Discussion:
Machine Learning and String Theory

Jim Halverson, Vishnu Jejjala,
and Andre Lukas

Topics to Review

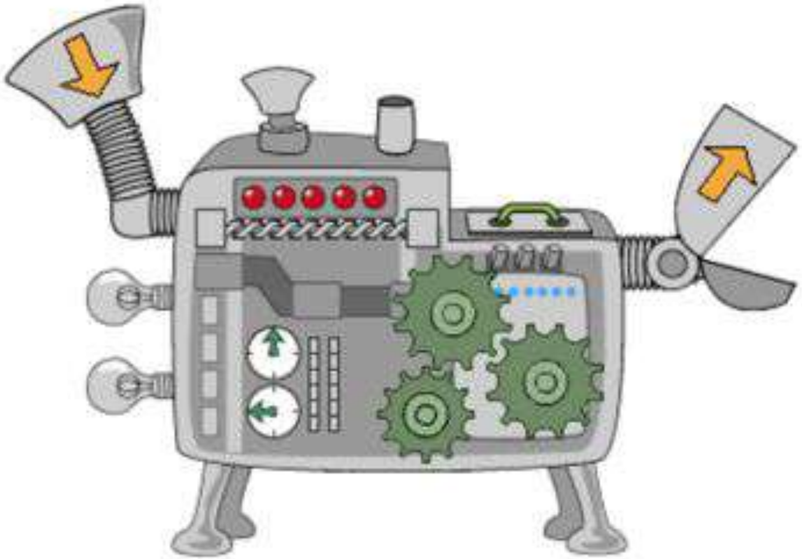
Supervised Learning (learning difficult-to-compute data, conjecture generation)

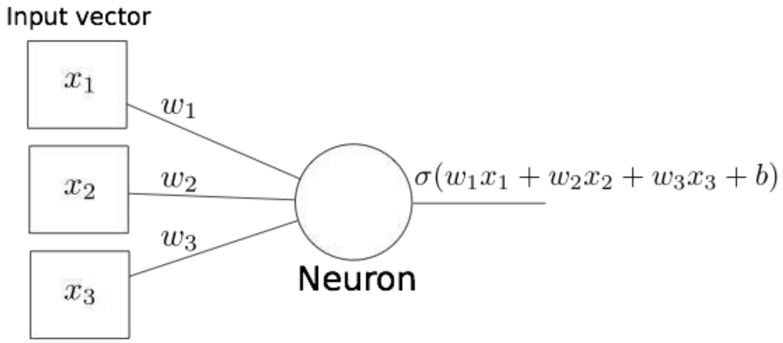
Self-Supervised Learning (CY metrics, wavefunctions, HYM, etc)

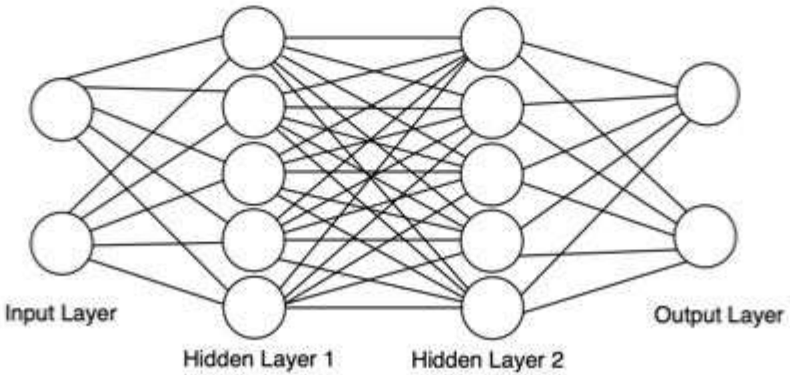
Reinforcement Learning (review, model-building + search)

Supervised

Neural networks







Supervised

Calabi–Yau data

CICY threefolds: He (2017), Bull, He, VJ, Mishra (2018, 2019), Erbin, Finotello (2020)

CICY fourfolds: He, Lukas (2020), Erbin, Finotello, Schneider, Tamaazousti (2021)

