1 Introduction

Deep beneath the Amundsen-Scott station on the geographic South Pole, lurks a telescope in the ice. This observatory hunts for some of the most elusive particles in our universe, hoping they may answer foundational questions about nature, and glean insight into new physics that the Genevan collider — or its potential successors — may never be able to.

The IceCube Neutrino Observatory was built between 2005 and 2010, but its heritage stretches back several decades. Spanning one cubic kilometer of Antarctic ice, the telescope consists of some 60 strings, hot water drilled into the ice, with each string containing a number of photomultiplier tubes, capable of detecting photons originating from processes involving neutrinos — the aforementioned elusive particles.

Measurements of certain neutrino properties may help us understand phenomena such as why the universe is matter-dominated[1] and dark matter[2]. IceCube is also capable of searching for point-like sources[3], and serves in the SuperNova Early Warning System[4]. But one thing is the ability to observe derived photons in Antarctic ice; another is reconstruction of this light in order to infer properties of the originating particles. To date, classical methods, such as PegLeg[5] and Retro Reco[*] have mostly been used for this purpose, but the burgeoning proliferation of the use of machine learning (ML) for scientific discovery[6] inspires the development of new methods for this cause.

Whereas classical methods require a hypothesis and thus preset logic, ML "changes these [problems] from logic problems to statistical problems"[†]. That is to say, given a large enough — or rather, a representative enough — dataset, an ML algorithm may find statistical similarities in disparate neutrino events sharing similar underlying properties, giving the network an ability to predict these same properties for previously unseen data. As a "machine learning model is almost like a pure function at inference time"[‡], owing to the universal approximation theorems, it should be possible to construct an ML reconstruction algorithm that is both faster, preciser, more extensible, captures a wider spectrum and is more future-proof than the current methods.

---

[*] https://github.com/IceCubeOpenSource/retro
[†] https://www.ben-evans.com/benedictevans/2019/9/6/face-recognition
[‡] Brennan Saeta, https://www.swiftbysundell.com/podcast/58/

The intention of this thesis is to explore that proposition, specifically for up-going muon neutrinos in the sub-teraelectronvolt range.

This problem is selected as it is the focus of the oscillations group at IceCube, studying neutrino oscillations. The current best reconstruction algorithm for this case, Retro Reco, can take minutes to reconstruct a neutrino, while an ML-based approach can be expected to reconstruct several thousand each second. An additional focus will be the IceCube upgrade[7], which will install additional improved detectors, and for which there is no current functioning reconstruction algorithm.

1.1 Flow

To weave a common thread through this document, and ensure the reader is informed about the big picture throughout, the following section presents the flow of the thesis. The story of an interaction in IceCube from physics to analysis proceeds at a very high level as follows: The universe is created with mechanisms that only allow a certain chirality of neutrinos to interact with forces other than gravity (section 2.1 on page 4). At some point an interaction will create such a particle (section 2.2 on page 4), which oscillates between different flavors (section 2.3 on page 4). As chance allows some neutrinos to interact in or around the instrumented ice, new particles are created that in turn send out a specific form of light (section 2.4 on page 4) which the photomultiplier tubes (section 3.1 on page 5) can detect. Today each in-ice digital optical module contains one photomultiplier tube. With the IceCube upgrade this changes (section 3.7 on page 5), and new optical modules will be fitted with additional photomultiplier tubes which should help with determining the direction of neutrinos. The detector is set up with certain triggers in differing levels (section 3.2 on page 5), to ensure that not just any light is recorded and saved to disk. Even further processing is done, as the bandwidth from pole to America is quite low, and so care must be taken in the choice of what to actually send down the wires for analysis. When the data arrives for analysis, different level of cleaning takes place to root out noise (section 3.4 on page 5). This pipeline also attempts to classify event types, using decision trees, until finally algorithms — none of them machine learning based — reconstruct neutrino properties of interest such as energy and direction (section 3.5 on page 5), and the events at all levels of cleaning are saved in an IceCube proprietary format, called I3 (section 3.6 on page 5). With this part done, we turn towards machine learning generalities. First the basics of machine learning in general (section 4.1 on page 6), then deep learning in particular (section 4.2 on page 6). Special attention is given to temporal convolutional neural networks (section 4.3 on page 6), as this is the architecture used. Because loss functions

are such an important part of any deep learning project, a section is devoted to this concept as well (section 4.4 on page 6), and for the most part, before data can be meaningfully trained in a neural network, some form of data transformation is required (section 4.5 on page 6).

The last part of the thesis concerns the work by the author. The data format provided by IceCube is not machine learning friendly, so a new solution fixing those shortcomings was built (section 5.3 on page 7). So that the reader is on the same page as the author, plots and metrics used for comparison are explained. The opponent algorithm is detailed (section 6.2 on page 8), before the structure of the data itself used in training is explained (section 6.3 on page 8) and the algorithm laid out (section 6.4 on page 8).

2 Particle physics

2.1 The weak interaction

2.2 Neutrino interactions

2.3 Neutrino oscillations

2.4 Cherenkov radiation

3 IceCube

3.1 DOMs

3.2 Triggers

3.3 Post processing

3.4 Levels

3.5 Current reconstruction methods

3.6 I3 file format

3.7 IceCube upgrade

4 Machine learning

4.1 ML basics

4.2 Deep learning

4.3 Temporal convolutional neural networks

4.4 Loss functions

4.5 Data transformation

5 Data

5.1 Monte Carlo simulation

5.2 OscNext

5.3 SQLite

6 Design

6.1 Plots and metrics

6.2 Opponents

6.3 Data structure

6.4 Algorithm

7 Results

8 Summary and outlook

Bibliography

[1] Hiroshi Nunokawa, Stephen Parke, and Jose W. F. Valle, "CP violation and neutrino oscillations," Progress in Particle and Nuclear Physics **60**, 338–402 (2008) (cited on page 1).

[2] Sebastian Baur, "Dark matter searches with the IceCube upgrade," arXiv:1908.08236 [astro-ph, physics:hep-ph] (2019) (cited on page 1).

[3] IceCube Collaboration et al., "Time-integrated neutrino source searches with 10 years of IceCube data," Physical Review Letters **124**, 051103 (2020) (cited on page 1).

[4] Lutz Köpke, "Supernova neutrino detection with IceCube," Journal of Physics: Conference Series **309**, 012029 (2011) (cited on page 1).

[5] Martin Leuermann, "Testing the neutrino mass ordering with IceCube DeepCore," PhD thesis (Rheinisch-Westfalische Technische Hochschule, 2018), 159 pages (cited on page 1).

[6] Maithra Raghu and Eric Schmidt, "A survey of deep learning for scientific discovery," arXiv:2003.11755 [cs, stat] (2020) (cited on page 1).

[7] Aya Ishihara, "The IceCube upgrade – design and science goals," arXiv:1908.09441 [astro-ph, physics:physics] (2019) (cited on page 2).