

How machines learn Galois groups

Seewoo Lee, UC Berkeley

Today's Goal

Let's understand how decision trees classify Galois groups of (Galois) number fields with Dedekind zeta coefficients!

This is a joint work with Kyu-Hwan Lee.

Recently, there has been a lot of work that uses “AI” for mathematics, including:

- AlphaTensor for matrix multiplication
- Reinforcement learning for group theory
- Transformers for ODEs
- LLMs for writing formal/informal proofs

Recently, there has been a lot of work that uses “AI” for mathematics, including:

- AlphaTensor for matrix multiplication
- Reinforcement learning for group theory
- Transformers for ODEs
- LLMs for writing formal/informal proofs

Today: ML for predicting number-theoretic invariants.

Case 1: Elliptic Curve (He–Lee¹–Oliver–Podznyakov)

An **elliptic curve** (over \mathbb{Q}) is a curve given by an equation of the form

$$E : y^2 = x^3 + ax + b, \quad a, b \in \mathbb{Q}.$$

The set of rational points

$$E(\mathbb{Q}) := \{(x, y) \in \mathbb{Q}^2 : y^2 = x^3 + ax + b\} \cup \{\infty\}$$

forms a finitely generated abelian group (Mordell–Weil).

¹Kyu-Hwan Lee, not Seewoo Lee

Rank of an elliptic curve

One can decompose $E(\mathbb{Q})$ as

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$$

The torsion subgroup $E(\mathbb{Q})_{\text{tors}}$ is well-understood (Nagell–Lutz, Mazur), but the rank r is **very hard** to compute!

Rank of an elliptic curve

One can decompose $E(\mathbb{Q})$ as

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$$

The torsion subgroup $E(\mathbb{Q})_{\text{tors}}$ is well-understood (Nagell–Lutz, Mazur), but the rank r is **very hard** to compute!

- BSD conjecture ($r = \text{ord}_{s=1} L(E, s)$)
- Is rank unbounded? (Elkies–Klagsbrun found rank ≥ 29 curve)

Rank of an elliptic curve

One can decompose $E(\mathbb{Q})$ as

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$$

The torsion subgroup $E(\mathbb{Q})_{\text{tors}}$ is well-understood (Nagell–Lutz, Mazur), but the rank r is **very hard** to compute!

- BSD conjecture ($r = \text{ord}_{s=1} L(E, s)$)
- Is rank unbounded? (Elkies–Klagsbrun found rank ≥ 29 curve)

What if we use machine learning to predict the rank?

Yang-Hui He, Kyu-Hwan Lee, and Thomas Oliver² used *Frobenius traces* $a_p(E)$ as features (inputs):

$$a_p(E) = p + 1 - |E(\mathbb{F}_p)|$$

where $|E(\mathbb{F}_p)|$ is the number of points on the reduced curve modulo p . Note that the rank is determined by $a_p(E)$ for *all* p .

They used logistic regression to distinguish rank 0 and rank 1 curves using a_{p_n} for $n \leq 300$.

²He–Lee–Oliver, *Machine learning invariants of arithmetic curves*, 2023

Quiz!

What was the accuracy of the experiment? Total 32000 curves (16000 for each rank) are used for training and testing.

- 1 0%
- 2 (0%, 50%]
- 3 (50%, 75%]
- 4 (75%, 90%]
- 5 (90%, 95%]
- 6 (95%, 100%)
- 7 100%

Quiz!

What was the accuracy of the experiment? Total 32000 curves (16000 for each rank) are used for training and testing.

- ❶ 0%
- ❷ (0%, 50%]
- ❸ (50%, 75%]
- ❹ (75%, 90%]
- ❺ (90%, 95%]
- ❻ (95%, 100%) (99.1%)
- ❼ 100%

Quiz!

What was the accuracy of the experiment? Total 32000 curves (16000 for each rank) are used for training and testing.

- 1 0%
- 2 (0%, 50%]
- 3 (50%, 75%]
- 4 (75%, 90%]
- 5 (90%, 95%]
- 6 (95%, 100%) (99.1%)
- 7 100%

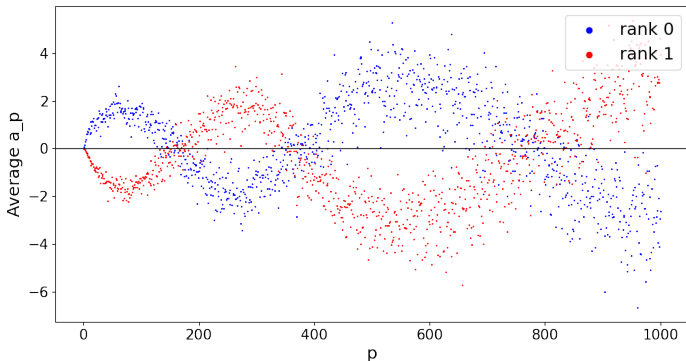
Why does it work so well?

Flying elliptic curves

While analyzing the model and the data, Podznyakov plotted the average of a_p for each p as a function of p :

Flying elliptic curves

While analyzing the model and the data, Podznyakov plotted the average of a_p for each p as a function of p :



No one plotted the graph before!

It is now called *murmuration*³, and has become an active area of study. The oscillating pattern is observed in other “families of L -functions”, and the “limit curve” has even been computed in many cases. Also related to the one-level density conjecture by Katz–Sarnak.

In this case, ML **motivated** mathematicians to find a new phenomenon.

³He–Lee–Oliver–Podznyakov, *Murmuration of elliptic curves*, 2024

Case 2: Number fields (He–Lee–Oliver, Lee–L.)

Definition

A **number field** is a finite extension K of \mathbb{Q} .

Case 2: Number fields (He–Lee–Oliver, Lee–L.)

Definition

A **number field** is a finite extension K of \mathbb{Q} .

