

Review

Topological Data Analysis: Foundations, Algorithms, and Emerging Applications

Dimitrios Georgiou ^{1,*}, Sotiris Kotsiantis ¹  and Fotini Sereti ²¹ Department of Mathematics, University of Patras, 26504 Patra, Greece; kotsiantis@upatras.gr² Department of Chemical Engineering, University of Western Macedonia, 50100 Kozani, Greece; aff01310@uowm.gr

* Correspondence: georgiou@math.upatras.gr

Abstract

Topological data analysis (TDA) has evolved into a flexible and robust paradigm for obtaining qualitative, geometry-inspired insights from high-dimensional, noisy, and complex data. Grounded in algebraic topology, geometry, statistics, and machine learning (ML), TDA provides multiscale descriptions through persistent homology, Mapper (a graph-based method that summarizes the shape of high-dimensional data), and related topological signatures that are often inaccessible to standard linear and metric methods. In recent years, and especially during 2024–2025, TDA has expanded rapidly across science, engineering, biomedical research, and socio-economic studies, while also being integrated with modern learning paradigms such as deep learning (DL) and graph learning. This survey summarizes recent developments in TDA using a carefully selected set of articles, with emphasis on 2024–2025. We first present the mathematical and computational foundations of TDA, covering simplicial complexes, filtrations, persistent homology, the Mapper algorithm, and computational advances such as data simplification, stability, and efficiency. We then review applications in time series and dynamical systems, biomedical imaging and precision medicine, engineering and physical sciences, finance and risk analysis, DL and interpretability, and security and critical infrastructure systems. Throughout, we highlight how TDA can extract informative features, function as a model component, and provide a conceptual lens for studying complex systems. However, the survey also emphasizes recurrent failure patterns: TDA performance is highly sensitive to filtration, embedding, and vectorization choices; aggressive simplification can dilute or remove informative topological signals; and integration into standard ML workflows still lacks uniform validation and reporting protocols. We conclude by outlining key challenges—including scalability, statistical foundations, interpretability, and compatibility with rapidly evolving artificial intelligence (AI) paradigms—and by identifying directions for future research. The survey also provides a unifying design perspective for TDA systems, highlighting methodological trade-offs and emerging research directions for integrating topology with modern ML.



Academic Editor: Miguel Ángel Sánchez-Granero

Received: 11 May 2026

Revised: 5 June 2026

Accepted: 16 June 2026

Published: 19 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: topological data analysis; algebraic topology; persistent homology; Mapper; machine learning applications

MSC: 54A05; 54H99; 55N31; 55U05; 55U10; 68T07; 68T09

1. Introduction

Data analysis is the process of inspecting, cleaning, transforming, and modeling data to extract useful information and support decision-making. However, handling

data with many features—especially high-dimensional data—remains one of its central challenges. TDA emerged to address this difficulty by using tools from topology to study the “geometry” of complex, high-dimensional, nonlinear, heterogeneous, and noisy data, including shape, structure, and connectivity. Algebraic topology is a core component of this approach, and its contribution to modern mathematics is widely recognized [1]. In many applications, data exhibit geometric and topological patterns that are difficult to capture with classical statistical models or purely geometric approaches. TDA addresses this gap by identifying structures that are robust to perturbations and invariant under continuous deformations. It does not replace classical data analysis; rather, the two approaches are complementary. Table 1 presents a basic comparison that highlights the strengths and weaknesses of each field.

Table 1. A basic comparison of the main features of TDA and classical data analysis in terms of aim, robustness, mathematical tools, and results.

Main Features	TDA	Classical Data Analysis
Aim	Understanding the geometry and structure of high-dimensional, complex, noisy, and nonlinear data	Inspecting, cleaning, transforming, and modeling data; identifying correlations and linear relationships
Robustness	Strong robustness in high-dimensional settings	Sensitive to high dimensionality
Mathematical tools	Persistent homology, Mapper algorithm, barcodes, persistence diagrams	Regression, clustering
Results	Visualizations based on topological characteristics	Data visualizations and predictive models

At the heart of TDA is persistent homology, which tracks the appearance and disappearance of topological features in a filtered simplicial complex and is supported by a substantial foundational literature [2–4]. It yields descriptors such as barcodes and persistence diagrams that summarize the geometry underlying the data in a concise way. Additional techniques—including Mapper (a graph-based summary built from a cover and a clustering step), zigzag persistence, multiparameter persistence, and graph and cubical variants—expand the TDA toolbox for analyzing dynamical systems, networks, images, and time series data. In particular, alpha complexes and cubical persistent homology are practical complements to Vietoris–Rips constructions, respectively offering geometry-faithful sparsification for Euclidean point clouds and native support for image or voxel data. Table 2 presents the main axes of this toolbox, while Table 3 provides a comparative taxonomy of representative methods and Figure 1 provides a compact end-to-end pipeline view from raw data to downstream learning and interpretation, which is discussed in detail in Section 4.

The literature collected for this review reflects an especially active period in the development and application of TDA. Recent works employ TDA to model financial markets and detect structural changes in time series [5–8], to identify clinically meaningful patient subgroups in precision medicine [9,10], to characterize fluid flows, material roughness, and other physical systems [11–13], and to assess resilience of infrastructure networks [14]. In parallel, there is accelerating integration of TDA with advanced learning architectures, including graph neural networks, interpretable DL methods, and federated or incremental learning frameworks [15–19].

Table 2. A compact “toolbox” view of common ingredients in TDA. The choices are indicative and meant as a reading guide for the remainder of the survey.

Ingredient	What It Provides	Typical Examples
Data-to-complex encoding	Discrete representation of “shape” from point clouds, grids, graphs, or sequences	Vietoris–Rips, Čech, alpha, cubical, clique/simplicial complexes
Filtration	Multiscale view of topology by varying a scale/threshold/time parameter	Radius or scale filtrations; density or function sublevel filtrations; time-indexed or sliding-window filtrations
Topological summary	Robust descriptors of connectivity and holes across scales	Barcodes, persistence diagrams, Betti curves, Euler characteristic curves
Vectorization and representation	Fixed-dimensional features compatible with ML and DL pipelines	Landscapes, persistence images, kernels, entropy, codebooks, signature features
Downstream learning and inference	Predictive task, clustering, monitoring, or scientific interpretation	Classification or regression; anomaly detection; change-point detection; phenotype discovery; segmentation

Table 3. Comparative taxonomy of representative TDA methods. Reported complexity levels are qualitative and indicate typical computational burden in practical workflows.

Method	Data Type	Complexity	Strengths	Typical Applications
Vietoris–Rips persistent homology	Point clouds	Exponential	Simple and widely used construction; robust multiscale descriptors	Shape analysis, geometric machine learning, anomaly detection
Čech complexes	Metric data	High	Strong geometric and topological guarantees	Geometric inference, manifold reconstruction
Mapper	High-dimensional datasets	Moderate	Interpretable graph summaries and exploratory visualization	Biomedical stratification, exploratory data analysis
Alpha complexes	Euclidean point clouds	Moderate to high	Geometrically faithful and often sparser than Vietoris–Rips	Shape reconstruction, geometric inference
Cubical persistent homology	Images, voxel grids, scalar fields	Moderate	Natural encoding for grid-structured data and efficient memory layout	Medical imaging, materials, spatio-temporal fields
Persistence images	Persistent homology outputs (diagrams)	Low	Fixed-size vectorization compatible with standard ML pipelines	Classification, regression, feature engineering
Multiparameter persistence	Multiscale, multifilter data	Very high	Richer structure beyond one-parameter filtrations	Dynamical systems, multimodal and heterogeneous data

The objectives of this review are threefold. First, we discuss the mathematical and algorithmic foundations of TDA in an expository and pedagogical manner. Second, we synthesize recent methodological advances, including computational improvements, statistical inference, and links to ML. Third, we provide a guided survey of applications in science, engineering, economics, biomedicine, and AI, highlighting emerging themes and open problems. The material assumes familiarity with basic linear algebra and ML concepts but

does not require prior background in algebraic topology. This review is intended to be useful to both newcomers and experts seeking an update on TDA in 2024 and 2025.

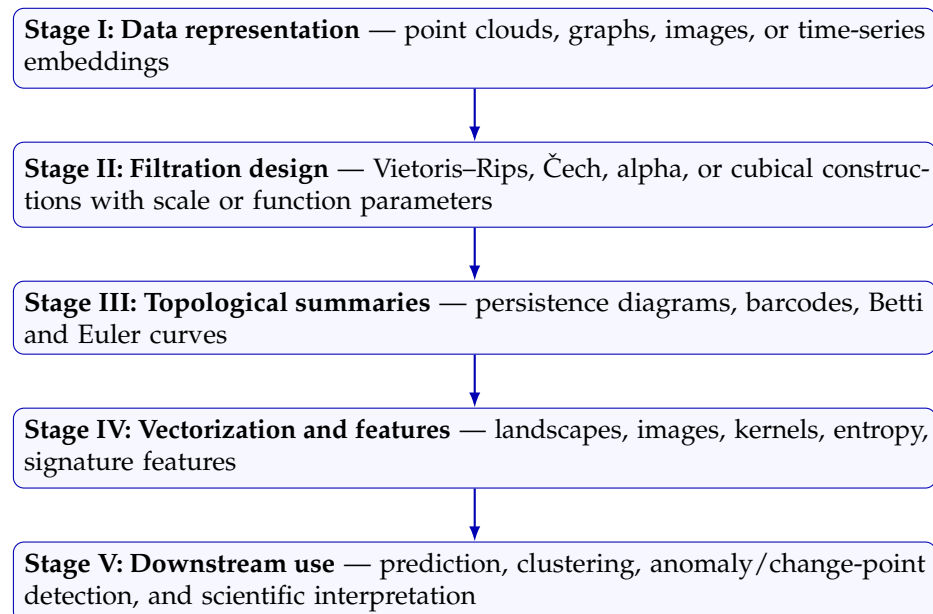


Figure 1. Conceptual TDA pipeline used throughout this survey.

To ensure consistent terminology throughout, we use TDA after its first occurrence, write Vietoris–Rips uniformly for complexes and filtrations, and use time series as the default wording.

The survey contribution is as follows:

- A curated and up-to-date synthesis of TDA developments, with primary emphasis on work published in 2024–2025.
- A unified presentation of mathematical foundations and computational advances, linking formal concepts to practical implementation choices.
- A cross-domain survey of applications (finance, biomedicine, engineering, dynamical systems, and AI), highlighting recurring design patterns.
- A pipeline-oriented perspective that clarifies where modeling decisions are made and how they affect robustness, interpretability, and performance.
- A critical discussion of open challenges—including scalability, statistical validity, and integration with modern AI—together with actionable future directions.

A Unifying Design Framework for TDA Systems

This survey introduces a unifying five-stage pipeline that clarifies how TDA-based systems are designed, implemented, and integrated with ML workflows. Figure 1 provides a compact visual summary of this stage-wise design view, while Table 3 summarizes representative methodological trade-offs. The framework makes modeling decisions explicit at each stage—from data representation and filtration construction to topological summarization, feature vectorization, and downstream learning—so that methodological assumptions can be traced and justified. It also enables systematic comparison of alternative design choices (e.g., complex family, filtration function, representation map, and learner) in terms of robustness, interpretability, and computational cost across application domains.

To position this manuscript against representative prior syntheses, Table 4 compares scope, years covered, domains, methods, and key strengths.

Table 4. Comparison with representative prior TDA surveys.

Survey	Scope	Years Covered	Domains	Methods Emphasized	Strengths
Edelsbrunner–Harer [2]	Foundational	Up to 2010	General topology, computation	Simplicial complexes, homology, persistence algorithms	Rigorous mathematical and algorithmic foundation.
Otter et al. [3]	Computational	~2002–2016	General; software-oriented	Persistent homology pipelines and algorithmic implementations	Clear map of computational choices and software ecosystem.
Wasserman [4]	Statistical	~2009–2017	Statistics/data science	Persistence summaries, inference, statistical viewpoints	Bridges TDA with modern statistical inference.
Chazal–Michel [20]	Introductory	Up to 2021	General data science	Persistent homology, vectorizations, practical workflow	Accessible entry point linking theory and practice.
Zhang et al. [15]	Domain-specific	Mainly 2017–2023	Deep learning interpretability	Mapper and topology-based deep neural network interpretation methods	Focused synthesis for neural network interpretation tasks.
This survey	Broad + integrative	Emphasis: 2024–2025	Finance, biomedicine, engineering, dynamical systems, AI	Foundations + computational advances + application design patterns	Cross-domain update with explicit limitations, reproducibility checklist, and open research questions.

Topological data analysis provides a complementary way to study complex data by focusing on shape, connectivity, and multiscale structure. In practical terms, it acts as a robust feature-extraction layer that can reveal patterns often missed by purely linear tools. The strongest results usually appear when TDA is combined with domain knowledge and modern learning models rather than used in isolation. The next sections explain the core mathematics, computational workflow, and evidence from applications.

2. Mathematical Background

The mathematical framework of TDA is inextricably linked with algebraic topology. Here, we present the main concepts commonly used in related studies. For more details, we refer to standard expositions and computational surveys [2–4]. This section introduces the core vocabulary needed to read TDA work: simplicial complexes, filtrations, homology, persistence, and Mapper. The practical idea is to observe data across multiple scales and retain features that remain stable as scale changes. Persistent summaries such as barcodes and diagrams convert complicated geometry into compact objects that can be compared across datasets. These concepts are flexible enough to support analysis of point clouds, images, networks, and time series.

2.1. Simplicial Complexes and Filtrations

A central objective in TDA is to approximate a dataset by a discrete combinatorial object that faithfully captures its underlying topology. This is achieved through simplicial complexes, which provide a flexible and computationally tractable representation of the geometric structure.

Definition 1 (Abstract simplicial complex). *Let S be a finite set of vertices. An abstract simplicial complex K on S is a collection of subsets $\sigma \subseteq S$ such that, whenever $\sigma \in K$ and $\tau \subseteq \sigma$, we also have*

$\tau \in K$. The elements of K are called *simplices*. A simplex $\sigma = \{v_0, \dots, v_k\}$ with $k + 1$ vertices is called a k -simplex and has dimension k . The dimension of K , denoted by $\dim K$, is

$$\dim K = \max\{\dim \sigma : \sigma \text{ is a simplex of } K\}.$$

Definition 2 (Geometric simplicial complex). A geometric simplex is the convex hull of a finite affinely independent set of points, and a face of such a simplex is the convex hull of a nonempty subset of its vertices. A geometric simplicial complex is a family P of geometric simplices satisfying the following two conditions:

- (i) if $\tau \in P$ and σ is a face of τ , then $\sigma \in P$;
- (ii) if $\sigma, \tau \in P$, then $\sigma \cap \tau$ is either empty or a face of both simplices.

Throughout this section, geometric complexes are considered in an ambient Euclidean space $\mathbb{R}^m$, with subspaces endowed with the induced topology.

In applications, the vertex set S typically consists of sampled data points equipped with a chosen metric $d : S \times S \rightarrow [0, \infty)$. This metric fixes the pairwise distances used to build complexes and filtrations. Common constructions include the following.

Definition 3 (Nerve of a cover). Let X be a topological space and let $\mathcal{U} = \{U_i\}_{i \in I}$ be a finite cover of X , meaning $X = \bigcup_{i \in I} U_i$. The nerve $N(\mathcal{U})$ is the abstract simplicial complex with vertex set I in which $\{i_0, i_1, \dots, i_k\}$ spans a k -simplex if and only if

$$U_{i_0} \cap U_{i_1} \cap \dots \cap U_{i_k} \neq \emptyset.$$

Thus, a simplex in the nerve records that the corresponding cover elements have a common intersection.

The nerve theorem is a central result because it allows complex, high-dimensional shapes to be replaced by simpler combinatorial structures under appropriate cover conditions. Recent work has investigated several versions and extensions of this theorem [21].

Theorem 1 (Nerve theorem). If $\mathcal{U}$ is a finite good cover of a topological space X —that is, every nonempty finite intersection $\bigcap_{j=0}^k U_{i_j}$ is contractible—then the underlying space of the nerve $N(\mathcal{U})$ is homotopy equivalent to X .

This is the standard nerve theorem; for a modern unified treatment and related variants, see Bauer et al. [21].

Other common constructions include:

- The Vietoris–Rips complex, where a set of points forms a simplex if all pairwise distances are at most a threshold ϵ .
- The Čech complex, where simplices correspond to nonempty intersections of closed balls of radius ϵ .
- Delaunay triangulations, in which a simplex with vertices in a finite point set P is included when its vertices lie on the boundary of a ball whose interior contains no point of P ; the resulting triangulation covers the convex hull of P under standard general-position assumptions.
- Alpha complexes, derived from Delaunay triangulations and sensitive to local geometry; for a scale ϵ , they retain Delaunay simplices whose circumscribing balls have a radius of at most ϵ .
- Cubical complexes, defined as collections of multidimensional cubes, such as vertices, edges, squares, and voxels, together with their faces.

For scales $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \dots$, varying ε produces a nested sequence of complexes,

$$K_{\varepsilon_1} \subseteq K_{\varepsilon_2} \subseteq \dots,$$

called a *filtration*. In Vietoris–Rips filtrations, for example, K_{ε_i} denotes the Vietoris–Rips complex formed at distance threshold ε_i . These filtrations are especially common because they depend only on pairwise distances, although they may grow exponentially in size [2,3]. Formally, Vietoris–Rips filtrations are nested collections of Vietoris–Rips complexes indexed by an increasing scale-distance parameter. Recent work has focused on refining these constructions to better reflect intrinsic geometric relationships, such as the topological overlap-based enrichments introduced in [10], which improve performance in high-dimensional genomic applications. Table 5 summarizes common simplicial and cubical constructions in TDA, describing how they are built, where they are used, and their typical size behavior.

Table 5. Common simplicial/cubical constructions used in TDA. “Complex size” is qualitative and refers to typical worst-case growth as the number of points/voxels increases.

Complex	Built From	Strengths/Typical Use	Complex Size
Vietoris–Rips	Point cloud with pairwise distances; add simplex if all edges have length $\leq \varepsilon$	Simple to define; widely used for persistent homology of point clouds	Can be very large (often exponential)
Čech	Point cloud; add simplex when ε -balls intersect	Good geometric meaning (nerve theorem)	Large; requires intersection tests
Alpha	Point cloud via Delaunay triangulation	Geometry-aware; often smaller than Rips in low dimensions	Moderate (dimension-dependent)
Cubical	Gridded data (images/volumes); cells on a lattice	Natural for images/voxel data; efficient implementations	Typically manageable
Clique/flag on graphs	Graphs or networks; simplices from cliques	Captures higher-order connectivity in networks	Depends on clique structure

2.2. Homology

Homology provides algebraic invariants that detect connected components, cycles, and holes. Given a simplicial complex K and a coefficient field $\mathbb{F}$, the chain group $C_k(K; \mathbb{F})$ is the vector space of formal finite linear combinations of k -simplices with coefficients in $\mathbb{F}$ [2,3]. For each k , the simplicial boundary map

$$\partial_k : C_k(K; \mathbb{F}) \rightarrow C_{k-1}(K; \mathbb{F})$$

sends an oriented k -simplex to the alternating sum of its $(k-1)$ -dimensional faces. The boundary maps satisfy the chain-complex condition

$$\partial_{k-1} \circ \partial_k = 0,$$

which permits the definition of the k -th homology group as

$$H_k(K; \mathbb{F}) = \ker(\partial_k) / \operatorname{im}(\partial_{k+1}).$$

The ranks of these groups, the *Betti numbers* $\beta_k = \dim H_k(K; \mathbb{F})$, count k -dimensional topological features: connected components (β_0), loops or tunnels (β_1), voids or enclosed cavities (β_2), and so forth. Homology can also be defined for cubical complexes, which arise naturally in image analysis, often yielding computational benefits. For example, ref. [22] uses cubical persistent homology and statistical methods to classify plant images.

