Deep Learning as Model Independence

Work in Progress, MCMP

June 15, 2023

Martin King
Munich Center for Mathematical Philosophy, LMU

Preface

“No, there is no crisis in particle physics. Except maybe in some people’s minds. But the thing is that we are living in times that we don’t know yet. We are at a real crossroads, if you want, and we can’t disentangle yet where the arrows pointing or which ones we should be following.”

(De Roeck, 03.07.19)

Introduction

1. No indications in favour of predictions from BSM models
2. No statistically significant deviations in SM precision measurements

Introduction

1. No indications in favour of predictions from BSM models
2. No statistically significant deviations in SM precision measurements

Sounds good.

Introduction

1. SM research programme is (in a sense) complete
   - SM as a theory is not
2. No new high-energy particle accelerator
   - theorists and experimentalists have to get clever

Introduction

So, where will new physics come from?

- well, we don't know, but discoveries
  - won't be done in the same way
  - won't be easy
- it will require an *openness*
  - an independence from traditional modelling assumptions

Introduction

The rise of “model-independent” methods

1. SM precision measurements
   - to search for any deviations whatever
2. SMEFT
   - to parametrise deviations given some assumptions
3. Deep Learning
   - to analyse low-level data with fewer high-level physics assumptions

Overview

Overview

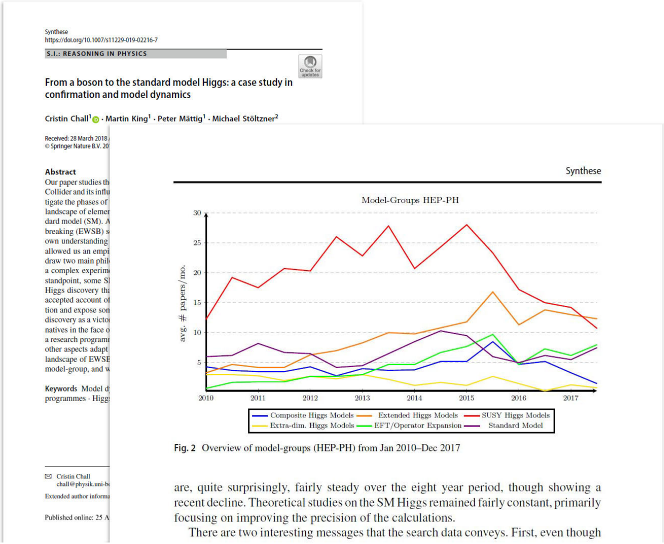
Uncertain Times

1. Identify, examine, and minimise the role of **biases**
   - explore new alternatives
   - re-evaluation of principles (e.g., naturalness)
   - model-based $\rightarrow$ model-independent (top-down $\rightarrow$ bottom-up)
2. Shift in cognitive division of labour

   "So just before the Higgs discovery, I was just doing model building for model building's sake. And I don't do that at all any more. I'm much more connected to experiment and I think that's true for most people."

Matthew McCullogh

Shifting Approaches



Synthese
https://doi.org/10.1007/s11229-019-02216-7
S.I.: REASONING IN PHYSICS

Check for updates

From a boson to the standard model Higgs: a case study in confirmation and model dynamics

Cristin Chall[1] · Martin King[1] · Peter Mättig[1] · Michael Stöltzner[2]

Received: 28 March 2018 /

Abstract
Our paper studies th
Collider and its influ
tigate the phases of
landscape of elemer
dard model (SM). A
breaking (EWSB) s
own understanding
allowed us an empi
draw two main phil
a complex experim
standpoint, some SI
Higgs discovery tha
accepted account of
tion and expose son
discovery as a victo
natives in the face o
a research programi
other aspects adapt
landscape of EWSE
model-group, and w

Keywords Model d
programmes · Higg

Cristin Chall
chall@physik.uni-b

Extended author informa

Published online: 25 A

Fig. 2 Overview of model-groups (HEP-PH) from Jan 2010–Dec 2017

are, quite surprisingly, fairly steady over the eight year period, though showing a recent decline. Theoretical studies on the SM Higgs remained fairly constant, primarily focusing on improving the precision of the calculations.

