Machine Learning Techniques in ANNIE

Evangelia Drakopoulou for the ANNIE Collaboration
NPML 2020 – 22 July, Chicago
Outline

- Physics and Technological Goals
- The ANNIE Experiment
- ANNIE Phase II
- High Energy Reconstruction
- Particle Identification & Ring Counting
Physics and Technological Goals
Two main goals:
• Measure the final state **neutron multiplicity** from Charged Current neutrino-nucleus interactions in water
  → Reduce systematic uncertainties on neutrino energy reconstruction in oscillation searches
  → Constrain backgrounds in proton decay searches
• Demonstrate the use of fast-timing Large Area Picosecond PhotoDetectors (**LAPPDs**) for event reconstruction

[arXiv:1603.01843 [physics.ins-det]]
The ANNIE Experiment
The ANNIE Detector

- **Accelerator Neutrino Neutron Interaction Experiment (ANNIE):** A 26-ton Gd-doped water Cherenkov detector installed in the Booster Neutrino Beam at Fermilab (flux peaks at 600 MeV).

![Diagram of the ANNIE Detector](image)
**Status of the Experiment**

**ANNIE – Phase I**
- Neutron background measurements in the detector site in Fermilab.
- Successful operation phase of the detector.

**ANNIE – Phase II**
- Measure the neutron yield from CCQE events in water
- First deployment and use of LAPPDs
ANNIE Phase II
• 3 m x 4 m tank filled with Gd (0.2%) loaded water

• ~132 PMTs + 5 LAPPDs:
  → LAPPDs will be placed downstream
  → Flexibility to add additional LAPPDs

• Fully instrumented MRD
  → 11 layers and 310 channels

• Upgraded electronics and readout
1. Charged Current neutrino interactions in fiducial volume
   → Cherenkov cone incident on PMTs and LAPPDs
   → Scintillation light from stopping muons in MRD
2. Final state neutrons thermalised and captured in Gd
3. Cascade of 8 MeV detected by PMTs
LAPPDs – A new technology tested in ANNIE

Micro-channel plate, fast-timing photodetectors
- Large-area: 20 × 20 cm
- Fast timing: <100 ps for a single photoelectron
- High quantum efficiency (QE): >20 %
- Position resolution: sub-mm
- Operable in a magnetic field

ANNIE Analysis flowchart

- MRD/FACC data
  - MRD/FACC efficiency
    - Track reconstruction
    - Vertex reconstruction
    - High energy reconstruction
    - Ring counting
    - Particle identification
  - Selected neutrino events
    - Machine Learning Techniques

- Data cleaning
  - Low energy reconstruction
  - Neutron detection efficiency
  - Neutron containment
  - Neutron tagging
  - Evaluate systematics
  - Neutron multiplicity measurement

- Event builder
- Cluster finder
- Tank PMT data

All events

Credit: V. Fischer
High Energy Reconstruction
Events with reconstructed vertex & direction

DNN to reconstruct the track length in the water tank

BDTG to reconstruct the muon energy
Track Length Reconstruction

- The track was initially reconstructed as the distance between the first and last Cherenkov photon emission point along the track (L).

- In order to improve the reconstructed track length in the water tank a Deep Learning Neural Network (DNN) (in Keras) was used.

- The parameters were optimised using GridSearchCV.

Input variables for DNN:
- All Cherenkov photons emission points
- Hit times
- The previously estimated track length, L
- The total number of hits in PMTs and LAPPDs for each event.

For more details on the reconstruction see: arXiv:1710.05668v3
Track Length Reconstruction

\[ \Delta R = L_{\text{reco}} - L_{\text{MC}} \]