2.3. Persistent Homology

Persistent homology extends classical homology by tracking the evolution of topological features across a filtration. Persistent homology extends classical homology by tracking the evolution of topological features across a filtration [2,3].

Definition 4 (Persistent homology, barcode, and persistence diagram). *Let $(K_\alpha)_{\alpha \in \mathbb{R}}$ be a filtration of simplicial complexes. For $\alpha \leq \beta$, the inclusion map $K_\alpha \hookrightarrow K_\beta$ induces a linear map on k -homology groups for each dimension k :*

$$H_k(K_\alpha) \rightarrow H_k(K_\beta).$$

The k -th persistent homology group is the image of this induced map, and the corresponding k -th persistent Betti number is its rank. A homology class that appears (is born) at scale α and disappears (dies) at scale β is represented by an interval $[\alpha, \beta)$. A barcode is the multiset of such intervals, and the interval length records the lifetime of the corresponding topological feature. Equivalently, a persistence diagram is the multiset of points (α, β) that record the same birth and death scales.

Persistence diagrams summarize the multiscale topology of data and enjoy fundamental stability theorems, guaranteeing robustness under perturbations of the input. These diagrams can be vectorized into forms suitable for statistical learning, such as persistence landscapes, persistence images, Betti curves, and entropy measures. Algebraically, a one-parameter persistence module over a field $\mathbb{F}$ is a sequence of vector spaces $(M_i)_{i \in \mathbb{N}}$ together with linear maps $M_i \rightarrow M_j$ for $i \leq j$ that are compatible with composition. It is of finite type when the vector spaces are finite-dimensional and only finitely many birth or death events occur. The structure theorem for such modules explains why persistent homology can be represented by barcodes or persistence diagrams; standard proofs and algorithmic treatments are given in [2].

Theorem 2 (Structure theorem for one-parameter persistence). *Any finite-type $\mathbb{N}$ -indexed persistence module over a field $\mathbb{F}$ decomposes as a direct sum of interval modules.*

Recent extensions include multiparameter persistence for data with several filtering functions [23], zigzag persistence for filtrations that grow and shrink over time, and Morse-theoretic approaches for structured signals such as corrosion-related ultrasonic waveforms [24]. From a purely algebraic perspective, Bjerkevik develops a stability theory for decompositions of multiparameter persistence modules that extends the one-parameter behavior described by the structure theorem [25]. Dey and Xin present an efficient algorithm for computing decompositions of finite multiparameter persistence modules [26]. Analytical work on random topology provides null models for interpreting topological signal strength in high dimensions [27].

2.4. Mapper and Related Constructions

Reeb graphs are useful TDA tools and are often regarded as precursors of the widely used Mapper algorithm. They visualize the core structure of an image, object, or dataset. Let X be a topological space and let $f : X \rightarrow \mathbb{R}$ be a continuous function. Define an equivalence relation on X by

$$x \sim y \quad \text{if} \quad f(x) = f(y) = t \quad \text{and there exists a continuous path } \gamma : [0, 1] \rightarrow f^{-1}(t) \text{ with } \gamma(0) = x, \gamma(1) = y.$$

The Reeb graph $R(X, f)$ is the quotient space $X / \sim$ equipped with the induced function $\hat{f} : R(X, f) \rightarrow \mathbb{R}$ defined by $\hat{f}([x]) = f(x)$, where $[x]$ is the equivalence class of x .

Recent work uses Reeb graphs to study medical artifacts and quality criteria in cerebral vascular trees affected by noise and subtle topological changes. Lepaire et al. introduce a Local-to-Global Reeb Graph (LGRG) variant [28], while Rahman et al. introduce Gradient-Aware Shortest Path (GASP), an algorithm for producing Reeb graph visualizations that satisfy boundary constraints, compactness, and gradient alignment [29].

The *Mapper* algorithm can be viewed as an applied nerve construction built from the pullback cover of a filtered topological space. Suppose that $f : X \rightarrow \mathbb{R}^m$ is continuous and that $\mathcal{W}$ is a finite open cover of $\mathbb{R}^m$. The pullback cover of X is

$$f^*\mathcal{W} = \{f^{-1}(W) : W \in \mathcal{W}\}.$$

This cover is refined by splitting each set $f^{-1}(W)$ into its path components; denote the resulting cover by $\widehat{f^*\mathcal{W}}$. The Mapper complex associated with (X, f) is the nerve of $\widehat{f^*\mathcal{W}}$. Consequently, Mapper provides a graph-based representation of data by combining:

1. a filter (or lens) function;
2. a cover of the filter range;
3. clustering within the preimage of each cover element.

The resulting graph captures connectivity between clusters and provides interpretable summaries of complex structures. Mapper has been especially successful in biomedical data analysis. For example, Loughrey et al. [9] introduce a hotspot detection framework for Mapper graphs that automates parameter selection and identifies clinically meaningful cancer subgroups. Kondapalli and Azarudeen [30] use Mapper to identify topological features of cancer cells that help distinguish healthy profiles from malignant ones. In addition, Bui et al. [31] use Mapper to visualize the evolution of neural network weights. A comprehensive review of Mapper and its applications is provided in [32].

Related constructions include:

- Ball Mapper, useful for point clouds with heterogeneous density and applied in credit risk and supply chain finance [33].
- Extensions such as V-Mapper for data with velocity fields [34] and Fuzzy-Mapper (F-Mapper), which handles cover intervals with an arbitrary overlap percentage [35].
- Euler-characteristic-based descriptors that integrate combinatorial counts with persistent information [36]. For a k -dimensional simplicial complex K , let $c_n(K)$ denote the number of n -simplices. The *Euler characteristic* is

$$\chi(K) = \sum_{n=0}^k (-1)^n c_n(K).$$

Equivalently, when homology is computed over a field, it can be written in terms of Betti numbers as

$$\chi(K) = \sum_{n=0}^k (-1)^n \beta_n$$

(see, for example, [37]).

These tools extend the reach of TDA to settings where filtrations alone may not capture the full shape of the data.

This section introduced the core vocabulary needed to read TDA work: simplicial complexes, filtrations, homology, persistence, and Mapper. The practical idea is to observe data across multiple scales and retain features that remain stable as scale changes. Persistent summaries such as barcodes and diagrams convert complicated geometry into compact objects that can be compared across datasets. These concepts are flexible enough to support analysis of point clouds, images, networks, and time series.

3. Algorithmic and Computational Advances

Theoretical foundations are a major strength of TDA for understanding complex datasets through algebraic-topological principles. At the same time, TDA has developed into a practical, scalable, and noise-robust framework for data analysis. Algorithmic and computational advances are central to this applied perspective.

3.1. Matrix Reduction and Cohomological Algorithms

The standard algorithm for computing persistent homology reduces a boundary matrix via column operations. Although worst-case complexity is cubic in the number of simplices, practical performance is often better, and many optimizations reduce both runtime and memory usage. Cohomology-based methods, such as persistent cohomology, are often faster in practice due to their dual formulation via cochain complexes. A concise review of persistent homology and persistent cohomology appears in [38]. Applications include the description of nongeometric information in molecular structures [39], while ref. [40] presents a new distributed algorithm for persistent cohomology.

Several strategies enhance large-scale computation:

- Clearing and compression techniques accelerate matrix reduction by eliminating unnecessary operations.
- Chunking partitions the boundary matrix to improve cache performance and parallelization.
- Ordering heuristics select simplex orderings that reduce fill-in and improve algorithmic stability.

Beyond classical matrix reduction, gradient-based optimization for functions defined on persistence diagrams has advanced significantly. For instance, Leygonie et al. propose gradient sampling methods for stratifiable objective functions derived from extended persistent homology over lower-star filtrations [41]. These methods enable more direct integration of persistent homology into optimization pipelines.

Furthermore, theoretical work on the sensitivity and stability of persistence-based descriptors has enabled the development of differentially private TDA mechanisms with near-optimal accuracy [42]. Complementary approaches, such as Euler-characteristic-based transforms for compressed topological signals [36], demonstrate that topological information can be extracted efficiently in low-memory environments.

3.2. Data Reduction and Stability

In TDA, stability is fundamental because it ensures that small perturbations of the input data produce small changes in the resulting topological summaries. Thus, if a dataset is slightly affected by noise, its persistence diagram should remain close to the diagram of the original dataset. This principle motivates stability theorems for filtrations, tame functions, and Lipschitz functions. For example, Atienza et al. study conditions under which persistent homology and persistent entropy remain stable [43]. A representative statement is the following.

Stability Theorem: Let K be a simplicial complex and let $f, g : K \rightarrow \mathbb{R}$ be monotone functions. For each dimension p , the bottleneck distance between the persistence diagrams $X = \text{Dgm}_p(f)$ and $Y = \text{Dgm}_p(g)$ is bounded above by the L_∞ -distance between the functions:

$$W_\infty(X, Y) \leq \|f - g\|_\infty.$$

In addition, data reduction is essential for scaling TDA to modern datasets, which may contain millions of points or high-dimensional embeddings. Choi et al. introduce the *Characteristic Lattice Algorithm* (CLA), a principled reduction method that preserves both

geometric and topological structure while providing stability bounds on the bottleneck distance between barcodes before and after reduction [44]. Such techniques significantly decrease computational burden without sacrificing essential topological information.

Data reduction also arises naturally in time-delay embeddings of time series. Sliding-window embeddings can produce large point clouds, making parameter selection (particularly the time delay) crucial. Myers et al. propose using persistent homology to select optimal delay parameters for permutation entropy calculations, replacing ad hoc choices with a topologically grounded criterion [45].

3.3. Enhanced Complexes and Overlapping Measures

Recent work has focused on modifying simplicial complexes to better reflect underlying data geometry. Mashatola et al. [10] introduce *topological overlapping measures* that enrich Vietoris–Rips complexes by incorporating shared-neighborhood information. In high-dimensional genomic settings, such enhancements yield persistent features that are more discriminative and more stable under noise.

Other methodological developments include:

- Expected-topology analysis, which studies the typical behavior of persistent diagrams for random point clouds and offers benchmarks for assessing significance [27].
- Shape-aware complex constructions that embed additional geometric information, improving sensitivity in engineered systems and biological structures [46].
- Cubical complexes for voxel and pixel data, which provide computational benefits for image-based TDA, as demonstrated in biomedical and material imaging contexts [22,47].

3.4. Statistical Limits and Inference

Despite its success, the statistical foundations of TDA are still under development. Vishwanath et al. demonstrate that persistent homology does not always provide sufficient statistics for inference, identifying conditions under which topological summaries are informative and models where they are fundamentally insufficient [48]. This work underscores the importance of understanding the theoretical limits of TDA-based inference.

Complementary contributions address inference and simulation, particularly in the analysis of dynamic dependence networks, such as those arising in neuroscience. El-Yaagoubi et al. introduce simulation frameworks for generating multivariate time series with known topological characteristics, enabling hypothesis testing and confidence-set construction for persistent homology applied to evolving networks [49,50].

3.5. Quantum Algorithms and Complexity

Quantum computation offers a potential pathway to accelerate persistent homology, particularly for datasets with large ambient dimensions. Berry et al. refine quantum TDA algorithms by improving Dicke-state preparation and projector-based eigenvalue estimation [51]. Their results show the possibility of superpolynomial speedups for structured instances with large Betti numbers.

However, complexity-theoretic work by Schmidhuber et al. reveals fundamental limitations: computing exact Betti numbers is $\#P$ -hard and even multiplicative approximation remains NP-hard in favorable regimes [52]. These results indicate that quantum speedups, while possible, are restricted to specific data regimes and cannot generally overcome worst-case combinatorial complexity. Hybrid quantum–classical TDA pipelines remain a promising but bounded open direction.

Algorithmic progress is what makes TDA usable on real datasets rather than only on toy examples. Improvements in reduction methods, cohomology-based computation,

sparsification, and statistical tooling help reduce runtime and improve reliability. At the same time, complexity results show that some computational barriers are fundamental, even with quantum ideas. For practitioners, the key is to balance scalability and accuracy while documenting computational choices clearly.

4. A Unifying Pipeline Taxonomy for TDA-Based Systems

Across domains, most TDA systems follow a common computational pattern. To make the literature comparable, we organize TDA-based approaches into a five-stage pipeline, with optional feedback loops between stages. Figure 2 shows the TDA workflow of this five-stage procedure. The key differences between papers typically arise from (i) how the raw data are encoded prior to topology, (ii) how the filtration is designed, and (iii) how topological information is represented for learning and decision-making.

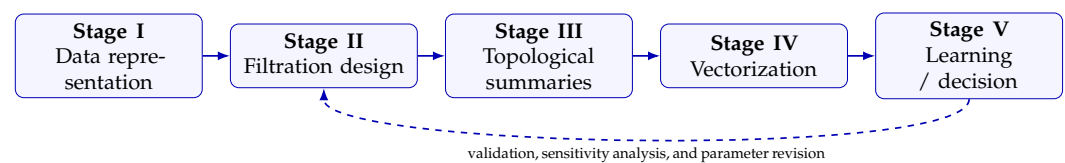


Figure 2. A revised high-level schematic of the TDA pipeline with readable stage labels and an explicit validation feedback loop.

Stage I: Data representation (pre-topological encoding).

The first stage converts domain data into a geometric or combinatorial object on which topology can be computed. Common choices include point clouds (e.g., measurements or embeddings), graphs (e.g., interaction networks), and images or fields (e.g., grayscale images or scalar fields). For time series, state-space reconstructions such as time-delay embeddings or sliding-window point clouds are typical, while for networks, one may work directly with weighted graphs or simplicial complex constructions that encode higher-order interactions [53,54]. This stage fixes the *ambient object* and therefore determines which topological summaries are meaningful and computationally feasible.

Stage II: Filtration design (the topological lens).

A filtration specifies how structure is revealed across scales and is often the most consequential modeling choice. Metric filtrations (e.g., Vietoris–Rips) are natural for point clouds, whereas function-based filtrations are standard for images and scalar fields (e.g., sublevel or superlevel filtrations). Temporal and time-resolved constructions provide topology-aware views of evolving systems and are widely used in change detection and monitoring applications, including finance [55,56]. Filtration design implicitly defines the notion of “scale” or “importance” and can encode application-specific priors (density, intensity, correlation strength, velocity, etc.).

Stage III: Topological feature extraction.

Given a filtration, one computes topological descriptors such as persistence barcodes and diagrams and derived invariants (Betti numbers across scale, representative cycles, etc.). This stage is mathematically uniform across application areas, but its computational burden varies substantially with the chosen complex and dataset size. For example, cubical complexes make image-based persistence tractable, whereas large Vietoris–Rips complexes may require sparsification or subsampling. This stage remains central for extracting meaningful topological evidence in applications. For instance, Wang et al. [57] propose a TDA-based framework for extracting structural features of materials, providing informative insights into structure–property relationships and predictive strategies.

Stage IV: Representation and vectorization.

Persistence diagrams are not directly compatible with most learning algorithms; hence, vectorization is required. Representative families include functional summaries such as persistence landscapes, information-theoretic summaries such as persistent entropy [43], and fixed-length embeddings such as persistence codebooks [58]. More algebraic constructions map persistence outputs to paths and apply signature features to obtain expressive, stable representations [59]. On the algorithmic side, improved diagram-comparison routines (e.g., efficient bottleneck-distance computation in special cases) support scalability when topology is used via distances [60]. Surveys and recent work further discuss how these representations can be combined with deep architectures and hybrid pipelines [61].

Stage V: Learning and decision layer.

Finally, topological representations are consumed by a downstream task: classification, clustering, regression, change-point detection, optimization, or monitoring. In finance and other dynamical settings, topology is frequently used for regime and change detection [6,55,56]. In biomedical and precision-medicine contexts, topology contributes robust, interpretable features for patient stratification and longitudinal analysis [62–64]. In engineering, topology can also constrain decision-making by enforcing validity domains during optimization [65] or by supporting monitoring of complex trajectories [66]. In [67], the use of TDA has been identified in smart manufacturing and industrial production.

Cross-cutting axes: role of topology, coupling strength, and temporal awareness.

Beyond the five stages, we find three cross-cutting axes that help classify the literature. First, the *role of topology* ranges from a descriptive summary (exploration and visualization) to a feature extractor for predictive models, a distance metric for comparing samples or trajectories, or an explicit constraint for optimization. Second, the *coupling strength* between topology and learning ranges from post hoc analysis (TDA computed independently) to hybrid pipelines (TDA features combined with standard machine learning) and, increasingly, to learned or end-to-end representations [61]. Third, *temporal awareness* distinguishes static pipelines from sliding-window approaches and explicitly time-indexed filtrations, which are central in monitoring and change-point problems [55,56].

As summarized in Table 6, differences between approaches primarily arise at the representation and filtration stages rather than in the topological computation itself. Thus, in this table, we focus on Stages I, II, IV, and V, as Stage III is widely applied for getting more topological and computational processes which are uniform in various applications.

Table 6. Representative TDA-based systems organized according to the proposed pipeline taxonomy. Each work is characterized by its data encoding (Stage I), filtration strategy (Stage II), representation or vectorization method (Stage IV), and downstream task (Stage V).

Reference	Stage I: Encoding	Stage II: Filtration	Stage IV: Representation	Stage V: Task
Zielinski et al. [58]	Persistence diagrams	Scale-based (PD)	Codebook vectors	Classification
Chevyrev et al. [59]	Persistence diagrams	Scale-based (PD)	Path signatures	Learning/classification
Atienza et al. [43]	Persistence diagrams	Scale-based (PD)	Persistent entropy	Pattern recognition
Guo et al. [55]	Time-delay embeddings	Sliding window	Landscape norms	Change detection
Katz & Biem [56]	Time series embeddings	Temporal filtration	Persistence landscapes	Early-warning detection
Dagliati et al. [62]	Patient records	Density-based	Topological summaries	Phenotype discovery
Lin [63]	Single-cell networks	Temporal filtration	Betti trajectories	Trajectory analysis
Bussola et al. [68]	Histopathology images	Cubical filtration	Persistence diagrams	Immune quantification
Frahi et al. [66]	Robot trajectories	Sliding window	Persistence summaries	Monitoring
Schweidtmann et al. [65]	Parameter samples	Distance-based	Topological constraints	Optimization

Method-Selection Guide (Data Type to Complex or Filtration to Topological Summary to Machine Learning Model)

To make the pipeline actionable for practitioners, Table 7 provides a compact starting guide that maps common data modalities to typical topological design choices and downstream learning models. The guide is not prescriptive; it is intended as a first-pass design template to be refined with domain-specific validation and sensitivity analysis.

How to use the guide. In practice, the following decision rules can be applied:

- (i) Select the simplest complex or filtration that matches the data geometry.
- (ii) Start with stable summaries, such as landscapes, Betti curves, or persistence images, before using higher-capacity representations.
- (iii) Benchmark topology-augmented models against strong non-topological baselines.
- (iv) Report sensitivity analyses for filtration and embedding hyperparameters.

Table 7. Method-selection guide for TDA pipelines: from data type to complex or filtration, topological summary, and machine learning model.

Data Type	Complex or Filtration (Stages I–II)	Topological Summary (Stages III–IV)	Machine Learning Model (Stage V)
Time series (regime shifts, early warning)	Time-delay or sliding-window point clouds; Vietoris–Rips or temporal filtration [6,55,56]	Persistence landscapes, Betti trajectories, diagram norms	Gradient boosting; temporal convolutional network (TCN) or long short-term memory (LSTM) model with topological features
Images and volumetric data	Cubical complexes; intensity/gradient sublevel filtrations [68,69]	Persistence images, Euler/Betti curves, persistent entropy	Convolutional neural network (CNN) with topological channel; support vector machine (SVM) or random forest (RF) in low-data regimes.
Point clouds and shape data	Vietoris–Rips, Čech, or alpha complexes with radius filtration [43,58]	Persistence diagrams, landscapes, kernel embeddings	Support vector machine or kernel methods; multilayer perceptron (MLP) with vectorized summaries.
Graphs and networks	Clique/flag complexes; weighted graph filtrations (edge weight, centrality, density) [14,70]	Persistence statistics, entropy, graph-topological signatures	Graph neural network (GNN) with topological features; tree ensembles.
Clinical and omics tabular data	Patient similarity networks or Mapper lenses with density-based filtration [9,10,62]	Mapper cluster descriptors, persistence summaries, subgroup markers	Tree ensembles, survival models, shallow neural models.
Streaming monitoring systems	Sliding temporal windows; zigzag or time-indexed filtrations [71,72]	Time-indexed Betti curves, persistence trajectories	Online anomaly detectors, temporal graph neural networks, sequential classifiers.

