Universes as Big Data

Yang-Hui He*

[1] *Merton College, University of Oxford, OX14JD, UK*
[2] *Department of Mathematics, City, University of London, EC1V 0HB, UK*
[3] *School of Physics, NanKai University, Tianjin, 300071, P.R. China*

Abstract

We briefly overview how, historically, string theory led theoretical physics first to precise problems in algebraic and differential geometry, and thence to computational geometry in the last decade or so, and now, in the last few years, to data science. Using the Calabi-Yau landscape - accumulated by the collaboration of physicists, mathematicians and computer scientists over the last 4 decades - as a starting-point and concrete playground, we review some recent progress in machine-learning applied to the sifting through of possible universes from compactification, as well as wider problems in geometrical engineering of quantum field theories. In parallel, we discuss the programme in machine-learning mathematical structures and address the tantalizing question of how it helps doing mathematics, ranging from mathematical physics, to geometry, to representation theory, to combinatorics, and to number theory.

*hey@math.ox.ac.uk

Invited review for IJMPA, based on various colloquia, seminars and conference talks in the 2019-2020 academic year.

Contents

1 Introduction

The twentieth century has firmly established that the correct language of fundamental theoretical physics is that of algebraic/differential geometry/topology, in all four combinations of these pairs of adjectives and nouns. Gravity and space-time, should be thought of as the metric and curvature of Riemannian manifolds; elementary particles, irreducible representations of the Lorentz group and gauge connections of

appropriate principal Lie-group fibrations, etc. (q.v. an attempted modern summary in [1]). In some sense, string theory is a brain-child of this tradition. Whether she will stand as the ultimate theory of everything remains to be seen, but her rôle in both bearing the torch and ploughing the field of the conversations between mathematics and physics is unquestionable.

The twenty-first century, with the dramatic progress in computing power and techniques, is bringing a new interlocutor to this dialogue. Already, in the first decade, software such as Macaulay2 [2], Singular [3], GAP [4], MAGMA [5], and the umbrella project of SageMath [6] (launched in 2005), and the increasing online mathematical databases – [7–9] to name but a few – are aiding pure mathematical research in an increasingly prominent way (q.v.. launch of ICMS in 2006 [10]). In parallel – and this is of course in tandem with experimental physics whose reliance on and interaction with computers has a rich history of its own – theoretical physics, and string theory in particular, has benefited from an algorithmic outlook [13–16].

In our present era of Big Data and AI, it is inevitable that machine-learning should have an ever-increasing presence in the second decade [17–21]. The purpose of this talk [11], aimed at a general physics audience, is to give an overview of some of this activity in the last few years, especially in the context of machine-learning applied to the string theoretic and geometric landscape, as well as to other mathematical structures within and beyond geometry (cf. an attempted pedagogical introduction in [22]). I will take a somewhat historical approach and start with Calabi-Yau data as a concrete playground; this is mainly due to the vastness of the subject in physics and in mathematics, which has consequently led to an abundance of data. The methodology should be applicable to much more general situations, under the rubric of machine-learning mathematical structures [12, 17, 23].

The organization of this talk is as follows. We begin by reviewing how 2 parallel traditions, one in theoretical physics and one in pure mathematics, converging around 1980s and both leading to the study of complex manifolds and Kähler geometry. This confluence initiated a concerted effort to construct Ricci-flat such spaces, viz., Calabi-Yau manifolds, in the last decade of the 20th century, continuing into the 21st, developing into an explosion of data. In §2, we take an overview of this mathematical data, being encouraged by its availability and plenitude, somehow daunted by the complexity of the algorithms needed to process them, and compelled by a thirst for techniques from the "Big Data" revolution and AI research. We then offer the

audience an invitation, in §3, to some modern data science, focusing on machine-learning and how it may be applied to problems in string phenomenology as well as algebraic geometry. We conclude with an outlook and report on some recent results in machine-learning fundamental structures in various branches of mathematics and relations to physics.

2 Trichotomy and Triadophilia

It is well known that string theory is a unified theory of gravity and elementary particles in high dimensions. Shortly after the First Revolution in 1984 with anomaly cancellation [25] and the discovery of the heterotic string [26], the subject of "string phenomenology" was born [27]. The reason for this excitement was that all at once, there was an anomaly-free quantum field theory which naturally contained the graviton as well as the E_8 gauge group. In other words, it presented a unified theory of quantum gravity - albeit in 10 space-time dimensions - which also, via the embedding $SU(3)\times SU(2)\times U(1) \subset SU(5) \subset SO(10) \subset E_6 \subset E_8$, could give rise to the standard model.

2.1 Complex, Kähler, Ricci-flat

The solution of [27] was to take inspiration from Kaluza-Klein [28] and treat the extra $10-4=6$ dimensions as small and space-like, in a *compactification* scenario with a 6-manifold M_6 on top of each point in our 4-dimensional space-time. Further conditions of supersymmetry [†] and vacuum Einstein solutions constrained the 6-manifold to be (1) complex, (2) Kähler and (3) Ricci-flat, i.e., respectively (1) a complex 3-fold, (2) the metric comes from a scalar potential $g_{\mu\bar{\nu}}(z,\bar{z}) = \partial_\mu \overline{\partial}_{\bar{\nu}} K(z,\overline{z})$ and (3) the Ricci curvature for $g_{\mu\bar{\nu}}$ vanishes. We emphasize that this is the simplest solution. In general, one has to solve the so-called *Hull-Strominger* system [31], which would lead

[†]Whilst it remains to be seen whether there is supersymmetry in Nature, it is undisputed from a theoretical perspective that quantum field theory with supersymmetry (SUSY) has much richer and tamable structure. A good (and in some sense rigourous) analogy would be that doing mathematics over $\mathbb{R}$ is difficult, and this is ameliorated by working over $\mathbb{C}$, the unique (commutative) algebraic closure. So too, do the theorems of Coleman-Mandula [29] and Haag-Łopuszański-Sohnius [30], guarantee SUSY as the unique extension to Poincaré symmetry in a field theory.

to a much wider variety of possible M_6. As we shall emphasize later, the Calabi-Yau landscape is only a corner of possible compactification scenarios.

To the theoretical physicist in the mid-1980s, perhaps only the word "Ricci-flat" was, because of general relativity, familiar. Meanwhile, for the mathematical community, this was also rather avant-garde. The story goes back to classical results of Euler, Gauß, and Riemann. Consider a surface Σ - we usually think of a sphere S^2 or the surface of a doughnut T^2 - and its possible **topological types**, i.e., equivalences up to topology. Restricting to the cases of smooth, compact (no punctures or boundaries) and orientable (nothing like Klein bottles or Möbius strips) surfaces, the familiar shapes, S^2, T^2, and those with increasing number of "holes" (genus) are all there is: any smooth, compact, orientable surface can be deformed continuously (topologically homeomorphic) to one of these. This single non-negative integer, the genus $g(\Sigma)$, classifies the topology of Σ. Closely related is the quantity $\chi(\Sigma) = 2 - 2g(\Sigma)$, called the **Euler characteristic** or Euler number. A high-light of the geometry of the 18th-19th centuries is the chain of equalities

$$2 - 2g(\Sigma) = \chi(\Sigma) = \sum_{i=0}^{\dim_{\mathbb{R}} \Sigma = 2} (-1)^i b^i(\Sigma) = \frac{1}{2\pi} \int_\Sigma R \ , \tag{2.1}$$

where one proceeds, from left to right, from topology, to combinatorics, to Gauß' Theorema Egregium for differential geometry. Here, b^i are the **Betti numbers**, counting the number of cycles in dimension i and R is the (Ricci) curvature. This whole setup above can be complexified where our 2-manifolds – named *Riemann surfaces* – become complex 1-folds, named *complex curves*. Furthermore, Σ are not just complex, but are also Kähler – one can check that the complex (Hermitian) metric on all such surfaces as complex 1-folds comes from a single potential.

To the equalities (2.1), early 20th century added another, viz., $\chi(\Sigma) = [c_1(T_\Sigma)]\cdot[\Sigma]$, where the last integral is re-interpreted as intersection theory between cohomology (here the Chern class c_1) and homology (here the class of the manifold). All of these are special cases of the index theorem of Atiyah-Singer and Grothendieck-Riemann-Roch (q.v. [32]) which are applicable to spaces of arbitrary dimension.

Having curvature controlling topology also gives us a natural **trichotomy**, which for Σ, is part of the Riemann Uniformization Theorem. Specifically, in complex

dimension 1, we have

$$R \quad \begin{cases} > 0: & g = 0, \text{ Spherical Geometry} \\ = 0: & g = 1, \text{ Flat Torus} \\ < 0: & g > 1, \text{ Hyperbolic Geometry} \end{cases} \tag{2.2}$$

Note that the $R \geq= 0$ cases are finite in topological type and the $R < 0$ cases are infinite.

Much of modern geometry is concerned with generalizing this beautiful story of complex dimension 1 to higher dimensions. Expectedly, the situation is much more involved and many questions still remain open conjectures. Nevertheless, for Kähler manifolds a conjecture of Calabi [34] dating to the 1950s does give the analogue of (2.2): essentially, it states that c_1, the first Chern class, uniquely controls the Ricci curvature for the Kähler metric. It was not until the Fields-Medal-deserving work in 1978 by Yau [35] that this was settled.

Fortuitously, Strominger, one of authors of [27] was visiting Yau at the IAS in 1985 and were neighbours. Thus, the object onto which physicists stumbled – Ricci-flat, Kähler manifolds – through string compactification, had the world-expert literally next door. In fact, such spaces were named **Calabi-Yau** manifolds by the physicists.

