektur und den Trainingsprozess der in dieser Bachelorarbeit verwendeten KNN. Die KNN werden komplett in Python programmiert und trainiert unter Verwendung der PyTorch Bibliothek [14]. PyTorch bietet dafür viele nützliche Features an, welche die Implementierung vereinfachen, wie z.B. die Differenzierungsbibliothek namens 'Autograd'. Diese Bibliothek bietet automatische Differenzierung, welches die Fehlerrückführung vereinfacht.

6.1 Netz Architektur

Die Architektur besteht aus einer Serie von vollständig verbundenen Schichten, die durch Batchnorm-Schichten ergänzt werden. Diese Architektur wurde gewählt, da sie eine einfache Implementierung und ein effektives Lernen ermöglicht. Die Netzarchitektur beginnt mit einer Input-Schicht, deren Größe durch die Anzahl der Eingabefeatures bestimmt wird. Darauf folgen mehrere versteckte Schichten, deren Anzahl und Größe flexibel anpassbar sind. Schließlich endet die Architektur in einer Output-Schicht, die die Vorhersage für die Klassen ermöglicht.

6.1.1 Aktivierungsfunktion und Dropout

Nach er Gewichtung durch die vollständig verbundenen Schicht und Normalisierung durch die Batchnorm-Schicht wird auf die Ausgabe die ReLU-Aktivierungsfunktion angewendet. Aktivierungsfunktionen werden generell verwendet, um es dem Netz zu ermöglichen, nichtlineare Zusammenhänge zu lernen. Die spezifische Wahl die ReLU-Funktion zu verwenden basiert auf der Fähigkeit von ReLU, sparsame Darstellung zu erzeugen und den Gradient-Absturz¹ zu vermeiden, was insbesondere bei tiefen Netzarchitekturen von Vorteil ist.

Auf die Aktivierungsfunktion folgt der Dropout. Dropout-Schichten sollen der Überanpassung entgegenwirken und damit die Generalisierungsfähigkeit des KNN verbessern.

 $^{^1 {\}rm verschwindende}$ Gradienten bei der Fehlerrückführung, welches das Training der vorderen Neuronen nahezu stoppt.

6.2 Training und Optimierung

Beim Training der KNN gibt es zwei distinkt unterschiedliche Prozesse. Einmal den normalen Trainings-Loop, welcher in jede Epoche, die das Netz verbessert unter verwendung der gewichtete Kreuzentropieverlust-Funktion und eines Optimierers. Gewichtet um nicht nur die im Datensatz dominierenden Klassen zu lernen. Dieses Taining kann ohne manuelle anpassungen der Hyperparameter nie besser werden als die Architektur des KNN es zulässt. Die zweite ist die übergreifende Bayesianische Optimierung aus 4.4. Die BO trainiert nicht nur das Modell, sie optimiert auch dessen Architektur und Hyperparameter. In jeder Iteration der BO wird ein Modell mit anderen Hyperparametern erstellt, trainiert und bewertet. Die Bewertung ist belibig wählbar und hängt von ziel des KNN ab. In dieser Arbeit wurde der F1-Score (siehe Abschnitt 4.3) verwendet. da für jeden test der Hyperparameter ein KNN trainiert werden muss ist der Rechenaufwand hoch. Um den entgegenzuwirken werden grundsätzlich für die zu trainierenden Parameter Interwalle definiert, auf die sie beschränkt sind. Dazu bietet die für die Optimierung verwendete Bibliotek Optuna" [1] Pruning als Feature. Pruning oder Beschneiden ist in diesem Kontext der Prozess indem die BO wärend des Trainings das KNN mit berreits fertig Trainierten Netzen vergleicht und bewertet ob der Ast, der gerade mit dem KNN erkundet wird vielversprechend ist oder abgeschnitten werden sollt. Mit Pruning lässt sich die geschwindigkeit der BO vervielfachen, es ist jedoch aufzupassen, Pruning in den ersten Iterationen zuzulassen hindert die BO den Hyperparameterraum anständich zu erfassen. Auch in den ersten Epochen des Trainings sollte es verboten sein um vielversprechende netze die einen langsamen oder instabilen Start haben zu bewaren.

