K.-H. Lee et al. *Res Math Sci(2026)13:60*
https://doi.org/10.1007/s40687-026-00620-3

Research in the Mathematical Sciences

RESEARCH

# Machines learn number fields, but how? The case of Galois groups

Kyu-Hwan Lee[1,2] and Seewoo Lee[3*]

*Correspondence:
seewoo5@berkeley.edu
[3]Department of Mathematics, University of California—Berkeley, Berkeley, CA 94720, USA
Full list of author information is available at the end of the article

**Abstract**

By applying interpretable machine learning methods such as decision trees, we study how simple models can classify the Galois groups of *Galois extensions* over $\mathbb{Q}$ of degrees *4, 6, 8, 9, and 10*, using Dedekind zeta coefficients. Our interpretation of the machine learning results allows us to understand how the distribution of zeta coefficients depends on the Galois group, and to prove new criteria for classifying the Galois groups of these extensions. Combined with previous results, this work provides another example of a new paradigm in mathematical research driven by machine learning.

## 1 Introduction

In recent years, machine learning (ML) has begun to influence the landscape of pure mathematics, echoing its transformative impact on the natural sciences. Around 2017, experimental efforts emerged to apply ML techniques to mathematical problems, revealing that machines can often detect structural patterns in mathematical data. Early examples arose in algebraic geometry [13,24,35,42], followed by applications in number theory [25–28], representation theory [19,36], and knot theory [19].

Several of these studies reported remarkable empirical successes, achieving high accuracy in classifying mathematical objects according to their invariants. Interpreting what the machine learns can sometimes lead to new discoveries or perspectives on the underlying mathematical structures. The recent discovery of *murmuration* phenomena in arithmetic is a prototypical example [17,28]. Together, these developments suggest a new paradigm for mathematical research [19,20]: Generate datasets, apply ML tools, interpret the results, gain new perspectives, formulate conjectures, and prove them.

The purpose of this paper is to complete a full cycle of this paradigm in the study of Galois groups. In earlier work by the first-named author with He and Oliver [25], machine learning algorithms were applied to the classification of number fields by invariants such as Galois groups, class numbers, and unit ranks, using input features derived from the coefficients of defining polynomials or Dedekind zeta functions. While many of these classification tasks achieved remarkably high accuracy, the mathematical rationale behind the machine's decision-making remained opaque. One notable case was the class numbers

of real quadratic fields, where subsequent efforts to interpret the ML results revealed that the algorithms were, in effect, rediscovering classical results from Gauss's genus theory [7].

In this paper, we focus on Galois groups as the next case study. We apply interpretable machine learning methods, particularly decision trees, to the problem of predicting Galois groups of number fields. By analyzing the structure of the resulting models, we are able to reverse-engineer the machine's logic and extract precise mathematical statements which we formulate into new conjectures about the determination of Galois groups. Unlike the case of class numbers of real quadratic fields and many earlier examples, these new conjectures are mathematically significant, and we prove them by traditional means. As an example, we prove the following proposition, motivated by an analysis of the classification logic of a decision tree model (see Example 3.10 for the proof).

**Proposition** *Let $K/\mathbb{Q}$ be a nonic Galois extension, and let $a_n(K)$ be the $n$-th coefficient of the Dedekind zeta function $\zeta_K(s)$. If $a_{1000}(K) \leq 4$, then $K/\mathbb{Q}$ is a cyclic extension.*

We obtain many other similar and more general results (see Corollaries 3.3, 3.16, 3.21, 3.25 and 3.35). Although Galois groups can be determined in various ways, our results offer very simple criteria that could be formulated only after analyzing the ML results. Our work demonstrates how machine learning can serve not merely as a black-box classifier, but as a tool for generating conjectures and aiding in the discovery of rigorous mathematical theorems. We anticipate that this approach will continue to be adopted and will lead to many fruitful discoveries in the future.

After this introduction, the next section presents preliminaries on the Dedekind zeta function and the ramification of local fields. Section 3 contains the main results, along with a detailed explanation of the ML experiments. The corresponding code is available in the GitHub repository https://github.com/seewoo5/ML-NF, which can be used to reproduce all the experiments in this paper. The final section offers concluding remarks.

## 2 Preliminaries

### 2.1 Dedekind zeta function

Let $K/\mathbb{Q}$ be a number field and $\mathcal{O}_K$ be the ring of integer of $K$. For each nonzero ideal $\mathfrak{a} \subset \mathcal{O}_K$, let $\mathrm{N}\mathfrak{a} \coloneqq [\mathcal{O}_K : \mathfrak{a}]$ be its ideal norm. The Dedekind zeta function $\zeta_K(s)$ is defined by

$$\zeta_K(s) \coloneqq \sum_{0 \neq \mathfrak{a} \subset \mathcal{O}_K} \frac{1}{\mathrm{N}\mathfrak{a}^s} = \sum_{n \geq 1} \frac{a_n(K)}{n^s}, \tag{1}$$

where $a_n(K) = \#\{\mathfrak{a} \subset \mathcal{O} : \mathrm{N}\mathfrak{a} = n\}$ counts the number of ideals with norm $n$, which is always a nonnegative integer. It admits an Euler product

$$\zeta_K(s) = \prod_{\mathfrak{p}} (1 - \mathrm{N}\mathfrak{p}^{-s})^{-1} = \prod_{p} \prod_{\mathfrak{p} \mid p} (1 - \mathrm{N}\mathfrak{p}^{-s})^{-1},$$

where the first product is over the prime ideals of $\mathcal{O}_K$, and the $p$-th Euler factor for a rational prime $p$ is defined by

$$\zeta_{K,p}(s) \coloneqq \prod_{\mathfrak{p} \mid p} (1 - \mathrm{N}\mathfrak{p}^{-s})^{-1}. \tag{2}$$

The zeta coefficients are multiplicative; that is, we have $a_{mn}(K) = a_m(K)a_n(K)$ for any coprime $m, n$. Moreover, the Euler factor $\zeta_{K,p}(s)$ is completely determined by the decomposition behavior of $p$ in $K$, as stated in the following standard proposition, whose proof we omit.

**Proposition 2.1** *Let $K/\mathbb{Q}$ be a number field and $p$ be a rational prime. Assume that the prime $p$ factors in $K$ as $(p) = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g}$, and let $f_i = f(\mathfrak{p}_i|p)$ be the inertia degree of $\mathfrak{p}_i$ over $p$, so that $\mathrm{N}\mathfrak{p}_i = p^{f_i}$. Then, the Euler factor can be written as*

$$\zeta_{K,p}(s) = \prod_{i=1}^{g} (1 - p^{-f_i s})^{-1}. \tag{3}$$

*In particular, when $K/\mathbb{Q}$ is Galois, all $e_i$'s and $f_i$'s are equal for $i = 1, 2, \ldots, g$, and*

$$\zeta_{K,p}(s) = (1 - p^{-fs})^{-g} = \sum_{k \geq 0} \binom{g + k - 1}{g - 1} p^{-fks}. \tag{4}$$

**Corollary 2.2** *Let $K/\mathbb{Q}$ be a Galois extension and $p$ be a rational prime. Assume that $p$ decomposes in $K$ as $(p) = (\mathfrak{p}_1 \cdots \mathfrak{p}_g)^e$ with $\mathrm{N}\mathfrak{p}_i = p^f$ for all $1 \leq i \leq g$. Then, the $p^a$-th zeta coefficient for $a \in \mathbb{Z}_{\geq 0}$ is given by*

$$a_{p^a}(K) = \begin{cases} \binom{g + \frac{a}{f} - 1}{g - 1} & \text{if } f \mid a, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

*Proof* If $f \nmid a$, then $a_{p^a}(K) = 0$ since $p^{-as}$ does not appear in (4). If $f \mid a$, take $k = \frac{a}{f}$. □

### 2.2 Ramification of local fields

We review ramification theory of local fields, which will be useful for studying decomposition behavior of primes. Let $L/K$ be a finite Galois extension of local fields with $G = \mathrm{Gal}(L/K)$. Let $\mathcal{O}_K, \pi_K, v_K, k$ be the ring of integer, uniformizer, valuation, and residue field of $K$, and similarly $\mathcal{O}_L, \pi_L, v_L, \ell$. For $i \geq -1$, define the $i$-th ramification group to be

$$G_i \coloneqq \{\sigma \in G : v_L(\sigma(x) - x) \geq i + 1 \ \forall x \in \mathcal{O}_L\}. \tag{6}$$

Then, $\{G_i\}_{i \geq -1}$ form a decreasing filtration of $G$. The group $G_0$ (resp. $G_1$) is often called the inertia subgroup (resp. wild inertia subgroup), and we denote their fixed fields as $L^u = L^{G_0}$ and $L^t = L^{G_1}$. We have $G_0/G_1 \hookrightarrow \mathcal{O}_L^\times/(1+\pi_L\mathcal{O}_L) \simeq \ell^\times$ and $G_i/G_{i+1} \hookrightarrow (1+\pi_L^i\mathcal{O}_L)/(1+\pi_L^{i+1}\mathcal{O}_L) \simeq \ell$ for $i \geq 1$, where the first embedding is given by $\sigma \mapsto \sigma(\pi_L)/\pi_L$ which is independent of choice of the uniformizer $\pi_L$. When $K = K'_{\mathfrak{p}}$ and $L = L'_{\mathfrak{P}}$ are localizations of a number field $K'$ and its extension $L'$, we may also denote the decomposition group and inertia group as $D_{\mathfrak{P}}$ and $I_{\mathfrak{P}}$. When $L'/K'$ is abelian, the decomposition and inertia groups only depend on $\mathfrak{p}$, and we denote them as $D_{\mathfrak{p}}$ or $I_{\mathfrak{p}}$ instead.

Let $e = e_{L/K}$ be the ramification degree ($\pi_K = \varepsilon\pi_L^e$ for some $\varepsilon \in \mathcal{O}_L^\times$) and $f = f_{L/K} = [\ell : k]$ be the inertia degree. Then, $L^u/K$ is an unramified extension of degree $f$, and $L^t/L^u$ is a tamely ramified extension of degree equal to the prime-to-$p$ part of $e$, and $L/L^t$ is a wildly ramified extension of degree equal to the $p$-part of $e$. We say that the extension $L/K$ is tamely ramified if $L^t = L$, i.e., $G_1 = \{1\}$ or equivalently, $(p, e) = 1$. Particularly,

$G_0 \hookrightarrow \ell^\times$ and the inertia group has to be cyclic of order $e$. Moreover, we have the following proposition.

**Proposition 2.3** ( Greenberg[23, Lemma 1], Newton [38, Proposition 2.2.9]) *Let $L/K$ be a tamely ramified abelian extension of local fields, and $e = e_{L/K}$ be its ramification degree. Let $k$ be the residue field of $K$. Then, $e$ divides $\#k^\times$.*

For a tamely ramified extension of $K$, the conjugation action of (a lift of) Frobenius on the tame inertia group is simply given by $\#k$-power map:

**Proposition 2.4** ( Iwasawa[29]) *Let $L/K$ be a tamely ramified extension with ramification degree $e$. Let $\tau$ be a generator of the inertia subgroup $G_0 \leq G = \mathrm{Gal}(L/K)$ and $\sigma \in G$ be a lift of a Frobenius. Then, $\sigma\tau\sigma^{-1} = \tau^q$, where $q = \#k$ is the size of the residue field of $K$.*

*Remark 2.5* For the later proofs, we will only use the case when $K = \mathbb{Q}_p$ and $q = p$ (see Lemma 3.14 and 3.19).

One may ask whether a local field with a given decomposition type can be *globalized*; that is, whether an extension $K/\mathbb{Q}_p$ arises as the localization of a number field $K'/\mathbb{Q}$. One may even ask whether we can choose $K'$ to have the same Galois group, i.e., $\mathrm{Gal}(K'/\mathbb{Q}) \simeq \mathrm{Gal}(K/\mathbb{Q}_p)$. When such a $K'$ exists and $K' \simeq \mathbb{Q}(\alpha)$, $K \simeq \mathbb{Q}_p(\alpha)$ for a root $\alpha$ of some $f(x) \in \mathbb{Z}[x]$, we call the polynomial $f(x)$ a *Galois splitting model*. Carrillo [14] proposed several methods to find it, including Panayi's root finding algorithm [40, Algorithm 8.4].

However, such a field $K'$ does not always exist. Indeed, Wang provided a famous counterexample to the incorrect proof of "Grunwald's theorem" [45,46]. In particular, Wang proved the following:

**Proposition 2.6** ( Wang [45]) *Let $s \geq 3$ and $K'/\mathbb{Q}$ be a cyclic extension of order $2^s$. Then, the prime 2 either totally ramifies in $K'$ or has at least two distinct prime factors.*

For example, when $K/\mathbb{Q}_2$ is the unramified cyclic extension of order 8, there exists no cyclic extension $K'/\mathbb{Q}$ of degree 8 that localizes to $K$.

### 2.3 Number fields and $p$-adic fields in `LMFDB`

`LMFDB` [37] is an online database that contains a wide range of mathematical objects in number theory and arithmetic geometry, including $L$-functions, modular forms, elliptic curves, number fields, $p$-adic fields, and representations. In this work, we used the number field database for machine learning experiments and the local field database for the proofs of some theorems.

The number field database in `LMFDB` is a compilation of several works including [10,12,33,34,39,44], where the detailed references can be found in the *Source and Acknowledgements* page for the database [3]. Since our focus is on Galois groups, we restrict our experiments to Galois extensions of $\mathbb{Q}$ of degrees 4, 6, 8, 9 and 10, among all number fields in the database. The corresponding statistics are summarized in Table 1.

For a given number field, there are infinitely many choices of defining polynomials. `LMFDB` uses the *normalized (reduced)* polynomial where the sum of the squares of the absolute values of all complex roots of a polynomial is minimized. If there is more than one such polynomial, choose the polynomials whose discriminant have minimal absolute value. If there are still multiple polynomials, say $P = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n$, choose

**Table 1** Galois extensions of degree 4, 6, 8, 9, 10 in LMFDB [37]

| Degree | Galois group | Number of fields | |
|---|---|---|---|
| 4 | $C_4$ (4T1) | 21873 (11.96%) | 182860 |
| | $C_2^2$ (4T2) | 160987 (88.04%) | |
| 9 | $C_9$ (9T1) | 284 (22.43%) | 1266 |
| | $C_3^2$ (9T2) | 982 (77.57%) | |
| 6 | $C_6$ (6T1) | 21146 (13.45%) | 157168 |
| | $S_3$ (6T2) | 136022 (86.55%) | |
| 10 | $C_{10}$ (10T1) | 1854 (33.86%) | 5476 |
| | $D_5$ (10T2) | 3622 (66.14%) | |
| 8 | $C_8$ (8T1) | 6206 (5.42%) | 114576 |
| | $C_4 \times C_2$ (8T2) | 19550 (17.06%) | |
| | $C_2^3$ (8T3) | 10767 (9.40%) | |
| | $D_4$ (8T4) | 27796 (24.26%) | |
| | $Q_8$ (8T5) | 50257 (43.86%) | |

the one where the vector $S(P) := (|a_1|, a_1, \ldots, |a_n|, a_n)$ has smallest lexicographic order (see [1] for more details). Note that these features (sum of squares of zeros, coefficients, etc.) are not field invariants, and the models trained on such data highly depend on such normalizations. See Sect. 3.3.2, 3.4.2, and *Polynomial coefficients* paragraphs of 3.6.1, and 3.6.2 for the experiments.

The $p$-adic field database in LMFDB is primarily based on the computations of Jones and Roberts [30–32], along with several other sources for specific data columns. Key columns include the ramification degree, inertia degree, Galois group, (wild) inertia group, and Galois splitting model [14]. As of now, the database contains all degree $n$ extensions of $\mathbb{Q}_p$ for $p < 200$ and $n \leq 23$; see the *completeness* page for $p$-adic fields [2].

## 3 Galois groups of number fields

In this main section, we investigate how a Galois group can be determined from a finite number of coefficients of the Dedekind zeta function and also from polynomial coefficients. This question is motivated and guided by machine learning (ML) experiments, in which these coefficients are used as features of the dataset. The details of the ML experiments are presented alongside theorems and proofs. While Galois groups can be determined by various classical methods, it is somewhat surprising that this particular approach of using zeta coefficients has not been explored in the existing literature. It is also worth noting that our criteria could only be formulated after analyzing the outcomes of the ML experiments.

### 3.1 Experimental details

As previously mentioned, we used the number field database from the LMFDB [37]. The zeta coefficients are separately computed using the function `zeta_coefficients` of `SageMath` [43] to make dataset for the experiments. For each experiment, we report average and standard deviation of accuracy, balanced accuracy, and per-class precision, recall, F1 over 5 different random splits of ratio 80 to 20 for training and test sets. All machine learning experiments were conducted using the default setup of the `scikit-learn` library [41], with fixed random state and no hyperparameter tuning except for setting the maximum number of iterations in logistic regression, which is set to be 10000.

For logistic regression models, they are trained with LBFGS solver with $L^2$-regularization. For decision trees, we trained them with Gini criterion and we did not set the maximum depths, so that the nodes are expanded until all leaves are pure. We also use `best` splitter, which uses midpoints of features values as thresholds. See the official `scikit-learn` documentations [5,6] for more details about default configurations.

### 3.2 From experiments to theorems

We briefly describe the overall workflow underlying our machine learning experiments and the subsequent process of formulating and proving theorems.

In the decision tree experiments involving zeta coefficients, we observe that coefficients indexed by perfect powers consistently appear near the top of the trees. In the cases of quartic and nonic extensions, the resulting decision trees (Figures 1 and 2) are remarkably simple. This allows us to readily conjecture that the vanishing of certain zeta coefficients characterizes cyclic extensions. We are able to prove both this conjecture *and* its converse (see Corollary 3.3).

For sextic, octic, and decic extensions, the decision trees become more intricate; nevertheless, the prominence of coefficients with perfect power indices remains evident. Motivated by this pattern, we conduct further exploratory data analysis (EDA), which reveals that specific values of zeta coefficients at prime-power indices occur only when the extension has a particular Galois group and when the corresponding prime satisfies certain congruence conditions. We subsequently establish these phenomena rigorously using Propositions 2.3, 2.4, and Theorem 2.6, together with additional group-theoretic arguments.

### 3.3 Quartic and nonic fields

Let $K/\mathbb{Q}$ be a Galois extension of degree 4. There are two possible Galois groups: cyclic group $C_4 = \mathbb{Z}/4$ or the Klein group $V_4 = C_2^2 = \mathbb{Z}/2 \times \mathbb{Z}/2$. It is completely known how to determine the Galois group of a splitting field of a quartic polynomial $f(x) \in \mathbb{Q}[x]$ in terms of discriminants and resolvents [21, p. 613]. Particularly, for a quartic Galois extension $K$, one has

$$\mathrm{Gal}(K/\mathbb{Q}) \simeq \begin{cases} V_4 & \text{if } \Delta_K \text{ is a perfect square,} \\ C_4 & \text{otherwise.} \end{cases} \tag{7}$$

Similarly, if $K/\mathbb{Q}$ is a degree 9 Galois extension, then there are only two possible Galois groups: cyclic group $C_9 = \mathbb{Z}/9$ or $C_3^2 = \mathbb{Z}/3 \times \mathbb{Z}/3$. In general, only possible Galois groups for degree $\ell^2$-Galois extensions for prime $\ell$ are $C_{\ell^2}$ or $C_\ell^2$, since those are the only groups of order $\ell^2$ up to isomorphism. In the rest of this subsection, we will use zeta and polynomial coefficients, respectively, to distinguish between the two cases.

#### *3.3.1 Learning Galois groups from zeta coefficients*

**3.3.1.1 Decision tree** We found that decision tree models can achieve almost perfect accuracy (100% on 4 out of 5 runs) for distinguishing Galois groups of quartic fields using zeta coefficients $a_n(K)$ up to $n \leq 1000$. Moreover, the splitting of the first few decision nodes (including the root node) has a form of $a_{m^2}(K) = 0$[1], i.e., whether $a_{m^2}(K)$ is zero

[1] In the actual experiments, the conditions are of the form $a_{m^2}(K) \leq 0.5$, since we are using the `best` splitter which chooses midpoints of feature values as thresholds [4].

