



City Research Online

City, University of London Institutional Repository

Citation: He, Y-H. (2023). From the String Landscape to the Mathematical Landscape: A Machine-Learning Outlook. Paper presented at the International Workshop on Lie Theory and Its Applications in Physics, 20-26 Jun 2021, Sofia, Bulgaria. doi: 10.1007/978-981-19-4751-3_2

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/32852/>

Link to published version: https://doi.org/10.1007/978-981-19-4751-3_2

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

From the String Landscape to the Mathematical Landscape: a Machine-Learning Outlook

Yang-Hui He

Abstract We review the recent programme of using machine-learning to explore the landscape of mathematical problems. With this paradigm as a model for human intuition - complementary to and in contrast with the more formalistic approach of automated theorem proving - we highlight some experiments on how AI helps with conjecture formulation, pattern recognition and computation.

1 The String Landscape

Perhaps the greatest theoretical challenge to string theory as a theory of everything is the vast proliferation of possible vacuum solutions, each of which is a possible 4-dimensional “universe” that descends from the 10 spacetime dimensions of the superstring. This is the so-called “vacuum degeneracy problem”, or the “string landscape problem”. The reason for this multitude is the vast number of possible geometries for the missing 6 dimensions. Whether we consider compactification, where the a Calabi-Yau manifold constitutes the missing dimensions, or configurations of branes whose world-volumes complement these dimensions, we are inevitably confronted with the heart of the problem: geometrical structures, often due to an underlying combinatorial problem, tend to grow exponentially with dimension.

We can see this from estimates of possible vacua, which engender such astronomical numbers as 10^{500} to 10^{10^5} [1–3]. These estimates come from

Yang-Hui He

London Institute for Mathematical Sciences, Royal Institution of Great Britain, 21 Albemarle St., London, W1S 4BS; Merton College, University of Oxford, OX1 4JD, UK; Department of Mathematics, City, University of London, EC1V0HB, UK; School of Physics, NanKai University, Tianjin, 300071, P.R. China.

e-mail: hey@maths.ox.ac.uk

tallying “typical” number of topologies of “typical” manifolds, as governed by the number of holes (or more strictly, algebraic cycles) of various dimensions within the manifolds. Such topological quantities are immanently combinatorial in nature.

Lacking a fundamental “selection principle” [4] - which would find *our universe* among the myriad - the traditional approaches have been statistical valuations [5], or brute-force searching for the Standard Model [6–11] as a needle in the haystack. Whilst these approaches have met some success, the overwhelming complexity (especially in the computational sense [12]) of, and the want of a canonical measure [19] on, the string landscape, beckon for a paradigmatically different method of attack.

As the Zeitgeist of Artificial Intelligence (AI) breathes over all disciplines of science [13] in recent times, and as we firmly enter the era of Big Data and Machine-Learning (ML), it is only natural that such a perspective be undertaken to explore the string landscape. This was indeed done in 2017 when ML was introduced into string theory [14–18]. In particular, the proposal of [14] was to see whether ML could be used to study the databases in algebraic geometry, which have been compiled over the last few decades for the sake of studying string theory in physics and concepts such as mirror symmetry in mathematics. To some details of this programme let us now turn.

1.1 Calabi-Yau Manifolds: from Geometry to Physics

Fig. 1 The trichotomy of (smooth, compact, boundary-less) surfaces, organized according to topological type and related curvature.

					...
$g(\Sigma) = 0$	$g(\Sigma) = 1$	$g(\Sigma) > 1$			
$\chi(\Sigma) = 2$	$\chi(\Sigma) = 0$	$\chi(\Sigma) < 0$			
Spherical	Ricci-Flat	Hyperbolic			
+ curvature	0 curvature	– curvature			

The classification of (compact, smooth, boundary-less) surfaces Σ goes back to at least Euler, who realized that a single integer, called *genus*, completely characterizes the topological type of Σ . Roughly, the genus g counts the number of “holes”: a sphere S^2 has genus 0, a torus $T^2 = S^1 \times S^1$ has genus 1, etc. The *Theorema Egregium* of Gauss then relates topology to metric geometry:

$$2 - 2g = \frac{1}{2\pi} \int_{\Sigma} R. \quad (1)$$

In the above, the combination $\chi = 2 - 2g$ is the Euler number and R is the (Gaussian) curvature. We therefore see a natural trichotomy of surfaces,

as summarized in Fig. 1: negative, zero and positive curvature, with the boundary case of $R = 0$, or Ricci-flatness, being the torus T^2 .

