



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Topological analysis of decision boundaries

Practical work

Maxim Mikhaylov

January 30, 2025

Supervisors: Prof. Dr. B. Gärtner, Dr. P. Schnider

Department of Computer Science, ETH Zürich

Abstract

This project investigates the application of topological data analysis (TDA) to the study of decision boundaries in machine learning classifiers in general, and how these boundaries evolve during training in particular. Decision boundaries partition the input space into regions corresponding to different class labels, and understanding their topology can provide deep insights into model behavior, particularly regarding overfitting and underfitting. Building upon the Labeled Vietoris-Rips complex framework, which was previously limited to binary classification, we extend it to handle multiclass classification problems and propose topological metrics that can be used to quantify the fitness of a classifier.

Contents

Contents	iii
1 Introduction	1
1.1 Background	1
1.2 Objectives	1
2 Related work and theoretical background	3
2.1 Related work	3
2.2 Theoretical background	4
2.2.1 Persistent homology	4
2.2.2 Labeled Čech complex	4
2.2.3 Labeled Vietoris-Rips complex	5
2.2.4 Dowker complex	6
3 Methodology	9
3.1 Datasets and models	9
3.2 Metrics and evaluation	10
3.3 Simplicial complex	10
4 Justification	13
4.1 Synthetic 2D and 3D data experiments	13
4.2 Sampling stability on MNIST	13
5 Results and discussion	15
5.1 Results on MNIST	15
5.2 Results on FashionMNIST	15
5.3 Results on multiclass classification	15
5.4 Comparison of CNN and MLP	15
A Dummy Appendix	17

CONTENTS

Bibliography	19
---------------------	-----------

Chapter 1

Introduction

1.1 Background

- ML is cool
- TDA is cool
- Decision boundaries are cool
- Why the intersection of these three may be cool

1.2 Objectives

Yoink from the proposal.

- Explore homological changes in decision boundaries during model training.
- Investigate connections between topological features and overfitting/underfitting.
- Extend the Labeled Vietoris-Rips complex to multiclass classification.

Chapter 2

Related work and theoretical background

2.1 Related work

Topological data analysis (TDA) has been applied to neural network research using a wide range of methods. As outlined in [2], these approaches can be broadly categorized according to which aspect of neural network they analyze:

1. *Structure of neural networks*, such as in [6], where two special homology groups are computed for graphs that represent feed-forward neural networks.
2. *Decision regions and boundaries*. In [18], the labeled Čech and Vietoris-Rips complexes are introduced to capture homology of neural network decision boundaries. This approach is augmented with active learning in [12] to improve sampling efficiency. Several other studies employ the *graph-based topological data analysis* (GTDA) algorithm [14], an extension of the Mapper algorithm [20] that handles graph inputs and generates Reeb networks.
3. *Activations and weights*. Methods in this category apply either the Mapper algorithm [5, 15, 10], or persistent homology [7, 11, 19] to neural network weights or activations.
4. *Training dynamics and loss functions*. In [16] the number of connected components and local valleys of a convex optimization target are studied in the context of fully connected feed-forward neural networks. In [24] and [4], the fractal dimensions of weight trajectories of neural networks during training are studied.

This research bridges the categories 2 and 4, presenting a novel investigation that combines topological analysis of decision boundaries with neural

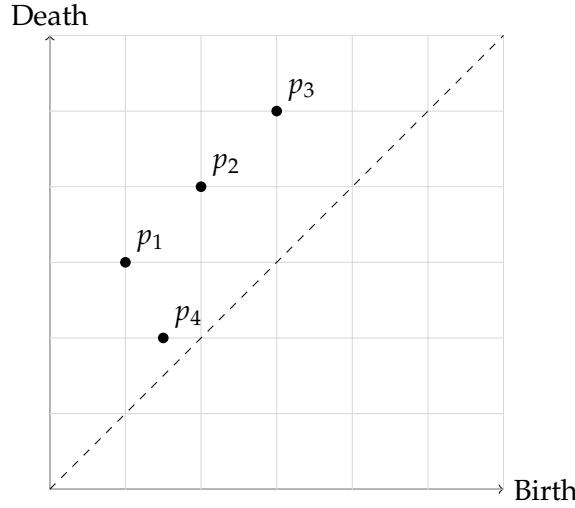


Figure 2.1: An example of a persistent diagram.

network training dynamics.

