

# Unveiling Topological Structures from Language: A Comprehensive Survey of Topological Data Analysis Applications in NLP

**Adaku Uchendu**  
MIT Lincoln Laboratory,  
Lexington, MA, USA  
adaku.uchendu@ll.mit.edu

**Thai Le**  
Indiana University,  
Bloomington, IN, USA  
tle@iu.edu

## Abstract

The surge of data available on the internet has led to the adoption of various computational methods to analyze and extract valuable insights from this wealth of information. Among these, the field of Machine Learning (ML) has thrived by leveraging data to extract meaningful insights. However, ML techniques face notable challenges when dealing with real-world data, often due to issues of imbalance, noise, insufficient labeling, and high dimensionality. To address these limitations, some researchers advocate for the adoption of Topological Data Analysis (TDA), a statistical approach that discerningly captures the intrinsic shape of data despite noise. Despite its potential, TDA has not gained as much traction within the Natural Language Processing (NLP) domain compared to structurally distinct areas like computer vision. Nevertheless, a dedicated community of researchers has been exploring the application of TDA in NLP, yielding **95 papers** we comprehensively survey in this paper. Our findings categorize these efforts into theoretical and non-theoretical approaches. Theoretical approaches aim to explain linguistic phenomena from a topological viewpoint, while non-theoretical approaches merge TDA with ML features, utilizing diverse numerical representation techniques. We conclude by exploring the challenges and unresolved questions that persist in this niche field. Resources and a list of papers on this topic can be found at: <https://github.com/AdaUchendu/AwesomeTDA4NLP><sup>1</sup>.

<sup>1</sup>DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. This material is based upon work supported by the Department of the Air Force under Air Force Contract No. FA8702-15-D-0001 or FA8702-25-D-B002. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of the Air Force. © 2025 Massachusetts Institute of Technology. Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work

## 1 Introduction

Proliferation of the Internet has given rise to the generation of massive amounts of data. These massive amounts of data when processed can solve many crucial issues plaguing our current society. Due to this well-established notion among stake-holding institutions, the Machine Learning (ML) field has been thriving as a tool that extracts trends and solutions to non-trivial problems. However, real-world data tends to be noisy, heterogeneous, imbalanced, have missing labels, contain high-dimensionality, etc., often making the adoption of ML techniques to such datasets non-trivial. Therefore, to extract meaningful findings from data, specifically real-world data, clever techniques that extract additional features, while *preserving the structure of the data need to be employed*. To that end, a small niche community for *Topological Data Analysis (TDA) applications in NLP* has emerged. Being promised as a technique that can extract and analyze the shape/topology of data, TDA has great potential in mitigating such issues witnessed in real-world data.

TDA is a “collection of powerful tools that can quantify shape and structure in data”<sup>2</sup> and is inspired by the algebra and geometry mathematical fields. The benefits of TDA are vast, including the ability to extract additional features, typically not captured by other feature extraction techniques (Uchendu et al., 2024; Michel et al., 2017). We call these features that TDA extracts - *topological features*. Unsurprisingly, since TDA is used to capture topological features, it has been applied to many tasks where data has distinct graphical structures (Papamarkou et al., 2024; Hensel et al., 2021). These include tasks that have obvious graph-like structures, such as protein classification (Dey

other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.

<sup>2</sup><https://www.indicative.com/resource/topological-data-analysis/>

Alternative Method	Description
Principal Component Analysis (PCA)	Reduces dimensionality while preserving variance using orthogonal transformations.
t-Distributed Stochastic Neighbor Embedding (t-SNE)	Projects high-dimensional data into 2D or 3D space while preserving local relationships.
Uniform Manifold Approximation and Projection (UMAP)	Similar to t-SNE but faster and better at preserving global structure.
Multidimensional Scaling (MDS)	Projects high-dimensional data into lower dimensions by preserving pairwise distances.
Autoencoders	A type of Neural network that learns compressed representations of data via encoding-decoding processes.
Manifold Learning (Isomap, LLE, etc.)	Captures intrinsic structures in high-dimensional data through graph-based techniques.
Graph-based Learning (Spectral Clustering, GNNs)	Uses graphs to model relationships and structure within data.
Clustering Methods (DBSCAN, K-Means, HDBSCAN)	Groups similar data points based on distance and density.
Kernel Methods (SVM, Kernel PCA)	Uses non-linear mappings to extract complex structures in data.
Geometric Deep Learning	Uses neural networks on non-Euclidean spaces like graphs and manifolds.
Geometric techniques (Delaunay triangulation, Convex hull)	Uses geometrical techniques to extract characteristics of data by analyzing spacial boundaries

Table 1: Alternatives to Topological Data Analysis in NLP

and Mandal, 2018; Lamine et al., 2023; Valeriani et al., 2024) and drug discovery (Alagappan et al., 2016); to those that are not so obvious such as diabetes classification (Wamil et al., 2023; Skaf and Laubenbacher, 2022), image classification (Horn et al., 2022; Trofimov et al., 2023), time series analysis (Petri and Leita, 2020; Tymochko et al., 2021; Gholizadeh and Zadrozny, 2018), and speech processing (Tulchinskii et al., 2023). However, since the shape of a text is not apparent, it has not gained as much attention in Natural Language Processing (NLP) as it has done in the Computer Vision (Horn et al., 2022; Trofimov et al., 2023) and Medical (Singh et al., 2023; Nielson et al., 2015) domains. Still, several researchers have found ways to extract additional features using TDA that the other typical numerical representation techniques in text such as TF-IDF, Word2Vec embeddings, BERT embeddings, etc have not captured. Figure 1 shows that there has been a gradual acceleration in the number of published works in TDA for NLP applications, and we project this trend will continue in the future.

TDA aims to answer the main question - *what is the true shape of a data?* We survey 95 papers that have attempted to find an answer through various approaches. We categorize all observed approaches that incorporate TDA in NLP into two categories, namely *theoretical* (Karlgrén et al., 2014; Port et al., 2018) and *non-theoretical* (Zhu, 2013; Doshi and Zadrozny, 2018). *Theoretical* approaches involve using TDA to explain linguistic phenomena by

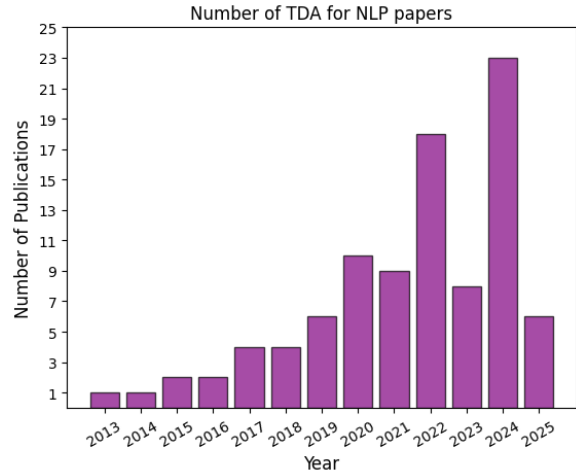


Figure 1: Number of NLP papers using TDA published each year from 2013 to May-2025

probing the topological space, shape, and evolution of topics. On the other hand, *Non-theoretical* approaches mainly discuss how to effectively apply existing numerical representation techniques in NLP to extract novel topological features with TDA. We will first discuss the principles behind TDA and the two main techniques employed for TDA feature extraction: *Persistent Homology* and *Mapper*.

## 2 Topological Data Analysis (TDA)

Topology is defined as “*the study of geometric properties and spatial relations unaffected by the continuous change of shape or size of figures,*” (Ox-

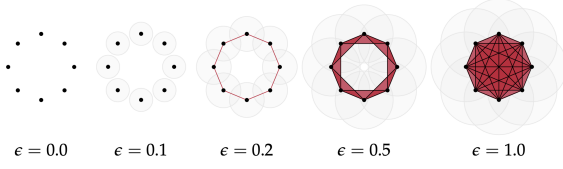


Figure 2: Illustration of the Persistent Homology technique using different radii to find the persistent features (Rieck, 2020).  $\epsilon$  is the ball diameter.

ford Dictionary). TDA is then a collection of powerful techniques that can quantify the shape and structure of data (Munch, 2017). While, there are alternatives to TDA (i.e., see Table 1), TDA is the only technique that can extract not only local but global features, preserve the shape and structure of the data, and is still effective even with insufficient data. Two main techniques are used to extract TDA features - *Persistent Homology* and *Mapper*.

## 2.1 Persistent Homology

Persistent Homology (PH) (Edelsbrunner et al., 2008) is the most popular TDA technique. It uses algebraic topology methods to extract topological signatures at different spatial dimensions. This process involves representing data as a point cloud and performing deformation or perturbation processes to extract the true “shape” of data after the noise has been removed. To achieve this, PH employs Vietoris-Rips complex (Munch, 2017). Vietoris-Rips complex is a way to build simplicial complexes which are used to represent data in a topological space. A simplicial complex is a topological space built by putting points, lines, and higher dimensional shapes together. These formations reveal features that are holes in different dimensions, represented as *beti numbers* ( $\beta_d$ ,  $d$ -dimension). Holes in the 0-dimension ( $\beta_0$ ) is represented as one vertex, 1-dimension ( $\beta_1$ ) is represented as an edge, 2-dimension ( $\beta_2$ ) is represented as a triangle. Further, these features are called connected components, loops/tunnels, and voids, respectively.

Using the method described above, data is represented as a point cloud, and circles are drawn around each point. Next, the radius of each circle is increased using a defined range of points, such that if the circles get bigger and touch, one of the points disappears and this is recorded as a *death*. Additionally, this process of perturbation in different dimensions can cause the *birth* of a new hole which is also recorded. Thus, due to these deformations, the following TDA features can be extracted

and recorded in a 3-column matrix, which consists of columns representing - the *birth* (formation of holes), *death* (deformation or the closing of holes), and persistence features. Persistence is defined as the length of time it took a feature to disappear or die ( $death - birth$ ). The *death* is recorded with the radii value at which the points overlap. Lastly, TDA features are typically visualized in a persistence diagram which is a visual representation of the 3-column matrix of TDA features. Figure 2 illustrates an example of the process of extracting TDA features using Persistent Homology. Other ways of visualizing TDA features include Persistence images (Adams et al., 2017) and Barcode plots (Ghrist, 2008). Application-wise, PH has been used to extract novel features to complement existing NLP representations and improve various classification performances (Doshi and Zadrozny, 2018; Uchendu et al., 2024; Wu et al., 2022).

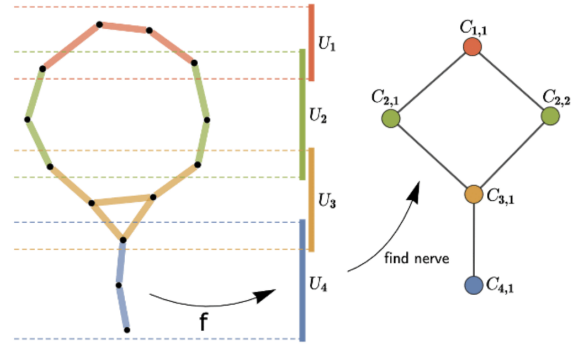


Figure 3: Illustration of Mapper from Murugan and Robertson (2019). The filter function  $f$  is a height function, which is a projection onto the y-axis. The cover of the projected space is the four intervals  $U_i$ . The Mapper graph on the right is a result of applying the rest of the Mapper algorithm and clustering each preimage in the nearest neighbor.