Common failure modes remain data-dependent: embedding-window sensitivity in time series, preprocessing dependence in imaging, combinatorial cost for large point clouds, graph-construction noise in network studies, cohort shift in clinical data, and drift in streaming environments.

The pipeline taxonomy turns the broad literature into a concrete checklist: encode data, build a filtration, compute topology, represent features, and perform learning or inference. This perspective clarifies where conclusions can change because of preprocessing, metric choice, or hyperparameters. It also explains why many successful papers report gains from thoughtful stage-by-stage integration rather than from a single algorithmic trick.

Reproducible TDA therefore depends on transparent reporting of all pipeline stages, not only final accuracy scores.

5. TDA, ML, and DL

AI, including ML, remains one of the most intensively studied areas in data science. In parallel, TDA is increasingly integrated with ML and DL, including multilayer neural architectures for tasks such as classification and regression.

5.1. Vectorizations of Persistence and Topological Feature Engineering

Persistent homology produces barcodes and persistence diagrams that typically need to be translated into fixed-dimensional vectors in order to be employed in machine learning. A wide range of vectorizations has been developed for persistence diagrams, such as persistence landscapes, persistence images, Betti curves, silhouette functions, kernel embeddings, and algebraic summary statistics. While there has been considerable debate on the relative merits of different persistence diagram representations in machine learning tasks, existing studies have found that barcode statistics perform competitively.

Recent work demonstrates the strong performance of TDA-derived features in noisy and nonlinear classification tasks. For example, TDA features extracted from voltage–current trajectories improve appliance identification in non-intrusive load monitoring over Principal Component Analysis (PCA)-based approaches, especially in high-noise conditions [73]. Similar effects are observed in chemical sensing: TDA-based features substantially improve accuracy in low-cost electronic nose systems, as demonstrated in [74]. In customer analytics, barcode statistics and persistence images have led to enhanced churn prediction without extensive hyperparameter tuning [75]. In materials science, persistence images derived from roughness surfaces enable quantitative comparison of micro-crack patterns in bonded assemblies [11].

Other developments include Euler-characteristic-based transforms, efficient compressed-vector schemes and differentially private embeddings of persistence diagrams [36,42]. These innovations expand the range of ML-compatible topological descriptors and support applications that require both privacy and interpretability.

A central challenge in topological data analysis is the transformation of persistence diagrams into representations suitable for statistical learning and downstream inference. Several theoretically grounded vectorization strategies have been proposed. Persistence codebooks provide a dictionary-based embedding of diagrams that enables efficient comparison and learning while preserving discriminative topological information [58]. Signature-based approaches interpret persistence diagrams as paths and apply tools from rough-path theory, yielding stable and expressive representations with strong theoretical guarantees [59].

Information-theoretic summaries have also been studied extensively. In particular, the stability of persistent entropy with respect to perturbations of the input data has been formally established, supporting its use as a robust scalar descriptor in noisy settings [43]. From a computational perspective, advances in efficient bottleneck-distance computation have significantly reduced the cost of diagram comparison, making TDA more scalable in practice [60]. Finally, comprehensive surveys have reviewed the integration of persistence-based representations with deep learning architectures, highlighting both theoretical foundations and practical design patterns [61].

Critical limitations: Persistence vectorization studies can conflate topological information with choices made later in the ML pipeline. Landscapes, images, kernels, entropy summaries, and codebooks require bandwidths, grids, weights, thresholds, or dictionary sizes, and these choices may either suppress localized features or create high-dimensional descriptors that overfit small benchmarks. Reported gains are therefore most convincing

when accompanied by ablations over vectorization parameters, comparisons with non-topological feature sets, calibration checks, and external or out-of-distribution validation.

5.2. Topology-Aware Deep Learning and Interpretability

Deep neural networks often lack interpretability and TDA has emerged as a powerful lens for probing their internal structure. Zhang et al. provide a unifying survey of TDA for DL explainability, highlighting applications ranging from analyzing data manifolds through persistent homology to visualizing decision boundaries via Mapper-based constructions [15]. At the activation level, persistent homology applied to correlation graphs of neuronal activations can predict the generalization gap of deep networks without requiring a test set [19]. These results suggest that network complexity and overfitting tendencies are reflected in persistent topological patterns.

Beyond interpretability, TDA has been incorporated directly into neural architectures. Examples include:

- Neural models augmented with topological features extracted from spatial cell layouts, achieving performance comparable to convolutional neural network (CNN) baselines in stem cell classification [22].
- TDA-SegUNet, a topological variant of the U-Net convolutional encoder–decoder segmentation architecture, which integrates persistence images as multiscale descriptors for brain tumor segmentation and achieves state-of-the-art results on brain tumor segmentation (BraTS) datasets [76].
- Hybrid geometric DL methods for protein pocket classification that combine global topological invariants with local graph neural network (GNN)-based representations [77].

These avenues reflect a broader trend: topology is increasingly viewed not just as an interpretability tool but as an architectural bias that enhances model robustness and representation quality.

Critical limitations: Topology-aware DL results are often sensitive to architecture, preprocessing, loss-weight selection, and dataset-specific annotation quality. Topological losses or persistence-image channels may add substantial computational cost, and apparent interpretability can remain qualitative unless it is tested against saliency, shape, or uncertainty baselines. Claims of state-of-the-art performance should therefore report runtime, memory, ablations without the topological component, and validation on independent cohorts or tasks.

5.3. Graph Neural Networks and Topology-Driven Representation Learning

Graphs play a central role in many domains, and graph neural networks (GNNs) often struggle to capture higher-order or multiscale topological structures. In parallel, Gavris et al. study a machine learning approach that uses message-passing GNNs to reduce non-physical transition zones in topology-optimization problems [78]. TDA offers complementary descriptors that can augment or constrain GNN message passing.

Pham introduces a fuzzy neural network coupled with topological graph learning for molecular property prediction, integrating uncertainty-aware modules with persistent homology-based descriptors [16]. Extending this idea, Pham et al. survey the rapidly growing literature on TDA-enhanced GNNs, including topological regularization, persistence-informed pooling, and homology-based graph kernels [17]. These methods often outperform classical GNN baselines, particularly on tasks that involve long-range dependencies or non-Euclidean structures.

TDA has also enriched learning paradigms beyond supervised settings. In federated and incremental learning, Gong et al. develop a TDA-based stability loss that mitigates catastrophic forgetting by preserving a global topological structure across client updates [18].

In multimodal recommendations, Bachiri et al. use persistent homology to construct robust cross-modality graph representations with improved ranking metrics and superior cold-start performance [79].

Beyond node- and edge-level descriptors, TDA provides tools for capturing higher-order connectivity patterns in complex networks. Simplicial and cell-complex representations enable the definition of topological centrality measures that generalize classical graph metrics to higher dimensions. Such constructions have been shown to reveal mesoscopic organization and functional roles in real-world networks, complementing standard GNN pipelines [53,54].

Together, these contributions illustrate how topological insights can guide the design of graph-based learning systems, leading to models that are both more expressive and more stable.

In ML and DL, TDA is most effective as a source of structured features or as a regularization signal. It can improve robustness and interpretability, especially in noisy, nonlinear, or data-scarce settings. However, gains are not automatic and depend on representation design and model–task alignment. A practical strategy is to compare topology-augmented models against strong baselines and report ablation analyses.

Critical limitations: TDA-enhanced GNN studies depend strongly on graph construction, edge weighting, filtration design, homology dimension, and pooling strategy. Improvements can reflect richer preprocessing or larger feature budgets rather than genuinely topological information. For large, dynamic, or heterogeneous graphs, scalability and stability under missing edges, noisy labels, and distribution shift remain only partially tested. Strong evidence requires comparisons with modern GNN baselines, parameter-sensitivity analysis, and task-level ablations that isolate the topological contribution.

Before moving to domain-specific sections, Figure 3 summarizes the application landscape covered in this review and highlights where topological summaries most often interact with modern learning pipelines.

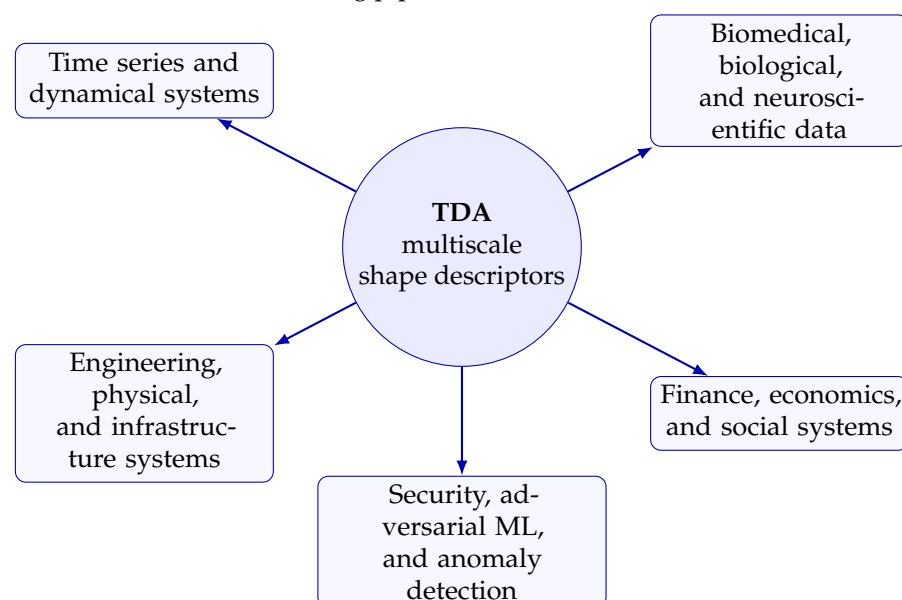


Figure 3. Domain map of applications synthesized in this survey.

6. Time Series and Dynamical Systems

Time series and dynamical systems are two strongly interacting fields where TDA is widely applied, ranging from financial environments to time series clustering pipelines. In this section, we review recent TDA applications in these areas.

6.1. Change-Point Detection and Financial Dynamics

Time series can be embedded into high-dimensional spaces using sliding windows or Takens' delay embedding, after which persistent homology can capture geometric signatures of dynamics. Yao et al. apply this approach to change-point detection in financial markets, defining TDA-based volatility indicators that align with known extreme events such as the European debt crisis, Brexit, the COVID-19 pandemic, and the Russia–Ukraine energy crisis [6]. Their TDA indicators outperform classical univariate and multivariate CPD methods in F1 score across different tolerance windows.

Nie studies nonlinear serial dependence in stock returns using topological cross-correlation measures and rolling-window TDA, revealing how geopolitical and policy events (e.g., Russia–Ukraine conflict, trade policies) impact the dependence structure of equity returns [5]. De Jesus et al. enhance univariate time series forecasting for financial instruments by integrating TDA-derived features (entropy, amplitude, counts from persistence diagrams) into N-BEATS models, achieving consistent improvements across cryptocurrencies and traditional assets [7]. Kulkarni et al. use persistent homology alongside geometry-inspired network measures (such as Ricci curvature) to assess fragility and systemic risk in Indian stock markets, finding persistent entropy to be a robust and informative topological measure [8]. Related work exploits topological properties of LPPLS-type trajectories to explain why TDA is particularly effective at detecting financial bubbles and early-warning signals [80] and extends persistence-based analyses to Chinese equity markets under major public events [81]. Md. Morshed Bin Shiraj et al. study the integration of the Mapper algorithm with DBSCAN clustering to detect anomalies in financial time series, showing that TDA provides a robust anomaly-detection framework, particularly in high-dimensional settings where classical clustering methods may fail to capture the global structure [82].

TDA has been increasingly applied to financial time series to capture structural changes that are difficult to detect using purely linear models. Persistent homology-based indicators have been proposed as early-warning signals for market crashes and regime shifts, demonstrating sensitivity to precursory changes in correlation structure and volatility patterns [55,56].

In related work, topological summaries have been incorporated into time series clustering and classification pipelines, where they improve discrimination between market states and asset behaviors [83,84]. TDA has also been explored in portfolio construction and enhanced indexing strategies, where topological features provide complementary information to classical risk–return metrics [85,86].

Beyond finance, Adami et al. study avalanche-size sequences in sandpile models via visibility graphs and persistent homology, uncovering scale-free behavior and power-law distributions for simplices and Betti numbers, with potential applications to earthquakes and other self-organized critical systems [87]. More generally, engineered topological features have been shown to outperform standard statistical and wavelet-based descriptors in classifying stochastic processes and time series with different noise properties [88,89].

Critical limitations: Consistent with the cautionary patterns summarized in Table 8, financial change-point and time-series studies are often retrospective and may tune embedding windows, distance choices, or persistence thresholds around historically known events. Several comparisons are made against classical CPD or forecasting baselines, but the evidence is less conclusive when stronger multivariate econometric models, modern sequence learners, transaction-cost-aware portfolio tests, or cross-market external validation are required. Reported improvements should therefore be interpreted as evidence that topology provides a useful structural signal, not as proof that TDA alone is a universally superior financial predictor.

Table 8. Critical synthesis of recurring evidence patterns across settings.

Setting	What Consistently Works	What Often Fails or Needs Caution
Noisy/nonstationary time series (finance, sensing, dynamical systems)	TDA captures regime shifts and transition structure when paired with sliding-window embeddings and conventional predictors [5,7,71]	Performance degrades with poorly tuned embedding dimensions and windows; false change points may appear under unstable preprocessing [8]
Biomedical imaging and omics with subgroup heterogeneity	Topological summaries and Mapper reveal phenotypic subgroups and morphology-related biomarkers, especially under class imbalance [9,10,69]	Clinical transfer is limited when cohorts are small, preprocessing differs across sites, or filtration choices are not externally validated [90]
Graph, infrastructure, and security monitoring	Persistent and graph-topological features improve anomaly detection and resilience analysis by encoding higher-order connectivity [14,70,72]	Benefits shrink when topology is used without domain constraints or when attacks/noise violate modeling assumptions [91]
Deep learning integration	Best results occur when topology acts as regularizer or auxiliary channel, improving robustness and interpretability [17–19]	End-to-end differentiable topology remains computationally heavy; gains are inconsistent without careful architecture-task matching [16]

6.2. Real-Time and Nonstationary Systems

Real-time state estimation in nonstationary mechanical systems poses difficult challenges. Razmarashooli et al. use TDA features derived from sliding-window embeddings to estimate moving boundary conditions in a testbed system, showing that maximum persistence in low-dimensional homology groups provides stable state estimates and can outperform short-time Fourier transform in rapidly changing regimes [92]. Similarly, Razmarashooli's work demonstrates the utility of TDA features for detecting impact-induced noise through higher-dimensional homology.

In building entropy-based measures of signal complexity, Myers et al. integrate TDA into permutation entropy frameworks to automatically select delay parameters, yielding parameter choices that align with expert recommendations and optimized settings across diverse dynamical systems [45]. Other pipelines exploit zigzag persistence and persistence-diagram-based change-point detection, such as PERsistence-diagram-based ChangE-PoinT detection (PERCEPT), a named online change-point detection framework for monitoring high-dimensional streams [93] and multilayer zigzag architectures for spatio-temporal meteorological forecasting [71].

Critical limitations: For real-time and nonstationary systems, the main weakness is that many evaluations rely on controlled testbeds, simulated regimes, or a small number of benchmark systems, so robustness under sensor drift, missing data, changing sampling rates, and strict latency constraints remains only partially established. Comparisons with short-time Fourier, entropy-based, and GNN baselines are informative, but they do not always include full ablations over embedding dimension, window length, homology degree, and vectorization. As in Table 8, the method is most credible when online computational cost and sensitivity analyses are reported explicitly.

6.3. Trajectory Analysis and Monitoring

Trajectory data (e.g., from hurricanes, animals, or vehicles) naturally encode spatio-temporal behavior. TDA offers a way to characterize the shape of trajectories rather than just pointwise positions. Esteve and Falco demonstrate that TDA-based representations can significantly improve trajectory classification accuracy, especially in hurricane trajectories and simulated scenarios [94]. In related work, they introduce tramoTDA, a named Python

library for TDA-based trajectory monitoring that provides user-friendly tools for visual and topological analysis of trajectories [95].

Topological methods have also been applied to Lagrangian orbits in convection flows, where persistent homology of alpha complexes reveals toroidal structures and transitions in flow regimes [12] and to orbit structures of more general flows on spherical surfaces via discrete representations and graph encodings that support TDA-driven classification [13].

Critical limitations: Trajectory studies demonstrate intuitive shape-based advantages, but their conclusions can depend strongly on resampling, normalization, distance metrics, and the availability of representative trajectory classes. Some results are obtained on simulated or domain-specific datasets, and comparisons with dynamic time warping, hidden Markov models, recurrent neural networks, or transformer-based sequence models are not always equally strong. External validation on different sensors, geographic regions, or flow regimes is needed before these topological descriptors can be regarded as generally transferable.

For time series and dynamical systems, TDA helps identify regime changes, transitions, and recurrent structure that can be difficult to capture with standard statistics alone. Embedding and windowing choices are central because they define the geometry on which topology is computed. When tuned carefully, topological summaries support forecasting, monitoring, and anomaly detection across finance, sensing, and physical systems. In practice, TDA works best as a complement to classical signal-processing and statistical methods.

7. Biomedical, Biological, and Neuroscientific Applications

The connection between mathematics and the life sciences is evident in many applications, including algebraic and vector-based representations of biological data that helped shape modern bioinformatics. This motivates the study of TDA in biological and medical settings, as reviewed in this section.

7.1. Precision Medicine and Gene Expression

TDA has become an important tool in biomedical data analysis, especially where heterogeneous and high-dimensional data are involved. Loughrey et al. develop a method for subgroup discovery in precision medicine using Mapper, with a focus on breast cancer data [9]. Their hotspot detection algorithm identifies homogeneous and geometrically compact subsets of patients with distinct clinical or molecular profiles and incorporates hotspot existence into Mapper parameter selection. The method reveals subgroups of estrogen receptor-positive patients with poor prognosis and specific expression signatures, validated on an independent dataset.

Mashatola et al. apply enhanced Vietoris–Rips complexes with topological overlapping measures to cancer gene expression data, showing that the resulting persistent features improve cancer phenotype prediction by up to 20% across multiple cancer types [10]. Narender et al. combine TDA, graph convolutional networks, and support vector machines (SVMs) for genomic expression classification-based phenotype prediction, achieving improvements over traditional approaches and highlighting TDA's role in high-dimensional feature extraction and network-based modeling [96]. Further applications include TDA-guided drug repurposing to tackle antibiotic resistance [97] and the analysis of antibody dynamics to stratify COVID-19 severity levels [98]. In [99], TDA delineates known breast cancer subtypes and identifies a new subtype within luminal B, together with its defining features. In [100], a TDA-radiomics investigation of ultrasound data suggests that a quantitative ultrasound risk-stratification score (US RSS) may improve the preoperative prediction of follicular carcinoma.

In biomedical data analysis, TDA offers a natural framework for modeling heterogeneous and longitudinal data. Temporal filtrations and topological summaries have been used to study single-cell dynamics, enabling the identification of cell-state transitions beyond traditional clustering methods [63]. In clinical settings, TDA has been applied to electronic health records to construct pseudo-time representations that capture disease progression trajectories [62].