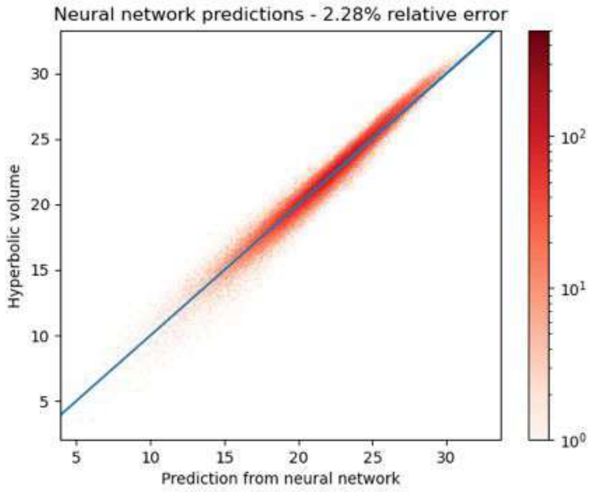
Line bundle cohomology

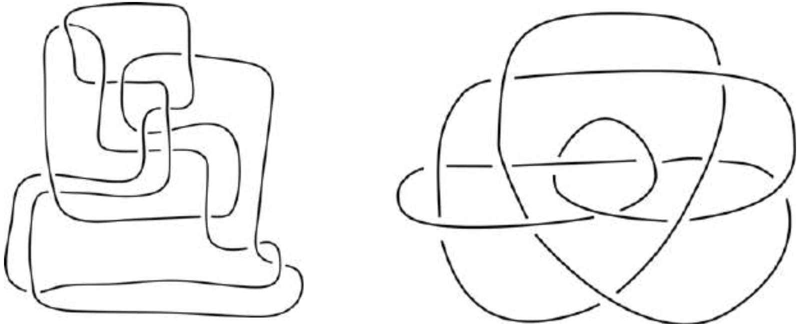
Klaewer, Schlechter (2018), Brodie, Constantin, Deen, Lukas (2019)

Supervised

Knot theory

Hughes (2016), VJ, Kar, Parrikar (2019), Levitt, Hajij, Sazdanovic (2019), Gukov, Halverson, Ruehle, Sulkowski (2020), Craven, VJ, Kar (2020)





Supervised

ML identifies associations — how does a machine learn?

Bridge this success to obtain new analytic results and methods

e.g., calculations of Hodge numbers scale polynomially rather than doubly exponentially — are there new ways to calculate in algebraic geometry?

Conjecture generation from simple algorithms

Carifio, Halverson, Krioukov, Nelson (2017)

Can machines do original physics or mathematics? How do $_{6}C$ and $_{14}Si$ collaborate?

Self-Generative Learning

- **crux:** let the NN be a variation ansatz for some function that you care about.

- **train:** loss is some function of NN itself, doesn't depend on labelled data.

- **examples:**
  - NN a CY metric, loss $|\mathrm{Ricci}|^2$.
  - NN quantum state, loss $|E|^2$. [Carleo, Troyer]
  - NN a PDE solution to $D f = 0$, D some op. loss is $|D f|^2$. e.g. HYM equations.

Theory Comments:

Good idea in principle b/c NN a universal approximator.

Doesn't guarantee learning dynamics that find the good solution.

Self-Generative Learning: Metrics

[Anderson, Gerdes, Gray, Krippendorf, Raghuram, Ruehle]
[Douglas, Qi] [Jejjala, Mayorga Pena, Mishra]

- Let the NN be a metric.
  Learn CY metric.

  1) use $L = |\text{Ricci}|^2$
  2) use Monge-Ampere loss
  3) other losses?

  Can build in moduli dependence.

- **Swampland:** given metric, study eigenvalues of KK modes, see level crossing, study rel. to Swampland dist. conj. [Ashmore, Ruehle].

- **Line bundle connections.**
  [Ashmore, Deen, He, Ovrut]

Outlook: Moving away from BPS

- two truly amazing results:
  1) Yau's theorem. Topology guarantees geometry, we have string backgrounds!

  2) Calibration. can compute some submanifold volumes, w/o a metric!

  Use CY metrics when we can't do the latter.

  - non BPS states, yes in CY, but also in M-theory on G2, crucial for gauge symmetry!