2.2 Theoretical background

2.2.1 Persistent homology

Persistent homology serves as a fundamental tool in TDA that tracks how topological features (connected components, holes) appear and disappear as a parameter is varied. Given a filtered simplicial complex $(K_\varepsilon)_{\varepsilon \geq 0}$, persistent homology computes pairs $(\varepsilon_b, \varepsilon_d)$ where a feature appears at ε_b (birth) and disappears at ε_d (death).

These birth–death pairs can be visualized in a *persistence diagram*, where each pair is shown as a point in the Euclidean plane. The distance of a point from the diagonal indicates the feature’s persistence, which is often interpreted as a measure of its significance. Figure 2.1 shows an example of a persistence diagram.

2.2.2 Labeled Čech complex

Definition 2.1 (Labeled Čech (LČ) complex, [18]) *Given a set of points X , a reference set Y , and parameters ε and γ , the labeled Čech complex contains an n -simplex formed by points $x_0, \dots, x_n \in X$ if and only if:*

1. $\bigcap_{i=0}^n B_\varepsilon(x_i) \neq \emptyset$
2. For each $i \in (0, \dots, n)$, there exists $y \in Y$ such that $\|x_i - y\| \leq \gamma$.

This simplicial complex is denoted as $L\check{C}_{X,Y}^{\varepsilon,\gamma}$.

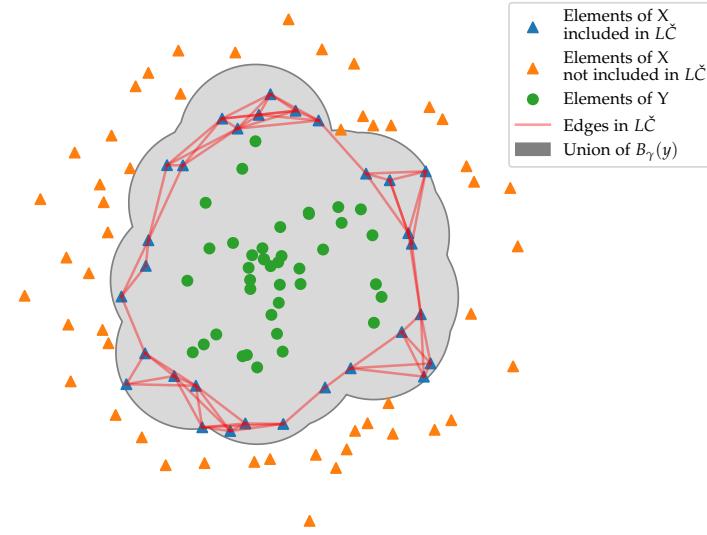


Figure 2.2: An example of the labeled Čech complex. Only 0- and 1-simplices are shown for clarity.

In [18], the *labeled Čech filtration* is obtained by varying ε , while keeping γ fixed.

In binary classification, X is set of points of one class, while Y is set of points of the other class. Then, $L\check{C}$ captures simplices on one side of the boundary. This can be seen in Figure 2.2.

[18] gives conditions that provide a probabilistic guarantee that the $L\check{C}$ complex recovers the homology of the decision boundary.

2.2.3 Labeled Vietoris-Rips complex

Definition 2.2 (Labeled Vietoris-Rips (LVR) complex, [18]) Given a set X with the set of associated labels c and a parameter ε , let the bipartite graph G_ε be a graph with X as its vertex set, and an edge $x_i - x_j$ be included if $\|x_i - x_j\| \leq \varepsilon$ and $c_i \neq c_j$. This adds all short enough edges between points in different classes. After adding these edges, all 2-hop neighbors are connected to induce simplices of order 2. The labeled Vietoris-Rips complex is built using the standard Vietoris-Rips induction on the resulting graph, i.e. an n -simplex is added if all of its $(n-1)$ -dimensional faces are included.

The *labeled Vietoris-Rips filtration* can be obtained by varying ε .

Analogous to the relationship between the standard Čech complex [1] and the standard Vietoris-Rips complex [21], the labeled Čech complex is more computationally expensive, while the labeled Vietoris-Rips complex is more

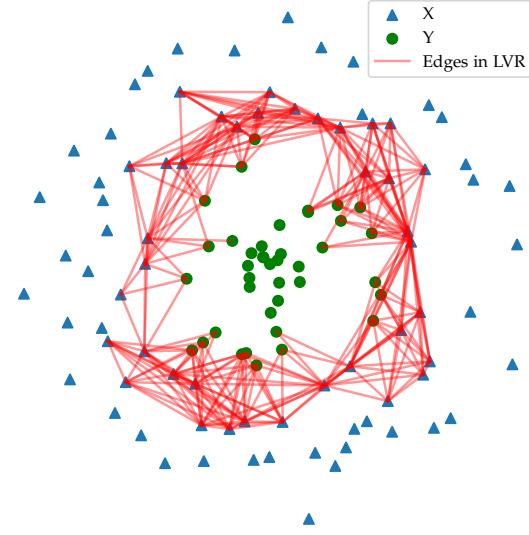


Figure 2.3: An example of the labeled Vietoris-Rips complex. Only 0- and 1-simplices are shown for clarity.

