Survey: Status of A.I. in Experimental Analysis

Yue Shi Lai
Nuclear Science Division

Thanks to: Wahid Bhimji, Eric Church, Mario Cromaz, Jacob Daughhetee, Markus Diefenthaler, Andrey Elagin, Cristiano Fanelli, Gian Michele Innocenti, Gaosong Li, Kate Scholberg, Nu Xu, Liang Yang
Disclaimer

- Both A.I. and NP are highly diverse fields
- The materials I will cover is largely limited to what I have received as response
- I will only briefly mention edge-cases such as Bayesian optimization
Overview, Terminology

**Artificial Intelligence (A.I.),** Andrew Moore 2017:

“Artificial intelligence is the science and engineering of making computers behave in ways that, until recently, we thought required human intelligence.”


**Machine learning (ML),** Arthur Samuel 1959:

“Programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort.”

A. Samuel, IBM J. Res. Dev. 3(3), 210–229 (1959)

**(Artificial) neural network [(A)NN],** Warren McCulloch, Walter Pitts 1943 (many thereafter):

Computation model loosely modeled after the biological axon-synapse-dendrite connections

**Why is it useful?**

- Eliminates the need for human programming
- Outperforms programming with structured human knowledge (“symbolic AI”)
A.I. vs. ML vs. NN

- **A.I.** (moving target)
- **ML** (expanding)
- **NN** (Regression)
Y. Lai (LBNL NSD)
Status of A.I. in Experimental Analysis
AI landscape

Supervised learning

Regression

Classification

Semi-supervised learning

Reinforcement learning
AI landscape

Supervised learning

- Calibration
- Regression
- Fast function evaluation

PID
- Classification
- Sig/bkg

Semi-supervised learning

Reinforcement learning
AI landscape

Unsupervised learning

Reinforcement learning

Semi-supervised learning

Clustering

Density estimation

Dimensionality reduction

Unsupervised learning
AI landscape

Unsupervised learning

Semi-supervised learning

Reinforcement learning

Density estimation

Generative

Dimensionality reduction

Clustering

Latent space for classification

Unsupervised learning
Use of AI in NP analyses

<table>
<thead>
<tr>
<th>Method</th>
<th>Heavy-ion</th>
<th>$0\nu\beta\beta$/neutrinos</th>
<th>Photoproduction</th>
<th>$\gamma$-ray tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calib.</td>
<td>✔ (TPC, jets, kin.)</td>
<td></td>
<td>✔ (RICH)</td>
<td></td>
</tr>
<tr>
<td>Multiv. Cut</td>
<td>✔ (HF)</td>
<td>✔ (PID)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN (superv.)</td>
<td>✔ (PID)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graph convol.</td>
<td>✔ (PID)</td>
<td></td>
<td></td>
<td>✔ (tracking)</td>
</tr>
<tr>
<td>RNN</td>
<td>✔ (PID)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-sup.</td>
<td>✔ (jets)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generative</td>
<td>✔ (jets)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimens. red.</td>
<td>✔ (PID)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
AI tools in NP analyses

**Infrastructure**

BDT

TMVA

XGBoost

**NN**

Keras/TensorFlow

PyTorch

SparseConvNet

Scikit-Learn

**Horovod**
Supervised: Calibration

ALICE, TPC charge distortion

Large distortions are expected in LHC Run-3 due to the presence of a intense flux of ions produced un-ionization processes:

- very slow ion drift speed (~0.16s compared to ~92μs for electrons)
- not uniform electric fields cause deviations from the expected field lines

Deep neural network algorithms being studied to perform fast calculation of the distortion maps based on the expected digital current measured by TPC readout

CERN-LHCC-2013-020 / ALICE-TDR-016

Y. Lai (LBNL NSD)
Status of A.I. in Experimental Analysis
Supervised: Calibration

Subtraction of underlying-event energy for jets at low $p_T$ very challenging in heavy-ion collisions:
  - large region-to-region fluctuations

Traditional techniques: event-by-event subtraction of the average energy density + unfolding

ML techniques for jet-by-jet background subtraction:
  - exploit the different properties of particles belonging to the event background and those belonging to the jet (jet constituents)
  - Regression algorithms based on shallow neural networks, random forest and linear regression

