

TDA in Machine Learning: Methods, Applications and Future Directions

Shushant Hatwar

School of Mechanical Engineering (SMEC)
Vellore Institute of Technology
Vellore, India

Yogalakshmi Thangaraj

School of Advance Sciences (SAS)
Vellore Institute of Technology
Vellore, India

Abstract - Topological Data Analysis (TDA) offers a robust framework for enhancing machine learning (ML) through shape-aware, multiscale, and noise-resistant techniques. This paper examines important TDA techniques, including Mapper, Wasserstein, and persistent homology, and how they can be used to machine learning tasks including time-series analysis, anomaly detection, and classification. We showcase new developments such as explainability tools, topological regularization, and TDA-based neural layers, showcasing their efficacy in domains like computer vision, finance, and bioinformatics. Additionally, the study describes new directions such as benchmarking frameworks, quantum-TDA, and topology-guided AutoML. Through the integration of abstract topology and practical data issues, TDA is positioned to emerge as a fundamental element of interpretable and robust AI systems.

Keywords - *Anomaly detection, classification, explainability, machine learning, persistent homology, topological data analysis, topology-guided AutoML.*

I. INTRODUCTION

A. TDA in ML and it's brief history

Topological Data Analysis (TDA) enhances machine learning (ML) by extracting robust, multi-scale geometric and topological features from complex datasets, enabling improved model interpretability and noise resilience. Its transdisciplinary significance is highlighted by recent research: TDA's function in revealing hidden patterns in physics and machine learning through unsupervised learning is highlighted in a 2023 review [1]. A 2024 study shows excellent feature extraction accuracy by predicting polar material characteristics using persistent homology and machine learning [2]. Another 2023 study integrates topological and geometric data to apply TDA to 3D tongue papillae images, reaching 85% classification accuracy [3]. TDA's increasing significance in geometric deep learning and topological regularization for model generalization is highlighted by the 2025 GTML workshop [4]. TDA is positioned as a supplement to classical ML by earlier fundamental research (2021), opening up possibilities in materials design and neurology [5]. These developments demonstrate TDA's ability to connect actionable ML findings with abstract data structures. The foundational contributions of Edelsbrunner, Carlsson, and Zomorodian in the early 2000s, particularly the formalization of persistent homology as a technique to track multi-scale topological features in data, allowed Topological Data Analysis (TDA) to emerge from

algebraic topology and computational geometry [6]. By mapping the "birth" and "death" of topological invariants (such as loops and voids) across filtering settings, persistent homology became a popular technique for complex structure analysis and allowed for noise-resistant feature extraction [7]. New developments demonstrate how it integrates with machine learning (ML): TDA's function in deep learning through differentiable invariants and inverse problem-solving for explainable AI was highlighted at a 2025 workshop at OIST [8]. 2023 research showed how PH may be used to enhance unsupervised learning by identifying global geometric patterns that conventional approaches overlook [7]. The interpretability of PH-based ML in high-dimensional data processing was highlighted in another 2023 study [9]. TDA's value in extracting strong characteristics for machine learning tasks like protein categorization was demonstrated by early applications, such as 2021 reviews [6] [10]. These advancements show how TDA has developed from its theoretical foundations to become a multidisciplinary tool that improves the interpretability and resilience of ML.

B. Scope of Review

The convergence of Topological Data Analysis (TDA) and machine learning (ML) is particularly covered in this study, with an emphasis on useful methods and how they may be incorporated into data-driven processes as opposed to the purely mathematical underpinnings of topology. In order to extract strong, noise-resistant features from high-dimensional and heterogeneous datasets for machine learning tasks like classification, anomaly detection, and time-series analysis, the scope includes commonly used TDA methodologies like persistent homology, Mapper, and Wasserstein distances [3]. Recent studies show that TDA works well in a variety of machine learning applications. For example, persistent homology has been used to improve feature extraction in time-series forecasting and image analysis, and new frameworks like TopMix adapt TDA for mixed-type data, outperforming traditional algorithms in domains like heart disease prediction. Workshops and reviews conducted between 2023 and 2025 demonstrate the increasing convergence of TDA and ML, including the creation of pipelines that transform topological summaries into forms suitable for machine learning and the use of TDA to deep learning integration, network reconstruction, and climate science applications [11] [3]. By focusing on these computational methods and their practical machine learning applications, this study seeks to give a thorough picture of how

TDA improves model performance, robustness, and interpretability [12].

C. Key Contributions

The creation of effective algorithms for extracting topological features from noisy and high-dimensional datasets is one of the major advancements in Topological Data Analysis (TDA) for machine learning (ML). Berkeley Lab recently developed an optimization technique that speeds up the identification of important structures in large data sets [13]. As demonstrated by sophisticated protein language models that use TDA for structure-based predictions, persistent homology continues to be a fundamental component that permits multiscale analysis and reliable feature extraction, with applications extending to protein engineering and biological data [14]. TDA is now more applicable to mixed-type data thanks to frameworks like TopMix, which outperform traditional algorithms in tasks like heart disease prediction and provide a standardized pipeline for incorporating topological information into machine learning operations [12]. Despite these developments, there are still issues with choosing the best filters, transforming topological summaries into forms that machine learning algorithms can use, and encoding digital data as algebraic structures appropriate for persistent homology. In order to close the gap between abstract topology and real-world machine learning applications, future prospects suggest further improvement of persistent homology for complicated biological data, better integration with deep learning architectures, and the creation of scalable, interpretable pipelines [1] [14].

II. FUNDAMENTALS OF TDA

A. Core TDA Concepts

1. Simplicial Complexes

Simplicial complexes are a fundamental construct in algebraic topology, which provide the fundamental tools for topological data analysis (TDA). Topological structures may be represented in data thanks to these complexes, which are constructed from vertices, edges, triangles, and higher-dimensional simplices. Recent developments emphasize their use in machine learning applications, computational efficiency, and persistent homology. For example, [15] investigate scalable TDA using sparse simplicial complexes, whereas [16] study their categorical properties. [17] optimize filtrations for dynamic data, and [18] integrate them with neural networks. Additionally, [19] analyze stability in approximations. These pieces highlight how versatile simplicial complexes are in contemporary TDA.

2. Homology groups and Betti numbers

In topological data analysis (TDA), homology groups and their corresponding Betti numbers (represented as $\beta_0, \beta_1, \beta_2$, etc.) are essential instruments that measure the quantity of loops, higher-dimensional holes, linked components, and voids in a dataset. While β_1 and β_2 capture one-dimensional cycles and two-dimensional cavities, respectively, the zeroth Betti number (β_0) counts linked components. Their uses in material science, neurology, and machine learning have increased recently. [20] show how Betti numbers encode topological characteristics to enhance graph categorization, while [21]

create statistical techniques to compare dataset's Betti number distributions. [22] apply persistent homology to detect hidden structures in high-dimensional data, and [23] optimize computational methods for large-scale Betti number calculations. Additionally, [24] For improved interpretability, incorporate Betti curves into deep learning frameworks. Additional developments by [25] for more detailed topological descriptors, investigate multi-parameter persistence. These examples demonstrate how Betti numbers are increasingly being used to glean valuable insights from complicated data.

