ML for number theory

Team NT

December 4, 2025

Goal

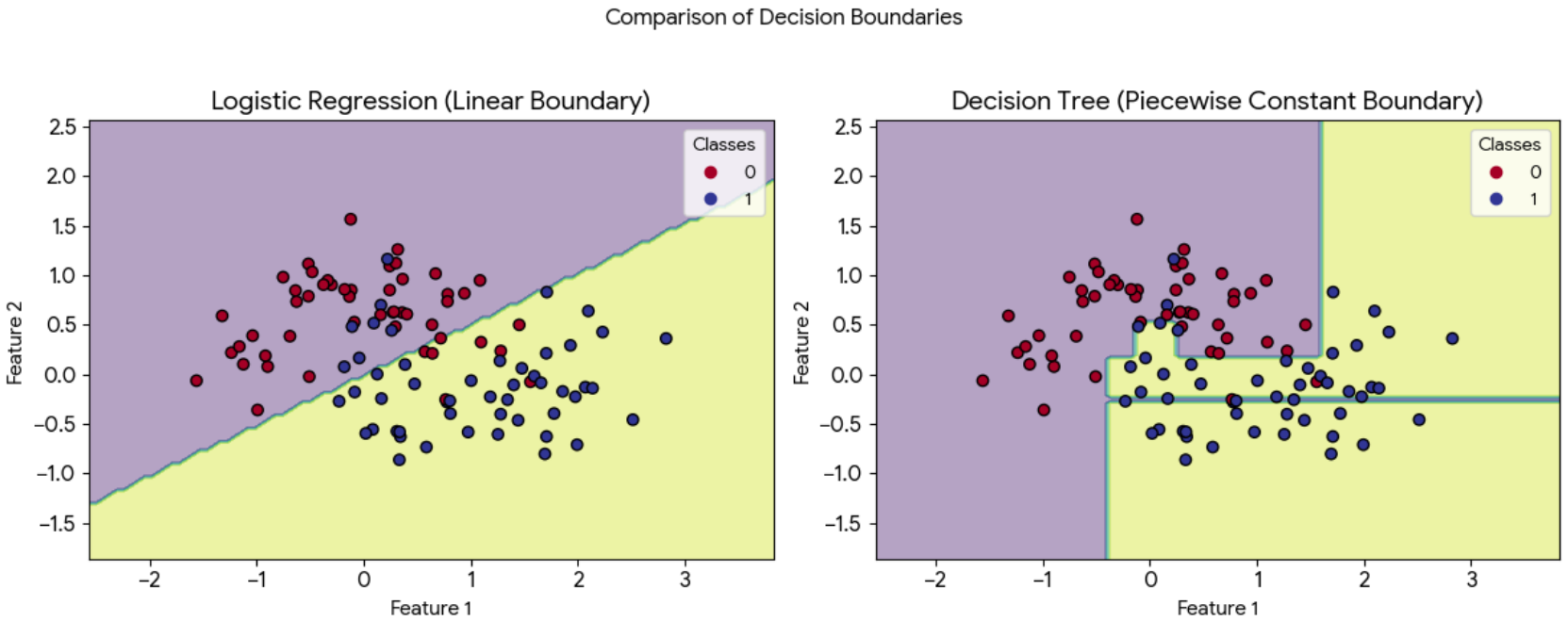
Use ML to solve classification problems in number theory.

Machine Learning

Classical ML models:

- Logistic regression: learn the best spearating hyperplane
- Decision tree: learn the best sequence of `if-else` rules, based on the size of features.

Logistic Regression vs Decision Tree



Data - LMFDB

LMFDB

⌂ → Artin representations → 3.1729.9t20.c.a

Citation · Feedback · Hide Menu

Artin representation 3.1729.9t20.c.a

Introduction
Overview · Random
Universe · Knowledge

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

Basic invariants

Dimension:	3
Group:	$C_3 \wr S_3$
Conductor:	$1729 = 7 \cdot 13 \cdot 19$
Artin stem field:	Galois closure of 9.3.738391927.1
Galois orbit size:	2
Smallest permutation container:	$C_3 \wr S_3$
Parity:	odd
Determinant:	1.1729.6t1.a.b
Projective image:	$C_3^2 : C_6$
Projective stem field:	Galois closure of 9.1.8936757492481.1

Defining polynomial

$$f(x) = x^9 - 2x^7 + 2x^5 - 5x^4 + 2x^3 + 3x^2 - 3x + 1 \quad .$$

The roots of f are computed in an extension of $\mathbb{Q}_{71}$ to precision 10.

Minimal polynomial of a generator a of K over $\mathbb{Q}_{71}$: $x^3 + 4x + 64$

Roots:

$$r_1 = 22a^2 + 52a + 43 + (a^2 + 39a + 43) \cdot 71 + (56a^2 + a + 13) \cdot 71^2 + (a^2 + 68a + 24) \cdot 71^3 + (38a^2 + 59a + 41) \cdot 71^4 + (33a^2 + 54a) \cdot 71^5 + (24a^2 + 50a + 68) \cdot 71^6 + (66a^2 + 43a + 53) \cdot 71^7 + (23a^2 + 12a + 13) \cdot 71^8 + (4a^2 + 17a + 2) \cdot 71^9 + O(71^{10})$$

$$r_2 = 70a^2 + 55a + 29 + (20a^2 + 36a + 1) \cdot 71 + (54a^2 + 34a + 9) \cdot 71^2 + (20a^2 + 49a + 51) \cdot 71^3 + (62a^2 + 25a + 58) \cdot 71^4 + (31a^2 + 32a + 19) \cdot 71^5 + (10a^2 + 48a + 54) \cdot 71^6 + (31a^2 + 15a + 54) \cdot 71^7 + (25a^2 + 30a + 17) \cdot 71^8 + (32a^2 + 43a + 53) \cdot 71^9 + O(71^{10})$$