Persistent homology features have also been employed for outcome prediction, including relapse risk in acute lymphoblastic leukemia patients [101]. Hybrid approaches combining neural networks with TDA have been proposed for tasks such as induced pluripotent stem cell colony classification, improving robustness and interpretability [102]. Additionally, classifiers based on topological summaries of repeated-measurement data have demonstrated strong performance in longitudinal biomedical studies [64].

In population genetics, TDA has been explored for quantifying recombination and cross-population gene flow by combining graph constructions (e.g., minimal spanning networks) with topological filtering and cycle-based summaries [103].

TDA has also been proposed as a theoretical and computational framework for vascular disease characterization, where persistent homology-derived indices can complement descriptors of stenosis geometry and vessel morphology [104]. In cardiovascular research, TDA supports the analysis of signals, such as electrocardiography, photoplethysmography, and arterial stiffness, and may improve diagnosis and prognosis [105]. TDA has also been used to assess coronary atherosclerosis by providing new techniques for characterizing calcified and noncalcified plaques [106].

Critical limitations: Precision-medicine and gene-expression applications are especially vulnerable to small cohort sizes, batch effects, missing clinical covariates, class imbalance, and multiple-testing bias. Mapper-based subgroups may change with the lens, cover, overlap, and clustering algorithm, while persistence-based predictors may capture cohort-specific artifacts rather than biology. Clinical claims should therefore be supported by independent validation cohorts, survival or outcome analyses, biological plausibility checks, and transparent sensitivity studies for all topological hyperparameters.

7.2. Imaging and Segmentation

TDA also plays a role in biomedical imaging. De Benedictis et al. combine TDA and low-rank tensor decomposition to enhance magnetic resonance imaging (MRI)-based brain tumor detection and classification, using persistent homology to identify critical regions and improve interpretability of ML predictions, achieving high classification accuracy [107]. Rahman et al. propose TDA-SegUNet, a U-Net-based segmentation model that integrates persistence images derived from 0-dimensional and 1-dimensional homology to encode local and global shape information for brain tumor segmentation, outperforming state-of-the-art models on brain tumor segmentation (BraTS) datasets [76]. Hybrid pipelines that pair persistent homology with DL backbones have also achieved state-of-the-art performance in skin cancer diagnosis, as in basal cell carcinoma classification using telangiectasia and lesion topology [108].

More broadly, Paige and Patrangenaru demonstrate how cubical persistent homology and statistical methods on non-Euclidean spaces can distinguish and classify images of leaves with minimal preprocessing, with cubical homology yielding superior performance compared with alternative descriptors [22], while Percival et al. use TDA to reveal conserved heteroblastic and ontogenetic programs in vining plant leaves that were not visible to PCA or linear discriminant analysis [109]. Ereemeev uses TDA-based barcodes and tree representations to detect repeated structures in satellite images, showing that persistent features can drive robust image analysis pipelines [110]. At finer scales, cubical persistent

homology supports feature detection and hypothesis testing in extremely noisy Transmission Electron Microscopy (TEM) images of nanoparticles [47] and automatic recognition of morphological structures in 3D vertebra models [111].

Topological features have been integrated into image analysis pipelines to enhance robustness to noise and geometric variability. In segmentation tasks, fractional calculus-based descriptors combined with persistent homology have been shown to improve boundary detection and region characterization [112]. In digital pathology, TDA has been applied to quantify immune-cell spatial organization, providing interpretable biomarkers that correlate with clinical outcomes [68]. Additionally, topological representations have been used as imaging biomarkers for ultrasound tumor diagnosis. Wei et al. propose wavelet-transform topological descriptors (WT-TD), an ultrasound topological representation method for distinguishing benign from malignant tumors and supporting clinical decision-making [113]. TDA has also been applied in radiological imaging, including tumor characterization, cardiovascular imaging, and COVID-19 detection [114].

Critical limitations: Imaging and segmentation results depend heavily on preprocessing, segmentation quality, image resolution, intensity normalization, scanner/site effects, and annotation protocols. Topological descriptors can improve shape sensitivity, but they do not by themselves guarantee clinical robustness or calibration. Benchmark improvements should be interpreted cautiously unless the same train–test protocol is used for all baselines, scanner or institution shifts are tested, and ablations show that topology adds information beyond standard CNN, radiomic, and morphology features.

7.3. Neuroscience and EEG

Electroencephalography (EEG) analysis has traditionally relied on statistical, spectral, and ML methods that may be sensitive to artifacts and noise. Ling et al. review TDA applications in EEG signal processing, summarizing TDA-based pipelines for disease diagnosis, brain state recognition, and perception evaluation, and highlighting strengths and limitations of TDA compared to conventional methods [115]. Zheng et al. [116] propose a TDA-based pipeline for multi-channel EEG, using Hilbert–Huang transforms to obtain instantaneous frequency and amplitude curves and then extracting TDA features for classification tasks. Their method achieves superior performance in Brain–Computer Interface (BCI) competitions and other EEG datasets, suggesting that topological features can capture informative structures in complex signals.

More specialized applications include emotion recognition from functional brain networks constructed via phase-locking values and analyzed with persistent homology to extract rich multiband topological descriptors [117], brain functional network analysis for image-quality assessment using Grey-TDA models [118], and TDA-based early prediction of ventricular fibrillation from ECG dynamics [119]. Catanzaro et al. show that TDA-based summaries of task-driven functional magnetic resonance imaging (fMRI) signals in the anterior cingulate cortex can outperform conventional vectorizations when classifying motor-task conditions [120]. Structural brain analyses have used TDA to quantify altered white-matter covariance networks in maltreated children [121]; related persistence-based survival models have also been used in political science, for example, in democracy-survival analysis with persistence homology-informed functional PCA [122]. Moreover, TDA has been applied to multi-channel EEG alterations in attention-deficit/hyperactivity disorder (ADHD) [123], EEG signals in children with sleep apnea [124], resting-state EEG data for Parkinson’s disease classification using entropy-based topological features [125], and schizophrenia classification in adolescents via EEG signal embeddings [126].

In neuroscience and biosignal analysis, TDA has been used to extract robust features from high-dimensional and noisy recordings. Studies on EEG data have investigated prin-

cipled parameter selection strategies for time-delay embeddings, improving the stability of topological features across subjects and sessions [127].

Persistent homology-derived descriptors have also been applied to cardiac electrophysiology, enabling the detection of ventricular fibrillation and tachycardia through topological signatures of ECG signals [128]. In affective computing, TDA-based representations have been proposed as interpretable alternatives to black-box deep models, supporting explainable emotion recognition from physiological signals [129].

Critical limitations: EEG, ECG, and fMRI studies are highly sensitive to filtering, artifact rejection, channel montage, referencing, time-delay parameters, window length, and subject-level leakage. Small cohorts and repeated measurements can inflate apparent accuracy if splits are not subject-independent. Credible evaluations should include artifact controls, cross-subject, and cross-device validation, comparisons with spectral, entropy-based, and deep sequence baselines, and interpretation that is linked to neurophysiological or physiological mechanisms rather than only classification scores.

7.4. Proteins, Molecular Structure, and Interaction Networks

At the molecular level, TDA has been used to analyze protein binding pockets, cryptic sites, and interaction networks. Jiang and Lugo-Martinez integrate TDA with geometric deep learning to characterize protein pockets, combining global topological invariants from TDA with local structural representations from GNNs to identify niches within pockets and improve classification tasks [77]. Koseki et al. introduce a TDA-based framework for quantifying structural and interaction changes due to amino acid mutations, capturing persistent changes in protein–protein interfaces [130], and leverage related ideas to detect cryptic binding sites via mixed-solvent simulations and TDA summaries [131].

Beyond individual proteins, Karthick et al. propose quantum graph-based differential models with fractional calculus and TDA to study dynamic protein–protein interaction networks, extracting persistent topological features and detecting critical transitions in network structures, with implications for systems biology and drug discovery [132]. At the cellular scale, TDA has been used to quantify collective motion patterns in mesenchymal cell populations via time-varying point clouds and Bayesian calibration of agent-based models [133], as well as to track emergent ring structures, filament organization, and remodeling in developmental and wound-healing contexts [134].

In computational biology, TDA has been used to compare and track evolving morphologies and emergent structures. Examples include distinguishing parameter regimes in angiogenesis simulations [135], tracking collective cell-motion interfaces over time [136] and detecting the onset/timing of ring-like structures in filament networks via time-resolved topological features [137]. TDA has also been applied to organismal behavior, where topological summaries of posture/locomotion enable quantitative comparisons across conditions and support interpretable behavioral motifs [138].

Critical limitations: Molecular and interaction network applications depend on atom selection, distance metrics, conformational sampling, solvent and protonation assumptions, graph construction, and threshold choices. A model may learn family-, assay- or database-specific biases rather than transferable molecular mechanisms. Strong validation should therefore use held-out protein families or interaction contexts, compare against established docking, affinity, or structural descriptors, test conformational ensembles, and connect persistent features to experimentally meaningful sites or functions.

Biomedical and biological datasets often contain heterogeneous subgroups, and this is where TDA is especially useful. The reviewed studies show that topological summaries can reveal clinically meaningful clusters, robust biosignal patterns, and interpretable molecular structure descriptors. These benefits are strongest when findings are validated on inde-

pendent cohorts or supported by complementary biological evidence. For non-specialists, TDA can be viewed as a structured way to uncover hidden organization in complex life science data.

8. Engineering, Physical, and Infrastructural Systems

It is commonly accepted that the fields of engineering and physical systems involve essential subbranches for investigation such as energy grids and infrastructure networks. At the same time, TDA is a robust mathematical framework increasingly applied to such areas, as the present section proves in the following.

8.1. Fluid Dynamics and Pattern Formation

In fluid dynamics and pattern formation, TDA offers a way to characterize complex spatial structures and their evolution. A work analyzes convection-driven flows in cylindrical geometries via Poincaré maps and persistent homology, characterizing transitions from quasi-periodic to chaotic regimes in terms of torus knots and cycle statistics [12]. Mototake et al. propose a TDA-based procedure, combined with machine learning, to interpret pattern formation in magnetic domain systems, linking TDA features to underlying physical mechanisms and suggesting reduced models that capture observed dynamics [139]. Topology-aware analyses of flows on spherical surfaces similarly exploit discrete encodings of orbit structures to support TDA-based comparisons and classification [13].

In fluid mechanics, TDA has been used to extract coherent structures from high-dimensional flow data. Persistent homology-based analyses have been applied to turbulent jet flows, where topological features capture the evolution and interaction of vortical structures across scales [140].

Critical limitations: Fluid-dynamics and pattern-formation studies often rely on simulated, controlled, or narrowly parameterized regimes. Persistent features can change with sampling density, mesh resolution, Poincaré-section choice, vortex extraction, time discretization, and noise level. Topological conclusions are strongest when they are checked against physical conservation laws, classical diagnostics such as POD/DMD or vorticity measures, parameter sweeps, and experimental data rather than being treated as standalone evidence of a flow mechanism.

8.2. Mechanical Systems and Structural Roughness

Structural health monitoring and materials characterization are natural domains for TDA. Canot et al. analyze roughness surfaces of bonded assemblies using persistent homology and persistence images, quantifying voids, micro-cracks, and peak amplitudes and relating them to adhesion and fracture resistance of adhesives used in aeronautics [11]. Pei et al. use TDA with Morse theory to extract topological features from ultrasonic-guided wave signals for corrosion characterization of steel strands, showing that topological features correlate linearly with cross-section loss and outperform traditional time, frequency, and time–frequency domain features in capturing corrosion development [24]. Miller et al. apply persistent homology to pulmonary arterial trees in murine models of pulmonary hypertension, revealing pruning and remodeling signatures in vascular topology [141].

Condition monitoring of rotating machinery has also benefited from TDA: Jeung and Kwon design a robust multivariate time series classification model that uses TDA to extract consistent features from multi-sensor vibration data, enabling fault-tolerant condition monitoring even when some sensor channels fail [142]. TDA-based feature extraction on stator current signals has been applied to induction motor eccentricity fault detection, with persistent homology-derived features feeding machine learning models that generalize across unseen fault levels [143]. At the instrumentation level, TDA combined with Takens

embeddings has been used to visualize and classify instrument outputs and rigid-body dynamics based on orbit topologies [144].

TDA has proven effective in analyzing complex dynamical systems encountered in engineering. Persistent homology features have been used to detect chaos and regime transitions in nonlinear mechanical systems [145], as well as to characterize human balance dynamics from motion data [146].

In robotics and autonomous systems, topological descriptors have been employed for trajectory monitoring and anomaly detection [66], including driver-assistance and human-machine interaction scenarios [147]. Related work has applied TDA to cluster parametric vibration modes and operational states [148]. Moreover, topology-aware validity-domain constraints have been introduced to guide optimization and model-based design under uncertainty [65].

Critical limitations: Mechanical, vibration, and roughness studies are frequently evaluated on laboratory faults, balanced datasets, or controlled operating conditions. In deployed systems, sensor drift, missing channels, changing loads, unobserved fault classes, and strict latency constraints can alter the topology of embedded trajectories. Practical claims should therefore report window length and sampling sensitivity, cross-machine or run-to-failure validation, robustness to sensor loss, computational cost, and comparisons with established time-frequency, physics-based, and condition-monitoring baselines.

8.3. Infrastructure Networks and Resilience

Complex infrastructures such as water distribution networks and power grids are critical and increasingly stressed by disturbances. Selicato et al. apply persistent homology to water distribution networks, proposing a new resilience metric based on topological features that complements existing graph-theoretic measures and provides a richer characterization of system robustness under failure scenarios [14]. Wang et al. integrate TDA with deep belief networks and decentralized control strategies to build short-term voltage prediction models in wind-integrated power systems, improving resilience assessment and control of small disturbances in renewable-rich grids [72]. Rail infrastructure has been analyzed with Mapper and Betti numbers to understand track geometry anomalies and maintenance needs [149], while TDA-based multimodal change detection methods have been used to track ecosystem state transitions in large river systems such as the Upper Mississippi [150].

Topological approaches have also been proposed for resilience assessment in infrastructure systems. By combining persistent homology with Wasserstein distances, TDA enables quantitative comparison of system states before and after disruptions, supporting the analysis of robustness and recovery dynamics in complex engineered networks [151].

In materials science and chemistry, persistent homology features have been used as compact structure descriptors for prediction and comparison tasks, including studies on high-temperature cuprate superconductors [152], comparisons of different graphene forms [153], and structure-property analysis of endohedral metallofullerenes [154], whereas in [155], a different investigation of automatically segmented large porous structures into local geometric features was performed; the shape and size of a pore or the curvature of a solid ligament, which affects the macroscopic properties of the material, has been succeeded using Morse theory. In [156], TDA has been treated as an unsupervised machine learning tool to uncover classification criteria in complex inorganic crystal chemistries via persistent homology, proving a hierarchical classification scheme.

Critical limitations: Infrastructure and resilience analyses depend on how physical systems are abstracted into graphs, weighted networks, or point clouds. Persistence or Wasserstein distances may miss hydraulic constraints, capacities, control policies, cascading

effects, maintenance costs, or regulatory requirements. Materials descriptors similarly need physical validation rather than only classification accuracy. Reliable studies should compare topological summaries with domain simulators and graph-theoretic baselines, propagate uncertainty in failure scenarios, and show that topological features correspond to actionable engineering mechanisms.

In engineering and physical sciences, TDA provides compact descriptors of evolving structure, from flow regimes to network resilience and material morphology. This is valuable when behavior is nonlinear and cannot be summarized well by single-point indicators. Topological features often improve monitoring and control when fused with domain-specific models. The practical message is that topology adds system-level context to conventional engineering analytics.

9. Finance, Economics, and Social Systems

In the modern era, the challenges are many, difficult, and constantly increasing. To a large extent, societies are faced with economic issues and the consequences of social phenomena like public health and social behavior. Hence, the workplace of TDA has its own applications and contributions to all these manifestations.

9.1. Financial Markets and Risk

As discussed earlier, TDA is increasingly used in financial time series analysis and market structure characterization. In addition to works on change-point detection and forecasting [6–8,80,81], Mojdehi et al. use Ball Mapper and GNNs for credit-risk assessment in supply chain finance, demonstrating that topological features and network-based representations improve accuracy and F1 scores in bankruptcy prediction [33]. Kheneifar and Amiri extend this line of work to maritime finance, using TDA over correlation-based networks of shipping firms to extract persistence features that capture nonlinear risk patterns and enable more accurate loan default prediction [157]. Topological descriptors have also supported portfolio-level risk modeling, bubble detection, and systemic fragility analysis through their sensitivity to clustering and voids in correlation structures [80,81].

Beyond finance, TDA has been applied to behavioral data such as clickstreams, where session dynamics are modeled as Markov chains and persistent homology summaries help identify intervention points and discriminate buyer vs. non-buyer browsing patterns [158].

Related work also uses TDA to refine risk stratification for firm distress and default. By mapping firms into the feature space of classical credit-risk indicators (e.g., Altman-style factors) and visualizing the resulting point cloud, TDA can reveal heterogeneous failure regions that are not easily separated by simple threshold rules [159].

Critical limitations: Financial market and risk applications face nonstationarity, survivorship bias, look-ahead leakage, class imbalance, and changing regulatory or macroeconomic regimes. Topological features extracted from correlations, credit indicators, or transaction behavior can be unstable when the market universe, sampling window, or normalization changes. Evidence should include strictly temporal validation, stress-period tests, transaction-cost or default-cost-aware metrics, calibrated probabilities, and comparisons with strong econometric, credit-scoring, and modern sequence-learning baselines.

9.2. Elections, Public Health, and Social Behavior

Mancilla et al. integrate TDA with machine learning and geostatistics to predict voting preferences in a gubernatorial election, constructing geospatial and non-geospatial models that incorporate TDA-derived features and achieving successful prediction of the election winner while enabling spatial exploration of voting patterns [160]. Dey and Kundu analyze county-level COVID-19 vaccine acceptance in the United States using network-

based models and TDA-derived clustering methods, uncovering macro-level communities with distinct vaccination patterns and linking them to sociodemographic factors such as education, income, and region [161]. TDA-based analyses of COVID-19 incidence and antibody dynamics highlight non-binary severity structures and multiscale spatial patterns, supporting more nuanced epidemiological modeling [98]. A related study applies TDA, specifically the Mapper algorithm, to COVID-19 data from China [162].

Mobility and migration have also been studied with TDA. Vittorietti et al. develop a topological measure for the attitude to mobility of Italian students and graduates, representing educational and occupational trajectories as graphs and ranking them via distances between persistence diagrams [163]. At the decision-making level, circular intuitionistic fuzzy TDA and related multi-criteria models have been used for AI-assisted evaluation in healthcare supply chains and uncertain logistics [164]. In political science, persistence homology-informed Bayesian survival models reveal topological heterogeneity in democracy survival data and support new regularization schemes for deep survival networks [122].

Finally, TDA has been explored in diverse interpretability-oriented settings: as a tool for solving or structuring certain classes of visual reasoning tasks (e.g., Bongard problems) [165], for uncovering organic relationships in legal corpora via case/precedent structure [166], and in decision-analytic settings where topology-inspired constructions are combined with fuzzy multi-criteria methods [167].

Critical limitations: Social, electoral, mobility, and public health studies are vulnerable to ecological fallacy, spatial aggregation effects, missing covariates, measurement bias, and privacy constraints. Topological clusters may reflect data collection, normalization, or geography rather than causal social mechanisms. Results should therefore be presented as exploratory or hypothesis-generating unless supported by causal designs, out-of-region validation, uncertainty quantification, demographic fairness checks, and careful communication of ethical limits.

In finance, economics, and social applications, TDA is mainly used to capture collective structures such as clustering, regime transitions, and systemic fragility. It has supported forecasting and risk analysis as well as community-level studies in public health and mobility. The evidence is promising, but interpretation depends on aligning topological patterns with domain mechanisms and context. For non-specialists, TDA is best seen as an early-warning and structure-discovery lens rather than a stand-alone predictor.