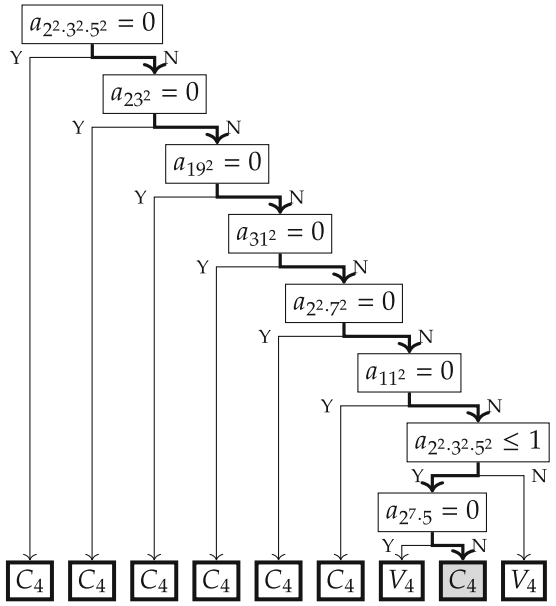



**Fig. 1** A decision tree predicts the Galois groups of quartic fields from zeta coefficients up to $a_{1000}(K)$

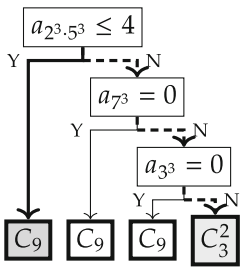



**Fig. 2** A decision tree predicts the Galois groups of nonic fields from zeta coefficients up to $a_{1000}(K)$

or not, and predicts as a cyclic extension when it is true (Figure 1). This strongly suggests the criterion: For a quartic Galois extension $K/\mathbb{Q}$, we have $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_4$ if $a_{m^2}(K) = 0$ for some $m \geq 2$.

Similarly, for degree 9 Galois extensions of $\mathbb{Q}$, decision tree models can achieve 100% accuracy on distinguishing Galois groups ($C_9$ or $C_3^2$) with zeta coefficients $a_n(K)$ up to $n \leq 1000$. More surprisingly, Fig. 2 shows that the model uses only three zeta coefficients $a_{1000}, a_{343}$ and $a_{27}$ out of 1000, and it strongly suggests that the prediction is based on the cube-indexed zeta coefficients. From these observations, it is natural to ask if one can distinguish two possible Galois groups of degree $\ell^2$ Galois extensions of $\mathbb{Q}$, using $\ell$-power coefficients. In particular, it is natural to conjecture the following: Let $K/\mathbb{Q}$ be a Galois extension of degree $\ell^2$. If $a_{p^\ell}(K) = 0$ for some prime $p$, then $K/\mathbb{Q}$ is a cyclic extension.

In Theorem 3.1, we show that the conjecture is true. More precisely, using Proposition 2.1 and Corollary 2.2, one can list all the possible decomposition types of a given prime $p$ in $K$ and the corresponding Euler factors and $p^\ell$-th zeta coefficients, as shown in Table 2. By proving that certain decomposition types cannot occur for $C_\ell^2$-extensions, we show that $p^\ell$-th zeta coefficients are nonvanishing for such extensions.

**Table 2** Decomposition types, Euler factors, and $a_p(K)$ and $a_{p^\ell}(K)$ for degree $\ell^2$ extensions

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

**Theorem 3.1** *Let $\ell$ be a prime and $K/\mathbb{Q}$ be a Galois extension with* $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_\ell^2$. *Then, $a_{p^\ell}(K) \neq 0$ for all prime $p$.*

*Proof* The last column of Table 2 shows that $a_{p^\ell}(K) = 0$ if and only if the decomposition type of $p$ is $(e,f,g) = (1, \ell^2, 1)$, i.e., totally inerts in $K$. Hence, it is enough to show that such decomposition type cannot occur for $C_\ell^2$-extensions. If it happens, we should have $I_p = 1$ and $D_p = C_\ell^2$, but

$$D_p \simeq D_p/I_p \simeq \mathrm{Gal}(\mathbb{F}_{p^{\ell^2}}/\mathbb{F}_p) \simeq C_{\ell^2},$$

which is a contradiction. □

In contrast, there are infinitely many primes $p$ for which $a_{p^\ell}(K) = 0$ when $K$ is a cyclic extension of degree $\ell^2$. We can prove the following general statement for all cyclic extensions of prime-power degree, which implies that the converse of Theorem 3.1 is also true.

**Theorem 3.2** *Let $\ell$ be a prime and $K/\mathbb{Q}$ be a cyclic extension of degree $\ell^m$ for some $m \geq 2$, i.e., $G = \mathrm{Gal}(K/\mathbb{Q}) \simeq C_{\ell^m}$. Then, for any $b \geq 1$ not divisible by $\ell^m$, the natural density of primes $p$ such that $a_{p^b}(K) = 0$ is positive. More precisely, if $\nu = \nu_\ell(b) < m$ is the $\ell$-adic valuation of $b$, then the density of primes with $a_{p^b}(K) = 0$ is $(\ell^{m-\nu}-1)/\ell^{m-\nu}$.*

*Proof* Consider unramified prime $p$ of decomposition type $(e,f,g) = (1, \ell^t, \ell^{m-t})$. The corresponding Euler factor is

$$\zeta_{K,p}(s) = (1-p^{-\ell^t s})^{-\ell^{m-t}} = \sum_{k\geq 0}\binom{\ell^{m-t}+k-1}{\ell^{m-t}-1}p^{-\ell^t ks}$$

, and this shows $a_{p^b}(K) = 0$ if and only if $t > \nu$, i.e., $\ell^t \nmid b$. For each $t > \nu$, $\mathrm{Frob}_p$ has order $\ell^t$ and there are $\varphi(\ell^t) = \ell^{t-1}(\ell-1)$-many elements in $C_{\ell^m}$ of order $\ell^t$, so the density of prime $p$ with $a_{p^b}(K) = 0$ is

$$\sum_{\nu<j\leq m}\frac{\ell^{j-1}(\ell-1)}{\ell^m} = \frac{\ell-1}{\ell^m}\frac{\ell^\nu(\ell^{m-\nu}-1)}{\ell-1} = \frac{\ell^{m-\nu}-1}{\ell^{m-\nu}}.$$

Note that the natural density and Dirichlet density coincides in Chebotarev density theorem. □

By combining two theorems, we obtain the following result suggested by the decision tree experiment.

**Corollary 3.3** *Let $\ell$ be a prime and $K/\mathbb{Q}$ be a Galois extension of degree $\ell^2$, hence* $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_{\ell^2}$ *or* $C_\ell^2$. *Then,* $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_{\ell^2}$ *if and only if there exists a prime $p$ such that $a_{p^\ell}(K) = 0$.*

In fact, for each Galois group and prime, we can analyze the possible decomposition types, as will be discussed below.

**Proposition 3.4** *Let $K/\mathbb{Q}$ be a quartic $C_4$-extension and $p \equiv 3 \pmod 4$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then, $(e,f,g) \neq (4,1,1)$.*

*Proof* Assume $(e,f,g) = (4,1,1)$. Since $(p,4) = 1$, the corresponding extension $K_\mathfrak{p}/\mathbb{Q}_p$ (for some $\mathfrak{p}|p$) is a tamely ramified extension, and hence, we should have $4 \mid p-1$ by Proposition 2.3. This cannot happen since $p \equiv 3 \pmod 4$. □

**Proposition 3.5** *Let $K/\mathbb{Q}$ be a quartic $C_2^2$-extension and $p$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then:*

1. $(e,f,g) \neq (1,4,1)$.
2. *If $p \neq 2$, then $(e,f,g) \neq (4,1,1)$.*

*Proof* 1) Since $C_4$ cannot be a quotient of $C_2^2$, we have $(e,f,g) \neq (1,4,1)$.

2) For odd $p$, the corresponding extension $K_\mathfrak{p}/\mathbb{Q}_p$ is tamely ramified and the inertia group has to be cyclic, so it cannot be totally ramified. □

*Example 3.6* ($C_4$, sufficient condition) Assume that a quartic Galois extension $K/\mathbb{Q}$ follows the bold branch in Fig. 1, so that the decision tree predicts the Galois group to be $C_4$. In this case, we have

$$a_{30^2} = a_{2^2} \cdot a_{3^2} \cdot a_{5^2} = 1 \quad \Leftrightarrow \quad a_{2^2} = a_{3^2} = a_{5^2} = 1,$$

and

$$a_{640} = a_{2^7} \cdot a_5 > 0 \quad \Rightarrow \quad a_5 \neq 0.$$

From Table 2, we can check that the only possible decomposition type of $p = 5$ satisfying $a_5 \neq 0$ and $a_{5^2} = 1$ is $(e,f,g) = (4,1,1)$, i.e., totally ramified. By Proposition 3.5, $K/\mathbb{Q}$ cannot be a $C_2^2$-extension and thus must be cyclic. This shows that the branch always gives correct prediction.

A similar analysis can be carried out for nonic extensions.

**Proposition 3.7** *Let $K/\mathbb{Q}$ be a nonic $C_9$-extension, and let $p \neq 3$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then:*

1. *If $p \equiv 4, 7 \pmod 9$, then $(e,f,g) \neq (9,1,1)$.*
2. *If $p \equiv 2 \pmod 3$, then $(e,f,g) \notin \{(3,1,3),(3,3,1),(9,1,1)\}$.*

*Proof* 1) Assume that $p \neq 3$ and $(e,f,g) = (9,1,1)$. Then, the corresponding extension $K_\mathfrak{p}/\mathbb{Q}_p$ (for $\mathfrak{p}|p$) is a totally and tamely ramified $C_9$-extension. By Proposition 2.3, this implies $p \equiv 1 \pmod 9$. Thus, $p \not\equiv 4, 7 \pmod 9$ and $p \not\equiv 2 \pmod 3$. In particular, this proves the first claim.

2) Now, assume $p \equiv 2 \pmod 3$. We already know $(e,f,g) \neq (9,1,1)$. Assume $(e,f,g) = (3,1,3)$. Then, the corresponding extension $K_{\mathfrak{p}}/\mathbb{Q}_p$ is a totally and tamely ramified $C_3$-extension. Then, Proposition 2.3 again shows $p \equiv 1 \pmod 3$, a contradiction. For the case of $(3,3,1)$, the corresponding extension $K_{\mathfrak{p}}/\mathbb{Q}_p$ has an unramified $C_3$-subextension $\tilde{K} \subset K_{\mathfrak{p}}$, where $K_{\mathfrak{p}}/\tilde{K}$ is totally and tamely ramified $C_3$-extension. Since the residue field of $\tilde{K}$ is $\mathbb{F}_{p^3}$, Proposition 2.3 implies that $p^3 \equiv 1 \pmod 3$, which contradicts to $p \equiv 2 \pmod 3$. □

**Proposition 3.8** *Let $K/\mathbb{Q}$ be a nonic $C_3^2$-extension, and let $p \neq 3$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then:*

1. *If $p \equiv 1 \pmod 3$, then $(e,f,g) \notin \{(1,9,1),(9,1,1)\}$*
2. *If $p \equiv 2 \pmod 3$, then $(e,f,g) \notin \{(1,9,1),(3,1,3),(3,3,1),(9,1,1)\}$*

*Proof* The decomposition type $(e,f,g) = (1,9,1)$ (resp. $(9,1,1)$) cannot occur since $C_9$ is not a quotient (resp. subgroup) of $C_3^2$. Now assume $p \equiv 2 \pmod 3$. If $(e,f,g) = (3,1,3)$, we have $(p) = \mathfrak{p}_1^3\mathfrak{p}_2^3\mathfrak{p}_3^3$, and a corresponding extension of local fields $K_{\mathfrak{p}}/\mathbb{Q}_p$ (for any $\mathfrak{p} = \mathfrak{p}_i$) is a totally and tamely ramified $C_3$-extension. By Proposition 2.3, we have $3 \mid p-1$, a contradiction. If $(e,f,g) = (3,3,1)$, we have $(p) = \mathfrak{p}^3$ with $f(\mathfrak{p}|p) = 3$. Let $\tilde{K} = K^{I_{\mathfrak{p}}}$ be the fixed field of the inertia group $I_{\mathfrak{p}}$, which is an unramified $C_3$-extension of $\mathbb{Q}_p$. Then, its residue field is $\mathbb{F}_{p^3}$, and $K/\tilde{K}$ is a totally and tamely ramified $C_3$-extension. By Proposition 2.3 again, we get $3 \mid p^3 - 1$, which contradicts to $p \equiv 2 \pmod 3$. □

By Proposition 3.8 and Table 2, we have:

**Corollary 3.9** *Let $K/\mathbb{Q}$ be a nonic extension with $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_3^2$, and let $p$ be a rational prime. If $p \equiv 2 \pmod 3$, then we have $a_{p^3}(K) \in \{3, 165\}$.*

*Example 3.10* ($C_9$, sufficient condition) Assume that a nonic Galois extension $K/\mathbb{Q}$ follows the bold branch in Figure 2, so that $a_{10^3} = a_{2^3} \cdot a_{5^3} \leq 4$. Then, $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_9$. Indeed, if $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_3^2$, then by Corollary 3.9, we should have $a_{10^3} = a_{2^3} \cdot a_{5^3} \geq 3 \cdot 3 = 9$. This explains the correctness of the prediction.

*Example 3.11* ( False positive for $C_3^2$) Although the decision tree (Figure 2) attains 100% accuracy on the test set, this does not imply that it is a perfect classifier. Indeed, there exists a $C_9$-extension that satisfies all the conditions of the right-most branch (the dashed bold branch). With brute-force search, we found a prime $q = 960643$ satisfying $q \equiv 1 \pmod 9$ and $p^{\frac{q-1}{9}} \equiv 1 \pmod q$ for $p = 2,3,5,7$. Let $L = \mathbb{Q}(\zeta_q)$, and let $K$ be the unique nonic subextension of $L$. Then, the Frobenius at $p = 2,3,5,7$ are all trivial on $K$, so these primes split completely in $K$. From Table 2, we find that $a_{2^3} = a_{3^3} = a_{5^3} = a_{7^3} = 165$, satisfying all the inequalities required by the decision tree, even though $K$ is not a $C_3^2$-extension. The extension is ramified only at 960643, and it is not currently listed in LMFDB [37].

Note that there are infinitely many such primes $q$ by the Chebotarev density theorem. More precisely, an odd prime $q$ satisfies $q \equiv 1 \pmod 9$ if and only if $\Phi_9(x) = x^6+x^3+1 \equiv 0 \pmod q$ has a solution. For such a prime $q$, we have $p^{\frac{q-1}{9}} \equiv 1 \pmod q$ if and only if the polynomial $f_p(x) \coloneqq x^9 - p$ has a solution modulo $q$. By the Chebotarev density theorem, the set of primes $q$ for which all of $\Phi_9(x), f_2(x), f_3(x), f_5(x), f_7(x)$ have solutions modulo $q$ has positive (natural) density; for example, the compositum of the splitting fields of these polynomials has a degree $39366 = 2 \cdot 3^9$ (which can be verified, for instance, using

`Magma` [11]). Hence, at least 1/39366 of all primes split completely in the compositum, and for such primes $q$, each polynomial has a solution modulo $q$.

**3.3.1.2 Logistic regression** We also trained logistic regression models and see if they can successfully distinguish Galois groups for quartic and nonic number fields. With 1000 zeta coefficients (resp. with standard normalization), they can obtain accuracy of 99.60% (resp. 99.41%) and 97.16% (resp. 78.51%) for quartic and nonic fields, respectively (Tables 15 and 16). We checked the indices with largest and smallest weights (Table 3). One can find that most of the indices for large weights are square or cube-indexed coefficients, which means that the model tends to predict output as *1*, which corresponds to $C_2^2$ or $C_3^2$, when those features are large. These results align well with Corollary 3.3.

**Table 3** Indices of the 10 largest and smallest weights (with actual weight values in parentheses) among 1000 logistic regression coefficients used to classify Galois groups of quartic and nonic fields. The last two rows with nonic* are the weights of the model trained with original zeta coefficients without normalization.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| quartic | top | **25 (8.4)** | **49 (7.2)** | **961 (6.1)** | **841 (5.5)** | **529 (5.3)** | **9 (5.1)** | **361 (4.3)** | **36 (3.2)** | **100 (3.0)** | **289 (3.0)** |
| | bottom | 343 (–10.6) | **625 (–7.6)** | 31 (–6.8) | **81 (–6.3)** | 29 (–5.7) | 23 (–5.5) | 19 (–4.7) | **324 (–3.7)** | **16 (–3.6)** | 17 (–3.4) |
| nonic | top | 392 (2.0) | **27 (1.9)** | 837 (1.9) | **8 (1.8)** | 999 (1.8) | 456 (1.6) | **216 (1.6)** | 248 (1.6) | **125 (1.6)** | 639 (1.5) |
| | bottom | 213 (–1.3) | 111 (–1.3) | 321 (–1.0) | 57 (–1.0) | 291 (–0.8) | 3 (–0.8) | 717 (–0.7) | 87 (–0.7) | 49 (–0.6) | 807 (–0.6) |
| nonic* | top | **216 (0.6)** | **1000 (0.5)** | **64 (0.2)** | 248 (0.2) | 513 (0.1) | **125 (0.1)** | **8 (0.1)** | 584 (0.1) | **27 (0.1)** | 424 (0.1) |
| | bottom | 625 (–0.5) | 199 (–0.2) | 256 (–0.2) | 267 (–0.2) | 569 (–0.2) | 941 (–0.1) | 677 (–0.1) | 619 (–0.1) | 200 (–0.1) | 241 (–0.1) |

Bold values correspond to square or cube-indexed coefficients

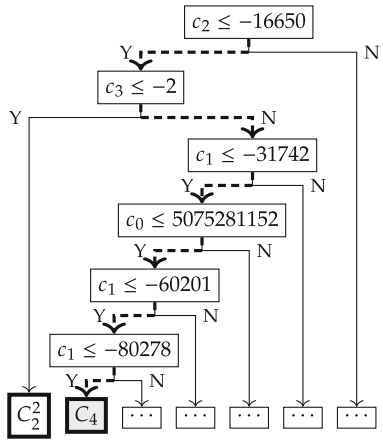



**Fig. 3** A decision tree predicts the Galois groups of quartic fields from polynomial coefficients

We also compared the distribution of the absolute values of the model weights for square-indexed (resp. cube-indexed) coefficients with the overall distribution. We found that the former tends to be higher on average than the latter (Figure 10). For quartic (resp. nonic) fields, the average absolute value of all weights is 0.19 (resp. 0.02), while the average for square-indexed (resp. cube-indexed) coefficients is 1.79 (resp. 0.20). This shows that the logistic regression models, like the decision tree models, are also able to learn that such coefficients are useful for prediction.

Tables 15 and 16 also include results for different subsets of zeta coefficients: using only square-indexed (or cube-indexed) coefficients or using only prime(-power)-indexed coefficients. Notably, using just 31 square-indexed coefficients still yields an accuracy of 99.99% (resp. 100%) for quartic fields (resp. nonic fields), further validating their importance. Moreover, the performance of models trained only on prime-power-indexed coefficients matches that of models using all coefficients, indicating that the logistic regression model is able to learn the importance of such coefficients for prediction, in agreement with the decision tree model. In contrast, the model achieves only 91.11% accuracy on quartic dataset with prime-indexed coefficients, which is not particularly high given the dataset's bias: 88.04% of quartic Galois fields have $C_2^2$ as their Galois group (Table 1).

*Remark 3.12* Mostly, the performance of logistic regression models with normalized zeta coefficients is similar or better than one trained with original zeta coefficients, e.g., quartic (15), sextic (21), and decic (26). However, this is not the case for nonic Galois extensions, where the accuracy with original zeta coefficient (97.60%) is much higher than with normalized (78.51%), where the latter barely matches the majority baseline (Table 1). It would be interesting to figure out why this happens specifically for nonic extensions.

#### *3.3.2 Learning Galois groups from polynomial coefficients*

We also conduct experiments on predicting Galois groups from the coefficients of defining polynomials, using decision tree and logistic regression models. Similar experiments were previously carried out in [25] using random forest models.

Let $c_0, \ldots, c_{d-1}$ be the coefficients of a monic polynomial of degree $d$, so that $f(x) = x^d + \sum_{k=0}^{d-1} c_k x^k$, where the polynomial is uniquely determined by the procedure described

**Table 4** Mean (left) and median (right) of polynomial coefficients for quartic extensions

| | $c_0$ | | $c_1$ | | $c_2$ | | $c_3$ | |
|---|---|---|---|---|---|---|---|---|
| $C_4$ | $1.0 \cdot 10^{11}$ | $2.1 \cdot 10^{6}$ | $-1.1 \cdot 10^{7}$ | 0 | $-4.1 \cdot 10^{4}$ | $-189$ | $-0.6$ | $-1$ |
| $C_2^2$ | $6.3 \cdot 10^{7}$ | $9.7 \cdot 10^{4}$ | 77.1 | 0 | $-4.6 \cdot 10^{2}$ | $-52$ | $-0.5$ | 0 |

**Table 5** Mean (left) and median (right) of polynomial coefficients for nonic extensions

| | $c_0$ | | $c_1$ | | $c_2$ | | $c_3$ | |
|---|---|---|---|---|---|---|---|---|
| $C_9$ | $-4.5 \cdot 10^{8}$ | $-1.9 \cdot 10^{5}$ | $8.1 \cdot 10^{8}$ | $8.1 \cdot 10^{6}$ | $-4.9 \cdot 10^{7}$ | $-5.0 \cdot 10^{4}$ | $-2.4 \cdot 10^{7}$ | $-4.6 \cdot 10^{6}$ |
| $C_3^2$ | $-1.5 \cdot 10^{9}$ | $-5.4 \cdot 10^{4}$ | $-8.4 \cdot 10^{8}$ | $-1.0 \cdot 10^{6}$ | $-1.2 \cdot 10^{8}$ | $-1.9 \cdot 10^{6}$ | $-9.4 \cdot 10^{6}$ | $-9.1 \cdot 10^{5}$ |

| | $c_4$ | | $c_5$ | | $c_6$ | | $c_7$ | | $c_8$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_9$ | $4.4 \cdot 10^{5}$ | $1.2 \cdot 10^{4}$ | $2.0 \cdot 10^{5}$ | $8.9 \cdot 10^{4}$ | $-1.3 \cdot 10^{3}$ | $-211$ | $-7.3 \cdot 10^{2}$ | $-657$ | $-0.5$ | 0 |
| $C_3^2$ | $6.2 \cdot 10^{5}$ | $6.1 \cdot 10^{4}$ | $1.1 \cdot 10^{5}$ | $3.5 \cdot 10^{4}$ | $-8.4 \cdot 10^{2}$ | $-346$ | $-4.6 \cdot 10^{2}$ | $-340$ | $-0.9$ | 0 |

in Section 2.3. Using these polynomial coefficients as features, decision tree models achieve reasonably high accuracies of 95.16% and 94.15% for quartic and nonic fields, respectively, although the resulting models are highly complex and not easily interpretable. Nevertheless, we found that the distributions of the coefficients vary significantly across Galois groups, and the models appear to exploit this variation for prediction (Figure 3). As shown in Table 4, the constant term $c_0$ tends to be much larger for polynomials with Galois group $C_4$ compared to those with $C_2^2$, while $c_1$ and $c_2$ (resp. $c_3$) are much (resp. slightly) smaller.