There are two interesting messages that the search data conveys. First, even though

Shifting Approaches

Studies in History and Philosophy of Science 92 (2022) 129–143

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

ELSEVIER

Bottoms up: The Standard Model Effective Field Theory from a model perspective

Philip Bechtle [a], Cristin Chall [a], Martin King [a,*], Michael Krämer [b], Peter Mättig [a], Michael Stöltzner [c]

[a] Physics Institute, University of Bonn, Germany
[b] Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen, Germany
[c] Department of Philosophy, University of South Carolina, USA

ARTICLE INFO

Keywords:
Scientific modelling
Beyond standard model
Particle physics
Effective field theory
Models and theories
Representation

ABSTRACT

Experiments in particle physics have hitherto failed to produce any significant evidence for the many explicit models of physics beyond the Standard Model (BSM) that had been proposed over the past decades. As a result, physicists have increasingly turned to model-independent strategies as tools in searching for a wide range of possible BSM effects. In this paper, we describe the Standard Model Effective Field Theory (SM-EFT) and analyse it in the context of the philosophical discussions about models, theories, and (bottom-up) effective field theories. We find that while the SM-EFT is a quantum field theory, assisting experimentalists in searching for deviations from the SM, in its general form it lacks some of the characteristic features of models. Those features only come into play if put in by hand or prompted by empirical evidence for deviations. Employing different philosophical approaches to models, we argue that the case study suggests not to take a view on models that is overly permissive because it blurs the lines between the different stages of the SM-EFT research strategies and glosses over particle physicists' motivations for undertaking this bottom-up approach in the first place. Looking at EFTs from the perspective of modelling does not require taking a stance on some specific brand of realism or taking sides in the debate between reduction and emergence into which EFTs have recently been embedded.

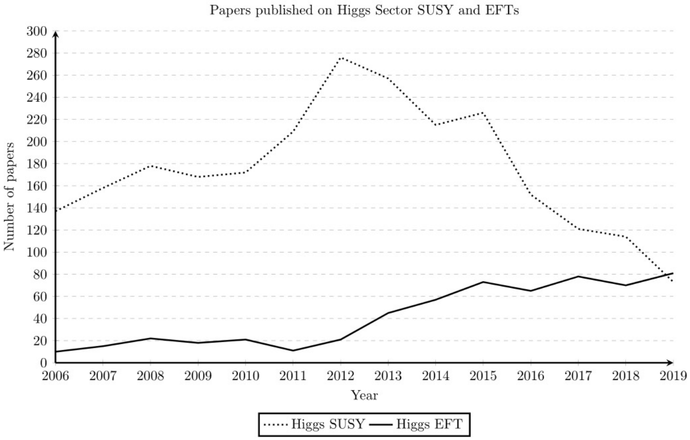
1. Introduction

To date, the Large Hadron Collider (LHC) has produced no significant evidence in favour of the many models beyond the Standard Model (BSM) that have been previously proposed. The data gathered since has increasingly reduced the parameter space in which these existing models dwell, while not indicating along which lines new models should be

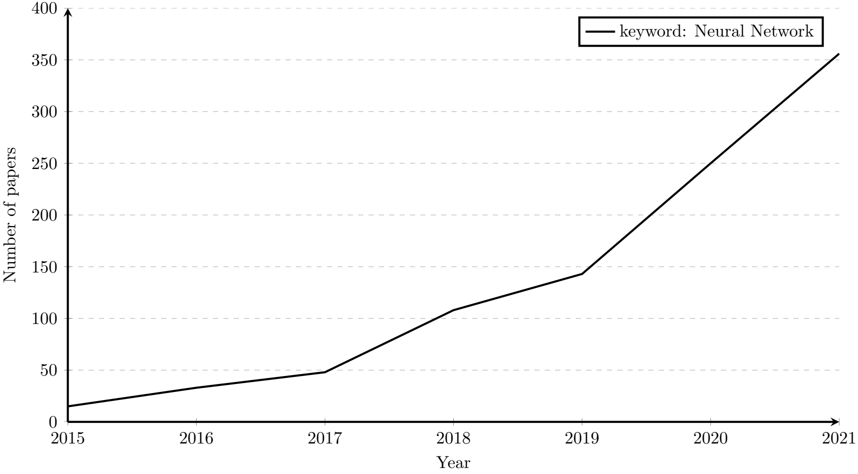
aspects of the use of the SM-EFT at the LHC. We want to ask: what is the SM-EFT and the bottom-up research strategy, what is model-independent about it, and how should we understand it in the context of the philosophical literature on models and effective field theories?

