Using machine learning to formulate mathematical conjectures

Marc Lackenby

February 2023

The goal of today's talk

The aim is to show how machine learning can be used
to discover new connections in mathematics and
to formulate new mathematical conjectures.

The goal of today's talk

The aim is to show how machine learning can be used
to discover new connections in mathematics and
to formulate new mathematical conjectures.

Joint work with Alex Davies, András Juhász, Nenad Tomasev

Supervised learning

Supervised learning

- Suppose that there is some subset S of $\mathbb{R}^n$ and some function $f\colon S \to \mathbb{R}^k$ which we can compute but do not know 'explicitly'.

Supervised learning

- Suppose that there is some subset S of $\mathbb{R}^n$ and some function $f \colon S \to \mathbb{R}^k$ which we can compute but do not know 'explicitly'.
- We are given various data points $v \in S$, as well as their images $f(v)$.

Supervised learning

- Suppose that there is some subset S of $\mathbb{R}^n$ and some function $f\colon S \to \mathbb{R}^k$ which we can compute but do not know 'explicitly'.
- We are given various data points $v \in S$, as well as their images $f(v)$.
- Machine learning algorithms provide a function $F\colon \mathbb{R}^n \to \mathbb{R}^k$ that is an approximation to f, at least at the given points v.

Supervised learning

- Suppose that there is some subset S of $\mathbb{R}^n$ and some function $f\colon S \to \mathbb{R}^k$ which we can compute but do not know 'explicitly'.
- We are given various data points $v \in S$, as well as their images $f(v)$.
- Machine learning algorithms provide a function $F\colon \mathbb{R}^n \to \mathbb{R}^k$ that is an approximation to f, at least at the given points v.
- This is essentially 'non-linear regression'.

Supervised learning

- Suppose that there is some subset S of $\mathbb{R}^n$ and some function $f\colon S \to \mathbb{R}^k$ which we can compute but do not know 'explicitly'.
- We are given various data points $v \in S$, as well as their images $f(v)$.
- Machine learning algorithms provide a function $F\colon \mathbb{R}^n \to \mathbb{R}^k$ that is an approximation to f, at least at the given points v.
- This is essentially 'non-linear regression'.
- BUT unlike unlike linear regression we do not get an 'explicit' output function F, but merely the ability to compute $F(w)$ for other inputs $w \in \mathbb{R}^n$.

Example

$n =$ the number of pixels of an input picture

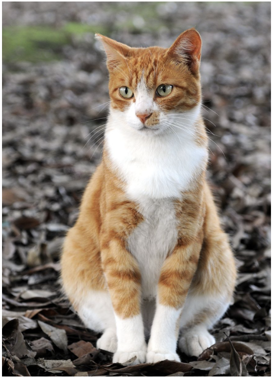
$v \in \mathbb{R}^n$ is the input picture (in grey-scale)

Example

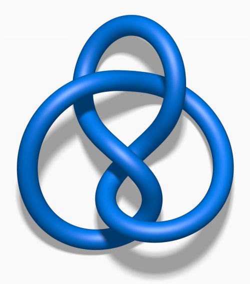
$n =$ the number of pixels of an input picture

$v \in \mathbb{R}^n$ is the input picture (in grey-scale)

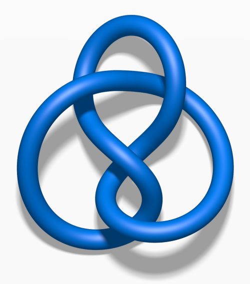
$$f(v) = \begin{cases} -1 & \text{if v is a picture of a cat} \\ 1 & \text{if v is a picture of a dog} \\ 0 & \textit{otherwise} \end{cases}$$



The branches of knot theory



The branches of knot theory



Knot theory is divided into three quite distinct subfields:

- hyperbolic knot theory
- gauge/Floer theory
- quantum topology

Invited speakers

- Ian Agol (U. C. Berkeley)
- Martin Bridson (Oxford U.)
- Jeff Brock (Yale U.)
- Ted Chinburg (U. Pennsylvania)
- Michelle Chu (U.I. Chicago)
- Jeff Danciger (U.T. Austin)
- Cameron Gordon (U.T. Austin)
- Ursula Hamenstadt (U. Bonn)
- Neil Hoffman (O.S.U.)
- Autumn Kent (U.W. Madison)
- Darren Long (U.C. Santa Barbara)
- Alex Lubotzky (Hebrew U.)
- Bruno Martelli (U. Pisa)
- Gaven Martin (Massey U.)
- Priyam Patel (U. Utah)
- Kate Petersen (U.M. Duluth)
- Jessica Purcell (Monash U.)
- Peter Sarnak (Princeton U.)
- Matt Stover (Temple U.)
- Sam Taylor (Temple U.)
- Genevieve Walsh (Tufts U.)
- Will Worden (Rice U.)