Hyperparameter	Zahlenart	Intervall
Anzahl der Schichten	\mathbb{N}	5 bis 15
Neutonen in der ersten versteckten Schicht	\mathbb{N}	32 bis 512
Exponent des Neuronenverfalls	\mathbb{R}	-3 bis 3
Dropout Rate	\mathbb{R}	0,1 bis $0,5$
Lernrate	$\log(\mathbb{R})$	10^{-6} bis 10^{-3}
Batchgröße	\mathbb{N}	32 bis 1028
L2-Norm	$\log(\mathbb{R})$	10^{-6} bis 10^{-2}

Tabelle 6.1: Hyperparameter der BO mit deren Intervallen und Zahlenart

In Tabelle 6.1 wurden die meißten Bereits in Kapitel 4 erklärt, da die Neuronen in der ersten versteckten Schicht (NVS) und der Neuronen zerfall mit der spezifischen architektur zusammenhängen wurden diese nicht erklärt, wobei die Anzahl der NVS selbsterklärend sind. der Neuronenzerfall wird beschrieben durch

$$n_{\text{Neuronen},j} = \begin{cases} \lfloor n_{\text{init}} - (n_{\text{init}} - n_{\text{out}}) \cdot \frac{1 - e^{-k \cdot L_k}}{1 - e^{-k}} \rfloor & k \neq 0\\ \lfloor n_{\text{init}} - (n_{\text{init}} - n_{\text{out}}) \cdot L_i \rfloor & \text{sonst} \end{cases},$$
(6.2.1)

dabei ist $n_{\text{Neuronen},j}$ die Anzahl der Neuronen der *j*-ten Schicht, n_{init} die Anzahl der NVS und n_{out} ist die Anzahl an Klassen in die das KNN klassifiziert. Der exponentielle Zerfallsfaktor aus der BO ist k und L ist die Anzahl der Schichten mit $L_j = j/L$. Die Idee dahinter ist, dem KNN möglichst viel Freiheit in der Bestimmung der Architektur zu geben und gleichzeitig starke Sprünge in der Menge an Neuronen pro Schicht zu vermeiden und die Gesamtmenge der Neutronen zu limitieren.

6.2.1 Optimierte Modelle

Es werden 3 verschiedene Modelle optimiert, die sich in den verwendeten Ereignislevel-Features unterscheiden. Modell 1 verwendet keine der Ereignis-Features, Modell 2 verwendet nur QE_{dep} , $n_{Neutron}$ und $n_{Elektron}$ aus der Ereignis-Features und Modell verwendet alle außer die defekten.

Die BO resultiert in folgenden Parametern:

Hyperparameter	Modell 1	Modell 2	Modell 3
Anzahl der Schichten	5	4	5
NVS	429	230	222
Neuronenverfalls	-0.39	5.2	-2,55
Dropout Rate	0.111	0.13	0.21
Lernrate	0.000637	0.000314	0.000235
Batchgröße	42	175	77
L2-Norm	0.00012	0.0021	0.001696

Tabelle 6.2: Hyperparameter der Modelle nach der BO.

6.3 Evaluierung

Die Parameter an sich haben keine Aussagekraft über das die Klassifikationsfähigkeit der Netze. Es werden die Konfusionsmatrizen verwendet, um ein Bild zu bekommen, wie die Optimierung gelaufen ist. Wir betrachten die Konfusionsmatrix, da diese gut für Multiklassifikationsprobleme sind.

In den folgenden Gleichungen steht e für Elektron, m für Myon, n für Neutrino, a für anti und cc und nc für die Art der Wechselwirkung.

In den Abbildungen 6.1-6.3 ist bereits eine zufriedenstellende Klassifikation der Ereignisse zu erkennen. Abbildung 6.1 weist von den drei dargestellten Modellen die geringste Genauigkeit auf, was zu erwarten war, da hier ausschließlich aggregierte Merkmale verwendet wurden. Dennoch ist die Diagonale bereits klar erkennbar.