An EFT efficiently describes phenomena on a specific energy scale by availing itself of a separation of scales and absorbing the physics at higher energies only into the parameters of lower scale physics that are deter-

Shifting Approaches



Shifting Approaches



Overview

Underdetermination

BSM models are underdetermined by data

1. theory assessment; non-empirical theory confirmation (Dawid, 2019, 2022)
2. novel ways of looking at empirical evidence; model-independent methods and DL

Whence New Physics?

Possible theoretical avenues:

1. big, beautiful theories
2. small, low-energy models connected to current experiments
3. reevaluating assumptions/questioning foundations
4. accept fine-tuning, ad hocness

Whence New Physics?

Possible theoretical avenues:

1. big, beautiful theories
2. small, low-energy models connected to current experiments
3. reevaluating assumptions/questioning foundations
4. accept fine-tuning, ad hocness

Forget a theory of everything,

“I’d be happy with a theory of anything,”

John Ellis

Whence New Physics?

Possible experimental avenues:

1. higher energy collider
2. astro-particle physics/cosmology
3. neutrino observatories
4. SM precision tests at HL-LHC, elsewhere
   - Higgs physics
   - b-quarks
   - muon g-2
5. low energy frontier
   - long-lived particles (FASER), axion experiments, etc.
6. DL on existing and new data

Whence New Physics?

Possible experimental avenues:

1. higher energy collider X
2. astro-particle physics/cosmology
3. neutrino observatories
4. SM precision tests at HL-LHC, elsewhere
   - Higgs physics
   - b-quarks
   - muon g-2
5. low energy frontier
   - long-lived particles (FASER), axion experiments, etc.
6. DL on existing and new data

Comb Through the Data

Spectrum of Model Independence

- **model-dependent:** full BSM models
  - search for processes and signatures in the context of a particular, well-defined BSM model
    - charged Higgs of Type-II 2HDM with mSUSY
  - Very specific, narrow focus

Spectrum of Model Independence

- **model-dependent:** full BSM models
  - search for processes and signatures in the context of a particular, well-defined BSM model
    - charged Higgs of Type-II 2HDM with mSUSY
  - Very specific, narrow focus
- **partially model-dependent:** simplified models
  - search for particles common to many BSM models
    - leptoquark, vector triplet, stop
  - specific, broadly applicable searches

Spectrum of Model Independence

- **model-dependent:** full BSM models
  - search for processes and signatures in the context of a particular, well-defined BSM model
    - charged Higgs of Type-II 2HDM with mSUSY
  - Very specific, narrow focus
- **partially model-dependent:** simplified models
  - search for particles common to many BSM models
    - leptoquark, vector triplet, stop
  - specific, broadly applicable searches
- **model-independent:** precision measurements, using SMEFT, e.g.
  - not to search for predictions of a model but search for deviations against the background

Model Independence

Model Independence is characterised by:

- a strong reduction of the influence of modelling biases

Model Independence

Model Independence is characterised by:

- a strong reduction of the influence of modelling biases

For BSM searches:

1. a lack of a well-defined *target model* or *target phenomenon*
2. where there is a well-defined background theory (SM) against which deviations can be observed
   - this can be relaxed

Deep Learning as Model Independence

Deep Learning

- provides the tools to make analyses based on 'low-level' data with minimal preprocessing and input from physicists
- can be *unsupervised*
- aim of finding patterns in data, without being told what patterns to look for
  - anomaly detection

Model Independence

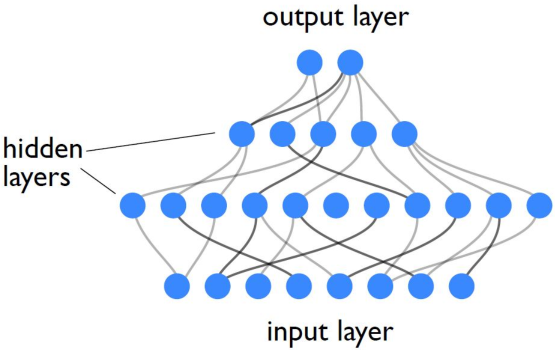
- doesn't have to stem from deep networks
- not qualitative difference, but significant quantitative

Overview

Neural Network

- automated decision algorithm of nodes and links, thought to resemble neurons in a brain
  - nodes are organized in layers from **input** through **hidden layers** to **output**
  - the outputs are *weighted* with free parameters (matrix)
  - through tuning the parameters to data, the model 'learns' to optimize the weights
- shallow or deep
  - hidden layers