## 2.2 Mapper

Mapper is a dimension reduction clustering technique, used to visualize TDA-extracted topological structures/signatures. It was proposed by Singh et al. (2007) and has been used extensively to visualize topological structures in data to create visually pleasing figures, as well as interpret model performance through data probing (Carls-son, 2020). The Mapper algorithm works in 4 steps<sup>3</sup> (Figure 3) following Murugan and Robertson (2019)’s instructions: (1) Transform the data

<sup>3</sup><https://www.quantmetry.com/blog/topological-data-analysis-with-mapper/>

to a lower-dimensional space using a filter function  $f$ , also known as a lens. This implies projecting from one space to another. Options for filter functions include PCA (Maćkiewicz and Ratajczak, 1993), UMAP (McInnes et al., 2018), and any other dimension-reduction algorithms; (2) Create a cover  $(U_i)_{i \in I}$  for the projected space, which is typically composed of overlapping intervals with a constant length; (3) Cluster the points in the preimage  $f^{-1}(U_i)$  into sets  $C_{i,1}, \dots, C_{i,k_i}$  per interval  $U_i$ ; (4) Create a graph where each vertex represents a cluster set, and there is an edge between two vertices if the corresponding clusters share common points. In this step, points in the same neighborhood are clustered using a defined clustering technique such as DBSCAN (Ester et al., 1996) to change a cluster of several points into a node of a graph.

The intrinsic nature of the Mapper algorithm makes it advantageous in preserving structure even with mapping from one dimension to another. Furthermore, the clustering techniques allow it to be used to explain model performance as the clusters, and colors have meaning that can be further explored. Finally, Mapper is more useful for exploratory data analysis, while PH is more useful for analyzing point clouds and examining the persistence of features.

### 3 Theoretical Approaches of TDA in NLP

Since the field of NLP is very interested in representing and analyzing texts or speech in meaningful ways, several theoretical approaches have been proposed to investigate how well these approaches align with linguistic principles. Thus to explain or confirm linguistic phenomena within the NLP paradigm, a few researchers have proposed topological approaches for probing NLP techniques. By employing TDA techniques - Persistent homology or Mapper to probe for linguistic phenomena, researchers aim to capture the *topological space* (both *semantic* and *syntactic* relationships) of texts, analyze and visualize *topological topic evolution* within texts, and extract the *topological shape* of words. See Table 4 in the Appendix for the theoretical approaches.

**Theoretical vs. Non-Theoretical approaches.** *The main difference between the theoretical and non-theoretical approaches in this context is - Theoretical approaches use TDA (sometimes in combination with other numerical techniques) to understand and explain linguistic phenomena, while Non-theoretical approaches use TDA to enhance or explain model performance. However, linguistic phenomena understood from the Theoretical approaches can inform Non-theoretical approaches.*

### 3.1 Topological Space

#### 3.1.1 Semantic Topological Space

A **semantic topological space** is a conceptual framework used to represent and analyze the relationships between meanings (semantics) of words, phrases, or other linguistic units in a topological or shape structure. This representation involves mapping these units into a mathematical space where the distance or structure between them reflects semantic similarity or other relationships (i.e., Euclidean space  $\rightarrow$  Topological space).

Due to the limitations of traditional clustering algorithms in capturing semantic relationships within texts, Chiang (2007) utilizes TDA and represents semantic space as a simplicial complex of Euclidean space. This technique captures the coherent topics of a document in the connected component dimension ( $\beta_0$ ), representing the semantic topological space (Chiang, 2007). Similarly, Karlgren et al. (2014) visualizes the topological semantic space of text using Mapper which identifies the topical density of the space. To capture topological properties, they train two semantic spaces in a specific topical domain (Karlgren et al., 2014). One space was trained only on articles of similar topics, and the other on introductory paragraphs of those same articles. Findings reveal that clusters of main concepts remained close for the space trained only on articles of similar topics. For the other topological space, the main concepts were randomly distributed (Karlgren et al., 2014).

In addition, Cavaliere et al. (2017) extracts main concepts from the texts by probing the context-aware semantic topological space built with simplicial complexes. Finally, Wagner et al. (2012) uses TF-IDF to numerically represent the top 10-50 words in a corpus and build a topological space that analyzes the structure of similarities within several documents. This topological space is built using discrete Morse theory and persistent homology to find meaningful topological patterns (Wagner et al., 2012).

**Insight (Semantic Space).** *The semantic topological space is explored by researchers to identify semantic linguistic principles captured in texts through a topological lens. Most of the applications in this section involve understanding the semantic similarity between text pairs from a topological lens.*

### 3.1.2 Syntactic Topological Space

A **syntactic topological space** is a theoretical framework used to represent and analyze the relationships between syntactic structures from a topological perspective. This concept is particularly relevant in linguistics, where it helps model and understand the structural aspects of language such as grammar, sentence construction, or the hierarchical organization of those linguistic units.

Therefore, [Port et al. \(2018\)](#) analyzes how syntactic parameters are distributed over different language families, including Indo-European, Niger-Congo, Austronesian, and Afro-Asiatic families. , and features in  $\beta_1$  capture syntactic differences between branches of families of languages, as well as the syntactic influences between them ([Port et al., 2018](#)). [Port et al. \(2018\)](#) show that the three persistent connected components ( $\beta_0$ ) in the Niger-Congo family represents its three subfamilies - Mande, Atlantic-Congo, and Kordofania. The syntactic topological structures of these languages also reveal the historical linguistic phenomena that the Hellenic branch played a role in the historical development of the Indo-European languages ([Port et al., 2018](#)).

Similarly, [Port et al. \(2022\)](#) probes the interpretability of the syntactic topological space. This is done by incorporating more explanation of linguistic structures by introducing so-called homoplasmy phenomena to explain persistent loops further. Homoplasmy phenomena in syntax are observed when dissimilar languages exhibit syntactic similarities ([Port et al., 2022](#)). Findings reveal that the Indo-European family languages - Czech, Lithuanian, Middle Dutch, and Swiss German have the same homoplasmy phenomena ([Port et al., 2022](#)).

**Insight (Syntactic Space).** *The syntactic topological space captures the syntactic structure of language (i.e., grammar, etc.) from a topological lens. Using this framework, researchers confirm linguistic phenomena in language families and subfamilies by exploring the syntactic relationship between languages. Thus, such a framework can be further utilized to discover novel linguistic phenomena.*

## 3.2 Topology of Topic Evolution

The **topology of topic evolution** refers to the study and representation of how topics, themes, or concepts develop and change over time within a given corpus of texts or discourses in a topological space/structure. This concept is particularly relevant in fields where understanding the temporal dynamics of topics can provide insights into trends, shifts in public opinion, or the development of scientific or cultural themes.

Thus, [Sami and Farrahi \(2017\)](#) utilizes TDA to visualize the relationship between words in a text block, words in a corpus, and text blocks in a corpus. They visualize both local context (i.e., each text block in a set of sentences), and global context (i.e., occurrence of extracted words in the corpus) features. These features are extracted by using the circular topology to represent words. Finally, findings reveal that using the circular topology in 2D space, core words from the corpus stay close to the center, and the explanatory words remain close to the circle's periphery.

**Insight (Topic Evolution).** *Exploring the topology of topic evolution is a novel framework for capturing the topology of topics in a corpus. The findings suggest that this framework can be adopted to evaluate the utility of a summarization, paraphrasing, or obfuscating model, by comparing the topology of the topic evolution of the original vs. the perturbed texts.*

## 3.3 Topological “Shape” of Words

The **topological “shape” of words** is a conceptual framework in linguistics and cognitive science that explores the structural properties of *words*. This framework leverages ideas from topology to capture the true shape of words in a linguistically meaningful way. Thus, using topological methods such as TDA, the structural properties of words can be extracted and analyzed.

[Draganov and Skiena \(2024\)](#) captures the “shape” of words for several languages by comparing the phylogenies or evolutionary history of language in the Indo-European language family. Initially, numerically representing the texts with FastText, they use persistent homology to construct language phylogenetic trees for over 81 Indo-European languages. Experiments reveal that: (1) the shape of the word embedding of a language carries historical and structural information, similar to [Port et al. \(2018, 2022\)](#)’s findings; and (2) TDA methods can successfully capture aspects of the shape

of language embeddings (Draganov and Skiena, 2024). Next, Fitz (2022) introduces a novel terminology - *word manifold*, which is a simplicial complex, whose topological space captures grammatical structure expressed by the corpus. This is done by implementing a technique for generating topological structure directly from strings of words. Experiments reveal that the homotopy type of the word manifold is also influenced by linguistic structure (Fitz, 2022). Similarly, Fitz et al. (2024) measures the topological complexity (known as *perforation*) of hidden representation of Transformer-based models to understand their topological shapes.

Similar to Port et al. (2018, 2022), Dong (2024) extracts the topological shapes of languages. Specifically South American languages - the Nuclear-Macro-Jê (NMJ) and Quechuan families using TDA. By using techniques like multiple correspondence analysis (MCA) for dimension reduction of the categorical-valued dataset and persistent homology, Dong (2024) visualizes each language in the selected families as a point cloud. This forms the topological shape of the South American languages, such that languages close together are more similar. By comparing the topological shapes of the languages, it is observed that there are major distinctions between the Jê-proper and the non-Jê-proper languages, as well as the northern and southern Quechuan languages (Dong, 2024). Finally, Bouazzaoui et al. (2021) explores the topological similarity of the shapes of two writing systems - Tifinagh and Phoenician scripts.

**Insight (Shape of Words).** *The topological “shape” of words is a concept that has interested several linguists, as it can be used to confirm and discover linguistic phenomena within languages. It is the most studied theoretical approach, with the main application focused on capturing the shape of several languages. This concept combines all other frameworks like the semantic and syntactic topological spaces to capture a linguistically informed topological shape of texts.*

## 4 Non-theoretical Approaches of TDA in NLP

There are several ways to categorize the applied/non-theoretical TDA applications in NLP. These applications can be categorized by *task*, *learning type*, *modality*, *TDA technique*, and *numerical representation*. We observe that categoriz-

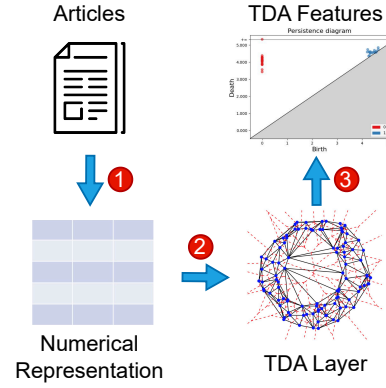


Figure 4: Illustration of the TDA feature extraction pipeline in NLP with three steps: ①-extracting numerical representations, ②-reformatting for TDA’s inputs, and ③-extracting TDA features.

ing these TDA applications by task and numerical representation is more meaningful than the other categories since those categories are binary and not very descriptive of the landscape.

Out of these dimensions, the numerical representation showcases the bottleneck for extracting useful TDA features. See Figure 4 for an illustration of the pipeline for extracting TDA features from numerically represented texts. Therefore, while we focus on both task and numerical representation, our main taxonomy for the non-theoretical applications is centered on how TDA features are extracted from different forms of numerical representations. Figure 5 illustrates the taxonomy of non-theoretical applications of TDA in NLP tasks. See Table 5 in the Appendix for the non-theoretical approaches.