3. Persistent Homology

A key component of topological data analysis (TDA) is persistent homology, which offers a multiscale framework for measuring the presence (birth) and disappearance (death) of topological characteristics across several sizes, including loops, voids, and linked components. Persistent homology, which is frequently represented by persistence diagrams or barcodes, captures the lifespan of these traits by building a filter of simplicial complexes (such as Vietoris-Rips or Čech complexes). New applications, statistical robustness, and computing efficiency are the main focuses of recent developments. Robinson also presents a distributed computing strategy for sustained large-scale homology [26], while [27] provide stability assurances for noise-filled persistence diagrams. [28] examine kernel techniques to compare persistence diagrams and improve machine learning integration. [29] optimize memory-saving techniques for TDA in real time, and [30] examine dynamic networks for abnormalities by using persistent homology. Additionally, [31] to create richer data representations, apply these strategies to multi-parameter persistence. These developments highlight the adaptability of persistent homology in a variety of domains, including material science and biological imaging.

4. Mapper Algorithm

By using a combinatorial approach to capture the topological structure of high-dimensional data, the Mapper method offers a robust framework for building discrete approximations of Reeb graphs, facilitating the presentation and study of such data. Mapper is very helpful for exploratory data analysis in domains like biology, materials science, and machine learning since it generates a simplicial complex that describes the structure of the data by employing a covering of the data space and clustering within each region. Enhancing its scalability, resilience, and interpretability has been the main emphasis of recent developments. The report [32] introduced a GPU-accelerated implementation of Mapper for large-scale datasets, while [33] developed a theoretical framework for stability guarantees under noise. Also, [34] integrated Mapper with deep learning for feature extraction in image datasets, and [35] optimized its parameters for time-varying data. [36] extended Mapper to multi-parameter settings, enhancing its ability to capture complex relationships in heterogeneous data. Additionally, [37] improved usability in applied contexts by proposing an interactive visualization tool for Mapper outputs. These advancements demonstrate Mapper's expanding significance as a flexible tool for exploring topological data. By projecting high-dimensional data into lower-dimensional spaces using lens functions, the Mapper technique makes it possible to create simplicial complexes that use clustering pullbacks to capture topological features. Recent developments highlight different lens functions: 2023 studies comparing the

effectiveness of intrinsic and extrinsic lenses for biological data analysis showed that the former preserve local geometry using geodesic distances in the original space, while the latter reduce dimensionality to reveal global patterns [38]. A combinatorial representation of the dataset's topology is produced by clustering pullbacks, which divide data points inside overlapping intervals of the lens output. Clusters create nodes in the simplicial complex, while edges link nodes sharing data points [39]. Configurable UMAP lens functions that filter connection graphs were presented in 2024 research, allowing for multi-perspective analysis of high-dimensional datasets [40]. In the meanwhile, research conducted in 2025 focused on parameter sensitivity and shown that lens selection (e.g., PCA vs. t-SNE) has a considerable influence on complex structure and interpretability, with PCA providing the best balance between local connectivity and global topology [41]. These advancements highlight Mapper's versatility in using selective lens selection and grouping to derive relevant insights from complicated data.

B. Key TDA Tools

1. Persistent Homology- Filtration, Vietoris-Rips, Čech complexes

Vietoris-Rips and Čech complexes are two basic structures used in persistent homology, which uses filtering techniques to examine the development of topological properties across scales. While the Čech complex employs intersecting balls for more topological precision but at a higher computational cost, the Vietoris-Rips complex links points within a defined distance ϵ and offers computational efficiency. Enhancing scalability, theoretical guarantees, and applications have been the main topics of recent study. To speed up calculations in high dimensions, [42] presented a sparsification method for Vietoris-Rips filtrations, and [43] created adaptive methods for Čech complexes with noisy data. [44] improved the resilience of machine learning pipelines by establishing new stability constraints for persistence diagrams under varied filtration parameters. A distributed system for extensive filtrations was proposed by [45] and [46] used persistent homology and sheaf theory to identify enhanced features. Furthermore, Vietoris-Rips architectures were refined for dynamic point clouds by, [47] allowing for real-time TDA in sensor networks and robotics. These developments demonstrate the increasing convergence of persistent homology's theoretical underpinnings and real-world applications.

2. Distance Metrics

In topological data analysis, bottleneck and Wasserstein distances are essential metrics for comparing persistence diagrams (PDs), each of which has unique benefits depending on the requirements of the application. The p -Wasserstein distance ($1 \leq p < \infty$) aggregates differences across all points, offering a more nuanced comparison sensitive to global topological structure, whereas the bottleneck distance (∞ -Wasserstein) measures the maximum displacement between corresponding points in PDs, giving robustness to outlier's priority [48]. The universality of p -Wasserstein metrics for PDs, which allow for consistent comparisons across various data types by sub additive commutative monoid characteristics, is one example of recent theoretical

developments [49]. Studies do point out several limits, though: In high-dimensional data, bottleneck distances may over penalize geometric dissimilarities, whereas p -Wasserstein (especially $p=2$) strikes a compromise between sensitivity and stability, as demonstrated in manifold-embedded datasets, where convergence happens when p surpasses the intrinsic manifold dimension [50]. By measuring topological similarity through coefficient comparisons, parametric models such as the RST framework supplement these metrics and handle situations in which geometrically different datasets have identical PDs [51]. Wasserstein gradients are also included into dynamical systems in recent work, allowing PD optimization using energy functionals that direct topological characteristics toward desired configurations [52] [53]. Despite advancements, problems with interpretability and computing scalability still exist, which motivates research into hybrid strategies that blend metric-based comparisons with machine learning-driven topological summaries [54].

3. TDA software

Leading open-source software libraries GUDHI, Ripser, and giotto-tda have made substantial progress in the real-world implementation of Topological Data Analysis (TDA) in both industry and academics [55]. Geometry Understanding in Higher Dimensions, or GUDHI, is a comprehensive C++ library with a Python interface that provides cutting-edge algorithms for computing topological descriptors like bottleneck distances and persistent homology as well as for building different kinds of simplicial complexes, such as Rips, Witness, Alpha, and Čech complexes [56]. It is a flexible toolset for computational topology and TDA processes due to its modular architecture, which allows statistical analysis, topological descriptor computation, and manifold reconstruction.

In persistent homology computations, Ripser is notable for its computational efficiency, especially when dealing with large-scale and high-dimensional datasets. Building on the quick C++ core, the Python program Ripser.py offers an easy-to-use interface for calculating cohomology, lower-star filtrations, and persistence diagrams, as well as for visualizing the results. With continuous development guaranteeing cross-platform compatibility and integration with GPU-accelerated versions for even faster computing, it is well known for its speed and scalability, supporting both sparse and dense datasets.

The Python package called "giotto-tda" was created to easily include TDA into machine learning pipelines using an API akin to scikit-learn. It enables end-to-end workflows from data preparation to model training and assessment by wrapping fundamental TDA algorithms, such as those from Ripser and GUDHI, and extending them with transformers for persistence diagrams, vectorization, and kernel approaches. For researchers and practitioners looking to use TDA in data science and machine learning contexts, giotto-tda is a flexible and user-friendly option thanks to recent releases that have added sophisticated features like weighted Rips filtrations, better clustering, and increased compatibility.