With Riemann enters complex geometry: Σ is not merely a real dimension 2 manifold, but a complex dimension 1 manifold. The trichotomy in this context manifests as Riemann Uniformization. Complexification allows us to employ the powers of algebraic geometry over \mathbb{C} and Σ can thus be realized as a complex algebraic curve. For instance, it can be the vanishing locus of a complex polynomial in the three projective variables $[x : y : z]$ of $\mathbb{C}\mathbb{P}^2$. The torus, in particular, can be realized as the famous cubic elliptic curve. In modern parlance, the Gaussian integral is thought of as the intersection theory between homology (the class $[\Sigma]$) with cohomology (the first Chern class $c_1(\Sigma)$). Likewise, χ , by the index theorem, is the alternating sum of dimensions of appropriate (co-)homology groups. We summarize this beautiful story, spanning the two centuries from Euler to Chern, Atiyah, Singer et al., in Fig. 2.

Fig. 2 The index theorem relating differential/algebraic geometry/topology for surfaces as complex algebraic curves.

$\chi(\Sigma) = 2 - 2g(\Sigma) =$			
$= [c_1(\Sigma)] \cdot [\Sigma] =$			
$= \frac{1}{2\pi} \int_{\Sigma} R =$			
$= \sum_{i=0}^2 (-1)^i h^i(\Sigma)$			
Topology	Algebraic Geometry	Differential Geometry	Index Theorem (co-)Homology
Invariants	Characteristic classes	Curvature	Betti Numbers

Generalizing Figs. 1 and 2 to complex dimension higher than 1 is, understandably, difficult. However, at least for a class of complex manifolds, called Kähler, whose (Hermitian) metric $g_{\bar{\mu}\nu}$ comes from a single scalar potential K as $g_{\bar{\mu}\nu} = \partial_{\bar{\mu}} \partial_{\nu} K$, the story does extend nicely: the Chern class governs the curvature. This is roughly the content and significance of the Calabi conjecture [21], which Yau proved some 20 years later in his Fields-Medal-winning work [22].

It is serendipitous that when string theorists worked out the conditions for compactification in the incipience of string phenomenology [23], one of the solutions (and today still standard) for the extra 6-dimensions is a complex, Kähler, Ricci-flat 3-fold. Furthermore, Strominger, one of the authors, was Yau's visitor at the IAS. And thus the world of high-energy theoretical physics intermingled with the world of complex algebraic geometry. In fact, the physicists named such manifolds "Calabi-Yau" (CY), and the rest, was history. The torus T^2 , is thus a premium example of a Calabi-Yau 1-fold, of complex dimension 1. The reader interested in further details of the Calabi-Yau landscape as a confluence between physics, mathematics and modern data science, is referred to the pedagogical book [24].

Over the decades since the mid-1980s, a host of activity ensued in creating large data-bases of CY manifolds for the intention of sifting through

to find the Standard Model. Perhaps it was unexpected that the number ¹ of CY 3-folds reached billions by the turn of the century (and still growing!) [25]. Furthermore, the sophisticated machinery of modern geometry, much of which was inherited from the Bourbaki School, was used to compute the various quantities (q.v. the classic [26] and for physicists, [27]), particularly the topological ones such as Euler, Betti and Hodge numbers, which have precise interpretation as Standard Model particles.