There are 146,288 quartic fields in the training set, of which 17,483 (resp. 128,805) have Galois group $C_4$ (resp. $C_2^2$). The first node of the decision tree checks whether $c_2 \leq -16,650$. Only 576 of the $C_2^2$-extensions (0.44% of the 128,805 $C_2^2$-extensions) satisfy this condition, while 3,025 of the $C_4$-extensions (17.3% of the 17,483 $C_4$-extensions) do. Following the remaining dashed path in Fig. 3, the tree predicts that a quartic field with

$$c_0 \leq 5{,}075{,}281{,}152, \quad c_1 \leq -80{,}278, \quad \text{and} \quad c_3 > -2$$

has Galois group $C_4$. This prediction is reasonable in light of the discrepancy in the distribution of polynomial coefficients shown in Table 4.

Similarly, we found that the distributions of polynomial coefficients differ significantly across Galois groups of nonic extensions (Table 5), and the decision tree model appears to use this information for prediction. The first node of the tree checks whether $c_1 \leq 233{,}597{,}936 \approx 2.3 \times 10^8$; if the condition is not satisfied, the model predicts the Galois group to be $C_9$. This is reasonable, considering that the values of $c_1$ for $C_9$-extensions are typically large and positive, while those for $C_3^2$-extensions tend to be negative. At present, we do not have a clear mathematical explanation for this discrepancy, and it would be interesting to explore one. We emphasize again that this phenomenon may strongly depend on how `LMFDB` selects a unique defining polynomial for each number field (see Section 2.3).

In contrast, the logistic regression model performed very poorly on quartic fields when using polynomial coefficients, achieving only 22.45% accuracy, which is significantly worse than the accuracy by a majority guess. This poor performance is due to the large magnitude of the polynomial coefficients; after standard normalization, the model's accuracy matches or higher than the majority baseline: 88.06% for quartic fields and 85.46% for nonic fields (the majority baselines are 88.04% and 77.57%, respectively).

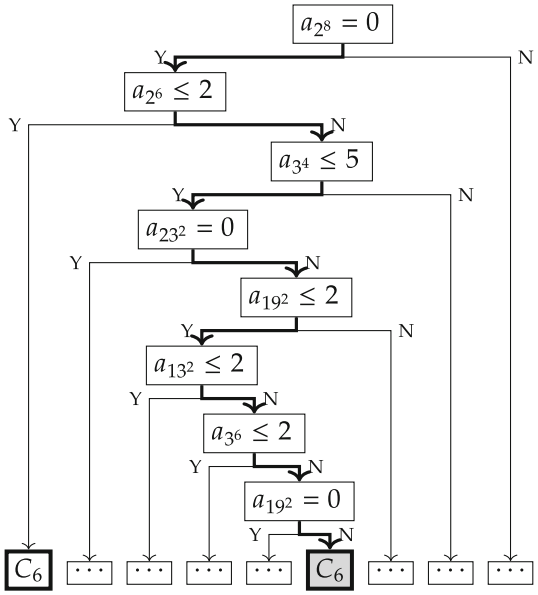



**Fig. 4** A decision tree predicts Galois groups of sextic fields from zeta coefficients up to $a_{1000}(K)$

Tables 15 and 16 summarize the performance of the models on quartic and nonic fields.

### 3.4 Sextic fields

For a degree 6 Galois extension $K/\mathbb{Q}$ (a *sextic* extension), there are two possible Galois groups: $C_6$ (cyclic) and $S_3$. In particular, the nonabelian group $S_3$ can arise as a Galois group, which makes the sextic case more interesting than the previously discussed cases of degree 4 and 9.

There have been various works on computing the Galois groups of sextic fields (including non-Galois extensions) from their defining polynomials [8,15,18]. As noted at the beginning of this Section 3, our interest lies in investigating how the Galois group can be determined from a finite number of Dedekind zeta function coefficients and polynomial coefficients, guided by the interpretation of machine learning (ML) experiments.

#### *3.4.1 Learning Galois groups from zeta coefficients*

**3.4.1.1 Decision tree** A decision tree model trained on zeta coefficients with $n = 1000$ achieves an accuracy of 98.96% in classifying the Galois groups of sextic fields. Similar to the quartic and nonic cases, we found that the top nodes of the tree correspond to square- or cube-indexed coefficients (see Fig. 4, where the first few nodes are $a_{2^8}$, $a_{2^6}$, $a_{3^6}$, $a_{3^4}$, $a_{18^2}$, $a_{20^2}$, etc.). Based on this observation, we trained a model using only the square- and cube-indexed coefficients up to $n = 1000$ (there are 38 such numbers) and obtained an even higher accuracy of 99.31%.

When we realized importance of square- and cube-indexed coefficients in decision trees' performances, we performed exploratory data analysis (EDA) on the dataset we used[2]. In particular, for each prime $p$, we investigated the possible values of $a_{p^2}(K)$ and $a_{p^3}(K)$ over

[2]This can be reproduced, for small primes, with `verify_galois.sage` in the GitHub repository.

**Table 6** Decomposition types, Euler factors, $p, p^2, p^3$-th zeta coefficients of sextic extensions

| $(e, f, g)$ | 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 |

all sextic Galois extension in `LMFDB`. Interestingly, we found that the possible values of $a_{p^2}(K)$ depend on the Galois group *and* $p \pmod 3$. Particularly, we found that when $p \equiv 1 \pmod 3$, then $a_{p^2}(K) = 1$ only if $K/\mathbb{Q}$ is a cyclic extension, and when $p \equiv 2 \pmod 3$, then $a_{p^2}(K) = 1$ only if $K/\mathbb{Q}$ is a sextic extension. Also, we observed that $a_{p^3}(K) = 1$ only if $K/\mathbb{Q}$ is a cyclic extension.

In the remainder of this subsection, we prove the above observation (Corollary 3.16) and explain how it can be used to show that the bold branch of the tree (Fig. 4) give *provably* correct predictions (Example 3.17). As in previous cases, we analyze the possible values of the zeta coefficients based on the decomposition types of primes. Table 6 lists the Euler factors and the corresponding zeta coefficients for each possible decomposition type.

**Proposition 3.13** *Let $K/\mathbb{Q}$ be a sextic extension with $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_6$, and let $p$ be a prime with decomposition type $(e, f, g)$ in $K$. If $p \equiv 5 \pmod 6$, then $(e, f, g) \notin \{(3, 1, 2), (3, 2, 1), (6, 1, 1)\}$.*

*Proof* Assume $(e, f, g) \in \{(3, 1, 2),\ (3, 2, 1),\ (6, 1, 1)\}$. In all cases, the decomposition group has order $ef = 3$ or 6, and the corresponding extension $K_{\mathfrak{p}}/\mathbb{Q}_p$ is a totally ramified $C_3$- or $C_6$-extension. (For $(e, f, g) = (3, 1, 2)$, we have $(p) = \mathfrak{p}_1^3\mathfrak{p}_2^3$, and we choose either $\mathfrak{p}_1$ or $\mathfrak{p}_2$ as $\mathfrak{p}$.) Let $\tilde{K} \subset K_{\mathfrak{p}}$ be the unique subextension of degree 3 over $\mathbb{Q}_p$. Then, $\tilde{K}/\mathbb{Q}_p$ is a totally and tamely ramified $C_3$-extension (note that $p \equiv 5 \pmod 6$ implies $p \neq 2, 3$, so the extension is tame), and Proposition 2.3 implies that $3 \mid (p-1)$, which contradicts the assumption $p \equiv 5 \pmod 6$. □

**Lemma 3.14** *Let $p \equiv 1 \pmod 6$ be a prime. Then, any degree 6 Galois extension of $\mathbb{Q}_p$ is cyclic.*

*Proof* Let $K/\mathbb{Q}_p$ be a sextic Galois extension with $G = \mathrm{Gal}(K/\mathbb{Q}_p) \simeq S_3$. Since $(p, 6) = 1$, the extension is tamely ramified, the inertia group $I$ is cyclic. Then, $G$ fits into the short exact sequence

$$1 \to C_e \to G \to C_f \to 1, \tag{8}$$

and the assumption $G \simeq S_3$ implies $(e, f) = (3, 2)$. Let $\tau$ be a generator of the inertia group $I \simeq C_3$, and let $\sigma$ be a lift of Frobenius, i.e., a generator of $C_f$. By Proposition 2.4 and $p \equiv 1 \pmod 3$, we have

$$\sigma\tau\sigma^{-1} = \tau^p = \tau,$$

which implies that $\tau$ and $\sigma$ commute. Thus, $G \simeq C_3 \times C_2 \simeq C_6$, contradicting $G \simeq S_3$. □

**Corollary 3.15** *Let $K/\mathbb{Q}$ be a sextic extension with $G = \mathrm{Gal}(K/\mathbb{Q}) \simeq S_3$, and let $p$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then:*

1. *For any $p$, $(e,f,g) \notin \{(1,6,1),(2,3,1)\}$.*
2. *If $p \equiv 1 \pmod 6$, then $(e,f,g) \notin \{(3,2,1),(6,1,1)\}$.*
3. *If $p \equiv 2 \pmod 3$, then $(e,f,g) \notin \{(3,1,2),(6,1,1)\}$.*

*Proof* 1) The cases $(e,f,g) = (1,6,1)$ and $(2,3,1)$ cannot occur for any prime $p$. Indeed, in either case, $G = D_{\mathfrak{p}}$ (for $\mathfrak{p} \mid p$) and $D_{\mathfrak{p}}/I_{\mathfrak{p}} \simeq C_6$ or $C_3$. However, this is impossible since $S_3$ admits no such quotients.

2) When $p \equiv 1 \pmod 6$, Lemma 3.14 implies that the decomposition group cannot have order 6 with $G \cong S_3$, i.e., $ef \neq 6$. This excludes $(3,2,1)$ and $(6,1,1)$.

3) Lastly, assume $p \equiv 5 \pmod 6$. Then, the decomposition type $(e,f,g) = (6,1,1)$ is not possible, since the inertia group of a tamely ramified extension must be cyclic, and $C_6$ is not a subgroup of $S_3$. If $(e,f,g) = (3,1,2)$, then $K_{\mathfrak{p}}/\mathbb{Q}_p$ is a totally ramified $C_3$-extension, which cannot exist when $p \equiv 5 \pmod 6$, as explained in the proof of Proposition 3.13.

For $p = 2$, assume it has a decomposition type of $(e,f,g) = (3,1,2)$. Since $2 \nmid 3$, the corresponding local extension $K_{\mathfrak{p}}/\mathbb{Q}_2$ is tamely ramified. By Proposition 2.3, we should have $e \mid p-1$, a contradiction. If it has decomposition type $(e,f,g) = (6,1,1)$, one can consider a cubic subextension of the corresponding local extension, which reduces to the case $(e,f) = (3,1)$. □

By Proposition 3.13, Corollary 3.15, and Table 6, we obtain the following result.

**Corollary 3.16** *Let $K/\mathbb{Q}$ be a sextic Galois extension and $p \geq 5$ be a prime.*

1. *If $p \equiv 1 \pmod 6$, then the set of possible values of $a_{p^2}(K)$ is given by*

$$a_{p^2}(K) \in \begin{cases} \{0,1,3,6,21\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_6, \\ \{0,3,6,21\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq S_3. \end{cases} \tag{9}$$

   *Particularly, if there exists a prime $p$ with $p \equiv 1 \pmod 6$ and $a_{p^2}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_6$.*
2. *If $p \equiv 2 \pmod 3$, then the set of possible values of $a_{p^2}(K)$ is given by*

$$a_{p^2}(K) \in \begin{cases} \{0,3,6,21\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_6, \\ \{0,1,3,6,21\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq S_3. \end{cases} \tag{10}$$

   *Particularly, if there exists a prime $p$ with $p \equiv 5 \pmod 6$ and $a_{p^2}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq S_3$.*
3. *The set of possible values of $a_{p^3}(K)$ is given by*

$$a_{p^3}(K) \in \begin{cases} \{0,1,2,4,10,56\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_6, \\ \{0,2,4,10,56\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq S_3. \end{cases} \tag{11}$$

   *Particularly, if there exists a prime $p$ with $a_{p^3}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_6$.*

*Example 3.17* ($C_6$, sufficient condition) Consider the bold branch in Figure 4. If a sextic field satisfies the inequalities, then we have $0 < a_{19^2} \leq 2$, and hence, $a_{19^2} = 1$ since $a_{19^2}$ cannot be 2 (Table 6). Since $19 \equiv 1 \pmod 6$, Corollary 3.16 implies that the field has to be a $C_6$-extension, which shows that the branch always gives a correct prediction.

One may ask whether Corollary 3.16 can distinguish between $C_6$- and $S_3$-extensions by searching for primes $p$ such that $a_{p^2}(K) = 1$. Table 21 shows that $a_{p^2}(K) = 1$ occurs only when the decomposition type of $p$ is (3, 2, 1) or (6, 1, 1), implying that $p$ must ramify in $K$. Consequently, for any Galois sextic extension $K$, verifying that $a_{p^2}(K) = 1$ for a ramifying prime $p$ with $p \equiv 1 \pmod 6$ suffices to conclude that $K$ is a $C_6$-extension. Table 17 presents the distribution of sextic fields according to the value of $a_{p^2}(K)$ for a fixed prime $p$ and Galois group $G \in \{C_6, S_3\}$. For example, about 7% of $C_6$-extensions on `LMFDB` satisfy $a_{19^2}(K) = 1$. In such cases, the above procedure correctly identifies the Galois group.

**3.4.1.2 Logistic regression** Using the first 1000 zeta coefficients (resp. with standard normalization), the logistic regression model achieves an accuracy of 98.40% (resp. 97.40%) (Table 21). Table 7 lists the indices with the largest and smallest weights, along with their corresponding values. Notably, most of the indices with large weights (except for 52 among the top 10) are products of prime powers with exponents that are multiples of 2 or 3.

As in the case of quartic and nonic fields (Section 3.3), we also compared the distribution of the absolute values of model weights for indices divisible by squares or cubes with the overall distribution. We found that the former is, on average, higher than the latter (Fig. 13a and b). The average absolute value of all weights is 0.63, while the average for indices divisible by square or cube factors is 3.94. This indicates that the logistic regression model aligns with the decision tree model in identifying the importance of such features.

### *3.4.2 Learning Galois groups from polynomial coefficients*

As in the case of quartic and nonic fields (Section 3.3), the decision tree model achieved a high accuracy of 97.56% with polynomial coefficients, although the trained model is very complicated and not easy to interpret the meaning of nodes. However, we found that it tends to classify a given sextic field as a $S_3$-extension when coefficients are "small." For example, there are 125,734 fields in the training set, where 17,029 (resp. 108,785) of them have a Galois group $C_6$ (resp. $S_3$). The first node of the tree asks whether $c_2 \leq 520,127$ or not, and 108442 of $S_3$-extensions (99.76% among 108785 $S_3$-extensions) satisfy the condition, while 5831 $C_6$-extensions (34.24% among 17,029 $C_6$-extensions) satisfy the opposite condition of $c_2 > 520,127$. Table 8 shows the mean and median of the polynomial coefficients $c_0, c_1, \ldots, c_5$ for each Galois group. The magnitudes $|c_0|, |c_1|, \ldots, |c_5|$ are, on average, much larger for $C_6$-extensions.

In contrast, the logistic regression model performed very poorly when using polynomial coefficients, achieving an accuracy of only 26.12%. Similarly to the discussion at the end of Sect. 3.3.2, this is due to the large magnitude of polynomial coefficients. After normalization, the model achieves 86.73% accuracy, which is slightly above the majority baseline (86.55%, Table 1). In particular, the balanced accuracy is only 50.72%, where the confusion matrix (d) of Fig. 11 shows that logistic regression models' predictions are highly biased toward $S_3$.

Table 21 summarizes the results for sextic extensions.

**Table 7** Indices of the 10 largest and smallest weights (with actual weight values in parentheses) among 1000 logistic regression coefficients used to classify Galois groups of sextic fields with normalized zeta coefficients

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| top | **125 (14.4)** | **25 (12.0)** | **4 (7.8)** | **512 (7.4)** | 128 (6.9) | 32 (6.3) | **27 (5.5)** | **100 (4.8)** | **36 (4.3)** | **196 (4.2)** |
| bottom | **625 (–20.9)** | **16 (–15.5)** | 243 (–9.5) | **64 (–9.3)** | 496 (–6.6) | 5 (–6.4) | 112 (–6.0) | 304 (–6.0) | **400 (–5.9)** | 48 (–5.9) |

Bold values correspond to indices that are products of prime powers with exponents that are multiples of 2 or 3

**Table 8** Mean (left) and median (right) of polynomial coefficients of sextic extensions

| | $c_0$ | | $c_1$ | | $c_2$ | | $c_3$ | | $c_4$ | | $c_5$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_6$ | $4.4 \cdot 10^{17}$ | $1.5 \cdot 10^5$ | $-3.7 \cdot 10^{14}$ | $-2.3 \cdot 10^3$ | $7.6 \cdot 10^{10}$ | $1.3 \cdot 10^5$ | $9.1 \cdot 10^6$ | $-2$ | $-3.1 \cdot 10^3$ | $-77$ | $-1.1$ | $-1$ |
| $S_3$ | $2.2 \cdot 10^9$ | $7.2 \cdot 10^5$ | $-2.4 \cdot 10^6$ | 0 | $2.0 \cdot 10^5$ | $2.3 \cdot 10^4$ | $8.6 \cdot 10^2$ | 0 | $-1.4 \cdot 10^2$ | $-50$ | $-0.7$ | 0 |

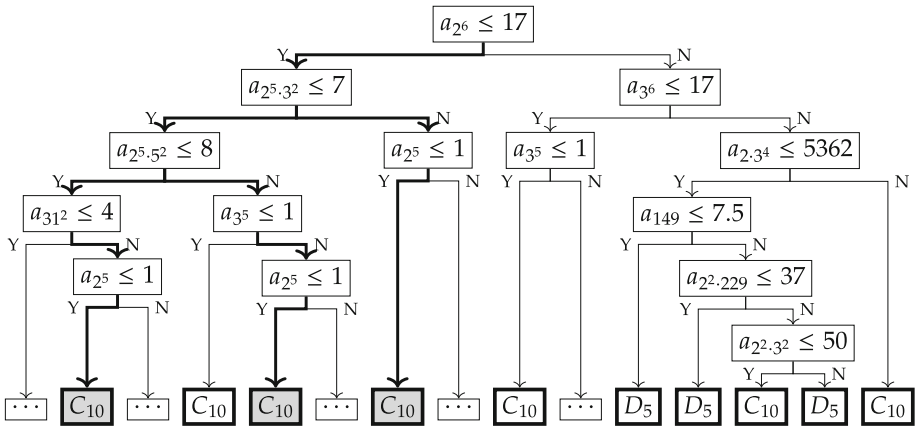



**Fig. 5** A decision tree predicts Galois groups of decic fields from zeta coefficients up to $a_{1000}(K)$

### 3.5 Decic fields

As in the sextic case, there are two possible Galois groups for decic Galois extensions $K/\mathbb{Q}$: $C_{10}$ (cyclic, abelian) and $D_5$ (dihedral, nonabelian). There are few prior works that focus on computing Galois groups of decic extensions, but several studies consider number fields of higher degrees, including [18,22,34].

#### *3.5.1 Learning Galois groups from zeta coefficients*

**3.5.1.1 Decision tree** A decision tree model trained on the first 1000 zeta coefficients achieves an accuracy of 97.19% in classifying the Galois groups of decic fields. Figure 5 shows the resulting tree, where the top nodes correspond to indices divisible by squares or fifth powers. Based on this observation, we trained a decision tree using only the 114 indices up to $n = 1000$ that are divisible by squares or fifth powers and obtained a higher accuracy of 98.56%.