$$r_3 = 31a^2 + 53a + 5 + (27a^2 + 48a + 47) \cdot 71 + (31a^2 + 41a + 55) \cdot 71^2 + (14a^2 + 45a + 60) \cdot 71^3 + (12a^2 + 40a + 66) \cdot 71^4 + (32a^2 + 26a + 2) \cdot 71^5 + (15a^2 + 47a + 33) \cdot 71^6 + (44a^2 + 36a + 41) \cdot 71^7 + (26a^2 + 32a + 41) \cdot 71^8 + (37a^2 + 66a + 62) \cdot 71^9 + O(71^{10})$$

$$r_4 = 50a^2 + 35a + 23 + (48a^2 + 65a + 4) \cdot 71 + (31a^2 + 34a + 67) \cdot 71^2 + (48a^2 + 24a + 53) \cdot 71^3 + (41a^2 + 56a + 3) \cdot 71^4 + (5a^2 + 54a + 68) \cdot 71^5 + (36a^2 + 42a + 27) \cdot 71^6 + (44a^2 + 11a + 19) \cdot 71^7 + (21a^2 + 28a + 31) \cdot 71^8 + (34a^2 + 10a + 58) \cdot 71^9 + O(71^{10})$$

$$r_5 = 21 + 4 \cdot 71 + 30 \cdot 71^2 + 30 \cdot 71^3 + 20 \cdot 71^4 + 48 \cdot 71^5 + 44 \cdot 71^6 + 52 \cdot 71^7 + 67 \cdot 71^8 + 69 \cdot 71^9 + O(71^{10})$$

$$r_6 = 54a^2 + 9a + 19 + (42a^2 + 64a + 64) \cdot 71 + (33a^2 + 15a + 37) \cdot 71^2 + (21a^2 + 47a + 8) \cdot 71^3 + (29a^2 + 40a + 65) \cdot 71^4 + (58a^2 + 70a + 1) \cdot 71^5 + (18a^2 + 5a + 42) \cdot 71^6 + (37a^2 + 31a + 46) \cdot 71^7 + (57a^2 + 22a + 5) \cdot 71^8 + (a^2 + 48a + 15) \cdot 71^9 + O(71^{10})$$

$$r_7 = 6 + 69 \cdot 71 + 29 \cdot 71^3 + 7 \cdot 71^4 + 7 \cdot 71^5 + 14 \cdot 71^6 + 45 \cdot 71^7 + 20 \cdot 71^8 + 20 \cdot 71^9 + O(71^{10})$$

$$r_8 = 40 + 26 \cdot 71 + 34 \cdot 71^2 + 28 \cdot 71^3 + 49 \cdot 71^4 + 33 \cdot 71^5 + 28 \cdot 71^6 + 3 \cdot 71^7 + 8 \cdot 71^8 + 49 \cdot 71^9 + O(71^{10})$$

$$r_9 = 57a^2 + 9a + 27 + (29a + 23) \cdot 71 + (6a^2 + 13a + 35) \cdot 71^2 + (35a^2 + 49a + 68) \cdot 71^3 + (29a^2 + 60a + 41) \cdot 71^4 + (51a^2 + 44a + 30) \cdot 71^5 + (36a^2 + 17a + 42) \cdot 71^6 + (60a^2 + 3a + 37) \cdot 71^7 + (57a^2 + 16a + 6) \cdot 71^8 + (31a^2 + 27a + 24) \cdot 71^9 + O(71^{10})$$

Properties

Label	3.1729.9t20.c.a
Dimension	3
Group	$C_3 \wr S_3$
Conductor	1729
Root number	not computed
Indicator	0

Related objects
Field 9.3.738391927.1
Field 9.1.8936757492481.1
Determinant 1.1729.6t1.a.b
Galois orbit 3.1729.9t20.c

Downloads
Underlying data

Learn more
Source and acknowledgments
Completeness of the data
Reliability of the data
Artin representations labels

Data - LMFDB

LMFDB

⌂ → Abelian varieties → Fq → 3 → 2 → ad_d_ac

Citation · Feedback · Hide Menu

Abelian variety isogeny class 3.2.ad_d_ac over $\mathbb{F}_2$

Introduction
Overview · Random
Universe · Knowledge

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

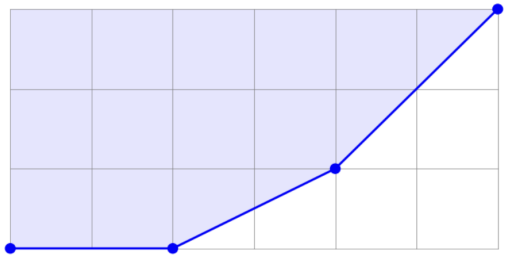
Database

Invariants

Base field:	$\mathbb{F}_2$
Dimension:	3
L-polynomial:	$(1 - 2x + 2x^2)(1 - x - x^2 - 2x^3 + 4x^4)$
	$1 - 3x + 3x^2 - 2x^3 + 6x^4 - 12x^5 + 8x^6$
Frobenius angles:	± 0.0516399385854, ± 0.250000000000, ± 0.718306605252
Angle rank:	1 (numerical)
Jacobians:	1
Cyclic group of points:	yes

This isogeny class is not simple, primitive, not ordinary, and not supersingular. It is principally polarizable and contains a Jacobian.