1.2 Machine-Learning Algebraic Geometry

The *point d'appui* of [14] was that these large sets of CY manifolds constituted labelled datasets ripe for machine-learning. In fact, the situation is even more general and *any* mathematical computation can be thought of this way. We shall not delve into the details of CY topological invariants or string phenomenology, but the idea can be construed as follows. The purpose of algebraic geometry is to realize a manifold as the vanishing loci of a system of multi-variate polynomials where the variables are the coordinates of some appropriate ambient space such as projective space. We can thus represent a manifold as a list (tensor) of coefficients ². Traditional methods such as exact sequences and Gröbner bases (q.v. [20] for ML on selecting S-pairs) then computes desired geometrical quantities such as Hilbert series or Betti numbers. In the special case of extracting topological quantities, the coefficients are irrelevant (topology does not depend on shape) and we have even simpler representations. For instance, one could record just the degrees of the various defining polynomials.

But a tensor can naturally be interpreted as a pixelated image (up to some normalization and padding if necessary), and thus the general statement of [14, 15] is that

OBSERVATION 1 *Computation in algebraic geometry is an image-recognition problem.*

To make this observation concrete, let us give an example. Suppose we are given a CY 3-fold ³, defined by the intersection of 8 polynomials in a product $(\mathbb{CP}^1)^6 \times (\mathbb{CP}^2)^2$ of projective spaces given by the configuration below ⁴.

¹ By contrast, a CY 1-fold can only be T^2 , a CY 2-fold can only be T^4 and K3. We therefore see the aforementioned exponential growth of possibilities as we increase in dimension. Nevertheless, it is a standing conjecture of Yau that the number of possible topological types of CY in every dimension is *finite*.

² These coefficients determine the “shape” of the manifold. In **Mathematica**, there is a convenient command for this, viz., `CoefficientList[]`.

³ Strictly, this is a family of manifolds since we are not specifying the coefficient which dictate complex structure (shape).

⁴ This is an example of a complete intersection CY in product of project spaces (CICY), which was possibly the first database in algebraic geometry [28]. To read it,

The topological quantity, a so-called Hodge number $h^{2,1}$ was computed (see [27]) to be 22 using long exact sequence in cohomology induced by an Euler sequence (quite a difficult and expensive computation!). However, we could associate 0 to, say, purple, green to 1 and red to 2. After padding with 0 (to normalize over the full CY dataset of which this is one case), and the computation of $h^{2,1}$ becomes an image-processing problem no different than hand-writing recognition:

$$h^{2,1}\left(\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}\right) = 22 \quad \text{becomes} \quad \begin{img alt="A small square image with a purple background and a few scattered green and red pixels." data-bbox="588 241 648 278"/> \longrightarrow 22 . \quad (2)$$

A surprising result of [14] is that such labeled data, consisting of typically around $10^5 \sim 10^6$ points, when fed into a standard ML algorithm, such as a fairly shallow feed-forward neural network (otherwise known as an MLP) with sigmoid activation functions, or a support vector machine (SVP), achieves over 90% accuracy in a standard 80-20 cross-validation ⁵ in a matter of seconds on an ordinary laptop. Since then, more sophisticated neural networks (NNs) have achieved over 99.9% accuracy [29–32] (q.v. recognition of elliptic fibration within the data using ML [33]). How could a relatively simple ML algorithm *guess* at a cohomology computation, without any knowledge of the underlying mathematics? At some level, this is the Universal Approximation Theorem of NNs at work [34], which states that at sufficiently large depth/width, a NN can approximate almost any map, much like the way a Taylor series can approximate any analytic function. Yet, the relative simplicity of the architecture of the NN is highly suggestive of a method which bypasses the sophistication and computational complexity of the standard algorithms of algebraic geometry. To this point let us now turn.

2 The Landscape of Mathematics

The great *utility* of our paradigm to the string landscape, and indeed to problems in theoretical physics, is obvious. Even when not reaching 100% accuracy, a rapid and highly accurate NN estimate could reduce practical computations, say, of searching the exact Standard Model within string string, many orders of magnitude faster. Utility aside, the unexpected success of machine learning of algebraic geometry beckons a deeper question: can one machine learn mathematics [24, 35]? By this we mean several levels: can ML/AI

each column is a defining polynomial. For example, the first column corresponds to a polynomial which is multi-linear in the first and second \mathbb{CP}^1 factors and also linear in the first \mathbb{CP}^2 factor.

⁵ In machine-learning, this means we take the full labelled data, train on 80% randomly selected, and validate - meaning we check what the output is as predicted by the NN versus the actual value - on the *unseen* 20%.