As in previous cases, we analyze the possible values of the zeta coefficients based on the decomposition types of primes. In particular, by performing EDA as in the sextic extension case, we found that $a_{p^2}(K)$ can be 1 only if $K$ has a certain Galois group, *depending on* $p \equiv 1 \pmod{10}$. Table 9 lists the Euler factors and the $p^2$- and $p^5$-th zeta coefficients corresponding to each possible decomposition type. The observation is proved in Corollary 3.21, and this can be used to prove that the bold branches of Figure 5 give *provably correct predictions* (see Example 3.22).

**Proposition 3.18** *Let $K/\mathbb{Q}$ be a decic Galois extension with Galois group* $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}$ *and $p$ be a rational prime. Denote by $(e, f, g)$ the decomposition type of $p$ in $K$. If $p = 2$ or $p \equiv 3, 7, 9 \pmod{10}$, then $(e, f, g) \notin \{(5, 1, 2), (5, 2, 1), (10, 1, 1)\}$.*

*Proof* The proof is similar to that of Proposition 3.13 and Corollary 3.15.

**Table 9** Decomposition types, Euler factors, and $p, p^2, p^5$-th zeta coefficients of decic extensions

| $(e,f,g)$ | Euler factor | $a_p(K)$ | $a_{p^2}(K)$ | $a_{p^5}(K)$ |
|---|---|---|---|---|
| (1, 1, 10) | $(1-p^{-s})^{-10} = \sum_{k\geq 0}\binom{k+9}{9}p^{-ks}$ | 10 | 55 | 2002 |
| (1, 2, 5) | $(1-p^{-2s})^{-5} = \sum_{k\geq 0}\binom{k+4}{4}p^{-2ks}$ | 0 | 5 | 0 |
| (1, 5, 2) | $(1-p^{-5s})^{-2} = \sum_{k\geq 0}(k+1)p^{-5ks}$ | 0 | 0 | 2 |
| (1, 10, 1) | $(1-p^{-10s})^{-1} = \sum_{k\geq 0}p^{-10ks}$ | 0 | 0 | 0 |
| (2, 1, 5) | $(1-p^{-s})^{-5} = \sum_{k\geq 0}\binom{k+4}{4}p^{-ks}$ | 5 | 15 | 126 |
| (2, 5, 1) | $(1-p^{-5s})^{-1} = \sum_{k\geq 0}p^{-5ks}$ | 0 | 0 | 1 |
| (5, 1, 2) | $(1-p^{-s})^{-2} = \sum_{k\geq 0}(k+1)p^{-ks}$ | 2 | 3 | 6 |
| (5, 2, 1) | $(1-p^{-2s})^{-1} = \sum_{k\geq 0}p^{-2ks}$ | 0 | 1 | 0 |
| (10, 1, 1) | $(1-p^{-s})^{-1} = \sum_{k\geq 0}p^{-ks}$ | 1 | 1 | 1 |

For $p = 2$ and $(e,f,g) = (5,1,2)$, assume such extension exists. Since $2 \nmid 5$, the corresponding local extension $K_{\mathfrak{p}}/\mathbb{Q}_2$ is tamely ramified, and Proposition 2.3 implies $e \mid p-1$, a contradiction. For $(e,f,g) = (5,2,1)$ or $(10,1,1)$, one can consider the unique quintic subextension of the local extension and reduce to the previous case $(e,f) = (5,1)$.

For $p \equiv 3,7,9 \pmod{10}$, assume $(e,f,g) \in \{(5,1,2),(5,2,1),(10,1,1)\}$. Then, the corresponding extension $K_{\mathfrak{p}}$ of $\mathbb{Q}_p$ (for some prime $\mathfrak{p}|p$) is totally ramified $C_5$- or $C_{10}$-extension. In either case, there exists a totally ramified $C_5$-subextension $\tilde{K} \subset K_{\mathfrak{p}}$. Since $(p,5) = 1$, the extension $\tilde{K}/\mathbb{Q}_p$ is tamely ramified, and Proposition 2.3 implies that $p \equiv 1 \pmod{5}$, which contradicts the assumption $p \equiv 3,7,9 \pmod{10}$. □

**Lemma 3.19** *Let $p \equiv 1 \pmod{10}$ be a prime. Then, any decic Galois extension of $\mathbb{Q}_p$ is cyclic.*

*Proof* The proof is similar to that of Lemma 3.14. If $K/\mathbb{Q}_p$ is a decic Galois extension with $G = \mathrm{Gal}(K/\mathbb{Q}_p) \simeq D_{10}$, then the extension is tamely ramified from $(p,10) = 1$, and hence, the inertia group $I$ is cyclic. By considering the same exact sequence (8) in Lemma 3.14, we should have $(e,f) = (5,2)$. Now, applying Proposition 2.4 gives a contradiction. □

**Proposition 3.20** *Let $K/\mathbb{Q}$ be a decic $D_5$-extension and $p$ be a rational prime. Let $(e,f,g)$ be the decomposition type of $p$ in $K$. Then:*

1. *If $p = 2$, then $(e,f,g) \notin \{(1,10,1),(2,5,1),(5,1,2),(5,2,1),(10,1,1)\}$.*
2. *If $p = 5$, then $(e,f,g) \notin \{(1,10,1),(2,5,1)\}$.*
3. *If $p \equiv 1 \pmod{10}$, then $(e,f,g) \notin \{(1,10,1),(2,5,1),(5,2,1),(10,1,1)\}$.*
4. *If $p \equiv 3,7 \pmod{10}$, then $(e,f,g) \notin \{(1,10,1),(2,5,1),(5,1,2),(5,2,1),(10,1,1)\}$.*
5. *If $p \equiv 9 \pmod{10}$, then $(e,f,g) \notin \{(1,10,1),(2,5,1),(5,1,2),(10,1,1)\}$.*

*Proof* The proof is similar to that of Corollary 3.15.

First, the case of (1, 10, 1) can be easily excluded for all $p$, since $C_{10}$ cannot arise as a quotient of $D_5$. Also, if a prime $p$ has decomposition type (2, 5, 1), then the decomposition group is isomorphic to $D_5$ and has a normal subgroup of order 2. This is impossible, since any such normal subgroup must lie in the center, whereas the center of $D_5$ is trivial as 5 is odd.

If $p \equiv 3,7,9 \pmod{10}$, then $(e,f,g) \neq (10,1,1)$, since the inertia group of a tamely ramified extension must be cyclic, but $C_{10}$ is not a subgroup of $D_5$. Similarly, the decomposition types (1, 10, 1) and (2, 5, 1) are excluded, as neither $C_{10}$ nor $C_5$ can occur as a quotient of

$D_5$. If $(e,f,g) = (5,1,2)$, then the corresponding extension $K_\mathfrak{p}/\mathbb{Q}_p$ is a totally and tamely ramified $C_5$-extension and Proposition 2.3 gives $p \equiv 1 \pmod 5$. Thus, $(e,f,g) = (5,1,2)$ is excluded for $p \equiv 3,7,9 \pmod{10}$. If $(e,f,g) = (5,2,1)$, a corresponding $D_5$-extension $K_\mathfrak{p}/\mathbb{Q}_p$ has an unramified quadratic extension $\widetilde{K} \subset K_\mathfrak{p}$ where $K_\mathfrak{p}/\widetilde{K}$ is a totally and tamely ramified $C_5$-extension. Since the residue field of $\widetilde{K}$ is $\mathbb{F}_{p^2}$, Proposition 2.3 implies that $p^2 \equiv 1 \pmod 5$, and hence, $p \not\equiv 3,7 \pmod{10}$. □

As a result, we obtain the following information about the values of $a_{p^2}(K)$ and $a_{p^5}(K)$ for a decic field $K$.

**Corollary 3.21** *1. If $p \equiv 1 \pmod{10}$, then the set of possible values of $a_{p^2}(K)$ is given by*

$$a_{p^2}(K) \in \begin{cases} \{0,1,3,5,15,55\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}, \\ \{0,3,5,15,55\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq D_5. \end{cases} \tag{12}$$

*Particularly, if there exists a prime $p$ with $p \equiv 1 \pmod{10}$ and $a_{p^2}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}$.*

*2. If $p \equiv 9 \pmod{10}$, then the set of possible values of $a_{p^2}(K)$ is given by*

$$a_{p^2}(K) \in \begin{cases} \{0,5,15,55\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}, \\ \{0,1,5,15,55\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq D_5. \end{cases} \tag{13}$$

*Particularly, if there exists a prime $p$ with $p \equiv 9 \pmod{10}$ and $a_{p^2}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq D_5$.*

*3. For $p \neq 5$, the set of possible values of $a_{p^5}(K)$ is given by*

$$a_{p^5}(K) \in \begin{cases} \{0,1,2,126,2002\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}, \\ \{0,2,126,2002\} & \text{if } \mathrm{Gal}(K/\mathbb{Q}) \simeq D_5. \end{cases} \tag{14}$$

*Particularly, if there exists a prime $p \neq 5$ such that $a_{p^5}(K) = 1$, then $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_{10}$.*

*Example 3.22* ($C_{10}$, sufficient condition) Assume that a decic extension $K/\mathbb{Q}$ follows one of the bold branches in Figure 5. Then, $a_{2^6}(K) \leq 17$ and $a_{2^5}(K) \leq 1$. If $(e,f,g) = (1,2,5)$, then by Corollary 2.2, we have $a_{2^6}(K) = 35 > 17$, so this decomposition type cannot occur. Furthermore, from Table 9, the condition $a_{2^5}(K) \leq 1$ implies $(e,f,g) \in \{(1,10,1),(2,5,1),(5,2,1),(10,1,1)\}$, which cannot be the case when $K/\mathbb{Q}$ is a $D_5$-extension by Proposition 3.20. Hence, $K/\mathbb{Q}$ must be a cyclic extension.

**3.5.1.2 Logistic regression** Using 1000 zeta coefficients (resp. with standard normalization), the logistic regression model achieves an accuracy of 93.81% (reps. 94.70%) (Table 26). Table 10 lists the indices with the largest and smallest weights, along with their corresponding values. One can observe that all the top ten indices with largest weights correspond to square indices. This reflects the importance of square- and quintic-indexed coefficients in the classification.

We also compared the distribution of the absolute values of model weights for indices divisible by squares or fifth powers with the overall distribution, and found that the former

**Table 10** Indices of the 10 largest and smallest weights (with actual weight values in parentheses) among 1000 logistic regression coefficients used to classify Galois groups of decic fields with normalized zeta coefficients

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| top | **25 (3.9)** | **49 (3.4)** | **961 (2.4)** | **121 (2.3)** | **9 (2.3)** | **841 (1.7)** | **529 (1.7)** | **289 (1.6)** | **169 (1.4)** | **361 (1.3)** |
| bottom | 343 (–2.7) | 31 (–2.1) | 11 (–2.1) | 5 (–1.6) | 125 (–1.6) | 17 (–1.5) | 29 (–1.5) | 23 (–1.3) | 13 (–1.3) | 19 (–1.2) |

Bold numbers correspond to indices that are squares

**Table 11** Mean (left) and median (right) of polynomial coefficients for decic extensions

| | $c_0$ | | $c_1$ | | $c_2$ | | $c_3$ | | $c_4$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{10}$ | $3.3 \cdot 10^{19}$ | $1.8 \cdot 10^5$ | $-1.0 \cdot 10^{18}$ | $-5.0 \cdot 10^5$ | $7.7 \cdot 10^{15}$ | $1.5 \cdot 10^7$ | $3.7 \cdot 10^{13}$ | 0 | $-5.7 \cdot 10^{11}$ | $7.8 \cdot 10^2$ |
| $D_5$ | $8.1 \cdot 10^6$ | $8.1 \cdot 10^5$ | $-2.0 \cdot 10^5$ | 0 | $7.7 \cdot 10^5$ | $1.3 \cdot 10^5$ | $-6.0 \cdot 10^3$ | 0 | $2.9 \cdot 10^4$ | $1.4 \cdot 10^4$ |
| | $c_5$ | | $c_6$ | | $c_7$ | | $c_8$ | | $c_9$ | |
| $C_{10}$ | $-2.7 \cdot 10^8$ | $-8.4 \cdot 10^3$ | $1.6 \cdot 10^7$ | $2.5 \cdot 10^4$ | $-4.5 \cdot 10^2$ | 0 | $-2.5 \cdot 10^2$ | $-55$ | $-1.6$ | $-1$ |
| $D_5$ | $-1.0 \cdot 10^3$ | 0 | $1.7 \cdot 10^3$ | $7.0 \cdot 10^2$ | $-9.2$ | 0 | $-1.2$ | $-4$ | $-1.3$ | 0 |

is higher on average (Fig. 13c and d). The average of all weights is 0.22, while the average of the weights for indices with square or quintic factors is 0.99. This indicates that the logistic regression model is able to learn the importance of such coefficients for prediction, in agreement with the decision tree model.

#### *3.5.2 Learning Galois groups from polynomial coefficients*

Using polynomial coefficients, the decision tree model achieved an accuracy of 86.71%. As in previous cases, the model's performance may stem from discrepancies in the distributions of the coefficients across Galois groups (Table 11). For example, the first node asks whether $c_2 \leq 8{,}101{,}335$, and approximately 99% of the $D_5$-extensions in the training set satisfy this condition. In contrast, the logistic regression model achieved only 46.25% accuracy, which improves to 66.18% after normalization, close to the proportion of $D_5$-extensions. Recall the discussion at the end of Sect. 3.3.2.

Table 26 summarizes the results for decic extensions.

### 3.6 Octic fields

For degree 8 Galois extensions, there are five possible Galois groups: Three abelian groups ($C_8$, $C_4 \times C_2$, and $C_2^3$) and two are nonabelian groups ($D_4$ and $Q_8$).

There are several works that study the Galois groups of octic fields. Most of these focus on fields defined by specific types of degree 8 polynomials, although they are not limited to Galois extensions. For example, Chen, Chin, and Tan provide a complete classification of possible Galois groups for *doubly even octic polynomials* (i.e., $f(x) = x^8 + ax^4 + b$) and *palindromic even octic polynomials* (i.e., $f(x) = x^8 + ax^6 + bx^4 + ax^2 + 1$ with $a \neq 0$) by analyzing linear resolvents [16]. Awtrey and Patane generalized their result for palindromic even octic polynomials over $\mathbb{Q}$ to arbitrary characteristic zero fields [9]. However, these results are not enough to give a complete characterization of Galois groups for arbitrary defining polynomials; for example, the quaternion group cannot occur as the Galois group of an octic field defined by a doubly even octic polynomial [16, Theorem 3.1].

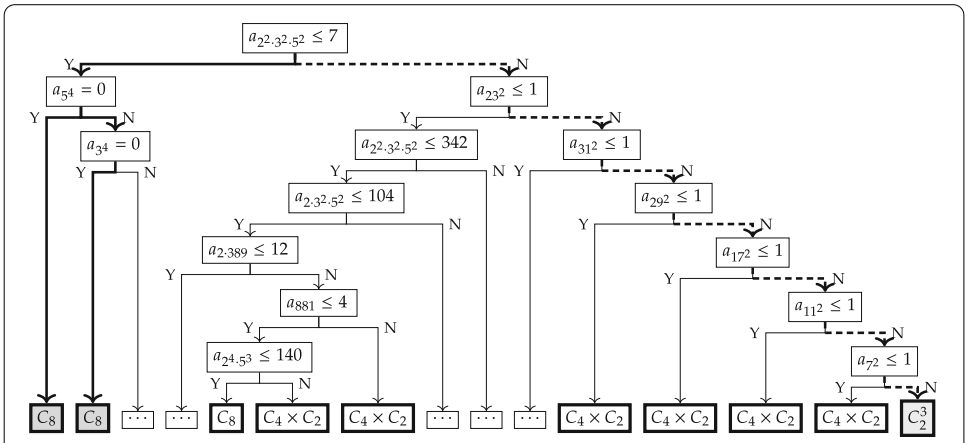



**Fig. 6** A decision tree predicts Galois groups of abelian octic fields from zeta coefficients up to $a_{1000}(K)$

**Table 12** Decomposition types, Euler factors, and $p^\ell$ and $p^{\ell^2}$-th zeta coefficients for degree $\ell^3$ extensions

| $(e, f, g)$ | Euler factor | $a_{p^\ell}(K)$ | $a_{p^{\ell^2}}(K)$ |
|---|---|---|---|
| $(1, 1, \ell^3)$ | $(1-p^{-s})^{-\ell^3} = \sum_{k\ge0}\binom{\ell^3+k-1}{\ell^3-1}p^{-ks}$ | $\binom{\ell^3+\ell-1}{\ell^3-1}$ | $\binom{\ell^3+\ell^2-1}{\ell^3-1}$ |
| $(1, \ell, \ell^2)$ | $(1-p^{-\ell s})^{-\ell^2} = \sum_{k\ge0}\binom{\ell^2+k-1}{\ell^2-1}p^{-\ell ks}$ | $\ell^2$ | $\binom{\ell^2+\ell-1}{\ell^2-1}$ |
| $(1, \ell^2, \ell)$ | $(1-p^{-\ell^2 s})^{-\ell} = \sum_{k\ge0}\binom{\ell+k-1}{\ell-1}p^{-\ell^2 ks}$ | 0 | $\ell+1$ |
| $(1, \ell^3, 1)$ | $(1-p^{-\ell^3 s})^{-1} = \sum_{k\ge0}p^{-\ell^3 ks}$ | 0 | 0 |
| $(\ell, 1, \ell^2)$ | $(1-p^{-s})^{-\ell^2} = \sum_{k\ge0}\binom{\ell^2+k-1}{\ell^2-1}p^{-ks}$ | $\binom{\ell^2+\ell-1}{\ell^2-1}$ | $\binom{2\ell^2-1}{\ell^2-1}$ |
| $(\ell, \ell, \ell)$ | $(1-p^{-\ell s})^{-\ell} = \sum_{k\ge0}\binom{\ell+k-1}{\ell-1}p^{-\ell ks}$ | $\ell$ | $\binom{2\ell-1}{\ell-1}$ |
| $(\ell, \ell^2, 1)$ | $(1-p^{-\ell^2 s})^{-1} = \sum_{k\ge0}p^{-\ell^2 ks}$ | 0 | 1 |
| $(\ell^2, 1, \ell)$ | $(1-p^{-s})^{-\ell} = \sum_{k\ge0}\binom{\ell+k-1}{\ell-1}p^{-ks}$ | $\binom{2\ell-1}{\ell-1}$ | $\binom{\ell+\ell^2-1}{\ell-1}$ |
| $(\ell^2, \ell, 1)$ | $(1-p^{-\ell s})^{-1} = \sum_{k\ge0}p^{-\ell ks}$ | 1 | 1 |
| $(\ell^3, 1, 1)$ | $(1-p^{-s})^{-1} = \sum_{k\ge0}p^{-ks}$ | 1 | 1 |

#### *3.6.1 Distinguishing between abelian extensions*

In [25, Section 5.2], the authors found that a random forest classifier trained on zeta coefficients up to $n = 1000$ can distinguish between different *abelian* octic extensions with high accuracy (95.47%). The three possible Galois groups are $C_8$, $C_4 \times C_2$, and $C_2^3$. We use a single decision tree for the classification task and find that it achieves a higher accuracy of 98.61%. As in the case of quartic extensions, we observe that the first few nodes of the decision tree focus on square-indexed zeta coefficients. Motivated by this, we train a decision tree using only square-indexed coefficients with $n \le 1000$, which results in an even higher accuracy of 98.91%.

Figure 6 shows that the model predicts the Galois group to be $C_8$ whenever a fourth-power-indexed coefficient $a_{m^4}(K)$ vanishes for some $m$. This strongly suggests that the Galois group is cyclic if there exists a prime $p$ such that $a_{p^4}(K) = 0$. Similarly, the right part of Fig. 6 shows that the model predicts the Galois group to be $C_4 \times C_2$ when the field is not cyclic and the square-indexed coefficients are small. Motivated by these observations, we studied the distribution of $a_{p^2}(K)$ and $a_{p^4}(K)$ for fixed primes $p$ and formulated a conjecture, which we are able to prove rigorously (Corollary 3.25). The proof follows from a couple of theorems established below.