  - WGC and minimal surfaces, a la [Demirtas, Long, McAllister, Stillman]

Self-Generative Learning

More Outlook

- **G2 metrics.** No analog of Yau's theorem that ensures a good starting place. Nevertheless!
- **Real fibrations:**
  Ambitious: SYZ fibration develops in limits of CY moduli, a test of mirror symmetry. [Strominger, Yau, Zaslow]

  Less known: five-manifold fibered by two-spheres in G2 manifolds with Coulomb-breaking of non-abelian theory. Follows from semi-classical physics of topological defects, exists in CY, but no calibration to help you in G2. [Halverson, Morrison] [Joyce, Karigiannis]

More Theory

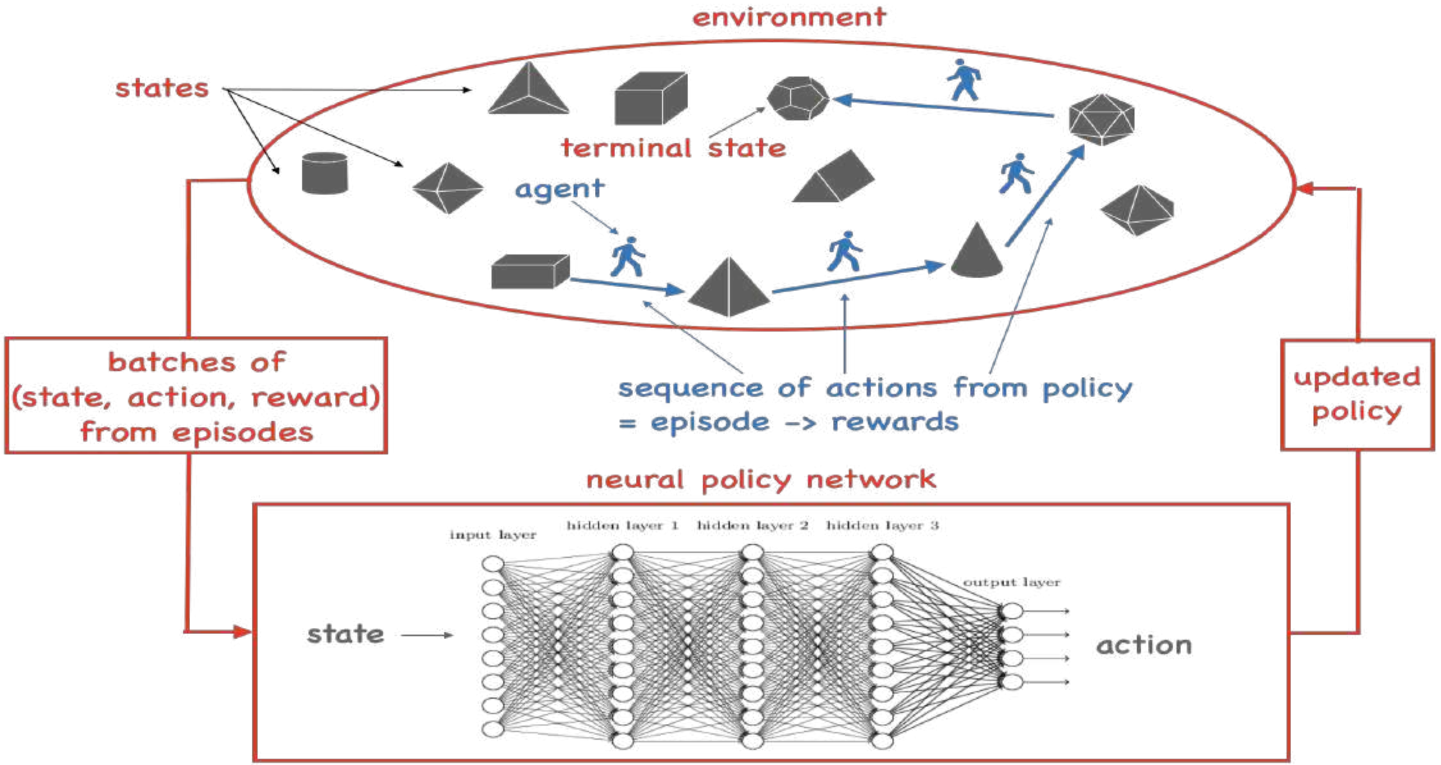
Neural tangent kernel:
Gradient flow governed of NNs in infinite width limit governed by t-independent, parameter-independent kernel, NTK.
[Jacot, Gabriel, Hongler]

L2 loss: analytic mean predictions for an infinite ensemble of infinitely wide trained NNs. **kern. reg.**

Infinite width self-generative learning WIP:
[Halverson, Luo], [Halverson, Ruehle]

Principled architecture design: tune NTK spectrum to encourage fast self-gen. learning, e.g. with RFFs.