(1) extract patterns from mathematical data, supervised and unsupervised, patterns which have not been noted by the human eye? (2) help formulate new conjectures and find easier formulae (q.v. recent collaboration on how AI can help with mathematical intuition [36])? (3) help with new pathways in a proof? (4) help understand the structure of mathematics across the disciplines?

It is expedient to digress momentarily on some speculations upon the nature of mathematics whilst we are planning to explore Her landscape. The turn of the 20th century witnessed a tension between two Schools of thought: (i) the logicism-formalism of Hilbert and (ii) the intuitionism-constructivism of Poincaré. The first, rested in the tradition of Leibniz, Frege, Peano, Russell-Whitehead, Wittgenstein, et al., and attempted to logically build all of mathematics, without contradiction, symbol by symbol. The second, propelled by Brouwer, Heyting, Poincaré et al., sought for a more “human” and experiential element to mathematics.

The advent of computers in mathematics has dawned a new era. More importantly, they are becoming more than a mere aid to computation. There is a growing number of major results - championed by e.g., the 4-colour theorem, or the classification of finite simple groups - which *could not* have been possible without computer work. The reason is simple: the rate of growth of mathematical knowledge and the requisite length of many a proof have perhaps already exceeded the capacity of the human mind. The full details of the proof of Fermat is hundreds of pages of highly technical mathematics understandable by a small community, that of the classification of simple groups, thousands. It is entirely conceivable that the proof of the Riemann Hypothesis will take longer than several human lifetimes to construct or digest, even if we take into account the cumulative nature of research.

Consequently, Buzzard, Davenport et al. [37, 38] have been emphasising how essential the Automated Theorem Proving programme (ATP) is to the future of mathematics. Software such as **Lean** is currently constructing all statements and proofs of mathematics, symbol by symbol, line by line. Their optimistic estimate is that within 10 years, all of undergraduate level mathematics will be built from scratch automatically. More strikingly, some at Google Deepmind suspect that as computers defeated humans at chess in the 1990s and Go in the 2010s, they will beat us at producing new mathematics by 2030.

The ATP programme can be thought of as being along the formalistic skein of Hilbert, and, to borrow terminology from physics, one could call this “bottom-up mathematics” [35]. Our foregoing discussion of using ML which attempts to extract patterns from data or extrapolate methods from heuristics, on the other hand, is much more along the intuitionistic line of Poincaré. Again, to borrow from physics, one could call this “top-down mathematics” [35]. These two threads should indeed be pursued in parallel and here, we shall summarize some recent experiments in the latter.

2.1 Methodology

For concreteness, let us focus on calculations of the form of (2), which should be ubiquitous in mathematics. We shall let $\mathcal{D} := \{T_i \rightarrow p_i\}$ be a set of input tensors T_i with output property p_i , typically obtained from some exhaustive and intensive computation. We then split $\mathcal{D} := \mathcal{T} \sqcup \mathcal{V}$ into training set \mathcal{T} and validation set \mathcal{V} where \mathcal{T} is a random sample of, say, 80%. Such data, representing “experience and intuition” of the practitioner, could then be passed to standard machine-learning algorithms such as neural classifiers/regressors, MLPs, SVMs, decisions trees, etc. Importantly, these algorithms have *no prior knowledge* of the mathematics⁶. Once validation reaches high precision (especially 100%), one could start formulating conjectures. On the other hand, if one could not reach any good results exhausting a multitude of algorithms, it would indicate an inherent difficulty in the problem whence the data came.

2.2 Across Disciplines

With this method of attack it is natural to scan through the available data of mathematics, as a reconnaissance onto the topography of Her territory. We saw in the above that algebraic geometry over \mathbb{C} responds well to ML and speculate that the reason for this is that all computations inherent thereto reduce to finding (co-)kernels of matrices. Over the past 5 years, there have been various excursions into a variety of disciplines and we shall highlight some representative cases, and refer the reader to the citations as well as the summary in [35].