Abbildung 6.2 und Abbildung 6.3 ähneln sich stark. Modell 3 scheint jedoch Myon-Neutrinos etwas besser zu klassifizieren als Modell 2 Elektron-Neutrinos oder NC-Ereignisse. Die Trennung zwischen Neutrino und Antineutrino gelingt jedoch bereits bei Modell 2 gut und verbessert sich mit zusätzlichen Ereignismerkmalen nicht signifikant.

Da JUNO lediglich 5000 bis 10.000 atmosphärische Neutrinos zu detektieren erwartet, ist neben der Genauigkeit auch die Sensitivität entscheidend (siehe Abschnitt 4.3), um einen signifikanten Beitrag zur Bestimmung der leisten zu können. Die Modelle werden daher anhand ihres durchschnittlichen F1-Scores bewertet. Modell 1 erreicht einen F1-Score von 0,45; der F1-Score reicht von 0 bis 1, wobei 1 einer perfekten Klassifikation entspricht. Modell 2 erzielt einen F1-Score von 0,66 und ist damit nur geringfügig schlechter als Modell 3, welches einen F1-Score von 0,69 erreicht.



Abbildung 6.1: Konfusionsmatrix des Modell 1, mit auf der x-Achse die vorhergesagten Klassen und auf der y-Achse die tatsächlichen Klassen



Abbildung 6.2: Konfusionsmatrix des Modell 2, mit auf der x-Achse die vorhergesagten Klassen und auf der y-Achse die tatsächlichen Klassen



Abbildung 6.3: Konfusionsmatrix des Modell 3, mit auf der x-Achse die vorhergesagten Klassen und auf der y-Achse die tatsächlichen Klassen

Kapitel 7

Zusammenfassung

Abschließend haben wir in dieser Arbeit die Instrumente und Methoden erkundet, die zur Untersuchung von atmosphärischen Neutrinos mit dem Jiangmen Underground Neutrino Observatory (JUNO) verwendet werden können. Wir haben detailliert beschrieben, wie der Detektor durch seine einzigartige Präzision und Empfindlichkeit zur Entschlüsselung der Massenhierarchie der Neutrinos beitragen kann.

Insbesondere haben wir den Einsatz von Klassifikations-Neuronalen Netzen zur Datenanalyse untersucht. Diese Netzwerke sind aufgrund ihrer Fähigkeit, komplexe Muster zu erkennen und zu klassifizieren, von entscheidender Bedeutung für die genaue Bestimmung der Neutrinooszillationsparameter. Wir haben erklärt, wie diese Netze trainiert und optimiert werden können, um die Vielfalt der Neutrinoereignisse präzise zu identifizieren.

Obwohl das JUNO-Experiment noch nicht mit der Datenerfassung begonnen hat, haben wir umfassend aufgezeigt, wie simulierte Daten zur Vorbereitung und Optimierung der Analyseprozesse verwendet werden. Unsere Untersuchungen durch Bayes'sche Optimierung führten zu vielversprechenden F1-Scores bei den getesteten Modellen, was die Effektivität unserer Methoden bestätigt.

Die durchgeführte Arbeit trägt nicht nur zur Vorbereitung auf die realen Datenerfassungen des JUNO-Detektors bei, sondern liefert auch wertvolle Erkenntnisse für zukünftige Forschungen auf dem Gebiet der Neutrinophysik. Die hier entwickelten Methoden und Modelle könnten die Grundlage für weitere Studien bilden, die darauf abzielen, eine der größten offenen Fragen der modernen Physik, die absolute Massenordnung der Neutrinos, zu lösen.