Nodes



$$y_{out} = F_w(y_{in}) \tag{1}$$

Processing

Steps

1. an input signal sent to some array of nodes
2. the nodes compute their **activation functions** and pass an output signal to the next layer along their links
3. these links have different weights w which modify outputs at every stage
4. the activation of the nodes propagates until it reaches the output layer of nodes
5. this is then decoded and taken as the network's decision for that input
6. typically there is a teaching or training goal, such that the network adjusts its weights to best achieve the goal

(Buckner, 2019)

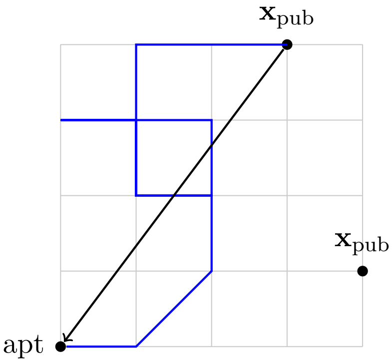
The Deep End

1. Convolution
   - linear operation passes a *kernal/filter* over the data to amplify desired features
   - exploits 'translational' symmetry, reducing the number of parameters
2. Rectification (ReLU)
   - threshold activation unit
3. Max Pooling
   - downsampling function, combining many kernals(filters), pooling only maximum values

Goals

The main practical goal is **generalisation**

1. learn a set of data
   - avoid underfitting
   - divide data into *training set* and *test set*
2. have it apply to new data
   - avoid overfitting
   - various methods of *regularization*
     - dropping out nodes, penalty terms, modifying input data, etc.



Reward is rare

1. conduct a bunch of trajectories
2. sum over rewards for each
3. apply gradient training

Backpropagation

- $\rightarrow$ true gradient
- approximate with stochastic gradient descent

Overview

Changing Game

"In the relatively few years that modern machine learning has existed, it has already made traditional collider physics obsolete. In the past, physicists, including me, would devote their efforts to understanding signatures of particular particles or processes from first-principles: why should a stream of pions coming from a W boson decay look different than a stream coming from an energetic gluon? Now we simply simulate the events, and let neural network learn to tell the two samples apart."

(Schwartz, 2021)

Curse of Dimensionality

The difficulty that deep learning can deal with is not the sheer amount of data, but the *high-dimensionality* of LHC data

Task in HEP

- estimate probability function $p(x|\theta)$ of observing some data x, given the model's parameters θ
- estimating is possible for low-dimensional data $(d < 5)$
- number of samples to estimate the function grows with the power of the dimensionality of data N^d

Must reduce the dimensionality of the data, $d = \mathcal{O}10^8$

(Guest et al., 2018)

Reducing Dimensionality

Strategy

- Construct higher-level, lower-dimensional objects and analyse those

Steps

1. reconstruction algorithms are used to process raw data into objects (clusters and tracks)
2. use this to estimate the energy and momentum of particles
3. identify particles
4. build *event-level* summaries
5. perform event selection for further analysis

The reconstruction and selection are traditionally based on *physicist-identified* features of the data (shape of the shower in the calorimeter)

Reducing Dimensionality

Strategy

- Construct higher-level, lower-dimensional objects and analyse those

Steps

1. reconstruction algorithms are used to process raw data into objects (clusters and tracks)
2. use this to estimate the energy and momentum of particles
3. identify particles
4. build *event-level* summaries
5. perform event selection for further analysis

The reconstruction and selection are traditionally based on *physicist-identified* features of the data (shape of the shower in the calorimeter)

- works pretty well, but not *optimal*
- DL helps optimize this process

Optimization

Most tasks can be formulated in terms of the optimization of a loss function $L[y, f(x)]$

- search $f(x)$ that optimizes move from high-dimensional space of observed data $\rightarrow$ low-dimension space
- the (reduced) space of functions searched is given by a series of transformations mapping inputs x onto hidden states h_i and then to the output y

$$h_{i+1} = g_i(W_i h_i + b_i) \tag{2}$$

where g_i is the activation function and a particular h_i is the ith transformation, called the embedding and the W's are matrices and the b's are biases

- the model is trained by optimizing the values of the weights
- training examples are used to calculate the gradient of the loss function with respect to the model parameters $\nabla_\phi L[f_\phi(x), y]$, often via backpropagation