2.2 Low-Energy Physics

Not only did [27] constrain the compactification manifold, they also established a dictionary between

$$\text{“Geometry of } X_6 \longleftrightarrow \text{ physics of } \mathbb{R}^{1,3}\text{.”}$$

Purely working from the group theory, the tangent bundle with its $SU(3)$ structure breaks the E_8 to an E_6 (SUSY) GUT theory. We will skip the details, but basically for the fundamental fermions (the choice of which is anti-generation and generation

is by convention):

$$\begin{aligned} \text{generations of particles} &\sim h^{2,1}(X) \ , \\ \text{anti-generations of particles} &\sim h^{1,1}(X) \ , \end{aligned} \tag{2.3}$$

where $h^{p,q}$ are the **Hodge numbers** of the Calabi-Yau manifold M_6, the complexified version (hence the double index, for complex and conjugate) of the Betti numbers mentioned earlier. The alternating sum of Betti numbers to the Euler number generalizes to a double sum, and in particular $\chi(M_6) = 2(h^{1,1}(M_6) - h^{2,1}(M_6))$. Since there are 3 generations of fermions, one of the original constraints of [27] is that

$$\left|h^{2,1}(X) - h^{1,1}(X)\right| = 3 \Rightarrow \chi(X) = \pm 6 \ . \tag{2.4}$$

Finding compact, smooth Calabi-Yau 3-folds with Euler number ± 6 was perhaps historically the first concrete challenge physicists gave to the algebraic geometry community.

Disclaimer: It must be emphasized that (2.3) is only for the so-called standard embedding for the heterotic string to get to E_6-GUT theories, the field has since evolved to far beyond merely computing Hodge numbers, but to computing equivariant cohomology of stable bundles (cf. [37–40]). In addition, there is a myriad of phenomenological approaches from other string/M-/F-theoretic constructions, which constitute the vastness of the "string landscape", the review of which is not our present intent. The reader is referred to the wonderful textbooks [36] in general, and to the classic [24] for an introduction to complex geometry for physicists, and, in the context of Calabi-Yau spaces, to [33] for a brief invitation and [22] for a pedagogical textbook.

Although the physics community is no longer searching for manifolds with property (2.4), there is an entire programme, especially led by Candelas, to look for Calabi-Yau manifolds of *small* Hodge numbers [42] which have interesting mathematics of its own. In any event, the search for a geometric interpretation, or origin, of 3 generations of particles has been dubbed "Triadophilia" [41]. In a way, the dictionary started by (2.3), where properties of our universe are purely phrased in the geometry of some manifold, is an elegant modern realization of Kepler's famous adage: "Ubi

materia, ibi geometria" [‡]. Perhaps for this reason by itself, it is worth studying string theory as a theory of physics, let alone its cross-fertilizations to mathematics and - as we will see - data science.

2.3 Early Constructions

One of the first questions which the physicists asked Yau was, indubitably, "how to construct an explicit Calabi-Yau 3-fold?" It is curious that students in theoretical physics are taught differential geometry first, before algebraic geometry, whereas the fundamental ideas of the latter - vanishing loci of polynomials - are certainly more familiar than that of the former - local patches and differentiable transition functions. We know how to construct shapes from Cartesian geometry since our early school days. For example, a quadratic equation in two real variables (x, y) is a conic section, such as a circle. Thus, the vanishing locus on a quadratic polynomial in two real variables gives a $2 - 1 = 1$ dimensional real manifold in an ambient $\mathbb{R}^2$. We have just created a simple **algebraic variety**.

Now, we are looking for complex manifolds, we thus construct them as the zero-locus of multiple polynomials in multiple complex variables. In this way, a Calabi-Yau 1-fold, a Riemann surface of zero curvature, viz. the torus $T^2 = S^1 \times S^1$, is realized as a cubic in two complex variables given by the so-called Weierstraß equation $T^2 \simeq \{x, y \in \mathbb{C} | y^2 = x^3 - g_2 x - g_4\} \subset \mathbb{C}^2$, where $g_{2,4}$ are complex constants. One can check by writing out (x, y) in their real and imaginary parts, and the Weierstraß equation becomes 2 real constraints in 4 real variables, which we can numerically plot by Monte Carlo to see a torus emerge. Next, compactness can be ensured by including the point at infinity, where $(x, y) = (\infty, \infty)$. One can do this by so-called *projectivization* where instead of $\mathbb{C}^2$, we introduce one more complex coordinate, z such that any point $(x, y, z) \in \mathbb{C}^3$ is identified with the scaled $\lambda(x, y, z)$ for non-zero $\lambda \in \mathbb{C}$. This scale-invariance brings the point at infinity to a finite point, rendering the resulting ambient space and the subsequent torus compact.

What we have done is to construct, from $\mathbb{C}^3$ with coordinates (x, y, z), the complex projective space $\mathbb{CP}^2$ with **homogeneous** coordinates $[x : y : z]$. More formally, we define $\mathbb{CP}^n$ from $\mathbb{C}^{n+1}$ with coordinates $(z_0, z_1, \ldots, z_n)$ as the quotient by the

[‡] "Where there is matter, there is geometry," from Johannes Kepler's Thesis XX from *De fundamentis astrologiae certioribus* 1602.

equivalence relation $\sim$

$$\mathbb{CP}^n := \mathbb{C}^{n+1}\backslash\{\vec{0}\} \Big/ (z_0, z_1, \ldots, z_n) \sim \lambda(z_0, \ldots, z_n) \ , \qquad \lambda \in \mathbb{C}\backslash\{0\} \ . \tag{2.5}$$

The complex n-fold $\mathbb{CP}^n$ is smooth, with $n+1$ homogeneous coordinates.

Thus, the Calabi-Yau 1-fold is realized as a so-called projective algebraic variety inside $\mathbb{CP}^2$

$$\{[x:y:z] | -y^2z + x^3 - 4g_2xz^2 - g_4z^3 = 0\} \subset \mathbb{CP}^2 \ , \tag{2.6}$$

a homogeneous cubic in the homogeneous coordinates $[x:y:z]$ of $\mathbb{CP}^2$. Luckily, complex projective space and the zero loci of any number homogeneous polynomials therein, are guaranteed to be Kähler. For $\mathbb{CP}^n$, the Kähler metric explicitly comes from the famous Fubini-Study potential $\log(1 + \sum_i |z_i|^2)$. This construction is valid in general: the hypersurface defined by a homogeneous polynomial of degree $n+1$ in $\mathbb{CP}^n$ is a Calabi-Yau $(n-1)$-fold. Thus we arrive at our first, and perhaps most famous, example of a Calabi-Yau 3-fold: the quintic hypersurface in $\mathbb{CP}^4$. There are many degree 5 monomials one could compose of 5 coordinates, the most well-studied is the so-called Fermat quintic:

$$Q := \{x_0^5 + x_1^5 + x_2^5 + x_3^5 + x_4^5 = 0\} \subset \mathbb{CP}^4_{[x_0:x_1:x_2:x_3:x_4]} \ . \tag{2.7}$$

What are the topological numbers of Q? The Hodge numbers turn out to be $h^{2,1}(Q) = 101$ and $h^{1,1}(Q) = 1$ so that $\chi(Q) = 2(1 - 101) = -200$.

Immediately, we also obtain 4 close relatives. Consider the intersection of 2 cubics in $\mathbb{CP}^5$; this is a complete intersection in that the number of defining polynomials - here 2 - is equal to the codimension – i.e., the dimension of the ambient $\mathbb{CP}^5$ minus the dimension of the required manifold, $5 - 3 = 2$. We denote this as $[5|3,3]$, much as we could denote the quintic as $[4|5]$. Note that the number of to the left of the bar is 1 less than the row-sum (Calabi-Yau condition) and also 3 more than the number of columns (complete intersection condition). A simple integer partition shows that there are 5 possibilities in total, including the quintic, viz.,

$$[4|5], \ [5|3,3], \ [5|2,4], \ [6|2,2,3], \ [7|2,2,2,2] \ . \tag{2.8}$$

These are called *cyclic* Calabi-Yau 3-folds, and are the only ones as complete intersections in a single projective space.

We need to emphasize that complete intersections are rare and most algebraic varieties are *not* so. In fact, there is a general result that (see [44])

THEOREM 1. *All Kähler 3-folds can be realized as vanishing loci of systems of polynomials in* $\mathbb{CP}^7$.

Therefore, one could in principle write all sorts of (non-complete-intersection) polynomials in 8 homogeneous variables and sift out the Calabi-Yau ones; but this is highly impractical.

Of the 5 immediate ones, none has the property (2.4), so the community turned to more general constructions. Again, as mentioned earlier, today physicists are no longer limited to (2.4) and (2.8) have all been met with renew zest. In the late 1980s, however, a different path was undertaken, and an industry of subsequent generalizations to (2.8) was initiated:

CICYs The first generalization is to take, instead of a single $\mathbb{CP}^n$, a product A of projective spaces. That is, let $A = \mathbb{CP}^{n_1} \times \ldots \times \mathbb{CP}^{n_m}$, of dimension $n = n_1+n_2+\ldots+n_m$ and each having homogeneous coordinates $[x_1^{(r)} : x_2^{(r)} : \ldots : x_{n_r}^{(r)}]$ with the superscript $(r) = n_1, n_2, \ldots, n_m$ indexing the projective space factors. The Calabi-Yau 3-fold is then defined as the complete intersection of $K = n-3$ homogeneous polynomials in the coordinates $x_j^{(r)}$. Succinctly [§], this information can be written into an $m \times K$ configuration matrix which generalize (2.8):

$$X = \left[\begin{array}{c|cccc} \mathbb{CP}^{n_1} & q_1^1 & q_2^1 & \ldots & q_K^1 \\ \mathbb{CP}^{n_2} & q_1^2 & q_2^2 & \ldots & q_K^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbb{CP}^{n_m} & q_1^m & q_2^m & \ldots & q_K^m \end{array}\right]_{m\times K} , \qquad \begin{array}{l} K = \sum_{r=1}^{m} n_r - 3 \ , \\ \sum_{j=1}^{K} q_j^r = n_r + 1 \ , \ \forall\, r = 1, \ldots, m \ . \end{array} \tag{2.9}$$

These manifolds defined by (2.9) were called CICYs (complete intersection Calabi-Yau manifolds) and were explicitly constructed by Candelas et al. [45]

[§]Importantly, the Chern classes and the Euler number can be read off the matrix configuration explicitly. The individual terms $(h^{1,1}, h^{2,1})$, however, cannot be deduced from the configuration matrix directly. This is one of the short-comings of the index theorem: the integral of curvature and the intersection of the Chern classes give only the alternating sum (Euler number) in (co-)homology, but not the individual terms.