Literatur

- [1] Takuya Akiba u.a. "Optuna: A Next-generation Hyperparameter Optimization Framework". In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2019.
- F. P. An u. a. "Observation of electron-antineutrino disappearance at Daya Bay". In: *Phys. Rev. Lett.* 108 (2012), S. 171803. DOI: 10.1103/PhysRevLett.108.171803. arXiv: 1203.1669 [hep-ex].
- [3] Fengpeng An u. a. "Neutrino physics with JUNO". In: Journal of Physics G: Nuclear and Particle Physics 43.3 (Feb. 2016), S. 030401. DOI: 10.1088/0954-3899/43/3/ 030401. URL: https://dx.doi.org/10.1088/0954-3899/43/3/030401.
- S. Bilenky. Neutrinos: Majorana or Dirac? 2020. arXiv: 2008.02110 [hep-ph].
 URL: https://arxiv.org/abs/2008.02110.
- JUNO Collaboration. "JUNO physics and detector". In: Progress in Particle and Nuclear Physics 123 (2022), S. 103927. ISSN: 0146-6410. DOI: https://doi.org/10. 1016/j.ppnp.2021.103927. URL: https://www.sciencedirect.com/science/ article/pii/S0146641021000880.
- [6] The ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: (Juli 2012). DOI: 10.1016/j.physletb.2012.08.020.
- Basudeb Dasgupta und Joachim Kopp. "Sterile neutrinos". In: *Physics Reports* 928 (Sep. 2021), S. 1–63. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2021.06.002.
 URL: http://dx.doi.org/10.1016/j.physrep.2021.06.002.
- [8] J. A. Formaggio und G. P. Zeller. "From eV to EeV: Neutrino cross sections across energy scales". In: *Reviews of Modern Physics* 84 (3 Sep. 2012), S. 1307–1341. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.84.1307.
- Y. Fukuda u. a. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (Aug. 1998), S. 1562–1567.
- [10] Mary K. Gaillard, Paul D. Grannis und Frank J. Sciulli. "The Standard Model of Particle Physics". In: (Dez. 1998). DOI: 10.1103/RevModPhys.71.S96.
- [11] Tao Lin u.a. "Simulation software of the JUNO experiment". In: *The European Physical Journal C* 83.5 (März 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11514-x. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11514-x.
- [12] L Michel. "Interaction between Four Half-Spin Particles and the Decay of the -Meson". In: Proceedings of the Physical Society. Section A 63 (5 Mai 1950), S. 514– 531. ISSN: 0370-1298. DOI: 10.1088/0370-1298/63/5/311.

- S. Navas u.a. "Review of Particle Physics". In: Phys. Rev. D 110 (3 Aug. 2024),
 S. 030001. DOI: 10.1103/PhysRevD.110.030001. URL: https://link.aps.org/ doi/10.1103/PhysRevD.110.030001.
- [14] Adam Paszke u.a. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: Advances in Neural Information Processing Systems 32. Curran Associates, Inc., 2019, S. 8024-8035. URL: http://papers.neurips.cc/paper/ 9015-pytorch-an-imperative-style-high-performance-deep-learninglibrary.pdf.
- [15] B. Pontecorvo. "Mesonium and anti-mesonium". In: Sov. Phys. JETP 26 (1968).
 [Zh. Eksp. Teor. Fiz. 53, 1717 (1968)], S. 984.
- [16] Nicolai Popov, William J. Briscoe und Igor Strakovsky. CP Violation Problem. 2024.
 arXiv: 2404.19123 [hep-ph]. URL: https://arxiv.org/abs/2404.19123.
- [17] Bobak Shahriari u. a. "Taking the Human Out of the Loop: A Review of Bayesian Optimization". In: *Proceedings of the IEEE* 104.1 (2016), S. 148–175. DOI: 10.1109/ JPROC.2015.2494218.
- [18] Luca Stanco. The neutrino mass hierarchy from oscillation. 2018. arXiv: 1803.
 08722 [hep-ph]. URL: https://arxiv.org/abs/1803.08722.
- [19] Oleksandr Tomalak u. a. "Radiative corrections to inverse muon decay for accelerator neutrinos". In: *Physical Review D* 107 (9 Mai 2023), S. 093005. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.107.093005.
- [20] Edoardo Vitagliano, Irene Tamborra und Georg Raffelt. "Grand unified neutrino spectrum at Earth: Sources and spectral components". In: *Reviews of Modern Phy*sics 92.4 (Dez. 2020). ISSN: 1539-0756. DOI: 10.1103/revmodphys.92.045006. URL: http://dx.doi.org/10.1103/RevModPhys.92.045006.
- [21] Christian Weinheimer und Kai Zuber. Neutrino Masses. 2013. DOI: https://doi. org/10.1002/andp.201300063. arXiv: 1307.3518 [hep-ex]. URL: https:// arxiv.org/abs/1307.3518.
- [22] L. Wolfenstein. "Neutrino oscillations in matter". In: *Phys. Rev. D* 17 (9 März 1978),
 S. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/ doi/10.1103/PhysRevD.17.2369.
- [23] K. Zuber. Neutrino Physics. 3. Aufl. https://doi.org/10.1201/9781315195612.
 CRC Press, 2020. DOI: 10.1201/9781315195612.

