



University
of Glasgow

Reusing Neural Networks: Lessons learned and Suggestions for the future

Tomasz Procter

Summary from the Les Houches reinterpretable ML working
group

INTRODUCTION

The problem

- Neural Nets are becoming more and more central features of many collider analyses.



- Use a wide variety of frameworks - tensorflow, scikit-learn, pytorch, ROOT TMVA, Neurobayes...
- Implies:
 - Wide variety of dependencies -> heavy codes.
 - Wide variety of output formats (not all human readable).
 - ML in industry is less interested in reproducibility - scary differences between version numbers.
- And anyway, it's rare that an analysis actually publishes their NN data...

Where we are



- Two publicly available LHC analysis ML networks - both from ATLAS SUSY:
 - ANA-SUSY-2019-04 was discussed at last RiF
 - ANA-SUSY-2018-30 seems to have worked a lot better in multiple frameworks - good to see!
- Nothing from CMS
- Nothing from ATLAS groups beyond SUSY.
- Clearly ATLAS SUSY group policies concerning simpleAnalysis have been very useful - do other groups need to take note?
- Personal comment: trying to get approval to add an .lwttn file to hepdata has been...prolonged..., and exact procedures seem unclear to everyone*.

* If you think you could help me get this out I'd appreciate it!

Summary of Les Houches wishlist - Analysis Design, Implementation and Validation (see full document for more detail)

Analysis Design

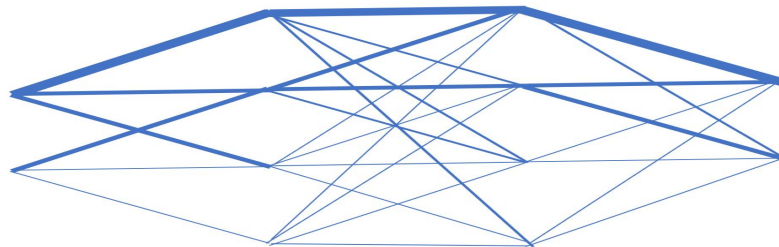
- Use an open-source framework (tensorflow, pytorch, etc)
- Ensure the network can be saved in a useful preservation format for inference (e.g. ONNX or lwttn).
 - Just leaving in a `.h5` file or `.pkl` file is unlikely to be stable
- Be considerate with choice of inputs - if a tagger depends entirely on detector level inputs, that's fine (but please provide detailed efficiencies – including misstags – or surrogates), but 10 truth-level quantities + pseudo-continuous b-score is frustrating.

Supplementary Material

- Like for variables in any other analysis, we need full definitions of all variables that go into and come out of the neural net.
 - This is even more important given the “black box” nature of a neural network.
- Definitions include:
 - Units (MeV v GeV)
 - Normalisations
 - Padding values
 - Phi conventions ($0 - 2\pi$ vs $-\pi - \pi$)
 - Input and output ordering
- A **validated** analysis code (rivet, simpleAnalysis) automatically supplies much of this info.
- Otherwise, a short, minimal note might need to be uploaded alongside the onnx/lwttn file.

Validation Material

- Where cuts depend on a neural net output, like for every other cut-based analysis, cutflows are a vital validation tool.
- Input and output plots (especially for most important features) could also be useful.
- When cutflows or extra plots are provided, just like for any other analysis, we really need to know the exact signal model that produced them (slha files, generator run cards, etc)
- Some understanding of feature importance is not only physically interesting, but can be essential in debugging.



Surrogates

Tomasz Procter, RIF, August 2023

What is a surrogate?

- What to do if an ML-model requires very experiment specific inputs (e.g. hits, exact btagging scores)?
(There are more and more such networks being used in analyses)
- Train *another* network:
 - Given truth/reinterpretation level inputs
 - Mimic output score of original model case by case
 - Probably with some randomness built in
- May or may not have access to the “true” answer (e.g. does the jet really contain a top quark?).
- If yes, effectively “parametrised efficiency on steroids”

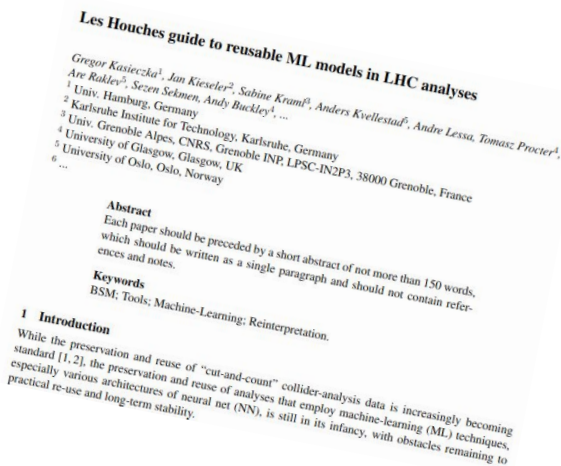
Surrogates - what next?

- No such surrogate networks currently available.
- Les Houches working group “proof of concept” project on CMS run-1 btagger using CMSOpenData.
- Would be great to see something originating in the experiments (who have all the data) - could be a really cool qualification task or similar.
- If you have a pet generative ML model you would like to throw at this, let me know!

Conclusions

Conclusions

- Experiments need to start publishing these, ideally on HEPdata -
 - If changes are needed to internal structures/procedures to help make this happen, do it.
 - This is a new type of experimental output.
 - The “institutional memory” required to properly preserve run-2 ML dependent analyses is fading fast.
 - If there are any analyses “essential” to reinterpreters that do depend on NN...
- The “Les Houches guidelines” are growing
 - <https://www.overleaf.com/8811915719zfjtnygcgdpv>
 - Welcome feedback/contributions.
 - Several WG members here, happy to discuss.
- Reinterpreters **will** be interested and **will** cite...



BONUS