Hierarchical Data Representation

	Image Recognition	**HEP**
High level	Shapes $\rightarrow$ Objects	Particles $\rightarrow$ Events
$\uparrow$		
Middle level	Edges $\rightarrow$ Shapes	Tracks $\rightarrow$ Particles
$\uparrow$		
Low level	Pixels $\rightarrow$ Edges	Hits $\rightarrow$ Tracks

A Third Benefit of DL

1. Outperforms traditional methods in many areas
2. More computationally efficient
3. Offers model-independent ways of analyzing data

Uses of DL in HEP

All Along Data Pipeline

1. Simulation
2. Improving trigger timing
3. Low-level: hit and track reconstruction
4. Mid-level: object identification
5. High-level: event classification

The Need for Speed

At HL-LHC, there will be 7x luminosity (20x run-2 data set), not enough more computing resources

- ML has been used for decades (multivariate analysis, boosted decision trees)
- DL solutions are needed

DL in Simulation

1. Detector
   - discovery requires comparing the detector response with SM and BSM expectations
   - detector response cannot be analytically computed
2. Events
   - Trillions of collisions need to be simulated in order to make accurate comparisons
   - DL can learn the physics and generate new events
3. Background
   - background can be learned more accurately (can be in situ)
   - aid in anomaly detection

Triggers

CMS/ATLAS have $\sim$100 million channels sampled every $\sim$25ns

- 60Tb/s raw data

40 million events/s must be searched in **real time**

1. L1 hardware trigger $\rightarrow$ 100k/s
   - calorimeter + muon spectrometer energy threshold (+topo)
   - decision in 200 nano secs
2. L2 software (high-level) trigger $\rightarrow$ 1000/s
   - jets, missing p_T
3. ATLAS has L3 trigger $\rightarrow$ 200 events

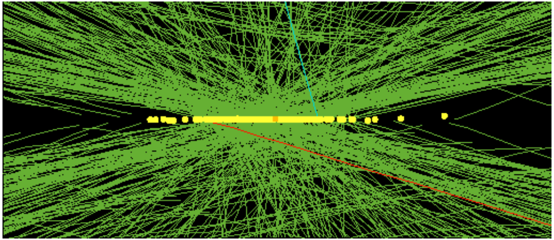
Triggers

DL offers the possibility of triggering on events that cannot currently be stored

- acceptance 1 in 40,000 events
  - can write 15Pb/y for offline analysis
- try to keep low-energy DM signatures, b and c quark physics, etc.
- event selection based on cuts may miss a lot of events of interest

(ATLAS Collaboration, 2021; CMS, 2018)

Low-level: Track Reconstruction



- may only be one event of interest in a bunch of 150-200 interactions
- the rest is called *pile up*
  - it can be removed by having the algorithm remove noise
- DNN can more faithfully and more quickly build tracks from hits

Low-level: Track Reconstruction

Process

1. clusters hits in innermost chamber (pixel sensors)
2. 10^8 pixel channels give 10^3 tracks
3. almost 100% efficient, but computationally very intensive
   - at HL-LHC increase in events/s leads to combinatorial increase in false seeds for tracks

Mid-level: Jet Classification

1. flavour tagging
   - top tagging: high mass, decays to 3 lighter quarks (jets), displaced vertex
   - ignore the physics, let the DNN decide which features to care about
2. substructure tagging
   - structure of calorimeters makes for a cylindrical projective surface on which the detectors can serve like pixels of image
   - leverage DL image recognition
   - typically requires preprocessing, but DL works better without

(Guest et al., 2018; Schwartz, 2021)

High-level: Event Selection

Traditionally

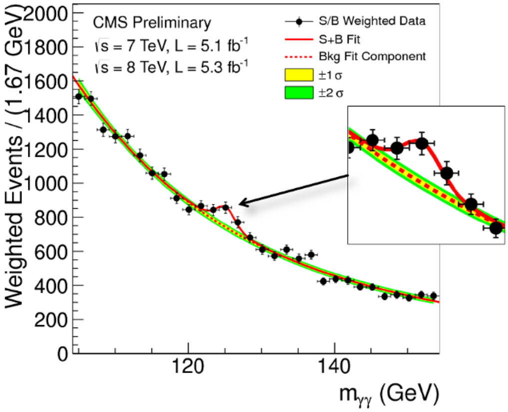
- analyses use reconstructed invariant masses of particles
- high-level physicist-processed data

DL end-to-end

- can use 4-momenta
- low-level raw data with minimal input and processing
- outperforms traditional classification

(Guest et al., 2018; Baldi, 2014)