Alan Reid's conference

Topics:

- Interplay of 3-dimensional and 4-dimensional Topology
- Floer homology theories and associated invariants
- Khovanov homology
- Geometric and analytic aspects of gauge theoretic equations

Speakers:
D. GABAI, Princeton University, USA
L. GOETTSCHE, ICTP, Italy
L. GUTH, MIT, USA
J. HOM, Georgia Tech, USA
C. HUGELMEYER, Princeton University, USA
*P. KRONHEIMER, Harvard University, USA
F. LIN, Princeton University, USA
R. LIPSHITZ, University of Oregon, USA
P. LISCA, Università di Pisa, Italy
C. MANOLESCU, Stanford University, USA
G. MATIC, University of Georgia, USA
R. MAZZEO, Stanford University, USA
M. MILLER, Stanford University, USA
E. MURPHY, Princeton University, USA
*J. PARDON, Princeton University, USA
L. PICCIRILLO, MIT, USA
J. PINZON CAICEDO, University of Notre Dame, USA
J. RASMUSSEN, Cambridge, UK
D. RUBERMAN, Brandeis University, USA
A. STIPSICZ, Central European University, Hungary
Z. SZABO, Princeton University, USA
*C.TAUBES, Harvard University, USA
D. WANG SUNY, Stony Brook, USA
J. WANG, Harvard University, USA
C. ZIBROWIUS, University of Regensburg, Germany

Trieste conference

Finding connections between these fields

Finding connections between these fields

Each field has plenty of knot invariants:

Hyperbolic invariants:

- Volume
- Cusp shape and volume
- Length spectrum
- Trace field ...

3/4-dimensional invariants:

- Heegaard Floer homology
- Instanton Floer homology
- s, τ, ϵ, Υ, ...

Finding connections between these fields

Each field has plenty of knot invariants:

Hyperbolic invariants:

- Volume
- Cusp shape and volume
- Length spectrum
- Trace field ...

3/4-dimensional invariants:

- signature
- Heegaard Floer homology
- Instanton Floer homology
- s, τ, ϵ, Υ, ...

Finding connections between these fields

Each field has plenty of knot invariants:

Hyperbolic invariants:

- Volume
- Cusp shape and volume
- Length spectrum
- Trace field ...

3/4-dimensional invariants:

- signature
- Heegaard Floer homology
- Instanton Floer homology
- s, τ, ϵ, Υ, ...

Goal: Find new connections between these invariants

Knot signature

The 3/4-dimensional invariant that we focused on was the signature.

This is defined by starting with a Seifert surface S for the knot K.

The symmetrised Seifert form for S is the bilinear form

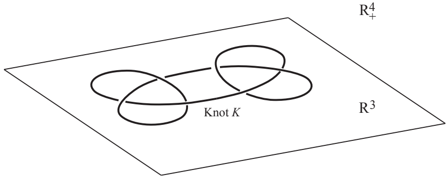
$$\begin{aligned} H_1(S) \times H_1(S) &\to \mathbb{Z} \\ (\ell_1, \ell_2) &\mapsto \mathrm{lk}(\ell_1, \ell_2^+) + \mathrm{lk}(\ell_2, \ell_1^+) \end{aligned}$$

where ℓ_2^+ is the push-off of ℓ_2 in the positive normal direction from S.

The signature $\sigma(K)$ is the signature of this bilinear form.