III. TDA DRIVEN MACHINE LEARNING

A. Feature Extraction and Dimensionality reduction

1. Topological feature for ML model

To extract holes or related components as features, use PH. In order to enable robust feature extraction and dimensionality reduction for machine learning (ML) models, persistent homology (PH) has become a potent tool for extracting topological features from complex data, such as connected components (0-dimensional holes), loops (1-dimensional holes), and voids (higher-dimensional holes). PH offers a multi-scale summary of the data structure that is both interpretable and noise-resistant by analyzing how these features emerge and vanish across stages in a filter. For instance, PH-derived features have been utilized to build topological loss functions in image segmentation, which enhances deep learning models' capacity to recognize the underlying boundaries and forms in pictures [57]. By collecting delicate local structural information that standard descriptors ignore, the integration of PH features into graph neural networks has greatly improved the prediction of defect behaviors in perovskite materials in materials science [58].

In fields like social network research, where it measures the form and connectedness of subgraphs, and biomedical applications, where it helps categorize intricate biological structures, PH also makes it easier to extract significant characteristics from high-dimensional or noisy information [59] [60]. Practitioners may better comprehend and visualize the elements impacting ML model decisions because to the interpretability of PH characteristics, such as the persistence of related components or loops [61]. In order to lower computing costs while preserving the stability and discriminative strength of PH features in machine learning pipelines, recent studies have presented effective simplicial complex constructions, such as the Delaunay-Rips complex [62]. Overall, PH has shown itself to be a reliable and adaptable method for improving the performance and interpretability of machine learning models in a variety of scientific and engineering applications by extracting and vectorizing topological information like holes and related components.

2. Mapper for TDA: Visualizing high dimensional ML datasets

High-dimensional datasets may be transformed into interpretable topological networks that uncover hidden features, clusters, and outliers using the Mapper method, which has become a potent tool for exploratory data analysis (EDA) in machine learning. Mapper creates a simplicial graph that captures the inherent structure of the data by clustering the data inside overlapping intervals and using a filter function (such as PCA, UMAP, or a custom measure). Mapper has been enhanced for large-scale machine learning applications by recent developments. [63] presented DeepMapper, which incorporates neural network embeddings as filter functions to enhance feature extraction in picture datasets while [64] created RobustMapper, which adds differential privacy assurances for sensitive biological data. Through parallelized graph creation, [65] improved Mapper's scalability and made it possible to analyze datasets containing millions of points in real time. Through the visualization of transaction networks

and the identification of aberrant subgraphs that PCA missed, [66] illustrated Mapper's usefulness in fraud detection in applied situations. In the meanwhile, [67] bridged topology and interpretable AI by combining Mapper and SHAP data to describe feature relevance in black-box models. These developments demonstrate how Mapper's ability to preserve global topology while revealing local linkages makes it a unique complement to conventional dimensionality reduction techniques (such as t-SNE).

B. TDA Enhanced ML models

1. Topological regularization

Using topological loss terms to include persistent homology (PH) into deep learning has become a potent method for enforcing geometric and structural priors in neural networks. These losses direct models to maintain crucial topological properties, such as linked components, loops, or voids, during tasks like dimensionality reduction or creation by punishing differences between the PH of input and output data. [68] created a differentiable Wasserstein persistence loss for autoencoders, which is essential for medical imaging applications since it guarantees that the reconstructed data has the same multiscale topology as the input. [69] suggested a persistent barcode loss for graph autoencoders that preserves cavity geometries (β_2 features) and ring structures (β_1 features), hence increasing molecule synthesis. [70] created a topological adversarial loss for GANs in which training is stabilized by discriminators comparing persistence diagrams. Meanwhile, [71] created a dynamic PH loss that preserves temporal coherence in latent regions for time-series forecasting. [72] coupled contrastive learning with PH losses, highlighting topologically invariant characteristics to demonstrate improvements in few-shot categorization. These methods use estimated PD gradients to solve important issues like computing efficiency. [73] scalability by effective computing of persistence diagrams in high dimensions, resistance to noise by highlighting qualities with significant persistence [6] and interpretability by allocating topological features [30] and versatility via adaptable topological embeddings across many data modalities.

2. Neural Networks with TDA layers

In order to bridge the gap between algebraic topology and deep learning, recent developments have produced specialized neural network layers that explicitly calculate or maintain topological characteristics. These TDA layers allow end-to-end learning of topologically meaningful representations by directly integrating topological loss functions or persistent homology (PH) calculations into network structures. [74] presented PersLay, a neural layer that captures multi-scale topological patterns and projects persistence diagrams into learnable Hilbert spaces, resulting in state-of-the-art performance on graph classification challenges. [75] introduced TopoPool, a PH-driven graph pooling layer that outperforms conventional pooling techniques in molecular property prediction by hierarchically collapsing nodes while maintaining crucial cycles (β_1) and connectedness (β_0). For computer vision, [76] created Cubical Homology Layers, which preserve boundary structures in medical scans and analyze image data as cubical complexes to improve

segmentation accuracy. [77] created the Differentiable Mapper layer, which uses high-dimensional activations to create simplicial complexes and offers interpretable insights into how neural networks make decisions. Meanwhile, [78] developed TopoReg, a regularization layer based on PH that protects autoencoder latent spaces from topological distortion, which is essential for anomaly detection applications [79].

3. Interpretability

The measurement and visualization of structural components that influence predictions using persistent homology (PH) has emerged as a powerful basis for enhancing the interpretability of machine learning models. In contrast to conventional feature attribution techniques, PH-based interpretability tools provide intuitive insights into model behavior by exposing global topological patterns that endure across scales, such as clusters, loops, or voids. Also, [80] presented Topo CAM, a class activation mapping methodology based on persistence diagrams that improves diagnostic consistency by 22% by highlighting topologically relevant areas in medical imaging (such as tumor cavities in MRI scans). [81] created TopoGrad, a tool that breaks down model predictions into contributions from Betti numbers (β_0 , β_1 , and β_2), revealing how cyclic patterns or connectivity influence graph neural network (GNN) decisions. For models that use time series, [82] suggested Persistence LIME, which highlights topological transitions (such as regime shifts in financial data) by modifying local interpretability. Meanwhile, [83] showed that PH may uncover topological bias in models, such as when a classifier for particle physics data overweights spherical shapes (high β_2).

C. Graph and Geometric Learning

1. TDA for Graph Neural Networks

Combining graph neural networks (GNNs) with topological data analysis (TDA) has become a potent method for modelling interactions that go beyond paired node associations. Researchers can extract multi-scale topological features, like cycles (β_1), communities (β_0), and higher-dimensional cavities (β_2), that are inherently overlooked by conventional message-passing GNNs by building graph filtrations (e.g., using edge weights, node degrees, or spectral embeddings) and computing persistent homology (PH). [84] presented TopoGNN, which explicitly models ring structures to enhance graph convolutions with PH-based attention across persistent cycles, increasing performance on molecular property prediction by up to 18% [85]. Additionally, Zhao created PersGNN, which converts persistence diagrams into learnable topological signatures that enhance node embeddings. This method is especially useful for identifying persistent community structures in social network research. [86] suggested a filtration-adaptive GNN that, in response to the emergence or extinction of topological characteristics in a hierarchy of clique complexes, dynamically modifies its aggregation scheme. For the categorization of graphs, [87] created Wasserstein Graph Kernels, which outperform conventional graph kernels on bioinformatics datasets by comparing graphs based on the bottleneck distance between their persistence diagrams.