tractable but lacks known recovery guarantees. While it is known that the VR and Čech complexes are log-interleaved through inclusions:

$$\check{C}_X^\varepsilon \subseteq VR_X^\varepsilon \subseteq \check{C}_X^{2\varepsilon}, \quad (2.1)$$

it is not clear how the labeled versions compare.

Definition 2.3 (Locally Scaled LVR (LS-LVR) complex, [18]) *The LS-LVR complex is constructed similarly to the LVR complex, but with the edge $x_i - x_j$ included in G_ε if $\|x_i - x_j\| \leq \varepsilon \sqrt{\rho_i, \rho_j}$, where ρ_i is the distance from x_i to its k -th closest neighbor, with k being a fixed parameter.*

The LS-LVR complex aims to recover homology groups of datasets with non-uniform density more accurately compared to the LVR complex.

2.2.4 Dowker complex

Definition 2.4 *Let X, Y be sets. The Dowker complex $D_{X,Y}^\varepsilon$ at parameter value ε is a simplicial complex with X as its vertex set, including the n -simplex x_0, \dots, x_n if and only there exists a $y \in Y$ such that $\|x_i - y\| \leq \varepsilon$ for all i .*

A *Dowker filtration* can be obtained by varying ε .

The Dowker complex captures simplices on one side of the decision boundary, similarly to the labeled Čech complex. Indeed, by definition,

$$LC_{X,Y}^{\varepsilon,\gamma} = \check{C}_X^\varepsilon \cap D_{X,Y}^\gamma. \quad (2.2)$$

2.2. Theoretical background

Although the roles of X and Y may appear arbitrary, the following theorem shows that they can be interchanged without changing the homology groups, up to isomorphism:

Theorem 2.5 (Dowker, [9]) *For any Dowker complex constructed from two sets X and Y , interchanging the roles of X and Y and keeping the same value of ε results in homology groups that are isomorphic to the original ones:*

$$H_k(\mathcal{D}_{X,Y}^\varepsilon) \cong H_k(\mathcal{D}_{Y,X}^\varepsilon). \quad (2.3)$$

Chapter 3

Methodology

3.1 Datasets and models

For evaluation we have used the MNIST [8] and FashionMNIST [22] datasets with the predefined training and testing splits. We have also used synthetic 2D and 3D binary classification datasets for more easily interpretable experiments. They are shown in Figures 3.1 and ??, with each dataset consisting of 10000 points, uniformly sampled from $[-1, 1]^2$ and split into training and testing sets with a ratio of 0.8 : 0.2.

TODO: Maybe drop 3D

TODO: Describe model for 2D/3D if we need that

For MNIST and FashionMNIST, we have used:

1. A simple feed-forward neural network with layer sizes 784, 512, 128, 16, 10, ReLU activations and a softmax output layer.
2. A convolutional neural network with a modification of the architecture given in <https://github.com/pytorch/examples/tree/main/mnist>. The modifications are:
 - a) Addition of the `scale` parameter, which allows us to change the neural network width. It is a multiplicative factor for the number of channels in the convolutional layers and number of neurons in the fully connected layers.
 - b) Addition of the `extra_cnn` parameter, which adds convolutional layers with $64 \cdot \text{scale}$ channels, 3×3 kernel size, 1 pixel padding and stride, as well as a ReLU activation.
 - c) Additon of the `extra_linear` parameter, which similarly adds a fully connected layer with $128 \cdot \text{scale}$ neurons, as well as a ReLU activation.

3. METHODOLOGY

Figure 3.1: Synthetic 2D datasets

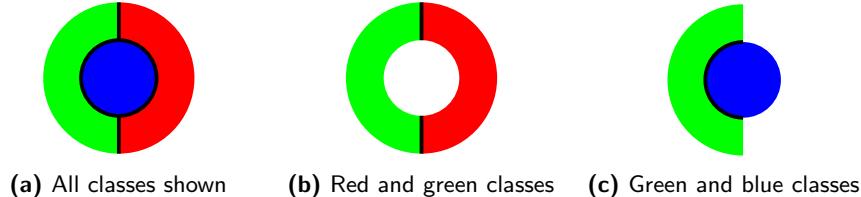


Figure 3.2: A 2D example that demonstrates the loss of topological information when binarizing a multiclass dataset. Red, green and blue denote different classes, the black line is the decision boundary.