(q.v. Hübsch's classic book [46]) in the early 1990s. The combinatorial problem for these integer matrices turned out to be rather non-trivial and one of the most powerful super-computers then available, the one at CERN, was recruited. To our knowledge, this might have been the first "data-base" in algebraic geometry.

Up to trivial equivalence such as row/column permutations as well as non-trivial ones such as so-called splitting, CICYs were shown to be finite in number, a total of 7890 configurations, with a maximum of 12 rows, a maximum of 15 columns, and all having entries $q_j^r \in [0,5]$. There are 266 distinct Hodge pairs $(h^{1,1}, h^{2,1}) = (1,65), \ldots, (19,19)$, giving 70 distinct Euler numbers $\chi \in [-200, 0]$.

WP4s Noticing that the CICY data is rather skewed in that all Euler numbers were non-positive, the constructions went on. The reason for this is that physicists knew about mirror symmetry by then, one of whose most salient features is the exchange of $h^{1,1} \leftrightarrow h^{2,1}$, which would flip the sign of χ. Another way to generalize projective space is to introduce weights [¶]. That is, one takes *weighted* projective space $\mathbb{CP}^4_{[d_0:\ldots:d_4]}$ as the ambient space A, which generalizes (2.5) by having integer "weights" $(d_0, d_1, d_2, d_3, d_4) \in \mathbb{Z}_+$ as

$$\mathbb{CP}^4_{[d_0:\ldots:d_4]} := \mathbb{C}^5 \backslash \{\vec{0}\} \Big/ \big((z_0, z_1, \ldots, z_4) \sim (\lambda^{d_0} z_0, \ldots, \lambda^{d_4} z_4)\big) \ , \qquad \lambda \in \mathbb{C}\backslash\{0\} \ . \tag{2.10}$$

Taking all weights $d_i = 1$ is the ordinary $\mathbb{CP}^4$. As with Q, if we embed a hypersurface of degree $d_0 + d_1 + \ldots + d_4$ into $\mathbb{CP}^4_{[d_0:\ldots:d_4]}$, it defines a CY_3. The classification of such manifolds was performed in [47] and a total of 7555 is found, with 2780 distinct Hodge pairs and a more balanced $\chi \in [-960, 960]$.

Reflexive Polytopes The next systematic generalization of weighted projective space is a **toric variety**, which, instead of having a single list of weights as in (2.10), has a list of m weights (giving a so-called charge-matrix) acting on $\mathbb{C}^{n+m}$ to give an n-fold. Based on the theorem of Batyrev-Borisov [48], Kreuzer and Skarke spent almost a decade explicitly constructing such Calabi-Yau manifolds, culminating in the early 2000s with the construction of the most extensive database of CY_3 so far, the **Toric Hypersurfaces** [13].

In brief, the ambient space is a toric 4-fold A, constructed from an integer polytope $\Delta \subset \mathbb{R}^4$ which is **reflexive**, meaning that Δ has a single interior point (which can be taken to be the origin) and all bounding hyperplanes are distance

[¶] In fact, products of projective spaces can also be thought of as a weighted projective space with vector-valued grading.

1 from this point. Furthermore, a particular hypersurface in the toric variety A is a CY with defining equation of the CY_3 is given by

$$X = \{\sum_{\vec{m}\in\Delta} c_{\vec{m}} \prod_{j=1}^{k} x_j^{\vec{m}\cdot\vec{v}_j+1} = 0\} \subset A \ , \tag{2.11}$$

with x_j coordinates of the ambient toric 4-fold, $c_{\vec{m}}$ complex coefficients, and $\vec{v}_j$ the (integer) vertices of Δ°. The weighted $\mathbb{CP}^4$ hypersurfaces are special cases of (2.11).

Thus the question of finding toric hypersurface CY_3 is the classification of reflexive integer 4-polytopes (up to $SL(4;\mathbb{Z})$, under which the toric 4-folds are equivalent). In $\mathbb{R}^1$, there is trivially 1 reflexive polytope (the pair of points ± 1). In $\mathbb{R}^2$, it is known at least to 19-th century mathematics, that there are 16 reflexive polygons up to $SL(2;\mathbb{Z})$. Unfortunately (and perhaps shockingly), the next number is already unknown until the work of Kreuzer-Skarke. They found 4319 reflexive polyhedra in $\mathbb{R}^3$. For $\mathbb{R}^4$, 6 months of computation on the best computer available to the late 1990s gave an astounding 473,800,776. Each of these gives [‖] a hypersurface Calabi-Yau 3-fold. Thus, our zoo of manifolds increased from 5, to some 10 thousand, and to some half-billion. Interestingly, the next number, that of reflexive polytopes in $\mathbb{R}^5$ up to $SL(5;\mathbb{Z})$, is unknown. It would be great to have a generating function for the sequence $1; 16; 4319; 473,800,776; \ldots$.

The KS dataset produced 30,108 distinct Hodge pairs and $\chi \in [-960, 960]$, with the extremal values of ± 960 being the weighted $\mathbb{CP}^4$ cases. No CY construction so far has ever produced an Euler number whose magnitude exceeds 960. A conjecture of Yau states that the topological type of (connected, smooth, compact) Calabi-Yau manifolds is *finite* in every dimension (we already see this in complex dimensions 1 and 2) and it could well be that 960 is the upper bound in dimension 3. There has been nice parallel directions of work in infinite families of Calabi-Yaus [52, 53] beyond topological type such as Gromow-Witten invariants, as well as in zooming in on special corners of small Hodge numbers [41–43].

[‖] It should be emphasized that most of these toric ambient spaces A (as with weighted $\mathbb{CP}^4$) are *not smooth*, and requires smoothing or resolution of singularities: different resolutions give rise to potentially different CY_3s. Thus, the actual number of Calabi-Yau 3-folds from this construction is estimated to be many orders of magnitude larger. For a given Δ, the Hodge pair will be the same, different resolutions will give different intersection numbers and Chern classes. Up to $h^{1,1} = 7$, this was done exhaustively in [49], while for the highest $h^{1,1} \sim 490$, this was done in [50]. The full list of CY_3s, after all the resolutions, has been recently estimated to be as large as 10^{10^5} [51].

Thus, by the turn of the century, there is an data-base of Calabi-Yau manifolds whose size is "big" even by today's standards. There is an internet meme, that "technically, Moses was the first person to download data from the cloud using a tablet" [54]. This amusing anachronism is a fitting analogy to how the age of "big data" in theoretical physics and algebraic geometry really goes back to the 1980s.

3 Data Explosion

Meanwhile, by the mid to late 1990s, in parallel to the heterotic programme outlined above, the discovery of D-branes [55], M-theory and G_2-compactification [56, 57], F-theory [58, 64], AdS/CFT [59], etc., as well as the wealth of dualities linking them begat the Second String Revolution. As with the First, this gave rise, and is still continuing to engender, a plethora of mathematical data, leading to various estimates of the "string landscape" [15,60], which was already anticipated in [61]. Numbers such as 10^{500} and, as aforementioned, today's 10^{10^5} began to enter the string and popular psyche.

In some sense, string theory has traded one difficult problem – the quantization of gravity – with another: the selection of the right vacuum. The latter is perhaps of more and certainly increasing interest to pure mathematicians, because the largess of data provides an inspiring playground for generating, testing and proving new conjectures.

Ultimately, whichever scenario one prefers to geometrically engineer (to use the phrase of [62]) one's preferred quantum field theory, including the standard model, the procedure can be algorithmized. Indeed, any problem in algebraic geometry (over $\mathbb{C}$) reduces to finding an appropriate Gröbner basis and then to finding (co-)kernels of integer matrices (in a corresponding monomial basis) [2].

Take, as an example, AdS/CFT from the point of view of computational geometry: this is a correspondence between a SUSY conformal QFT and a (non-compact) Calabi-Yau cone M over a Sasaki-Einstein manifold X. Moreover specifically, this is a mapping between a quiver representation and the geometric data of X (see e.g., [63] for a quick review). When M is toric, for instance, the graph data of the quiver and the combinatorial data of M are both amenable to an algorithmic treatment.

3.1 The Good, The Bad, and The ?

Taking stock of the progress up to the second decade of this century, we hope to have given the reader a glimpse of how computational and algorithmic geometry has enriched the classical dialogue between physics and mathematics. The ever increasing number of (freely available) mathematical databases online (typically of size $\sim 1-10$ Gb) is augmented by ever-more efficient software developed to address them (especially the umbrella project of SageMath [6]) as well as by the growing power of the personal laptop. This, certainly can be considered "the Good."