### 4.1 Tasks, Learning Types, Modality, and TDA techniques

We categorize the NLP tasks to which TDA has been applied into seven categories:

1. **Classification:** The most popular application is deepfake text detection (Løvlie, 2023; Tulchinskii et al., 2024; Kushnareva et al., 2024, 2021; Uchendu et al., 2024; Wei et al., 2025).
2. **Clustering and Topic Modeling:** The most popular application is document clustering and topic modeling (Holmes, 2020; Guan et al., 2016).
3. **Sentiment and Semantic Analysis:** The most popular applications are linguistic/grammatical acceptability (Cherniavskii et al., 2022; Jain et al., 2024), word sense induction and disambiguation (Rawson et al., 2022; Temčinas, 2018), and polysemy word classification (Jakubowski et al., 2020;

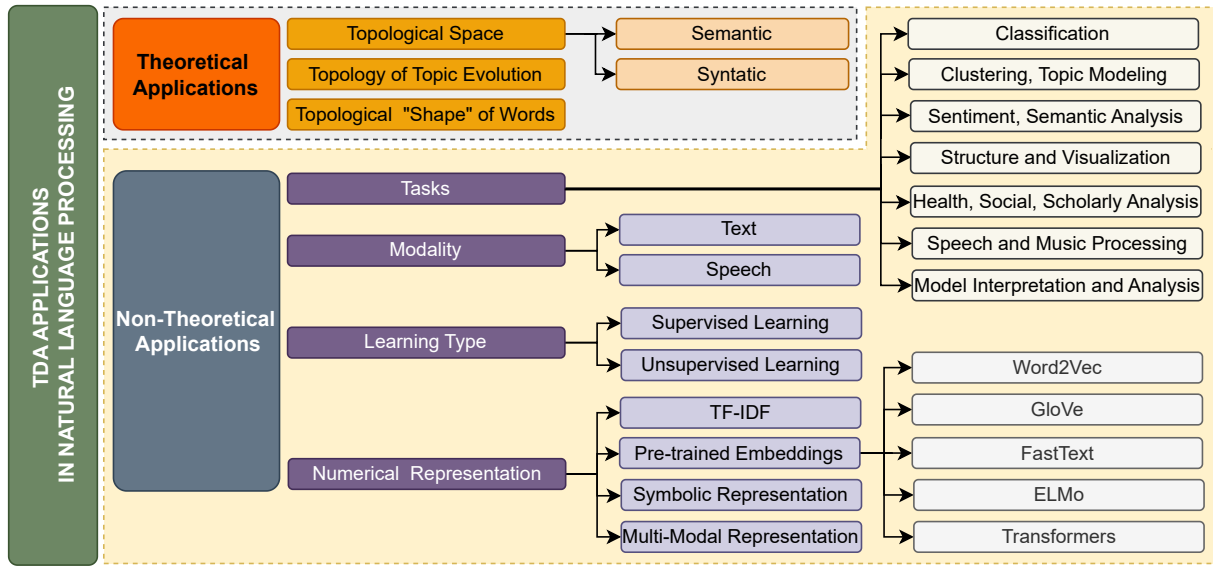


Figure 5: Taxonomy of Topological Data Analysis (TDA) for Natural Language Processing (NLP) Applications

Shehu, 2024).

4. **Structure and Visualization:** The most popular is using Mapper to visualize model hidden weights (Garcia, 2022; Rathore et al., 2023).
5. **Health, Social, and Scholarly Analysis:** Since this is not a popular application for TDA, the most interesting applications are - prediction of epidemics (Petri and Leitao, 2020) and categorization of lonely people (Effah, 2017).
6. **Speech and Music Processing:** The most popular applications are studying vocalizations (Bonafos et al., 2023, 2024) and music classification (Bergomi, 2015; Sassone et al., 2022).
7. **Model Interpretation and Analysis:** The most popular applications are model probing to reveal behavior in hidden weights (Kostenok et al., 2023; Gourgoulis et al., 2024).

**Categories (Tasks).** *The seven categories of Tasks are selected based on surveying the types of problems researchers have attempted to solve by employing TDA techniques.*

Finally, we can also categorize non-theoretical TDA applications in NLP by *learning types*, *Modality*, and *TDA techniques*. *Learning types* has supervised (Elyasi and Moghadam, 2019; Lavery et al., 2024), and unsupervised (Spannaus et al., 2024; Bonafos et al., 2024); *Modality* has text (Triki, 2021; Kostenok et al., 2023), and Speech (Sassone et al., 2022); and *TDA techniques*, has Persistent Homology (Torres-Tramón et al., 2015; Cherniavskii et al., 2022), and Mapper (Holmes, 2020;

Elyasi and Moghadam, 2019).

**Categories (Learning Types, Modality, & TDA techniques).** *Due to the binary nature of these three categories, they are not very informative and thus unsuitable as the main taxonomy for the non-theoretical applications.*

## 4.2 Numerical Representation

### 4.2.1 TF-IDF

TF-IDF (Term Frequency - Inverse Document Frequency) is a well-known statistical formula that calculates the importance of words relative to a corpus. A few works investigated the extraction of topological features from TF-IDF representations as part of the pipeline illustrated in Fig. 4. For instance, *SIFT*, a persistent homology-based model with TF-IDF, is developed to differentiate between child and adolescent writings (Zhu, 2013). This model represented the TF-IDF features as a time series, and then extracted topological features to enhance text classification. Several other researchers applied this model to other classification tasks, such as deepfake text detection (Løvlie, 2023), presidential election speech attribution (Huang, 2022), and movie genre classification (Doshi and Zadrozny, 2018; Shin, 2019).

**Insight (TFIDE).** *Topological features extracted from TF-IDF have only been applied to the text classification task. It is possible that topological features extracted from TF-IDF are not rich enough to perform the other six tasks.*

#### 4.2.2 Pre-trained Non-contextual Embeddings

**Word2Vec Embeddings.** Word2Vec embeddings are a type of word representation that allows words with similar meanings to have similar vector representations (Mikolov et al., 2013). Unlike TF-IDF, topological features extracted from word2vec have been applied to several tasks, including structure & visualization, where Haghighatkah et al. (2022) creates story trees to trace story lines. Next, we observe applications in sentiment & semantic analysis task, where TDA is applied to novel problems such as the creation of a topological search engine (Cornell, 2020), measuring distance between the literary style of Spanish poets (Paluzo Hidalgo et al., 2019), word sense induction and disambiguation (Rawson et al., 2022; Temčinas, 2018), derivation of the correlation between sentence vectors and their semantics (Sun and Nelson, 2023), and comparing text and speech embeddings (Yessenbayev and Kozhirkbayev, 2022, 2024). Thirdly, for the classification task, researchers detect fraudulent papers (Tymochko et al., 2021), and topological loops in logical statements (Tymochko et al., 2020).

Additionally, we observe applications in the health, social, and scholarly analysis task - disease prediction from epidemic curves (Petri and Leitao, 2020), topic modeling tasks (Holmes, 2020; Wright and Zheng, 2020). Finally, for the model interpretation and analysis task, Feng et al. (2024) uses both topological and geometrical features to investigate the strength of LLM-enhanced data augmentation.

**Insight (Word2Vec).** Topological features extracted from Word2vec are considered rich, such that researchers have attempted six out of seven tasks. The only tasks that have not been attempted are speech and music processing, since different embedding is needed for audio data. Finally, we observe that the sentiment and semantic analysis task is the most popular task for word2vec.

**GloVe Embeddings.** GloVe or Global Vectors for Word Representation is another technique for numerically representing texts as embeddings. The tasks we observe that are attempted using topological features extracted from GloVe embeddings are - classification, where subtasks include novel author attribution (Gholizadeh et al., 2018), deepfake text detection (Løvlie, 2023); sentiment and semantic analysis (Gholizadeh et al., 2020; Tymochko et al., 2020; Zadrozny, 2021b); model interpreta-

tion and analysis, where Haim Meirom and Bobrowski (2022); Michel et al. (2017) compare text representations and embeddings, Spannaus et al. (2024) explains model performance, and Zadrozny (2021a) tests the manifestation of intelligence and understanding in models; health, social, and scholarly analysis (Byers, 2021; Novak, 2019); and clustering and topic modeling (Deng and Duzhin, 2022)

**Insight (GloVe).** Topological features extracted from GloVe are considered rich, such that researchers have attempted all tasks, except (1) speech and music processing and (2) structure and visualization. The most popular application is model interpretation and analysis, suggesting that glove embeddings can be sufficiently probed using TDA techniques to excavate explanations for performance.

**FastText Embeddings.** FastText embeddings build on the Word2Vec approach by incorporating subword information, improving the representation of rare words, and allowing for embedding out-of-vocabulary words (Bojanowski et al., 2017). This type of embedding is not a popular feature researchers use to enhance topological features as only two tasks are attempted - sentiment and semantic analysis (Jakubowski et al., 2020; Triki, 2021; Shehu, 2024), specifically for word sense induction & disambiguation; and text classification, where Tymochko et al. (2021) detects fraudulent papers.

**Insight (FastText).** FastText is not popularly adopted for topological applications, with only a few researchers applying it to two tasks - text classification and sentiment & semantic analysis.

#### 4.2.3 Pre-trained Contextual Embeddings

**Transformer Embeddings.** Researchers evaluated the strength of the TDA features extracted from Transformer-based (Vaswani et al., 2017) embeddings. Using the idea of self-attention, the neural network can encode more semantic and syntactic features than previous embeddings which should allow for richer TDA features to be extracted. To incorporate TDA features for various tasks, several researchers have investigated the efficacy of using other outputs of encoder and decoder models - CLS Embedding output, hidden weights, and attention weights to extract high-quality additional features. **CLS Embedding Output.** Researchers have applied these features on text classification tasks, specifically for deepfake text detection (Tulchinskii et al., 2024; Kushnareva et al., 2024; Wei et al., 2025;

Guilinger et al., 2025), fake news detection (Lavery et al., 2024), and TEDtalk public speaking ratings classification (Das et al., 2021).

Additionally, we observe applications to the topic modeling task (Byrne et al., 2022), and the model interpretation and analysis task (Gourgoulas et al., 2024; Proskura and Zaytsev, 2024). Gourgoulas et al. (2024) probes LLMs to estimate class separability of text datasets, and Proskura and Zaytsev (2024) uses topological information from encoder models to select the best models to create an ensemble. Finally, Rathore et al. (2023) performs the structure and visualization task by visualizing the training process of transformer-based models.

**Insight (Transformer-CLS).** *CLS embedding is applied to four tasks - text classification (the most popular), topic modeling, model interpretation & analysis, and structure & visualization. The most interesting subtasks are probing LLMs to estimate class separability of text datasets, and visualizing the training process of transformer-based models.*

**Hidden Weights.** Classification tasks include deepfake text detection (Uchendu et al., 2024), and LLM hallucination detection (Bazarova et al., 2025). Next, researchers explore a combination of sentiment & semantic analysis and structure & visualization tasks by using Mapper to visualize polysemous words in the hidden representations of the BERT transformer model (Garcia, 2022). Similarly, Alexander and Wang (2023) combines the health analysis and visualization tasks to visualize GPT-3’s embeddings of hate speech, misinformation, and psychiatric disorder texts with Mapper. Also, Ruppik et al. (2024) performs the clustering and topic modeling task through dialogue term extraction. Finally, model interpretation and analysis is the most popular task, where Gardinazzi et al. (2024) proposes a novel metric - *persistence similarity* to prune redundant layers in LLMs; Balderas et al. (2025) proposes Persistent BERT Compression and Explainability (PBCE) to compress BERT by pruning redundant layers; Sun and Nelson (2023) probes the correlation between sentence vectors and their semantics; Chauhan and Kaul (2022) proposes a novel scoring metric - *persistence scoring function* which captures the homology of the hidden representations of BERT; García-Castellanos et al. (2024) performs zero-shot model stitching by employing *topological densification*; and Kudriashov et al. (2024) probes BERT’s hidden weights on new grammatical features, known

as *polypersonality*.

