The Shape of Digits

A Bayesian Topological Data Analytic Approach to Classification of Handwritten Digits

Thomas Reinke Theophilus A. Bediako Daniel Lim

August 16, 2025

Table of contents

1 Abstract

This paper ...[1]

2 Introduction

The motivation for this work "Data has shape, shape has meaning, and meaning brings value." - Gunnar Carlsson

3 Background and Related Work

about mnist MNIST Overview

- A Bayesian Framework for Persistent Homology(Maroulas, Nasrin, and Oballe 2020)
- Classification based on Topological Data Analysis(Kindelan et al. 2021)
- A topological data analysis based classification method for multiple measurements(Riihimäki et al. 2019)
- Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival(Nicolau, Levine, and Carlsson 2011)
- A Topological "Reading" Lesson: Classification of MNIST using TDA(Garin and Tauzin 2019)
- Classification and Evaluation of Machine Learning Algorithms on the MNIST Dataset (Yeboah 2025)

4 Methodology

4.1 Traditional Machine Learning

4.1.1 Neural Networks

Dropout

Ridge

Lasso

Multinomial Logistic Regression as ML

[1] Example footnote

4.2 Bayesian TDA Methodology

latex flowchart

4.2.1 Cubical Complexes & Persistent Homology

4.2.2 Marked PPP

4.2.3 Bayes Update & Guassian representation

4.3 TDA + ML

4.3.1 Filtering

5 Experiments

5.1 Data Summary

5.1.1 EDA

5.2 NN dropout

5.3 NN Ridge

5.4 NN Lasso

5.5 Multinomial Logistic

5.6 Bayes TDA

5.7 TDA + ML

6 Discussion and Analysis

Comparing the models

7 Conclusion

8 References

Garin, Adélie, and Guillaume Tauzin. 2019. "A Topological "Reading" Lesson: Classification of MNIST Using TDA." *CoRR* abs/1910.08345. http://arxiv.org/abs/1910.08345.

Kindelan, Rolando, José Frías, Mauricio Cerda, and Nancy Hitschfeld. 2021. "Classification Based on Topological Data Analysis." *CoRR* abs/2102.03709. https://arxiv.org/abs/2102.03709.

Maroulas, Vasileios, Farzana Nasrin, and Christopher Oballe. 2020. "A Bayesian Framework for Persistent Homology." *SIAM J. Math. Data Sci.* 2 (1): 48–74.

Nicolau, Monica, Arnold J. Levine, and Gunnar Carlsson. 2011. "Topology Based Data Analysis Identifies a Subgroup of Breast Cancers with a Unique Mutational Profile and Excellent Survival." *Proceedings of the National Academy of Sciences* 108 (17): 7265–70. https://doi.org/10.1073/pnas.1102826108.

Riihimäki, Henri, Wojciech Chachólski, Jakob Theorell, Jan Hillert, and Ryan Ramanujam. 2019. "A Topological Data Analysis Based Classification Method for Multiple Measurements." *CoRR* abs/1904.02971. http://arxiv.org/abs/1904.02971.

Yeboah, Felix. 2025. "Classification and Evaluation of Machine Learning Algorithms on the MNIST Dataset."