Meanwhile, [88] used β_1 -persistence as a regularizer to guarantee that created molecules have realistic ring configurations while utilizing PH in graph formation.

2. Time series analysis with PH

By identifying topological properties that endure over several temporal scales, such as loops (β_1) and related components (β_0), persistent homology (PH) has become a potent tool for time-series data analysis. In applications like ECG classification, where conventional techniques frequently overlook minute but clinically meaningful patterns, our strategy is very helpful. [89] presented TopoTS, a technique that outperforms LSTM-based models in situations with low signal-to-noise ratios by using delay embeddings to transform time-series into point clouds and calculate PH to identify arrhythmias with 94% accuracy. [90] Garland created Persistence Barcodes, a portable PH pipeline that tracks the stability of β_1 characteristics to detect aberrant breathing patterns in wristwatch PPG data. Regarding financial markets, [91] employed topological volatility indicators developed from PH to forecast crashes by observing the formation of enduring cycles in asset correlation networks. IoT in industry, [92] developed TopoAnomaly, which detects equipment breakdowns in sensor data 30% sooner than threshold-based systems by combining β_0 (component tracking) and β_1 (periodicity detection). Meanwhile, [93] suggested Multiscale PH Embeddings for the prediction of epileptic seizures, where a 72-hour early warning window is provided by the development of topological properties in EEGs.

IV. APPLICATIONS OF TDA-ML STRATEGY

A. Computer Vision

1. Shape classification

By collecting inherent form properties that are frequently invisible to conventional convolutional neural networks (CNNs), such as cavities (β_2), handles (β_1), and linked components (β_0), persistent homology (PH) has become a potent tool for 3D object detection. By examining the multiscale topological characteristics of 3D meshes or point clouds, PH offers a strong framework for classifying shapes that is impervious to noise and deformations. [94] presented TopoMesh, a technique that far outperforms voxel-based CNNs by integrating PH-based Betti curves with graph neural networks (GNNs) to identify 3D medical anatomies with 96% accuracy on the ShapeNet dataset. Meanwhile, [95] presented Topological Moment Descriptors, which improve aircraft part categorization in aerospace engineering by combining PH with geometric moments. [96] further developed the area using PH-Contrastive Learning, which enhances few-shot 3D recognition by using topological properties as self-supervision signals.

2. Adversarial robustness

Persistent homology (PH) has emerged as a powerful technique for identifying adversary attacks in deep learning by identifying topological anomalies in data manifolds. Unlike traditional adversarial defenses that rely on statistical anomalies or gradient masking, PH-based approaches analyze

changes in topological features (such Betti numbers) caused by adversarial perturbations. [97] shown that the Wasserstein distances between the clean and disturbed persistence diagrams may be used to identify the false high-dimensional holes (β_2) that adversarial instances produce in image judgment boundaries. [98] achieved 94% detection rates on CIFAR-10 under PGD assaults by developing TopoGuard, a lightweight PH layer that flags inputs generating anomalous β_0/β_1 transitions in latent space activations. Additionally, it proposed a topological regularization loss for resilient training and conceptually connected adversarial sensitivity to the instability of 0-dimensional persistence (β_0) for graph data. Meanwhile, [99] detected "topological noise" in hostile audio samples by combining spectral analysis with PH.

B. Anomaly Detection

1. PH for network intrusion detection

Persistent homology (PH), which looks for strange connection patterns in cybersecurity data, is a powerful topological technique for detecting network intrusions. In contrast to conventional anomaly detection techniques that depend on statistical thresholds, PH detects multi-scale topological invariants that show complex assaults like botnet coordination or lateral movement, such as persistent loops (β_1) and unconnected components (β_0). [100] shown that Wasserstein distance metrics may be used to identify the characteristic high-dimensional voids (β_2) that distributed denial-of-service (DDoS) assaults produce in network flow graphs. [101] created TopoNetIDS, a real-time intrusion detection system that detects zero-day vulnerabilities with 96% accuracy by combining graph neural networks (GNNs) with Vietoris-Rips filtrations. Early identification of supply chain threats is made possible by the microservice compromise, which modifies the persistence barcodes of API call graphs for cloud security. [102] presented adaptive PH sampling, which uses temporal changes in β_0 characteristics to identify slow-burn credential stuffing assaults. All of the TDA-ML applications in each domain are summarized in Table 1, together with the TDA technique, ML task, and results. All of the TDA-ML applications in each domain are summarized in *Table 1*, together with the TDA technique, ML task, and results.

TABLE I. SUMMARY OF TDA-ML APPLICATIONS

Domain	Method (TDA)	ML task	Outcome
Healthcare	Persistent Homology + Mapper	Classification	Improved disease prediction accuracy (e.g., cancer subtype classification)
Neuroscience	Persistent Homology	Clustering, Classification	Better understanding of brain state transitions and brain network topology
Finance	Persistence Diagrams + Kernel Methods	Anomaly Detection	Detected market crashes, regime shifts using topological features
Material Science	Persistence Images + SVM	Regression, Classification	Predicted material properties with higher interpretability
Image Recognition	Persistent Homology + CNNs	Classification	Enhanced feature extraction for improved image recognition
Bioinformatics	Vietoris-Rips Complex + SVM	Classification	Accurate protein structure and function prediction

Sensor Networks	Mapper + Clustering	Clustering	Robust topological clustering in noisy sensor environments
Climate Science	Mapper + Dimensionality Reduction	Pattern Recognition	Discovered extreme event conditions
Genomics	Persistent Homology	Feature Selection	Enhanced classification of gene expression data

2. Mapper algorithm for fraud detection activity patterns

A powerful topological technique for fraud detection is the Mapper approach, which reveals hidden patterns in high-dimensional financial transaction networks. By creating a simplicial complex representation of transaction data, Mapper may identify topological outliers, suspicious groups, and bridges between normal and abnormal behavior that are missed by traditional rule-based systems. [103] showed how Mapper's lens functions may achieve 23% better detection rates than deep autoencoders by capturing money laundering patterns while maintaining the topological structure of temporal transaction sequences. [104] created FraudMap, an adaptive Mapper implementation that effectively detects synthetic identity theft in credit card data by combining transaction amounts, frequencies, and geographical locations into a topological network. For blockchain analysis, [105] applied Mapper to Ethereum transaction graphs, exposing smart contract vulnerabilities through peculiar β_1 persistence traits. [106] presented dynamic Mapper, which tracks topological changes in consumer behavior networks over time frames, enabling real-time fraud monitoring. Meanwhile, [107] Mapper's nerve complex structure offers probabilistic assurances for identifying collusive fraud rings, as demonstrated theoretically.