Task: distinguish signal from background



Discovery

A signal is often a peak due to resonance

- in that region the background is negligible
- elsewhere the signal is negligible

Probability distribution

$$P_{data} = \alpha_S P_S + \alpha_B P_B \tag{3}$$

and then determine the coefficients and see if α_S is zero

- it takes many, many events to be able to constrain α_S
- analyses typically done on *one* variable at a time (histogram)

Discovery

Deep Learning

- run through the entire data with minimal assumptions
- make decisions based on more than one variable at the same time
- identify background without knowing the SM
- identify anomalous events
  - rank by typicality

Model-Independent Strategy

If you want to tag a top quark ($t\bar{t}$ pair production)

- don't tell the model that the top is heavy, that it decays to three other jets, that it has a slightly displaced vertex, etc.
- feed it tons of simulated data and it will have a higher success rate

Overview

Epistemic Issues with DL

1. lack of theory
   - no help in design
2. lack of opacity
   - which features is it latching onto?
3. lack of explainability
   - how are decisions justified?
4. predictions without explanation
   - getting an answer is nice, but will it give us satisfying explanations?

Opacity

Given any decision, we don't know along what lines it was made

- model decision output with understandable model (xAI)
  - need to show over which domain the net is reliable and accurate, where it breaks down, and where its behaviour is uncertain
  - can't always be done
  - no guarantee that it is picking up on the same features
  - means more computation
- vary input
  - are we varying it in the right way?

(Boge, 2021; Mittelstadt et al., 2019)

Explanation

1. explaining the model's decisions
   - in many contexts one may be asked to justify to decision
   - all one can point to is the reliability of the outcome
2. using the input-output in explanation
   - like a black-box or complex simulation
   - limited use
     - no view of the process/components
     - no causal or mechanical information
3. Not an ordinary black box
   - there is an in-principle opacity

(Boge, 2021; Schubbach, 2019; Sullivan, 2019)

Explainable AI

There are two broad aims in xAI

1. transparency—how a model functions internally
   - simulatability
   - decomposability
   - algorithmic transparency
2. post hoc interpretation—how the model behaves
   - input-output relations
   - relative influence of components
   - local explanations by retrofitting a simplified model, or assessing its robustness
   - identifying most similar training data

(Mittelstadt et al., 2019)

Overview

Epistemic Difficulties

1. No theory of DL
2. Reliance on Simulated Data
3. Mismodelling
4. Interpretability and Understanding

1. No Theory

The mathematical theory of statistical learning lags far behind the success of DL models

- little to offer in terms of design and refinement of algorithms and new techniques
- matter of heuristics and trial and error
- No theoretical guarantee of reliability or optimality

(Bahri et al., 2020; Belkin, 2021)

2. Reliance on Simulation

Simulations and simulated data is used *everywhere* in HEP

(Morrison, 2015)

Quantum data The quantum nature of the real event means that simulated data is different

- there are no ground truths
- for a given event, you cannot say ‘Higgs’ or ‘no Higgs’ decisively
- all one can do is collect statistics and try to reject the background only hypothesis

(Schwartz, 2021)

3. Mismodelling

Trusted simulations work in low-dimensional spaces with high-level physics features

- correlations in a high-dimensional space may not be faithful

Just because a model has a better false-positive rate doesn't mean it will perform better in particle identification

- e.g. if it works better on simulated data than real
- performance will be up to systematic uncertainties, signal-to-background ratio, etc.
  - things that are traditionally accounted for, but ignored in DL

4. Interpretation and Understanding

Progress on computation is often a trade off with understanding the model's decisions

> *"Generically, we should anticipate a trade-off between performance and interpretability"*

(Guest et al., 2018, p. 175)

It is not a like an analytic function, where the structure of the solution can be analysed

- reverse engineering the classification strategy is impossible due to the high-dimensional nature of the data

The decisions may be faster, more accurate, more computationally tractable, but as long as we don't understand the decisions, they will be of limited benefit

- maybe the features being picked up on could be insightful, but we won't know

Specificity

- Perhaps something learned in a network trained on top tagging could be useful in bottom tagging, but we won't know
- Sharing with other experiments may not be possible, since they often run different software

Non-traditional

- not a cut-based analysis given in histograms and natural language description

Issues but not Problems

1. No theory is kind of okay, as long as it works
2. Sufficient data for reliability
3. Post hoc interpretation—the model's don't do everything

Justification is not explanation

- justification comes from success
- its decisions can be easily justified, but not easily explained

The proof is in the pudding

- theoretical guarantees still rely on statistical arguments that are essentially about reliability of results

2. Simulated and Real Data

No need In principle, one can proceed in HEP entirely without simulations.