Reinforcement Learning (RL)



(Reinforce, actor-critic, Q-learning, . . .)

RL

String model search and model building

Environment →	family of string models/string data
states →	specific models
action →	``small" modification of model
reward →	measure for how much desirable features of model improve

Goals:

- Explore large classes of models, identify desirable ones
- Use trained network to build models
- Identify model-building strategies

RL

- IIA intersection brane models -> new model-building strategy
  (J. Halverson, B. Nelson, F. Ruehle, 1903.11616)
- Heterotic line bundle models -> explore large classes, scaling with $h^{1,1}(X)$
  (M. Larfors, R. Schneider, 2003.04817)
- Knots -> RL learns to "unknot"
  (S. Gukov, F. Ruehle, P. Sulkowski, 2010.16263)
- Heterotic monad bundles -> large environments, new models
  (A. Constantin. T. Harvey, AL, 2108.07316)
- Heterotic monad bundles -> comprehensive scan, checked with GA
  (A. Abel, A. Constantin. T. Harvey, AL, 2110.14029)

Remarks:

- RL changes approach to string/QFT model building
- Complete search of string landscape ($10^{\mathcal{O}(10)\, h^{1,1}(X)}$) possible? Scaling with $h^{1,1}(X)$?
- Too many states ($\sim 10^{h^{1,1}(X)}$) with "good" spectrum. Need to refine what we are searching for . . .

Discussion Time

Extending These Topics + Discussing Your Topics

Extending These Topics

Supervision?

Learn mathematical structures
(what questions in mathematics/physics are learnable)
(failed experiments)

How machines learn – analogy to RG

Electron mass, ...

Embed symmetries in architecture

Self-Generative?

KK-modes?

Non-BPS objects?

Real fibrations?

G2?

NTK?

More CY Examples?

RL?

Large h11?
Complete searches?

Find models w/
realistic particles +
cosmo

Detailed model
requirements?

F-theory
searches?

Your topics!

sub-idea 1?

Your Big Idea

sub-idea 2?

sub-idea 3?

---

sub-idea 1?

Your Big Idea

sub-idea 2?

sub-idea 3?

---

sub-idea 1?

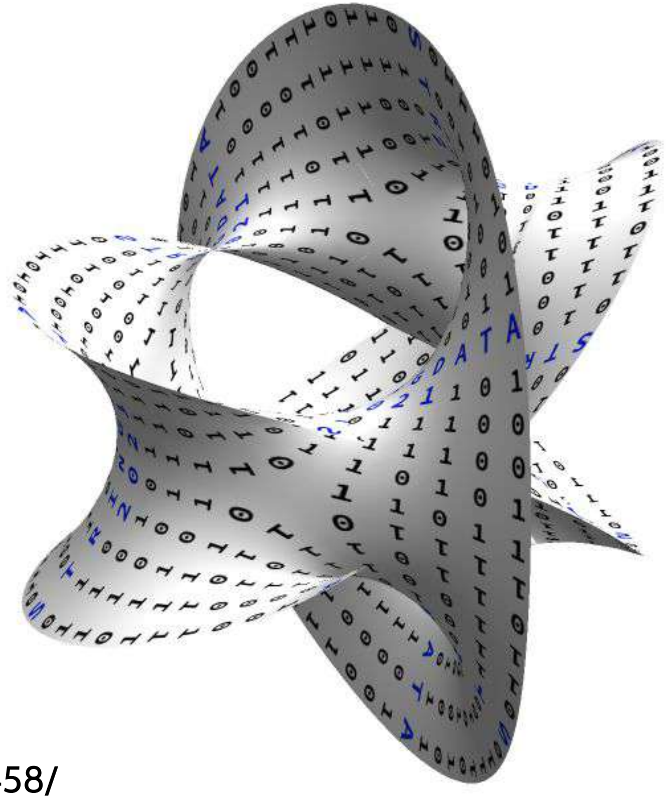
Your Big Idea

sub-idea 2?

sub-idea 3?

string_data_2021

13-17 December 2021



https://indico.cern.ch/event/1065458/