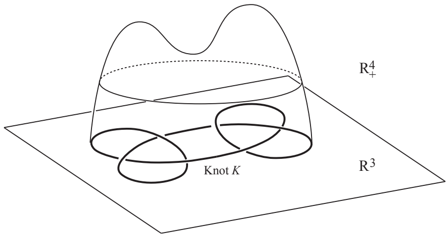
Connections with dimension 4

View $\mathbb{R}^3$ as the boundary of $\mathbb{R}^4_+ = \{(x_1, x_2, x_3, x_4) : x_4 \geq 0\}$.



Connections with dimension 4

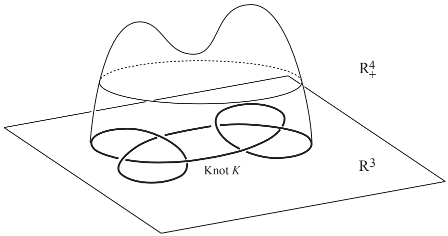
View $\mathbb{R}^3$ as the boundary of $\mathbb{R}^4_+ = \{(x_1, x_2, x_3, x_4) : x_4 \geq 0\}$.



The 4-ball genus of a knot K is the minimal genus of a (topological locally-flat) surface in $\mathbb{R}^4_+$ with boundary equal to K.

Connections with dimension 4

View $\mathbb{R}^3$ as the boundary of $\mathbb{R}^4_+ = \{(x_1, x_2, x_3, x_4) : x_4 \geq 0\}$.



The 4-ball genus of a knot K is the minimal genus of a (topological locally-flat) surface in $\mathbb{R}^4_+$ with boundary equal to K.

Theorem: [Murasugi 1965] $g_4(K) \geq |\sigma(K)|/2$.

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.
- This was the Regina census of 1,700,000 knots with ≤ 16 crossings plus $1,000,000$ randomly chosen knots with ≤ 80 crossings.

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.
- This was the Regina census of 1,700,000 knots with ≤ 16 crossings plus $1,000,000$ randomly chosen knots with ≤ 80 crossings.
- We randomly divided them into two groups: a training set and a test set.

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f \colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.
- This was the Regina census of 1,700,000 knots with ≤ 16 crossings plus $1,000,000$ randomly chosen knots with ≤ 80 crossings.
- We randomly divided them into two groups: a training set and a test set.
- We trained a neural network to predict the signature from the hyperbolic invariants.

Using machine learning in knot theory

Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.
- This was the Regina census of 1,700,000 knots with ≤ 16 crossings plus $1,000,000$ randomly chosen knots with ≤ 80 crossings.
- We randomly divided them into two groups: a training set and a test set.
- We trained a neural network to predict the signature from the hyperbolic invariants.
- We then tested this network using the test set.

Using machine learning in knot theory

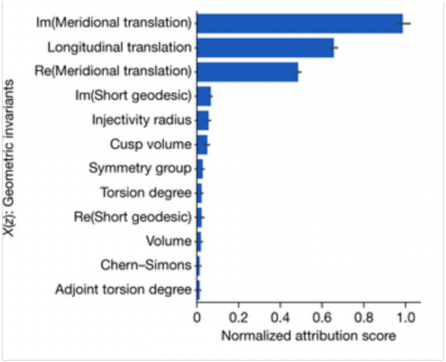
Goal: can we predict the signature from hyperbolic invariants?

In other words, is there a function $f\colon \{\text{hyperbolic knot invariants}\}(\subseteq \mathbb{R}^n) \to \mathbb{R}$ that outputs a knot's signature (or at least a good approximation to it)?

- Using snappy, we created a sample set of 2,700,000 hyperbolic knots.
- This was the Regina census of 1,700,000 knots with ≤ 16 crossings plus $1,000,000$ randomly chosen knots with ≤ 80 crossings.
- We randomly divided them into two groups: a training set and a test set.
- We trained a neural network to predict the signature from the hyperbolic invariants.
- We then tested this network using the test set.
- The network could predict the signature with impressive accuracy.

Saliency

The main hyperbolic invariants that were used to predict signature:

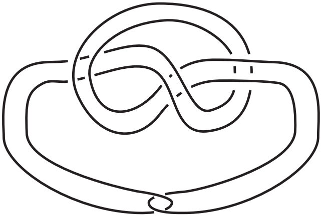


Hyperbolic structures

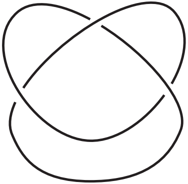
A hyperbolic structure on a knot complement is a complete finite-volume Riemannian metric of constant curvature -1.