3.2 Metrics and evaluation

TODO: Explain how and why we use those

Definition 3.1 For $q \geq 1$, the Wasserstein distance between two persistence diagrams $\text{Dgm}_p(\mathcal{F})$ and $\text{Dgm}_p(\mathcal{G})$ is defined as

$$d_{W,q}(\text{Dgm}_p(\mathcal{F}), \text{Dgm}_p(\mathcal{G})) = \left[\inf_{\pi \in \Pi} \left(\sum_{x \in \text{Dgm}_p(\mathcal{F})} \|x - \pi(x)\|_\infty^q \right) \right]^{1/q} \quad (3.1)$$

Definition 3.2 The total bar length of a persistence diagram $\text{Dgm}_p(\mathcal{F})$ is defined as

$$\sum_{(b,d) \in \text{Dgm}_p(\mathcal{F})} d - b \quad (3.2)$$

3.3 Simplicial complex

We use the Labeled Vietoris-Rips complex, as presented in [18]. To improve computational performance, we used the `ripser++` [23] and the `giotto-ph` [17] packages for the Vietoris-Rips persistence barcode computation. Both of these methods have shown lower runtime than the original `ripser` [3] package. However, due to GPU memory limitations, we use `giotto-ph` in all experiments.

In previous research [18, 13]**TODO: maybe more**, classification datasets with $n > 2$ classes were split into $\binom{n}{2}$ binary classification problems. This process loses topological information, as shown in an example dataset in Figure 3.2. The decision boundary for all three classes (3.2a) has a non-trivial first homology group H_1 . This is not the case for any binary classification problem for this dataset (3.2b, 3.2c). Moreover, the decision boundary is disconnected for the red and green subdataset, but connected for the full dataset.

3.3. Simplicial complex

To avoid this information loss, we propose a generalization of the Labeled Vietoris-Rips complex to multiclass classification. For a labeled discrete sample $\{(z_1, c_1), \dots, (z_n, c_n)\}$, where $z \in \mathbb{R}^d$ is a data point and $c \in \{1, \dots, n\}$ is its class, we define a bipartite graph G_θ by adding an edge between z_i and z_j iff $c_i \neq c_j$ and $\|z_i - z_j\| \leq \theta$. **TODO: Say how this is a no-op compared to LVR** **TODO: Say we don't generalize LC**

To combat issues shown in ??, we propose *circumcircle filtering* (CC). CC removes an edge between points A and B from the simplicial complex if exists a point C such that $|AB|^2 > (|AC|^2 + |BC|^2)\theta$, where $\theta \in [0, 2]$ is a parameter. Intuitively, this approach aims to make a simplicial complex more similar to the Alpha complex **TODO: cite** by removing some (but not all) non-Delaunay edges. **TODO: Explain this better**

Alternatively, we propose usage of the *Dowker complex* to combat the same issues. However, computing it requires $\mathcal{O}(n^3 \cdot d)$ runtime and $\mathcal{O}(n^3)$ memory for n points in \mathbb{R}^d , which is prohibitively expensive for the MNIST and FashionMNIST datasets. Additionally, the Dowker complex is not applicable in the multiclass setting. Nonetheless, we demonstrate its effectiveness on 2D data in ??.

TODO: Confirm complexity

Chapter 4

Justification

4.1 Synthetic 2D and 3D data experiments

- Show what isn't captured correctly
- Show that CC filtering or the Dowker complex can help

In **TODO: ref**, we can see that the LVR complex includes edges and triangles that cross the decision boundary multiple times. This is a problem, because these edges and triangles

4.2 Sampling stability on MNIST

Show that with the default parameters, the persistence diagrams change too much under sampling. Show it's stable under $N=10$.

Chapter 5

Results and discussion

5.1 Results on MNIST

- Accuracy of the model correlates with the Wasserstein distance and the total bar length
- “Elbow” behavior — first epochs have high changes in topological metrics, then it’s flatter. Probably easier to show for Wasserstein distance, as the total bar length may be both high and low. Alternatively, show the difference between total bar length and GT total bar length.
- Underfitting and overfitting leads to worse topological metrics
- More accurate models have better topological metrics
- Size of the model doesn’t impact the topological metrics (as long as the accuracy is the same for them)

5.2 Results on FashionMNIST

Hopefully, just show it’s the same.