10. Security, Adversarial Machine Learning, and Anomaly Detection

Data poisoning and adversarial attacks pose substantial threats to ML-based systems. Monkam et al. propose a TDA-based approach to detect data poisoning in network intrusion detection systems, using topological features and clustering to isolate clusters of poisoned data before training, thereby improving security without relying solely on classifier robustness [70]. Ferrara's game-theoretic analysis, mentioned earlier, provides a complementary theoretical perspective [91].

Topological features also appear in anomaly detection more broadly, such as in power system monitoring, meteorological forecasting, and sensor networks. For example, Ma et al. use zigzag persistence and supra-graph constructions in the ZPDSN model, the authors' named topological graph neural network architecture for spatio-temporal meteorological forecasting, capturing high-order structural information and outperforming conventional GNN-based methods on multiple meteorological variables [71]. Wang et al. incorporate TDA-derived information into voltage prediction and reactive power control strategies for wind farms, enhancing detection and mitigation of small disturbances [72]. At the data-management level, TDA has been explored as a tool to automatically detect data quality

faults and duplicate entities in large databases, providing an unsupervised complement to rule-based and record-matching approaches [168].

Topological summaries have also been explored in security-related analytics. For example, TDA has been used to visualize and characterize malicious packet patterns in darknet monitoring data, providing an interpretable global picture of attack activity beyond conventional traffic statistics [169]. In Natural Language Processing (NLP) security, TDA-derived features have been investigated as a complement to deep models for fake news detection, with particular gains reported in regimes where labeled training data are scarce [170].

Critical limitations: Security and anomaly-detection studies often rely on fixed benchmark datasets, known attack types, or controlled simulations, whereas deployed adversaries adapt to detection rules and concept drift. Topological summaries may be costly to update online and may generate false positives when benign traffic changes shape. Operational evaluations should include latency, memory use, adaptive attacks, drift scenarios, false-alarm costs, and comparisons with production intrusion detection, graph analytics, and sequence model baselines.

In security workflows, TDA helps characterize global structural patterns that can be missed by local feature checks. This can improve detection of poisoning behavior and spatio-temporal anomalies in complex systems. The approach is particularly useful when labels are limited or attack patterns evolve over time. In practice, topology is most effective as a complementary signal inside a broader defense pipeline.

11. Software Ecosystem and Practical Considerations

The growing TDA ecosystem includes both general-purpose libraries and domain-specific tools. In addition to widely used packages such as the Geometry Understanding in Higher Dimensions (GUDHI) library, Ripser, and Giotto-TDA, recent developments include:

- tramoTDA, a Python library for trajectory monitoring and classification that leverages persistent homology and TDA-based distances, designed for both technical and non-technical users [95].
- The multipers library and Core Delaunay constructions for multiparameter persistence, used in shape recognition experiments that distinguish synthetic shapes such as circles, spheres, tori, and coffee cups [23].
- Specialized pipelines that integrate TDA with wavelet transforms, Hilbert–Huang transforms, and domain-specific preprocessing steps in EEG, financial, and environmental applications [7,116,171].

Practical deployment of TDA-based workflows raises several issues:

- (a) *Parameter selection*: Choosing scale parameters, filter functions, and embedding windows is nontrivial. Methods such as TDA-based delay selection [45] and Mapper parameter exploration guided by hotspot detection [9] represent important steps, but automated and principled parameter selection remains an open challenge.
- (b) *Scalability*: Data reduction methods such as CLA [44] and enhanced complexes [10] help manage computational load, but scaling to truly massive datasets (e.g., billions of points, large streaming graphs) is still difficult.
- (c) *Interpretability*: While persistence diagrams and Mapper graphs are conceptually interpretable, translating them into domain-specific insights often requires collaboration with subject-matter experts and careful experimental design.

On the software side, giotto-tda provides a scikit-learn-compatible toolkit that makes common TDA pipelines (preprocessing, persistent homology, and vectorizations) easier to integrate into ML workflows [172]. For interactive analysis at scale, topology-driven

visualization and aggregation methods have been proposed to enable exploration of high-dimensional model behavior on very large scientific datasets [173].

The software ecosystem is now mature enough for non-specialists to prototype TDA using standard data-science tools. However, libraries do not remove the need for careful parameter choice and sensitivity checks. Reproducible conclusions still depend on transparent preprocessing, filtration design, and evaluation. A practical starting point is to use simple documented pipelines before moving to specialized high-complexity variants.

12. Additional Recent Applications of TDA

To avoid a long catalog of loosely connected examples, this section now shows only representative cases that clarify the five-stage design patterns in Section 4; the broader reference list is moved to Appendix A. The first pattern is *topology as a task-specific classifier*. Kindelan et al. provide a compact example because the topological component is part of the classification strategy itself rather than only an auxiliary visualization [174]. Related imaging and manufacturing examples, such as Matrix-Assisted Laser Desorption/Ionization (MALDI) tumor typing and wafer-defect recognition, are treated as domain-specific variants of the same pattern [69,175].

The second pattern is *topology as a morphology-aware biomedical or signal descriptor*. Representative fMRI, ECG/sleep, and functional network studies show the same workflow: construct a biologically meaningful image, signal, or network representation; choose a filtration compatible with that representation; vectorize persistence outputs; and test whether topology adds information beyond conventional descriptors [117,119,176,177]. These cases are more useful for the main argument than a longer biomedical list because they expose where design choices enter the pipeline.

The third and fourth patterns are *topology as a global transition detector* and *topology as an exploratory lens for structured non-Euclidean data*. Physical, ecological, geophysical, political, and mobility studies illustrate transition, roughness, connectivity, and regime heterogeneity questions [11,122,139,150,163,178]; spatial, cultural, and symbolic data studies show how the same pipeline can be transferred once a defensible representation and filtration are defined [179–182]. The important point is not the number of application domains, but whether each example clarifies a reusable design pattern and is supported by appropriate validation.

13. From Catalog to Critical Synthesis: What Works, What Fails, and Where

The domain-by-domain review above can be condensed into a smaller set of recurring empirical patterns. Across finance, biomedicine, engineering, and security applications, TDA is most reliable when it is used as a *structural prior* that complements statistical or neural models, rather than as a stand-alone replacement for the full learning pipeline [5,9,70,71]. The purpose of this section is to synthesize consistent evidence and to separate robust effects from context-dependent claims. Table 8 presents the corresponding special investigation of this critical synthesis.

13.1. What Consistently Works

- *Multiscale summarization of complex structure*: persistent homology and related summaries repeatedly provide compact, noise-tolerant descriptors of nontrivial geometry in time series, graphs, and imaging data [7,10,69].
- *Hybridization with classical or deep models*: the most stable performance gains occur when topological summaries are fused with domain features and learned representations, especially in forecasting, diagnosis, and anomaly detection [8,16,18,70].

- *Interpretability at mesoscopic scale:* Mapper/persistence-based analysis is consistently useful for identifying subgroups, transition regimes, and failure regions that are difficult to detect with purely local statistics [9,159,161].

13.2. What Repeatedly Fails or Becomes Fragile

- *Hyperparameter sensitivity:* filtration design, metric choice, and embedding parameters can change conclusions qualitatively; this is a major source of instability when parameter sweeps are weakly justified [183,184].
- *Scalability bottlenecks:* large Vietoris–Rips constructions remain expensive, and aggressive approximations can remove informative fine-scale structures [44,185].
- *Signal dilution after vectorization:* converting diagrams into fixed-length vectors is often necessary for ML pipelines, but can weaken geometric meaning and reduce portability across datasets [90,186].
- *Evaluation fragility:* positive results are frequently reported on small or domain-specific benchmarks with limited ablation and sensitivity analysis, making cross-study comparisons difficult [187].

13.3. In Which Settings Each Pattern Is Most Reliable

Across domains, the strongest evidence supports using TDA as a complementary structural component, not as a universal stand-alone solution. Reliable gains appear when datasets are noisy, multiscale, or weakly labeled and when topological features are integrated with domain-aware models. Fragility appears when hyperparameters are weakly justified, evaluations are narrow, or scalability shortcuts remove informative structures. The practical rule is to use TDA to enrich existing pipelines and verify robustness with clear baselines and sensitivity analyses.

13.4. Expanded Discussion of Cross-Application Results

The applications reviewed in this survey point to a common interpretation: TDA is most useful when the research question concerns structure, transitions, or heterogeneity rather than only pointwise prediction accuracy. In financial and other nonstationary time series settings, the main result is not simply that persistence-based features can improve forecasting, but that sliding-window topology can expose regime changes, market instability, and evolving dependence patterns that are difficult to summarize with local statistics alone [5–8]. The same principle appears in dynamical systems and sensor applications, where topological summaries function as early-warning or monitoring descriptors rather than as isolated classifiers [71,92,94].

In biomedical and biological applications, the strongest contribution of TDA is its ability to reveal subgroup structure and shape-driven biomarkers under high dimensionality, class imbalance, and noisy measurement. Mapper-based precision-medicine studies, enhanced Vietoris–Rips constructions for gene-expression data, and imaging or segmentation pipelines all illustrate that topology can support stratification and morphology-aware learning [9,10,69,76]. However, these results should be interpreted as evidence for hypothesis generation and model enrichment unless they are supported by external validation cohorts, stable preprocessing, and clinically meaningful sensitivity analyses.

Engineering, physical science, infrastructure, and security studies show a second recurring pattern: topology is valuable when global connectivity or multiscale geometry carries operational meaning. In materials, fluid flows, surface roughness, and structural monitoring, persistent summaries can encode geometric organization that is not captured by scalar descriptors alone [11,12,92,155]. In infrastructure and cybersecurity, graph-topological features help characterize resilience, anomaly patterns, or poisoned clusters by preserving higher-order relationships among system components [14,70,72,91]. These applications

indicate that TDA is most convincing when topological features can be linked back to a domain mechanism, such as connectivity, recurrence, roughness, or failure propagation.

Across ML and DL applications, the main result is that topology usually works best as an auxiliary representation, regularizer, or interpretability layer. Persistence images, landscapes, graph-topological features, and differentiable topological losses can improve robustness or reveal model behavior, but their benefit depends strongly on architecture, vectorization, and evaluation design [15–19]. Consequently, the overall conclusion from the application survey is balanced: TDA offers a transferable language for complex structures, but its practical value is highest when it is paired with domain knowledge, strong non-topological baselines, repeated sensitivity checks, and transparent reporting of computational choices.

14. Design Choices, Limitations, and Common Pitfalls

Beyond the application-driven studies discussed above, several works address methodological, theoretical, and evaluation-oriented aspects of topological data analysis. These include analyses of robustness and uncertainty in topological summaries [186,188], as well as investigations into stability and convergence properties of persistence-based constructions under different sampling and noise regimes [183,184]. Related contributions examine comparative evaluation and benchmarking considerations for TDA pipelines in applied settings [90,187] and the integration of topological descriptors into multi-criteria decision-making frameworks [189]. Together, these studies complement application-focused work by clarifying the limits, assumptions, and practical reliability of topology-based methods.

While topological data analysis has been shown to have a wide range of applications in many fields, its successful application depends on a number of non-technical choices. This section will highlight the common difficulties that have been identified in the literature, grouped according to the steps in the pipeline discussed in the previous section.

14.1. Sensitivity to Data Representation and Filtration Design

The choice of data encoding (Stage I) and filtration (Stage II) often controls downstream performance. Small changes in distance metrics, density estimation, or thresholding can lead to qualitatively different topological patterns. This is particularly evident in time series embeddings and graph filtrations, where scale parameters are implicitly tied to modeling choices. Filtration designs are rarely portable across datasets, which limits out-of-the-box use.

14.2. Stability–Expressivity Trade-Offs

Persistent homology is robust to perturbations of the input data. However, this robustness comes at the cost of reduced expressivity, as there are topological representations that ignore geometric and metric details potentially useful for learning. In contrast, more expressive representations, such as learned or high-dimensional vector representations, might improve performance but are likely to break the robustness guarantee of persistence. Therefore, balancing robustness, discriminative ability, and interpretability is still an open problem.

14.3. Scalability and Computational Constraints

Despite progress in algorithmic improvements, computational scalability remains a bottleneck, especially for large point clouds and high-dimensional data used to build Vietoris–Rips complexes. In practice, pipelines often rely on subsampling, sparsification, or approximate algorithms, which may introduce bias or lose fine-scale information. Although cubical complexes alleviate these issues for image data, scalable algorithms for general metric spaces remain an active research area.

14.4. Vectorization and Loss of Interpretability

Most learning applications rely on vectorizing persistence outputs (Stage IV). While summaries such as landscapes, entropy, and codebooks enable compatibility with conventional ML frameworks, they can weaken the geometric interpretability of topological features. This trade-off directly affects one of the original motivations of topological data analysis.

14.5. Evaluation, Validation, and Reproducibility

A persistent concern in the literature is the lack of standardized evaluation protocols. Reported performance is often constrained by small or domain-specific datasets. Moreover, sensitivity analyses for filtration parameters and embedding choices are not always well documented. Progress in this area requires clearer reporting of design decisions and benchmark datasets tailored to topological approaches.

Compact reproducibility checklist.

Table 9 summarizes a minimum reporting standard for practical reproducibility in TDA studies.

Illustrative application of the checklist.

To make Table 9 operational rather than merely prescriptive, Table 10 applies it to three representative articles that are central to the application survey and correspond to references [9,10,56] in the current reference list. The audit is intentionally conservative: “not clearly reported” means that the item is not sufficiently explicit in the article-level reporting reviewed here to permit direct reproduction by an independent reader, not that the information was necessarily unavailable to the original authors.

Table 9. Compact reproducibility checklist for TDA studies.

Checklist Item	Minimum Information to Report
Filtration parameters	Complex type, metric, filtration function, parameter search range, selected values, and selection rationale.
Random seeds	All seeds used for splitting, subsampling, model initialization, and repeated runs.
Data splits	Exact training, validation, and test protocol, stratification or grouping constraints, and sample counts per split.
Compute budget	Hardware details (central processing unit (CPU), graphics processing unit (GPU), or tensor processing unit (TPU)), wall-clock runtime, and peak memory for training and inference.
Code availability	Repository or archive link, commit hash/version tag, and scripts for reproducing the main tables and figures.

Across this small audit, all three representative studies describe the scientific dataset or domain setting and all provide at least a conceptual description of the TDA construction. However, none clearly satisfy the full checklist in Table 9: the open executable code, random seeds, complete parameter grids, hardware/runtime information, and scripts for regenerating the main figures or tables are not consistently reported. This reinforces the practical message of the survey: transparent TDA reporting should include not only diagrams and accuracy values but also the exact filtration design, stochastic choices, computational budget, and reproducible analysis scripts.

Summary.

These limitations show that TDA is usually best treated as an inductive bias rather than a replacement for full statistical modeling. Small design choices in representation, filtration, and vectorization can materially change conclusions. Credible TDA studies therefore require transparent reporting, sensitivity analyses, and external validation whenever

possible. For non-specialists, the central message is that methodological discipline matters as much as the topological tool itself.

Table 10. Illustrative reproducibility audit for representative TDA application studies, using the checklist in Table 9.

Representative Study	Data and Splits	Filtration Parameters	Seeds and Compute	Code and Main Gap
Precision-medicine Mapper study [9]	Reported: breast cancer case study and independent validation cohort.	Partial: Mapper and hotspot logic described; full executable grid not explicit.	Not clearly reported.	Code not clearly reported; exact filters, clustering settings, seeds, and regeneration scripts would be needed.
Enhanced Vietoris–Rips study in genomics [10]	Reported: cancer gene-expression datasets and preprocessing context.	Partial but strong: enhanced Vietoris–Rips construction described; complete threshold/search configuration not fully reusable.	Not clearly reported.	Code not clearly reported; end-to-end scripts and model configurations would strengthen reproduction.
Time-resolved market-instability study [56]	Partial: market data and crisis windows described; no prospective split.	Partial: time-resolved design described; reusable windows, metrics, and thresholds not packaged.	Mostly deterministic; compute not clear.	Code not clearly reported; data-access instructions and executable parameters would be required.

15. Open Challenges and Future Directions

Despite rapid progress, several fundamental questions remain. In this section, we organize the main open challenges into thematic categories.

15.1. Open Problems and Research Directions

We group current research priorities into five interconnected themes.

15.1.1. Scalability

A central challenge is to make TDA reliable for large datasets through sparse or approximate constructions, GPU-accelerated computation, and streaming TDA pipelines that update summaries online.

15.1.2. Statistical Foundations

Key questions concern calibrated confidence sets for persistence diagrams and rigorous hypothesis testing under finite-sample and dependent-data regimes.

15.1.3. Integration with AI

An active direction is coupling TDA components with DL and graph neural networks so that topological signals improve generalization without destabilizing training.

15.1.4. Interpretability

Topological descriptors can support explainable machine learning by linking model predictions to robust geometric and structural patterns in data.

15.1.5. Multiparameter Persistence

Major open problems include representational choices, invariant design, and severe computational complexity in multifiltration settings.

15.2. Scalability and Streaming Data

Persistent homology computation for large Vietoris–Rips complexes is still a bottleneck. For example, Manu Aggarwal et al. presented an efficient and scalable algorithm for computing persistent homology of sparse Vietoris–Rips complexes on larger datasets, including a high-resolution human-genome application based on a genome-wide Hi-C dataset containing approximately three million points [185].

The central practical issue is the accuracy–efficiency trade-off. Exact Vietoris–Rips persistence is attractive because it is simple and reproducible, but the number of simplices grows combinatorially with the number of points, the filtration threshold, and the maximum homology dimension. Approximation strategies reduce this burden in different ways: sampling and data reduction methods replace the full point cloud by a smaller representative set; skeletonization fixes a low maximum dimension and discards high-dimensional simplices; sparse Vietoris–Rips constructions prune edges or cofaces; and witness complexes use landmarks to summarize a much larger set of witnesses. These choices improve runtime and memory use, but they may suppress short-lived or spatially localized features. Therefore, approximation should be reported together with stability checks, such as repeated landmark selections, bottleneck or Wasserstein distances between diagrams, and downstream ablation tests [3,44,185].

Table 11 suggests a practical decision rule for end users. Start with an exact Ripser computation in low dimension on a controlled subset to establish a reproducible baseline. If memory or runtime becomes prohibitive, first lower the maximum dimension and filtration diameter, then compare at least two approximations—for example, CLA or farthest-point sampling versus sparse Rips or witness complexes—and keep the cheapest option whose persistence diagrams and downstream scores remain stable. Use GUDHI when the data naturally require alpha, cubical, witness, or custom filtrations, and use giotto-tda when the main requirement is integration with a machine learning pipeline. In all cases, the report should include the number of points, retained landmarks or edges, maximum homology dimension, filtration threshold, software version, hardware, wall-clock time, peak memory, and sensitivity analysis.

Table 11. Indicative scalability regimes and trade-offs for common persistent homology strategies, including point counts, maximum dimensions, and typical computation time/memory behavior. The ranges are empirical rules of thumb rather than hardware-independent guarantees; runtime depends strongly on metric structure, filtration threshold, coefficient field, retained edges or simplices, implementation version, and available memory.

Scaling Strategy and Software	Typical Input Size	Typical Max. Dimension	Typical Computation Time/Memory Profile	Accuracy Trade-Off and User Advice
Exact Vietoris–Rips with Ripser	10^3 – 5×10^4 points when thresholds are sparse	H_0 – H_2 ; higher only for small data	Seconds–minutes for sparse H_0 – H_1 jobs; minutes–hours or memory-limited for dense H_2 ; usually fastest for pure Vietoris–Rips benchmarks	Use as a baseline for small/medium metric data; cap the diameter and dimension before increasing n [3,190].
General complexes with GUDHI	10^3 – 10^5 points depending on alpha, cubical, witness, or sparse Rips construction	H_0 – H_2 in most applied workflows	Seconds–minutes for alpha or cubical cases; minutes–hours when explicit Rips or witness complexes become large; memory grows with represented simplices	Prefer when geometry-specific complexes, cubical data, witness complexes, or custom filtrations are needed [191].