By Mostow rigidity, if such a metric exists, it is a unique up to isometry.

Thurston's theorem: The complement of a non-trivial knot K has a hyperbolic structure if and only if K is not a torus knot or a satellite knot.



satellite knot



torus knot

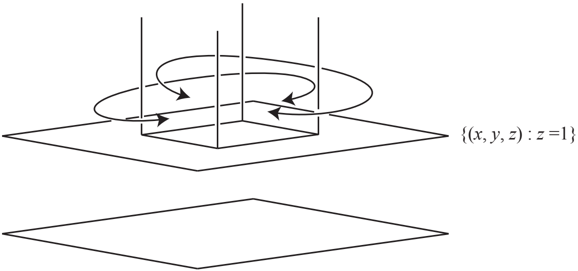
Cusp geometry

Any knot complement has an end of the form $T^2 \times [1, \infty)$.

When the knot is hyperbolic, this has a canonical geometry and is called a cusp.

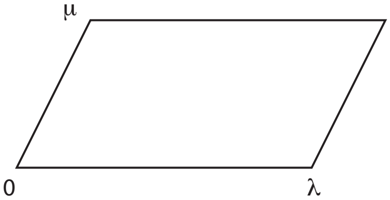
Let $\mathbb{H}^3$ be upper-half space $\{(x, y, z) : z > 0\}$. Let H be the horoball $\{z \geq 1\}$.

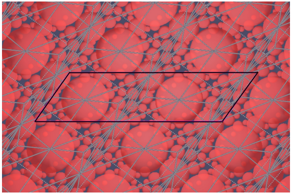
Then the cusp is formed $H/\langle\text{group of Euclidean translations}\rangle$.



The cusp boundary

The boundary of the cusp is a Euclidean torus $\mathbb{C}/\Lambda$ for a lattice Λ. We normalise Λ so that the longitude λ is real and positive, and the meridian μ has positive imaginary part.

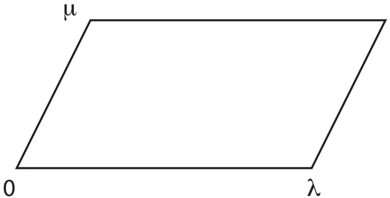


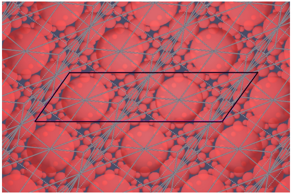


Cusp torus for 6_1

The cusp boundary

The boundary of the cusp is a Euclidean torus $\mathbb{C}/\Lambda$ for a lattice Λ. We normalise Λ so that the longitude λ is real and positive, and the meridian μ has positive imaginary part.

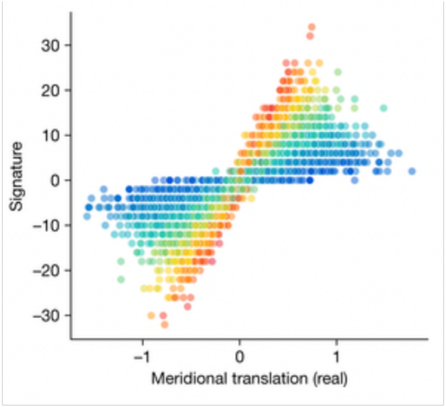




Cusp torus for 6_1

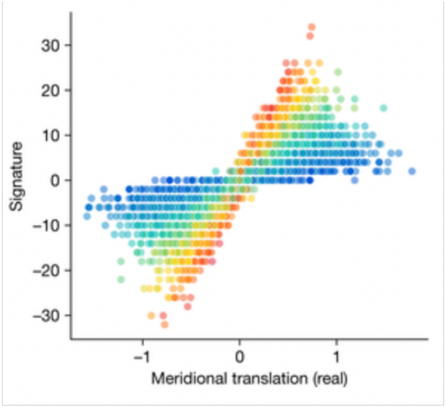
The three main features that the machine learning algorithms used to predict signature were λ, $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$.