Newton polygon



p-rank: 2
Slopes: $[0, 0, 1/2, 1/2, 1, 1]$

Point counts

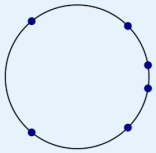
Point counts of the abelian variety

r	1	2	3	4	5
$A(\mathbb{F}_{q^r})$	1	35	208	6475	30791

Point counts of the curve

Properties

Label 3.2.ad_d_ac



Base field	$\mathbb{F}_2$
Dimension	3
p-rank	2
Ordinary	no
Supersingular	no
Simple	no
Geometrically simple	no
Primitive	yes
Principally polarizable	yes
Contains a Jacobian	yes

Related objects
L-functions

Downloads
All stored data to text
Curves to text
Underlying data

Learn more
Source and acknowledgments
Completeness of the data
Reliability of the data
Labeling convention

Artin Representation

Artin Representation

Definition

Artin representation is a finite dimensional complex representation of a Galois group, i.e. a group homomorphism

$$\rho : \mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$$

For each (unramified) prime p, there is a conjugacy class $\mathrm{Frob}_p \subset \mathrm{Gal}(K/\mathbb{Q})$ called the **Frobenius conjugacy class** at p. We denote by a_p for the trace of $\rho(\mathrm{Frob}_p)$ ($a_p = 0$ if p is ramified).

Task

a_p are important, e.g. they define the Artin L-function $L(\rho, s)$.

Question

Can we predict other invariants of ρ with a_p's?

Candidates:

- Parity
- Frobenius–Schur indicator
- **Projective image** (image in $\mathrm{PGL}_n(\mathbb{C})$)

3-dim representations

- Total $\approx 30,000$
- S_4 vs A_4.
- Logistic regression
- Decision tree

3-dim representation, further data analysis

Conjecture

Let $\rho : \mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{GL}_3(\mathbb{C})$ *be an irreducible Artin representation with projective image isomorphic to* A_4*. If* $a_p = -3$ *for some* p*, then* $a_{p'} = 1$ *for some* p'*.*

True for all data in LMFDB.

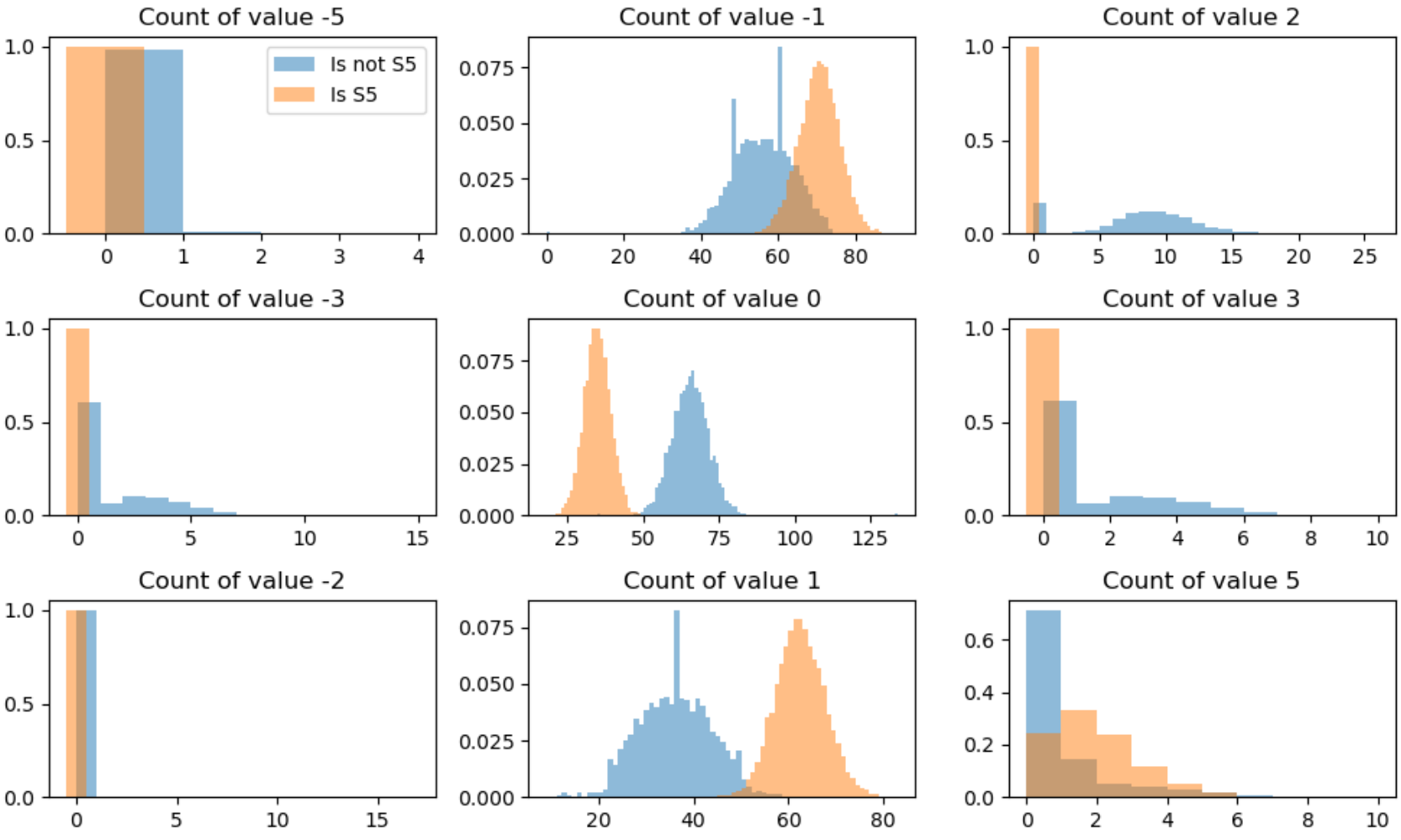
5-dim representations

- Total $\approx 20,000$
- Majority of them are S_5, A_6, S_6
- Logistic regression, one-hot encoding: 99.5% accuracy.
- Decision tree: $> 92\%$ accuracy.