**Theorem 3.23** *Let $\ell$ be a prime, and let $K/\mathbb{Q}$ be a Galois extension of degree $\ell^3$ with $G = \mathrm{Gal}(K/\mathbb{Q}) \simeq C_\ell^3$. Then for any prime $p$ and any integer $b \geq 1$, we have $a_{p^{b\ell}}(K) \neq 0$; that is, the $p^{b\ell}$-th zeta coefficients are always nonzero.*

*Proof* Since every element of $C_\ell^3$ has order at most $\ell$, and $D_p/I_p \subset G/I_p$ is a cyclic group of order $f$ (when $p$ has decomposition type $(e, f, g)$ in $K$), we must have $f = 1$ or $\ell$. Therefore, only 7 cases from Table 12 are possible, and the corresponding Euler factors are all of the form $(1 - p^{-s})^{-g}$ or $(1 - p^{-\ell s})^{-g}$. The nonvanishing of the $p^{b\ell}$-th coefficient then follows immediately. □

**Theorem 3.24** *Let $K/\mathbb{Q}$ be a degree $\ell^3$ abelian extension with $G = \mathrm{Gal}(K/\mathbb{Q}) \simeq C_{\ell^2} \times C_\ell$. Then, the set of rational primes $p$ for which $a_{p^\ell}(K) = 0$ has positive (natural) density. Moreover, $a_{p^{\ell^2}}(K) \neq 0$ for all primes $p$; that is, the $p^{\ell^2}$-th zeta coefficients are always nonzero.*

*Proof* From Table 12, any unramified prime $p$ with decomposition type $(1, \ell^2, \ell)$ or $(1, \ell^3, 1)$ (i.e., inert in $K$) satisfies $a_{p^\ell}(K) = 0$. The set of such primes has positive density by the Chebotarev density theorem. For the second part, from Table 12 again, it suffices to show that the decomposition type $(1, \ell^3, 1)$ cannot occur in $K$. This follows from the fact that $C_{\ell^3}$ is not a subquotient of $C_{\ell^2} \times C_\ell$. □

**Corollary 3.25** *Let $K/\mathbb{Q}$ be an abelian Galois extension of degree 8. Then, the following holds:*

1. *We have $a_{p^2}(K) \neq 0$ for all primes $p$ if and only if $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_2^3$.*
2. *We have $a_{p^4}(K) = 0$ for some prime $p$ if and only if $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_8$.*

*Proof* 1. ($\Rightarrow$) If $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_8$, then Theorem 3.2 with $\ell = 2$, $m = 3$, and $b = 2$ shows that there are infinitely many primes $p$ such that $a_{p^2}(K) = 0$. Theorem 3.24 implies the same conclusion when $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_4 \times C_2$. ($\Leftarrow$) This is precisely Theorem 3.23 with $\ell = 2$ and $b = 1$.

2. ($\Rightarrow$) If $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_2^3$, then Theorem 3.23 shows that $a_{p^4}(K)$ is always nonzero, by taking $\ell = b = 2$. The case $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_4 \times C_2$ follows from Theorem 3.24. ($\Leftarrow$) This follows from Theorem 3.2 with $\ell = 2$, $m = 3$, and $b = 4$. In particular, there are infinitely many such primes $p$. □

*Example 3.26* ($C_8$, sufficient condition) Consider the bold branch in Figure 6. If an abelian octic field satisfies the inequalities along this branch, then either $a_{5^4} = 0$ or $a_{3^4} = 0$. By Corollary 3.25, this implies that the field must be a $C_8$-extension. Therefore, the branch always yields a correct prediction.

As before, we can provide a list of the decomposition types that do not occur for each Galois group and prime $p$. (Propositions 3.27–3.29)

**Proposition 3.27** *Let $K/\mathbb{Q}$ be a $C_8$-extension and $p$ be a rational prime with decomposition type $(e, f, g)$ in $K$. Then:*

1. *If $p = 2$, then $(e, f, g) \notin \{(1, 8, 1), (2, 4, 1), (4, 2, 1)\}$.*
2. *If $p \equiv 5 \pmod 8$, then $(e, f, g) \neq (8, 1, 1)$.*
3. *If $p \equiv 3 \pmod 4$, then $(e, f, g) \notin \{(4, 1, 2), (4, 2, 1), (8, 1, 1)\}$.*

*For $p \equiv 1 \pmod 8$, each decomposition type $(e,f,g)$ occurs in some example.*

*Proof* $p = 2$ case is a direct consequence of Proposition 2.6; if $g = 1$, then 2 totally ramifies in $K$ and the only possibility is $(e,f,g) = (8,1,1)$.

For odd primes $p$, Proposition 2.3 shows that a totally ramified $C_8$-extension of $\mathbb{Q}_p$ exists only if $p \equiv 1 \pmod 8$. Hence, such an extension is not possible when $p \equiv 5 \pmod 8$ or $p \equiv 3 \pmod 4$. Similarly, if $p \equiv 3 \pmod 4$, there is no totally and tamely ramified $C_4$-extension of $\mathbb{Q}_p$, which excludes $(e,f,g) = (4,1,2)$ and $(4,2,1)$. □

**Proposition 3.28** *Let $K/\mathbb{Q}$ be a $C_4 \times C_2$-extension and $p$ be a rational prime with a decomposition type $(e,f,g)$ in $K$. Then:*

1. $(e,f,g) \neq (1,8,1)$.
2. *If* $p \equiv 1 \pmod 4$, *then* $(e,f,g) \neq (8,1,1)$.
3. *If* $p \equiv 3 \pmod 4$, *then* $(e,f,g) \notin \{(4,1,2),(4,2,1),(8,1,1)\}$.

*Proof* The largest possible cyclic quotient of $C_4 \times C_2$ is $C_4$. Hence, $f \leq 4$ and this excludes $(e,f,g) = (1,8,1)$.

For odd primes $p$, the corresponding extension $K_\mathfrak{p}/\mathbb{Q}_p$ is tamely ramified, so both the inertia subgroup $I_\mathfrak{p}$ and the quotient $D_\mathfrak{p}/I_\mathfrak{p}$ are cyclic, of orders $e$ and $f$, respectively. In particular, the decomposition type $(e,f,g) = (8,1,1)$ is impossible, since $C_8$ cannot embed into $C_4 \times C_2$. If $(e,f,g) = (4,1,2)$ or $(4,2,1)$, then Proposition 2.3 implies $p \equiv 1 \pmod 4$, so such types are excluded when $p \equiv 3 \pmod 4$. □

**Proposition 3.29** *Let $K/\mathbb{Q}$ be a $C_2^3$-extension and $p$ be a rational prime with decomposition type $(e,f,g)$ in $K$. Then:*

1. *If* $p = 2$, *then* $(e,f,g) \notin \{(1,4,2),(1,8,1),(2,4,1),(8,1,1)\}$.
2. *If $p$ is odd, then* $(e,f,g) \notin \{(1,4,2),(1,8,1),(2,4,1),(4,1,2),(4,2,1),(8,1,1)\}$.

*Proof* For $p = 2$ and $(e,f) = (1,4)$, there is no such extension, since decomposition group of such extension should be $C_4$, which cannot be a subgroup of $C_2^3$. Similarly, if $ef = 8$, then the decomposition group is $C_2^3$. Since $\mathrm{Gal}(\mathbb{F}_{2^f}/\mathbb{F}_2) \simeq C_f$ has to be a quotient of the decomposition group, we have $f \leq 2$ and this rules out $(e,f) = (1,8)$ and $(2,4)$.

In the case of $(8,1)$, assume that there exists a totally ramified $C_2^3$-extension $K_\mathfrak{p}$ of $\mathbb{Q}_2$. There are 7 quadratic extensions of $\mathbb{Q}_2$, which are $\mathbb{Q}_2(\sqrt{-1})$, $\mathbb{Q}_2(\sqrt{\pm 2})$, $\mathbb{Q}_2(\sqrt{\pm 3})$, $\mathbb{Q}_2(\sqrt{\pm 6})$, including the unique unramified extension $\mathbb{Q}_2(\sqrt{-3})$. Since there are exactly 7 (normal) subgroups of $C_2^3$ of index 2 (which can be counted by the number of nonzero vectors in $\mathbb{F}_2^3$ corresponding to the normal vectors of the corresponding hyperplanes), all of the quadratic extensions of $\mathbb{Q}_2$ appear as a subextension of $K_\mathfrak{p}$. In particular, it contains the unramified extension $\mathbb{Q}_2(\sqrt{-3})$ and cannot be totally ramified.

When $p$ is odd, the corresponding local field extensions $K_\mathfrak{p}/\mathbb{Q}_p$ (for $\mathfrak{p} \mid p$) are tamely ramified. The assertion then follows from the fact that the only cyclic subgroups or quotients of $C_2^3$ are trivial or isomorphic to $C_2$, so we must have $e,f \leq 2$. □

*Example 3.30* ($C_4 \times C_2$) Based on Propositions 3.27–3.29, one way to distinguish $C_4 \times C_2$ from the other two possibilities using zeta coefficients is to show that the decomposition type at $p = 2$ is $(2,4,1)$. For instance, this is enforced when $a_{2^2} = a_{2^3} = 0$ and $a_{2^4} = 1$, as

one can see from Table 12. However, we were not able to identify a branch in the decision tree (Fig. 6) that uses this condition to predict the Galois group as $C_4 \times C_2$.

*Example 3.31* ( False positive for $C_2^3$) We also have an example of a "false positive" branch in Fig. 6, similar to Example 3.11.

The only branch that predicts the Galois group as $C_2^3$ in Figure 6 is the right-most branch (dashed bold branch). In fact, the model achieves 100% recall and precision on the test set when the Galois group is $C_2^3$, i.e., there are no false positives or negatives for $C_2^3$. However, there exist non-$C_2^3$ abelian octic extensions that satisfy all the inequalities along this path but are not included in our dataset. Indeed, using a brute-force search, we find a prime $q = 14836487689$ such that $q \equiv 1 \pmod 8$ and $p^{\frac{q-1}{8}} \equiv 1 \pmod q$ for all $p = 2, 3, 5, 7, 11, 17, 23, 29, 31$. Then, all these primes $p$ completely split in the unique $C_8$-subextension $K$ of $\mathbb{Q}(\zeta_q)$, so $a_{p^2} = 36$ from Table 12 and the field satisfies all the inequalities on the dashed path, but is not a $C_2^3$-extension. Note that this field is not listed in `LMFDB` [37]; there is no number field ramified at $q$). As in Example 3.11, there are infinitely many such $q$ by the Chebotarev density theorem. Although these are the only examples we found from the trees, it would be interesting to find more such examples in other branches or tasks.

**3.6.1.1 Logistic regression** With 1,000 zeta coefficients (resp. with standard normalization) as input, the logistic regression model achieved 93.81% (resp. 92.88%) accuracy in classifying abelian octic Galois groups (see Table 29). Table 13 lists, for each Galois group, the indices corresponding to the 10 largest and 10 smallest model weights, along with the weight values. Since this is a 3-way classification task, the model has a total of $3 \times 1000$ weights and 3 bias terms. In comparison, all the other logistic regression experiments in this study were binary classification tasks, each involving 1,000 weights and one bias.

A notable pattern is that many indices associated with significantly large or small weights are squares. For instance, among the 10 largest weights for the Galois group $C_2^3$, all of them correspond to square indices. Also, 25 appears as the index of the seventh and sixth smallest weight for $C_8$ and $C_4 \times C_2$, while it is the index of the second largest weight for $C_2^3$. This indicates that a large value of $a_{25}$ pushes the model toward predicting $C_2^3$. The average absolute values of all weights are 0.32, 0.20, and 0.27 for the three Galois groups. However, when restricted to square-indexed coefficients, these averages increase significantly to 2.36, 1.27, and 2.50. This indicates that the model identifies square-indexed coefficients as particularly informative, consistent with the interpretations drawn from the decision tree model.

**3.6.1.2 Polynomial coefficients** Using polynomial coefficients, both the decision tree and logistic regression models perform poorly, achieving accuracies of 77.21% and 19.11%, respectively. The latter improves to 64.89% after normalization. Recall the discussion at the end of Section 3.3.2.

### *3.6.2 Distinguishing between nonabelian extensions*

We trained a decision tree to classify nonabelian octic Galois extensions, i.e., distinguishing between $D_4$- and $Q_8$-extensions. The model achieved a high accuracy of 98.67%, and we observed that the top nodes in the tree correspond to square-indexed coefficients. However, the previous approach does not yield a criterion based on zeta coefficients to

**Table 13** Indices of the 10 largest and smallest weights (with actual weight values in parentheses) among 1000 logistic regression coefficients used to classify Galois groups of abelian octic fields with normalized zeta coefficients

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_8$ | top | 125 (6.5) | 343 (5.7) | 31 (5.2) | **16 (4.4)** | **729 (4.2)** | 23 (4.0) | 19 (3.9) | 29 (3.8) | 11 (2.7) | 243 (2.6) |
| | bottom | **81 (–6.2)** | **961 (–5.3)** | **49 (–5.2)** | **529 (–4.3)** | **841 (–4.2)** | **361 (–4.2)** | **25 (–3.8)** | **289 (–2.8)** | **121 (–2.8)** | 162 (–2.7) |
| $C_4 \times C_2$ | top | **81 (9.2)** | **625 (5.4)** | 162 (3.8) | **324 (3.2)** | 567 (2.8) | 5 (1.8) | 891 (1.8) | 405 (1.5) | 648 (1.3) | 488 (1.3) |
| | bottom | 125 (–5.1) | **729 (–3.2)** | 243 (–2.8) | 189 (–2.7) | 27 (–2.0) | **25 (–2.0)** | 486 (–1.9) | **16 (–1.8)** | 135 (–1.6) | 54 (–1.5) |
| $C_2^3$ | top | **961 (5.8)** | **25 (5.8)** | **49 (5.7)** | **841 (5.0)** | **529 (4.8)** | **361 (4.6)** | **9 (3.7)** | **289 (3.4)** | **121 (2.9)** | **36 (2.6)** |
| | bottom | 343 (–6.4) | 31 (–5.9) | 29 (–4.7) | 23 (–4.4) | 19 (–4.4) | **625 (–4.1)** | 17 (–3.3) | 11 (–3.1) | **81 (–3.0)** | **16 (–2.6)** |

Bold numbers correspond to indices that are squares

distinguish between the two groups, since the possible decomposition types for a given prime are identical for both $D_4$- and $Q_8$-extensions. In particular, for primes $p \leq 17$, we checked that all the decomposition types that are *not* ruled out by Proposition 3.34 can occur for some number field in LMFDB.[3]

Proposition 3.34 describes the decomposition types that cannot occur in either $D_4$- or $Q_8$-extensions. We first present two lemmas needed for the proof.

**Lemma 3.32** *Let $p \equiv 1 \pmod 4$ be a prime. Then, any degree 8 Galois extension of $\mathbb{Q}_p$ has Galois group either $C_8$ or $C_4 \times C_2$.*

*Proof* The proof is similar to that of Lemma 3.14. Let $K/\mathbb{Q}_p$ be an octic Galois extension with Galois group $G$. Since $(p, 2) = 1$, the extension is tamely ramified, and the argument in Lemma 3.14 shows that $G \simeq C_e \times C_f$. Thus, $G$ must be abelian. Moreover, $G$ cannot be isomorphic to $C_2^3$, since at least one of $e$ or $f$ must be 4, and hence, the exponent of $G$ must be at least 4. □

By a similar argument, we also obtain the following:

**Lemma 3.33** *Let $p$ be an odd prime. Then, any degree 8 Galois extension of $\mathbb{Q}_p$ with decomposition type $(e, f) = (2, 4)$ is abelian.*

*Proof* If $p \equiv 1 \pmod 4$, then Lemma 3.32 shows the claim. If $p \equiv 3 \pmod 4$, the extension is still tamely ramified, and the same argument as in Lemma 3.14 shows that the short exact sequence must split, and hence, $G$ must be abelian. □

**Proposition 3.34** *Let $K/\mathbb{Q}$ be an octic Galois extension with Galois group $D_4$ or $Q_8$. Let $p$ be a rational prime with decomposition type $(e, f, g)$ in $K$. Then:*

1. *For all prime $p$, $(e, f, g) \notin \{(1, 8, 1), (2, 4, 1)\}$.*
2. *If $p \equiv 1 \pmod 4$, then $(e, f, g) \notin \{(4, 2, 1), (8, 1, 1)\}$.*
3. *If $p \equiv 3 \pmod 4$, then $(e, f, g) \notin \{(4, 1, 2), (8, 1, 1)\}$.*

*Proof* First of all, we have $(e, f, g) \neq (1, 8, 1)$ for any prime $p$, since $C_8$ cannot arise as a quotient of either $D_4$ or $Q_8$. Assume that a prime $p$ has decomposition type $(e, f, g) = (2, 4, 1)$. Then, the decomposition group $D_p$ has a normal subgroup $I_p$ of order $e = 2$, whose quotient has to be cyclic (isomorphic to $\mathrm{Gal}(\mathbb{F}_{p^4}/\mathbb{F}_p) \simeq C_4$). However, the only normal subgroup of order 2 in $D_4$ (resp. $Q_8$) is the center, i.e., $\langle r^2 \rangle$ where $r$ is the rotation by 90 degree (resp. $\langle -1 \rangle$), and the quotients by the subgroup are isomorphic to $C_2^2$, since the square of any element belongs to the center. This proves $(e, f, g) \notin \{(1, 8, 1), (2, 4, 1)\}$ for any $p$.

For odd $p$, a corresponding local extension $K_\mathfrak{p}/\mathbb{Q}_p$ has to be tame and so both $I_\mathfrak{p}$ and $D_\mathfrak{p}/I_\mathfrak{p}$ are cyclic of order $e$ and $f$, respectively, which excludes the possibilities $(1, 8, 1)$ and $(8, 1, 1)$. For $p \equiv 1 \pmod 4$, $(e, f, g) \neq (4, 2, 1)$ follows from Lemma 3.32. For $p \equiv 3 \pmod 4$, Lemma 3.33 shows $(e, f, g) \neq (2, 4, 1)$. If $p \equiv 3 \pmod 4$ has a decomposition type of $(e, f, g) = (4, 1, 2)$, then a corresponding local extension is totally and tamely ramified quartic extension, so Proposition 2.3 shows $p \equiv 1 \pmod 4$ and gives a contradiction. □

Since the proposition above applies to both nonabelian octic extensions, it does not help distinguish between the two cases. Nonetheless, we observe that the distributions of

[3] See `src/verify_galois.sage` of the GitHub repository.

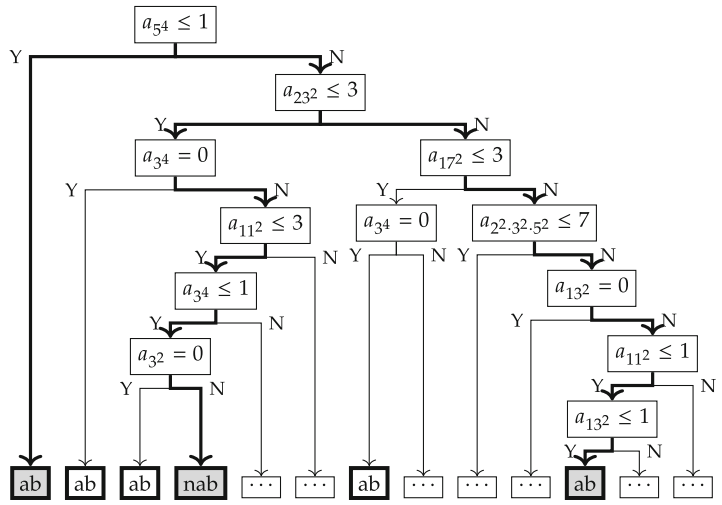



**Fig. 7** A decision tree predicts commutativity of Galois groups of octic fields from zeta coefficients up to $a_{1000}(K)$

square-indexed coefficients differ significantly between $D_4$- and $Q_8$-extensions (Table 27). This discrepancy partially explains why the decision tree model achieved high accuracy using square-indexed zeta coefficients. It would be interesting to derive an explicit formula for the probability $\mathbb{P}[a_{p^2}(K) = a \mid \mathrm{Gal}(K/\mathbb{Q}) \simeq G]$ and $\mathbb{P}[\mathrm{Gal}(K/\mathbb{Q}) \simeq G \mid a_{p^2}(K) = a]$ over all Galois extensions $K$ with Galois group $G$ (ordered by their discriminants or heights) in terms of $a$ and $p$.

**3.6.2.1 Polynomial coefficients** With polynomial coefficients, the decision tree model yields a high accuracy of 98.76%. By examining the first few nodes, we reach a conclusion similar to the sextic case (Sect. 3.4.2): The model tends to classify a given octic field as a $D_4$-extension when the coefficients are "small." For example, among the 62442 fields in the training set, 22307 (resp. 40135) have Galois group $D_4$ (resp. $Q_8$). The first node of the decision tree checks whether $c_4 \le 167573$; 20953 $D_4$-extensions (93.93% of 22307) satisfy this condition, while 39594 $Q_8$-extensions (98.65% of 40135) satisfy the opposite inequality $c_4 > 167573$. We also computed the mean and median of the polynomial coefficients $c_0, c_1, \ldots, c_7$ for each Galois group and found that the magnitudes $|c_0|, |c_1|, \ldots, |c_6|$ are significantly larger for $Q_8$-extensions (Table 14).

In contrast, the logistic regression model achieves only 68.21% (resp. 84.33%) accuracy with polynomial coefficients (resp. normalized polynomial coefficients); see Table 30.