Signature and cusp geometry



A plot of signature against $\mathrm{Re}(\mu)$ coloured by λ

Signature and cusp geometry

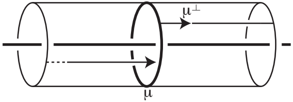


A plot of signature against $\mathrm{Re}(\mu)$ coloured by λ

Initial observation: the signs of the signature and $\mathrm{Re}(\mu)$ are highly correlated.

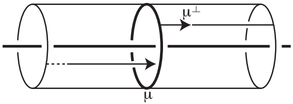
The natural slope

The natural slope



- Pick a geodesic representative μ for the meridian.
- Fire a geodesic $\mu^{\perp}$ orthogonally from it.
- Eventually, it will return to the meridian.
- In that time, it will have gone along one longitude and some number s of meridians.

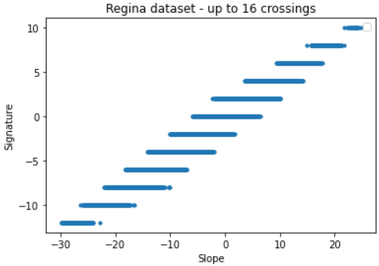
The natural slope

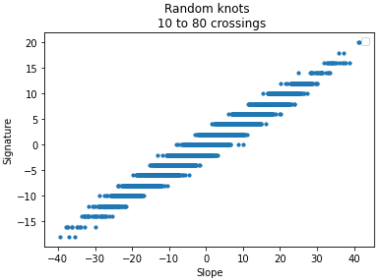


- Pick a geodesic representative μ for the meridian.
- Fire a geodesic $\mu^\perp$ orthogonally from it.
- Eventually, it will return to the meridian.
- In that time, it will have gone along one longitude and some number s of meridians.
- Define the natural slope to be $-s$.

$$\mathrm{slope}(K) = \mathrm{Re}(\lambda/\mu).$$

Slope and signature





First conjectures

Conjecture: There is a constant c_0 such that

$$\sigma(K) \simeq c_0 \,\mathrm{slope}(K).$$

First conjectures

Conjecture: There is a constant c_0 such that

$$\sigma(K) \simeq c_0 \operatorname{slope}(K).$$

Conjecture: There are constants c_0 and c_1 such that

$$|\sigma(K) - c_0 \operatorname{slope}(K)| \leq c_1 \operatorname{vol}(K).$$

Highly twisted knots

Theorem: Let K be a knot, and let $C_1, \dots, C_n$ be curves in the complement that bound disjoint discs in S^3. Suppose $K \cup C_1 \cup \cdots \cup C_n$ is hyperbolic. Let $K(q_1, \dots, q_n)$ be the knot obtained from K by adding q_i full twists along each C_i.

Highly twisted knots

Theorem: Let K be a knot, and let $C_1, \dots, C_n$ be curves in the complement that bound disjoint discs in S^3. Suppose $K \cup C_1 \cup \dots \cup C_n$ is hyperbolic. Let $K(q_1, \dots, q_n)$ be the knot obtained from K by adding q_i full twists along each C_i. Let $\ell_i = \mathrm{lk}(K, C_i)$. Suppose $\ell_1, \dots, \ell_m$ are even and $\ell_{m+1}, \dots, \ell_n$ are odd.

Highly twisted knots

Theorem: Let K be a knot, and let $C_1, \ldots, C_n$ be curves in the complement that bound disjoint discs in S^3. Suppose $K \cup C_1 \cup \cdots \cup C_n$ is hyperbolic. Let $K(q_1, \ldots, q_n)$ be the knot obtained from K by adding q_i full twists along each C_i. Let $\ell_i = \mathrm{lk}(K, C_i)$. Suppose $\ell_1, \ldots, \ell_m$ are even and $\ell_{m+1}, \ldots, \ell_n$ are odd. Then there is a constant k such that if each $|q_i| >> 0$,

$$\left| \mathrm{slope}(K(q_1, \ldots, q_n)) + \sum_{i=1}^{n} \ell_i^2 q_i \right| \leq k$$

$$\left| \sigma(K(q_1, \ldots, q_n)) + \left(\frac{1}{2} \sum_{i=1}^{m} \ell_i^2 q_i + \frac{1}{2} \sum_{i=m+1}^{n} (\ell_i^2 - 1) q_i \right) \right| \leq k$$

$$\mathrm{vol}(K(q_1, \ldots, q_n)) \leq k.$$

So the conjectures are false!