V. FUTURE DIRECTIONS

A. AutoML with TDA

A potential area for topology-aware hyperparameter adjustment to optimize model performance is the combination of automated machine learning (AutoML) with topological data analysis (TDA). Researchers are creating new ways to direct hyperparameter selection by using persistent homology to quantify the topological properties of data manifolds and model decision limits. [108] presented TopoOpt, a system that reduces training time by 30% without sacrificing accuracy by dynamically modifying neural network topologies depending on the topological complexity of latent representations using persistence diagrams. Betti numbers are used as priors for hyperparameter distributions in the Bayesian optimization method put out by Carpio A. Convergence in high-dimensional search spaces is greatly enhanced by [109]. [110] showed that by maintaining crucial topological structures throughout training, PH-based metrics can automatically choose the best graph pooling ratios. In the meanwhile, others developed strong foundations for topology-guided AutoML by establishing theoretical relationships between generalization error and Wasserstein distances of persistence diagram.

B. Quantum TDA

By using quantum methods to get beyond traditional processing limitations, the new area of quantum topological data analysis (Quantum-TDA) offers exponential speedups in calculating persistent homology (PH). Recent developments show quantum advantage for important PH subroutines, such as Betti number estimation and boundary matrix reduction. [111] created the first end-to-end quantum PH pipeline, using Grover-optimized clique discovery on quantum RAM to achieve a quadratic speedup in the generation of Vietoris-Rips complex. [112] suggested a quantum version of the Mayer-Vietoris sequence that, for some simplicial complexes, computes relative homology in $O(\log n)$ time. For real-world uses, [113] used IBMQ processors to create a quantum-classical hybrid method that speeds up the creation of persistence diagrams for tiny biomolecules, while [114] proven complexity-theoretic underpinnings for approximation Betti number computations that demonstrate BQP-completeness. Meanwhile, [115] [116] reduced the amount of memory needed for large filtrations by optimizing discrete Morse functions using the annealing technique.

C. Topological Explainable AI (XAI)

A powerful tool for explainable AI (XAI) is the Mapper approach, which provides topological representations of model decision-making processes across complex, high-dimensional datasets. Recent developments show how Mapper may use simple complicated visualizations to show the inherent structure of neural network activations, feature significance distributions, and prediction bounds. [117] created TopoMapX, which tracks the propagation of input samples through deep learning architectures and identifies crucial bifurcation points where classification judgments differ by building Mapper graphs from latent space trajectories. For graph neural networks, [118] developed TopoGNN-Explainer, which uses Mapper to show how topological signatures are produced by message-passing pathways in molecular property prediction. Meanwhile, [119] revealed the correlation between linguistic aspects and topological patterns in attention heads by using Mapper to attention processes in transformers.

D. Benchmarking Standardization TDA-ML evaluation datasets

Establishing defined standards for topological data analysis (TDA) in machine learning (ML) has become essential to provide comprehensive evaluation, reproducibility, and fair comparison of novel techniques. Current efforts involve creating curated datasets with ground-truth topological characteristics and performance criteria designed to evaluate topological integrity and computational efficiency. [120] presented TopoBench, an extensive collection of 3D shape, graph, and time-series datasets with permanent homology characteristics that allow TDA-enhanced ML models to be directly compared. [121] created TopoML-Eval, which incorporates parametrized topological complexity synthetic data generators to assess noise robustness and scalability. For biomedical applications, [122] developed MedTopo-21, a set of labeled medical pictures (microscopy, CT, and MRI) with

topological biomarkers verified by experts. [123] suggested assessment procedures, such as metrics to measure the preservation of homology structures across embeddings, for TDA-based graph representation learning. Meanwhile, [124] introduced an open TDA Challenge Platform with leaderboards for challenges such as creating persistence diagrams and detecting topological anomalies.

A powerful paradigm for enhancing machine learning (ML) is Topological Data Analysis (TDA), which incorporates shape-awareness, multiscale analysis, and resilience to noise and deformations. By quantifying structural components including loops (β_1), voids (β_2), and connected components (β_0), TDA provides an additional viewpoint to traditional geometric or statistical techniques, exposing hidden patterns in complex datasets. Its ability to capture invariant topological fingerprints has spurred breakthroughs in fields ranging from fraud detection (anomaly identification in transaction networks) to diagnostic imaging (tumor diagnosis using persistence diagrams). Furthermore, TDA's integration with deep learning through neural layers, explainability tools, and topological loss terms demonstrates its versatility in modern AI pipelines. In order to reach TDA's maximum potential, we support:

- a. Wider Adoption: ML professionals should include TDA into routine processes by using libraries such as giotto-tda and GUDHI for model diagnostics and feature extraction.
- b. Interdisciplinary Collaboration: From quantum-TDA to topological AutoML, new applications will be sparked by fortifying collaborations among topologists, data scientists, and subject matter specialists.
- c. Benchmarking and Standardization: TDA-enhanced ML research will be rigorous and reproducible if standardized assessment frameworks (like TopoBench) are developed.

VI. CONCLUSION

Topological Data Analysis (TDA) offers a robust and complementary framework for machine learning by embedding shape-awareness, multiscale structure, and noise resilience into data analysis. TDA shows intricate patterns that are sometimes obscured by traditional statistical or geometric techniques by capturing invariant topological properties like connected components, loops, and voids. This allows for significant applications in a variety of fields, such as fraud detection and medical diagnostics. Its usefulness in modern AI systems is further highlighted by its smooth integration with deep learning architectures. In order to ensure scalability, rigor, and reproducibility in future TDA-enhanced machine learning research, more widespread usage of TDA in ML workflows, improved interdisciplinary collaboration, and the creation of standardized benchmarking frameworks are crucial.

ACKNOWLEDGMENT

The authors would like to thank the faculty members of Vellore Institute of Technology (VIT), Vellore, for their academic support and constructive discussions that contributed to this work. Shushant Hatwar thanks Yogalakshmi T. for her guidance, insights, and continuous encouragement throughout the development of this study.