Table 11. Cont.

Scaling Strategy and Software	Typical Input Size	Typical Max. Dimension	Typical Computation Time/Memory Profile	Accuracy Trade-Off and User Advice
Pipeline-oriented giotto-tda/giotto-ph	10^2 – 10^4 points per sample, often repeated over many samples in ML pipelines	H_0 – H_2	Seconds–minutes per sample; total time multiplies over cross-validation folds but parallel execution reduces wall-clock cost	Use for cross-validation, model selection, and reproducible ML workflows; still tune filtration thresholds and validate vectorization choices [172,192].
Sampling or data reduction, including CLA or farthest-point subsampling	Reduce n to $m \ll n$, often $m = 10^2$ – 10^4 landmarks	H_0 – H_2 after reduction	Preprocessing plus persistence is often seconds–minutes once m is fixed; memory is driven by m^2 distances or retained simplices	Efficient first option for very large point clouds, but repeat the sampling and quantify diagram variation to avoid losing rare features [44].
Skeletonization and sparse Vietoris–Rips	10^4 – 10^6 points when the retained graph is sparse	Usually H_0 – H_1 , sometimes H_2	Minutes–hours for large sparse graphs; storage scales with retained edges or cofaces rather than all possible simplices	Good for global structure and streaming settings; report the sparsification rule and compare against an exact small-subset baseline [185,190].
Witness or landmark complexes	Large witness set n with 10^2 – 10^4 landmarks	H_0 – H_2	Landmark selection and witness–landmark distances often dominate; persistence on landmarks is commonly seconds–minutes	Useful when interpretability and memory limits dominate; sensitivity to landmark choice should be tested explicitly [3,191].

In view of these trade-offs, future research directions should include:

- Randomized and sketch-based approximations of persistent diagrams.
- Multi-resolution and hierarchical methods that focus computational effort on informative regions of parameter space.
- Streaming algorithms that update topological summaries incrementally as data arrive.

Integration with hardware accelerators, such as graphics processing units (GPUs) and tensor processing units (TPUs), and distributed frameworks is likely to be crucial for real-time applications such as online risk monitoring and cyber–physical system control.

15.3. Statistical Foundations and Limits

Works such as [48] emphasize that TDA-based summaries are not universally informative for statistical inference. Key open problems include:

- Developing confidence sets and hypothesis tests for persistence diagrams, Betti curves, and related summaries.
- Understanding identifiability and sufficiency conditions for topological statistics.
- Designing Bayesian and likelihood-based models that incorporate topological information in a principled way.

Further progress will likely draw on advances in random topology [27], dependence network simulation and inference [49,50], and the study of topology-aware loss functions and regularizers in deep and graphical models [17,18].

15.4. Integration with Modern AI Architectures

The rapid development of TDA and AI, both separately and in combination, is creating new opportunities for methodology and applications. The literature surveyed here reports successful integrations of TDA with deep learning, graph neural networks, and hybrid neuro-symbolic architectures [15–19,193]. Future directions include:

- Differentiable persistent homology layers with better gradient properties and scalability.
- Topological regularizers that enforce global constraints (e.g., connectivity, number of holes) in generative and discriminative models.

- Topology-aware transformers and foundation models that operate over graphs, manifolds, and multimodal data.
- End-to-end pipelines where topological and neural components co-adapt during training.

Quantum-inspired and quantum-accelerated TDA methods [51,52] may also play a role in large-scale or resource-constrained scenarios, provided that data-loading and complexity constraints can be adequately addressed.

Future progress in TDA depends on advances in scalability, stronger statistical foundations, and better integration with modern AI architectures. Faster algorithms alone are not enough unless uncertainty quantification and validation standards also improve. Likewise, topology-aware neural models are promising only if they remain interpretable and computationally practical. The field is moving from proof-of-concept successes toward robust and deployable methodology.

16. Conclusions

From a specialized mathematical toolkit, topological data analysis has evolved into a broader framework used across data science, ML, and scientific computing. Historically, the core notions of persistent homology, persistence diagrams, and barcodes emerged between the 1990s and 2010s, and recent technological advances have made both theoretical and practical study of TDA increasingly important. The literature surveyed here demonstrates the versatility and relevance of TDA, including applications to change-point detection in financial markets, improved forecasting and recommendation systems, analysis of biological and neuronal networks, interpretation of DL systems, defense against poisoning attacks, and resilience analysis of infrastructure networks.

At the same time, a critical synthesis of recent evidence indicates that TDA works most consistently as an inductive bias and structural descriptor, not as a universal stand-alone predictor. Strong outcomes are most frequent in settings with multiscale geometry, heterogeneous subpopulations, and noisy or partially labeled data, especially when topological features are fused with domain-informed models. Conversely, unstable filtration design, limited benchmarking, and scalability bottlenecks remain recurrent failure modes. Addressing these issues requires coordinated progress in algorithms, statistical validation, and topology-aware AI architectures, together with closer collaboration between mathematicians, domain scientists, and ML practitioners.

To conclude with an actionable agenda, we replace broad claims with concrete research questions that can guide near-term work:

- (Q1) How can persistent homology pipelines for large Vietoris–Rips complexes be made truly scalable (e.g., in streaming and distributed settings) while retaining provable approximation guarantees and bounded memory?
- (Q2) Which statistical procedures provide calibrated uncertainty quantification for persistence-based summaries (diagrams, landscapes, Betti curves) under realistic assumptions such as dependence, heteroskedastic noise, and finite samples?
- (Q3) How should filtrations and hyperparameters be selected in a data-driven yet interpretable way, and what sensitivity analysis protocols should be reported as a minimum standard for reproducible TDA studies?
- (Q4) What benchmark design (datasets, tasks, metrics, ablations) best isolates the added value of topological features over strong non-topological baselines across domains?
- (Q5) How can differentiable topological modules be integrated into modern deep and graph architectures so that training remains stable, computationally feasible, and scientifically interpretable?

Answering these questions would move TDA from promising demonstrations toward robust, validated, and routinely deployable methodology across scientific and engineering

applications. Practitioners should nevertheless treat TDA as a sensitivity-dependent component rather than a plug-and-play predictor: filtration and embedding hyperparameters can change the resulting diagrams, vectorization can dilute localized topological signals, and the absence of standardized validation protocols makes cross-study comparisons fragile.

Author Contributions: Conceptualization, S.K. and D.G.; methodology, F.S.; formal analysis, D.G.; investigation, F.S.; data curation, S.K.; writing—original draft preparation, S.K. and F.S.; writing—review and editing, D.G.; visualization, F.S.; supervision, S.K.; project administration, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Extended Catalog of Recent TDA Applications

This appendix preserves representative recent examples that support the cross-domain scope of the survey without interrupting the critical synthesis in the main text. The examples are grouped by methodological role rather than by application domain, because the main transferable issue is how the data representation, filtration, topological summary, vectorization, and downstream validation are aligned.

- Classification and recognition tasks: TDA has been used in task-specific classifiers for biomedical and industrial recognition problems, including MALDI tumor typing, wafer-defect recognition, and related imaging settings, where persistence summaries provide morphology-sensitive features [69,174,175].
- Biomedical and physiological descriptors: fMRI, ECG/sleep, and functional network studies use topological summaries to represent multiscale shape or connectivity patterns before downstream statistical or ML analysis [117,119,176,177].
- Transitions in physical and ecological systems: Applications in physics, ecology, and geoscience show how persistence-based summaries can mark roughness, connectivity, phase transitions, or regime changes in complex empirical data [11,139,150,178].
- Structured social, cultural, and symbolic data: Political, mobility, spatial, cultural, and symbolic data studies illustrate the use of TDA as an exploratory lens once a defensible metric, graph, or embedding has been specified [122,163,179–182].

Across these examples, the recurring lesson is that TDA contributes most clearly when the topological construction is tied to a domain mechanism and evaluated against non-topological baselines. The catalog is therefore not intended as an exhaustive list of applications; it records additional evidence for the design patterns summarized in Section 4.

References

1. Akomodj, J.O. Exploring the Applications of Algebraic Topology in Modern Mathematics and Data Analysis. *Int. J. Recent Innov. Acad. Res.* **2025**, *9*, 334–351.
2. Edelsbrunner, H.; Harer, J. *Computational Topology: An Introduction*; American Mathematical Society: Providence, RI, USA, 2010.
3. Otter, N.; Porter, M.A.; Tillmann, U.; Grindrod, P.; Harrington, H.A. A roadmap for the computation of persistent homology. *EPJ Data Sci.* **2017**, *6*, 17. [CrossRef] [PubMed]
4. Wasserman, L. Topological data analysis. *Annu. Rev. Stat. Its Appl.* **2018**, *5*, 501–532. [CrossRef]
5. Nie, C. Unveiling complex nonlinear dynamics in stock markets through topological data analysis. *Phys. A Stat. Mech. Its Appl.* **2025**, *680*, 131025. [CrossRef]
6. Yao, J.; Li, J.; Wu, J.; Yang, M.; Wang, X. Change Point Detection in Financial Market Using Topological Data Analysis. *Systems* **2025**, *13*, 875. [CrossRef]

7. de Jesus, L.C.; Fernández-Navarro, F.D.A.; Carbonero-Ruz, M. Enhancing financial time series forecasting through topological data analysis. *Neural Comput. Appl.* **2025**, *37*, 6527–6545. [\[CrossRef\]](#)
8. Kulkarni, S.; Pharasi, H.K.; Vijayaraghavan, S.; Kumar, S.; Chakraborti, A.; Samal, A.K. Investigation of Indian stock markets using topological data analysis and geometry-inspired network measures. *Phys. A Stat. Mech. Its Appl.* **2024**, *643*, 129785. [\[CrossRef\]](#)
9. Loughrey, C.F.; Maguire, S.L.; Dłotko, P.; Bai, L.; Orr, N.J.; Jurek-Loughrey, A. A novel method for subgroup discovery in precision medicine based on topological data analysis. *Bmc Med. Inform. Decis. Mak.* **2025**, *25*, 60. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Mashatola, L.; Kader, Z.; Abdulla, N.; Kaur, M. Enhancing the Vietoris–Rips simplicial complex for topological data analysis: Applications in cancer gene expression datasets. *Int. J. Data Sci. Anal.* **2025**, *20*, 1383–1400. [\[CrossRef\]](#)
11. Canot, H.; Durand, P.; Frénod, E.; Hassoune, B.; Nassiet, V.; Tramis, O. Topological data analysis for roughness surfaces of bonding assembly. *Discret. Contin. Dyn. Syst.-Ser. S* **2025**, *18*, 1743–1766. [\[CrossRef\]](#)
12. Núñez, J.; González, A.; Ramos, E.S. Topological data analysis of Lagrangian orbits in natural convection flows confined in a cylinder. *Phys. Rev. Fluids* **2022**, *7*, 123501. [\[CrossRef\]](#)
13. Sakajo, T.; Yokoyama, T. Discrete representations of orbit structures of flows for topological data analysis. *Discret. Math. Algorithms Appl.* **2023**, *15*, 2250143. [\[CrossRef\]](#)
14. Selicato, L.; Pagano, A.; Esposito, F.; Icardi, M. Topological data analysis for resilience assessment of water distribution networks. *Math. Comput. Simul.* **2025**, *231*, 62–70. [\[CrossRef\]](#)
15. Zhang, B.; He, Z.; Lin, H. A comprehensive review of deep neural network interpretation using topological data analysis. *Neurocomputing* **2024**, *609*, 128513. [\[CrossRef\]](#)
16. Pham, P. An Integrated Fuzzy Neural Network and Topological Data Analysis for Molecular Graph Representation Learning and Property Forecasting. *Mol. Inform.* **2025**, *44*, e202400335. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Pham, P.; Bui, Q.T.; Nguyen, N.T.; Kozma, R.; Yu, P.S.; Vo, B. Topological Data Analysis in Graph Neural Networks: Surveys and Perspectives. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 9758–9776. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Hu, K.; Gong, S.; Li, L.; Luo, Y.; Li, Y.; Jiang, S. Federated Incremental Learning algorithm based on Topological Data Analysis. *Pattern Recognit.* **2025**, *158*, 111048. [\[CrossRef\]](#)
19. Ballester, R.; Clemente, X.A.I.; Casacuberta, C.; Madadi, M.; Corneanu, C.A.; Escalera, S. Predicting the generalization gap in neural networks using topological data analysis. *Neurocomputing* **2024**, *596*, 127787. [\[CrossRef\]](#)
20. Chazal, F.; Michel, B. An introduction to topological data analysis: Fundamental and practical aspects for data scientists. *Front. Artif. Intell.* **2021**, *4*, 667963. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Bauer, U.; Kerber, M.; Roll, F.; Rolle, A. A Unified View on the Functorial Nerve Theorem and its Variations. *Expo. Math.* **2023**, *41*, 125503. [\[CrossRef\]](#)
22. Paige, R.L.; Patrangenaru, V.P. Novel Statistical and Topological Data Analyses of 2D Electronic Images. *J. Stat. Theory Pract.* **2025**, *19*, 53. [\[CrossRef\]](#)
23. Janes, F.; Mwanzalima, M.; Mpimbo, M. On shapes recognition in topological data analysis. *Res. Math.* **2025**, *12*, 2573577. [\[CrossRef\]](#)
24. Pei, Y.; Zhang, J.; Yuan, Y.; Yang, D.; Au, F.T.K. Ultrasonic guided wave topological data analysis for corrosion characterization of steel strand based on Morse theory. *Mech. Syst. Signal Process.* **2025**, *237*, 113119. [\[CrossRef\]](#)
25. Bjerkevik, H.B. Stabilizing Decomposition of Multiparameter Persistence Modules. *Found. Comput. Math.* **2025**, *26*, 939–998. [\[CrossRef\]](#)
26. Dey, T.K.; Xin, C. Generalized persistence algorithm for decomposing multiparameter persistence modules. *J. Appl. Comput. Topol.* **2022**, *6*, 271–322. [\[CrossRef\]](#)
27. Gluzberg, V.E.; Katz, Y.A. Topological data analysis of noise: Uniform unimodal distributions. *Commun. Nonlinear Sci. Numer. Simul.* **2023**, *121*, 107216. [\[CrossRef\]](#)
28. Lepaire, C.; Belhaouari, H.; Pascual, R.; Meseure, P. Exploitation of Local Adjacencies for Parallel Construction of a Reeb Graph Variant: Cerebral Vascular Tree Case. *J. WSCG* **2024**, *32*, 79–90. [\[CrossRef\]](#)
29. Rahman, S.E.; Athawale, T.M.; Rosen, P. GASP: Gradient-Aware Shortest Path Algorithm for Boundary-Confining 2-Manifold Reeb Graph Visualization. In *Proceedings of the IEEE Workshop on Topological Data Analysis and Visualization (TopoInVis)*; IEEE Press: Piscataway, NJ, USA, 2025. [\[CrossRef\]](#)
30. Kondapalli, L.S.; Azarudeen, S. Topological Data Analysis of Breast Cancer Using the Mapper Algorithm. In *Proceedings of the Fifth International Conference on Emerging Trends in Mathematical Sciences & Computing (IEMSC-24)*; Springer Nature: Berlin/Heidelberg, Germany, 2024; pp. 312–320.
31. Bui, Q.T.; Tien, T.N.; Nguyen, T.B.; Nguyen, S.P.; Vo, B. Understanding Conventional Deep Learning Models Through the Lens of Topological Data Analysis Using the Mapper Algorithm. In *Proceedings of the Sixth International Conference on Real-Time Intelligent Systems (RTIS 2024)*; Springer Nature: Berlin/Heidelberg, Germany, 2025; pp. 73–85.

32. Madukpe, V.N.; Ugoala, B.C.; Zulkepli, N.F.S. A Comprehensive Review of the Mapper Algorithm, a Topological Data Analysis Technique, and Its Applications Across Various Fields (2007–2025). *Int. J. Data Sci. Anal.* **2026**, *21*, 56.
33. Mojdehi, K.F.; Amiri, B.; Haddadi, A. A Novel Hybrid Model for Credit Risk Assessment of Supply Chain Finance Based on Topological Data Analysis and Graph Neural Network. *IEEE Access* **2025**, *13*, 13101–13127. [\[CrossRef\]](#)
34. Imoto, Y.; Hiraoka, Y. V-Mapper: Topological data analysis for high-dimensional data with velocity. *Nonlinear Theory Its Appl. IEICE* **2023**, *14*, 92–105. [\[CrossRef\]](#)
35. Bui, Q.T.; Vo, B.; Do, H.A.N.; Hung, N.Q.V.; Snášel, V. F-Mapper: A Fuzzy Mapper Clustering Algorithm. *Knowl.-Based Syst.* **2020**, *189*, 105107. [\[CrossRef\]](#)
36. Hacquard, O.; Lebovici, V. Euler Characteristic Tools for Topological Data Analysis. *J. Mach. Learn. Res.* **2024**, *25*, 1–39.
37. Smith, A.; Zavala, V.M. The Euler Characteristic: A General Topological Descriptor for Complex Data. *Comput. Chem. Eng.* **2021**, *154*, 107463. [\[CrossRef\]](#)
38. Okediji, B.A. Persistent Homology and Persistent Cohomology: A Review. *Earthline J. Math. Sci.* **2024**, *14*, 349–378. [\[CrossRef\]](#)
39. Cang, Z.; Wei, G.W. Persistent Cohomology for Data With Multicomponent Heterogeneous Information. *SIAM J. Math. Data Sci.* **2020**, *2*, 396–418. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Nigmatov, A.; Morozov, D. Distributed Computation of Persistent Cohomology. In *Proceedings of the SIAM Symposium on Algorithm Engineering and Experiments (ALENEX)*; SIAM: Philadelphia, PA, USA, 2026; pp. 194–206. [\[CrossRef\]](#)
41. Leygonie, J.; Carrière, M.; Lacombe, T.; Oudot, S.Y. A gradient sampling algorithm for stratified maps with applications to topological data analysis. *Math. Program.* **2023**, *202*, 199–239. [\[CrossRef\]](#)
42. Kang, T.; Kim, S.; Sohn, J.; Awan, J.A. Differentially Private Topological Data Analysis. *J. Mach. Learn. Res.* **2024**, *25*, 1–42.
43. Atienza, N.; Gonzalez-Diaz, R.; Soriano-Trigueros, M. On the stability of persistent entropy and new summary functions for topological data analysis. *Pattern Recognit.* **2020**, *107*, 107509. [\[CrossRef\]](#)
44. Choi, S.; Oh, J.; Park, J.R.; Yang, S.Y.; Yun, H. Effective data reduction algorithm for topological data analysis. *Appl. Math. Comput.* **2025**, *495*, 129302. [\[CrossRef\]](#)
45. Myers, A.D.; Chumley, M.M.; Khasawneh, F.A. Delay Parameter Selection in Permutation Entropy Using Topological Data Analysis. *Matematica* **2024**, *3*, 1103–1136. [\[CrossRef\]](#)
46. Martínez-Cadena, J.A. Topological data analysis for cluster-based dengue forecasting in Mexico. *Spat. Spatio-Temporal Epidemiol.* **2026**, *56*, 100784. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Thomas, A.M.; Crozier, P.A.; Xu, Y.; Matteson, D.S. Feature Detection and Hypothesis Testing for Extremely Noisy Nanoparticle Images using Topological Data Analysis. *Technometrics* **2023**, *65*, 590–603. [\[CrossRef\]](#)
48. Vishwanath, S.; Fukumizu, K.; Kuriki, S.; Sriperumbudur, B.K. On the limits of topological data analysis for statistical inference. *Found. Data Sci.* **2025**, *7*, 502–535. [\[CrossRef\]](#)
49. El-Yaagoubi, A.B.; Chung, M.K.; Ombao, H.C. Topological Data Analysis for Multivariate Time Series Data. *Entropy* **2023**, *25*, 1509. [\[CrossRef\]](#) [\[PubMed\]](#)
50. El-Yaagoubi, A.B.; Chung, M.K.; Ombao, H.C. Statistical inference for dependence networks in topological data analysis. *Front. Artif. Intell.* **2023**, *6*, 1293504. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Berry, D.W.; Su, Y.; Gyurik, C.; King, R.; Basso, J.; Barba, A.D.T.; Rajput, A.; Wiebe, N.; Dunjko, V.; Babbush, R. Analyzing Prospects for Quantum Advantage in Topological Data Analysis. *PRX Quantum* **2024**, *5*, 010319. [\[CrossRef\]](#)
52. Schmidhuber, A.; Lloyd, S. Complexity-Theoretic Limitations on Quantum Algorithms for Topological Data Analysis. *PRX Quantum* **2023**, *4*, 040349. [\[CrossRef\]](#)
53. Hernández Serrano, D.; Sánchez Gómez, D. Centrality measures in simplicial complexes: Applications of topological data analysis to network science. *Appl. Math. Comput.* **2020**, *382*, 125331. [\[CrossRef\]](#)
54. Hernández Serrano, D.; Hernández-Serrano, J.; Sánchez Gómez, D. Simplicial degree in complex networks. applications of topological data analysis to network science. *Chaos Solitons Fractals* **2020**, *137*, 109839. [\[CrossRef\]](#)
55. Guo, H.; Xia, S.; An, Q.; Zhang, X.; Sun, W.; Zhao, X. Empirical study of financial crises based on topological data analysis. *Phys. A Stat. Mech. Its Appl.* **2020**, *558*, 124956. [\[CrossRef\]](#)
56. Katz, Y.A.; Biem, A.E. Time-resolved topological data analysis of market instabilities. *Phys. A Stat. Mech. Its Appl.* **2021**, *571*, 125816. [\[CrossRef\]](#)
57. Wang, B.; Feng, B.; Lv, L.; Li, S.; Pan, F. Structural Feature Extraction via Topological Data Analysis. *J. Phys. Chem. Lett.* **2025**, *16*, 8056–8067. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Zieliński, B.; Lipinski, M.; Juda, M.; Zeppelzauer, M.; Dłotko, P. Persistence codebooks for topological data analysis. *Artif. Intell. Rev.* **2021**, *54*, 1969–2009. [\[CrossRef\]](#)
59. Chevyrev, I.; Nanda, V.; Oberhauser, H. Persistence Paths and Signature Features in Topological Data Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 192–202. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Ignacio, P.S.; Bulauan, J.A.; Uminsky, D.T. LUMÁWIG: An efficient algorithm for dimension zero bottleneck distance computation in topological data analysis. *Algorithms* **2020**, *13*, 291. [\[CrossRef\]](#)