Theorems

Theorem 1: There is a constant c_1 such that

$$|\sigma(K) - (1/2)\,\mathrm{slope}(K)| \leq c_1\,\mathrm{vol}(K)\,\mathrm{inj}(K)^{-3}.$$

Here, $\mathrm{inj}(K)$ is $\inf\{\mathrm{inj}_x(S^3 - K) : x \in (S^3 - K) - \text{cusp}\}$.

Theorems

Theorem 1: There is a constant c_1 such that

$$|\sigma(K) - (1/2)\,\mathrm{slope}(K)| \leq c_1 \,\mathrm{vol}(K)\,\mathrm{inj}(K)^{-3}.$$

Here, $\mathrm{inj}(K)$ is $\inf\{\mathrm{inj}_x(S^3 - K) : x \in (S^3 - K) - \text{cusp}\}$.

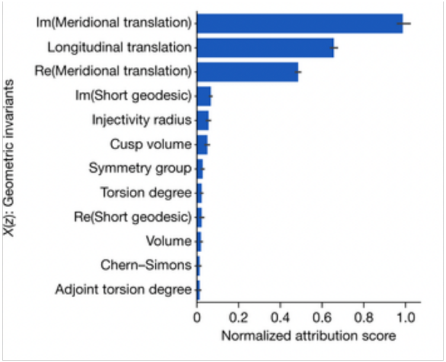
Theorem 2: $\sigma(K)$ and

$$(1/2)\,\mathrm{slope}(K) + \sum_{\gamma \in \mathrm{OddGeo}} \kappa(\gamma)$$

differ by at most $c_2 \mathrm{vol}(K)$ for some constant c_2.

Here, OddGeo is the set of geodesics with length at most 0.1 and that have odd linking number with K, and $\kappa(\gamma)$ is a correction term defined in terms of the complex length of γ.

The machine knew all along!



Items 4 and 5 are the terms appearing in Theorems 1 and 2.

Difficulties with this method

Difficulties with this method

- Finding a formula for F currently requires human input

Difficulties with this method

- Finding a formula for F currently requires human input
- What about inputs other than real numbers?

Difficulties with this method

- Finding a formula for F currently requires human input
- What about inputs other than real numbers?
- ML tends to ignore outliers

Further examples

The Jones polynomial $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ of K is a mysterious invariant.

Further examples

The Jones polynomial $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ of K is a mysterious invariant.

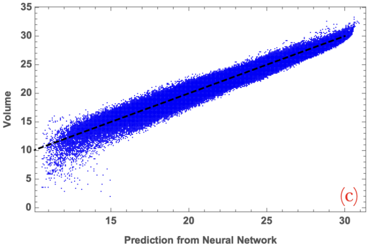
Is it related to other invariants?

Further examples

The Jones polynomial $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ of K is a mysterious invariant.

Is it related to other invariants?

[Jejjalaa, Kar, Parrikar]: The Jones polynomial seems to encode information about the hyperbolic volume

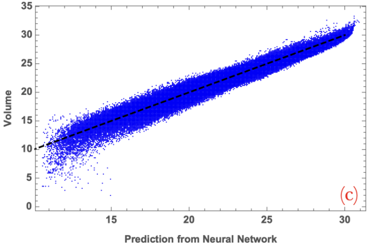


Further examples

The Jones polynomial $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ of K is a mysterious invariant.

Is it related to other invariants?

[Jejjalaa, Kar, Parrikar]: The Jones polynomial seems to encode information about the hyperbolic volume



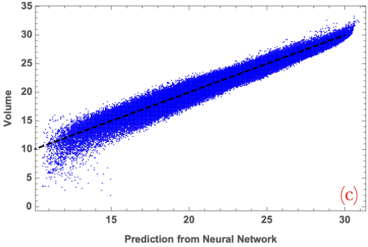
But it seems hard to encapsulate this into a conjecture.

Further examples

The Jones polynomial $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ of K is a mysterious invariant.

Is it related to other invariants?

[Jejjalaa, Kar, Parrikar]: The Jones polynomial seems to encode information about the hyperbolic volume



But it seems hard to encapsulate this into a conjecture.

Many other connections found by [Craven, Hughes, Jejjala, Kar].