- Mean: 13.11, Std: 23.70
- Previous: Mean: 46.12, Std: 24.43

\[ \Delta R = L_{\text{reco}} - L_{\text{MC}} \text{ [cm]} \]

~~~

\[ \text{NEW} \]

\[ \text{Previous} \]

\[ \text{DNN-CC0}\pi \]

~ 1000 events used for the track length prediction and 1000 for the DNN training
Muon Energy Reconstruction

- The track length in water and the track length in the MRD are used among other variables as inputs to a Boosted Decision Tree (BDTG) to reconstruct the muon energy.
- The parameters were optimised using GridSearchCV.
- Results for: Reconstructed CCQE events with $E_{\nu}<2\text{GeV}$ passing through the MRD.
- Figure of merit: $\Delta E/E = 100 \times \left( E_{\text{MC}} - E_{\text{reco}} \right) / E_{\text{MC}}$

Input Variables for BDTG:
- Track length in water
- Reconstructed track length in the MRD
- Angle difference between the reconstructed z direction and the beam direction at (0,0,1)
- The total number of hits in PMTs and LAPPDs for each event
- The reconstructed vertex coordinates
- The distances of the reconstructed vertex from the detector walls ($D_R$, $D_y$)
Muon Energy Reconstruction

For more details on the reconstruction see: arXiv:1710.05668v3

95.4% of events in $\Delta E/E < 10$
4.3% of events in $10\% \leq \Delta E/E < 20$
0.3% of events in $\Delta E/E \geq 20$

For more details on the reconstruction see: arXiv:1710.05668v3
Particle Identification & Ring Counting

Multi-Layer Perceptron NN by Michael Nieslony
Classification Tasks in ANNI

**Ring Counting**
- Classification: single- vs multi-rings
- Problems in a small (ANNIE-like) Water-Cherenkov detector: Cherenkov rings not well separable if direction similar
- Event sample: BNB simulation sample (GENIE+WCSim)
- Methods: Multi-Layer Perceptron, Convolutional Neural Network

**Particle ID**
- Classification: electrons vs. muons
- Problems in a small (ANNIE-like) Water-Cherenkov detector: Observe Cherenkov disks instead of rings → very similar
- Event sample: Electron + muon simulation sample (WCSim)
- Methods: Multi-Layer-Perceptron, Convolutional Neural Network

**Multi-ring event**

**Single-ring event**

**500 MeV electron**

**500 MeV muon**
Multi-Layer Perceptron Structure

**Software package:**

- MLPClassifier from `sklearn.neural_network` package
- Activation Layer: RELU
- Training set / test set split: 60%/40%

**Input elements:**
- Physics-driven mean properties from PMT charge & time values
- Additional input variables from LAPPDs available

**Hidden layers:**
- Two hidden layers
- Respective width of 100 nodes

**Output Layer:**
- Binary problem (muons/electrons)
- Pred. Prob ($\mu$) = 1 - Pred. Prob. (e)

**Output:**
- muon probability
- electron probability
Particle ID - Multi-Layer Perceptron

**Input variables - Particle ID:**

- Total recorded charge
- Fraction of downstream charge
- Fraction of highest charge PMT
- Number of hit PMTs
- Variance of time profile
- Fraction of PMT hits at large angles
- Variance of angular profiles
- Number of LAPPD hits
- Total recorded LAPPD charge
- Number of MRD clusters
- Number of paddles/layers in MRD clusters
- + a couple more similar variables

**Dataset information:**

- ~60,000 muons & ~60,000 electrons simulated with beam-like properties
- Require event vertex in Fiducial Volume
- Minimum charge for PMT: 10 p.e.
A few correlations between input variables:

Clearly visible overlaps in the input parameter space.
Particle ID - Multi-Layer Perceptron

Performance:
Recall (muons): 94%
Recall (electrons): 89%
Overall accuracy: 91%

Energy dependence:
### Input variables – Ring Counting:

- Total recorded charge
- Fraction of downstream charge
- Fraction of highest charge PMT
- Number of hit PMTs
- Variance of time profile
- Fraction of PMT hits at large angles
- Variance of angular profiles
- Number of LAPPD hits
- Total recorded LAPPD charge
- Number of MRD clusters
- Number of paddles/layers in MRD clusters
- + a couple more similar variables

### Dataset information:

- Complete simulation of BNB neutrino interactions in ANNIE → ~67,000 single-rings + ~39,000 multi-rings
- Require event vertex in Fiducial Volume
- Minimum charge for PMT: 10 p.e.
A few correlations between input variables:

Even stronger overlaps in the input parameter space.
Ring Counting - Multi-Layer Perceptron

Performance:

Recall (1-ring): 88%
Recall (multi-ring): 78%
Overall accuracy: 83%

Energy dependence:

Single-rings better classified at lower energies
Multi-rings better classified at higher energies

Multi-ring events favored for high-charge events

Ring Classification performance highly correlated to pion energy
Particle Identification & Ring Counting

CNN by David Maksimovic
Particle ID - CNN

Time values are flipped for the network (high values mean earlier arrival time)
Distinguishing Features:

• The muon creates a well-defined Cherenkov cone

• The multiple electrons from the electron shower give rise to multiple overlapping Cherenkov cones leading to a diffuse disk

20k Test Set
Particle ID - CNN

Visible Energy

PMT:
We see that the Network has classification problems for low light electron events

Muons are better classified in the relevant energy range from 0 to 1500 MeV

If a selection for at least one MRD clusters is chosen, the accuracy of the CNN increases to 99%
Ring Counting - CNN

Single Rings

Multi Rings

Charge

Normalized charge

Normed charge

Phi