5-dim representations, Logistic regression

- S_5 vs rest
- One-hot encoding: for each p and $a \in \{-5, -3, -2, -1, 0, 1, 2, 3, 5\}$, we create a feature which is 1 if $a_p = a$ and 0 otherwise.
- Number of p's with $a_p = 0$ is important, which can be explained by Chebotarev density theorem.

5-dim representations, a_p statistics



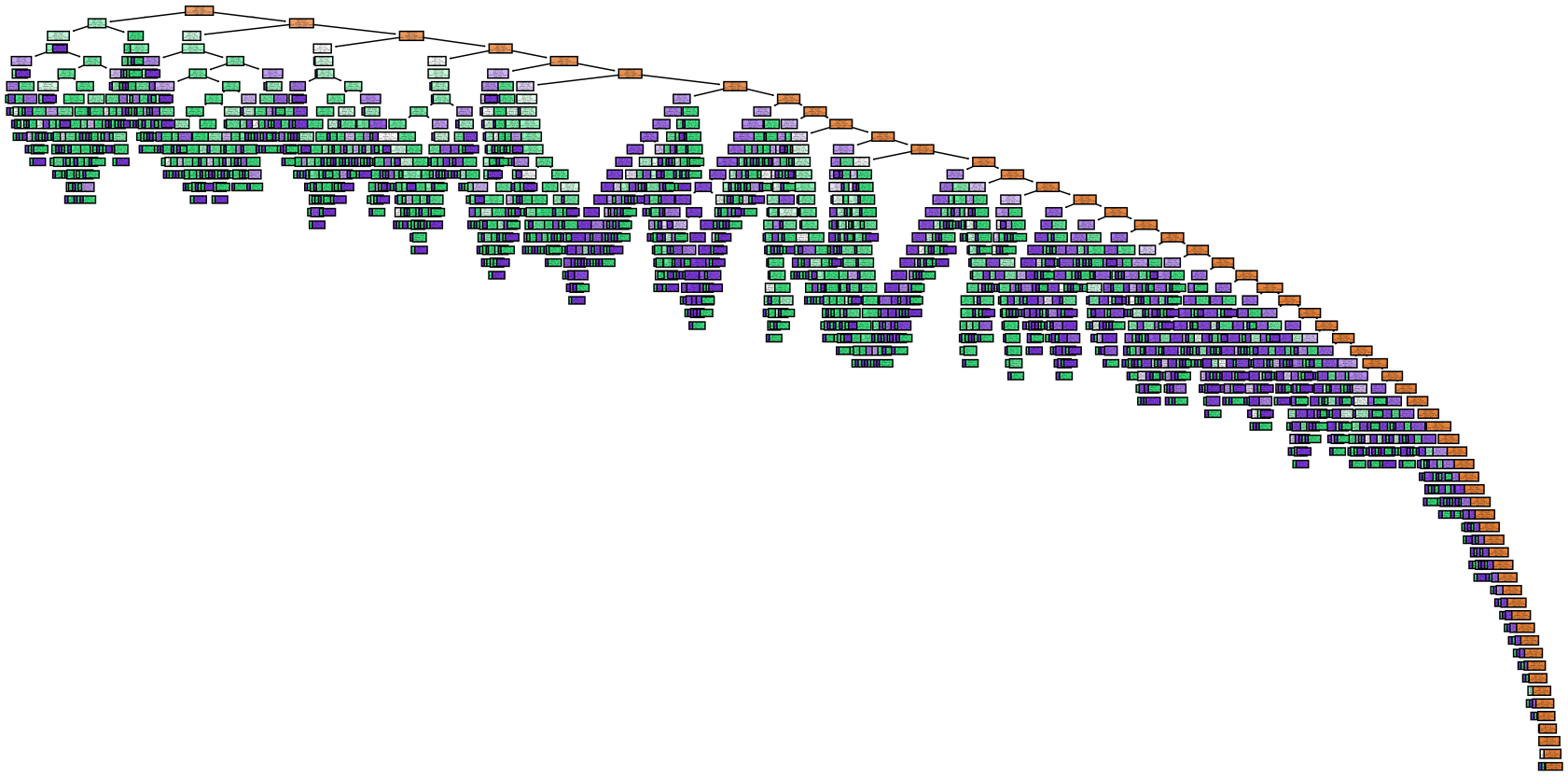
5-dim representations, Decision tree

- S_6 vs rest
- Existence of p with $a_p = 2$ is important, which can be explained by checking the character table.