- \mathbb{Q}
- $\mathbb{Q}(i) = \{a + bi : a, b \in \mathbb{Q}\}$
- $\mathbb{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} : a, b, c \in \mathbb{Q}\}$
- and many more ...

Definition

Galois group of a number field K over \mathbb{Q} is defined as

$$\text{Gal}(K/\mathbb{Q}) = \{\text{automorphisms of } K \text{ fixing } \mathbb{Q}\}$$

Galois group

Definition

Galois group of a number field K over \mathbb{Q} is defined as

$$\text{Gal}(K/\mathbb{Q}) = \{\text{automorphisms of } K \text{ fixing } \mathbb{Q}\}$$

$$\text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) = \{\text{id}, (a + bi \mapsto a - bi)\} \simeq C_2.$$

Galois group

Definition

Galois group of a number field K over \mathbb{Q} is defined as

$$\text{Gal}(K/\mathbb{Q}) = \{\text{automorphisms of } K \text{ fixing } \mathbb{Q}\}$$

$$\text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) = \{\text{id}, (a + bi \mapsto a - bi)\} \simeq C_2.$$

Definition

A number field K is called **Galois** (or normal) if $\text{Gal}(K/\mathbb{Q})$ has size equal to the degree $[K : \mathbb{Q}]$.

Galois group

Definition

Galois group of a number field K over \mathbb{Q} is defined as

$$\text{Gal}(K/\mathbb{Q}) = \{\text{automorphisms of } K \text{ fixing } \mathbb{Q}\}$$

$$\text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) = \{\text{id}, (a + bi \mapsto a - bi)\} \simeq C_2.$$

Definition

A number field K is called **Galois** (or normal) if $\text{Gal}(K/\mathbb{Q})$ has size equal to the degree $[K : \mathbb{Q}]$.

Any quadratic extension is Galois.

$\mathbb{Q}(\sqrt[3]{2})$ is not Galois, since there's no nontrivial automorphism fixing \mathbb{Q} , while degree is 3.

Question

How do we compute $\text{Gal}(K/\mathbb{Q})$?

Question

How do we compute $\text{Gal}(K/\mathbb{Q})$?

It is not an easy problem.

Quiz!

$$K = \mathbb{Q}[x]/(x^4 + 2x^2 + 4)$$

The Galois group $\text{Gal}(K/\mathbb{Q})$ is isomorphic to

- ☒ C_4
- ☐ C_2^2

Quiz!

$$K = \mathbb{Q}[x]/(x^4 + 2x^2 + 4) \simeq \mathbb{Q}(\sqrt{2}, \sqrt{-3})$$

The Galois group $\text{Gal}(K/\mathbb{Q})$ is isomorphic to

- ☒ C_4
- ☐ C_2^2

Quiz!

$$K = \mathbb{Q}[x]/(x^8 - 2x^7 + 2x^6 - 2x^5 + 7x^4 - 10x^3 + 8x^2 - 4x + 1)$$

LMFDB 8.0.2985984.1. Abelian or non-abelian?

- 1 Abelian (C_8 or $C_4 \times C_2$ or C_2^3)
- 2 Non-abelian (D_4 or Q_8)

Quiz!

$$K = \mathbb{Q}[x]/(x^8 - 2x^7 + 2x^6 - 2x^5 + 7x^4 - 10x^3 + 8x^2 - 4x + 1)$$

LMFDB 8.0.2985984.1. Abelian or non-abelian?

- 1 Abelian (C_8 or $C_4 \times C_2$ or C_2^3)
- 2 Non-abelian (D_4 or Q_8)

Question

Can we use ML to predict the Galois group?

Dedekind zeta function

The analogue of $a_p(E)$ for number fields are the coefficients of the *Dedekind zeta function*.

Definition

For a number field K , the **Dedekind zeta function** $\zeta_K(s)$ is defined as

$$\zeta_K(s) = \sum_{\mathfrak{a} \subset \mathcal{O}_K} \frac{1}{(N\mathfrak{a})^s}$$

where the sum is over all nonzero ideals \mathfrak{a} of the ring of integers \mathcal{O}_K of K , and $N\mathfrak{a} = |\mathcal{O}_K/\mathfrak{a}|$ is the norm of \mathfrak{a} .

Dedekind zeta function

It can be rewritten as

$$\zeta_K(s) = \sum_{n \geq 1} \frac{a_n(K)}{n^s}$$

for $a_n(K) = |\{\mathfrak{a} \subset \mathcal{O}_K : N\mathfrak{a} = n\}|$. These will be used as features.

Dedekind zeta function

It can be rewritten as

$$\zeta_K(s) = \sum_{n \geq 1} \frac{a_n(K)}{n^s}$$

for $a_n(K) = |\{\mathfrak{a} \subset \mathcal{O}_K : N\mathfrak{a} = n\}|$. These will be used as features.

It admits an Euler product:

$$\zeta_K(s) = \prod_p \zeta_{K,p}(s) = \prod_p \left(1 + \frac{a_p(K)}{p^s} + \frac{a_{p^2}(K)}{p^{2s}} + \dots \right).$$

In particular, $a_{mn}(K) = a_m(K)a_n(K)$ for $\gcd(m, n) = 1$.

Examples

$$K \rightarrow (a_n(K))_{n \geq 1}$$

$$\mathbb{Q} \rightarrow (1, 1, 1, 1, 1, 1, 1, 1, \dots)$$

$$\mathbb{Q}(i) \rightarrow (1, 1, 0, 1, 2, 0, 0, 1, 1, \dots)$$

$$\mathbb{Q}[x]/(x^4 + 2x^2 + 4) \rightarrow (1, 0, 0, 1, 0, 0, 4, 0, 1, \dots)$$

He, Lee, and Oliver⁴ used classical ML algorithms to predict invariants of number fields, including Galois groups, using Dedekind zeta coefficients. They could distinguish among abelian degree 8 extensions with an accuracy of 95%.