Algebra: In [46], the question was posed as to whether one could “see” a finite group being simple or not, by direct inspection of its Cayley multiplication table. Surprisingly, an SVM could do so to more than 0.98 precision, instigating the curious conjecture that simple and non-simple groups could be separable when plotting their flattened Cayley tables. For continuous groups, the tensor decomposition into irreps for simple Lie algebras of type $ABCDG_2$ is computationally exponential as one goes up in weight. Yet, numerical quantities such as the number of terms in the decomposition can be quickly machine-learned by an MLP with only a few layers to 0.96 precision [47]. In [48], MLPs, decision trees and graph NNs could distinguish

⁶ Of course, building activation functions which *know* some of the underlying theory is effective and computationally helpful, as was done in, e.g., [39–45], but the true surprises lie in blind tests. This was performed in the initial experiments of [14] and the ones we shall shortly report, could lead to conjectures unfathomed by human thought.

table/non-table ideals to 100% accuracy, whereby suggesting the existence of a yet-unknown formula.

Graphs & Combinatorics: Various properties of finite graphs, such as cyclicity, genus, existence of Euler or Hamilton cycles, etc., were explored by “looking” at the the adjacency matrix with MLPs and SVMs [49]. The algorithms determining some of these quantities are quite involved indeed. For instance, Hamilton cycle detection is that of the traveling salesman problem, which is NP-hard. Typically, for these problems, one could reach 80-90% accuracies, which could be related to the fact that detecting matrix permutations - and hence graph isomorphism - is currently a challenge to ML. However, when more structures are put in, such as quiver representations [50], or tropical geometry [51], accuracies in the high 90s can once more be attained. Explorations in lattice polytopes [52,53] and knot invariants [54,55] also yield good results.

Analytic Number Theory: As one might imagine, uncovering patterns in arithmetic functions, such as prime characteristic, or the likes of Mobius μ and Liouville λ , would be very hard. And it turns out to be so not only for the human eye, but also for any standard ML algorithm [14, 24, 35]. Likewise, one would imagine finding new patterns in the Riemann zeta function [56, 57] to be a formidable challenge.

Arithmetic Geometry: Yet, with a mixture of initial astonishment and a posteriori reassurance, problems in arithmetic geometry are very much amenable to ML. Properties such as the arithmetic of L-functions [58, 59], degree of Galois extensions for dessins d’enfants [62], or even the quantities pertaining to the strong Birch-Swinnerton-Dyer conjecture [60, 61] (interestingly, the most difficult Tate-Shafarevich group is the least responsive) can all be learnt to high accuracies. Indeed, as exemplified by countless historical cases, translating Diophantine problems to geometry, especially that of (hyper-)elliptic curves, renders them much more tractable. In this sense, our ML methodology and results on the data are consistent with this notion that arithmetic geometry is closer to geometry than to arithmetic.

With these experiments, we conclude with the remark and speculation that there is a “hierarchy” of mathematical problems, perhaps in tune with our expectations:

OBSERVATION 2 *Across the disciplines of mathematics,*

$$[\textit{numerical analysis}] < [\textit{algebraic geometry over } \mathbb{C} \sim \textit{arithmetic geometry}] < \\ [\textit{algebra/representation theory}] < [\textit{combinatorics}] < [\textit{analytic number theory}]$$

where $a < b$ means patterns from problem from a are more easily extractable than those from b , or indeed that problems in a are more easily solvable.

Above all, we encourage the readers to take their favourite problems and data and see how well ML performs on them.



Acknowledgements *Ad Musam Alphabeticam meam tenebris luscis siderium.*

The author is grateful to STFC UK for grant ST/J00037X/2, as well as the kind hospitality - virtual and in person (!) during 2021 - of Cambridge (Winter School on ML), Cairo (BSM 22), Sofia (Lie Theory in Physics XIV), Toulouse (Geometry, Topology & AI), Lisboa (BH, BPS, & QI), Trento (ML for HEP), Singapore (M-theory & Beyond), Johannesburg (String Data 22), Bangalore (KAWS 22), Hang-Zhou, Pisa, Minnesota (IMA), Tokyo (NaapingClass), Leicester, and Galway, on the talks at which this review is based.