2-dim odd representations, image in $\mathrm{GL}_2(\mathbb{C})$

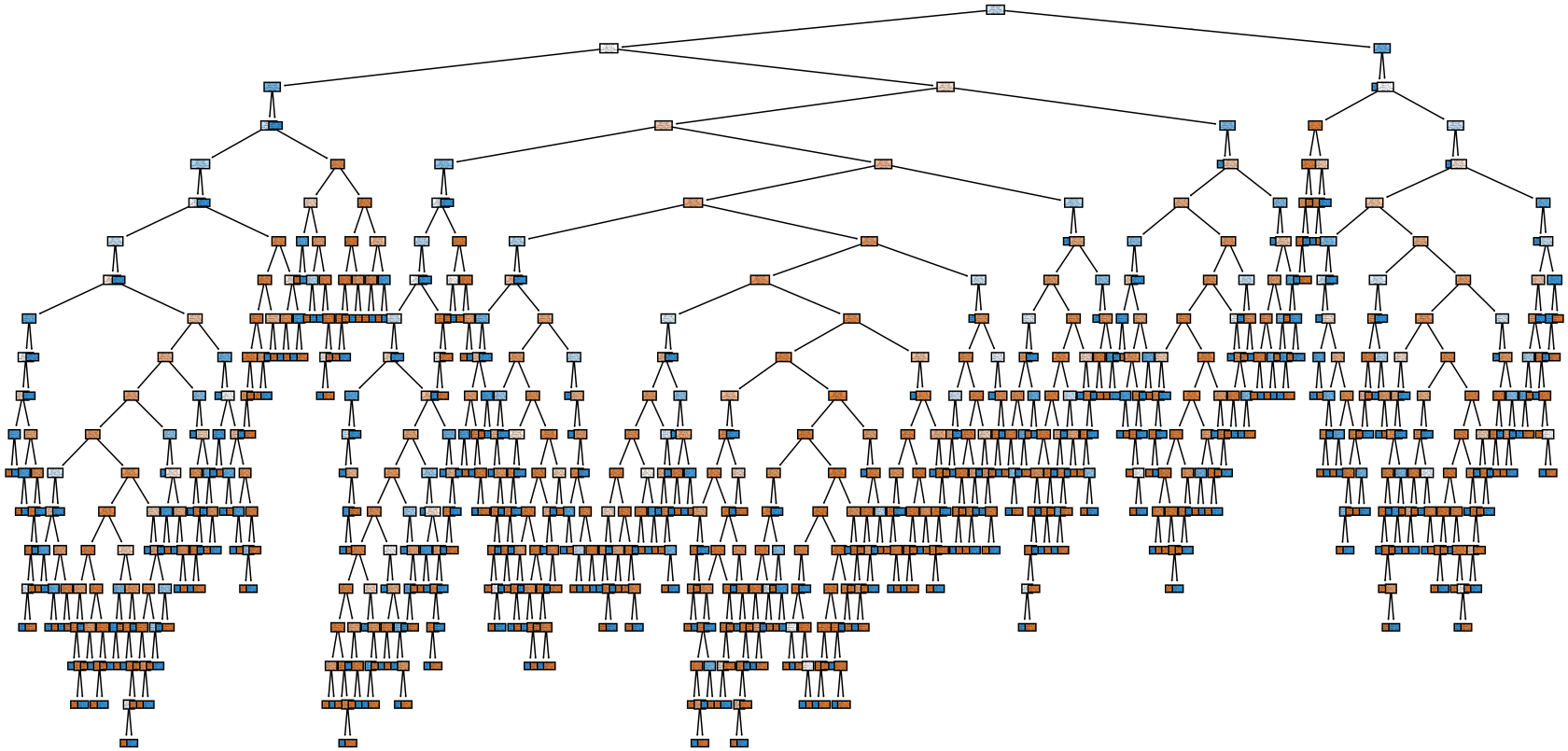
- Total $\approx 220,000$
- S_3 vs D_4 vs D_6
- Decision tree: $> 99\%$ accuracy.

Tree (S_3 vs D_4 vs D_6)



If $a_p = -2$ for some p, then the image is *not* S_3. Can be explained using character table.

Tree (D_4 vs D_6)



Future plans

- More interpretation of decision trees for 2-dim and 5-dim cases.
- Prove conjecture for 3-dim case.

Abelian Varieties over finite fields

Abelian Varieties

Definition

Abelian variety is a **variety** with an **abelian group** structure.

e.g. Elliptic curves (1-dim abelian varieties). $y^2 = x^3 + ax + b$.

L-polynomial

Definition

L-polynomial of an abelian variety $A_{/\mathbb{F}_q}$ is defined as

$$L_A(x) = \det(1 - x\mathrm{Frob}_q^{-1} | \mathrm{H}^1_{\text{ét}}(A_{\overline{\mathbb{F}}_q}; \overline{\mathbb{Q}}_\ell))$$

which is an integer polynomial of degree $2g$ where $g = \dim A$.

It has a form of

$$L_A(x) = 1 + a_1 x + \cdots + a_{2g-1} x^{2g-1} + q^g t^{2g}$$

and $a_{g+1}, a_{g+2}, \ldots, a_{2g}$ are multiples of p.

(Deligne) Zeros of $L_A(t)$ have an absolute value $q^{-1/2}$.

(Honda–Tate) Isogeny classes $\leftrightarrow$ L-polynomials.

Task

Question

Can we predict other invariants of A with a_i's?

Candidates:

- **p-rank**
- Principally polarizable
- **Jacobian**

p-rank

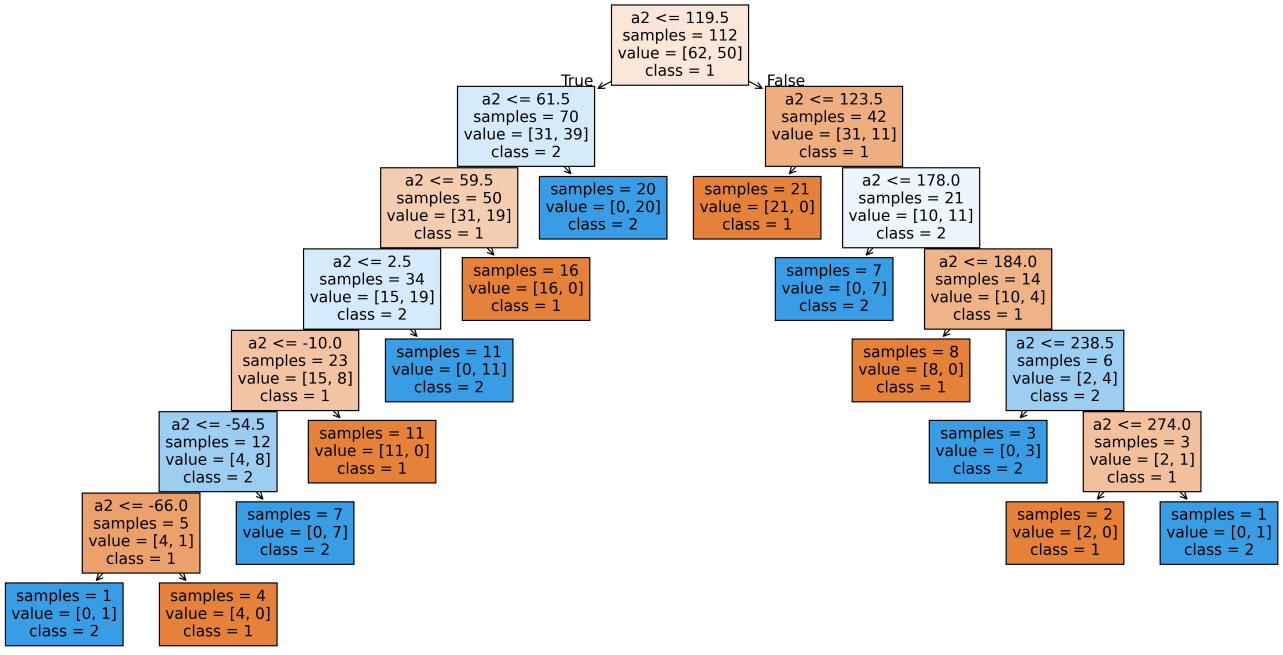
Definition

$$\mathrm{rank}_p(A) = \dim_{\mathbb{F}_p}(A(\overline{\mathbb{F}}_p)[p])$$

Possible values: $0, 1, \ldots, g = \dim A$.

When $g = 2$, there are very few rank 0 cases, so we focus on rank 1 vs rank 2 classification.

Tree



- Most of the nodes sees a_2. (a_1 and a_3 are not important.)
- When $a_2 = 61k + r$ with *small* $|r|$, then the p-rank is 1.

Why?

In fact, p-rank is determined by the *Newton polygon* of $L_A(t)$: the lower convex hull of the points

$$(0,0),\ (1, v_p(a_1)),\ \ldots,\ (2g-1, v_p(a_{2g-1})),\ (2g, g)$$

The rank is equal to the number of segments of slope 0.

When $g = 2$, $v_p(a_2) = 0 \Leftrightarrow a_2 \not\equiv 0 \pmod{p}$ implies p-rank 2.

Jacobian

For any genus g curve $C_{/\mathbb{F}_q}$, its Jacobian variety J_C is a g-dim abelian variety.

Question

For a given isogeny class of abelian variety, does it contain a Jacobian variety? Can we predict it from $L_A(t)$?

(Howe–Nart–Ritzenthaler) A complete answer for $g = 2$.

For $g \geq 3$, some entries in the LMFDB marked as *unknown*.

$g = 2$, split

p-rank of $\mathcal{A}$	Condition on p and q	Conditions on s and t
—	—	$\lvert s-t\rvert = 1$
2	—	$s = t$ and $t^2 - 4q \in \{-3, -4, -7\}$
	$q = 2$	$\lvert s\rvert = \lvert t\rvert = 1$ and $s \neq t$
1	q square	$s^2 = 4q$ and $s - t$ squarefree
0	$p > 3$	$s^2 \neq t^2$
	$p = 3$ and q nonsquare	$s^2 = t^2 = 3q$
	$p = 3$ and q square	$s - t$ is not divisible by $3\sqrt{q}$
	$p = 2$	$s^2 - t^2$ is not divisible by $2q$
	$q = 2$ or $q = 3$	$s = t$
	$q = 4$ or $q = 9$	$s^2 = t^2 = 4q$

Table 1.1. Conditions that ensure that the split isogeny class with Weil polynomial $(x^2 - sx + q)(x^2 - tx + q)$ does not contain a Jacobian. Here we assume that $\lvert s\rvert \geqslant \lvert t\rvert$.

Weil polynomial $= x^{2g} L_A(1/x)$

$g = 2$, split

Proposition (Howe–Nart–Ritzenhaler)

For non-square q and $p \neq 2, 3$, if L_A factors as $L_A(x) = (1 + sx + qx^2)(1 + tx + qx^2)$ and $|s - t| > 1$, then the isogeny class contains a Jacobian.

$s - t = \pm 1$ would work as a decision boundary. Decision tree cannot easily learn such "diagonal" boundary.