### *3.6.3 Distinguishing between abelian and nonabelian extensions*

We also conducted experiments on distinguishing between abelian ($C_8$, $C_4 \times C_2$, $C_2^3$) and nonabelian ($D_4$, $Q_8$) extensions as a binary classification task. Using 1000 zeta coefficients, the decision tree model achieved an accuracy of 96.58%. We observed that the model primarily focuses on square-indexed coefficients (Fig. 7). By combining Propositions 3.27–3.29 and 3.34, one can derive criteria that distinguish abelian from nonabelian extensions.

**Corollary 3.35** *Let $K/\mathbb{Q}$ be an octic Galois extension, and let $p$ be a rational prime.*

1. *If $a_{p^4}(K) = 0$, then $\mathrm{Gal}(K/\mathbb{Q})$ is a $C_8$-extension (hence abelian).*

**Table 14** Mean (left) and median (right) of polynomial coefficients of nonabelian octic extensions

| | $c_0$ | | $c_1$ | | $c_2$ | | $c_3$ | | $c_4$ | | $c_5$ | | $c_6$ | | $c_7$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $D_4$ | $1.3 \cdot 10^{11}$ | $1.2 \cdot 10^{5}$ | $5.5 \cdot 10^{8}$ | 0 | $-1.7 \cdot 10^{8}$ | $1.8 \cdot 10^{3}$ | $2.0 \cdot 10^{6}$ | 0 | $3.3 \cdot 10^{5}$ | $1.7 \cdot 10^{3}$ | 78.4 | 0 | −104.6 | 4.0 | −1.2 | 0 |
| $Q_8$ | $2.3 \cdot 10^{18}$ | $1.5 \cdot 10^{14}$ | $-2.3 \cdot 10^{11}$ | 0 | $-1.8 \cdot 10^{11}$ | $-1.0 \cdot 10^{9}$ | $1.1 \cdot 10^{8}$ | 0 | $7.9 \cdot 10^{9}$ | $7.7 \cdot 10^{8}$ | $2.4 \cdot 10^{3}$ | 0 | $-8.8 \cdot 10^{3}$ | $-1.3 \cdot 10^{4}$ | −0.6 | 0 |

2. *For* $p \equiv 1 \pmod 4$, *if* $a_{p^4}(K) = 1$ *or* $a_{p^2}(K) = 1$, *then* $\mathrm{Gal}(K/\mathbb{Q})$ *is abelian.*
3. *If* $p \equiv 3 \pmod 4$, $a_{p^4}(K) = 1$, *and* $a_{p^2}(K) > 0$, *then* $\mathrm{Gal}(K/\mathbb{Q})$ *is nonabelian.*

*Proof* Let $(e, f, g)$ be the decomposition type of $p$ in $K$.

1. If $a_{p^4}(K) = 0$, Table 12 shows that we must have $(e, f, g) = (1, 8, 1)$, which implies that $\mathrm{Gal}(K/\mathbb{Q})$ is isomorphic to the decomposition group ($g = 1$), and the decomposition group is isomorphic to $\mathrm{Gal}(\mathbb{F}_{p^8}/\mathbb{F}_p) \simeq C_8$ ($e = 1$). This proves the claim. (Note that there is no $C_8$-extension of $\mathbb{Q}$ with $p = 2$ which inerts or unramifies by Proposition 2.6.)
2. From Table 12, the condition $a_{p^2}(K) = 1$ or $a_{p^4}(K) = 1$ implies that the possible decomposition types of $p$ are (1, 8, 1), (2, 4, 1), (4, 2, 1), or (8, 1, 1). By Proposition 3.34, all of these possibilities are excluded for nonabelian extensions, and thus the extension must be abelian.
3. The condition $a_{p^4}(K) = 1$ implies that the possible decomposition types of $p$ are (2, 4, 1), (4, 2, 1), or (8, 1, 1). Also, $a_{p^2}(K) > 0$ implies that the decomposition type cannot be (2, 4, 1), and hence, we must have $(e, f, g) \in \{(4, 2, 1), (8, 1, 1)\}$. However, by Propositions 3.27–3.29, such decomposition types cannot occur for abelian octic extensions when $p \equiv 3 \pmod 4$, so $\mathrm{Gal}(K/\mathbb{Q})$ must be nonabelian.

□

*Example 3.36* (abelian, sufficient condition) Assume that an octic Galois extension $K/\mathbb{Q}$ follows the leftmost bold branch of Fig. 7, so that $a_{5^4}(K) \leq 1$. Then, $a_{5^4}(K) = 0$ or 1 and Corollary 3.35 shows that the extension must be abelian.

Similarly, consider the right-most bold branch of Fig. 7. If an octic Galois extension satisfies all the inequalities along this path, then $a_{13^2}(K) = 1$, which again forces $K$ to be an abelian extension by Corollary 3.35.

*Example 3.37* (nonabelian, sufficient condition) Consider the middle bold branch of Figure 7. If an octic Galois extension $K/\mathbb{Q}$ satisfies all the corresponding inequalities, then we have $a_{3^4}(K) = 1$ and $a_{3^2}(K) > 0$. Then, Corollary 3.35 implies that $K$ must be a nonabelian extension, i.e., the branch always yields a correct prediction.

In contrast, we found that logistic regression models perform poorly in distinguishing between abelian and nonabelian extensions, whether using zeta coefficients or polynomial coefficients. The model achieved only 76.27% (resp. 75.83%) accuracy with 1000 zeta coefficients (resp. with standard normalization).

Tables 29, 30, and 31 in the Appendix summarize the results for octic extensions.

## 4 Conclusion

In this paper, we apply interpretable machine learning (ML) methods to the problem of classifying Galois groups and prove several new theorems based on the resulting interpretations. These results offer different perspectives that emerged only through the analysis of ML outputs. Our study highlights the potential of ML not merely as a predictive tool, but as a means to generate mathematical conjectures and contribute to the discovery of rigorous mathematical theorems. We hope that this emerging methodology will continue to gain momentum and contribute to meaningful advances in mathematics.

We conclude this paper by proposing possible directions for future research on Galois groups. First, it is natural to ask if there are similar criteria in other degrees for determining Galois groups from Dedekind zeta coefficients. For example, Yang-Hui He asked whether the following statement is true: *For a given degree $d$, there exist (possibly infinitely many) pairs $(n, a)$ of natural numbers such that, for any Galois extension $K/\mathbb{Q}$ of degree $d$, the condition $a_n(K) = a$ implies that $K$ is a cyclic extension.*

Next, one can consider more general extensions—including non-Galois extensions—and attempt to predict the Galois groups of their splitting fields. For example, up to isomorphism there are five possible Galois groups of quartic fields ($C_4$, $C_2 \times C_2$, $D_4$, $A_4$, and $S_4$), and it would be interesting to see whether simple ML algorithms, such as decision trees, can still classify these groups effectively. One can also study extensions of number fields whose base field is not necessarily $\mathbb{Q}$, using the zeta function coefficients of both the base and extension fields to predict the corresponding Galois groups. Ultimately, it would be very interesting if these ML experiments led to new efficient algorithms for computing Galois groups of number fields.

**Acknowledgements**
We thank Jordan Ellenberg, Tony Feng, Yang-Hui He, Hyukpyo Hong, and Sug Woo Shin for helpful discussions. We are also grateful to John Jones for answering questions on local fields in `LMFDB` and to David Roe for helping us use the `lmfdb-lite` library. We also thank ChatGPT for assisting us in writing LaTeX scripts for the decision tree figures and in searching the literature.

**Data availability**
All the data used in the experiments come from `LMFDB` [37] with further computations using SageMath [43]. Anyone can get the same data using `nf_query.sage` file in the GitHub repository.

## Declarations

**Conflict of interest**
The authors declare no conflicts of interest regarding this manuscript

**Author details**
[1]Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA, [2]„ Korea Institute for Advanced Study, Seoul 02455, Republic of Korea, [3]Department of Mathematics, University of California—Berkeley, Berkeley, CA 94720, USA.

## A Tables and confusion matrices

### A.1 Quartic and nonic fields

See Tables 15, 16.

**Table 15** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of quartic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of $C_4$ (resp. $C_2^2$). The majority baseline is 88.04% ($C_2^2$)

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 100.00 ± 0.00 | 99.99 ± 0.02 | 100.00 ± 0.00 | 99.98 ± 0.04 | 99.99 ± 0.02 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | squares (31) | 99.99 ± 0.01 | 99.97 ± 0.02 | 99.97 ± 0.03 | 99.95 ± 0.04 | 99.96 ± 0.03 |
| | | | | | 99.99 ± 0.01 | 100.00 ± 0.00 | 99.99 ± 0.00 |
| | | primes (168) | 91.11 ± 0.09 | 79.27 ± 0.30 | 62.64 ± 0.94 | 63.69 ± 0.54 | 63.16 ± 0.69 |
| | | | | | 95.06 ± 0.06 | 94.84 ± 0.11 | 94.95 ± 0.05 |
| | | prime powers (193) | 100.00 ± 0.00 | 99.99 ± 0.02 | 100.00 ± 0.00 | 99.98 ± 0.04 | 99.99 ± 0.02 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | PC (4) | 95.16 ± 0.18 | 87.82 ± 0.32 | 80.73 ± 0.93 | 78.18 ± 0.61 | 79.43 ± 0.65 |
| | | | | | 97.05 ± 0.06 | 97.46 ± 0.16 | 97.25 ± 0.11 |
| LR | ZC | all (1000) | 99.60 ± 0.03 | 98.88 ± 0.09 | 98.76 ± 0.19 | 97.93 ± 0.19 | 98.34 ± 0.09 |
| | | | | | 99.72 ± 0.03 | 99.83 ± 0.03 | 99.78 ± 0.01 |
| | | all (normalized, 1000) | 99.41 ± 0.03 | 98.27 ± 0.09 | 98.26 ± 0.02 | 96.77 ± 0.16 | 97.51 ± 0.14 |
| | | | | | 99.56 ± 0.02 | 99.77 ± 0.03 | 99.66 ± 0.02 |
| | | squares (31) | 98.05 ± 0.06 | 92.29 ± 0.25 | 98.85 ± 0.08 | 84.72 ± 0.05 | 91.24 ± 0.27 |
| | | | | | 97.96 ± 0.08 | 99.87 ± 0.01 | 98.91 ± 0.04 |
| | | primes (168) | 88.04 ± 0.17 | 50.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| | | | | | 88.04 ± 0.00 | 100.00 ± 0.00 | 0.00 ± 0.00 |
| | | prime powers (193) | 99.65 ± 0.02 | 98.92 ± 0.01 | 99.16 ± 0.22 | 97.94 ± 0.06 | 98.55 ± 0.08 |
| | | | | | 99.72 ± 0.01 | 99.89 ± 0.03 | 99.80 ± 0.01 |
| | | PC (4) | 22.45 ± 0.57 | 44.34 ± 0.43 | 10.53 ± 0.17 | 73.12 ± 1.41 | 18.40 ± 0.30 |
| | | | | | 80.98 ± 0.66 | 15.56 ± 0.86 | 26.10 ± 1.23 |
| | | PC (normalized, 4) | 88.06 ± 0.18 | 50.08 ± 0.05 | 100.00 ± 0.00 | 0.16 ± 0.09 | 0.32 ± 0.18 |
| | | | | | 88.06 ± 0.18 | 100.00 ± 0.00 | 93.65 ± 0.10 |

**Table 16** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of nonic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of $C_9$ (resp. $C_3^2$). The majority baseline is 77.57% ($C_3^2$)

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | cubes (10) | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | primes (168) | 92.50 ± 0.66 | 88.60 ± 1.68 | 84.27 ± 3.78 | 81.57 ± 4.21 | 82.76 ± 2.16 |
| | | | | | 94.75 ± 1.25 | 95.62 ± 1.23 | 95.17 ± 0.51 |
| | | prime powers (193) | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | PC (9) | 94.15 ± 2.58 | 92.16 ± 3.48 | 86.19 ± 6.74 | 88.38 ± 5.38 | 87.18 ± 5.29 |
| | | | | | 96.48 ± 1.81 | 95.94 ± 2.12 | 96.20 ± 1.70 |
| LR | ZC | all (1000) | 97.16 ± 0.92 | 94.66 ± 1.67 | 97.24 ± 3.90 | 90.01 ± 3.70 | 93.36 ± 1.93 |
| | | | | | 97.10 ± 1.38 | 99.30 ± 0.89 | 98.18 ± 0.63 |
| | | all (normalized, 1000) | 78.51 ± 2.84 | 67.19 ± 2.42 | 52.57 ± 4.16 | 46.52 ± 4.72 | 49.13 ± 2.91 |
| | | | | | 84.91 ± 3.15 | 87.86 ± 2.29 | 86.32 ± 2.16 |
| | | cubes (10) | 95.97 ± 0.52 | 91.12 ± 0.48 | 99.59 ± 0.82 | 82.33 ± 0.98 | 90.14 ± 0.62 |
| | | | | | 95.13 ± 0.77 | 99.90 ± 0.21 | 97.45 ± 0.42 |
| | | primes (168) | 73.70 ± 1.80 | 56.77 ± 2.15 | 37.88 ± 9.01 | 26.13 ± 2.87 | 30.73 ± 4.79 |
| | | | | | 80.41 ± 2.34 | 87.42 ± 2.23 | 83.73 ± 1.35 |
| | | prime powers (193) | 96.52 ± 0.77 | 93.57 ± 1.05 | 96.13 ± 1.93 | 88.16 ± 2.12 | 91.94 ± 1.30 |
| | | | | | 96.59 ± 1.00 | 98.98 ± 0.56 | 97.77 ± 0.55 |
| | | PC (9) | 84.52 ± 2.95 | 68.65 ± 2.50 | 83.28 ± 5.46 | 39.65 ± 5.08 | 53.44 ± 4.55 |
| | | | | | 84.71 ± 3.31 | 97.65 ± 1.01 | 90.69 ± 1.98 |
| | | PC (normalized, 9) | 85.46 ± 3.37 | 68.72 ± 3.27 | 95.69 ± 3.81 | 37.93 ± 6.79 | 53.86 ± 6.74 |
| | | | | | 84.54 ± 3.81 | 99.50 ± 0.45 | 91.36 ± 2.15 |

### A.2 Sextic and decic fields

See Tables 17, 18, 19, 20, 21, 22, 23, 24, 25, 26

**Table 17** Distribution of $a_{p^2}(K)$ with Galois group $G \in \{C_6, S_3\}$ for sextic extensions and $p = 2, 3, 5, 7, 17, 19$

| | $a_{2^2}$ | | $a_{3^2}$ | | $a_{5^2}$ | | $a_{7^2}$ | | $a_{17^2}$ | | $a_{19^2}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ |
| 0 | 0.78 | 0.04 | 0.53 | 0.22 | 0.75 | 0.27 | 0.50 | 0.33 | 0.72 | 0.31 | 0.64 | 0.36 |
| 1 | – | 0.26 | 0.29 | 0.11 | – | 0.06 | 0.23 | – | – | 0.01 | 0.07 | – |
| 3 | 0.08 | 0.16 | 0.10 | 0.31 | 0.11 | 0.42 | 0.17 | 0.42 | 0.15 | 0.50 | 0.17 | 0.45 |
| 6 | 0.09 | 0.50 | 0.05 | 0.30 | 0.05 | 0.17 | 0.02 | 0.13 | 0.02 | 0.06 | 0.01 | 0.05 |
| 21 | 0.05 | 0.03 | 0.03 | 0.06 | 0.10 | 0.08 | 0.07 | 0.12 | 0.11 | 0.12 | 0.12 | 0.15 |

**Table 18** Distribution of Galois group $G \in \{C_6, S_3\}$ for sextic extensions and given $a_{p^2}(K)$ with $p = 2, 3, 5, 7, 17, 19$.

| | $a_4$ | | $a_9$ | | $a_{25}$ | | $a_{49}$ | | $a_{289}$ | | $a_{361}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ |
| 0 | 0.75 | 0.25 | 0.27 | 0.73 | 0.30 | 0.70 | 0.19 | 0.81 | 0.27 | 0.73 | 0.22 | 0.78 |
| 1 | – | 1.00 | 0.30 | 0.70 | – | 1.00 | 1.00 | – | - - | 1.00 | 1.00 | – |
| 3 | 0.07 | 0.93 | 0.05 | 0.95 | 0.04 | 0.96 | 0.06 | 0.94 | 0.05 | 0.95 | 0.05 | 0.95 |
| 6 | 0.03 | 0.97 | 0.02 | 0.98 | 0.04 | 0.96 | 0.03 | 0.97 | 0.04 | 0.96 | 0.04 | 0.96 |
| 21 | 0.18 | 0.82 | 0.07 | 0.93 | 0.15 | 0.85 | 0.09 | 0.91 | 0.12 | 0.88 | 0.11 | 0.89 |

**Table 19** Distribution of $a_{p^3}(K)$ with Galois group $G \in \{C_6, S_3\}$ for sextic extensions and $p = 2, 3, 5, 7, 17, 19$

| | $a_{2^3}$ | | $a_{3^3}$ | | $a_{5^3}$ | | $a_{7^3}$ | | $a_{17^3}$ | | $a_{19^3}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ |
| 0 | 0.37 | 0.42 | 0.32 | 0.31 | 0.45 | 0.47 | 0.37 | 0.39 | 0.51 | 0.51 | 0.47 | 0.44 |
| 1 | 0.26 | – | 0.35 | 0.09 | 0.12 | – | 0.18 | – | 0.03 | – | 0.06 | – |
| 2 | 0.22 | 0.04 | 0.19 | 0.23 | 0.28 | 0.27 | 0.24 | 0.33 | 0.33 | 0.31 | 0.29 | 0.36 |
| 4 | – | – | 0.06 | 0.01 | – | – | 0.11 | 0.04 | – | – | 0.04 | 0.01 |
| 10 | 0.09 | 0.50 | 0.05 | 0.30 | 0.05 | 0.17 | 0.02 | 0.13 | 0.02 | 0.06 | 0.01 | 0.05 |
| 56 | 0.05 | 0.04 | 0.03 | 0.06 | 0.10 | 0.08 | 0.07 | 0.12 | 0.11 | 0.12 | 0.12 | 0.15 |

**Table 20** Distribution of Galois group $G \in \{C_6, S_3\}$ for sextic extensions and given $a_{p^3}(K)$ with $p = 2, 3, 5, 7, 17, 19$

| | $a_{2^3}$ | | $a_{3^3}$ | | $a_{5^3}$ | | $a_{7^3}$ | | $a_{17^3}$ | | $a_{19^3}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ | $C_6$ | $S_3$ |
| 0 | 0.12 | 0.88 | 0.14 | 0.86 | 0.13 | 0.87 | 0.13 | 0.87 | 0.13 | 0.87 | 0.14 | 0.86 |
| 1 | 1.00 | – | 0.37 | 0.63 | 1.00 | – | 1.00 | – | 1.00 | – | 1.00 | - |
| 2 | 0.46 | 0.54 | 0.11 | 0.89 | 0.14 | 0.86 | 0.10 | 0.90 | 0.14 | 0.86 | 0.11 | 0.89 |
| 4 | – | – | 0.48 | 0.52 | – | – | 0.32 | 0.68 | – | – | 0.39 | 0.61 |
| 10 | 0.03 | 0.97 | 0.02 | 0.98 | 0.04 | 0.96 | 0.03 | 0.97 | 0.04 | 0.96 | 0.04 | 0.96 |
| 56 | 0.18 | 0.82 | 0.07 | 0.93 | 0.15 | 0.85 | 0.09 | 0.91 | 0.12 | 0.88 | 0.11 | 0.89 |

**Table 21** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of sextic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of $C_6$ (resp. $S_3$). The majority baseline is 86.55% ($S_3$)