References

1. S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, “De Sitter vacua in string theory,” *Phys. Rev. D* **68** (2003), 046005 [arXiv:hep-th/0301240 [hep-th]].
2. W. Taylor and Y. N. Wang, “The F-theory geometry with most flux vacua,” *JHEP* **12** (2015), 164 [arXiv:1511.03209 [hep-th]].
3. J. Halverson, C. Long, B. Sung, “Algorithmic universality in F-theory compactifications,” *Phys. Rev. D* **96** (2017) no.12, 126006 [arXiv:1706.02299 [hep-th]].
4. P. Candelas, X. de la Ossa, Y. H. He and B. Szendroi, “Triadophilia: A Special Corner in the Landscape,” *Adv. Theor. Math. Phys.* **12** (2008) no.2, 429-473 [arXiv:0706.3134 [hep-th]].
5. M. R. Douglas, “The Statistics of string / M theory vacua,” *JHEP* **05** (2003), 046 [arXiv:hep-th/0303194 [hep-th]].
6. V. Braun, Y. H. He, B. A. Ovrut and T. Pantev, “The Exact MSSM spectrum from string theory,” *JHEP* **05** (2006), 043 [arXiv:hep-th/0512177 [hep-th]].
7. V. Bouchard and R. Donagi, “An SU(5) heterotic standard model,” *Phys. Lett. B* **633** (2006), 783-791 [arXiv:hep-th/0512149 [hep-th]].
8. F. Gmeiner, R. Blumenhagen, G. Honecker, D. Lust and T. Weigand, “One in a billion: MSSM-like D-brane statistics,” *JHEP* **01** (2006), 004 [arXiv:hep-th/0510170 [hep-th]].
9. L. B. Anderson, J. Gray, A. Lukas and E. Palti, “Two Hundred Heterotic Standard Models on Smooth Calabi-Yau Threefolds,” *Phys. Rev. D* **84** (2011), 106005 [arXiv:1106.4804 [hep-th]].
10. M. Cvetič, J. Halverson, L. Lin and C. Long, “Constraints on Standard Model Constructions in F-theory,” *Phys. Rev. D* **102** (2020) no.2, 026012 [arXiv:2004.00630 [hep-th]].
11. A. Constantin, Y. H. He and A. Lukas, “Counting String Theory Standard Models,” *Phys. Lett. B* **792** (2019), 258-262 [arXiv:1810.00444 [hep-th]].
12. J. Halverson and F. Ruehle, “Computational Complexity of Vacua and Near-Vacua in Field and String Theory,” *Phys. Rev. D* **99** (2019) no.4, 046015 [arXiv:1809.08279 [hep-th]].
13. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature* 521 no. 7553, (2015) 436–444.
14. Y. H. He, “Deep-Learning the Landscape,” [arXiv:1706.02714 [hep-th]]. q.v., Feature article, M. Hutson, *Science*, 2019, doi: 10.1126/science.aay9534
15. Y. H. He, “Machine-learning the string landscape,” *Phys. Lett. B* **774** (2017), 564-568
16. D. Krefl and R. K. Seong, “Machine Learning of Calabi-Yau Volumes,” *Phys. Rev. D* **96** (2017) no.6, 066014 [arXiv:1706.03346 [hep-th]].
17. J. Cariffo, J. Halverson, D. Krioukov and B. D. Nelson, “Machine Learning in the String Landscape,” *JHEP* **09** (2017), 157 [arXiv:1707.00655 [hep-th]].
18. F. Ruehle, “Evolving neural networks with genetic algorithms to study the String Landscape,” *JHEP* **08** (2017), 038 [arXiv:1706.07024 [hep-th]].
19. Y. H. He, S. Lal and M. Z. Zaz, “The World in a Grain of Sand: Condensing the String Vacuum Degeneracy,” [arXiv:2111.04761 [hep-th]].
20. D. Peifer, M. Stillman, D. Halpern-Leistner, “Learning selection strategies in Buchberger’s algorithm”, [arXiv:2005.01917]
21. Calabi, Eugenio, “The space of Kähler metrics”, *Proc. Internat. Congress Math. Amsterdam*, 2, pp. 206 - 207 (1954); “On Kähler manifolds with vanishing canonical class”, in Fox, Spencer, Tucker, *Algebraic geometry and topology. A symposium in honor of S. Lefschetz*, Princeton Mathematical Series, 12, PUP, pp. 78 - 89 (1957).
22. S.-T. Yau, “Calabi’s conjecture and some new results in algebraic geometry,” *Proc. Nat. Acad., USA*, 74 (5), pp 1798-9, (1977); “On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation I”, *Comm. Pure and Applied Maths*, 31 (3), pp 339-411, (1978).
23. P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, “Vacuum Configurations for Superstrings,” *Nucl. Phys. B* **258**, 46 (1985).
24. Y. H. He, “The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning,” *Springer LNM* 2293, doi:10.1007/978-3-030-77562-9, 2021 [arXiv:1812.02893 [hep-th]].