- if it can be simulated, it can be done with pen and paper
- it would be slow

Use real data

- DL can also be run on real data (without ground truths), or mix of real and sim
- there is roughly the same amount of sim and real data at LHC (trillion events)

Sufficient data

- in HEP there is *too much* data and computing resources
  - train and test until the cows come home
  - statistical errors are small, *systematics* become the issue
    - DL can help with this too! $(error_{sys} \to \nu)$

Varied approaches

- data is public and there are open competitions
- new approaches and architectures are tried across the globe

Maybe its just not for that

- it is not already a model of new physics
- if an anomaly is discovered, the deep CNN (or whatever it is) will provide no understandable model in terms of intuitive physics

4. After Discovery

Maybe its just not for that

- it is not already a model of new physics
- if an anomaly is discovered, the deep CNN (or whatever it is) will provide no understandable model in terms of intuitive physics

But it will indicate where new physics is

- then traditional model-building will turn on
- similar to EFT approach
- similar to xAI

It is can be tool to identify correlations, but it will not likely help in interpreting the data

Conclusion: the Big Data Approach

"statistical algorithms find patterns where science cannot"

"We usually don't know about causation, and we often don't necessarily care... the objective is more to predict than it is to understand the world... It just needs to work; prediction trumps explanation."

(Kitchin, 2014)

Conclusion: Role for DL in HEP

- three important feature of DL: performance, efficiency, and model independence

Conclusion: Role for DL in HEP

- three important feature of DL: performance, efficiency, and model independence
- has many practical and epistemic difficulties to overcome

Conclusion: Role for DL in HEP

- three important feature of DL: performance, efficiency, and model independence
- has many practical and epistemic difficulties to overcome
- but these are *issues* not *problems*

Thank you

Martin King
m.king@lmu.de
www.philphys.com
Munich Center for Mathematical Philosophy, LMU

This work was supported by the DFG.

References:

ATLAS Collaboration (2021). *JINST*, 16.

Bahri, Y., Kadmon, J., Pennington, J., Schoenholz, S. S., Sohl-Dickstein, J., and Ganguli, S. (2020). Statistical mechanics of deep learning. *Annual Review of Condensed Matter Physics*, 11(1):501–528.

Baldi, P., S. P. . W. D. (2014). Searching for exotic particles in high-energy physics with deep learning. *Nature Commun*, 5:4308.

Belkin, M. (2021). Fit without fear: remarkable mathematical phenomena of deep learning through the prism of interpolation.

Boge, F. J. (2021). Two dimensions of opacity and the deep learning predicament. *Minds and Machines*, 32(1):43–75.

Buckner, C. (2019). Deep learning: A philosophical introduction. *Philosophy Compass*, 14(10):e12625. e12625 PHCO-1206.R1.

CMS (2018). Usage of tracking in the cms level-1 trigger for the high luminosity lhc upgrade.

Dawid, R. (2019). The significance of non-empirical confirmation in fundamental physics. In Dardashti, R., Dawid, R., and Thebault, K., editors, *Why Trust a Theory? Epistemology of ModernPhysics*, pages 99–119. Cambridge University Press.

Dawid, R. (2022). Meta-empirical confirmation: Addressing three points of criticism. *Studies in History and Philosophy of Science Part A*, 93(C):66–71.

Guest, D., Cranmer, K., and Whiteson, D. (2018). Deep learning and its application to lhc physics. *Annual Review of Nuclear and Particle Science*, 68(1):161–181.

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1):2053951714528481.

Mittelstadt, B., Russell, C., and Wachter, S. (2019). Explaining explanations in ai. *Proceedings of the Conference on Fairness, Accountability, and Transparency*.

Morrison, M. (2015). *Reconstructing Reality: Models, Mathematics, and Simulations*. Oup Usa.

Schubbach, A. (2019). Judging machines: Philosophical aspects of deep learning. *Synthese*, 198(2):1807–1827.

Schwartz, M. D. (2021). Modern machine learning and particle physics. *Harvard Data Science Review*, 3(2).

Sullivan, E. (2019). Understanding from machine learning models. *British Journal for the Philosophy of Science*, pages 1–28.