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 98.96 ± 0.02 | 97.95 ± 0.07 | 95.75 ± 0.18 | 96.57 ± 0.14 | 96.16 ± 0.08 |
| | | | | | 99.47 ± 0.03 | 99.33 ± 0.02 | 99.40 ± 0.01 |
| | | squares and cubes (38) | 99.31 ± 0.04 | 98.60 ± 0.13 | 97.24 ± 0.32 | 97.64 ± 0.28 | 97.44 ± 0.15 |
| | | | | | 99.63 ± 0.05 | 99.57 ± 0.05 | 99.60 ± 0.03 |
| | | primes (168) | 89.66 ± 0.17 | 80.38 ± 0.54 | 60.33 ± 0.69 | 67.69 ± 1.11 | 63.79 ± 0.63 |
| | | | | | 94.88 ± 0.22 | 93.08 ± 0.13 | 93.97 ± 0.10 |
| | | prime powers (193) | 99.04 ± 0.04 | 98.21 ± 0.09 | 95.81 ± 0.37 | 97.09 ± 0.22 | 96.45 ± 0.13 |
| | | | | | 99.55 ± 0.04 | 99.34 ± 0.06 | 99.44 ± 0.02 |
| | | PC (6) | 97.56 ± 0.08 | 94.65 ± 0.18 | 91.13 ± 0.35 | 90.67 ± 0.36 | 90.90 ± 0.26 |
| | | | | | 98.55 ± 0.06 | 98.63 ± 0.06 | 98.59 ± 0.05 |
| LR | ZC | all (1000) | 98.40 ± 0.07 | 95.90 ± 0.19 | 95.47 ± 0.38 | 92.48 ± 0.40 | 93.95 ± 0.24 |
| | | | | | 98.84 ± 0.06 | 99.32 ± 0.07 | 99.08 ± 0.04 |
| | | all (normalized, 1000) | 97.80 ± 0.10 | 94.35 ± 0.21 | 93.75 ± 0.45 | 89.63 ± 0.40 | 91.65 ± 0.33 |
| | | | | | 98.40 ± 0.07 | 99.07 ± 0.08 | 98.73 ± 0.06 |
| | | squares and cubes (38) | 96.51 ± 0.13 | 89.79 ± 0.50 | 92.50 ± 0.18 | 80.60 ± 1.02 | 86.14 ± 0.59 |
| | | | | | 97.04 ± 0.15 | 98.98 ± 0.03 | 98.00 ± 0.08 |
| | | primes (168) | 91.34 ± 0.15 | 69.36 ± 0.29 | 91.56 ± 0.67 | 39.27 ± 0.58 | 54.97 ± 0.56 |
| | | | | | 91.33 ± 0.19 | 99.44 ± 0.04 | 95.21 ± 0.09 |
| | | prime powers (193) | 98.48 ± 0.03 | 95.99 ± 0.13 | 95.98 ± 0.29 | 92.58 ± 0.30 | 94.25 ± 0.07 |
| | | | | | 98.85 ± 0.05 | 99.40 ± 0.05 | 99.12 ± 0.02 |
| | | PC (6) | 26.22 ± 4.75 | 50.69 ± 2.70 | 13.58 ± 0.74 | 84.15 ± 12.45 | 23.36 ± 1.54 |
| | | | | | 90.48 ± 6.47 | 17.22 ± 7.35 | 27.97 ± 10.44 |
| | | PC (normalized, 6) | 86.73 ± 0.50 | 50.72 ± 1.22 | 99.26 ± 1.48 | 1.44 ± 2.45 | 2.73 ± 4.59 |
| | | | | | 86.71 ± 0.47 | 99.99 ± 0.01 | 92.88 ± 0.26 |

**Table 22** Distribution of $a_{p^2}(K)$ with Galois group $G \in \{C_{10}, D_5\}$ for decic extensions and $p = 2, 3, 5, 7, 11, 19$.

| | $a_{2^2}$ | | $a_{3^2}$ | | $a_{5^2}$ | | $a_{7^2}$ | | $a_{11^2}$ | | $a_{19^2}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ |
| 0 | 0.92 | 0.30 | 0.91 | 0.33 | 0.66 | 0.38 | 0.86 | 0.41 | 0.39 | 0.40 | 0.92 | 0.41 |
| 1 | – | – | – | – | 0.18 | 0.06 | – | – | 0.39 | – | – | 0.02 |
| 3 | – | – | – | – | 0.06 | 0.01 | – | – | 0.16 | 0.04 | – | – |
| 5 | 0.03 | 0.34 | 0.04 | 0.38 | 0.04 | 0.38 | 0.06 | 0.43 | 0.03 | 0.44 | 0.04 | 0.46 |
| 15 | 0.03 | 0.32 | 0.02 | 0.24 | 0.02 | 0.13 | 0.02 | 0.11 | 0.00 | 0.06 | 0.00 | 0.05 |
| 55 | 0.02 | 0.04 | 0.03 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.03 | 0.06 | 0.04 | 0.06 |

**Table 23** Distribution of Galois group $G \in \{C_{10}, D_5\}$ for decic extensions and given $a_{p^2}(K)$ with $p = 2, 3, 5, 7, 11, 19$

| | $a_{2^2}$ | | $a_{3^2}$ | | $a_{5^2}$ | | $a_{7^2}$ | | $a_{11^2}$ | | $a_{19^2}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ |
| 0 | 0.61 | 0.39 | 0.58 | 0.42 | 0.47 | 0.53 | 0.52 | 0.48 | 0.33 | 0.67 | 0.53 | 0.47 |
| 1 | – | – | – | – | 0.61 | 0.39 | – | – | 1.00 | – | – | 1.00 |
| 3 | – | – | – | – | 0.77 | 0.23 | – | – | 0.67 | 0.33 | – | – |
| 5 | 0.04 | 0.96 | 0.05 | 0.95 | 0.05 | 0.95 | 0.07 | 0.93 | 0.04 | 0.96 | 0.05 | 0.95 |
| 15 | 0.04 | 0.96 | 0.05 | 0.95 | 0.08 | 0.92 | 0.07 | 0.93 | 0.03 | 0.97 | 0.01 | 0.99 |
| 55 | 0.25 | 0.75 | 0.29 | 0.71 | 0.34 | 0.66 | 0.37 | 0.63 | 0.18 | 0.82 | 0.25 | 0.75 |

**Table 24** Distribution of $a_{p^5}(K)$ with Galois group $G \in \{C_{10}, D_5\}$ for decic extensions and $p = 2, 3, 5, 7$

| | $a_{2^5}$ | | $a_{3^5}$ | | $a_{5^5}$ | | $a_{7^5}$ | |
|---|---|---|---|---|---|---|---|---|
| $G$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ |
| 0 | 0.35 | 0.34 | 0.39 | 0.38 | 0.35 | 0.39 | 0.47 | 0.43 |
| 1 | 0.32 | – | 0.22 | – | 0.26 | 0.05 | 0.09 | – |
| 2 | 0.28 | 0.30 | 0.33 | 0.33 | 0.26 | 0.38 | 0.37 | 0.41 |
| 6 | – | – | – | – | 0.06 | 0.01 | – | – |
| 126 | 0.03 | 0.32 | 0.02 | 0.24 | 0.02 | 0.13 | 0.02 | 0.11 |
| 2002 | 0.02 | 0.04 | 0.03 | 0.04 | 0.04 | 0.04 | 0.06 | 0.05 |

**Table 25** Distribution of Galois group $G \in \{C_{10}, D_5\}$ for sextic extensions and given $a_{p^5}(K)$ with $p = 2, 3, 5, 7$

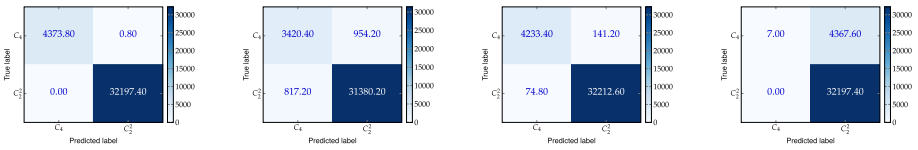
| | $a_{2^5}$ | | $a_{3^5}$ | | $a_{5^5}$ | | $a_{7^5}$ | |
|---|---|---|---|---|---|---|---|---|
| $G$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ | $C_{10}$ | $D_5$ |
| 0 | 0.35 | 0.65 | 0.35 | 0.65 | 0.32 | 0.68 | 0.36 | 0.64 |
| 1 | 1.00 | – | 1.00 | – | 0.72 | 0.28 | 1.00 | – |
| 2 | 0.32 | 0.68 | 0.34 | 0.66 | 0.26 | 0.74 | 0.32 | 0.68 |
| 6 | – | – | – | – | 0.77 | 0.23 | - | - |
| 126 | 0.04 | 0.96 | 0.05 | 0.95 | 0.08 | 0.92 | 0.07 | 0.93 |
| 2002 | 0.25 | 0.75 | 0.29 | 0.71 | 0.34 | 0.66 | 0.37 | 0.63 |

**Table 26** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of decic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of $C_{10}$ (resp. $D_5$). The majority baseline is 66.14% ($D_5$)

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all ($n \leq 1000$) | 97.19 ± 0.47 | 96.92 ± 0.62 | 95.61 ± 0.66 | 96.11 ± 1.18 | 95.85 ± 0.71 |
| | | | | | 98.01 ± 0.57 | 97.73 ± 0.39 | 97.87 ± 0.36 |
| | | squares and quintics (33) | 98.56 ± 0.24 | 98.46 ± 0.26 | 97.60 ± 0.52 | 98.16 ± 0.45 | 97.88 ± 0.35 |
| | | | | | 99.06 ± 0.21 | 98.76 ± 0.30 | 98.91 ± 0.18 |
| | | primes (168) | 85.50 ± 0.43 | 84.77 ± 0.46 | 76.50 ± 0.66 | 82.53 ± 0.67 | 79.40 ± 0.56 |
| | | | | | 90.68 ± 0.48 | 87.02 ± 0.40 | 88.81 ± 0.39 |
| | | prime powers (193) | 98.14 ± 0.43 | 97.85 ± 0.57 | 97.50 ± 0.58 | 96.97 ± 1.02 | 97.23 ± 0.67 |
| | | | | | 98.46 ± 0.50 | 98.73 ± 0.28 | 98.60 ± 0.32 |
| | | PC (10) | 86.71 ± 0.96 | 85.32 ± 1.07 | 79.96 ± 1.58 | 81.06 ± 1.69 | 80.50 ± 1.40 |
| | | | | | 90.24 ± 0.83 | 89.59 ± 0.98 | 89.91 ± 0.77 |
| LR | ZC | all (1000) | 93.81 ± 0.54 | 93.51 ± 0.60 | 89.53 ± 0.77 | 92.57 ± 0.91 | 91.02 ± 0.65 |
| | | | | | 96.12 ± 0.55 | 94.45 ± 0.54 | 95.28 ± 0.46 |
| | | all (normalized, 1000) | 94.70 ± 0.63 | 94.20 ± 0.70 | 91.77 ± 1.01 | 92.65 ± 0.94 | 92.21 ± 0.97 |
| | | | | | 96.23 ± 0.44 | 95.75 ± 0.51 | 95.99 ± 0.47 |
| | | squares and quintics (33) | 94.47 ± 1.23 | 93.91 ± 1.36 | 91.55 ± 1.91 | 92.18 ± 1.89 | 91.86 ± 1.78 |
| | | | | | 95.98 ± 0.99 | 95.64 ± 1.02 | 95.81 ± 0.94 |
| | | primes (168) | 85.88 ± 0.50 | 82.53 ± 0.80 | 83.91 ± 1.83 | 72.16 ± 1.90 | 77.56 ± 1.16 |
| | | | | | 86.71 ± 0.72 | 92.91 ± 0.89 | 89.70 ± 0.31 |
| | | prime powers (193) | 96.57 ± 0.45 | 96.14 ± 0.53 | 95.06 ± 1.23 | 94.82 ± 1.26 | 94.93 ± 0.61 |
| | | | | | 97.36 ± 0.63 | 97.46 ± 0.72 | 97.40 ± 0.35 |
| | | PC (10) | 46.25 ± 4.62 | 56.92 ± 4.79 | 37.71 ± 2.99 | 89.82 ± 8.35 | 53.05 ± 3.96 |
| | | | | | 83.46 ± 10.91 | 24.02 ± 5.97 | 36.84 ± 7.05 |
| | | PC (normalized, 10) | 66.18 ± 0.79 | 50.06 ± 0.11 | 20.00 ± 40.00 | 0.11 ± 0.22 | 0.22 ± 0.45 |
| | | | | | 66.17 ± 0.77 | 100.00 ± 0.00 | 79.64 ± 0.55 |

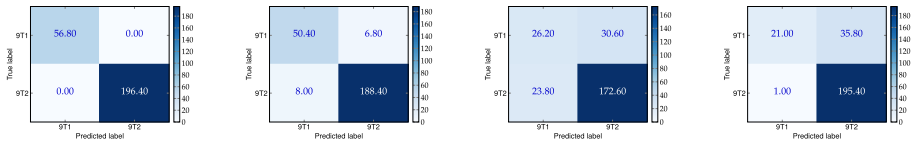
### A.3 Octic fields

See Figs. 8, 9 and Tables 27, 28, 29, 30, 31




*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 8** Confusion matrices for the quartic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients




*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 9** Confusion matrices for the nonic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

**Table 27** Distribution of $a_{p^2}(K)$ for nonabelian octic extensions and $p = 2, 3, 5, 7, 11, 13$. Note that $a_{p^2}(K) \neq 3$ for odd $p$.

| | $a_{2^2}$ | | $a_{3^2}$ | | $a_{5^2}$ | | $a_{7^2}$ | | $a_{11^2}$ | | $a_{13^2}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ |
| 0 | 0.11 | 0.24 | 0.21 | 0.30 | 0.26 | 0.38 | 0.24 | 0.41 | 0.28 | 0.46 | 0.30 | 0.50 |
| 1 | 0.32 | 0.48 | 0.10 | 0.50 | 0.13 | 0.09 | 0.04 | 0.38 | 0.02 | 0.32 | 0.06 | 0.07 |
| 2 | 0.17 | 0.06 | 0.18 | 0.10 | 0.05 | 0.41 | 0.10 | 0.08 | 0.07 | 0.07 | 0.01 | 0.29 |
| 3 | 0.10 | 0.15 | – | – | – | – | – | – | – | – | – | – |
| 4 | 0.17 | 0.03 | 0.37 | 0.03 | 0.41 | 0.05 | 0.50 | 0.05 | 0.52 | 0.07 | 0.51 | 0.06 |
| 10 | 0.11 | 0.01 | 0.11 | 0.03 | 0.09 | 0.02 | 0.07 | 0.02 | 0.04 | 0.02 | 0.04 | 0.02 |
| 36 | 0.02 | 0.03 | 0.04 | 0.03 | 0.05 | 0.05 | 0.06 | 0.06 | 0.07 | 0.07 | 0.08 | 0.06 |

**Table 28** Distribution of Galois group $G \in \{D_4, Q_8\}$ for octic nonabelian extensions and given $a_{p^2}(K)$ with $p = 2, 3, 5, 7, 11, 13$

| | $a_{2^2}$ | | $a_{3^2}$ | | $a_{5^2}$ | | $a_{7^2}$ | | $a_{11^2}$ | | $a_{13^2}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ | $D_4$ | $Q_8$ |
| 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 29** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of abelian octic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, each of three rows correspond to the class of $C_8$, $C_4 \times C_2$, $C_2^3$. The majority baseline is 54.03% ($C_4 \times C_2$).

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 98.61 ± 0.16 | 98.17 ± 0.26 | 96.06 ± 0.84 | 95.79 ± 0.79 | 95.92 ± 0.48 |
| | | | | | 98.66 ± 0.27 | 98.74 ± 0.22 | 98.70 ± 0.14 |
| | | | | | 99.96 ± 0.07 | 100.00 ± 0.00 | 99.98 ± 0.04 |
| | | squares (31) | 98.91 ± 0.18 | 98.66 ± 0.25 | 96.59 ± 0.55 | 97.11 ± 0.63 | 96.85 ± 0.49 |
| | | | | | 99.08 ± 0.22 | 98.88 ± 0.17 | 98.98 ± 0.17 |
| | | | | | 99.94 ± 0.07 | 100.00 ± 0.00 | 99.97 ± 0.03 |
| | | primes (168) | 77.43 ± 0.49 | 74.65 ± 0.45 | 70.84 ± 0.66 | 65.27 ± 1.37 | 67.93 ± 0.78 |
| | | | | | 80.50 ± 0.30 | 81.34 ± 0.81 | 80.91 ± 0.48 |
| | | | | | 75.37 ± 1.41 | 77.34 ± 0.78 | 76.34 ± 0.98 |
| | | prime powers (193) | 98.82 ± 0.16 | 98.48 ± 0.29 | 96.49 ± 0.71 | 96.56 ± 0.94 | 96.52 ± 0.45 |
| | | | | | 98.91 ± 0.32 | 98.89 ± 0.23 | 98.90 ± 0.15 |
| | | | | | 100.00 ± 0.00 | 100.00 ± 0.00 | 100.00 ± 0.00 |
| | | PC (8) | 77.21 ± 0.56 | 77.48 ± 0.60 | 83.02 ± 0.74 | 82.86 ± 0.74 | 82.94 ± 0.51 |
| | | | | | 78.72 ± 0.55 | 79.16 ± 0.43 | 78.94 ± 0.42 |
| | | | | | 71.05 ± 1.07 | 70.42 ± 1.50 | 70.73 ± 1.17 |
| LR | ZC | all (1000) | 93.81 ± 0.13 | 92.04 ± 0.19 | 89.03 ± 1.10 | 83.45 ± 0.65 | 86.14 ± 0.21 |
| | | | | | 93.75 ± 0.31 | 94.94 ± 0.41 | 94.34 ± 0.16 |
| | | | | | 96.46 ± 0.19 | 97.71 ± 0.37 | 97.08 ± 0.24 |
| | | all (normalized, 1000) | 92.88 ± 0.20 | 90.51 ± 0.15 | 85.31 ± 1.25 | 78.99 ± 0.74 | 82.02 ± 0.29 |
| | | | | | 92.56 ± 0.28 | 94.38 ± 0.58 | 93.46 ± 0.22 |
| | | | | | 97.50 ± 0.28 | 98.16 ± 0.28 | 97.83 ± 0.21 |
| | | squares (31) | 85.08 ± 0.73 | 81.87 ± 0.95 | 83.45 ± 2.29 | 64.86 ± 1.13 | 72.98 ± 1.37 |
| | | | | | 86.29 ± 1.12 | 86.36 ± 0.52 | 86.32 ± 0.64 |
| | | | | | 83.83 ± 0.45 | 94.41 ± 1.96 | 88.79 ± 0.66 |
| | | primes (168) | 57.99 ± 0.39 | 46.31 ± 0.45 | 49.11 ± 2.42 | 19.04 ± 0.54 | 27.43 ± 0.80 |
| | | | | | 59.73 ± 0.36 | 80.72 ± 0.64 | 68.65 ± 0.28 |
| | | | | | 54.81 ± 1.33 | 39.18 ± 1.00 | 45.69 ± 1.10 |
| | | prime powers (193) | 94.86 ± 0.28 | 92.99 ± 0.35 | 90.70 ± 1.38 | 84.06 ± 0.91 | 87.24 ± 0.56 |
| | | | | | 94.39 ± 0.42 | 96.15 ± 0.46 | 95.26 ± 0.29 |
| | | | | | 97.93 ± 0.30 | 98.77 ± 0.28 | 98.35 ± 0.22 |
| | | PC (8) | 19.11 ± 2.19 | 32.43 ± 0.42 | 16.45 ± 0.64 | 77.52 ± 20.78 | 26.85 ± 1.94 |
| | | | | | 31.72 ± 25.96 | 0.28 ± 0.25 | 0.55 ± 0.50 |
| | | | | | 34.93 ± 5.83 | 19.51 ± 19.91 | 17.39 ± 13.96 |

**Table 29** continued

| Model | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|
| | PC (normalized, 8) | 64.89 ± 0.44 | 54.28 ± 0.42 | 90.86 ± 0.31 | 34.75 ± 1.16 | 50.26 ± 1.23 |
| | | | | 62.31 ± 0.63 | 88.26 ± 0.47 | 73.05 ± 0.46 |
| | | | | 66.40 ± 1.01 | 39.82 ± 0.45 | 49.78 ± 0.44 |

**Table 30** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing Galois groups of nonabelian octic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of $D_4$ (resp. $Q_8$). The majority baseline is 64.39% ($Q_8$)

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 98.67 ± 0.08 | 98.52 ± 0.10 | 98.25 ± 0.05 | 98.00 ± 0.19 | 98.13 ± 0.10 |
| | | | | | 98.89 ± 0.12 | 99.04 ± 0.04 | 98.96 ± 0.07 |
| | | squares (31) | 99.17 ± 0.09 | 99.09 ± 0.11 | 98.82 ± 0.08 | 98.84 ± 0.20 | 98.83 ± 0.11 |
| | | | | | 99.36 ± 0.12 | 99.35 ± 0.05 | 99.35 ± 0.07 |
| | | primes (168) | 84.50 ± 0.43 | 83.16 ± 0.48 | 78.07 ± 0.42 | 78.52 ± 0.87 | 78.30 ± 0.51 |
| | | | | | 88.08 ± 0.58 | 87.80 ± 0.31 | 87.94 ± 0.38 |
| | | prime powers (193) | 98.85 ± 0.06 | 98.73 ± 0.07 | 98.46 ± 0.24 | 98.32 ± 0.20 | 98.39 ± 0.08 |
| | | | | | 99.07 ± 0.12 | 99.15 ± 0.13 | 99.11 ± 0.05 |
| | | PC (8) | 98.76 ± 0.11 | 98.64 ± 0.13 | 98.29 ± 0.14 | 98.23 ± 0.21 | 98.26 ± 0.16 |
| | | | | | 99.02 ± 0.12 | 99.05 ± 0.07 | 99.04 ± 0.08 |
| LR | ZC | all (1000) | 97.51 ± 0.13 | 97.16 ± 0.16 | 97.07 ± 0.14 | 95.91 ± 0.25 | 96.49 ± 0.19 |
| | | | | | 97.75 ± 0.15 | 98.40 ± 0.07 | 98.08 ± 0.11 |
| | | all (normalized, 1000) | 97.54 ± 0.10 | 97.16 ± 0.14 | 97.21 ± 0.21 | 95.84 ± 0.33 | 96.52 ± 0.12 |
| | | | | | 97.72 ± 0.21 | 98.48 ± 0.10 | 98.09 ± 0.08 |
| | | square (31) | 77.73 ± 0.21 | 72.99 ± 0.18 | 74.76 ± 0.66 | 56.55 ± 0.46 | 64.39 ± 0.31 |
| | | | | | 78.82 ± 0.39 | 89.44 ± 0.30 | 83.79 ± 0.21 |
| | | primes (168) | 71.93 ± 0.23 | 64.78 ± 0.10 | 68.04 ± 0.85 | 39.94 ± 0.43 | 50.33 ± 0.23 |
| | | | | | 72.96 ± 0.43 | 89.62 ± 0.37 | 80.43 ± 0.24 |
| | | prime powers (193) | 97.76 ± 0.04 | 97.39 ± 0.05 | 97.55 ± 0.14 | 96.12 ± 0.15 | 96.83 ± 0.03 |
| | | | | | 97.87 ± 0.09 | 98.67 ± 0.09 | 98.27 ± 0.04 |
| | | PC (8) | 68.21 ± 0.47 | 55.37 ± 0.18 | 99.57 ± 0.30 | 10.78 ± 0.36 | 19.45 ± 0.59 |
| | | | | | 66.95 ± 0.48 | 99.97 ± 0.02 | 80.20 ± 0.35 |
| | | PC (normalized, 8) | 84.33 ± 0.38 | 87.78 ± 0.30 | 69.51 ± 0.61 | 99.76 ± 0.04 | 81.93 ± 0.42 |
| | | | | | 99.82 ± 0.03 | 75.80 ± 0.59 | 86.17 ± 0.38 |