Unfortunately, most algorithms needed to compute anything, whether it be finding Gröbner bases, obtaining triangulations of polytopes, or extracting dual cones, are exponential in complexity. Thus, if one aims to sift through vacua to find the standard model or to understand the minimal model approach to algebraic varieties, case-by-case checks is impossible, even with the best HPC available. This, certainly needs to be rendered as "the Bad".

While the statistics of the vacuum degeneracy in string theory had been considered in the last decade [15], it is only expedient, given the breath-taking speed with which the Big Data Revolution is taking over every aspect of civilization, especially in this decade of the new millenium, that one should apply the most recent techniques to address the landscape of mathematical data.

In many ways, the search of the standard model within the string landscape reminds us of the hunt for exo-planets. The latter scans the heavens for habitable earths and the former, for universes akin to ours. The latter accumulates more and more real data with the betterment of technology and the former, theoretical data with furtherance of methodology. Whether one believes our universe is "special" by some anthropic argument, or by a selection principle, or is a mere point in the multiverse, is currently still a matter of debate, but the big data of mathematical universes beckon exploration.

4 Deep-Learning the Landscape

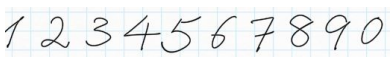
A question which instinctively occurred whilst contemplating the big data of universes [17, 18], was that the typical problem in string theory, or, in algebraic geometry for that matter, is of the form

$$\underset{\boxed{\text{integer tensor}}}{\overset{INPUT}{}} \longrightarrow \underset{\boxed{\text{integer}}}{\overset{OUTPUT}{}} \quad .$$

That one has integer output, especially in string theory, is because presently much of the field is still at the stage of finding quantities such as number generations, or the charges of particles. The fact that there is no known non-trivial (compact) Calabi-Yau metric analytically (Yau's proof of the Calabi Conjecture is famously non-constructive and relies on subtle existence statements of Monge-Ampère PDEs) hinders questions such as finding masses **. In geometry, much of the field is concerned with finding topological invariants such as indices or Betti numbers (as mentioned in introduction in Gauß' theorema egregium) because when complicated integrals become integers there is usually some deep mathematics going on. On the other hand, one has integer tensor input is seen in a multitude of examples above: whether we are dealing with polytopes or CICY configurations or quiver adjacency matrices. Of course, in general cases (such as numerical metrics), the integer condition case be relaxed to numerical tensor input going to some numerical output.

As discussed repeatedly, the machinery of computational geometry has developed sophisticated algorithms to obtain the output from the given input, even though the generic such algorithm is expensive. Nevertheless, physicists and mathematicians have bitten the bullet over the last 20 years or so and computed extensive examples. For instance, all Hodge numbers for the 1/2-billion Kreuzer-Skarke Calabi-Yau manifolds have been calculated using combinatorics of polytopes [65], likewise, all those for CICYs have been obtained by chasing exact sequences [46].

The situation is rather reminiscent of hand-writing recognition, the archetypal problem in machine-learning. For example, I write 0 to 9 as follows

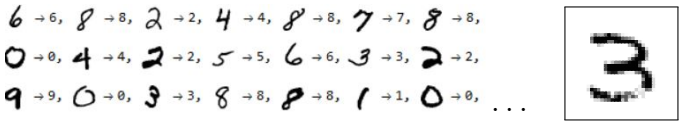


(4.12)

**We will later address some of the recent advances in numerical metrics and connections.

and we wish to let the computer recognize them. The input is an image, which is an $m \times n$ matrix (indexing the pixels in a 2-dimensional grid) each entry of which is a 3-vector of a real value between 0 and 1, denoting the percentage of RGB values. If we only wish to keep gray-scale information, each entry is then a real number between 0 and 1. Or, if we only want black-white, the input is just a binary matrix. The output is an integer from 0 to 9, called a *10-channel output*.

As mathematicians or theoretical physicists, we might solve this problem by exploiting the geometry and find, say, a clever Morse function as we scan the input matrix row-wise and column-wise and detect the critical points. This is, of course, very expensive. What Google or your smart-phone does, is to turn to *labeled data*. Such data has been painfully collected over the years by NIST (National institute of Standards), and look like the following (each is given as, for example, by a 28×28 pixelated image):



$$\ldots \qquad 28 \times 28 \times (RGB) \tag{4.13}$$

The difficult part of labeling each image with the correct channel has been done, and still adjusting with new usage by new users.

In summary, what happens is the following:

Data Acquisition: the collection of known cases (input $\to$ output), such as (4.13), gives us *training data*;

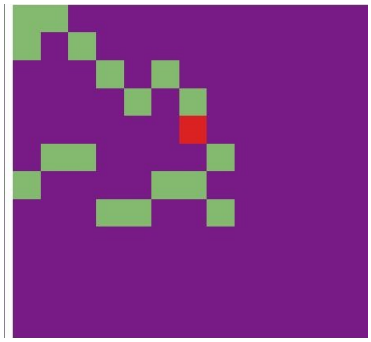
Machine-Learning: setting up some algorithm to optimize parameters which does the classification best;

Data Validation: once the machine has "learnt" the training data, we can take a set of *validation data*, which, importantly, the machine has *not* seen before. This is in the same format as the training data, with given input and output and we check the actual with the predicted outputs.

How different, really, is a problem in algebraic geometry? For example, computing the Hodge number of a given CICY, after all the work in long exact sequences in

cohomology, gives the association rule

$$X = \begin{pmatrix} 1&1&0&0&0&0&0&0 \\ 1&0&1&0&0&0&0&0 \\ 0&0&0&1&0&1&0&0 \\ 0&0&0&0&1&0&1&0 \\ 0&0&0&0&0&0&2&0 \\ 0&1&1&0&0&0&0&1 \\ 1&0&0&0&0&1&1&0 \\ 0&0&0&1&1&0&0&1 \end{pmatrix}, \quad h^{2,1}(X) = 22; \qquad \to 22\ . \tag{4.14}$$



To the right, we have purposefully represented the CICY configuration as a pixelated image, since all CICYs can be embedded, after right-bottom zero-padding, into a 12×15 integer matrix with entries $\in [0, 5]$. We have 7890 labeled data-points, from which we can take, say, 80% for training, to be validated on the remaining 20%. The programme of machine-learning algebraic geometry was thus initiated [17, 18].

It is timely, that in 2017 (the same year that Sophia, the AI robot, became the first non-human citizen of a country), 4 independent groups were thinking about various aspects of machine-learning the string landscape [17, 19–21]. Thinking back, let us see the sequence of the starting year of annual series of conferences in the string community: "Strings" (1986-), 'StringPheno" (2002-), "NSF String Vacuum Project" (2006 - 2010), "String-Math" (2011-), session of stringy mathematics and physics at "SIAM" (2014-), and now, "String-Data" (2017-).

4.1 An Invitation to Machine-Learning

We refer the reader to the now classic introduction to machine-learning in [66] as well as a wonderful new monograph for physicists in [67]. Here, it is expedient to give a rapid taster of this vast subject.

Contrary to expectations, the field of machine-learning and neural networks goes as far back as cybernetics in the 1940s. In 1957, the first *perceptron* was set up by MIT-Cornell, where a wall of CdS photo-receptors was set up to emulate neurons firing. In the 1980 - 90s, artificial neural networks went under the philosophy of connectivism where computational power emerged from inter-connectivity. Slowly the

word "artificial" disappeared and such algorithms were simply called *neural networks* (NNs). By 2006, the phrase "Deep" NN came into being, a term which we will explain shortly.

In general, sorting data into discrete categories is done by **classifiers** and predicting continuous values, **regressors**. Given data, machine-learning (ML) roughly fall under the headings of **unsupervised**, where patterns are to be extracted, and **supervised** where *labeled data*, such as the ones in (4.13) and (4.14), where the ML algorithms are trained to associate input to output. Examples of unsupervised ML include clustering analysis, auto-encoders, principle component analyses (PCA), etc., and those of supervised ML include support vector machines (SVM), neural regressors and neural classifiers, etc. In this talk, I will concentrate, because of the nature of the problem, on supervised ML.

Let us start with a single neuron (the perceptron), which consists of a (ususally analytic) function $f(z_i)$ called the *activation function*, for some input tensor z_i with multi-index i. We then consider $f(w_i z_i + b)$ with weights w_i and bias b. Typical activation functions include: (1) Logistic Sigmoid: $(1 + e^{-x})^{-1}$; (2) Hyperbolic tangent: $\tanh(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}}$; (3) Softplus: $\log(1 + e^x)$, a "softened" version of ReLu (Rectified Linear Unit): $\max(0, x)$; (4) Softmax: $x_i \to \frac{e^{x_i}}{\sum_i e^{x_i}}$; (5) Identity: $x_i \to x_i$ (which, with weights and biases, becomes the general affine transformation).