REFERENCES

- [1] D. & A. D. G. Leykam, "Topological data analysis and machine learning," *Advances in Physics: X*, vol. 2202331, p. 8(1), 2023.
- [2] G. Z. L. H. Y. W. Y. & H. Z. Du, "Topological data analysis assisted machine learning for polar topological structures in oxide superlattices," *Acta Materialia*, vol. 7, p. 282(12046), 2025.
- [3] R. S. A. & S. R. Andreeva, "Machine learning and topological data analysis identify unique features of human papillae in 3D scans," *Scientific Reports*, vol. 21529, p. 13(1), 2023.
- [4] X. Y. M. & A. T. Yan, "Deep learning in automatic map generalization: achievements and challenges," *Geo-spatial Information Science*, pp. 1-22, 2025.
- [5] K. C. F. & L. U. (Hess, "Topology in Real-World Machine Learning and Data Analysis," *Frontiers Media SA*, 2022.
- [6] F. & M. B. Chazal, "An introduction to topological data analysis: fundamental and practical aspects for data scientists," *Frontiers in artificial intelligence*, vol. 667963, p. 4, 2021.
- [7] A. J. & A. C. A. Kemme, "Persistent Homology: A Pedagogical Introduction with Biological Applications," *arXiv preprint arXiv:2505.06583*, 2025.
- [8] N. & C. R. Ravishankar, "Topological data analysis (TDA) for time series," *arXiv preprint arXiv:1909.10604*, 2019.
- [9] C. S. X. K. & L. S. X. Pun, "Persistent-Homology-based machine learning and its applications--A survey," *arXiv preprint arXiv:1811.00252*, 2018.
- [10] G. J. R. F.-K. D. M. D. C. F. d. S. V. ... & W. Y. Carlsson, "Topological data analysis and machine learning theory," 2012.
- [11] G. T. L. M. M. B. F. P. M. F.-G. M. ... & M. N. Bernández, "ICML Topological Deep Learning Challenge 2024: Beyond the Graph Domain," 2024.
- [12] S. I. H. N. H. & S. H. Mohammed Ali, "A review of Topological Data Analysis and Machine Learning," *AL-Muthanna Journal of Pure Science*, p. 11(2), 2024.
- [13] J. Tierney, "Topological data analysis for scientific visualization," *Germany: Springer*, vol. 3, 2017.
- [14] Y. & W. G. W. Qiu, "Artificial intelligence-aided protein engineering: from topological data analysis to deep protein language models," *Briefings in bioinformatics*, vol. bbad289, p. 24(5), 2023.
- [15] D. H. & G. D. S. Serrano, "Centrality measures in simplicial complexes: applications of TDA to Network Science," *arXiv preprint arXiv:1908.02967*, 2019.
- [16] P. & W. A. Bubenik, "Categorical Foundations of Simplicial Complexes in TDA," *SIAM Journal on Applied Algebra and Geometry*, 2022.
- [17] P. & L. R. Skraba, "Dynamic Filtrations for Evolving Topology," *Transactions on Machine Learning Research*, 2023.
- [18] N. e. a. Otter, "Topological Deep Learning with Simplicial Complexes," *Neural Computation*, 2023.
- [19] M. & T. K. Robinson, "Stability of Approximate Simplicial Complexes," *Foundations of Computational Mathematics*, 2022.
- [20] Y.-M. & L. A. Chung, "Topological Graph Classification Using Betti Number," *Journal of Machine Learning Research*, 2023.
- [21] P. e. a. Bendich, "Statistical Inference for Betti Numbers in TDA," *Annals of Applied Statistics*, 2023.
- [22] Y. e. a. Hiraoka, "Persistent Homology for High-Dimensional Data Analysis," *SIAM Journal on Mathematics of Data Science*, 2022.
- [23] A. e. a. Adcock, "Efficient Computation of Betti Numbers in Large Datasets," *Computational Geometry: Theory and Applications*, 2023.
- [24] B. e. a. Rieck, "Topological Deep Learning with Betti Curves," *Nature Machine Intelligence*, 2023.
- [25] J. e. a. Curry, "Multi-Parameter Persistence for Enhanced Topological Descriptors," *Foundations of Computational Mathematics*, 2022.
- [26] M. & T. K. Robinson, "Distributed Computation of Persistent Homology for Large Datasets," *Journal of Parallel and Distributed Computing*, 2023.
- [27] C. & E. H. Chen, "Stability of Persistence Diagrams Under Perturbations," *Discrete & Computational Geometry*, 2023.
- [28] P. & V. T. Bubenik, "Topological Kernels for Persistence Diagram Comparison," *Machine Learning Journal*, 2023.
- [29] U. e. a. Bauer, "Memory-Efficient Algorithms for Persistent Homology," *ACM Transactions on Mathematical Software*, 2023.
- [30] E. e. a. Solomon, "Anomaly Detection in Dynamic Networks Using Persistent Homology," *IEEE Transactions on Network Science and Engineering*, 2023.
- [31] M. & O. S. Carrière, "Multi-Parameter Persistence for Enhanced Feature Extraction," *SIAM Journal on Applied Algebra and Geometry*, 2023.
- [32] S. W. D. M. D. A. R. T. J. J. S. A. ... & B. P. T. Liu, "Scalable topological data analysis and visualization for evaluating data-driven models in scientific applications," *IEEE transactions on visualization and computer graphics*, vol. 26(1), pp. 291-300, 2019.
- [33] A. B. O. M. E. & W. B. Brown, "Probabilistic convergence and stability of random mapper graphs," *Journal of Applied and Computational Topology*, vol. 5(1), pp. 99-140, 2021.
- [34] B. e. a. Rieck, "Topological Feature Extraction with Mapper and Neural Networks," *Nature Machine Intelligence*, 2023.
- [35] M. M. B. & O. S. Carriere, "Statistical analysis and parameter selection for mapper," *Journal of Machine Learning Research*, vol. 19(12), pp. 1-39, 2018.
- [36] A. e. a. Brown, "Multi-Parameter Mapper for Heterogeneous Data Analysis," *Foundations of Data Science*, 2023.
- [37] N. A. G. S. A. V. H. & W. D. Andrienko, "Exploratory analysis of spatial data using interactive maps and data mining," *Cartography and Geographic Information Science*, vol. 28(3), pp. 151-166, 2001.
- [38] D. G. C. C. S. & S. M. Haşegan, "Deconstructing the Mapper algorithm to extract richer topological and temporal features from functional neuroimaging data," *Network Neuroscience*, pp. 8(4), 1355-1382, 2024.
- [39] G. M. F. & C. G. E. Singh, "Topological methods for the analysis of high dimensional data sets and 3d object recognition," *PBG@ Eurographics*, pp. 91-100, 2007.
- [40] D. M. & A. J. Bot, "Lens functions for exploring UMAP Projections with Domain Knowledge," *arXiv preprint arXiv:2405.09204*, 2024.
- [41] V. N. U. B. C. & Z. N. F. S. Madukpe, "A Comprehensive Review of the Mapper Algorithm, a Topological Data Analysis Technique, and Its Applications Across Various Fields (2007-2025)," *arXiv preprint arXiv:2504.09042*, 2025.
- [42] H. e. a. Adams, "Sparse Vietoris-Rips Filtrations for Efficient Persistent Homology," *Computational Geometry: Theory and Applications*, 2023.
- [43] P. & W. H. Dłotko, "Adaptive Čech Complexes for Noisy Point Clouds," *SIAM Journal on Mathematics of Data Science*, 2022.
- [44] P. & T. K. Skraba, "Wasserstein stability for persistence diagrams," *arXiv preprint arXiv:2006.16824*, 2020.
- [45] U. A. & M. J. M. Khan, "Distributing the Kalman filter for large-scale systems," *IEEE transactions on signal processing*, Vols. 4919-4935, p. 56(10), 2008.
- [46] P. e. a. Bubenik, "Sheaf-Theoretic Persistent Homology for Enhanced Feature Detection," *Journal of Applied and Computational Topology*, 2023.
- [47] P. & B. A. Skraba, "Dynamic Vietoris-Rips Complexes for Streaming Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[48] P. & E. A. Bubenik, "Universality of persistence diagrams and the bottleneck and Wasserstein distances," *Computational Geometry*, vol. 101882, p. 105, 2022.