⁴He–Lee–Oliver, *Machine Learning Number Fields*, 2022

He, Lee, and Oliver⁴ used classical ML algorithms to predict invariants of number fields, including Galois groups, using Dedekind zeta coefficients. They could distinguish among abelian degree 8 extensions with an accuracy of 95%.


But there was no explanation for **why** it worked so well.

⁴He–Lee–Oliver, *Machine Learning Number Fields*, 2022

Experimental setup

- Fix degree $\in \{4, 6, 8, 9, 10\}$.
- Input (feature): $\{a_n(K)\}_{n \leq N}$, e.g., $N = 1000$.
- Output (target): the Galois group
- Model: **Decision tree (easy to interpret)**
- **We'll only focus on Galois extensions.**

Data (LMFDB + Sage)



[Introduction](#)
[Overview](#) [Random Universe](#)
[L-functions](#)
[Rational](#) [All](#)
[Modular forms](#)
[Classical](#) [Maass](#)
[Hilbert](#) [Bianchi](#)
[Varieties](#)
[Elliptic curves over \$\mathbb{Q}\$](#)
[Elliptic curves over \$\mathbb{Q}\(\alpha\)\$](#)
[Genus 2 curves over \$\mathbb{Q}\$](#)
[Higher genus families](#)
[Abelian varieties over \$\mathbb{F}_q\$](#)
[Belyi maps](#)
[Fields](#)
[Number fields](#)
[p-adic fields](#)
[Representations](#)
[Dirichlet characters](#)
[Artin representations](#)
[Groups](#)
[Galois groups](#)
[Sato-Tate groups](#)
[Abstract groups](#)
[Database](#)

[\$\hat{\mathbb{O}} \rightarrow\$ Number fields \$\rightarrow\$ 9.9.301789003173921081.1](#)

Number field 9.9.301789003173921081.1

[Normalized defining polynomial](#)

$$x^9 - 3x^8 - 120x^7 + 14x^6 + 5109x^5 + 12303x^4 - 59793x^3 - 316281x^2 - 504861x - 269801$$

[Show commands: Magma / Oscar / Pari/GP / SageMath](#)

[Invariants](#)

Degree: 9
Signature: $[9, 0]$
Discriminant: $301789003173921081 = 3^{12} \cdot 7^6 \cdot 13^6$
Root discriminant: 87.54
Galois root discriminant: $3^{4/3} 7^{2/3} 13^{2/3} \approx 87.5365180904025$
Ramified primes: 3, 7, 13
Discriminant root field: \mathbb{Q}
 $\text{Aut}(K/\mathbb{Q}) = \text{Gal}(K/\mathbb{Q})$: C_3^2
This field is Galois and abelian over \mathbb{Q} .
Conductor: $819 = 3^2 \cdot 7 \cdot 13$
Dirichlet character group: $\{\chi_{819}(352, \cdot), \chi_{819}(1, \cdot), \chi_{819}(646, \cdot), \chi_{819}(295, \cdot), \chi_{819}(235, \cdot), \chi_{819}(529, \cdot), \chi_{819}(562, \cdot), \chi_{819}(211, \cdot), \chi_{819}(445, \cdot)\}$
This is not a CM field.
This field has no CM subfields.

[Integral basis \(with respect to field generator \$\alpha\$ \)](#)

$$1, \alpha, \alpha^2, \alpha^3, \frac{1}{110}\alpha^4 - \frac{1}{11}\alpha^2 - \frac{1}{11}\alpha - \frac{1}{11}, \frac{1}{61576445}\alpha^5 - \frac{1}{61576445}\alpha^3 - \frac{1}{61576445}\alpha^2 - \frac{1}{61576445}\alpha - \frac{1}{61576445}, \frac{1}{246306179}\alpha^6 - \frac{1}{246306179}\alpha^4 - \frac{1}{246306179}\alpha^3 - \frac{1}{246306179}\alpha^2 - \frac{1}{246306179}\alpha - \frac{1}{246306179}, \frac{1}{53979675}\alpha^7 - \frac{1}{53979675}\alpha^5 - \frac{1}{53979675}\alpha^4 - \frac{1}{53979675}\alpha^3 - \frac{1}{53979675}\alpha^2 - \frac{1}{53979675}\alpha - \frac{1}{53979675}, \frac{1}{1201530650}\alpha^8 - \frac{1}{1201530650}\alpha^6 - \frac{1}{1201530650}\alpha^5 - \frac{1}{1201530650}\alpha^4 - \frac{1}{1201530650}\alpha^3 - \frac{1}{1201530650}\alpha^2 - \frac{1}{1201530650}\alpha - \frac{1}{1201530650}$$

Monogenic: No
Index: Not computed
Inessential primes: 2

[Class group and class number](#)

[Citation](#) · [Feedback](#) · [Hide Menu](#)

[!\[\]\(37f115ba2c702ea4bb164007f296097c_img.jpg\)](#)

[!\[\]\(3aa883258a9fdc29c652ea3fde92670b_img.jpg\)](#)

Label: 9.9.301789003173921081.1

Degree: 9
Signature: $[9, 0]$
Discriminant: 3.018×10^{17}
Root discriminant: 87.54
Ramified primes: 3, 7, 13
Class number: 3 (GRH)
Class group: $[3]$ (GRH)
Galois group: C_3^2 (as 9T2)

[Related objects](#)

Galois group
Discriminant root field
Dirichlet character group

[Downloads](#)

Stored data to gp
Magma commands
Oscar commands
Pari/GP commands
SageMath commands
Underlying data

[Learn more](#)

Source and acknowledgments
Completeness of the data
Reliability of the data
Number field labels

Use Sage's `zeta_coefficients()` to compute $a_n(K)$.