25. M. Kreuzer and H. Skarke, "On the classification of reflexive polyhedra," *Commun. Math. Phys.* **185**, 495 (1997) [hep-th/9512204]; "Reflexive polyhedra, weights and toric Calabi-Yau fibrations," *Rev. Math. Phys.* **14**, 343 (2002) [math/0001106 [math-ag]].
26. Robin Hartshorne, "Algebraic Geometry", GTM, 1997, ISBN 13: 9780387902449.
27. T. Hubsch, *Calabi-Yau manifolds: A Bestiary for physicists*, WS, 1994, ISBN 9810206623
28. P. Candelas, A. M. Dale, C. A. Lutken, R. Schimmrigk, "Complete Intersection Calabi-Yau Manifolds," *Nucl. Phys. B* **298**, 493 (1988).
29. K. Bull, Y. H. He, V. Jejjala and C. Mishra, "Machine Learning CICY Threefolds," *Phys. Lett. B* **785** (2018), 65-72 [arXiv:1806.03121 [hep-th]].
30. S. Krippendorf, M. Syvaeri, "Detecting Symmetries with Neural Networks," [arXiv:2003.13679 [physics.comp-ph]].
31. H. Erbin, R. Finotello, "ML for complete intersection Calabi-Yau manifolds: a methodological study," *PRD* **103** (2021) no.12, 126014 [arXiv:2007.15706 [hep-th]].
32. R. Altman, J. Carifio, X. Gao and B. Nelson, "Orientifold Calabi-Yau Threefolds with Divisor Involutions and String Landscape," [arXiv:2111.03078 [hep-th]].
33. Y. H. He and S. J. Lee, "Distinguishing elliptic fibrations with AI," *Phys. Lett. B* **798** (2019), 134889 [arXiv:1904.08530 [hep-th]].
34. K. Hornik, M. Tinchcombe, H. White, "Multilayer Feedforward Networks are Universal Approximators". *Neural Networks*. 2. Pergamon Press. pp. 359 - 366.
35. Y. H. He, "Machine-Learning Mathematical Structures," [arXiv:2101.06317 [cs.LG]].
36. Davies, A., Velickovic, P., Buesing, L. et al. Advancing mathematics by guiding human intuition with AI. *Nature* **600**, 70–74 (2021). <https://doi.org/10.1038/s41586-021-04086-x>
37. K. Buzzard, "The future of Mathematics", https://wwwf.imperial.ac.uk/~buzzard/one_off_lectures/msr.pdf; <https://www.youtube.com/watch?v=Dp-mQ3HxgDE>; The Xena Project, <https://wwwf.imperial.ac.uk/~buzzard/xena/>
38. J. Davenport, B. Poonen, J. Maynard, H. Helfgott, P. H. Tiep, L. Cruz-Filipe, "Machine-Assisted Proofs", ICM Panel 2018
39. R. Deen, Y. H. He, S. J. Lee, A. Lukas, "ML String SMs," [arXiv:2003.13339 [hep-th]].
40. K. Heal, A. Kulkarni, E. Sertoz, "Deep Learning Gauss-Manin Connections", arXiv:2007.13786 [cs.LG]
41. V. Jejjala, D. K. Mayorga Pena and C. Mishra, "Neural Network Approximations for Calabi-Yau Metrics," [arXiv:2012.15821 [hep-th]].
42. M. R. Douglas, S. Lakshminarasimhan and Y. Qi, "Numerical Calabi-Yau metrics from holomorphic networks," [arXiv:2012.04797 [hep-th]].