Given Training data: $\mathcal{D} = \{(x_i^{(j)}, d^{(j)}\}$ with input x_i and known output $d^{(j)}$, we minimize some appropriate **cost/loss function** to find optimal w_i and b (this is the "learning"). Then, with parameters fixed, we can check against Validation Data. Common cost functions include SEL (squared-error-loss)

$$SEL := \sum_j \left[f\left(\sum_i w_i x_i^{(j)} + b \right) - d^{(j)} \right]^2 \tag{4.15}$$

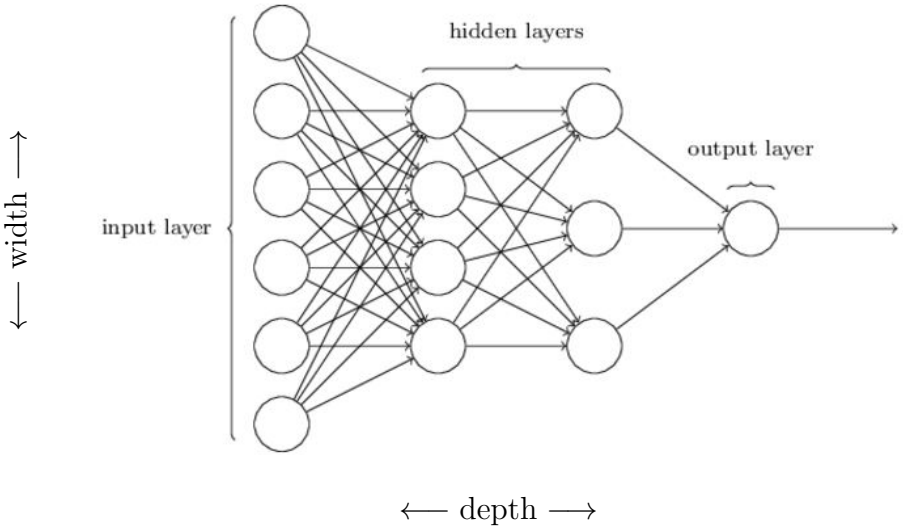
for continuous output, and XC (cross-entropy)

$$XE := -\frac{1}{n} \sum_j \left[d^{(j)} \log f(x^{(j)}) + (1 - d^{(j)}) \log(1 - f(x^{(j)})) \right] \tag{4.16}$$

for discrete (categoritcal) data.

The astute reader would recognize that we have done is precisely (non-linear)

regression. With a single neuron, supervised ML is exactly that. When we link up a multitude of neurons into a directed graph, complexity emerges through connectivity in a gestalt-philosophical way; this is the NN. A common type of NN is when the graph organizes into "layers" as in



(4.17)

which is called a *forward-feeding* NN (or, a more dated acronym, MLP, for multi-layer perceptron). The MLP is composed of an input layer, an output layer, and a number of hidden layers. The total number of layers is called the **depth** and the rough number of neurons per layer, the **width**. ML with large depth NNs is, for obvious reasons, called *deep learning*.

The precise choice of activation functions and inter-connectivity of the NN is called the *architecture*. The various parameters - not the variables like weights and biases to be optimized during training - such as depth, width, learning-rate (this the step-size for any gradient descent method using for finding minima), batch-size (the training data is usually passed in batches at a time), etc., are called *hyper-parameters*. As one can imagine, there is a variety of **universal approximation theorems** which essentially state that for sufficiently large width, or depth, any output can be approximated to arbitrary precision. In fact, a forward-feeding fully-connected NN with only ReLU activation is good enough to approximate any integrable function.

As with all models of statistical prediction, it is good to have a measure of "goodness of fit". Some standard ones are as follows

Naïve Precision: This is particularly useful when the output is discrete (and be-

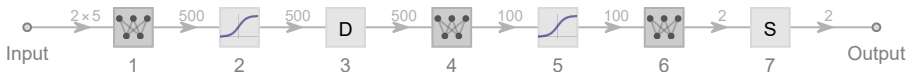
longing to a relatively small number of categories), and we simply compute the percentage of agreed cases between the predicted and actual.

R-squared: For continuous output, suppose on validation dataset $\mathcal{V} = \{x_i^{(j)} \longrightarrow d^{(j)}\}_{j=1,2,\ldots,m}$, the predicted values are $\{x_i^{(j)} \longrightarrow \hat{d}^{(j)}\}_j$. Then the Coefficient of Determination, or simply R-squared, is defined to be $R^2 := 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$, where data variance is $SS_{\text{tot}} := \sum_j (d^{(j)} - \overline{d^{(j)}})^2$ for mean $\overline{d^{(j)}}$, and residual sum of squares is $SS_{\text{res}} := \sum_j (d^{(j)} - \hat{d}^{(j)})^2$. A bad fit is when R^2 is close to 0, and a perfect fit, when $R^2 = 1$.

Confusion Matrix: For discrete output (say n categories), we can establish an $n \times n$ matrix with the (i, j)-th entry being the number of cases predicted to be j while the actual value is i. Ideally, we wish this to be a diagonal matrix. A measure of how close to the diagonal is the **Matthews' ϕ-coefficient** defined to be $\sqrt{\chi^2/n}$ where χ^2 is the Chi-square of the matrix treated as a contingency table. A value of $\phi = 0$ means the correlation is random and $\phi = 1$ is a perfect fit (incidentally, $\phi = -1$ mean complete anti-correlation); thus we can use ϕ as a measure of confidence (avoiding false positives and false negatives) in addition to the naïve precision.

4.2 Initial Experiments

Armed with the appropriate mathematical data and the technique of ML, implemented via either Python's Keras/TensorFlow [68] or Wolfram's Mathematica version > 11.0 [69]. One can take a simple MLP of the form (the hyper-parameters and architecture differ for specific cases and the following is only an illustration of a typical case)



(4.18)

where the hidden layers are (1) a fully-connected linear layer of 500 nodes; (2) element-wise sigmoid activation $\sigma(z) := (1 + e^{-z})^{-1}$; (3) dropout layer (we switch off neurons with some probability in order not to over-fit); (3) linear layer of 100 nodes; (4) sigmoid; (5) linear layer of 2 nodes; (6) Softmax to output. Such an NN was found to estimate the size of Hodge numbers for CICYs and WP4 hypersurfaces very well

in a matter of seconds on an ordinary laptop [17]. One could think of this setup, a fully connected neural networks of depth d, as the following composition of maps:

$$\mathbb{R}^{n_0} \xrightarrow{L_{n_1}} \mathbb{R}^{n_1} \xrightarrow{f} \mathbb{R}^{n_1} \xrightarrow{L_{n_2}} \cdots \xrightarrow{L_{n_d}} \mathbb{R}^{n_d} \to \mathbb{R} \ , \tag{4.19}$$

where L_n are activation functions (such as sigmoids) with trainable weights and biases and co-domain dimension n and the last layer outputs some real value or discrete value. The power of the MLP is the harnessing of the ultimately complicated (not even necessarily analytic) structure of the composite map.

A detailed analysis was carried out in [70] where the 19-way classifier/regressor in an architecture similar to (4.18) as well as an SVM, performed to about 90% accuracy in an 80-20% training-validation split [††]. It is interesting that in these experiments, one never exploited the *matrix* structure of the input: e.g., the CICY configuration was flattened a long vector of integers. This is quite contrary to the image processing of (4.13) where a convolutional network (CNN) would be used which "convolves" with nearest neighbours. Such CNNs were indeed tried more recently and $> 99\%$ accuracies were reached [73].

4.3 More Success Stories in String/Geometry

One can imagine that all computational problems in string phenomenology and more generally in computational algebraic geometry could benefit from the paradigm of machine-learning. Indeed, the initial explorations of [19] on Calabi-Yau volumes, of [20] on line-bundle cohomology, and of [21] on F-theory compactifications, in conjunctions with [17], launched the String Data conference series from 2017 on-wards.

Though it is difficult to review all the works since, I will give a bird's-eye-view of the the various directions taken, first within string/geometry, and then more generally to other branches of mathematics. In the former, some major directions and success stories (with 0.90 accuracies with relatively simple architectures) have included

Heterotic: Selection of MSSM from heterotic orbifold constructions [92,93], distinguishing standard models from heterotic line bundles [96,99]. Machine-learning

[††]We remark that for CICY 4-folds, where the data is about a million, accuracies to around 96% was achieved [71].

of bundle cohomology of surfaces [95] as well as toric hypersurfaces [97]. One points out [98] where exact formulae were found for line-bundle cohomology through an MLP exploration of the regions in moduli space.

F/M-Theory: Finding gauge groups [88] and matter-content [89] within F-theory compactifications. Distinguishing elliptically fibered manifolds within the CICYs [91]. Decidability issue of diophantine systems in Kähler stabilization [90].

Type II: topological data analysis [84] of, and genetic algorithms for searching within [85] flux vacua in type II. Reinforcement learning explorations of IIA brane configurations [86] and IIB landscape [87]. Seiberg duality in type IIB quiver theories [112].

Physical Symmetries: symmetries in various physical systems (including representations of CICYs) [105], and CFT symmetries [106].

Metric: As mentioned several times, there is no known analytic Calabi-Yau metric on a non-trivial compact Kähler manifold. Donaldson developed an efficient numerical algorithm using the method of balanced metrics from a potential formed by increasing powers of monomial sections [100], which were then nicely implemented in [101, 102] (q.v.. also the functional method of [103]). It was shown in [104] that Donaldson's algorithm can be machine-learnt (and to 10-100-fold increase in efficiency).

Cosmology: The cosmic landscape [80], especially vacuum selection from cosmology constraints [81], were studied. Interesting network structures were found in [81] and [83] studied certain accessibility measures in inflation and [82], machine learning in inflation.

"Meta" Physics: A fun experiment was undertaken in [107] where all titles from hep-th and four related sections of the arXiv: hep-ph, hep-lat, gr-qc, and math-ph were downloaded since the beginning and fed into the NN *Word2Vec* (about 10^6 titles). Interesting linear syntactical identities such as "holography + quantum + string + ads = extremal-black-hole" presented themselves and the syntactical structure of the different sections were indeed found to be distinct.

Of particular note are the striking ideas in [74–77] where the fundamentals of quantum field theory, holography and renormalization group flow, are phrased in

terms of appropriate neural networks. Indeed, the reader is also referred to the recent works of [78, 79] on the possible computational nature of reality itself.