Let K be a degree 9 Galois extension of \mathbb{Q} . We have two possibilities for the Galois group:

$$C_9, \quad C_3^2$$

Quiz!

There are 1266 Galois degree 9 fields in the LMFDB; 22% are C_9 and 78% are C_3^2 . Split them randomly into a training (80%) and a test (20%) set. What was the accuracy of the decision tree model?

- 1 0%
- 2 (0%, 50%]
- 3 (50%, 75%]
- 4 (75%, 90%]
- 5 (90%, 95%]
- 6 (95%, 100%)
- 7 100%

Quiz!

There are 1266 Galois degree 9 fields in the LMFDB; 22% are C_9 and 78% are C_3^2 . Split them randomly into a training (80%) and a test (20%) set. What was the accuracy of the decision tree model?

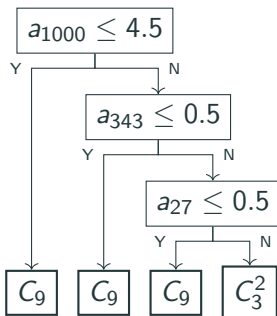
- ☐ 0%
- ☐ (0%, 50%]
- ☐ (50%, 75%]
- ☐ (75%, 90%]
- ☐ (90%, 95%]
- ☐ (95%, 100%)
- ☒ 100%

Why? Let's look inside the tree.

Decision trees are great since they are often easy to interpret. Here is the tree that achieves 100% accuracy:

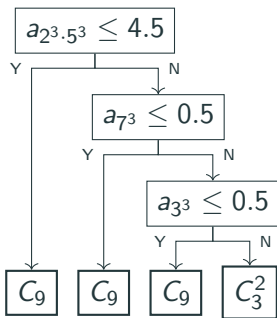
Why? Let's look inside the tree.

Decision trees are great since they are often easy to interpret. Here is the tree that achieves 100% accuracy:



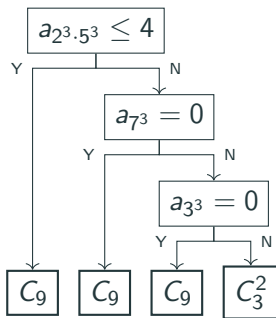
Why? Let's look inside the tree.

Decision trees are great since they are often easy to interpret. Here is the tree that achieves 100% accuracy:

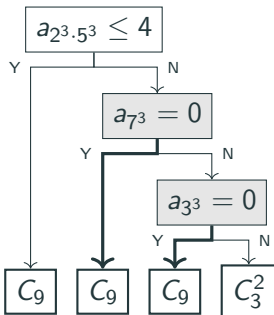


Why? Let's look inside the tree.

Decision trees are great since they are often easy to interpret. Here is the tree that achieves 100% accuracy:



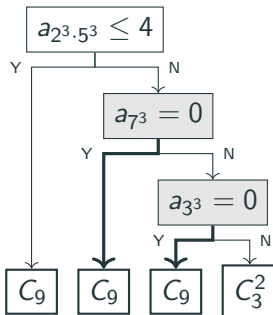
Tree's logic



Prediction

Let K/\mathbb{Q} be a degree 9 Galois extension. If $a_{p^3}(K) = 0$ for some prime p , then $\text{Gal}(K/\mathbb{Q}) \simeq C_9$.

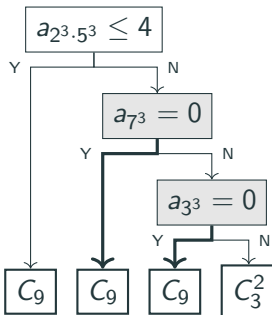
Tree's logic



Proposition (Lee-L.)

Let K/\mathbb{Q} be a degree 9 Galois extension. Then $a_{p^3}(K) = 0$ for some prime p if and only if $\text{Gal}(K/\mathbb{Q}) \simeq C_9$.

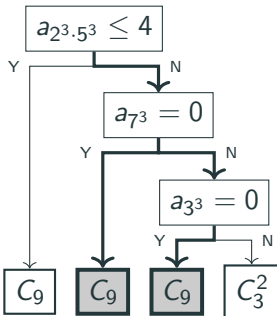
Tree's logic



Proposition (Lee–L.)

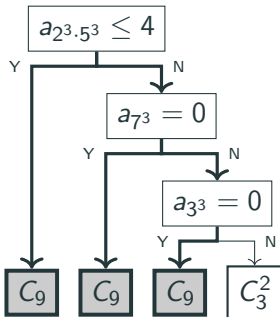
Let ℓ be a prime and K/\mathbb{Q} be a degree ℓ^2 Galois extension. Then $a_{p^\ell}(K) = 0$ for some prime p if and only if $\text{Gal}(K/\mathbb{Q}) \simeq C_{\ell^2}$.

Provable prediction



The bold paths give **provably correct predictions!**

Provable prediction



The bold paths give **provably correct predictions!**

Proof (short)

(\Rightarrow) The local Euler factor of $\zeta_K(s)$ at a prime p is

$$\zeta_{K,p}(s) = \prod_{\mathfrak{p}|p} (1 - N\mathfrak{p}^{-s})^{-1} = (1 - p^{-fs})^{-g} = \sum_{k \geq 0} \binom{g+k-1}{g-1} p^{-fks}$$

where $p\mathcal{O}_K = \prod_{i=1}^g \mathfrak{p}_i^e$ with $N\mathfrak{p}_i = p^f$. We have 6 possible combinations for (e, f, g) :

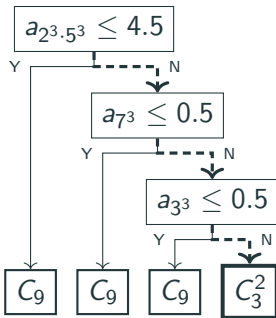
$$(1, 1, \ell^2), (1, \ell, \ell), (1, \ell^2, 1), (\ell, 1, \ell), (\ell, \ell, 1), (\ell^2, 1, 1)$$

and $a_{p^\ell}(K) = 0$ only occurs when $(e, f, g) = (1, \ell^2, 1)$. Thus $\text{Gal}(K/\mathbb{Q}) \simeq \text{Gal}(K_{\mathfrak{p}}/\mathbb{Q}_p) \simeq \text{Gal}(\mathbb{F}_{p^{\ell^2}}/\mathbb{F}_p) \simeq C_{\ell^2}$.