→ push jet measurements in heavy-ion down to lower momenta

Y. Lai (LBNL NSD)
Status of A.I. in Experimental Analysis

Rüdiger Haake, Constantin Loizides

Supervised: Calibration

EIC

Use both electron and hadronic final state (overdetermined)

Optimally determine ep collision kinematics

Abdullah Farhat, Yuesheng Xu (Department of Mathematics, Old Dominion University)
Markus Diefenthaler, Dmitry Romanov, Douglas Higinbotham (EIC Center, Jefferson Lab)
Andrii Verbytskyi (Max-Planck-Institut für Physik)
Bayesian optimization: Calibration

GlueX, Cristiano Fanelli

Real Offsets
3-seg mirror: 
$\theta_x, \theta_y, \theta_z = (0.25, 0.50, 0.15)$ deg, $y = 0.5$ mm; 
$\text{bar } z = 2.0$ mm;  
PMT $(r, \theta) = (1.5$ mm, $1.0$ deg)

Minimum at 3-seg mirror: 
$\theta_x, \theta_y, \theta_z = (0.2485, 0.5832, 0.1171)$ deg, 
$y = 0.5894$ mm;  
$\text{bar } z = 2.0788$ mm;  
PMT $(r, \theta) = 1.8690$ mm, $1.3544$ deg

Recipe: For each call of the optimizer, M offset points are explored using N different particles (for each call). The total number of calls is $T = 120, M = 10, N = 125$

Particles used = 15000
Points explored = 1200

FoM = LogL normalized to a default alignment

3-seg mirror offsets (most critical for alignment) found within the tolerances.
Supervised: Cut-based reco./analyses

ALICE, Heavy Flavor ($\Lambda_c$)

Gian Michele Innocenti

BDT techniques (XGBoost) have been exploited to select heavy-flavor decays in central PbPb collisions:

- E.g. $\Lambda_c$ selection ($c\tau \sim 50\mu m$)
- TOF and TPC variables used in the training to exploit correlations with topological selections
- increase in significance up to 3-4 at low $p_T$ with respect to standard analysis techniques
Supervised: Cut-based reco./analyses

STAR, also HF/Λ_c, Nu Xu

**QM 2017 —> QM 2018:**
- Total statistics doubled with combined 2014+2016 data
- More than 50% improvement in signal significance using Boosted Decision Trees!
  Effectively providing 2x additional data!
Supervised: Cut-based reco./analyses

EXO-200

$0\nu\beta\beta$ with background from detector material

2018 result using BDT: 15% better than energy-only

https://doi.org/10.1103/PhysRevLett.120.072701

2019 result using BDT: 25% better than energy-only

https://doi.org/10.1103/PhysRevLett.123.161802

Liang Yang, Gaosong Li

Y. Lai (LBNL NSD)

Status of A.I. in Experimental Analysis
Supervised: Image-based (CNN)

- Greatly improved background rejection when used on simulated events
- A. Li and C. Grant (Boston University) are implementing this network model and a more sophisticated spherical CNN on the KamLAND-Zen and SNO+ experiments

**A. Li et al. NIM A 947 (2019) 162604**

Time evolution of the PMT hits for a simulated 0νββ event and 10C background event at the center of the KamLAND-Zen detector

KamiLAND-Zen, Christopher Grant
Supervised: Image-based (CNN)

NEXT

2D CNN classification of Xe TPC
Planned semantic segmentation of decay candidates
 Runs on OLCF Summit with PyTorch + SparseConvNet and Horovod

Eric Church

FRIIB Hall B
Michelle Kouchera

NEXT’s 3D images – lend themselves naturally to DNNs

Two-neutrino double beta decay candidates

NEXT-White data
Topologically identified and energy-separated from double escape peaks

PRE-TRAINED ON IMAGENET DATA!
• Searching for rare events in liquid noble gas detectors requires excellent background rejection.
• Backgrounds (gamma rays and betas) produce differently shaped waveforms than CEvNS signals (nuclear recoils).
• **Convolutional Neural Network** can be trained on waveform images to provide classification.
• Provides a tool to discriminate event types at detector-thresholds where conventional analysis fails.