Anhang A

Anhang

Akronyme

- **AKW** Atomkraftwerk. 15
- **BO** Bayesianische Optimierung. 26, 40, 41
- CC Charged Current. 9, 30, 31
- **CP-Verletzung** Verletzung der Ladungs- und Raumsymmetrie (Charge and Parity). 5
- **DQM** Differenz der quadrierten Massen. 13, 14
- FHT "First Hit Time". 30, 34
- **FS** Flüssigszintillator. 15–17, 29, 30, 33, 34
- **IO** invertierte Massenordnung. 13, 14, 18
- JUNO Jiangmen Underground Neutrino Observatory. 15–17, 28–31, 36, 37, 41
- **KNN** Künstliche Neuronale Netze. 19–23, 26, 28, 30, 34, 36, 37, 39, 40
- LPMT großen Photonenvervielfacher. 30, 31, 33, 34
- MC Monte-Carlo. 28
- ML Mittelpunkt der Ladung. 31, 33
- NC Neutral Current. 8, 9, 30, 37, 38, 41
- NO normale Massenordnung. 13, 14, 18
- **NPE** "Number of Photoelectrons". 30, 31, 33
- **PMNS-Matrix** Pontecorvo-Maki–Nakagawa–Sakata-Matrix U_{PMNS} . 5, 11
- PMT Photoelektronenvervielfacher. 16, 17, 29, 33, 34
- SM Standardmodell der Teilchenphysik. 4
- SWW schwache Wechselwirkung. 4, 8, 13

Glossar

- K-Meson Ein K-Mesonen (auch Kaonen genannt) bestehen aus Quarks und Antiquarks, und es gibt zwei Hauptarten: das neutrale K-Meson (K⁰) und das geladene K-Meson (K⁺ und K⁻). Diese bilden sich aus Kombinationen von Up- oder Down-Quarks mit Strange-Quarks. 11
- **Lorentz-Faktor** Der Lorentz-Faktor beschreibt die relativistische Zeitdilatation und Längenkontraktion im Zusammenhang mit der Geschwindigkeit relativ zu Lichtgeschwindigkeit mit $\beta = v/c$. Er wird durch die Formel $\gamma = 1/\sqrt{1-\beta}$ oder $\gamma = E/(m \cdot c^2)$ definiert. 16
- Matter-Antimatter-Diskrepanz Die Materie-Antimaterie-Diskrepanz bezieht sich auf das beobachtete Ungleichgewicht zwischen Materie und Antimaterie im Universum. Nach den Gesetzen der Physik sollten Materie und Antimaterie in etwa gleichen Mengen entstanden sein, doch das beobachtete Universum besteht fast ausschließlich aus Materie. 11
- natürliche Einheiten Die Naürlichen Einheiten sind ein Einheitensystem, bei denen die grundlegenden Konstanten $c = \hbar = 1$ sind. Es wird prominent in der Teilchenphysik verwendet.. 6

A.1 Danksagung

Vielen Dank erstmal an Prof. Dr. Caren Hagner, dass ich in ihrer Arbeitsgruppe schreiben durfte und dass Sie das Erstgutachten bearbeiten und Danke an Dr. Daniel Bick für die Bearbeitung des Zweitgutachtens. Im Weiteren danke ich für die schöne Arbeitsumgebung, die durch alle Kollegen verbreitet wird. Speziellen Danksagungen gehen dabei an Rosmarie Wirth für die Betreuung meiner Arbeit und Malthe Stender für die Unterstützung bis zur letzten Minute. Ich freue mich sehr in dieser Arbeitsgruppe meinen Bachelor geschrieben zu haben und hoffe auf weitere zukünftige Zusammenarbeiten.