Proof (short)

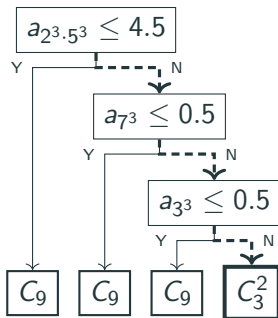
(\Leftarrow) By the Chebotarev density theorem, there are infinitely many (indeed, of positive density) primes p totally inert in K/\mathbb{Q} , i.e. $(e, f, g) = (1, \ell^2, 1)$. For such p , we have $a_{p^\ell}(K) = 0$.

The other branch



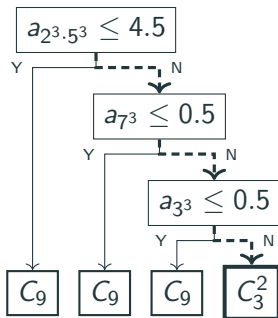
How about the rightmost path?

The other branch



How about the rightmost path? The unique degree 9 subextension K of the cyclotomic field $\mathbb{Q}(\zeta_{960643})$ satisfies $a_{1000} = 27225$ and $a_{343} = a_{27} = 165$, so the decision tree predicts $\text{Gal}(K/\mathbb{Q}) \simeq C_3^2$, but in fact $\text{Gal}(K/\mathbb{Q}) \simeq C_9$.

The other branch



How about the rightmost path? The unique degree 9 subextension K of the cyclotomic field $\mathbb{Q}(\zeta_{960643})$ satisfies $a_{1000} = 27225$ and $a_{343} = a_{27} = 165$, so the decision tree predicts $\text{Gal}(K/\mathbb{Q}) \simeq C_3^2$, but in fact $\text{Gal}(K/\mathbb{Q}) \simeq C_9$. But it is not in the LMFDB!

So trees are helpful

Degree $4 = 2^2$ case is similar.

In both cases, we could easily interpret the decision tree, formulate a prediction (conjecture), and prove it.

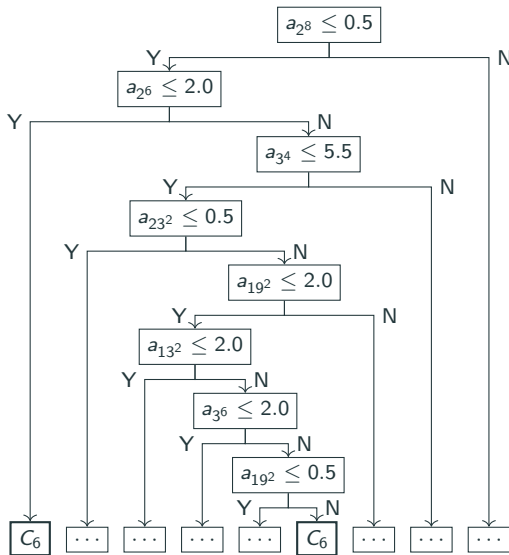
Degree 6

Degrees 6, 8, and 10 are more interesting since nonabelian Galois groups can occur.

Degrees 6, 8, and 10 are more interesting since nonabelian Galois groups can occur.

For degree 6, decision tree achieved an accuracy of 98.98% (C_6 vs S_3). But the tree is more complicated than the degree 9 case.

Degree 6



Degree 6 - exploratory data analysis

However, the tree indicates that $a_n(K)$ for **square** n are more important than the other $a_n(K)$'s.

Degree 6 - exploratory data analysis

However, the tree indicates that $a_n(K)$ for **square** n are more important than the other $a_n(K)$'s.

Based on the observation, we studied **data**, and focused on the possibility of classifying Galois groups using $a_{p^2}(K)$ for prime p .

Degree 6 - exploratory data analysis

And we indeed found that such coefficients are useful. For example, the possible values of $a_{25}(K)$ for C_6 -extensions are

$$0, \quad 3, \quad 6, \quad 21,$$

while those for S_3 -extensions are

$$0, \quad \mathbf{1}, \quad 3, \quad 6, \quad 21,$$

so we expect $a_{25}(K) = 1 \Rightarrow \text{Gal}(K/\mathbb{Q}) \simeq S_3$.

Degree 6 - exploratory data analysis

However, the possible values of $a_{49}(K)$ for C_6 -extensions are

$$0, \quad 1, \quad 3, \quad 6, \quad 21,$$

while those for S_3 -extensions are

$$0, \quad 3, \quad 6, \quad 21,$$

so we expect $a_{49}(K) = 1 \Rightarrow \text{Gal}(K/\mathbb{Q}) \simeq C_6$.

Prediction

Let K/\mathbb{Q} be a degree 6 Galois extension.

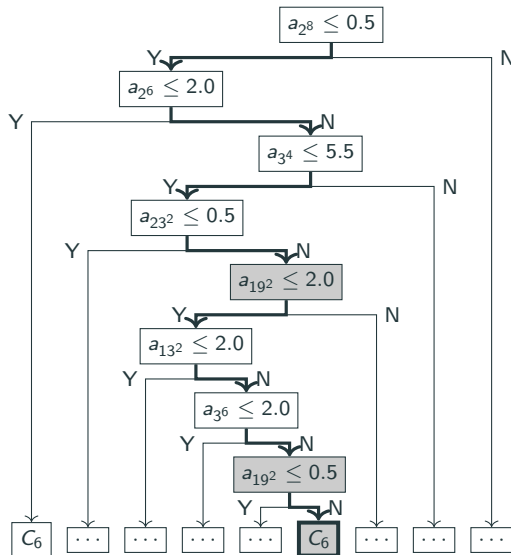
- ❶ *If there exists a prime $p \equiv 1 \pmod{6}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_6$.*
- ❷ *If there exists a prime $p \equiv 5 \pmod{6}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq S_3$.*
- ❸ *If there exists a prime p with $a_{p^3}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_6$.*

Proposition (Lee–L.)