**Insight (Transformer-Hidden).** *Researchers have applied TDA features extracted from Hidden weights in to all tasks, except health, social, & scholarly analysis and speed & music processing tasks. From the various novel model interpretation and analysis applications, it seems that TDA features extracted from the Hidden weights are able to make these black-box models more transparent.*

**Attention Weights.** BERT’s attention weights have been transformed to both directed and undirected graphs, on top of which different TDA features are extracted for text classification tasks such as deepfake text detection (Kushnareva et al., 2021; Uchendu et al., 2024), robustness evaluation of TDA features (Perez and Reinauer, 2022), authorship attribution of Japanese texts (Sakurai et al., 2025), out-of-distribution detection (OOD) (Polano et al., 2024; Perez and Reinauer, 2022), and vulnerability detection in code (Snopov and Golubinskiy, 2024). Next, this framework is applied to the sentiment and semantic analysis task, specifically on human linguistic competence (i.e., grammatical acceptability judgment) (Cherniavskii et al., 2022; Proskurina et al., 2023; Perez and Reinauer, 2022), dialog term extraction (Vukovic et al., 2022), and document coherence (Jain et al., 2024). Similarly, with the same framework, a speech and music processing application (Tulchinskii et al., 2023). Finally, researchers attempt the model interpretation and analysis task, where Kostenok et al. (2023) uses the topological features extracted from the attention weights to estimate uncertainty in the encoder models, and Proskura and Zaytsev (2024) performs dynamic weighting for creating ensemble models.

**Insight (Transformer-Attention).** *Given the wealth of information contained in the attention weights, researchers have applied topological features extracted from these weights to several tasks. The most interesting applications are observed on the sentiment and semantic tasks - grammatical acceptability judgment and document coherence analysis. Finally, we observe that for all the Transformer-based weights, the most explored application is deepfake text detection.*

**ELMo Embeddings.** ELMo embeddings are a type of word representation that captures both the meaning of words and their usage in context (Peters et al., 2018). Similar to other embeddings, ELMo has also been leveraged to extract topolog-

ical features. [Tymochko et al. \(2021\)](#) performs text classification by applying these embeddings for detecting fraudulent papers by examining their titles and abstracts. This is done in comparison of other embeddings (Word2Vec, GloVe, FastText, and Frequency Time Series) to ascertain the best embeddings to extract strong topological features.

**Insight (ELMo).** *ELMo embeddings is applied to one task - text classification by one researcher in comparison to other embeddings.*

#### 4.2.4 Symbolic Representations

Symbolic representations in the context of AI and cognitive science refer to the use of symbols such as letters, numbers, tokens, or abstract entities to represent concepts, objects, relationships, and rules within a system. These symbols can be manipulated according to predefined rules to perform reasoning, problem-solving, and decision-making. Symbolic representation contrasts with sub-symbolic representations, such as neural network-based embeddings, which do not explicitly use symbols or rules. This section then discusses the creation of symbolic representations by using *principles of letter coding (PLC)* and *principles of speech sound coding (PSSC)* which are then used to extract topological features.

PLC refers to rules and methods used to encode letters that fuel various communication systems, cryptography techniques, or linguistic analyses. Letter coding transforms letters or characters into different symbols, numbers, or other forms. PSSC is similar to PLC but for extracting topological features from speech sounds. One particular application of PLC and PSSC is the study of tongue twisters, such as for Ukrainian ([Yurchuk and Gurnik, 2023](#); [Kovaliuk et al., 2024](#)). These applications fall under the text classification task for PLC and speech & music processing task.

**Insight (Symbolic).** *Symbolic representations are a novel technique for numerically representing text. The only applications of this technique are on Uranian tongue twisters, both for the classification and speech & processing tasks. This suggests that such a technique could be applied to processing low-resource languages, where numerical representation may be insufficient.*

#### 4.2.5 Multi-Modal Representations

TDA features have also been extracted from other representations of NLP-related features, including

multimedia data such as audio and video. In this section, the only task performed by researchers is speech & music processing, where applications include - studying human vowels and infant vocalizations ([Bonafos et al., 2023, 2024](#)), emotion recognition from audio speech ([Gonzalez-Diaz et al., 2019](#)) and videos ([Paluzo-Hidalgo et al., 2022](#)), depression detection from audio clips ([Tlachac et al., 2020](#)), recognizing voiced and voiceless consonants in speech ([Zhu et al., 2024](#)), music classification ([Bergomi, 2015](#); [Sassone et al., 2022](#)), and assessing the adversarial robustness of image-text multi-modal models by measuring topological consistency ([Vu et al., 2025](#)).

**Insight (Multi-Modal).** *Multi-modal representations are another novel technique for numerically representing. The only task performed by researchers using this embedding is the speech & music processing task. This is because of the nature of the data being audio. Finally, the most interesting applications are the emotion recognition from audio speech, and depression detection from audio clips. All researchers show that topological features extracted from such embeddings experience high gains.*

### 5 Case Studies of TDA for NLP Tasks

We showcase the utility of TDA in non-trivial NLP tasks by applying persistent homology and map-per using the *Theoretical* and *Non-theoretical* approaches. See below:

#### 5.1 Theoretical Approaches.

**(Semantic) Topological Space.** Although applications of this framework have yielded interesting findings, which are discussed in Section 3.1.1, it might still be interesting to answer the question - *does topological features of text capture semantic features as well?* Thus, we perform a small analysis to illustrate a real-world application. We perform this analysis on a relevant topic - distinguishing between watermarked and non-watermarked LLM-generated texts.

We extract topological and linguistic features for a correlation test. The text is initially represented numerically with BERT's attention embeddings, then as an undirected graph, similar to [Kushnareva et al. \(2021\)](#) to extract topological features. Next, using the undirected graph as input, we calculate topological features in the 0-dimension, such as *the number of non-zero values*, *maximum value*, *mean of values*, *number of values*, and *persistent entropy*. Also, for the linguistic features, we extract

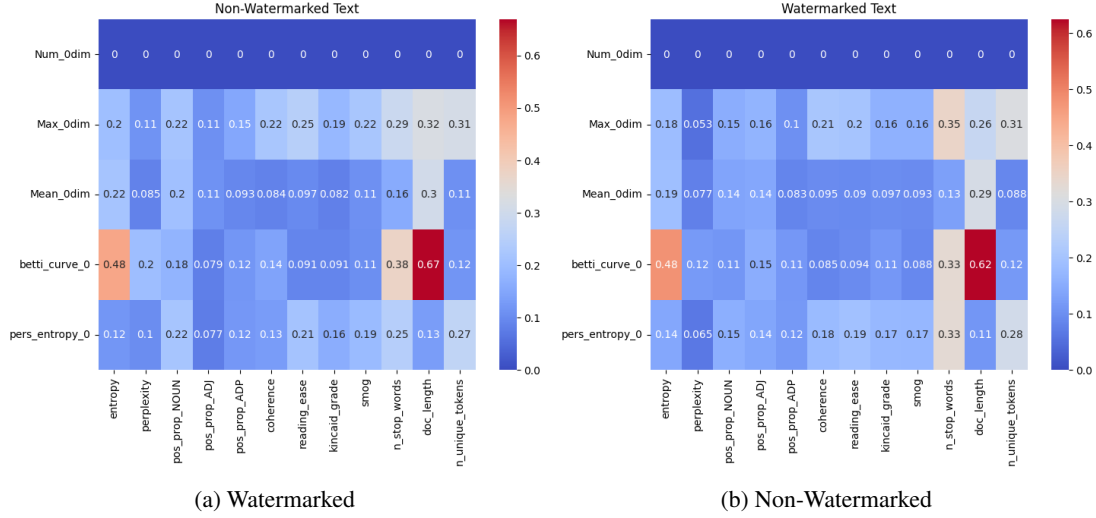


Figure 6: Using Watermarked vs. Non-Watermarked Texts, (a) & (b) are the distance correlation matrix of TDA features vs. Linguistic features heatmaps

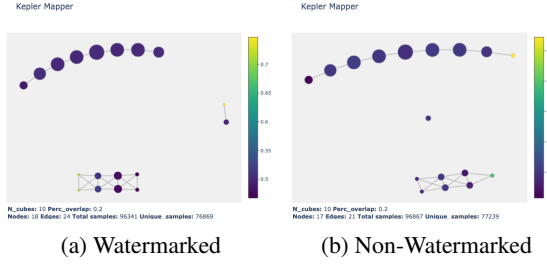


Figure 7: Using Watermarked vs. Non-Watermarked Texts, (a) & (b) are Mapper plots to visualize the shape of data.

entropy, perplexity, position of noun & adjective, coherence, flesh reading ease score & grade, number of stopwords, length of text, and number of unique words. Finally, we use distance correlation to measure the statistical dependence of the two non-linear sets of features. See Figures 6a & 6b for the correlation matrix heatmap. Many of the TDA features correlate with the linguistic features, with the highest correlation being the length of text (`doc_length`) and number of topological values in  $\beta_0$  (`betti_curve_0`). Lastly, these results suggest that topological features can capture semantic features, however, a more comprehensive study is needed to make this a strong finding. In addition, such a finding can inform the non-theoretical approaches we adopt for real-world NLP tasks.

**Topological “Shape” of Words.** Using the same method as above to represent the Watermarked vs. Non-Watermarked texts, we visualize the “shape” of their data using Mapper to observe the similarity

or distinctness between the text types. See Figures 7a and 7b for the Mapper plots of the Watermarked and Non-Watermarked texts, respectively. We observe from the Figures that these different texts are very similar, although the watermarked text has slightly more nodes and edges than the non-watermarked texts, suggesting more connectivity. Thus, we can conclude that these texts are topologically similar, which could suggest that the watermarking technique does not significantly deviate from the style and quality of the non-watermarked version. However, to further explain these results, a more comprehensive study is required to interpret topological features extracted from text linguistically. This further stresses the validity of the theoretical study in Section 5.1. In addition, such findings could indicate that the watermarking technique does not break the semantics, style, and syntax of the non-watermarked pre-trained model.

## 5.2 Non-Theoretical Approaches

Distance	Value (Avg $\pm$ std)
Euclidean	$1.412 \pm 0.0249$
Wasserstein	$0.054 \pm 0.0204$

Table 2: Average Euclidean and Wasserstein distances and standard deviation of the English-French text pairs

**Model Interpretation and Analysis.** To illustrate an interesting task for applying TDA to NLP, we aim to investigate whether the embeddings of English and non-English text pairs are the same.

This analysis can provide insights into the ability of text embeddings to capture the same information but in different languages. The idea is to represent an English text with a monolingual English embedding model and a non-English text with the same meaning with a multilingual embedding model, and compare the similarities between these embeddings. This similarity is captured by calculating both the Euclidean and topological distances. For the topological distance, we use Wasserstein distance, a similarity metric between two probability distributions, typically used in computational topology (Panaretos and Zemel, 2019; Shin, 2019; Draganov and Skiena, 2024). Specifically, we evaluate the Euclidean and Wasserstein distances between the embeddings of English-French translation text pairs. For the Euclidean distance, we compare the embeddings from the English model and the multilingual model. To calculate the Wasserstein distance, we first extract topological features from the embeddings, represent them as persistent diagrams, and then calculate the distance. Euclidean and Wasserstein distance range is  $[0, \infty)$ . We further normalize the embeddings such that both distance metrics are comparable. See Table 2 for the distance values. Thus, we obtain an average distance of 1.412 and 0.054 for the Euclidean and Wasserstein distances, respectively. These values suggest that these embeddings capture the similarity between the text pairs, with Wasserstein being almost zero, meaning that the text pair embeddings are topologically similar. Finally, we propose a new way of evaluating the strength of multilingual models by observing their topological similarity to monolingual models.