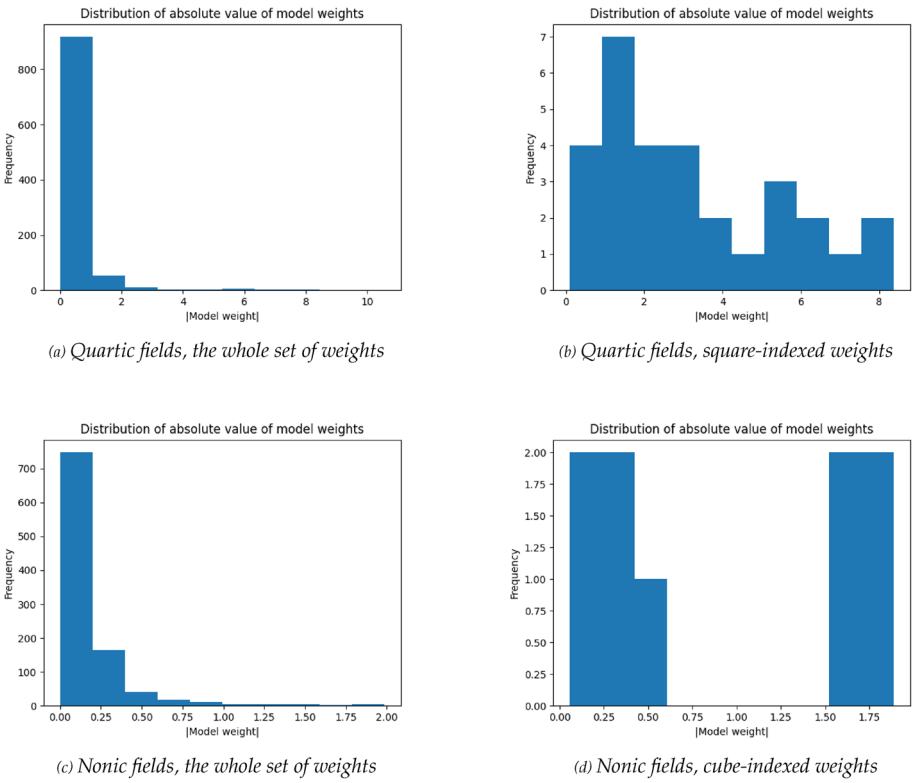
**Table 31** Performance of decision tree (DT) and logistic regression (LR) models on distinguishing abelian and nonabelian octic extensions, using zeta coefficients (ZC) or polynomial coefficients (PC) as features, averaged over 5 different random splits. We report accuracy (ACC), balanced accuracy (BACC), per-class precision, recall, and F1. For the latter three metrics, top (resp. bottom) row corresponds to the class of abelian (resp. nonabelian) extensions. The majority baseline is 67.89% (nonabelian).

| Model | | Feature | ACC | BACC | Precision | Recall | F1 |
|---|---|---|---|---|---|---|---|
| DT | ZC | all (1000) | 96.58 ± 0.10 | 96.14 ± 0.15 | 94.37 ± 0.28 | 94.93 ± 0.37 | 94.65 ± 0.13 |
| | | | | | 97.62 ± 0.20 | 97.35 ± 0.11 | 97.48 ± 0.08 |
| | | squares (31) | 97.20 ± 0.09 | 96.91 ± 0.12 | 95.16 ± 0.29 | 96.10 ± 0.25 | 95.63 ± 0.16 |
| | | | | | 98.17 ± 0.13 | 97.71 ± 0.11 | 97.94 ± 0.07 |
| | | primes (168) | 77.65 ± 0.21 | 75.11 ± 0.33 | 64.06 ± 0.58 | 68.09 ± 0.88 | 66.01 ± 0.40 |
| | | | | | 84.62 ± 0.45 | 82.13 ± 0.40 | 83.35 ± 0.17 |
| | | prime powers (193) | 96.64 ± 0.07 | 96.25 ± 0.13 | 94.36 ± 0.19 | 95.15 ± 0.33 | 94.75 ± 0.14 |
| | | | | | 97.72 ± 0.14 | 97.34 ± 0.10 | 97.53 ± 0.05 |
| | | PC (8) | 84.08 ± 0.26 | 81.92 ± 0.28 | 74.59 ± 0.26 | 75.94 ± 0.42 | 75.26 ± 0.32 |
| | | | | | 88.64 ± 0.29 | 87.89 ± 0.17 | 88.27 ± 0.23 |
| LR | ZC | all (1000) | 76.27 ± 0.25 | 69.61 ± 0.47 | 66.61 ± 0.87 | 51.23 ± 0.92 | 57.91 ± 0.84 |
| | | | | | 79.41 ± 0.25 | 87.99 ± 0.24 | 83.48 ± 0.14 |
| | | all (normalized, 1000) | 75.83 ± 0.26 | 68.92 ± 0.43 | 66.00 ± 0.85 | 49.87 ± 0.78 | 56.81 ± 0.76 |
| | | | | | 78.95 ± 0.26 | 87.98 ± 0.22 | 83.22 ± 0.16 |
| | | squares (31) | 67.88 ± 0.27 | 51.12 ± 0.15 | 46.38 ± 2.25 | 4.88 ± 0.20 | 8.83 ± 0.37 |
| | | | | | 68.63 ± 0.32 | 97.36 ± 0.14 | 80.50 ± 0.19 |
| | | primes (168) | 70.69 ± 0.23 | 57.43 ± 0.29 | 61.94 ± 1.55 | 20.87 ± 0.44 | 31.21 ± 0.62 |
| | | | | | 71.74 ± 0.28 | 94.00 ± 0.27 | 81.37 ± 0.16 |
| | | prime powers (193) | 76.01 ± 0.38 | 69.02 ± 0.65 | 66.54 ± 0.40 | 49.73 ± 1.56 | 56.91 ± 1.07 |
| | | | | | 78.97 ± 0.58 | 88.30 ± 0.29 | 83.37 ± 0.23 |
| | | PC (8) | 66.41 ± 0.26 | 49.51 ± 0.10 | 25.85 ± 1.33 | 2.88 ± 0.11 | 5.18 ± 0.20 |
| | | | | | 67.90 ± 0.31 | 96.13 ± 0.13 | 79.59 ± 0.00 |
| | | PC (normalized, 8) | 68.02 ± 0.31 | 50.01 ± 0.08 | 29.98 ± 11.85 | 0.31 ± 0.20 | 0.61 ± 0.40 |
| | | | | | 68.13 ± 0.33 | 99.71 ± 0.07 | 80.95 ± 0.22 |

# B Figures

## B.1 Quartic and nonic fields

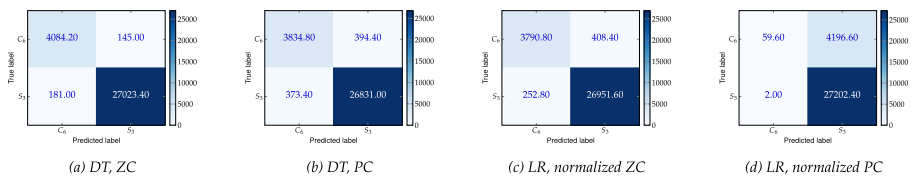
See Fig. 10, 11, 12




*(a) Quartic fields, the whole set of weights* *(b) Quartic fields, square-indexed weights*

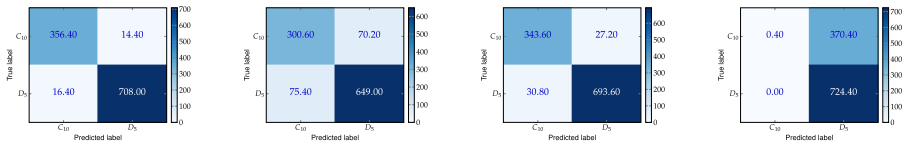
*(c) Nonic fields, the whole set of weights* *(d) Nonic fields, cube-indexed weights*

**Fig. 10** Distribution of the absolute values of the logistic regression model weights trained on normalized zeta coefficients of quartic and nonic Galois extensions




*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 11** Confusion matrices for the sextic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

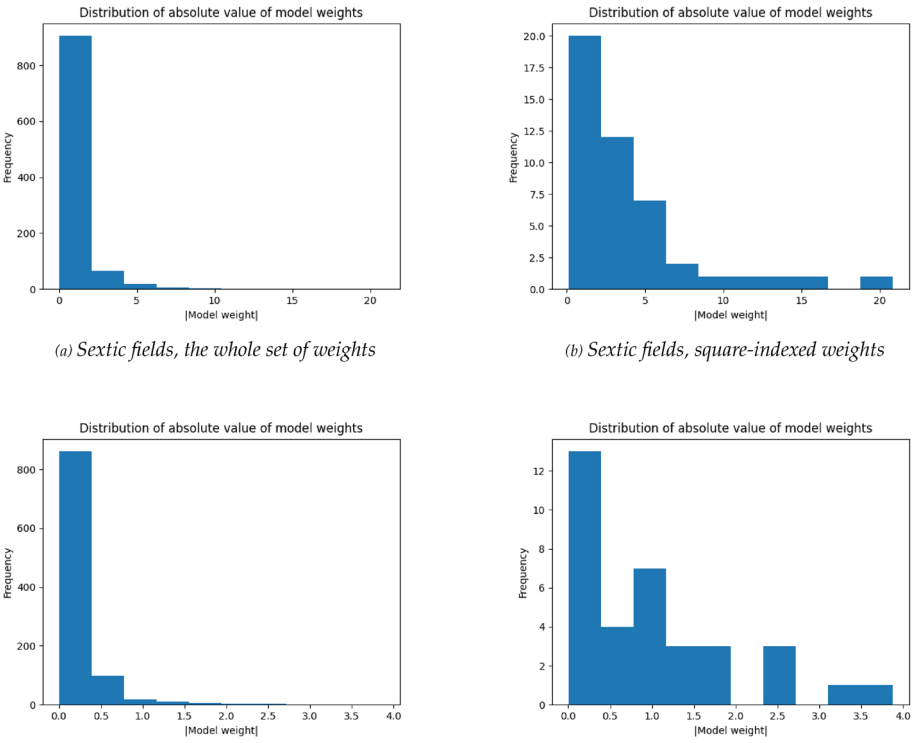



*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 12** Confusion matrices for the decic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

## B.2 Sextic and decic fields

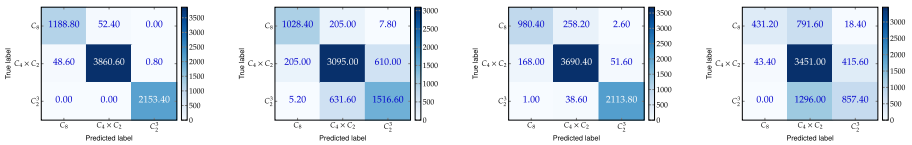
See Figs. 13, 14, 15




*(a) Sextic fields, the whole set of weights* *(b) Sextic fields, square-indexed weights*

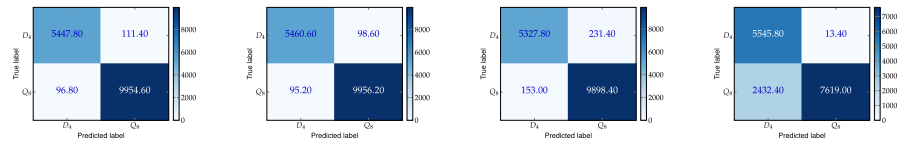
*(c) Decic fields, the whole set of weights* *(d) Decic fields, cube-indexed weights*

**Fig. 13** Distribution of the absolute values of the logistic regression model weights trained on normalized zeta coefficients of sextic and decic Galois extensions




*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 14** Confusion matrices for the nonabelian octic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

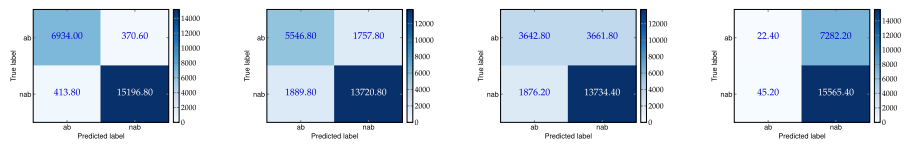



*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 15** Confusion matrices for the nonabelian octic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

### B.3 Octic fields

See Fig. 16




*(a) DT, ZC* *(b) DT, PC* *(c) LR, normalized ZC* *(d) LR, normalized PC*

**Fig. 16** Confusion matrices for the abelian and nonabelian octic Galois field experiments with **a** decision tree, using all ($n \leq 1000$) zeta coefficients, **b** decision tree, using polynomial coefficients, **c** logistic regression, using all ($n \leq 1000$) normalized zeta coefficients, **d** logistic regression, using normalized polynomial coefficients

Received: 13 August 2025 Accepted: 24 March 2026
Published online: 06 July 2026

**References**

1. LMFDB, Canonical defining polynomial for number fields. https://www.lmfdb.org/knowledge/show/nf.polredabs (2026). Accessed 10 March 2026
2. LMFDB, Completeness of *p*-adic field database. https://www.lmfdb.org/padicField/Completeness (2026). Accessed 10 March 2026
3. LMFDB, Source and Acknowledgements for number field database. https://www.lmfdb.org/NumberField/Source (2026). Accessed 10 March 2026
4. Scikit-Learn, Decision Trees. https://scikit-learn.org/stable/modules/tree.html. Accessed 10 March 2026
5. Scikit-Learn, DecisionTreeClassifier. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html (2026). Accessed 10 March 2026
6. Scikit-Learn, LogisticRegression. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html (2026). Accessed 10 March 2026
7. Amir, M., He, Y.-H., Lee, K.-H., Oliver, T., Sultanow, E.: Machine learning class numbers of real quadratic fields. Int. J. Data Sci. Mathematical Sci. **1**(02), 107–134 (2023)

8. Awtrey, C., French, R., Jakes, P., Russell, A.: Irreducible sextic polynomials and their absolute resolvents. Minnesota J. Undergraduate Mathematics **1**, 1 (2015)
9. Awtrey, C., Patane, F.: On the Galois group of a reciprocal even octic polynomial. Comm. Algebra **52**(7), 3018–3026 (2024)
10. Battistoni, F.: On small discriminants of number fields of degree 8 and 9. J. Théor. Nombres Bordeaux **32**(2), 489–501 (2020)
11. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system I The user language. J. Symbolic Comput. **24**(3–4), 235–265 (1997)
12. Buchmann, J., Ford, D., Pohst, M.: Enumeration of quartic fields of small discriminant. Math. Comput. **61**(204), 873–879 (1993)
13. Carifio, J., Halverson, J., Krioukov, D., Nelson, B.D.: Machine learning in the string landscape. J. High Energy Phys. **2017**(9), 157 (2017)
14. Carrillo, B.: Finding Galois splitting models to compute local invariants. J. Number Theory **261**, 241–251 (2024)
15. Cavallo, A.: An elementary computation of the Galois groups of symmetric sextic trinomials. Preprint at arXiv preprint arXiv:1902.00965 (2019)
16. Chen, M.H.W., Chin, A.Y.M., Tan, T.S.: Galois groups of certain even octic polynomials. J. Algebra Appl. **22**(12), 2350263 (2023)
17. Chiou, L.: Elliptic curve 'murmurations' found with AI take flight. Quanta Magazine (2024)
18. Cohen, H.: Number theory: Volume II: Analytic and modern tools, vol. 240. Springer Science & Business Media (2007)
19. Davies, A., Veličković, P., Buesing, L., Blackwell, S., Zheng, D., Tomašev, N., Tanburn, R., Battaglia, P., Blundell, C., Juhász, A., et al.: Advancing mathematics by guiding human intuition with AI. Nature **600**(7887), 70–74 (2021)
20. Douglas, M.R., Lee, K.-H.: Mathematical data science. Advances in Theoretical and Mathematical Physics (2025). To appear
21. Dummit, D.S., Foote, R.M., et al.: Abstract algebra, vol. 3. Wiley Hoboken (2004)
22. Fieker, C., Klüners, J.: Computation of Galois groups of rational polynomials. LMS J. Comput. Math. **17**(1), 141–158 (2014)
23. Greenberg, M.: An elementary proof of the kronecker-weber theorem. Am. Math. Mon. **81**(6), 601–607 (1974)
24. He, Y.-H.: Machine-learning the string landscape. Phys. Lett. B **774**, 564–568 (2017)
25. He, Y.-H., Lee, K.-H., Oliver, T.: Machine-learning number fields. Mathematics, Comput. Geometry Data **2**(1), 49–66 (2022)
26. He, Y.-H., Lee, K.-H., Oliver, T.: Machine-learning the Sato-Tate conjecture. J. Symbolic Comput. **111**, 61–72 (2022)
27. He, Y.-H., Lee, K.-H., Oliver, T.: Machine learning invariants of arithmetic curves. J. Symbolic Comput. **115**, 478–491 (2023)
28. He, Y.-H., Lee, K.-H., Oliver, T., Pozdnyakov, A.: Murmurations of elliptic curves. Experimental Mathematics 1–13 (2024)
29. Iwasawa, K.: On Galois groups of local fields. Trans. Am. Math. Soc. **80**(2), 448–469 (1955)
30. Jones, J.W., Roberts, D.P.: Nonic 3-adic fields. In: International Algorithmic Number Theory Symposium, Springer, pp. 293–308 (2004)
31. Jones, J.W., Roberts, D.P.: A database of local fields. J. Symb. Comput. **41**(1), 80–97 (2006)
32. Jones, J.W., Roberts, D.P.: Octic 2-adic fields. J. Number Theory **128**(6), 1410–1429 (2008)
33. Jones, J.W., Roberts, D.P.: A database of number fields. LMS J. Comput. Math. **17**(1), 595–618 (2014)
34. Klüners, J., Malle, G.: A database for field extensions of the rationals. LMS J. Comput. Math. **4**, 182–196 (2001)
35. Krefl, D., Seong, R.-K.: Machine learning of Calabi-Yau volumes. Phys. Rev. D **96**(6), 066014 (2017)
36. Lee, K.-H.: Data-scientific study of Kronecker coefficients. Experimental Mathematics (2025)
37. LMFDB Collaboration, T.: The L-functions and modular forms database. https://www.lmfdb.org (2025). Accessed 19 July 2025
38. Newton, R.: Explicit local reciprocity for tame extensions. In: Mathematical Proceedings of the Cambridge Philosophical Society vol. 152, Cambridge University Press, pp. 425–454 (2012)
39. Olivier, M.: The computation of sextic fields with a cubic subfield and no quadratic subfield. Math. Comp. **58**(197), 419–432 (1992)
40. Pauli, S., Roblot, X.-F.: On the computation of all extensions of a $p$-adic field of a given degree. Math. Comput. **70**(236), 1641–1659 (2001)
41. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
42. Ruehle, F.: Evolving neural networks with genetic algorithms to study the string landscape. J. High Energy Phys. **2017**(8), 1–20 (2017)
43. The Sage Developers. SageMath, the Sage Mathematics Software System (2022). https://doi.org/10.5281/zenodo.6259615
44. Voight, J.: Enumeration of totally real number fields of bounded root discriminant. In Algorithmic Number Theory: 8th International Symposium, ANTS-VIII Banff, Canada 8, Springer, pp. 268–281 (2008)
45. Wang, S.: A counter-example to Grunwald's theorem. Ann. Math. **49**(4), 1008–1009 (1948)
46. Wang, S.: On Grunwald's theorem. Ann. Math. **51**(2), 471–484 (1950)


## Publisher's Note