Let K/\mathbb{Q} be a degree 6 Galois extension.

- ❶ *If there exists a prime $p \equiv 1 \pmod{6}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_6$.*
- ❷ *If there exists a prime $p \equiv 5 \pmod{6}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq S_3$.*
- ❸ *If there exists a prime p with $a_{p^3}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_6$.*

Degree 6



If $a_{19^2} \leq 2.0$ and $a_{19^2} > 0.5$, then $a_{19^2} = 1$ ($a_{19^2} = 2$ is not possible). From $19 \equiv 1 \pmod{6}$, the theorem implies $\text{Gal}(K/\mathbb{Q}) \simeq C_6$. So the bold path gives a provably correct prediction!

Proofs are based on the following results and case-by-case analysis.

- (Greenberg, Newton) Let L/K be an abelian extension of local fields with ramification index e and residue field k of K of characteristic p . If $p \nmid e$, then $e \mid |k^\times|$.
- (Iwasawa) Let K/\mathbb{Q}_p be an extension with ramification index e . If $p \nmid e$, then a lift of the Frobenius in $\text{Gal}(K/\mathbb{Q}_p)$ acts on the inertia subgroup by raising to the p -th power.

These can be used to rule out some ramification types, e.g.,

- Let K/\mathbb{Q} be a cyclic degree 6 extension and let p be a rational prime with ramification type (e, f, g) in K . If $p \equiv 2 \pmod{3}$, then

$$(e, f, g) \notin \{(3, 1, 2), (3, 2, 1), (6, 1, 1)\}.$$

- If $p \equiv 1 \pmod{6}$, then any degree 6 Galois extension of \mathbb{Q}_p is cyclic.

Proposition (Lee–L.)

Let K/\mathbb{Q} be a degree 10 Galois extension.

- ❶ *If there exists a prime $p \equiv 1 \pmod{10}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_{10}$.*
- ❷ *If there exists a prime $p \equiv 9 \pmod{10}$ such that $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq D_5$.*
- ❸ *If there exists a prime $p \neq 5$ such that $a_{p^5}(K) = 1$, then $\text{Gal}(K/\mathbb{Q}) \simeq C_{10}$.*

Proofs are similar to the degree-6 case. Decision tree achieved an accuracy of 97.63%.

For degree 8 Galois extensions, we have five possible Galois groups:

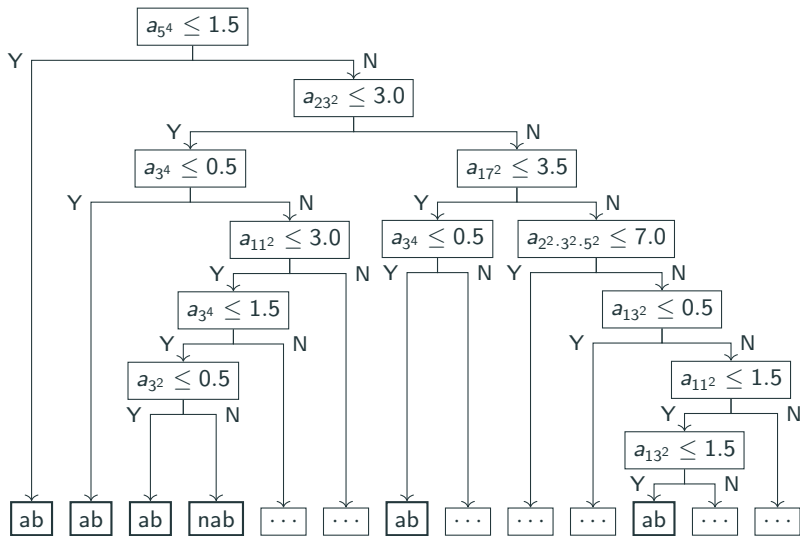
$$C_8, \quad C_4 \times C_2, \quad C_2^3, \quad D_4, \quad Q_8$$

We conducted three experiments:

- 1 Distinguish abelian extensions ($C_8, C_4 \times C_2, C_2^3$)
- 2 Distinguish nonabelian extensions (D_4, Q_8)
- 3 Abelian vs nonabelian

We achieved accuracies $> 96\%$ in all three experiments.

Degree 8, abelian vs nonabelian



Degree 8, abelian vs nonabelian

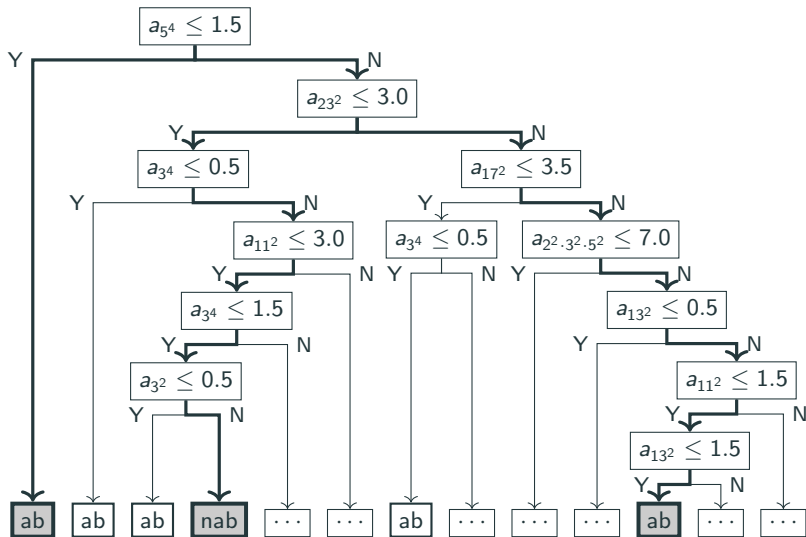
Again, it is a large tree, but we can see that square-indexed coefficients are more important. We prove:

Proposition (Lee–L.)