## 6 Open Problems and Future Directions

### 6.1 Pros and Cons of TDA in NLP

There are several benefits and challenges to adopting TDA techniques to solve non-trivial problems. We outline these pros and cons for NLP problems in Table 3. The most glaring benefit and challenge are that TDA is effective in low-resource settings and the high computational cost of employing TDA, respectively. This benefit makes TDA particularly attractive for problems with noisy, insufficient, heterogeneous, and high-dimensional datasets. Although the high computational cost presents challenges for the adoption of TDA, recent improvements in computational resources mitigate this challenge. Finally, below we discuss how researchers can leverage these benefits to address the

open problems and future directions.

### 6.2 LLMs to Convert TDA Concepts to Codes

One of the major challenges in applying TDA to NLP tasks is the steep learning curve associated with its mathematical foundations, which are often accessible only to expert audiences. Moreover, the theorists who develop and understand these advanced concepts do not always collaborate with computational scientists to translate them into executable code. To address this gap, Liu et al. (2024) proposes leveraging ChatGPT to generate Python code for TDA concepts by training it on these mathematical foundations. Their findings suggest that ChatGPT can alleviate this bottleneck, particularly for complex TDA concepts like hypergraphs, digraphs, and persistent harmonic space, which have not been as heavily explored as the Vietoris-Rips complex (Liu et al., 2024). Similarly, experts can develop specialized code generators, such as fine-tuning models like Code-Llama (Roziere et al., 2023) on TDA concepts. Creating a dedicated LLM for TDA code generation could significantly lower the entry barrier, encouraging the NLP community to explore TDA applications more innovatively.

### 6.3 Theoretical Approaches Connecting TDA Features with Linguistic Principles

There is a need for theoretical approaches that better tie TDA features to linguistic phenomena. For instance, Draganov and Skiena (2024) investigate the shape of words and their embeddings in Indo-European languages and find similar conclusions as Port et al. (2018) and Port et al. (2022), who investigate the syntactic topological space of such languages. They found that TDA features represent historical facts as in languages clustered closely together are similar or influenced by the other. Finally, we observe only 12 theoretical TDA works in NLP, compared to over 80 non-theoretical ones. Thus, we need more theoretical TDA approaches as the depth and understanding of performance from a topological perspective without more investigations.

### 6.4 TDA for Interpretability

Interpreting TDA features in NLP problems given its non-intuitive nature is very challenging. This is evident in the fact that most TDA for explainability applications is mostly in Computer Vision (Saul and Arendt, 2018), where the structure is distinct. Consequently, the interpretation of TDA features

Pros	Cons
<p><b>Captures Global &amp; Local Text Structure</b> – TDA models relationships between words, sentences, or documents in a topological space, uncovering hierarchical and contextual structures.</p> <p><b>Robust to Noise &amp; Variability</b> – Persistent homology can filter out minor textual variations while retaining essential linguistic structures.</p> <p><b>Effective for Low-Resource Settings</b> – Unlike deep learning, which requires large datasets, TDA can work with smaller corpora by leveraging topological structures rather than statistical frequency-based methods.</p> <p><b>Detects Complex Relationships</b> – TDA can uncover semantic relationships and linguistic patterns that traditional word embeddings may miss.</p> <p><b>No Need for Explicit Feature Engineering</b> – Unlike many traditional methods, TDA analyzes raw data without requiring predefined features.</p> <p><b>Works Well with Word Embeddings &amp; Transformers</b> – Can be integrated with word2vec, and Encoder &amp; Decoder model embeddings to enhance understanding of text structures and improve classification tasks.</p> <p><b>Provides Geometric Insights into Syntax &amp; Semantics</b> – Helps visualize the shape of linguistic structures, which is valuable in understanding complex texts.</p>	<p><b>High Computational Cost</b> – Applying TDA to large text corpora requires significant processing power, especially when constructing high-dimensional topological features.</p> <p><b>Interpretability Challenges</b> – Persistence diagrams and barcodes are not intuitive for NLP practitioners, requiring additional processing to extract meaningful linguistic insights.</p> <p><b>Limited Software &amp; NLP-Specific Tools</b> – Most TDA tools are designed for point clouds and image data, requiring adaptations for NLP applications.</p> <p><b>Not Yet Standard in NLP Pipelines</b> – TDA is still experimental in NLP, lacking standardized frameworks for integration with common NLP libraries like spaCy, Transformers, or NLTK.</p> <p><b>Limited Adoption</b> – TDA is still an emerging field, meaning fewer case studies and industrial applications in NLP compared to traditional methods.</p> <p><b>Requires Specialized Expertise</b> – Implementing TDA in NLP requires both topology and NLP expertise, making adoption difficult for standard NLP practitioners.</p> <p><b>Difficult to Benchmark</b> – Unlike traditional NLP metrics (e.g., BLEU, perplexity), there is no clear evaluation standard for TDA-based NLP models.</p>

Table 3: Pros and Cons of Topological Data Analysis (TDA) in NLP

for text or speech data remains an open problem. There are currently two main tasks in this space - (1) explain model performance by interpreting TDA features; (2) explain model performance by using TDA to probe the prediction space or data. Either tasks require a deeper understanding of TDA such that intuitive explanations can be used to tie topology to linguistic phenomena. Specifically, we need novel approaches that link TDA features to linguistic phenomena, for instance, disentangling  $\beta_0$ , and  $\beta_1$ 's representations to different properties of natural texts such as coherency, and writing style. This can be done either through visualizing the prediction space of a model (Rathore et al., 2023) or probing the prediction space of models with TDA (Gardinazzi et al., 2024; Xenopoulos et al., 2022; Solunke et al., 2024).

## 6.5 Novel Applications of TDA

When we have more theoretical approaches of TDA and issues barring the application of TDA on interpretable NLP tasks are mitigated, we can hope that TDA can be applied to even more novel, diverse and important tasks. From Section 4.1, we can see that TDA has been applied to 7 categories on non-theoretical NLP tasks. While many of the tasks are interesting, especially the speech & music processing and health applications, there are still

nuanced niche fields that could benefit from TDA. One glaring application is on multi-lingual tasks, just as Haim Meirom and Bobrowski (2022) and García-Castellanos et al. (2024). Due to the benefits of TDA which include performing robustly on heterogeneous, imbalanced, and noisy data, its application on multi-lingual tasks is necessary. Other applications include: semantic and syntactic structure, forensic authorship, embedding space, and multi-modal (e.g., language-vision model) analysis

## 6.6 Improvement in TDA Feature Extraction

Unlike some other data modalities that have a distinct shape, texts take the shape of their numerical representation. However, different numerical representations capture different linguistic features, making it challenging to intuitively interpret their topological differences. Therefore, we must find novel ways of representing texts numerically, for example, the use of *symbolic representations* or better ways to use the numerical representations that exist in such a way that is advantageous for extracting the best TDA features.

## 6.7 Adversarial Robustness of TDA Features

Robustness to noise, particularly adversarial perturbations, has been an important research topic in NLP. While such robustness of TDA features is

promising, there have been only a few works in this direction (Perez and Reinauer, 2022; Chauhan and Kaul, 2022). For instance, Perez and Reinauer (2022) show that their topologically-augmented BERT model is much more robust than the base BERT model when tested against perturbations generated by TextAttack (Morris et al., 2020). Chauhan and Kaul (2022) also shows that there are some weak correlations between persistent homology features of a trained BERT model and its adversarial robustness against several state-of-the-art attackers. However, all existing works only evaluate on BERT model with simple attack mechanism, missing other security scenarios such as poisoned or backdoor attacks. Therefore, we call for a more comprehensive robustness evaluation of TDA on NLP models, especially from the NLP, ML, and security communities.

## 6.8 Topological Deep Learning for NLP

Due to the benefits of TDA and deep learning, a new niche field is born - Topological Deep Learning (TDL) a “the collection of ideas and methods related to the use of topological concepts in deep learning” (Papamarkou et al., 2024). TDL allows a deep learning model to be integrated more deeply with concepts of algebraic topology, such as the introduction of simplicial neural networks (NNs), which are NNs with layers made up of simplicial complexes. This deeper integration of TDA into NNs makes TDL particularly useful for the explosion of high-dimensional data. These high-dimensional data need better tools for processing as the current tools shrink the dimension, resulting in information loss. In NLP, one particular approach to integrate TDA with high-dimensional NLP embeddings has been the utilization of text in graphical forms, which have been shown to yield better results than directly using texts as a sequence of tokens (Zhong et al., 2020; Liu et al., 2023; Phan et al., 2023). Nevertheless, more research is still needed to validate such an approach.

## 7 Conclusion

Our world is currently experiencing an explosion of data and an equally explosion of computational techniques to process such data. Machine Learning (ML) is the most popular of these computational methods; however, while its benefits are numerous, it has a few limitations. The biggest of the limitations of ML is its inability to sufficiently

process data that is high-dimensional, imbalanced, noisy, and scarce. Therefore, a small community of NLP researchers emerged to tackle this limitation by proposing using TDA to tackle difficult NLP tasks. These researchers employ two TDA techniques - Persistent Homology and Mapper to solve NLP tasks using theoretical and non-theoretical approaches. This yielded 95 papers which we comprehensively surveyed in this paper. Finally, we conclude that while the applications of TDA in NLP have greatly improved since 2013, there is still room for improvement, specifically in reducing the barrier to entry for non-TDA experts to apply it to their NLP tasks.

## Acknowledgments

Adaku would like to thank Dr. Charlie Dagli for his encouragement, reading the paper drafts, and providing invaluable recommendations.