61. Ali, D.; Asaad, A.T.; Jimenez, M.J.; Nanda, V.; Paluzo-Hidalgo, E.; Soriano-Trigueros, M. A Survey of Vectorization Methods in Topological Data Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 14069–14080. [[CrossRef](#)] [[PubMed](#)]
62. Dagliati, A.; Geifman, D.N.; Peek, N.B.; Holmes, J.H.; Sacchi, L.L.; Bellazzi, R.; Sajjadi, S.E.; Tucker, A.J. Using topological data analysis and pseudo time series to infer temporal phenotypes from electronic health records. *Artif. Intell. Med.* **2020**, *108*, 101930. [[CrossRef](#)] [[PubMed](#)]
63. Lin, B. Topological Data Analysis in Time Series: Temporal Filtration and Application to Single-Cell Genomics. *Algorithms* **2022**, *15*, 371. [[CrossRef](#)]
64. Riihimäki, H.; Chachólski, W.; Theorell, J.; Hillert, J.A.; Ramanujam, R. A topological data analysis based classification method for multiple measurements. *BMC Bioinform.* **2020**, *21*, 322. [[CrossRef](#)] [[PubMed](#)]
65. Schweidtmann, A.M.; Weber, J.M.; Wende, C.; Netze, L.; Mitsos, A. Obey validity limits of data-driven models through topological data analysis and one-class classification. *Optim. Eng.* **2022**, *23*, 855–876. [[CrossRef](#)]
66. Frahi, T.; Sancarlos, A.; Galle, M.; Beaulieu, X.; Chambard, A.; Falcó Montesinos, A.; Cueto, E.; Chinesta, F.S. Monitoring Weeder Robots and Anticipating Their Functioning by Using Advanced Topological Data Analysis. *Front. Artif. Intell.* **2021**, *4*, 761123. [[CrossRef](#)] [[PubMed](#)]
67. Uray, M.; Giunti, B.; Kerber, M.; Huber, S. Topological Data Analysis in Smart Manufacturing: State of the Art and Future Directions. *J. Manuf. Syst.* **2024**, *76*, 75–91. [[CrossRef](#)]
68. Bussola, N.; Papa, B.; Melaiu, O.; Castellano, A.; Fruci, D.; Jurman, G. Quantification of the immune content in neuroblastoma: Deep learning and topological data analysis in digital pathology. *Int. J. Mol. Sci.* **2021**, *22*, 8804. [[CrossRef](#)] [[PubMed](#)]
69. Klaila, G.; Vutov, V.; Stefanou, A. Supervised topological data analysis for MALDI mass spectrometry imaging applications. *BMC Bioinform.* **2023**, *24*, 273. [[CrossRef](#)] [[PubMed](#)]
70. Monkam, G.F.; de Lucia, M.J.; Bastian, N.D. A topological data analysis approach for detecting data poisoning attacks against machine learning based network intrusion detection systems. *Comput. Secur.* **2024**, *144*, 103929. [[CrossRef](#)]
71. Ma, T.; Su, Y.; Abdelwahab, M.M.; Khalil, A.A.E. ZPDSN: Spatio-temporal meteorological forecasting with topological data analysis. *Appl. Intell.* **2025**, *55*, 9. [[CrossRef](#)]
72. Wang, H.; Li, T.; Dong, Z. Research on the Construction of a Short-Term Voltage Prediction Model Integrating Topological Data Analysis and Deep Neural Network under the Power System Resilience Assessment Framework. *EAI Endorsed Trans. Energy Web* **2024**, *12*, e8896. [[CrossRef](#)]
73. Han, Y.; Li, K.; Cai, H.; Luan, W.; Zhao, B.; Liu, B. An adaptive appliance identification method leveraging topological data analysis. *Eng. Appl. Artif. Intell.* **2025**, *156*, 111024. [[CrossRef](#)]
74. Shylaja, R.; Nedumaran, D.; Venkateswaran, C. The synergy of topological data analysis and machine learning for low-cost e-nose systems. *Microsyst. Technol.* **2025**, *31*, 2353–2369. [[CrossRef](#)]
75. Sagming, M.N.; Heymann, R.; Visaya, M.V.V. Using topological data analysis and machine learning to predict customer churn. *J. Big Data* **2024**, *11*, 160. [[CrossRef](#)]
76. Rahman, A.; Satti, A.; Shahid, A.R.; Shafi, Q.M.; Farooq, K.; Ali Safi, A. TDA SegUNet: Topological Data Analysis-Based Shape-Aware Brain Tumor Segmentation. *IEEE Access* **2025**, *13*, 36190–36200. [[CrossRef](#)]
77. Jiang, P.; Lugo-Martinez, J. Combined Topological Data Analysis and Geometric Deep Learning Reveal Niches by the Quantification of Protein Binding Pockets. *J. Comput. Biol.* **2025**, *32*, 659–674. [[CrossRef](#)] [[PubMed](#)]
78. Gavris, G.B.; Sun, W. Topology Optimization with Graph Neural Network Enabled Regularized Thresholding. *Extrem. Mech. Lett.* **2024**, *71*, 102215. [[CrossRef](#)]
79. Bachiri, K.; Ali, Y.; Malek, M.; Rogovschi, N. Topological Data Analysis and Graph-Based Learning for Multimodal Recommendation. *IEEE Access* **2025**, *13*, 108934–108954. [[CrossRef](#)]
80. Akingbade, S.W.; Gidea, M.; Manzi, M.; Nateghi, V. Why topological data analysis detects financial bubbles? *Commun. Nonlinear Sci. Numer. Simul.* **2024**, *128*, 107665. [[CrossRef](#)]
81. Guo, H.; Ming, Z.; Xing, B. Topological data analysis of Chinese stocks' dynamic correlations under major public events. *Front. Phys.* **2023**, *11*, 1253953. [[CrossRef](#)]
82. Shiraj, M.M.B.; Rahman, M.M.; Al-Imran, M.; Liza, M.Z.A.; Murshed, M.M.; Akhter, N. Anomaly Detection in Financial Time Series Data via Mapper Algorithm and DBSCAN Clustering. *World J. Adv. Eng. Technol. Sci.* **2024**, *13*, 70–84. [[CrossRef](#)]
83. Majumdar, S.; Laha, A.K. Clustering and classification of time series using topological data analysis with applications to finance. *Expert Syst. Appl.* **2020**, *162*, 113868. [[CrossRef](#)]
84. Karan, A.; Kaygun, A. Time series classification via topological data analysis. *Expert Syst. Appl.* **2021**, *183*, 115326. [[CrossRef](#)]
85. Goel, A.; Pasricha, P.; Mehra, A. Topological data analysis in investment decisions. *Expert Syst. Appl.* **2020**, *147*, 113222. [[CrossRef](#)]
86. Yen, T.; Cheong, S. Using Topological Data Analysis (TDA) and Persistent Homology to Analyze the Stock Markets in Singapore and Taiwan. *Front. Phys.* **2021**, *9*, 572216. [[CrossRef](#)]
87. Adami, V.; Masoomy, H.; Najafi, M.N. Topological data analysis of the visibility graphs of the BTW avalanche-size sequences. *J. Phys. A Math. Theor.* **2025**, *58*, 205002. [[CrossRef](#)]

88. Güzel, İ.; Kaygun, A. Classification of stochastic processes with topological data analysis. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e7732. [\[CrossRef\]](#)
89. Singh, M.K.; Chaube, S.; Pant, S.; Singh, S.K.; Kumar, A. An integrated image visibility graph and topological data analysis for extracting time series features. *Decis. Anal. J.* **2023**, *8*, 100253. [\[CrossRef\]](#)
90. Cleveland, E.; Zhu, A.; Sandstede, B.; Volkening, A. Quantifying Different Modeling Frameworks Using Topological Data Analysis: A Case Study with Zebrafish Patterns. *SIAM J. Appl. Dyn. Syst.* **2023**, *22*, 3233–3266. [\[CrossRef\]](#)
91. Ferrara, M. Modeling by topological data analysis and game theory for analyzing data poisoning phenomena. *AIMS Math.* **2025**, *10*, 15457–15475. [\[CrossRef\]](#)
92. Razmarashooli, A.; Chua, Y.; Barzegar, V.; Salazar, D.; Laflamme, S.; Hu, C.; Downey, A.R.; Dodson, J.C.; Schrader, P.T. Real-time state estimation of nonstationary systems through dominant fundamental frequency using topological data analysis features. *Mech. Syst. Signal Process.* **2025**, *224*, 112048. [\[CrossRef\]](#)
93. Zheng, X.; Mak, S.; Xie, L.; Xie, Y. PERCEPT: A New Online Change-Point Detection Method using Topological Data Analysis. *Technometrics* **2023**, *65*, 162–178. [\[CrossRef\]](#)
94. Esteve, M.; Falcó Montesinos, A. Trajectory Classification Through Topological Data Analysis Perspectives. *IEEE Access* **2025**, *13*, 32458–32469. [\[CrossRef\]](#)
95. Esteve, M.; Falcó Montesinos, A. tramoTDA: A trajectory monitoring system using Topological Data Analysis. *SoftwareX* **2024**, *28*, 101953. [\[CrossRef\]](#)
96. Narender, M.; Mohsin, K.S.; Ragunthar, T.; Papasani, A.; Ayasrah, F.T.; Naik, A.R. Machine Learning for Genomic Expression Classification-Based Phenotype Prediction in Topological Data Analysis. *J. Mach. Comput.* **2024**, *4*, 1152–1157. [\[CrossRef\]](#)
97. Tarín-Pelló, A.; Suay-García, B.; Forés-Martos, J.; Falcó Montesinos, A.; Perez-Gracia, M.T. Computer-aided drug repurposing to tackle antibiotic resistance based on topological data analysis. *Comput. Biol. Med.* **2023**, *166*, 107496. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Blanco-Rodríguez, R.; Ordoñez-Jiménez, F.; Almocera, A.E.S.; Chinney-Herrera, G.; Hernandez-Vargas, E.A. Topological data analysis of antibody dynamics of severe and non-severe patients with COVID-19. *Math. Biosci.* **2023**, *361*, 109011. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Rostami, Z.; Fooshee, D.; Carlsson, G.; Subramaniam, S. Topological Data Analysis Reveals a Subgroup of Luminal B Breast Cancer. *IEEE Open J. Eng. Med. Biol.* **2025**, *6*, 465–471. [\[CrossRef\]](#) [\[PubMed\]](#)
100. Thomas, A.M.; Lin, A.C.; Deng, G.; Xu, Y.; Ranvier, G.F.; Taye, A.; Matteson, D.S.; Lee, D. A Proof-of-Concept Investigation into Predicting Follicular Carcinoma on Ultrasound Using Topological Data Analysis and Radiomics. *Imaging* **2025**, *17*, 39–48. [\[CrossRef\]](#)
101. Chulián, S.; Stolz, B.J.; Martínez-Rubio, Á.; Blázquez-Goñi, C.B.; Rodríguez-Gutiérrez, J.F.; Velázquez, T.C.; Molinos-Quintana, Á.; Ramirez-Orellana, M.; Castillo-Robleda, A.; Soler, J.L.F.; et al. The shape of cancer relapse: Topological data analysis predicts recurrence in paediatric acute lymphoblastic leukaemia. *PLoS Comput. Biol.* **2023**, *19*, e1011329. [\[CrossRef\]](#) [\[PubMed\]](#)
102. de Perez, A.R.; Anderson, P.E.; Dimitrova, E.S.; Kemp, M.L. Neural network approaches, including use of topological data analysis, enhance classification of human induced pluripotent stem cell colonies by treatment condition. *PLoS Comput. Biol.* **2025**, *21*, e1012801. [\[CrossRef\]](#) [\[PubMed\]](#)
103. Migdałek, G.; Zelawski, M. Measuring population-level plant gene flow with topological data analysis. *Ecol. Inform.* **2022**, *70*, 101740. [\[CrossRef\]](#)
104. Nicponski, J.; Jung, J.-H. Topological Data Analysis of Vascular Disease: A Theoretical Framework. *Front. Appl. Math. Stat.* **2020**, *6*, 34. [\[CrossRef\]](#)
105. Hernández-Lemus, E.; Miramontes, P.; Martínez-García, M. Topological Data Analysis in Cardiovascular Signals: An Overview. *Entropy* **2024**, *26*, 67. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Me, Y.S.; Hathaway, Q.A.; Farrelly, C.; Budoff, M.J.; Erickson, B.; Collins, J.D.; Blaha, M.J.; Leiner, T.; Lopez-Jimenez, F.; Rozenblit, J.; et al. Topological Data Analysis in the Assessment of Coronary Atherosclerosis: A Comprehensive Narrative Review. *Mayo Clin. Proc. Digit. Health* **2025**, *3*, 100199. [\[CrossRef\]](#)
107. De Benedictis, S.G.; Gargano, G.; Settembre, G. Enhanced MRI brain tumor detection and classification via topological data analysis and low-rank tensor decomposition. *J. Comput. Math. Data Sci.* **2024**, *13*, 100103. [\[CrossRef\]](#)
108. Maurya, A.; Stanley, R.J.; Lama, N.; Nambisan, A.K.; Patel, G.; Saeed, D.; Swinfard, S.; Smith, C.; Jagannathan, S.; Hagerty, J.R.; et al. Hybrid Topological Data Analysis and Deep Learning for Basal Cell Carcinoma Diagnosis. *J. Imaging Inform. Med.* **2024**, *37*, 92–106. [\[CrossRef\]](#) [\[PubMed\]](#)
109. Percival, S.; Onyenedum, J.G.; Chitwood, D.H.; Husbands, A.Y. Topological data analysis reveals core heteroblastic and ontogenetic programs embedded in leaves of grapevine (Vitaceae) and maracuyá (Passifloraceae). *PLoS Comput. Biol.* **2024**, *20*, e1011845. [\[CrossRef\]](#) [\[PubMed\]](#)
110. Ereemeev, S.V. Detection of Repeated Structures in an Image Based on Topological Data Analysis. *Pattern Recognit. Image Anal.* **2024**, *34*, 936–939. [\[CrossRef\]](#)