$g = 2$, split

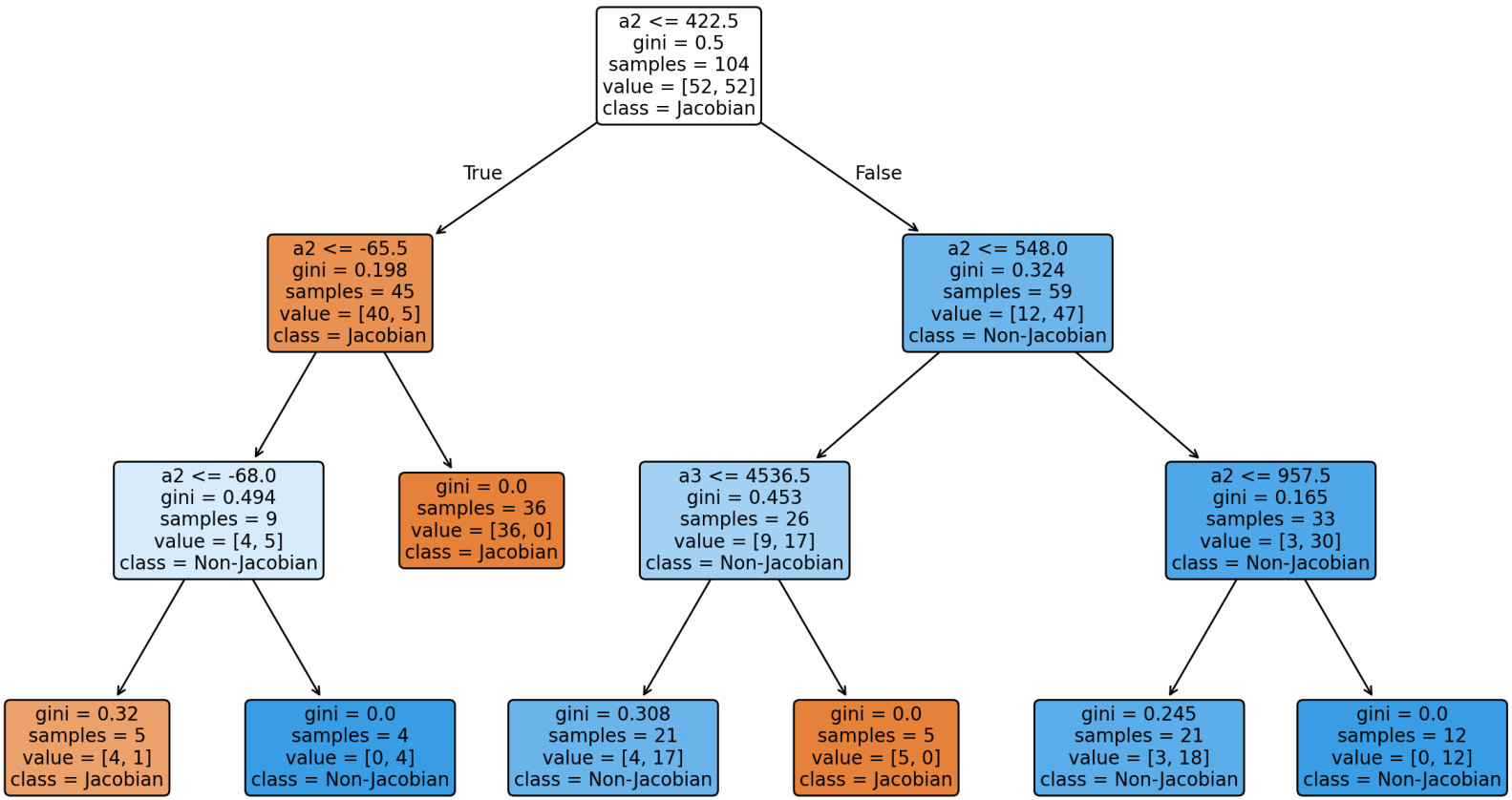
Proposition (Howe–Nart–Ritzenhaler)

For non-square q and $p \neq 2, 3$, if L_A factors as $L_A(x) = (1 + sx + qx^2)(1 + tx + qx^2)$ and $|s - t| > 1$, then the isogeny class contains a Jacobian.

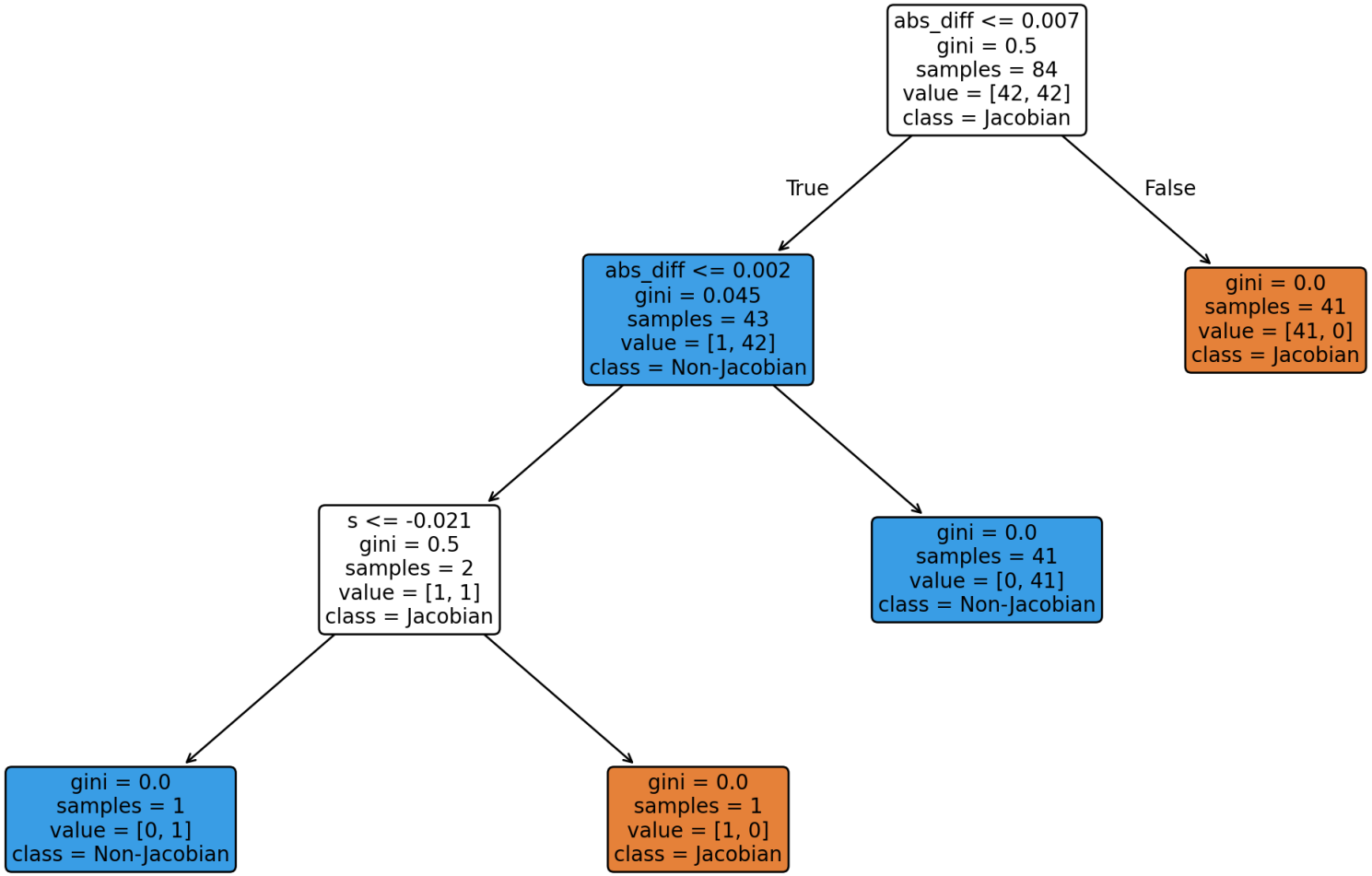
$s - t = \pm 1$ would work as a decision boundary. Decision tree cannot easily learn such "diagonal" boundary.