5 Outlook: ML Mathematical Structures

Given the efficacy of ML in so many directions in string/geometry, it is natural to ask whether and how different problems in mathematics respond to ML. We leave a detailed discussion of this to [23], but for now, it is perhaps fitting that we conclude this talk with some conducive experiments which have been performed as well as some speculations for the future. Let us approximately group the successful experiments by subject:

Algebraic Geometry over $\mathbb{C}$: Most of the problems mentioned above fall under this heading. We need to emphasize that we work over $\mathbb{C}$, an algebraically closed number field. Any problem in computational algebraic geometry essentially boils down to finding kernels and co-kernels of integer matrices (in appropriate monomial bases), something quite adaptable to ML. A recent work on using reinforcement learning to perform the key step of finding S-pairs in constructing Gröbner bases was done in [118].

Representation Theory: Preliminary investigation on whether SVMs and MLPs can distinguish finite groups and finite rings from random matrix structures was initiated in [108]; more surprising was the fact that *simple groups* seemed to be distinguishable. For continuous groups, lengths of branching rules and tensor decomposition in Lie algebras can also be learned by looking at weight vectors [109]; this is obviously also of importance to particle physics.

Knot Theory: Jones polynomials and complementary volume of knots are studied from ML in [110] and letting ML find configurations of knots themselves, in [111].

Graph Theory and Combinatorics: Cluster mutation on quivers was studied in [112]. On a more basic level, properties of finite simple graphs, such as whether it possesses Euler or Hamilton cycles, whether it is flat (there is a notion of Ricci-flatness for finite graphs), etc. were studied in [113].

Number Theory: As one might imagine, a direct attack on predicting the next prime number by ML is most likely unfruitful [17, 22]. Likewise, predicting quantities relevant to the Birch-Swinnerton-Dyer Conjecture was also difficult [114]. Surprisingly, however, problems in arithmetic geometry, ranging from dessins d'enfant [115] (> 0.9 accuracy), to arithmetic properties of hyper-elliptic curves [116] ($\sim 0.99 - 1.00$ accuracies) and Galois number field extensions of the rationals [117] (> 0.9 accuracy) behaved very well to simple classifiers such as Naïve Bayes.

Symbolic Manipulation: Recent advances in generating new identities in calculus [119] and continued fractions [120] have met with impressive success. So too, have there been tools to extract fundamental laws [121] and formulae [122] of physics.

With these tantalizing thoughts let us conclude my talk here. We have seen how into the alembic of mathematics and fundamental physics is now infused, over the last few years, new techniques of the data revolution, especially the predictive power of machine-learning and neural networks. We are, of course, only at the early stage. Having an ML predict a result, even to 100% accuracy, does not always mean one could obtain analytic information as to why. What we hope for, is what the physicist Max Tegmark calls "intelligible intelligence", where we can formulate new results, or at least conjecture precise statements, when we are given an ML algorithm which performs superbly well. When I shared my initial excitement back in 2017, of the prospects to machine-learn problems ranging from geometry to algebra, to my friend the logician Boris Zilber, he astutely remarked: "now you have syntax, it would be good to find the semantics."

Acknowledgments

We are grateful for the kind invitations, in person and over Zoom, of the various institutions over this most extraordinary year of 2020 – the hospitality and conversations before the lock-down and the opportunity for a glimpse of the outside world during: Harvard University, Tsinghua University/BIMSA, Universidad Católica del Norte Chile, London Institute of Mathematical Sciences, Queen's Belfast, King's College London, University of Connecticut, "Clifford Algebra & Applications 2020" at

UST China, "String Maths 2020" at Capetown, "International Congress Mathematical Software 2020" at Braunschweig, University of Torino, "SageMath/M2 - an Open Source Initiative" at the University of Minnesota, "East Asia Strings" at Taipei-Seoul-Tokyo, Nankai University, Imperial College London, and Nottingham University. The work, as always, is indebted to STFC UK for grant ST/J00037X/1 and Merton College, Oxford for a quiet corner of paradise.

References

[1] C. N. Yang, M. L. Ge and Y. H. He, Ed. "Topology and Physics," with contributions from Atiyah, Penrose, Witten, et al., WS 2019. ISBN: 978-981-3278-49-3 https://doi.org/10.1142/11217

[2] D. Grayson, M. Stillman, "Macaulay2, a software system for research in algebraic geometry", Available at https://faculty.math.illinois.edu/Macaulay2/

[3] W. Decker, G-M. Greuel, G. Pfister, H. Schönemann, SINGULAR, A computer algebra system for polynomial computations. http://www.singular.uni-kl.de

[4] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.9.2*; 2018, https://www.gap-system.org

[5] Magma Comp. Algebra System, http://magma.maths.usyd.edu.au/

[6] SageMath, "the Sage Mathematics Software System", The Sage Developers, http://www.sagemath.org

[7] The Graded Ring Database, http://www.grdb.co.uk/
The C^3NG collaboration: http://geometry.ma.ic.ac.uk/3CinG/index.php/team-members-and-collaborators/ Data at: http://geometry.ma.ic.ac.uk/3CinG/index.php/data/ http://coates.ma.ic.ac.uk/fanosearch/

[8] The Knots Atlas, http://katlas.org/wiki/Main_Page

[9] The L-functions & Modular Forms Database, http://www.lmfdb.org/

[10] International Congress on Math. Software, http://icms-conference.org/

[11] Some videos of this talk can be found at Oxford ML&Physics seminar, https://www.youtube.com/watch?v=nMP2f14gYzc
StringMaths 2020, https://www.youtube.com/watch?v=GqoqxFsaogY

[12] Y-H. He, P. Dechant, A. Kaspryzyk, A. Lukas, Ed. "Machine-learning mathematical structures," topical collection for Advances in Applied Clifford Algebras, Birkhäuser, Springer, call open: https://www.springer.com/journal/6/updates/18581430

[13] M. Kreuzer and H. Skarke, "Complete classification of reflexive polyhedra in four-dimensions," Adv. Theor. Math. Phys. **4**, 1209 (2002) [hep-th/0002240].

[14] F. Gmeiner, R. Blumenhagen, G. Honecker, D. Lust and T. Weigand, "One in a billion:

MSSM-like D-brane statistics," JHEP **0601**, 004 (2006) [hep-th/0510170].

[15] D. Lust, "The landscape of string theory (orientifolds and their statistics, D-brane instantons, AdS(4) domain walls and black holes)," Fortsch. Phys. **56**, 694 (2008).
F. Denef and M. R. Douglas, "Computational complexity of the landscape. I.," Annals Phys. **322** (2007), 1096-1142 [arXiv:hep-th/0602072 [hep-th]].

[16] L. B. Anderson, Y. H. He and A. Lukas, "Heterotic Compactification, An Algorithmic Approach," JHEP **0707**, 049 (2007) [hep-th/0702210].

[17] Y. H. He, "Deep-Learning the Landscape," arXiv:1706.02714 [hep-th]. *Science*, vol 365, issue 6452, Aug 2019.

[18] Y. H. He, "Machine-learning the string landscape," Phys. Lett. B **774**, 564 (2017).

[19] 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]].

[20] F. Ruehle, "Evolving neural networks with genetic algorithms to study the String Landscape," JHEP **08** (2017), 038 [arXiv:1706.07024 [hep-th]].

[21] J. Carifio, J. Halverson, D. Krioukov and B. D. Nelson, "Machine Learning in the String Landscape," JHEP **1709**, 157 (2017) [arXiv:1707.00655 [hep-th]].

[22] Y. H. He, "The Calabi-Yau Landscape: from Geometry, to Physics, to Machine-Learning," [arXiv:1812.02893 [hep-th]]. To appear, Springer.

[23] Y. H. He, "Machine-Learning Mathematical Structures," to appear.

[24] P. Candelas, "Lectures On Complex Manifolds," in *Trieste 1987, proceedings, superstrings '87*, pp1-88.

[25] M. B. Green and J. H. Schwarz, "Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory," Phys. Lett. B **149** (1984), 117-122

[26] D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, "The Heterotic String," Phys. Rev. Lett. **54**, 502 (1985).

[27] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, "Vacuum Configurations for Superstrings," Nucl. Phys. B **258**, 46 (1985).

[28] T. Kaluza, "Zum Unitätsproblem in der Physik". Sitzungsber. Preuss. Akad. Wiss. Berlin. (Math. Phys.): 966 - 972 (1921)
O. Klein, "Quantentheorie und fünfdimensionale Relativitätstheorie". Z. für Physik A. 37 (12): 895 - 906 (1926); "The Atomicity of Electricity as a Quantum Theory Law". Nature. 118 (2971): 516. (1926)

[29] S. Coleman, J. Mandula, "All Possible Symmetries of the S Matrix", Physical Review. 159 (5): 1251 (1967)

[30] R. Haag, M. Sohnius, J. Łopuszański, "All possible generators of supersymmetries of the S-matrix", Nuclear Physics B, 88: 257–274 (1975).

[31] C. Hull, "Superstring compactifications with torsion and space-time supersymmetry," in Turin 1985 Proceedings, *Superunification and Extra Dimensions* (1986), 347 (375).
A. Strominger, "Superstrings with Torsion," Nucl. Phys. B **274**, 253 (1986).

[32] Robin Hartshorne, "Algebraic Geometry", GTM, Springer 1997, ISBN 13: 9780387902449.

[33] Y. H. He, "Calabi-Yau Spaces in the String Landscape," [arXiv:2006.16623 [hep-th]], entry to the Oxford Research Encyclopaedia of Physics, OUP, 2020.

[34] 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).

[35] 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).