Let K/\mathbb{Q} be an degree 8 Galois extension.

- ❶ *If $a_{p^4}(K) = 0$, then $\text{Gal}(K/\mathbb{Q})$ is a C_8 -extension (hence abelian).*
- ❷ *For $p \equiv 1 \pmod{4}$, if $a_{p^4}(K) = 1$ or $a_{p^2}(K) = 1$, then $\text{Gal}(K/\mathbb{Q})$ is abelian.*
- ❸ *If $p \equiv 3 \pmod{4}$, $a_{p^4}(K) = 1$, and $a_{p^2}(K) > 0$, then $\text{Gal}(K/\mathbb{Q})$ is nonabelian.*

Degree 8, provably correct predictions



Review Quiz

$$K = \mathbb{Q}[x]/(x^8 - 2x^7 + 2x^6 - 2x^5 + 7x^4 - 10x^3 + 8x^2 - 4x + 1)$$

Review Quiz

$$K = \mathbb{Q}[x]/(x^8 - 2x^7 + 2x^6 - 2x^5 + 7x^4 - 10x^3 + 8x^2 - 4x + 1)$$

Dedekind zeta coefficients of K are

$$(1, 0, 0, 1, 0, \dots, a_9 = 1, \dots, a_{81} = 1, \dots)$$

Hence it is a nonabelian extension.

- Using polynomial coefficients as features also gives decent accuracy, but it is not easy to interpret. This likely depends on how the defining polynomials in the LMFDB are chosen.
- Logistic regression also works well. The trained weights reflect similar number-theoretic phenomena.

Takeaways

- ML is useful to **motivate** mathematicians to formulate new conjectures. In particular, decision trees are great for interpretability.
- Even if the model is not easy to interpret, it can tell you which features are important, and further exploratory data analysis can help discover interesting patterns in the data.
- Sometimes, we can **prove** the predictions made by ML models.

What's next?

Some questions

Question

Can our work be used as a new and more efficient algorithm for computing Galois groups?

Some questions

Question

Can our work be used as a new and more efficient algorithm for computing Galois groups?

Unfortunately, the answer is no.

- ❶ We only considered Galois extensions, so we need to know that the extension is Galois beforehand, without knowing the actual Galois group.
- ❷ However, one can consider non-Galois extensions, and it should be possible to derive similar theorems (determining the Galois group from the a_n 's) as above.
- ❸ Depends on computing a_n 's or ramification types of primes.

Some questions

For degree 8, a decision tree can distinguish between D_4 and Q_8 extensions with 98.8% accuracy (for $n \leq 1000$), even when the possible values of a_{p^k} are the same for both groups. But their *distributions* are quite different, e.g., we have

$$\mathbb{P}[\mathrm{Gal}(K/\mathbb{Q}) \simeq D_4 | a_{5^2}(K) = 3] \approx 0.07$$

$$\mathbb{P}[\mathrm{Gal}(K/\mathbb{Q}) \simeq Q_8 | a_{5^2}(K) = 3] \approx 0.93$$

using data from the LMFDB.

Question

Can we compute these conditional probabilities theoretically?

cf. Wood; Shankar–Varma; Koymans–Pagano.

Some questions

We can ask whether there's a general criterion for predicting Galois groups from the a_n 's. After reading our draft, Yang-Hui He asked:

Question (Yang-Hui He)

For a given group G of order d , are there infinitely many (n, a) such that $a_n(K) = a$ implies $\text{Gal}(K/\mathbb{Q}) \simeq G$ for any degree d Galois number field K ?

The answer is negative for $G = D_4$ and Q_8 , as we just saw. But it seems to be positive, e.g., for cyclic groups.

Some questions

Question

Can we scale up the experiments? Using modern ML/DL models (e.g., transformers)? What would they learn?

Some questions

Question

Can we scale up the experiments? Using modern ML/DL models (e.g., transformers)? What would they learn?

Question

Can we predict other number-theoretic objects using trees?

Some questions

Question

Can we scale up the experiments? Using modern ML/DL models (e.g., transformers)? What would they learn?

Question

Can we predict other number-theoretic objects using trees?

(In progress, with Banwait, Huang, Lee, Oliver, Podznyakov)

Predict the minimal Weierstrass equation of an elliptic curve from Frobenius traces.

Some questions

Question

Can we scale up the experiments? Using modern ML/DL models (e.g., transformers)? What would they learn?

Question

Can we predict other number-theoretic objects using trees?

(In progress, with Banwait, Huang, Lee, Oliver, Podznyakov)

Predict the minimal Weierstrass equation of an elliptic curve from Frobenius traces.

Spoiler: the decision tree does a great job again!

Thank you!