To mitigate this, we can simply add $|s - t|$ as a new feature.

Tree, without $|s - t|$ ($q = 211$, acc = 80.7%)



Tree, with $|s - t|$ ($q = 211$, acc = 100%)



$g = 3$, split

Question

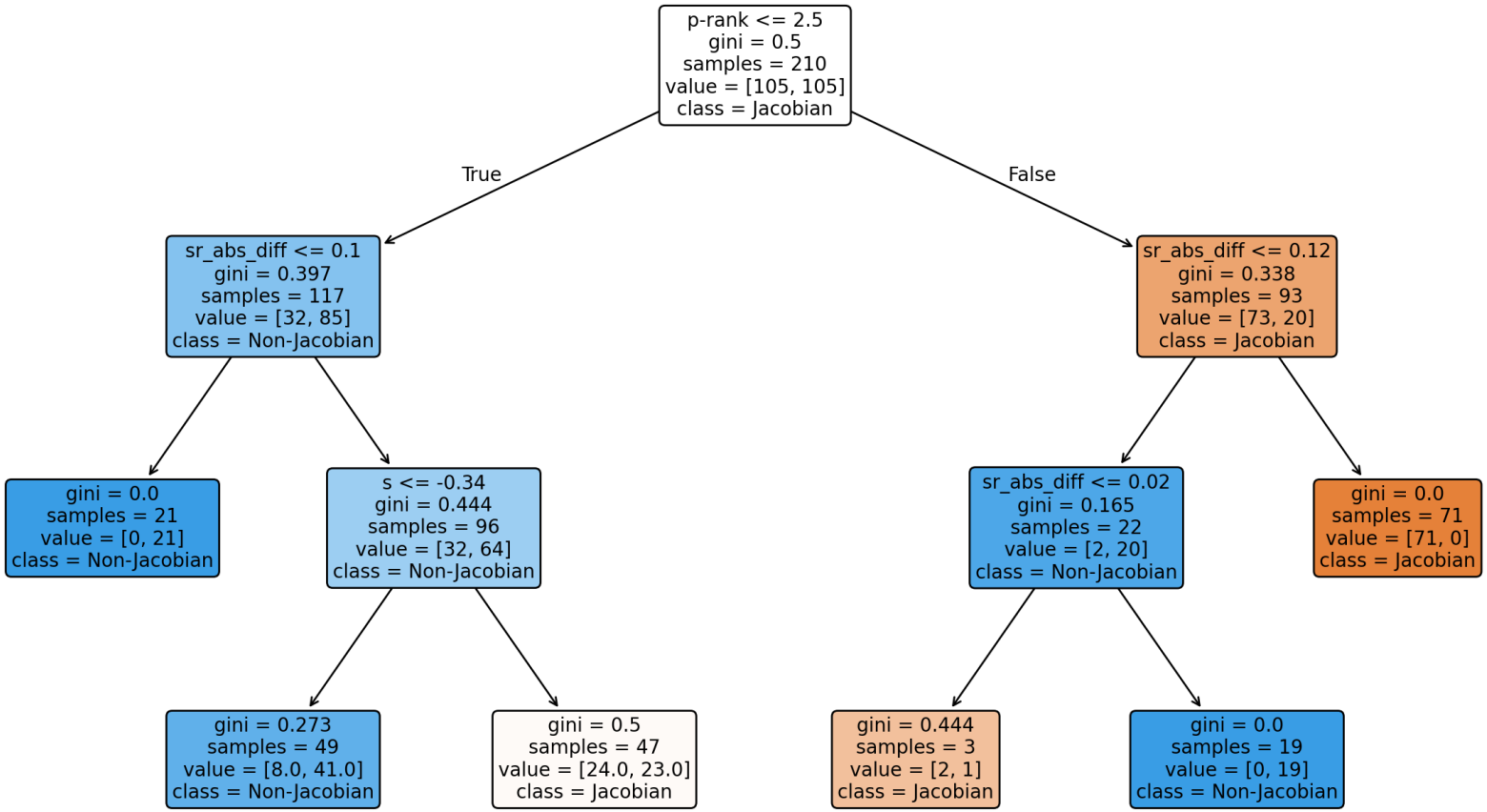
How about $g = 3$ split case? If

$$L_A(x) = (1 + sx + qx^2)(1 + tx + qx^2)(1 + rx + qx^2)$$

can we predict whether the isogeny class contains a Jacobian, using s, t, u?

We assume $s \leq t \leq r$. We will add $|s - t|, |t - r|, |r - s|$ as new features.

Tree ($q = 25$)



If $\text{rank}_p(A) = 3$ and $|r - s|$ is large, then the tree predicts that the isogeny class contains a Jacobian.

Future plans

- *Prove* the heuristics suggested from ML experiments and data for $g = 3$.
- Non-split cases?
- Data visualization for better understanding.
- $g \geq 4$?

Summary

- ML models tell us *rough* strategy to predict invariants. In particular, decision trees are good for
  - interpretability
  - when output depends on smallness/bigness of inputs
- Further investigation of data may lead to *precise* conjectures.
- If we are lucky, we can
  - prove the conjecture
  - come up with more efficient algorithms
  - filling missing entries in databases

Use ML for your own problem!