[36] Green, M. B.; Schwarz, J. H.; Witten, E., Superstring Theory. Vol. 1&2, Cambridge Monographs on Mathematical Physics, 1987.
J. Polchinski, "String Theory", Vol. 1& 2. CUP, 1998
Barton Zwiebach, *A First Course in String Theory,* CUP, 2004, ISBN-9780511841682
Elias Kiritsis, *String Theory in a Nutshell,* 2nd Ed., PUP 2019.
Katrin Becker, Melanie Becker, John H. Schwarz, *String Theory and M-Theory: A Modern Introduction*, CUP 2007.
M. Dine, *Supersymmetry and String Theory: Beyond the Standard Model*, CUP 2007.
Luis E. Ibáñez, Angel M. Uranga, *String Theory and Particle Physics: An Introduction to String Phenomenology*, CUP, 2012

[37] 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]].

[38] V. Bouchard and R. Donagi, "An SU(5) heterotic standard model," Phys. Lett. B **633** (2006), 783-791 [arXiv:hep-th/0512149 [hep-th]].

[39] L. B. Anderson, J. Gray, A. Lukas and E. Palti, "Heterotic Line Bundle Standard Models," JHEP **06** (2012), 113 [arXiv:1202.1757 [hep-th]].

[40] 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]].

[41] 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]] (New Scientist, Jan, 5, 2008 feature).

[42] P. Candelas, A. Constantin and C. Mishra, "Calabi-Yau Threefolds with Small Hodge Numbers," Fortsch. Phys. **66** (2018) no.6, 1800029 [arXiv:1602.06303 [hep-th]].
P. Candelas and R. Davies, "New Calabi-Yau Manifolds with Small Hodge Numbers," Fortsch. Phys. **58** (2010), 383-466 [arXiv:0809.4681 [hep-th]].

[43] H. Schenck, M. Stillman, B. Yuan "Calabi-Yau threefolds in pn and gorenstein rings",

arXiv:2011.10871

[44] M. Gross, D. Huybrechts, D. Joyce *Calabi-Yau manifolds and related geometries.* Springer 2012.

[45] P. Candelas, A. M. Dale, C. A. Lutken, R. Schimmrigk, "Complete Intersection Calabi-Yau Manifolds," Nucl. Phys. B **298**, 493 (1988).
P. Candelas, C. A. Lutken, R. Schimmrigk, "Complete Intersection Calabi-Yau Manifolds. 2. Three Generation Manifolds," Nucl. Phys. B **306**, 113 (1988).
M. Gagnon, Q. Ho-Kim, "An Exhaustive list of complete intersection Calabi-Yau manifolds," Mod. Phys. Lett. A **9** (1994) 2235.

[46] T. Hubsch, *Calabi-Yau manifolds: A Bestiary for physicists*, World Scientific, 1994, ISBN 9810206623

[47] P. Candelas, M. Lynker and R. Schimmrigk, "Calabi-Yau Manifolds in Weighted P(4)," Nucl. Phys. B **341**, 383 (1990).

[48] V.V.Batyrev and L.A.Borisov, "On Calabi-Yau complete intersections in toric varieties", alg-geom/9412017.
V. V. Batyrev, "Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties", J. Alg. Geom. 3 (1994) 493 - 545, [alg-geom/9310003].

[49] R. Altman, J. Gray, Y. H. He, V. Jejjala and B. D. Nelson, "A Calabi-Yau Database: Threefolds Constructed from the Kreuzer-Skarke List," JHEP **1502**, 158 (2015) [arXiv:1411.1418 [hep-th]].

[50] A. P. Braun, C. Long, L. McAllister, M. Stillman and B. Sung, "The Hodge Numbers of Divisors of Calabi-Yau Threefold Hypersurfaces," arXiv:1712.04946 [hep-th].

[51] R. Altman, J. Carifio, J. Halverson and B. D. Nelson, "Estimating Calabi-Yau Hypersurface and Triangulation Counts with Equation Learners," JHEP **1903**, 186 (2019) [arXiv:1811.06490 [hep-th]].

[52] P. Berglund and T. Hübsch, "On Calabi–Yau generalized complete intersections from Hirzebruch varieties and novel $K3$-fibrations," Adv. Theor. Math. Phys. **22** (2018), 261-303 [arXiv:1606.07420 [hep-th]].
A. Garbagnati and B. van Geemen, "A remark on generalized complete intersections," Nucl. Phys. B **925** (2017), 135-143 [arXiv:1708.00517 [math.AG]].
P. Berglund and T. Hubsch, "A Generalized Construction of Calabi-Yau Models and Mirror Symmetry," SciPost Phys. **4** (2018) no.2, 009 [arXiv:1611.10300 [hep-th]].

[53] E. Ballico, E. Gasparim, B. Suzuki, "Infinite dimensional families of Calabi–Yau threefolds and moduli of vector bundles," J. Pure and Applied Algebra, Vol 225, 4, 2021.

[54] Meme: `https://jr.co.il/humor/pass144.htm`

[55] J. Polchinski, "Dirichlet Branes and Ramond-Ramond charges," Phys. Rev. Lett. **75**, 4724 (1995) [hep-th/9510017].

[56] P. Horava and E. Witten, "Eleven-dimensional supergravity on a manifold with boundary," Nucl. Phys. B **475** (1996), 94-114 [arXiv:hep-th/9603142 [hep-th]].

[57] M. Atiyah and E. Witten, "M theory dynamics on a manifold of G(2) holonomy," Adv. Theor. Math. Phys. **6**, 1 (2003) [hep-th/0107177].

[58] C. Vafa, "Evidence for F theory," Nucl. Phys. B **469** (1996), 403-418 [arXiv:hep-th/9602022 [hep-th]].

[59] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Int. J. Theor. Phys. **38**, 1113 (1999) [Adv. Theor. Math. Phys. **2**, 231 (1998)] [hep-th/9711200].

[60] S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, "De Sitter vacua in string theory," Phys. Rev. D **68**, 046005 (2003) [hep-th/0301240].

[61] W. Lerche, D. Lust and A. N. Schellekens, "Ten-dimensional Heterotic Strings From Niemeier Lattices," Phys. Lett. B 181, 71 (1986).
A. N. Schellekens, "The Landscape 'avant la lettre'," physics/0604134.

[62] S. H. Katz, A. Klemm and C. Vafa, "Geometric engineering of quantum field theories," Nucl. Phys. B **497** (1997), 173-195 [arXiv:hep-th/9609239 [hep-th]].

[63] Y. H. He, "Calabi-Yau Varieties: from Quiver Representations to Dessins d'Enfants," [arXiv:1611.09398 [math.AG]], in *Proc. Grothendieck-Teichmuller Theories and Impact*, 2016.

[64] S. Katz, D. R. Morrison, S. Schafer-Nameki and J. Sully, "Tate's algorithm and F-theory," JHEP **08** (2011), 094 [arXiv:1106.3854 [hep-th]].
Y. Kimura, "Nongeometric heterotic strings and dual F-theory with enhanced gauge groups," JHEP **02** (2019), 036 [arXiv:1810.07657 [hep-th]].
T. Weigand, "TASI Lectures on F-theory," arXiv:1806.01854 [hep-th].

[65] M. Kreuzer and H. Skarke, "PALP: A Package for analyzing lattice polytopes with applications to toric geometry," Comput. Phys. Commun. **157** (2004) 87 [arXiv:0204356 [math.NA]]. A. P. Braun and N. O. Walliser, "A New offspring of PALP," arXiv:1106.4529 [math.AG]. A. P. Braun, J. Knapp, E. Scheidegger, H. Skarke and N. O. Walliser, "PALP - a User Manual," arXiv:1205.4147 [math.AG].

[66] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, ISBN: 9780262035613, MIT Press, 2016.

[67] F. Ruehle, "Data science applications to string theory," Phys. Rept. **839** (2020), 1-117

[68] Python Software Foundation, `http://www.python.org`
G. van Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.

[69] Wolfram Research, Inc., Mathematica, Champaign, IL (2018). `www.wolfram.com`

[70] 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]].
–, "Getting CICY High," PLB **795** (2019), 700-706 [arXiv:1903.03113 [hep-th]].

[71] Y. H. He and A. Lukas, "Machine Learning Calabi-Yau Four-folds," [arXiv:2009.02544 [hep-th]].

[72] H. Erbin and R. Finotello, "Inception Neural Network for Complete Intersection Calabi-Yau 3-folds," [arXiv:2007.13379 [hep-th]].

[73] H. Erbin and R. Finotello, "Machine learning for complete intersection Calabi-Yau manifolds: a methodological study," [arXiv:2007.15706 [hep-th]].
–, "Machine learning for complete intersection Calabi-Yau manifolds: a methodological study," [arXiv:2007.15706 [hep-th]].

[74] K. Hashimoto, S. Sugishita, A. Tanaka and A. Tomiya, "Deep learning and the AdS/CFT correspondence," Phys. Rev. D **98** (2018) no.4, 046019 doi:10.1103/PhysRevD.98.046019 [arXiv:1802.08313 [hep-th]].

[75] K. Hashimoto, "AdS/CFT correspondence as a deep Boltzmann machine," Phys. Rev. D **99** (2019) no.10, 106017 [arXiv:1903.04951 [hep-th]].

[76] E. d. Koch, R. de Mello Koch and L. Cheng, "Is Deep Learning a Renormalization Group Flow?," [arXiv:1906.05212 [cs.LG]].

[77] J. Halverson, A. Maiti and K. Stoner, "Neural Networks and Quantum Field Theory," [arXiv:2008.08601 [cs.LG]].

[78] V. Vanchurin, "The world as a neural network." Entropy 22.11 (2020): 1210. arXiv:2008.01540.