111. Cheng, P.; Cao, X.; Yang, Y.; Zhang, G.; He, Y. Automatically recognize and segment morphological features of the 3D vertebra based on topological data analysis. *Comput. Biol. Med.* **2022**, *149*, 106031. [[CrossRef](#)] [[PubMed](#)]
112. Jain, T.K.; Joshi, A.; Keswani, A.; Kumar, A.; Dadheech, P. Mathematics driven methods with fractional calculus and topological data analysis in image segmentation. *J. Interdiscip. Math.* **2025**, *28*, 637–646. [[CrossRef](#)]
113. Wei, Q.; Zeng, S.; Jiang, Z.; Xu, F.; Shi, S.; Wen, W.; Qin, Z.; Lou, Z.; Li, K. Topological Representation Based on Wavelet Transform as a Novel Imaging Biomarker for Tumor Diagnosis in Ultrasound Images: A Comprehensive Study. *Comput. Methods Programs Biomed.* **2025**, *269*, 108859. [[CrossRef](#)] [[PubMed](#)]
114. Singh, Y.; Quaia, E. Unraveling the Invisible: Topological Data Analysis as the New Frontier in Radiology's Diagnostic Arsenal. *Tomography* **2025**, *11*, 6. [[CrossRef](#)] [[PubMed](#)]
115. Ling, C.Y.F.; Piau, P.; Liew, S.H. Topological data analysis in EEG signal processing: A review. *Commun. Math. Biol. Neurosci.* **2025**, *2025*, 9511. [[CrossRef](#)]
116. Zheng, J.; Feng, Z.; Ekstrom, A.D. Towards Analysis of Multivariate Time Series Using Topological Data Analysis. *Mathematics* **2024**, *12*, 1727. [[CrossRef](#)]
117. Wang, Z.M.; Li, S.; Zhang, J.; Liang, C. Emotion recognition based on phase-locking value brain functional network and topological data analysis. *Neural Comput. Appl.* **2024**, *36*, 7903–7922. [[CrossRef](#)]
118. Liu, C.; Ma, X.; Zhang, H.; Xie, S.; Yu, D. Dynamic Neuropsychological Approach for Multi-Quality Image Assessment Using Grey-Topological Data Analysis. *IEEE Access* **2024**, *12*, 139609–139619. [[CrossRef](#)]
119. Ling, T.; Zhu, Z.; Zhang, Y.; Jiang, F. Early Ventricular Fibrillation Prediction Based on Topological Data Analysis of ECG Signal. *Appl. Sci.* **2022**, *12*, 10370. [[CrossRef](#)]
120. Catanzaro, M.J.; Rizzo, S.; Kopchick, J.; Chowdury, A.Z.; Rosenberg, D.R.; Bubenik, P.; Diwadkar, V.A. Topological Data Analysis Captures Task-Driven fMRI Profiles in Individual Participants: A Classification Pipeline Based on Persistence. *Neuroinformatics* **2024**, *22*, 45–62. [[CrossRef](#)] [[PubMed](#)]
121. Chung, M.K.; Azizi, T.; Hanson, J.L.; Alexander, A.L.; Pollak, S.D.; Davidson, R.J. Altered topological structure of the brain white matter in maltreated children through topological data analysis. *Netw. Neurosci.* **2024**, *8*, 355–376. [[CrossRef](#)] [[PubMed](#)]
122. Arfi, B. The promises of persistent homology, machine learning, and deep neural networks in topological data analysis of democracy survival. *Qual. Quant.* **2024**, *58*, 1685–1727. [[CrossRef](#)]
123. Cai, T.; Zhao, G.; Zang, J.; Zong, C.; Zhang, Z.; Xue, C. Topological Feature Search Method for Multichannel EEG: Application in ADHD Classification. *Biomed. Signal Process. Control* **2025**, *100*, 107153. [[CrossRef](#)]
124. Sathyanarayana, A.; Manjunath, S.; Perea, J.A. Topological Data Analysis Based Characteristics of Electroencephalogram Signals in Children With Sleep Apnea. *J. Sleep Res.* **2025**, *34*, e70017. [[CrossRef](#)] [[PubMed](#)]
125. Goodarzi, N.; Yaghmae, A.; Bahrami, F. Topological Data Analysis of Resting-State EEG for On/Off Medication Classification of Parkinson's Disease Patients. In *Proceedings of the 2024 31st National and 9th International Iranian Conference on Biomedical Engineering (ICBME)*; IEEE Press: Piscataway, NJ, USA, 2024. [[CrossRef](#)]
126. Poetto, S.; Duch, W. Classification of Schizophrenia EEG Recording Using Homological Features. In *Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN)*; IEEE Press: Piscataway, NJ, USA, 2024. [[CrossRef](#)]
127. Altindis, F.; Yilmaz, B.; Borisenok, S.; İçöz, K. Parameter investigation of topological data analysis for EEG signals. *Biomed. Signal Process. Control* **2021**, *63*, 102196. [[CrossRef](#)]
128. Mjahad, A.; Frances-Villora, J.V.; Bataller-Mompeán, M.; Rosado-Muñoz, A. Ventricular Fibrillation and Tachycardia Detection Using Features Derived from Topological Data Analysis. *Appl. Sci.* **2022**, *12*, 7248. [[CrossRef](#)]
129. Elhamdadi, H.; Canavan, S.J.; Rosen, P.A. AffectiveTDA: Using Topological Data Analysis to Improve Analysis and Explainability in Affective Computing. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 769–779. [[CrossRef](#)] [[PubMed](#)]
130. Koseki, J.; Hayashi, S.; Kojima, Y.; Hirose, H.; Shimamura, T. Topological data analysis of protein structure and inter/intra-molecular interaction changes attributable to amino acid mutations. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 2950–2959. [[CrossRef](#)] [[PubMed](#)]
131. Koseki, J.; Motono, C.; Yanagisawa, K.; Kudo, G.; Yoshino, R.; Hirokawa, T.; Imai, K. CrypToth: Cryptic Pocket Detection through Mixed-Solvent Molecular Dynamics Simulations-Based Topological Data Analysis. *J. Chem. Inf. Model.* **2025**, *65*, 5567–5575. [[CrossRef](#)] [[PubMed](#)]
132. Karthick, V.; Paulraj Jayasimman, I.; Dhilshath, S.; Arasu, R.; Chinnadurai, K.; Suganthi, J. Quantum graph-based differential models with fractional calculus and topological data analysis for dynamic characterization of protein-protein interaction networks. *Asia Pac. J. Math.* **2025**, *12*, 70. [[CrossRef](#)]
133. Nguyen, K.C.; Jameson, C.D.; Baldwin, S.A.; Nardini, J.T.; Smith, R.C.; Haugh, J.M.; Flores, K.B. Quantifying collective motion patterns in mesenchymal cell populations using topological data analysis and agent-based modeling. *Math. Biosci.* **2024**, *370*, 109158. [[CrossRef](#)] [[PubMed](#)]
134. Dawson, M.; Dudley, C.; Omoma, S.; Tung, H.R.; Ciocanel, M.V. Characterizing emerging features in cell dynamics using topological data analysis methods. *Math. Biosci. Eng.* **2023**, *20*, 3023–3046. [[CrossRef](#)] [[PubMed](#)]

135. Nardini, J.T.; Stolz, B.J.; Flores, K.B.; Harrington, H.A.; Byrne, H.M. Topological data analysis distinguishes parameter regimes in the Anderson-Chaplain model of angiogenesis. *PLoS Comput. Biol.* **2021**, *17*, e1009094. [[CrossRef](#)] [[PubMed](#)]
136. Bonilla, L.L.; Carpio, A.; Trenado-Yuste, C. Tracking collective cell motion by topological data analysis. *PLoS Comput. Biol.* **2020**, *16*, e1008407. [[CrossRef](#)] [[PubMed](#)]
137. Ciocanel, M.V.; Juenemann, R.; Dawes, A.T.; McKinley, S.A. Topological Data Analysis Approaches to Uncovering the Timing of Ring Structure Onset in Filamentous Networks. *Bull. Math. Biol.* **2021**, *83*, 17. [[CrossRef](#)] [[PubMed](#)]
138. Thomas, A.; Bates, K.E.; Elchesen, A.; Hartsock, I.; Lu, H.; Bubenik, P. Topological Data Analysis of *C. elegans* Locomotion and Behavior. *Front. Artif. Intell.* **2021**, *4*, 668395. [[CrossRef](#)] [[PubMed](#)]
139. Mototake, Y.I.; Mizumaki, M.; Kudo, K.; Fukumizu, K. Procedure to reveal the mechanism of pattern formation process by topological data analysis. *Phys. D Nonlinear Phenom.* **2024**, *470*, 134359. [[CrossRef](#)]
140. González, A.; Hernández, B.; Acosta-Zamora, K.P.; Ramos, E.; Núñez, J. Topological data analysis of three dimensional orbits in a convective flow. *Phys. D Nonlinear Phenom.* **2025**, *481*, 134841. [[CrossRef](#)]
141. Miller, M.J.; Johnston, N.; Livengood, I.; Spinelli, M.; Sazdanović, R.; Olufsen, M.S. A topological data analysis study on murine pulmonary arterial trees with pulmonary hypertension. *Math. Biosci.* **2023**, *364*, 109056. [[CrossRef](#)] [[PubMed](#)]
142. Jeung, S.y.; Kwon, J. A Robust Multivariate Time Series Classification Approach Based on Topological Data Analysis for Channel Fault Tolerance. *Sensors* **2025**, *25*, 2709. [[CrossRef](#)] [[PubMed](#)]
143. Wang, B.; Lin, C.; Inoue, H.; Kanemaru, M. Induction Motor Eccentricity Fault Detection and Quantification Using Topological Data Analysis. *IEEE Access* **2024**, *12*, 37891–37902. [[CrossRef](#)]
144. Chukanov, S.N.; Chukanov, I.S. Using Topological Data Analysis to Visualize Instrument Output. *Sci. Vis.* **2023**, *15*, 11–21. [[CrossRef](#)]
145. Tempelman, J.R.; Khasawneh, F.A. A look into chaos detection through topological data analysis. *Phys. D Nonlinear Phenom.* **2020**, *406*, 132446. [[CrossRef](#)]
146. Siegrist, K.W.; Kramer, R.M.; Chagdes, J.R. Investigating the nonlinear dynamics of human balance using topological data analysis. *J. Comput. Nonlinear Dyn.* **2020**, *15*, 101002. [[CrossRef](#)]
147. Frahi, T.; Chinesta, F.S.; Falcó Montesinos, A.; Badías-Herbera, A.; Cueto, E.; Choi, H.; Han, M.; Duval, J.L. Empowering advanced driver-assistance systems from topological data analysis. *Mathematics* **2021**, *9*, 634. [[CrossRef](#)]
148. Frahi, T.; Falcó Montesinos, A.; Mau, B.V.; Duval, J.L.; Chinesta, F.S. Empowering advanced parametric modes clustering from topological data analysis. *Appl. Sci.* **2021**, *11*, 6554. [[CrossRef](#)]
149. Woldemariam, P.; Attoh-Okine, N.O. Topological Data Analysis for Railway Track Geometry Safety and Maintenance. *J. Comput. Civ. Eng.* **2025**, *39*, CPENG-6168. [[CrossRef](#)]
150. Larson, D.M.; Bungula, W.T.; McKean, C.; Stockdill, A.; Lee, A.; Miller, F.F.; Davis, K.T. Quantifying ecosystem states and state transitions of the Upper Mississippi River System using topological data analysis. *PLoS Comput. Biol.* **2023**, *19*, e1011147. [[CrossRef](#)] [[PubMed](#)]
151. Pereira, L.M.; Torres, L.C.; Amini, M.H. Topological Data Analysis for Network Resilience Quantification. *Oper. Res. Forum* **2021**, *2*, 26. [[CrossRef](#)]
152. Torshin, I.Y.; Rudakov, K.V. Topological Data Analysis in Materials Science: The Case of High-Temperature Cuprate Superconductors. *Pattern Recognit. Image Anal.* **2020**, *30*, 264–276. [[CrossRef](#)]
153. Karthik, J.; Jose, N.; Kumar, M.S.S.; Somashekar, R.K. Topological data analysis and comparison of physical parameters of different forms of graphene. *Int. J. Comput. Mater. Sci. Surf. Eng.* **2022**, *11*, 73–83. [[CrossRef](#)]
154. Zhao, Y.; Wang, Y.; Ding, Y.; Han, H. Topological data analysis for the energy and stability of endohedral metallofullerenes. *J. Math. Chem.* **2022**, *60*, 337–352. [[CrossRef](#)]
155. Vasudevan, A.; Prieto, J.Z.; Zorkaltsev, S.; Haranczyk, M. tda-segmentor: A Tool to Extract and Analyze Local Structure and Porosity Features in Porous Materials. *Comput. Phys. Commun.* **2024**, *305*, 109344. [[CrossRef](#)]
156. Broderick, S.; Dongol, R.; Zhang, T.; Rajan, K. Classification of Apatite Structures via Topological Data Analysis: A Framework for a Materials Barcode Representation of Structure Maps. *Sci. Rep.* **2021**, *11*, 11599. [[CrossRef](#)] [[PubMed](#)]
157. Kheneifar, M.A.; Amiri, B. A Novel Hybrid Model for Loan Default Prediction in Maritime Finance Based on Topological Data Analysis and Machine Learning. *IEEE Access* **2025**, *13*, 81474–81493. [[CrossRef](#)]
158. Lakshminarayan, C.K.; Yin, M. Topological data analysis in digital marketing. *Appl. Stoch. Model. Bus. Ind.* **2020**, *36*, 1014–1028. [[CrossRef](#)]
159. Qiu, W.; Rudkin, S.T.; Dłotko, P. Refining understanding of corporate failure through a topological data analysis mapping of Altman's Z-score model. *Expert Syst. Appl.* **2020**, *156*, 113475. [[CrossRef](#)]
160. Mancilla, S.; Wences, G.; Hernandez-Lopez, E.; Cohen, I. Sub-spatial prediction of votes integrating socioeconomic, educational, and age strata with machine learning and topological data analysis. *J. Big Data* **2025**, *12*, 11112. [[CrossRef](#)]
161. Dey, A.K.; Kundu, S. Complex Network and Topological Data Analysis Methods for County Level COVID-19 Vaccine Acceptance Analysis in the United States. *Stat. Med.* **2025**, *44*, e70109. [[CrossRef](#)] [[PubMed](#)]

162. Lu, Z.; Liu, H. A Topological Data Analysis Approach to the COVID-19. In *Proceedings of the 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*; IEEE Press: Piscataway, NJ, USA, 2022. [\[CrossRef\]](#)
163. Vittorietti, M.; Giambalvo, O.; Genova, V.G.; Aiello, F. A new measure for the attitude to mobility of Italian students and graduates: A topological data analysis approach. *Stat. Methods Appl.* **2023**, *32*, 509–543. [\[CrossRef\]](#)
164. Riaz, M.; Qamar, F.; Tariq, S.; Alsager, K.M. AI-Driven LOPCOW-AROMAN Framework and Topological Data Analysis Using Circular Intuitionistic Fuzzy Information: Healthcare Supply Chain Innovation. *Mathematics* **2024**, *12*, 3593. [\[CrossRef\]](#)
165. Bouazzaoui, H.; Mamouni, M.I.; Elomary, M.A. Bongard problems: A topological data analysis approach. *WSEAS Trans. Syst. Control* **2020**, *15*, 131–140. [\[CrossRef\]](#)
166. Kim, S.; Jeong, J. Organic relationship between laws based on judicial precedents using topological data analysis. *Korean J. Math.* **2021**, *29*, 649–664. [\[CrossRef\]](#)
167. Riaz, M.; Kausar, R.; Jameel, T.; Pamucar, D.S.S. Cubic picture fuzzy topological data analysis with integrating blockchain and the metaverse for uncertain supply chain management. *Eng. Appl. Artif. Intell.* **2024**, *131*, 107827. [\[CrossRef\]](#)
168. Tudoreanu, M.E. Exploring the use of topological data analysis to automatically detect data quality faults. *Front. Big Data* **2022**, *5*, 931398. [\[CrossRef\]](#) [\[PubMed\]](#)
169. Narita, M. An empirical study on darknet visualization based on topological data analysis. *Int. J. Networked Distrib. Comput.* **2021**, *9*, 52–58. [\[CrossRef\]](#)
170. Deng, R.; Duzhin, F.S. Topological Data Analysis Helps to Improve Accuracy of Deep Learning Models for Fake News Detection Trained on Very Small Training Sets. *Big Data Cogn. Comput.* **2022**, *6*, 74. [\[CrossRef\]](#)
171. Raqibul Hasan, M.; Hossain, M.J.; Waliullah, M.; Hannan, A.; Rahman, M.M. Topological Data Analysis and Wavelet-Unsupervised Machine Learning Approaches to Identifying the Flooding and Non-Flooding Zones. *IEEE Access* **2025**, *13*, 111710–111721. [\[CrossRef\]](#)
172. Tauzin, G.; Lupo, U.; Tunstall, L.; Perez, J.B.; Caorsi, M.; Medina-Mardones, A.M.; Dassatti, A.; Hess, K.P. giotto-tda: A topological data analysis toolkit for machine learning and data exploration. *J. Mach. Learn. Res.* **2021**, *22*, 1–6.
173. Liu, S.; Gaffney, J.A.; Peterson, J.L.; Robinson, P.B.; Bhatia, H.; Pascucci, V.; Spears, B.K.; Bremer, P.T.; Wang, D.; Maljovec, D.; et al. Scalable Topological Data Analysis and Visualization for Evaluating Data-Driven Models in Scientific Applications. *IEEE Trans. Vis. Comput. Graph.* **2020**, *26*, 291–300. [\[CrossRef\]](#) [\[PubMed\]](#)
174. Kindelan, R.; Frías, J.; Cerda, M.; Hitschfeld, N. A topological data analysis based classifier. *Adv. Data Anal. Classif.* **2024**, *18*, 493–538. [\[CrossRef\]](#)
175. Ko, S.; Koo, D. A novel approach for wafer defect pattern classification based on topological data analysis. *Expert Syst. Appl.* **2023**, *231*, 120765. [\[CrossRef\]](#)
176. Zhang, X.; Gao, Y.; Zhang, Y.; Li, F.; Li, H.; Lei, F. Identification of Autism Spectrum Disorder Using Topological Data Analysis. *J. Imaging Inform. Med.* **2024**, *37*, 1023–1037. [\[CrossRef\]](#) [\[PubMed\]](#)
177. Chung, Y.; Huang, W.K.; Wu, H.T. Topological data analysis assisted automated sleep stage scoring using airflow signals. *Biomed. Signal Process. Control* **2024**, *89*, 105760. [\[CrossRef\]](#)
178. Lee, D.; Bresten, C.L.; Youm, K.; Seo, K.; Jung, J.H. Model discrepancy of Earth polar motion using topological data analysis and convolutional neural network analysis. *Int. J. Mod. Phys. C* **2020**, *31*, 2050117. [\[CrossRef\]](#)
179. Vuillien, M.; Adamo, D.; Vila, E.; Agraw, A.; Argant, T.; Helmer, D.; Mashkour, M.; Moussous, A.; Notter, O.; Rossoni-Notter, E.; et al. Topological Data Analysis and Multiple Kernel Learning for Species Identification of Modern and Archaeological Small Ruminants. *J. Comput. Appl. Archaeol.* **2025**, *8*, 170–187. [\[CrossRef\]](#)
180. Sekuloski, P.; Kitanovski, D.; Goshev, I.; Mishev, K.; Simjanoska Misheva, M.; Dimitrievska Ristovska, V. Exploring the Potential of Topological Data Analysis for Explainable Large Language Models: A Scoping Review. *Mathematics* **2026**, *14*, 378. [\[CrossRef\]](#)
181. Tran, M.L.; Park, C.; Jung, J.H. Topological data analysis of Korean music in Jeongganbo: A cycle structure. *J. Math. Music* **2023**, *17*, 403–432. [\[CrossRef\]](#)
182. Liu, J. Application of topological data analysis method based on calligraphy creation schema and its limitation analysis. *Appl. Math. Nonlinear Sci.* **2024**, *9*, 20230736. [\[CrossRef\]](#)
183. Kalisnik, S.; Lehn, C.; Limic, V. Geometric and probabilistic limit theorems in topological data analysis. *Adv. Appl. Math.* **2021**, *131*, 102244. [\[CrossRef\]](#)
184. Sasaki, K.; Bruder, D.; Hernandez-Vargas, E.A. Topological data analysis to model the shape of immune responses during co-infections. *Commun. Nonlinear Sci. Numer. Simul.* **2020**, *85*, 105228. [\[CrossRef\]](#) [\[PubMed\]](#)
185. Aggarwal, M.; Periwal, V. Dory: Computation of Persistence Diagrams up to Dimension Two for Vietoris–Rips Filtrations of Large Data Sets. *J. Comput. Sci.* **2024**, *79*, 102290. [\[CrossRef\]](#) [\[PubMed\]](#)
186. Stolz, B.J.; Emerson, T.H.; Nahkuri, S.; Porter, M.A.; Harrington, H.A. Topological data analysis of task-based fMRI data from experiments on schizophrenia. *J. Phys. Complex.* **2021**, *2*, 035006. [\[CrossRef\]](#)
187. Corcoran, P.; Jones, C.B. Topological data analysis for geographical information science using persistent homology. *Int. J. Geogr. Inf. Sci.* **2023**, *37*, 712–745. [\[CrossRef\]](#)

188. Smith, A.D.; Dłotko, P.; Zavala, V.M. Topological data analysis: Concepts, computation, and applications in chemical engineering. *Comput. Chem. Eng.* **2021**, *146*, 107202. [[CrossRef](#)]
189. Riaz, M.; Batool, S.; Almalki, Y.; AHMAD, D. Topological Data Analysis with Cubic Hesitant Fuzzy TOPSIS Approach. *Symmetry* **2022**, *14*, 865. [[CrossRef](#)]
190. Bauer, U. Ripser: Efficient computation of Vietoris–Rips persistence barcodes. *J. Appl. Comput. Topol.* **2021**, *5*, 391–423. [[CrossRef](#)]
191. Maria, C.; Boissonnat, J.D.; Glisse, M.; Yvinec, M. The GUDHI Library: Simplicial Complexes and Persistent Homology. In *Proceedings of the Mathematical Software—ICMS 2014; Lecture Notes in Computer Science; Springer Nature: Berlin/Heidelberg, Germany, 2014; Volume 8592*, pp. 167–174. [[CrossRef](#)]
192. Burella Pérez, J.; Hauke, S.; Lupo, U.; Caorsi, M.; Dassatti, A. giotto-ph: A Python Library for High-Performance Computation of Persistent Homology of Vietoris–Rips Filtrations. *arXiv* **2021**, arXiv:2107.05412.
193. Anang, A.N.; Chukwunweike, J.N. Leveraging Topological Data Analysis and AI for Advanced Manufacturing: Integrating Machine Learning and Automation for Predictive Maintenance and Process Optimization. *Int. J. Comput. Appl. Technol. Res.* **2024**, *13*, 27–39. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.