## References

- Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. 2017. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18(8):1–35.
- Muthuraman Alagappan, Dadi Jiang, Nicholas Denko, and Albert C Koong. 2016. A multimodal data analysis approach for targeted drug discovery involving topological data analysis (tda). In *Tumor Microenvironment: Study Protocols*, pages 253–268. Springer.
- Andrew Alexander and Hongbin Wang. 2023. Topological data mapping of online hate speech, misinformation, and general mental health: A large language model based study. *arXiv preprint arXiv:2309.13098*.
- Luis Balderas, Miguel Lastra, and José M Benítez. 2025. A green ai methodology based on persistent homology for compressing bert. *Applied Sciences*, 15(1):390.
- Alexandra Bazarova, Aleksandr Yugay, Andrey Shulga, Alina Ermilova, Andrei Volodichev, Konstantin Polev, Julia Belikova, Rauf Parchiev, Dmitry Simakov, Maxim Savchenko, et al. 2025. Hallucination detection in llms via topological divergence on attention graphs. *arXiv preprint arXiv:2504.10063*.
- Mattia Giuseppe Bergomi. 2015. *Dynamical and topological tools for (modern) music analysis*. Ph.D. thesis, Université Pierre et Marie Curie-Paris VI; Università degli studi (Milan, Italie).
- Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146.
- Guillem Bonafos, Clara Bourot, Pierre Pudlo, Jean-Marc Freyermuth, Laurence Reboul, Samuel Tronçon, and Arnaud Rey. 2024. Dirichlet process mixture model based on topologically augmented signal representation for clustering infant vocalizations. In *Proc. Interspeech 2024*, pages 3605–3609.
- Guillem Bonafos, Jean-Marc Freyermuth, Pierre Pudlo, Samuel Tronçon, and Arnaud Rey. 2023. Topological data analysis of human vowels: Persistent homologies across representation spaces. *arXiv preprint arXiv:2310.06508*.
- Hajar Bouazzaoui, Mohamed Abdou Elomary, and My Ismail Mamouni. 2021. An application of persistent homology and the graph theory to linguistics: The case of tiffinagh and phoenician scripts. *Statistics in Transition new series*, 22(3):141–156.
- Morgan Byers. 2021. *The Hidden Shape of Data: Topological Data Analysis for Anxiety Detection in Text*. Ph.D. thesis, Texas State University.
- Ciarán Byrne, Danijela Horak, Karo Moilanen, and Amandla Mabona. 2022. Topic modeling with topological data analysis. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 11514–11533.
- Gunnar Carlsson. 2020. Topological methods for data modelling. *Nature Reviews Physics*, 2(12):697–708.
- Danilo Cavaliere, Sabrina Senatore, and Vincenzo Loia. 2017. Context-aware profiling of concepts from a semantic topological space. *Knowledge-Based Systems*, 130:102–115.
- Jatin Chauhan and Manohar Kaul. 2022. Bertops: Studying bert representations under a topological lens. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.
- Daniil Cherniavskii, Eduard Tulchinskii, Vladislav Mikhailov, Irina Proskurina, Laida Kushnareva, Ekaterina Artemova, Serguei Barannikov, Irina Piontkovskaya, Dmitri Piontkovski, and Evgeny Burnaev. 2022. Acceptability judgements via examining the topology of attention maps. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 88–107.
- I-Jen Chiang. 2007. Discover the semantic topology in high-dimensional data. *Expert Systems with Applications*, 33(1):256–262.
- Filip Cornell. 2020. [An explainable topological search engine with giotto-tda](#). In *gt-da-challenge-2020*.
- Shouman Das, Syed A Haque, and Md Iftekhar Tanveer. 2021. Persistence homology of tedtalk: Do sentence embeddings have a topological shape? *arXiv preprint arXiv:2103.14131*.
- Ran Deng and Fedor Duzhin. 2022. Topological data analysis helps to improve accuracy of deep learning models for fake news detection trained on very small training sets. *Big Data and Cognitive Computing*, 6(3):74.
- Tamal K Dey and Sayan Mandal. 2018. Protein classification with improved topological data analysis. In *18th International Workshop on Algorithms in Bioinformatics (WABI 2018)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik.
- Rui Dong. 2024. Linguistics from a topological viewpoint. *arXiv preprint arXiv:2403.15440*.
- Pratik Doshi and Wlodek Zadrozny. 2018. Movie genre detection using topological data analysis. In *Statistical Language and Speech Processing: 6th International Conference, SLSP 2018, Mons, Belgium, October 15–16, 2018, Proceedings 6*, pages 117–128. Springer.
- Ondřej Draganov and Steven Skiena. 2024. The shape of word embeddings: Quantifying non-isometry with topological data analysis. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12080–12099.

- Herbert Edelsbrunner, John Harer, et al. 2008. Persistent homology-a survey. *Contemporary mathematics*, 453(26):257–282.
- Bright Effah. 2017. *Topological data analysis of open-ended responses*. Ph.D. thesis, University of Cape Coast.
- Naierreh Elyasi and Mehdi Hosseini Moghadam. 2019. An introduction to a new text classification and visualization for natural language processing using topological data analysis. *arXiv preprint arXiv:1906.01726*.
- Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231.
- Sherry JH Feng, Edmund MK Lai, and Weihua Li. 2024. Geometry of textual data augmentation: Insights from large language models. *Electronics*, 13(18):3781.
- Stephen Fitz. 2022. The shape of words-topological structure in natural language data. In *Topological, Algebraic and Geometric Learning Workshops 2022*, pages 116–123. PMLR.
- Stephen Fitz, Peter Romero, and Jiyan Jonas Schneider. 2024. Hidden holes: topological aspects of language models. *arXiv preprint arXiv:2406.05798*.
- Jason S Garcia. 2022. *Applications of topological data analysis to natural language processing and computer vision*. Ph.D. thesis, Colorado State University.
- Alejandro García-Castellanos, Giovanni Luca Marchetti, Danica Kragic, and Martina Scolamiero. 2024. Relative representations: Topological and geometric perspectives. In *UniReps: 2nd Edition of the Workshop on Unifying Representations in Neural Models*.
- Yuri Gardinazzi, Giada Panerai, Karthik Viswanathan, Alessio Ansuini, Alberto Cazzaniga, and Matteo Biagetti. 2024. Persistent topological features in large language models. *arXiv preprint arXiv:2410.11042*.
- Shafie Gholizadeh, Armin Seyeditabari, and Wlodek Zadrozny. 2018. Topological signature of 19th century novelists: Persistent homology in text mining. *big data and cognitive computing*, 2(4):33.
- Shafie Gholizadeh, Armin Seyeditabari, and Wlodek Zadrozny. 2020. A novel method of extracting topological features from word embeddings. *arXiv preprint arXiv:2003.13074*.
- Shafie Gholizadeh and Wlodek Zadrozny. 2018. A short survey of topological data analysis in time series and systems analysis. *arXiv preprint arXiv:1809.10745*.
- Robert Ghrist. 2008. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75.
- Rocio Gonzalez-Diaz, Eduardo Paluzo-Hidalgo, and José F Quesada. 2019. Towards emotion recognition: a persistent entropy application. In *Computational Topology in Image Context: 7th International Workshop, CTIC 2019, Málaga, Spain, January 24-25, 2019, Proceedings 7*, pages 96–109. Springer.
- Kostis Gourgoulas, Najah Ghalyan, Maxime Labonne, Sean Moran, Joseph Sabelja, et al. 2024. Estimating class separability of text embeddings with persistent homology. *Transactions on Machine Learning Research*.
- Hui Guan, Wen Tang, Hamid Krim, James Keiser, Andrew Rindos, and Radmila Sazdanovic. 2016. A topological collapse for document summarization. In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE.
- Ann Guiling, Eli Best, and Vinay Awasthi. 2025. Topological data analysis for distinguishing human written and ai generated abstracts. *preprints.org*.
- Pantea Haghighatkah, Antske Fokkens, Pia Sommerauer, Bettina Speckmann, and Kevin Verbeek. 2022. Story trees: Representing documents using topological persistence. In *Proceedings of the Thirteenth LREC 2022*, pages 2413–2429.
- Shaked Haim Meirom and Omer Bobrowski. 2022. [Un-supervised geometric and topological approaches for cross-lingual sentence representation and comparison](#). In *Proceedings of the 7th Workshop on Representation Learning for NLP*, pages 173–183, Dublin, Ireland. Association for Computational Linguistics.
- Felix Hensel, Michael Moor, and Bastian Rieck. 2021. A survey of topological machine learning methods. *Frontiers in Artificial Intelligence*, 4:681108.
- Wesley J Holmes. 2020. Topological analysis of averaged sentence embeddings. Master’s thesis, Wright State University.
- Max Horn, Edward De Brouwer, Michael Moor, Yves Moreau, Bastian Rieck, and Karsten Borgwardt. 2022. Topological graph neural networks. In *International Conference on Learning Representations*.
- Jici Huang. 2022. A tda approach of analyzing election speeches with nlp techniques. *Journal of Computing Sciences in Colleges*, 38(3):215–215.
- Samyak Jain, Rishi Singhal, Sriram Krishna, Yaman K Singla, and Rajiv Ratn Shah. 2024. Beyond words: A topological exploration of coherence in text documents. In *The Second Tiny Papers Track at ICLR 2024*.
- Alexander Jakubowski, Milica Gasic, and Marcus Zibrowius. 2020. [Topology of word embeddings: Singularities reflect polysemy](#). In *Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics*, pages 103–113, Barcelona, Spain (Online). Association for Computational Linguistics.

- Jussi Karlgren, Martin Bohman, Ariel Ekgren, Gabriel Isheden, Emelie Kullmann, and David Nilsson. 2014. Semantic topology. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1939–1942.
- Elizaveta Kostenok, Daniil Cherniavskii, and Alexey Zaytsev. 2023. Uncertainty estimation of transformers’ predictions via topological analysis of the attention matrices. *arXiv preprint arXiv:2308.11295*.
- Tetiana Kovaliuk, Iryna Yurchuk, and Olga Gurnik. 2024. Topological structure of ukrainian tongue twisters based on speech sound analysis. *6th International Workshop on Modern Data Science Technologies*.
- Sergei Kudriashov, Veronika Zykova, Angelina Stepanova, Jacob Raskind, and Eduard Klyshinsky. 2024. The more polypersonal the better—a short look on space geometry of fine-tuned layers. In *International Conference on Neuroinformatics*, pages 13–22. Springer.
- Ankit Kumar and Apurba Sarkar. 2022. Extractive text summarization using topological features. In *International Workshop on Combinatorial Image Analysis*, pages 105–121. Springer.
- Laida Kushnareva, Daniil Cherniavskii, Vladislav Mikhailov, Ekaterina Artemova, Serguei Barannikov, Alexander Bernstein, Irina Piontkovskaya, Dmitri Piontkovski, and Evgeny Burnaev. 2021. Artificial text detection via examining the topology of attention maps. In *Proceedings of the 2021 EMNLP*, pages 635–649.
- Laida Kushnareva, Tatiana Gaintseva, German Magai, Serguei Barannikov, Dmitry Abulkhanov, Kristian Kuznetsov, Eduard Tulchinskii, Irina Piontkovskaya, and Sergey Nikolenko. 2024. Ai-generated text boundary detection with roft. In *1st Conference on Language Modeling (COLM)*, volume 2024.
- Zakaria Lamine, My Ismail Mamouni, and Mohammed Wadia Mansouri. 2023. A topological data analysis of the protein structure. *International Journal of Analysis and Applications*, 21:136–136.
- Rachel Lavery, Anna Jurek-Loughrey, and Lu Bai. 2024. Combining topological signature with text embeddings: Multi-modal approach to fake news detection. In *2024 35th Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE.
- Jian Liu, Li Shen, and Guo-Wei Wei. 2024. Chatgpt for computational topology. *Foundations of data science (Springfield, Mo.)*, 6(2):221.
- Xiaoming Liu, Zhaohan Zhang, Yichen Wang, Hang Pu, Yu Lan, and Chao Shen. 2023. Coco: Coherence-enhanced machine-generated text detection under low resource with contrastive learning. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 16167–16188.
- Bendik Løvlie. 2023. Text classification via topological data analysis. Master’s thesis, Norwegian University of Science and Technology (NTNU).
- Lamis Maadarani and Sayonita Ghosh Hajra. 2020. The shape of poems. In *Fall Poster Forum*.
- Andrzej Maćkiewicz and Waldemar Ratajczak. 1993. Principal components analysis (pca). *Computers & Geosciences*, 19(3):303–342.
- Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. 2018. Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29):861.
- Paul Michel, Abhilasha Ravichander, and Shruti Rijhwani. 2017. Does the geometry of word embeddings help document classification? a case study on persistent homology-based representations. In *Proceedings of the 2nd Workshop on Representation Learning for NLP*, pages 235–240.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.
- John Morris, Eli Lifland, Jin Yong Yoo, Jake Grigsby, Di Jin, and Yanjun Qi. 2020. Textattack: A framework for adversarial attacks, data augmentation, and adversarial training in nlp. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 119–126.
- Elizabeth Munch. 2017. A user’s guide to topological data analysis. *Journal of Learning Analytics*, 4(2):47–61.
- Jeff Murugan and Duncan Robertson. 2019. An introduction to topological data analysis for physicists: From lgm to frbs. *arXiv preprint arXiv:1904.11044*.
- Jessica L Nielson, Jesse Paquette, Aiwen W Liu, Cristian F Guandique, C Amy Tovar, Tomoo Inoue, Karen-Amanda Irvine, John C Gensel, Jennifer Kloke, Tanya C Petrossian, et al. 2015. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. *Nature communications*, 6(1):8581.
- Lance Novak. 2019. Network and topological analysis of scholarly metadata: A platform to model and predict collaboration. Master’s thesis, Purdue University.
- Eduardo Paluzo-Hidalgo, Rocio Gonzalez-Diaz, and Guillermo Aguirre-Carranza. 2022. Emotion recognition in talking-face videos using persistent entropy and neural networks. *Electronic Research Archive*, 30(2):644–660.