Theta

Time

Normed time

Early

Late

Muons

Normed time

1.0

0.0

-150  -100  -50    0    50    100   150

-150  -100  -50    0    50    100   150

-150  -100  -50    0    50    100   150

-150  -100  -50    0    50    100   150
Ring Counting - CNN

10k Test Set

Ring Counting 83% acc on testset
PMT+LAPPD 5x5

True label

SR
0.934
0.066

MR
0.317
0.683

Predicted label

Preliminary

NN Prediction Distribution

Counts

10^{-1}

10^{0}

10^{1}

10^{2}

10^{3}

0.0

0.2

0.4

0.6

0.8

1.0

Probability

True SR

True MR
Ring Counting - CNN

Visible Energy

For the CNN, low energy pions look like background within a single ring event rather then multi ring components.

Neutron produces within the interactions do not influence the CNN.

We see that the Network has only classification problems for high energetic single ring events.
Schematic for best PID- and RC-model:
Kernels per conv. layer: 400
Kernel size: (3x3)
Double conv. layers: 3
Dense layers: 2

Batch normalisation + Dropout 20%

# of kernels
Feature Maps
Feature Maps
Max-pooled Maps
Feature Maps
Feature Maps
Max-pooled Maps

Flattening
Width= 500

# of dense layers

Input layer

Kernel size
Conclusions

- ANNIE will measure the neutron yield as a function of the momentum transfer from neutrino-nucleus interactions in water.
- To fulfill its scientific goals ANNIE will use LAPPDs and Gd-doped water.
- Machine learning techniques are widely used in ANNIE improving the event reconstruction.
- DNNs and BDTGs are used for the energy reconstruction and MLP and CNN are used for the particle identification and the ring counting.
Thank you!
Backup
As neutrino-antineutrino event-rate comparisons are important for $\delta_{\text{CP}}$ measurements, the relative neutron composition of final hadronic states is significant.

NuSTEC white paper

**Signal/Background separation:**

Multiplicity and absence of neutrons is also a strong handle for signal-background separation in a number of physics analyses!

ANNIE Letter of Intent

arXiv:1707.08222 [physics.ins-det]

Atmospheric neutrino interactions in water may produce final state neutrons!

The knowledge of neutron yield will reduce background for:

Proton Decay searches, Diffuse Supernova Background (DSNB)
Incom Inc has now produced multiple LAPPD devices [http://www.incomusa.com](http://www.incomusa.com)

- A number of tiles have been produced and tested → gain, timing and QE
- Purchased tile #25 from INCOM
  → Thorough testing ongoing at ISU
  → Expected to be deployed in ANNIE Phase II

**Gain - in units of elementary charge**

- single-PE gain: \(2.54 \times 10^6\)

**Transit Time Spread**

- time resolution: \(64\) psec
LAPPDs enable the ANNIE physics:

- Neutrons created in ANNIE can drift up to 2 m:
  - drift is symmetric in the direction transverse to beam
  - drift is mostly forward in the beam direction
- Given ANNIE’s small size it is crucial to maximize the fiducial volume
- A vertex resolution of ~10 cm is needed to properly identify events in the fiducial volume.

- Such resolution is beyond the capability of traditional PMTs!
- Precise timing-based reconstruction enabled by LAPPDs is essential.
LAPPDs – A new technology tested in ANNIE

- Glass body, minimal feedthroughs
- MCPs made using atomic layer deposition
- Transmission line anode
- Fast and economical front-end electronics
- Large area, flat panel photocathodes
LAPPD detectors:
• Thin-films on borosilicate glass
• Glass vacuum assembly
• Simple, pure materials
• Scalable electronics
• Designed to cover large areas

Conventional MCPs:
• Conditioning of leaded glass (MCPs)
• Ceramic body
• Not designed for large area applications
The LAPPD Concept

Microchannel Plate (MCP):
- a thin plate with microscopic (typically <50 μm) pores
- pores are optimized for secondary electron emission (SEE).
- Accelerating electrons accelerating across an electric potential strike the pore walls, initiating an avalanche of secondary electrons.

- An MCP–PMT is, sealed vacuum tube photodetector.
- Incoming light, incident on a photocathode can produce electrons by the photoelectric effect.
- Microchannel plates provide a gain stage, amplifying the electrical signal by a factor typically above $10^6$.
- Signal is collected on the anode.
Momentum Transfer: $Q^2$

- Momentum transfer for CCQE events: the primary interaction channel in ANNIE
- CCQE events are completely described by the energy of the incoming neutrino and the energy and momentum of the outgoing muon.

$$Q_{QE}^2 = 2E_{\nu}^{QE}(E_\mu - p_\mu \cos \theta_\mu) - m_\mu^2$$

- 1-$\sigma$ $Q^2$ resolution for four bins in true $Q^2$
- The addition of LAPPDs considerably improves the $Q^2$ resolution.
Vertex and Track Reconstruction

Steps:

1. “Simple vertex” fit $\rightarrow (x, y, z, t)$
   - Consider a point source at a hypothesised location, emitting Cherenkov light
   - For each hit calculate the timing residual and timing Figure of Merit (FOM)
   - Adjust the four hypothesised parameters to maximise FOM

2. Extended vertex fit $\rightarrow (x, y, z, t, \theta, \varphi)$
   - Start with position from simple vertex fit and add hypothesised track direction
   - For each hit calculate extended time residual including muon travel time
   - Calculate cone FOM by comparing predicted to measured Cherenkov cone
   - Adjust all six parameters to maximise total FOM (time FOM + cone FOM)