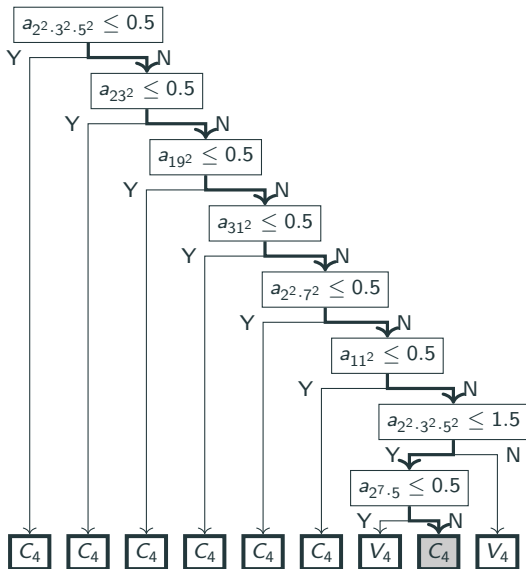
(a) paper



(b) code

Appendix

Tree for degree 4



Euler factors and zeta coefficients for degree ℓ^2

(e, f, g)	Euler factor	$a_p(K)$	$a_{p^\ell}(K)$
$(1, 1, \ell^2)$	$(1 - p^{-s})^{-\ell^2} = \sum_{k \geq 0} \binom{\ell^2+k-1}{\ell^2-1} p^{-ks}$	ℓ^2	$\binom{\ell^2+\ell-1}{\ell^2-1}$
$(1, \ell, \ell)$	$(1 - p^{-\ell s})^{-\ell} = \sum_{k \geq 0} \binom{\ell+k-1}{\ell-1} p^{-\ell ks}$	0	ℓ
$(1, \ell^2, 1)$	$(1 - p^{-\ell^2 s})^{-1} = \sum_{k \geq 0} p^{-\ell^2 ks}$	0	0
$(\ell, 1, \ell)$	$(1 - p^{-s})^{-\ell} = \sum_{k \geq 0} \binom{\ell+k-1}{\ell-1} p^{-ks}$	ℓ	$\binom{2\ell-1}{\ell-1}$
$(\ell, \ell, 1)$	$(1 - p^{-\ell s})^{-1} = \sum_{k \geq 0} p^{-\ell ks}$	0	1
$(\ell^2, 1, 1)$	$(1 - p^{-s})^{-1} = \sum_{k \geq 0} p^{-ks}$	1	1

Euler factors and zeta coefficients for degree 6

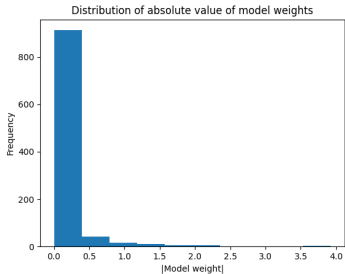
(e, f, g)	Euler factor	$a_p(K)$	$a_{p^2}(K)$	$a_{p^3}(K)$
$(1, 1, 6)$	$(1 - p^{-s})^{-6} = \sum_{k \geq 0} \binom{k+5}{5} p^{-ks}$	6	21	56
$(1, 2, 3)$	$(1 - p^{-2s})^{-3} = \sum_{k \geq 0} \binom{k+2}{2} p^{-2ks}$	0	3	0
$(1, 3, 2)$	$(1 - p^{-3s})^{-2} = \sum_{k \geq 0} (k+1) p^{-3ks}$	0	0	2
$(1, 6, 1)$	$(1 - p^{-6s})^{-1} = \sum_{k \geq 0} p^{-6ks}$	0	0	0
$(2, 1, 3)$	$(1 - p^{-s})^{-3} = \sum_{k \geq 0} \binom{k+2}{2} p^{-ks}$	3	6	10
$(2, 3, 1)$	$(1 - p^{-3s})^{-1} = \sum_{k \geq 0} p^{-3ks}$	0	0	1
$(3, 1, 2)$	$(1 - p^{-s})^{-2} = \sum_{k \geq 0} (k+1) p^{-ks}$	2	3	4
$(3, 2, 1)$	$(1 - p^{-2s})^{-1} = \sum_{k \geq 0} p^{-2ks}$	0	1	0
$(6, 1, 1)$	$(1 - p^{-s})^{-1} = \sum_{k \geq 0} p^{-ks}$	1	1	1

Logistic regression, degree 6

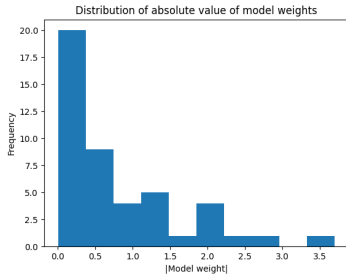
top	4	16	125	36	25	108	100	52	200	512
bottom	50	1	5	54	8	29	23	31	17	13

Table 1: Indices of the 10 largest and smallest weights among 1000 logistic regression coefficients.

Logistic regression, degree 6



The whole set of weights



Indices $n = \prod_i p_i^{e_i}$ with e_i divisible by 2 or 3

Conditional probabilities for degree 8 nonabelian extensions

	a_{2^2}		a_{3^2}		a_{5^2}		a_{7^2}		a_{11^2}		a_{13^2}	
G	D_4	Q_8	D_4	Q_8	D_4	Q_8	D_4	Q_8	D_4	Q_8	D_4	Q_8
0	0.21	0.79	0.28	0.72	0.28	0.72	0.24	0.76	0.25	0.75	0.25	0.75
1	0.27	0.73	0.10	0.90	-	-	0.05	0.95	0.03	0.97	-	-
2	0.62	0.38	0.49	0.51	0.45	0.55	0.41	0.59	0.35	0.65	0.33	0.67
3	0.27	0.73	-	-	0.07	0.93	-	-	-	-	0.03	0.97
4	0.77	0.23	0.86	0.14	0.83	0.17	0.84	0.16	0.81	0.19	0.82	0.18
10	0.80	0.20	0.68	0.32	0.66	0.34	0.64	0.36	0.52	0.48	0.54	0.46
36	0.29	0.71	0.38	0.62	0.38	0.62	0.36	0.64	0.36	0.64	0.41	0.59

Table 2: $\mathbb{P}[\text{Gal}(K/\mathbb{Q}) \simeq G | a_{p^2}(K) = a]$ for $G \in \{D_4, Q_8\}$ among degree 8 nonabelian Galois number fields.