Jacob Daughhetee, U. of Tennessee, Knoxville
Supervised: Graph convolution

GRETA
Gamma-ray tracking
Tracks from multiple-scattering
Several recent technologies considered:
Bayesian search for Ge-transport + CNN (for interaction point)
Graph convolution (for tracking)
Supervised: RNN

LAPPD for THEIA $0\nu\beta\beta$, Andrey Elagin

Performance for accepting $0\nu\beta\beta$ (positive) and rejecting $^8\text{B}$ background (negative)

LSTM used due to high sparsity/mostly empty pixels

Traditional rejection only works with known vertex

double beta decay

AUC = 0.78

Performance for accepting $0\nu\beta\beta$ (positive) and rejecting $^8\text{B}$ background (negative)

LSTM used due to high sparsity/mostly empty pixels

Traditional rejection only works with known vertex
Generative: Simulation

LHC ATLAS, Wahid Bhimji (NERSC)

- Particle physics uses detailed micro-physics detector simulations (e.g. with Geant4)
  - >~50% LHC computing budget (10^9 CPU hours)
  - Much of this compute time in calorimeter ‘shower’
- CaloGAN models a 3-layer calorimeter detector inspired by that of the ATLAS LHC experiment
- Custom NN design
  - sparsity
  - high dynamic range
  - highly location-dependent features


Application for EIC?
Generative: Automated analysis

LBNL LDRD, YSL, Mateusz Ploskon, Felix Ringer

- Discovery of new observables by NN (arXiv:1810.00835)
- Discovery of theoretical models via automated analysis
  - A generator that mimicks the quark/gluon passing the plasma, and learns data-driven from experiment how to parton-shower and hadronize
  - A discriminator looks at both the output of the generator, and real measurement data, and decide if this was simulated or is the reality/reference

Experimentalist NN learning how to measure plasma with Jewel

NN reduced to 13 algebraic terms
Fanelli & Pomponi proved that deepRICH can reach the PID performance of established algorithms. This depends only on the available resources for improvement by ~4 orders of magnitude the reconstruction time.

<table>
<thead>
<tr>
<th>Kinematics</th>
<th>DeepRICH</th>
<th>FastDIRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 GeV/c</td>
<td>99.74</td>
<td>98.18</td>
</tr>
<tr>
<td>4.5 GeV/c</td>
<td>98.78</td>
<td>95.21</td>
</tr>
<tr>
<td>5 GeV/c</td>
<td>96.64</td>
<td>91.13</td>
</tr>
</tbody>
</table>

Cristiano Fanelli

Y. Lai (LBNL NSD)
Status of A.I. in Experimental Analysis
Challenges

Methodological challenges using current AI techniques:

• How to find the least complex model that fits the job description? “Model parameter efficiency”

• Uncertainty quantification (UQ): Reliability data vs. MC and how to avoid rare, but catastrophically wrong results

Infrastructure challenges:

• Lack of GPU at high-throughput computing (HTC) facilities, HTC projects require separate high-performance computing (HPC) allocation for GPU training
Opportunities for NP

Area where NP benefits greatly from AI application:

• Fast turn-around using simulation, and without investing in many person-years of custom algorithm development

• Performance that are difficult/humanly impossible to achieve with manually crafted algorithms

• Fast reconstruction without manually tuned GPU/FPGA code

• Many experiments already benefit from off-the-shelf methods
Opportunities beyond NP

Some of the projects/goals are unique to NP and solves challenges that are cross-cutting to other fields

- Large (measurement) sample size allowing generative models, automated learning of underlying physics principles.
- Some of the data sets are not first-principles calculable: strong interaction physics, many-body physics
- Problems demanding innovative reconstruction algorithms

Cross-cutting goals where collaboration with wider AI field would be beneficial:

- Study of physics behavior: NN with built-in known symmetries/conservation laws
  - Interpretability
  - Data efficiency
Summary

• There is already wide adoption of AI techniques in NP
• AI for NP promises:
  - Fast development cycle
  - Efficient utilization of commercial hardware
• It comes with a few pitfalls
• There is uniqueness of NP and the dataset we are dealing with, for AI application
  - Potential synergies and collaboration with the wider AI field for cross-cutting questions
Thanks

Wahid Bhimji, Eric Church, Mario Cromaz, Jacob Daughhetee, Markus Diefenthaler, Andrey Elagin, Cristiano Fanelli, Gian Michele Innocenti, Gaosong Li, Kate Scholberg, Nu Xu, Liang Yang

ALICE, COHERENT, EXO, GlueX, GRETA, KamiLAND-Zen, NEXT, STAR