- Eduardo Paluzo Hidalgo, Rocío González Díaz, and Miguel Ángel Gutiérrez Naranjo. 2019. Summary and distance between sets of texts based on topological data analysis. *arXiv preprint arXiv:1912.09253*.
- Victor M Panaretos and Yoav Zemel. 2019. Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, 6(1):405–431.
- Theodore Papamarkou, Tolga Birdal, Michael M Bronstein, Gunnar E Carlsson, Justin Curry, Yue Gao, Mustafa Hajij, Roland Kwitt, Pietro Lio, Paolo Di Lorenzo, et al. 2024. Position: Topological deep learning is the new frontier for relational learning. In *Forty-first International Conference on Machine Learning*.
- Ilan Perez and Raphael Reinauer. 2022. The topological bert: Transforming attention into topology for natural language processing. *arXiv preprint arXiv:2206.15195*.
- Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237.
- Giovanni Petri and Antonio Leita. 2020. [Prediction of disease type from topological features of time series](#). In *gt-da-challenge-2020*.
- Huyen Trang Phan, Ngoc Thanh Nguyen, and Dosam Hwang. 2023. Fake news detection: A survey of graph neural network methods. *Applied Soft Computing*, 139:110235.
- Andres Pollano, Anupam Chaudhuri, and Anj Simmons. 2024. Detecting out-of-distribution text using topological features of transformer-based language models. *The IJCAI-2024 AISafety Workshop*.
- Alexander Port, Iulia Gheorghita, Daniel Guth, John M Clark, Crystal Liang, Shival Dasu, and Matilde Marcolli. 2018. Persistent topology of syntax. *Mathematics in Computer Science*, 12(1):33–50.
- Alexander Port, Taelin Karidi, and Matilde Marcolli. 2022. Topological analysis of syntactic structures. *Mathematics in Computer Science*, 16(1):2.
- Polina Proskura and Alexey Zaytsev. 2024. Beyond simple averaging: Improving nlp ensemble performance with topological-data-analysis-based weighting. In *2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–8. IEEE.
- Irina Proskurina, Ekaterina Artemova, and Irina Piontkovskaya. 2023. Can bert eat rucola? topological data analysis to explain. In *Proceedings of the 9th Workshop on Slavic Natural Language Processing 2023 (SlavicNLP 2023)*, pages 123–137.
- Archit Rathore, Yichu Zhou, Vivek Srikumar, and Bei Wang. 2023. Topobert: Exploring the topology of fine-tuned word representations. *Information Visualization*, 22(3):186–208.
- Michael Rawson, Samuel Dooley, Mithun Bharadwaj, and Rishabh Choudhary. 2022. Topological data analysis for word sense disambiguation. *arXiv preprint arXiv:2203.00565*.
- Bastian Rieck. 2020. Topological data analysis for machine learning. *ECML-PKDD 2020 (Tutorial)*.
- Baptiste Roziere, Jonas Gehring, Fabian Gloeckle, Sten Sootla, Itai Gat, Xiaoqing Ellen Tan, Yossi Adi, Jingyu Liu, Romain Sauvestre, Tal Remez, et al. 2023. Code llama: Open foundation models for code. *arXiv preprint arXiv:2308.12950*.
- Benjamin Matthias Ruppik, Michael Heck, Carel van Niekerk, Renato Vukovic, Hsien-chin Lin, Shutong Feng, Marcus Zibrowius, and Milica Gašić. 2024. Local topology measures of contextual language model latent spaces with applications to dialogue term extraction. *Proceedings of the 25th Meeting of the Special Interest Group on Discourse and Dialogue*.
- Wataru Sakurai, Masato Asano, Daisuke Imoto, Masakatsu Honma, and Kenji Kurosawa. 2025. Authorship attribution by attention pattern of bert with topological data analysis and umap. In *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pages 0296–0301. IEEE.
- Ishrat Rahman Sami and Katayoun Farrahi. 2017. A simplified topological representation of text for local and global context. In *Proceedings of the 25th ACM International Conference on Multimedia*, pages 1451–1456.
- Luca Sassone, Marco Manetti, Mattia G Bergomi, and Massimo Ferri. 2022. *Bridging Topological Persistence and Machine Learning for Music Information Retrieval*. Ph.D. thesis, Sapienza – University of Rome.
- Nathaniel Saul and Dustin L Arendt. 2018. Machine learning explanations with topological data analysis. In *Demo at the Workshop on Visualization for AI explainability (VISxAI)*.
- Ketki Savle, Wlodek Zadrozny, and Minwoo Lee. 2019. Topological data analysis for discourse semantics? In *Proceedings of the 13th International Conference on Computational Semantics-Student Papers*, pages 34–43.
- Denis Shehu. 2024. An analysis of the effect of polysemy on the topology of the latent manifold. Master’s thesis, Eindhoven University of Technology.
- Kevin Shin. 2019. Genre classification: A topological data analysis approach. *kevin-shin.com*.

- Gurjeet Singh, Facundo Mémoli, Gunnar E Carlsson, et al. 2007. Topological methods for the analysis of high dimensional data sets and 3d object recognition. *PBG@ Eurographics*, 2:091–100.
- Yashbir Singh, Colleen M Farrelly, Quincy A Hathaway, Tim Leiner, Jaidip Jagtap, Gunnar E Carlsson, and Bradley J Erickson. 2023. Topological data analysis in medical imaging: current state of the art. *Insights into Imaging*, 14(1):58.
- Yara Skaf and Reinhard Laubenbacher. 2022. Topological data analysis in biomedicine: A review. *Journal of Biomedical Informatics*, 130:104082.
- Pavel Snopov and Andrey Nikolaevich Golubinskiy. 2024. Vulnerability detection via topological analysis of attention maps. *arXiv preprint arXiv:2410.03470*.
- Parikshit Solunke, Vitoria Guardieiro, Joao Rulff, Peter Xenopoulos, Gromit Yeuk-Yin Chan, Brian Barr, Luis Gustavo Nonato, and Claudio Silva. 2024. Mountaineer: Topology-driven visual analytics for comparing local explanations. *IEEE Transactions on Visualization and Computer Graphics*.
- Blaž Sovdat. 2016. Text mining via homology. Master’s thesis, UNIVERSITY OF LJUBLJANA.
- Adam Spannaus, Heidi A Hanson, Georgia Tourassi, and Lynne Penberthy. 2024. Topological interpretability for deep learning. In *Proceedings of the Platform for Advanced Scientific Computing Conference*, pages 1–11.
- Tianyi Sun and Bradley Nelson. 2023. Topological interpretations of gpt-3. *arXiv preprint arXiv:2308.03565*.
- Tadas Temčinas. 2018. Local homology of word embeddings. *arXiv preprint arXiv:1810.10136*.
- ML Tlachac, Adam Sargent, Ermal Toto, Randy Paffenroth, and Elke Rundensteiner. 2020. Topological data analysis to engineer features from audio signals for depression detection. In *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 302–307. IEEE.
- Pablo Torres-Tramón, Hugo Hromic, and Bahareh Rahmazadeh Heravi. 2015. Topic detection in twitter using topology data analysis. *Current Trends in Web Engineering*, pages 186–197.
- Jonas Folkvord Triki. 2021. Analysis of word embeddings: A clustering and topological approach. Master’s thesis, The University of Bergen.
- Ilya Trofimov, Daniil Cherniavskii, Eduard Tulchinskii, Nikita Balabin, Evgeny Burnaev, and Serguei Barannikov. 2023. Learning topology-preserving data representations. In *ICLR 2023 International Conference on Learning Representations*.
- Eduard Tulchinskii, Kristian Kuznetsov, Daniil Cherniavskii, Serguei Barannikov, Sergey Nikolenko, and Evgeny Burnaev. 2023. Topological data analysis for speech processing. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pages 311–315.
- Eduard Tulchinskii, Kristian Kuznetsov, Laida Kushnareva, Daniil Cherniavskii, Sergey Nikolenko, Evgeny Burnaev, Serguei Barannikov, and Irina Piontkovskaya. 2024. Intrinsic dimension estimation for robust detection of ai-generated texts. *Advances in Neural Information Processing Systems*, 36.
- Sarah Tymochko, Julien Chaput, Timothy Doster, Emilie Purvine, Jackson Warley, and Tegan Emerson. 2021. Con connections: Detecting fraud from abstracts using topological data analysis. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 403–408. IEEE.
- Sarah Tymochko, Zachary New, Lucius Bynum, Emilie Purvine, Timothy Doster, Julien Chaput, and Tegan Emerson. 2020. Argumentative topology: Finding loop (holes) in logic. *arXiv preprint arXiv:2011.08952*.
- Adaku Uchendu, Thai Le, and Dongwon Lee. 2024. Topformer: Topology-aware authorship attribution of deepfake texts with diverse writing styles. *ECAI 2024*.
- Lucrezia Valeriani, Diego Doimo, Francesca Cuturello, Alessandro Laio, Alessio Ansuini, and Alberto Cazzaniga. 2024. The geometry of hidden representations of large transformer models. *Advances in Neural Information Processing Systems*, 36.
- Hendrik Jacob van Veen. 2020. Novel topological shapes of model interpretability. In *TDA and Beyond at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- Minh Vu, Geigh Zollicoffer, Huy Mai, Ben Nebgen, Boian Alexandrov, and Manish Bhattarai. 2025. Topological signatures of adversaries in multimodal alignments. *arXiv preprint arXiv:2501.18006*.
- Renato Vukovic, Michael Heck, Benjamin Ruppik, Carel van Niekerk, Marcus Zibrowius, and Milica Gasic. 2022. Dialogue term extraction using transfer learning and topological data analysis. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 564–581.
- Hubert Wagner, Paweł Dłotko, and Marian Mrozek. 2012. Computational topology in text mining. In *Computational Topology in Image Context: 4th International Workshop, CTIC 2012, Bertinoro, Italy, May 28-30, 2012. Proceedings*, pages 68–78. Springer.

Malgorzata Wamil, Abdelaali Hassaine, Shishir Rao, Yikuan Li, Mohammad Mamouei, Dexter Canoy, Milad Nazarzadeh, Zeinab Bidel, Emma Copland, Kazem Rahimi, et al. 2023. Stratification of diabetes in the context of comorbidities, using representation learning and topological data analysis. *Scientific Reports*, 13(1):11478.