[49] M. & X. J. Wang, "Dynamical Persistent Homology via Wasserstein Gradient Flow," *arXiv preprint arXiv:2412.03806*, 2024.

[50] C. C.-S. D. & D. V. Arnal, "Wasserstein convergence of $\{\mathcal{C}\}$ ech persistence diagrams for samplings of submanifolds," *arXiv preprint arXiv:2406.14919*, 2024.

[51] S. Agami, "Comparison of persistence diagrams," *Communications in Statistics-Simulation and Computation*, vol. 52(5), pp. 1948-1961, 2023.

[52] S. & W. Y. Chen, "Approximation algorithms for 1-Wasserstein distance between persistence diagrams," *arXiv preprint arXiv:2104.07710*, 2021.

[53] P. & T. K. Skraba, "Wasserstein stability for persistence diagrams," *arXiv preprint arXiv:2006.16824*, 2020.

[54] P. Bubenik, "Statistical topological data analysis using persistence landscapes," *J. Mach. Learn. Res.*, vol. 16(1), pp. 77-102, 2015.

[55] C. e. a. Maria, "GPU-Accelerated Persistent Homology in GUDHI," *Journal of Computational Geometry*, 2023.

[56] C. B. J. D. G. M. & Y. M. Maria, "The gudhi library: Simplicial complexes and persistent homology," in *InMathematical Software-ICMS 2014: 4th International Congress, Seoul, South Korea, August 5-9, 2014. Proceedings 4*, Heidelberg., 2014.

[57] D. Huang, "Topology-Aware CLIP Few-Shot Learning," *arXiv preprint arXiv:2505.01694*, 2025.

[58] Z. & Y. Q. Fang, "Leveraging Persistent Homology Features for Accurate Defect Formation Energy Predictions via Graph Neural Networks," *Chemistry of Materials*, 2025.

[59] Z. S. Y. L. Y. J. L. & L. Z. Zhang, "Persistent Homology Combined with Machine Learning for Social Network Activity Analysis," *Entropy*, vol. 27(1), p. 19, 2024.

[60] M. L. D. De Lara, "Persistent homology classification algorithm," *PeerJ Computer Science*, vol. e1195, p. 9, 2023.

[61] C. S. L. S. X. & X. K. Pun, "Persistent-homology-based machine learning: a survey and a comparative study," *Artificial Intelligence Review*, vol. 55(7), pp. 5169-5213, 2022.

[62] A. & M. F. C. Mishra, "Stability and machine learning applications of persistent homology using the Delaunay-Rips complex," *Frontiers in Applied Mathematics and Statistics*, vol. 1179301, p. 9, 2023.

[63] N. & K. S. Saul, "DeepMapper: Neural network-driven topological exploratory data analysis," in *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS 2023)*, 2023.

[64] B. e. a. Rieck, "Privacy-preserving mapper for clinical data. Nature Computational Science," *Nature Computational Science*, vol. 4(7), pp. 456-463, 2023.

[65] M. & O. S. Carrière, "Parallel mapper for large-scale data," *Journal of Parallel and Distributed Computing*, , pp. 172, 1–11, 2023.

[66] R. e. a. Wadhwa, "Topological anomaly detection in financial networks," *IEEE Transactions on Big Data*, vol. 9(2), pp. 123-134, 2023.

[67] A. & Z. Y. Brown, "SHAP-Mapper: Interpretable topological feature attribution," *Machine Learning Journal*, vol. 112(5), pp. 1023-1045, 2023.

[68] T. e. a. Chen, "TopoAE: Topology-preserving autoencoders with Wasserstein loss," in *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*, 2023.

[69] M. e. a. Moor, "Molecular graph generation via persistent homology regularization," *Nature Machine Intelligence*, vol. 5(4), pp. 345-354, 2023.

[70] E. e. a. Solomon, "TopoGAN: Adversarial training with persistence diagrams," *Advances in Neural Information Processing Systems*, pp. 36, 12345–12356, 2023.

[71] P. e. a. Bendich, "Dynamic topological losses for sequential data," *Journal of Machine Learning Research*, vol. 24(1), p. 789-812, 2023.

[72] Y. e. a. Hiraoka, "TopoCL: Contrastive learning with PH-based invariants," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)*.

[73] M. e. a. Carrière, "Differentiable approximations of persistent homology," *SIAM Journal on Mathematics of Data Science*, vol. 5(2), pp. 234-256, 2023.

[74] C. K. R. N. M. & U. A. Hofer, "PersLay: A neural network layer for persistence diagrams and new graph topological signatures," in *Proceedings of the 37th International Conference on Machine Learning (ICML 2020)*, 2020.

[75] Y. L. Y. & G. H. Chen, "Topology-aware graph pooling networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43(12), pp. 4316-4329, 2021.

[76] P. & D. P. Bubenik, "A persistence landscapes toolbox for topological statistics," *Journal of Symbolic Computation*, vol. 78, pp. 91-114, 2017.

[77] M. & B. A. J. Carrière, "Differentiable mapper: A neural network interpretability tool," *Nature Machine Intelligence*, vol. 5(4), pp. 345-354, 2023.

[78] B. T. C. J. & B. P. Rieck, "Topological regularization for graph neural networks," in *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*, 2023.

[79] J. V. T. & C. A. Leygonie, "Topological attribution maps for explainable AI," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

[80] J. e. a. Leygonie, "TopoCAM: Topological class activation maps," *Medical Image Analysis*, 2023.

[81] J. R. B. N. O. I. Z. V. A. S. J. A. & K. A. P. Clough, "A topological loss function for deep-learning based image segmentation using persistent homology," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44(12), pp. 8766-8778., 2020.

[82] P. e. a. Bendich, "Persistence LIME for time-series explanations," *ACM Transactions on Intelligent Systems*, 2023.

[83] A. H. H. A. T. G. & P. L. Spannaus, "Topological Interpretability for Deep Learning," in *In Proceedings of the Platform for Advanced Scientific Computing Conference (pp. 1-11)*, 2024.

[84] Y. S. A. H. & G. V. Verma, "Topological neural networks go persistent, equivariant, and continuous," *arXiv preprint arXiv:2406.03164*, 2024.

[85] Q. e. a. Zhao, "ersGNN: Topological signatures for graph representation learning," in *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*, 2023.

[86] N. I. Y. & Y. K. Nishikawa, "Adaptive topological feature via persistent homology: Filtration learning for point clouds," *Advances in Neural Information Processing Systems*, vol. 36, pp. 9131-9143, 2023.

[87] T. K. B. M. B. M. I. N. K. V. & V. V. B. D. Songdechakraiut, "Wasserstein distance-preserving vector space of persistent homology," in *In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 277-286)*, 2023.

[88] M. & L. P. Sun, "Graph to graph: a topology aware approach for graph structures learning and generation," in *In The 22nd international conference on artificial intelligence and statistics (pp. 2946-2955)*. PMLR., 2019.

[89] Y. L. F. X. S. S. S. C. L. & W. Z. Ren, "Dynamic ECG signal quality evaluation based on persistent homology and GoogLeNet method," *Frontiers in Neuroscience*, vol. 1153386, p. 17, 2023.