[79] S. Wolfram, `https://www.wolframphysics.org/`

[80] J. Liu, "Artificial Neural Network in Cosmic Landscape," JHEP **12** (2017), 149 [arXiv:1707.02800 [hep-th]].

[81] J. Carifio, W. J. Cunningham, J. Halverson, D. Krioukov, C. Long and B. D. Nelson, "Vacuum Selection from Cosmology on Networks of String Geometries," Phys. Rev. Lett. **121** (2018) no.10, 101602 [arXiv:1711.06685 [hep-th]].

[82] T. Rudelius, "Learning to Inflate," JCAP **02** (2019), 044 [arXiv:1810.05159 [hep-th]].

[83] J. Khoury, "Accessibility Measure for Eternal Inflation: Dynamical Criticality and Higgs Metastability," [arXiv:1912.06706 [hep-th]].

[84] A. Cole and G. Shiu, "Topological Data Analysis for the String Landscape," JHEP **03** (2019), 054 [arXiv:1812.06960 [hep-th]].

[85] A. Cole, A. Schachner and G. Shiu, "Searching the Landscape of Flux Vacua with Genetic Algorithms," JHEP **11** (2019), 045 [arXiv:1907.10072 [hep-th]].

[86] J. Halverson, B. Nelson and F. Ruehle, "Branes with Brains: Exploring String Vacua with Deep Reinforcement Learning," JHEP **06** (2019), 003 [arXiv:1903.11616 [hep-th]].

[87] V. M. Mehta, M. Demirtas, C. Long, D. J. E. Marsh, L. McAllister and M. J. Stott, "Superradiance Exclusions in the Landscape of Type IIB String Theory," [arXiv:2011.08693 [hep-th]].

[88] Y. N. Wang and Z. Zhang, "Learning non-Higgsable gauge groups in 4D F-theory," JHEP **08** (2018), 009 [arXiv:1804.07296 [hep-th]].

[89] M. Bies, M. Cvetič, R. Donagi, L. Lin, M. Liu and F. Ruehle, "Machine Learning and Algebraic Approaches towards Complete Matter Spectra in 4d F-theory,"

[arXiv:2007.00009 [hep-th]].

[90] J. Halverson, M. Plesser, F. Ruehle, J. Tian, "Kähler Moduli Stabilization & Propagation of Decidability," PRD **101** (2020) no.4, 046010 [arXiv:1911.07835 [hep-th]].

[91] Y. H. He and S. J. Lee, "Distinguishing elliptic fibrations with AI," Phys. Lett. B **798** (2019), 134889 [arXiv:1904.08530 [hep-th]].

[92] E. Parr and P. K. S. Vaudrevange, "Contrast data mining for the MSSM from strings," Nucl. Phys. B **952** (2020), 114922 [arXiv:1910.13473 [hep-th]].

[93] E. Parr, P. K. S. Vaudrevange and M. Wimmer, "Predicting the orbifold origin of the MSSM," Fortsch. Phys. **68** (2020) no.5, 2000032 [arXiv:2003.01732 [hep-th]].

[94] M. Larfors and R. Schneider, "Explore and Exploit with Heterotic Line Bundle Models," Fortsch. Phys. **68** (2020) no.5, 2000034 [arXiv:2003.04817 [hep-th]].

[95] C. R. Brodie, A. Constantin, R. Deen and A. Lukas, "Machine Learning Line Bundle Cohomology," Fortsch. Phys. **68** (2020) no.1, 1900087 [arXiv:1906.08730 [hep-th]].

[96] H. Otsuka and K. Takemoto, "Deep learning and k-means clustering in heterotic string vacua with line bundles," JHEP **05** (2020), 047 [arXiv:2003.11880 [hep-th]].

[97] D. Klaewer, L. Schlechter, "Machine Learning Line Bundle Cohomologies of Hypersurfaces in Toric Varieties," PLB **789** (2019), 438-443 [arXiv:1809.02547 [hep-th]].

[98] A. Constantin and A. Lukas, "Formulae for Line Bundle Cohomology on Calabi-Yau Threefolds," Fortsch. Phys. **67** (2019) no.12, 1900084 [arXiv:1808.09992 [hep-th]].

[99] R. Deen, Y. H. He, S. J. Lee and A. Lukas, "Machine Learning String Standard Models," [arXiv:2003.13339 [hep-th]].

[100] S. Donaldson, "Scalar curvature and projective embeddings, I", J. Differential Geom. 59 3, (11, 2001) 479 - 522; "Some numerical results in complex differential geometry", ArXiv:math/0512625.

[101] M. R. Douglas, R. L. Karp, S. Lukic and R. Reinbacher, "Numerical Calabi-Yau metrics," J. Math. Phys. **49** (2008), 032302 [arXiv:hep-th/0612075 [hep-th]].
–, "Numerical solution to the hermitian Yang-Mills equation on the Fermat quintic," JHEP **12** (2007), 083 [arXiv:hep-th/0606261 [hep-th]].

[102] V. Braun, T. Brelidze, M. R. Douglas and B. A. Ovrut, "Calabi-Yau Metrics for Quotients and Complete Intersections," JHEP **05** (2008), 080 [arXiv:0712.3563 [hep-th]].
– "Eigenvalues and Eigenfunctions of the Scalar Laplace Operator on Calabi-Yau Manifolds," JHEP **07** (2008), 120 [arXiv:0805.3689 [hep-th]].
L. B. Anderson, V. Braun, R. L. Karp and B. A. Ovrut, "Numerical Hermitian Yang-Mills Connections and Vector Bundle Stability in Heterotic Theories," JHEP **06** (2010), 107 [arXiv:1004.4399 [hep-th]].

[103] M. Headrick and T. Wiseman, "Numerical Ricci-flat metrics on K3," Class. Quant. Grav. **22** (2005), 4931-4960 [arXiv:hep-th/0506129 [hep-th]].
M. Headrick and A. Nassar, "Energy functionals for Calabi-Yau metrics," Adv. Theor. Math. Phys. **17** (2013) no.5, 867-902 [arXiv:0908.2635 [hep-th]].

[104] A. Ashmore, Y. H. He and B. A. Ovrut, "Machine learning Calabi-Yau metrics," Fortsch. Phys. **68** (2020) no.9, 2000068 [arXiv:1910.08605 [hep-th]].

[105] S. Krippendorf and M. Syvaeri, "Detecting Symmetries with Neural Networks," [arXiv:2003.13679 [physics.comp-ph]].

[106] H. Y. Chen, Y. H. He, S. Lal and M. Z. Zaz, "Machine Learning Etudes in Conformal Field Theories," [arXiv:2006.16114 [hep-th]].

[107] Y. H. He, V. Jejjala and B. D. Nelson, "hep-th," [arXiv:1807.00735 [cs.CL]].

[108] Y. H. He and M. Kim, "Learning Algebraic Structures: Preliminary Investigations," [arXiv:1905.02263 [cs.LG]].

[109] H. Y. Chen, Y. H. He, S. Lal and S. Majumder, "Machine Learning Lie Structures & Applications to Physics," [arXiv:2011.00871 [hep-th]].

[110] V. Jejjala, A. Kar and O. Parrikar, "Deep Learning the Hyperbolic Volume of a Knot," Phys. Lett. B **799** (2019), 135033 [arXiv:1902.05547 [hep-th]].

[111] S. Gukov, J. Halverson, F. Ruehle and P. Sułkowski, "Learning to Unknot," [arXiv:2010.16263 [math.GT]].

[112] J. Bao, S. Franco, Y. H. He, E. Hirst, G. Musiker and Y. Xiao, "Quiver Mutations, Seiberg Duality and Machine Learning," Phys. Rev. D **102** (2020) no.8, 086013 [arXiv:2006.10783 [hep-th]].

[113] Y. H. He and S. T. Yau, "Graph Laplacians, Riemannian Manifolds and their Machine-Learning," [arXiv:2006.16619 [math.CO]].

[114] L. Alessandretti, A. Baronchelli and Y. H. He, "Machine Learning meets Number Theory: The Data Science of Birch-Swinnerton-Dyer," [arXiv:1911.02008 [math.NT]].

[115] Y. H. He, E. Hirst and T. Peterken, "Machine-Learning Dessins d'Enfants: Explorations via Modular and Seiberg-Witten Curves," [arXiv:2004.05218 [hep-th]].

[116] Y. H. He, K. H. Lee and T. Oliver, "Machine-Learning the Sato–Tate Conjecture," [arXiv:2010.01213 [math.NT]].

[117] Y. H. He, K. H. Lee and T. Oliver, "Machine-Learning Number Fields," [arXiv:2011.08958 [math.NT]].

[118] Dylan Peifer, Michael Stillman, Daniel Halpern-Leistner, "Learning selection strategies in Buchberger's algorithm", [arXiv:2005.01917]

[119] G. Lample, F. Charton "Deep Learning for Symbolic Maths", arXiv:1912.01412 [cs.SC]

[120] G. Raayoni, S. Gottlieb, G. Pisha, Y. Harris, Y. Manor, U. Mendlovic, D. Haviv, Y. Hadad, I. Kaminer, "The Ramanujan Machine: Automatically Generated Conjectures on Fundamental Constants", arXiv:1907.00205 [cs.LG]

[121] R. Iten, T. Metger, H. Wilming, L. del Rio, R. Renner, "Discovering Physical Concepts with Neural Networks", Phys. Rev. Lett. 124, 010508, 2020

[122] S. M. Udrescu and M. Tegmark, "AI Feynman: a Physics-Inspired Method for Symbolic Regression," [arXiv:1905.11481 [physics.comp-ph]].