43. L. B. Anderson, M. Gerdes, J. Gray, S. Krippendorf, N. Raghuram and F. Ruehle, "Moduli-dependent Calabi-Yau and SU(3)-structure metrics from Machine Learning," *JHEP* **05** (2021), 013 [arXiv:2012.04656 [hep-th]].
44. M. Larfors, A. Lukas, F. Ruehle and R. Schneider, "Learning Size and Shape of Calabi-Yau Spaces," [arXiv:2111.01436 [hep-th]].
45. X. Gao, H. Zou, "ML to the Orientifold CY with String Vacua," [arXiv:2112.04950 [hep-th]].
46. Y. H. He and M. Kim, "Learning Algebraic Structures: Preliminary Investigations," [arXiv:1905.02263 [cs.LG]].
47. H. Y. Chen, Y. H. He, S. Lal, S. Majumder, "Machine learning Lie structures & applications to physics," *Phys. Lett. B* **817** (2021), 136297 [arXiv:2011.00871 [hep-th]].
48. L. Amoros, O. Gasanova, L. Jakobsson, "A machine learning approach to commutative algebra: Distinguishing table vs non-table ideals", arXiv:2109.11417 [math.AC]
49. Y. H. He and S. T. Yau, "Graph Laplacians, Riemannian Manifolds and their Machine-Learning," [arXiv:2006.16619 [math.CO]].
50. J. Bao, S. Franco, Y. H. He, E. Hirst, G. Musiker, Y. Xiao, "Quiver Mutations, Seiberg Duality and ML," *PRD* **102** (2020) no.8, 086013 [arXiv:2006.10783 [hep-th]].
51. J. Bao, Y. H. He, E. Hirst, "Neurons on Amoebae," [arXiv:2106.03695 [math.AG]].
52. J. Bao, Y. H. He, E. Hirst, J. Hofscheier, A. Kasprzyk, S. Majumder, "Polytopes and ML," [arXiv:2109.09602 [math.CO]].
53. P. Berglund, B. Campbell and V. Jejjala, "Machine Learning Kreuzer-Skarke Calabi-Yau Threefolds," [arXiv:2112.09117 [hep-th]].
54. S. Gukov, J. Halverson, F. Ruehle and P. Sulkowski, "Learning to Unknot," *Mach. Learn. Sci. Tech.* **2** (2021) no.2, 025035 [arXiv:2010.16263 [math.GT]].
55. J. Craven, M. Hughes, V. Jejjala and A. Kar, "Learning knot invariants across dimensions," [arXiv:2112.00016 [hep-th]].
56. O. Shanker, "NN prediction of Riemann zeta zeros", *Adv. Model. & Optim.*, **14**, 3, 2012.
57. J. Kampe, A. Vysogorets, "Predicting Zeros of the Riemann Zeta Function Using Machine Learning: A Comparative Analysis", <http://www.sci.sdsu.edu/math-reu/2018-2.pdf>
58. Y. H. He, K. H. Lee, T. Oliver, "ML the Sato-Tate Conjecture," [arXiv:2010.01213 [math.NT]]. *J. Symbolic Computation* **111**, 61.
59. Y. H. He, K. H. Lee, T. Oliver, "ML Number Fields," [arXiv:2011.08958 [math.NT]].
60. L. Alessandretti, A. Baronchelli, Y. H. He, "ML meets Number Theory: The Data Science of BSD," [arXiv:1911.02008 [math.NT]], to appear, in *Stemma Journal*.
61. Y. H. He, K. H. Lee, T. Oliver, "ML Arithmetic Curves," [arXiv:2012.04084 [math.NT]].
62. Y. H. He, E. Hirst, T. Peterken, "ML dessins d'enfants: explorations via modular and Seiberg-Witten curves," *J. Phys. A* **54** (2021) no.7, 075401 [arXiv:2004.05218 [hep-th]].