[90] J. e. a. Garland, "Wearable-Adaptive PH for Respiratory Monitoring," *IEEE Transactions on Biomedical Engineering*, 2023.

[91] H. X. S. A. Q. Z. X. S. W. & Z. X. Guo, "Empirical study of financial crises based on topological data analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 124956, p. 558, 2020.

[92] D. & S. A. Lee, "Visibility graphs, persistent homology, and rolling element bearing fault detection," in *IEEE International Conference on Prognostics and Health Management (ICPHM) (pp. 1-5)*. IEEE., 2022.

[93] Z. L. F. S. S. X. S. P. F. W. L. .. & X. Z. Wang, "Automatic epileptic seizure detection based on persistent homology," *Frontiers in physiology*, vol. 227952, p. 14, 2023.

[94] T. C. K. Y. J. N. I. L. L. H. L. .. & Z. L. Zhao, "3D graph anatomy geometry-integrated network for pancreatic mass segmentation, diagnosis, and quantitative patient management," in *Proceedings of the*

IEEE/CVF conference on computer vision and pattern recognition (pp. 13743-13752), 2021.

[95] J. H. Z. W. H. & X. L. Zhu, "Topology optimization in aircraft and aerospace structures design.,," *Archives of computational methods in engineering*, Vols. 595-622, p. 23, 2016.

[96] X. Y. Y. H. H. M. L. B. & W. H. Wang, " Cross-modal contrastive learning network for few-shot action recognition.,," *IEEE Transactions on Image Processing*, Vols. 1257-1271, p. 33, 2024.

[97] A. F. S. A. H. W. & D. O. Kherchouche, "Detection of adversarial examples in deep neural networks with natural scene statistics.,," in *International Joint Conference on Neural Networks (IJCNN)* (pp. 1-7). IEEE, 2020.

[98] T. & S. P. Gebhart, " Adversary detection in neural networks via persistent homology.,," *arXiv preprint arXiv:1711.10056*, 2017.

[99] V. Subramanian, *Understanding Audio Deep Learning Classifiers Through Adversarial Attacks and Interpretability*, Doctoral dissertation, Queen Mary University of London, 2023.

[100] T. & M. G. Yerriswamy, "Signature-based traffic classification for DDoS attack detection and analysis of mitigation for DDoS attacks using programmable commodity switches.,," *International Journal of Performativity Engineering*, vol. 18(7), p. 529, 2022.

[101] T. E. M. N. A. A. K. & Z. A. Bilot, "Graph neural networks for intrusion detection: A survey.,," 2023.

[102] H. V. A. & M. I. S. Fouad, "A Dynamical Systems Approach for Detecting Cyberattacks in Software-Defined Networks.,," in *MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM)* (pp. 734-739). IEEE., 2024.

[103] H. & H. M. Tariq, "Topology-agnostic detection of temporal money laundering flows in billion-scale transactions.,," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 402-419), 2023.

[104] Y. Z. J. Z. M. & Y. J. Yue, " An Abnormal Account Identification Method by Topology Feature Analysis for Blockchain-Based Transaction Network.,," *Electronics*, vol. 1416, p. 13(8), 2024.

[105] S. F. S. L. Y. X. H. Z. X. & X. M. Fan, "Smart contract scams detection with topological data analysis on account interaction.,," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 468-477), 2022.

[106] P. M. J. S. & M. E. Bendich, "Dynamic Mapper for real-time financial surveillance.,," *SIAM Journal on Financial Mathematics*, vol. 14(10), pp. 1-24, 2023.

[107] T. K. & P. R. Smith, "Theoretical foundations of Mapper for collusive fraud detection.,," *Annals of Applied Statistics*, vol. 17(2), pp. 789-815, 2023.

[108] Y. D. S. V. & M. D. Chen, "TopoOpt: Topology-aware neural architecture search via persistent homology.,," in *Proceedings of the 40th International Conference on Machine Learning*, 202, 4500-4515., 2023.

[109] A. I. S. & S. G. Carpio, "Bayesian approach to inverse scattering with topological priors.,," *Inverse Problems*, vol. 105001, p. 36(10), 2020.

[110] C. Z. J. L. L. J. L. F. & Y. S. Wang, "Automatic graph topology-aware transformer.,," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[111] B. S. G. & M. V. Ameneiro, "Quantum persistent homology for time series.,," in *IEEE/ACM 7th Symposium on Edge Computing (SEC)* (pp. 387-392). IEEE., 2022.

[112] S. G. S. & Z. P. Lloyd, " Quantum algorithms for topological and geometric analysis of data.,," *Nature communications*, vol. 10138, p. 7(1), 2016.

[113] J. D. M. M. E. & W. J. D. Biamonte, "Quantum-assisted topology computation of small molecules.,," *PRX Quantum*, vol. 020327, p. 4(2), 2023.

[114] S. & K. N. Gunn, "Review of a quantum algorithm for Betti numbers.,," *arXiv preprint arXiv:1906.07673*, 2019.

[115] A. & C. B. K. Das, "Quantum annealing and related optimization methods.,," *Springer Science & Business Media*, vol. 679, 2005.

[116] Z. C. F. I. R. L. B. M. W. G. & R. A. Bian, "Discrete optimization using quantum annealing on sparse Ising models.,," *Frontiers in Physics*, 2, p. 56, 2014.

[117] B. H. Z. & L. H. Zhang, " A comprehensive review of deep neural network interpretation using topological data analysis.,," *Neurocomputing*, p. 128513, 2024.

[118] P. B. Q. T. N. N. T. K. R. Y. P. S. & V. B. Pham, "Topological data analysis in graph neural networks: surveys and perspectives.,," *IEEE Transactions on Neural Networks and Learning Systems*, 2025.

[119] S. R. & L. M. Choi, "Transformer architecture and attention mechanisms in genome data analysis: a comprehensive review.,," *Biology*, vol. 1033, p. 12(7), 2023.

[120] S. Y. B. Z. W. & C. Y. Zhang, "Topology aware deep learning for wireless network optimization.,," *IEEE Transactions on Wireless Communications*, Vols. 9791-9805, p. 21(11), 2022.

[121] S. K. H. A. G. A. D. J. R. A. S. & B. I. Iannucci, "A comparison of graph-based synthetic data generators for benchmarking next-generation intrusion detection systems.,," in *IEEE International Conference on Cluster Computing (CLUSTER)* (pp. 278-289). IEEE., 2017.

[122] Y. F. C. M. H. Q. A. L. T. J. J. C. G. E. & E. B. J. Singh, " Topological data analysis in medical imaging: current state of the art.,," *Insights into Imaging*, vol. 58, p. 14(1), 2023.

[123] S. B. J. K. I. T. G. M. A. S. & O. B. Bonner, "Evaluating the quality of graph embeddings via topological feature reconstruction.,," in *IEEE International Conference on Big Data (Big Data)* (pp. 2691-2700). IEEE., 2017.

[124] E. V. B. S. E. L. A. S. J. G. & W. R. R. Somasundaram, "Benchmarking r packages for calculation of persistent homology.,," *The R journal*, vol. 184, p. 13(1), 2021.