5.3 Results on multiclass classification

Show it’s the same (?)

5.4 Comparison of CNN and MLP

If there’s anything interesting

Appendix A

Dummy Appendix

You can defer lengthy calculations that would otherwise only interrupt the flow of your thesis to an appendix.

Bibliography

- [1] Paul Alexandroff. Über kombinatorische eigenschaften allgemeiner kurven. *Mathematische Annalen*, 96(1):512–554, Dec 1927.
- [2] Rubén Ballester, Carles Casacuberta, and Sergio Escalera. Topological data analysis for neural network analysis: A comprehensive survey, 2024.
- [3] Ulrich Bauer. Ripser: efficient computation of Vietoris-Rips persistence barcodes. *J. Appl. Comput. Topol.*, 5(3):391–423, 2021.
- [4] Tolga Birdal, Aaron Lou, Leonidas Guibas, and Umut Şimşekli. Intrinsic dimension, persistent homology and generalization in neural networks, 2021.
- [5] Rickard Brüel Gabrielsson and Gunnar Carlsson. Exposition and interpretation of the topology of neural networks. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 1069–1076, 2019.
- [6] Samir Chowdhury, Thomas Gebhart, Steve Huntsman, and Matvey Yutin. Path homologies of deep feedforward networks. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, page 1077–1082. IEEE, December 2019.
- [7] Ciprian A. Corneanu, Meysam Madadi, Sergio Escalera, and Aleix M. Martinez. What does it mean to learn in deep networks? and, how does one detect adversarial attacks? In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4752–4761, 2019.
- [8] Li Deng. The mnist database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, 29(6):141–142, 2012.

BIBLIOGRAPHY

- [9] C. H. Dowker. Homology groups of relations. *Annals of Mathematics*, 56(1):84–95, 1952.
- [10] Maxime Gabella. Topology of learning in feedforward neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8):3588–3592, 2021.
- [11] Thomas Gebhart, Paul Schrater, and Alan Hylton. Characterizing the shape of activation space in deep neural networks, 2019.
- [12] Weizhi Li, Gautam Dasarathy, Karthikeyan Natesan Ramamurthy, and Visar Berisha. Finding the homology of decision boundaries with active learning. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 8355–8365. Curran Associates, Inc., 2020.
- [13] Weizhi Li, Gautam Dasarathy, Karthikeyan Natesan Ramamurthy, and Visar Berisha. Finding the homology of decision boundaries with active learning. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 8355–8365. Curran Associates, Inc., 2020.
- [14] Meng Liu, Tamal K. Dey, and David F. Gleich. Topological structure of complex predictions. *Nature Machine Intelligence*, 5(12):1382–1389, Dec 2023.
- [15] Ephy Love, Benjamin Filippenko, Vasileios Maroulas, and Gunnar E. Carlsson. Topological convolutional neural networks. In *TDA & Beyond*, 2020.
- [16] Quynh Nguyen. On connected sublevel sets in deep learning, 2019.
- [17] Julián Burella Pérez, Sydney Hauke, Umberto Lupo, Matteo Caorsi, and Alberto Dassatti. giotto-ph: A python library for high-performance computation of persistent homology of vietoris–rips filtrations, 2021.
- [18] Karthikeyan Natesan Ramamurthy, Kush R. Varshney, and Krishnan Mody. Topological data analysis of decision boundaries with application to model selection, 2018.
- [19] Rieck, Bastian Alexander, Togninalli, Matteo, Bock, Christian, Moor, Michael, Horn, Max, Gumsch, Thomas, and Borgwardt, Karsten. Neural persistence: A complexity measure for deep neural networks using algebraic topology. 2023.

Bibliography

- [20] Gurjeet Singh, Facundo Memoli, and Gunnar Carlsson. Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition . In M. Botsch, R. Pajarola, B. Chen, and M. Zwicker, editors, *Eurographics Symposium on Point-Based Graphics*. The Eurographics Association, 2007.
- [21] L. Vietoris. Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. *Mathematische Annalen*, 97(1):454–472, Dec 1927.
- [22] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms, 2017.
- [23] Simon Zhang, Mengbai Xiao, and Hao Wang. Gpu-accelerated computation of Vietoris-Rips persistence barcodes. In *36th International Symposium on Computational Geometry (SoCG 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.
- [24] Umut Şimşekli, Ozan Sener, George Deligiannidis, and Murat A Erdogan. Hausdorff dimension, heavy tails, and generalization in neural networks*. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124014, December 2021.

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):

First name(s):

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the '[Citation etiquette](#)' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

Signature(s)

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.