Dongjun Wei, Minjia Mao, Xiao Fang, and Michael Chau. 2025. Short-phd: Detecting short llm-generated text with topological data analysis after off-topic content insertion. *arXiv preprint arXiv:2504.02873*.

Matthew Wright and Xiaojun Zheng. 2020. Topological data analysis on simple english wikipedia articles. *The PUMP Journal of Undergraduate Research*, 3:308–328.

Xiangcheng Wu, Xi Niu, and Ruhani Rahman. 2022. Topological analysis of contradictions in text. In *Proceedings of the 45th International ACM SIGIR*, pages 2478–2483.

Peter Xenopoulos, Gromit Chan, Harish Doraiswamy, Luis Gustavo Nonato, Brian Barr, and Claudio Silva. 2022. Gale: Globally assessing local explanations. In *Topological, Algebraic and Geometric Learning Workshops 2022*, pages 322–331. PMLR.

Zhandos Yessenbayev and Zhanibek Kozhimbayev. 2022. Comparison of word embeddings of unaligned audio and text data using persistent homology. In *International Conference on Speech and Computer*, pages 700–711. Springer.

Zhandos Yessenbayev and Zhanibek Kozhimbayev. 2024. Use of riemannian distance metric to verify topological similarity of acoustic and text domains. In *International Conference on Artificial Neural Networks*, pages 368–380. Springer.

Iryna Yurchuk and Olga Gurnik. 2023. Tongue twisters detection in ukrainian by using tda. In *CEUR Workshop Proceedings*, pages 163–172.

Wlodek W Zadrozny. 2021a. Abstraction, reasoning and deep learning: A study of the "look and say" sequence. *arXiv preprint arXiv:2109.12755*.

Wlodek W Zadrozny. 2021b. A note on argumentative topology: Circularity and syllogisms as unsolved problems. *arXiv preprint arXiv:2102.03874*.

Wanjun Zhong, Duyu Tang, Zenan Xu, Ruize Wang, Nan Duan, Ming Zhou, Jiahai Wang, and Jian Yin. 2020. Neural deepfake detection with factual structure of text. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2461–2470.

Xiao Jin Zhu. 2013. Persistent homology: An introduction and a new text representation for natural language processing. In *IJCAI*, pages 1953–1959.

Y Zhu, P Feng, S Yi, Q Qu, and Z Yu. 2024. Topology-enhanced machine learning for consonant recognition. *Research Square*.

## A Case Studies of TDA for NLP Tasks

We showcase the utility of TDA in non-trivial NLP tasks by applying persistent homology and mapper using the *Theoretical* and *Non-theoretical* approaches. See below:

### A.1 Theoretical Approaches.

**(Semantic) Topological Space.** For the analysis, we use a subset of the C4 dataset that has watermarked and non-watermarked LLM-generated texts<sup>4</sup>. Next, we use the *ripser* python package to extract the topological features with persistent homology and *texdescriptives* Python package<sup>5</sup> to extract the linguistic features.

**Topological “Shape” of Words.** Using the same C4 dataset, we create Mapper diagrams using the *Mapper* Python package<sup>6</sup>.

### A.2 Non-Theoretical Approaches

**Model Interpretation and Analysis.** We evaluate the Euclidean and Wasserstein distances between the embeddings of English-French translation text pairs<sup>7</sup> (using the first 500 text pairs). For the Euclidean distance, we compare the embeddings from the English model<sup>8</sup> and the multilingual model<sup>9</sup>. To calculate the Wasserstein distance, we first extract topological features from the embeddings, represent them as persistent diagrams, and then calculate the distance. Euclidean and Wasserstein distance range is  $[0, \infty)$ .

## B List of papers

See list of Theoretical and Non-Theoretical applications of TDA in NLP in Tables 4 and 5, respectively.

<sup>4</sup>[https://huggingface.co/datasets/acmc/watermarked\\_c4\\_dataset](https://huggingface.co/datasets/acmc/watermarked_c4_dataset)

<sup>5</sup><https://github.com/HLasse/TextDescriptives?tab=readme-ov-file>

<sup>6</sup><https://github.com/scikit-tda/kepler-mapper/tree/master>

<sup>7</sup><https://huggingface.co/datasets/aircrypto/English-French-Translations>

<sup>8</sup><https://huggingface.co/sentence-transformers/stsb-xlm-r-multilingual>

<sup>9</sup><https://huggingface.co/sentence-transformers/msmarco-roberta-base-v2>

Name	Category	Task	TDA Technique
(Chiang, 2007)	Topological Space (Semantic)	Capturing semantic relationships within texts	Persistent Homology
(Karlgren et al., 2014)	Topological Space (Semantic)	Identify topical density of the space	Mapper
(Cavaliere et al., 2017)	Topological Space (Semantic)	Extracting extracts main concepts from text	Persistent Homology
(Wagner et al., 2012)	Topological Space (Semantic)	Analyzes document similarities	Persistent Homology
(Port et al., 2018)	Topological Space (Syntactic)	Analyzes syntactic parameters of different language families	Persistent Homology
(Port et al., 2022)	Topological Space (Syntactic)	Explaining linguistic structures with homoplasy phenomena	Persistent Homology
(Sami and Farrahi, 2017)	Topology of Topic Evolution	Topic evolution within documents	Persistent Homology
(Draganov and Skiena, 2024)	Topological “Shape” of Words	Investigate the “shape” of language phylogenies in the Indo-European language family	Persistent Homology
(Fitz, 2022)	Topological “Shape” of Words	Capturing grammatical structure expressed by corpus using <i>word manifold</i>	Persistent Homology
(Fitz et al., 2024)	Topological “Shape” of Words	Measures the topological complexity of Transformer-based hidden representations	Persistent Homology
(Dong, 2024)	Topological “Shape” of Words	Investigates the topological shapes of South American languages - the Nuclear-Macro-Jê (NMJ) and Quechuan families	Persistent Homology
(Bouazzaoui et al., 2021)	Topological “Shape” of Words	Investigates the similarity between the topological shapes of the Tifinagh and Phoenician scripts	Persistent Homology

Table 4: Theoretical Applications of TDA in NLP.

Name	Task	Numerical Representation	Learning Type	Modality	TDA Technique
SIFT (Zhu, 2013)	child vs. adolescent writing detection	TF-IDF	Supervised	Text	Persistent Homology
(Løvlie, 2023)	deepfake text detection	TF-IDF, GloVe	Supervised	Text	Persistent Homology
(Huang, 2022)	election speech feature extraction	TF-IDF	Supervised	Text	Persistent Homology
(Doshi and Zadrozny, 2018; Shin, 2019)	movie genre classification	TF-IDF	Supervised	Text	Persistent Homology
(Sovdat, 2016)	distinguishing between languages	TF-IDF	Supervised	Text	Persistent Homology
(Savle et al., 2019)	legal document entailment	TF-IDF	Supervised	Text	Persistent Homology
(Michel et al., 2017)	clustering and sentiment analysis	TF-IDF, GloVe	Supervised	Text	Persistent Homology
DoCollapse (Guan et al., 2016)	keyphrase extraction	TF-IDF	Unsupervised	Text	Persistent Homology
TOPOL (Torres-Tramón et al., 2015)	Twitter topic detection	TF-IDF	Supervised	Text	Persistent Homology
(Kumar and Sarkar, 2022)	text summarization	TF-IDF	Unsupervised	Text	Persistent Homology
(van Veen, 2020)	Propaganda tweet classification	TF-IDF	Unsupervised	Text	Mapper
(Elyasi and Moghadam, 2019)	classify Persian poems	TF-IDF	Supervised	Text	Persistent Homology & Mapper
(Effah, 2017)	age group categorization of lonely people	TF-IDF	Supervised	Text	Persistent Homology & Mapper
(Maadarani and Hajra, 2020)	nursery rhyme classification from different continents - Australia, Asia, Africa, Europe, and North America	TF-IDF	Supervised	Text	Persistent Homology
(Haghighatkah et al., 2022)	building document structure	Pre-trained (Word2Vec)	Unsupervised	Text	Persistent Homology
BERT+TDA (Wu et al., 2022)	contradiction detection	Pre-trained (Word2Vec)	Supervised	Text	Persistent Homology
(Yessenbayev and Kozhirkbayev, 2022)	speaker recognition & text processing	Pre-trained (Word2Vec)	Unsupervised	Speech	Persistent Homology
(Yessenbayev and Kozhirkbayev, 2024)	speaker recognition & text processing	Pre-trained (Word2Vec)	Unsupervised	Speech	Persistent Homology
(Cornell, 2020)	building a topological search engine	Pre-trained (Word2Vec)	Unsupervised	Text	Mapper
(Holmes, 2020)	document clustering and topic modeling tasks	Pre-trained (Word2Vec)	Supervised	Text	Mapper
(Rawson et al., 2022)	word sense induction and disambiguation	Pre-trained (Word2Vec)	Unsupervised	Text	Persistent Homology
(Temčinas, 2018)	word sense induction and disambiguation	Pre-trained (Word2Vec, GloVe)	Unsupervised	Text	Persistent Homology
(Feng et al., 2024)	Geometry of textual data augmentation	Pre-trained (Word2Vec)	Supervised	Text	Persistent Homology
(Tymochko et al., 2021)	fraudulent paper detection	Pre-trained (Word2Vec, GloVe, ElMo)	Supervised	Text	Persistent Homology
(Petri and Leita, 2020)	disease epidemic prediction	Pre-trained (Word2Vec)	Supervised	Text	Persistent Homology
(Tymochko et al., 2020)	finding topological loops in logical statements	Pre-trained (Word2Vec)	Unsupervised	Text	Persistent Homology
(Wright and Zheng, 2020)	distinguish subsets in data	Pre-trained (Word2Vec)	Unsupervised	Text	Persistent Homology
(Paluzo Hidalgo et al., 2019)	measuring the distance between the literary style of Spanish poets	Pre-trained (Word2Vec)	Supervised	Text	Persistent Homology
(Gholizadeh et al., 2018)	extract the topological signatures of novelists	Pre-trained (GloVe)	Supervised	Text	Persistent Homology
TIES (Gholizadeh et al., 2020)	document categorization & sentiment analysis	Pre-trained (GloVe)	Unsupervised	Text	Persistent Homology
(Spannaus et al., 2024)	phenotype prediction and news group categorization	Pre-trained (GloVe)	Unsupervised	Text	Mapper
(Zadrozny, 2021b)	finding topological loops in logical statements	Pre-trained (GloVe)	Unsupervised	Text	Persistent Homology
(Byers, 2021)	social anxiety detection	Pre-trained (GloVe)	Supervised	Text	Persistent Homology
(Haim Meiron and Bobrowski, 2022)	compare cross-lingual sentence representations	Pre-trained (GloVe)	Unsupervised	Text	Persistent Homology
(Zadrozny, 2021a)	investigates the manifestations of intelligence and understanding in neural networks	Pre-trained (GloVe)	Supervised	Text	Persistent Homology
(Deng and Duzhin, 2022)	fake news detection	Pre-trained (GloVe, BERT)	Supervised	Text	Persistent Homology
(Novak, 2019)	analyzing scholarly network	Pre-trained (GloVe, BERT)	Unsupervised	Text	Persistent Homology
(Jakubowski et al., 2020)	(1) polysemy word classification, and (2) word sense induction & disambiguation	Pre-trained (FastText)	Unsupervised	Text	Persistent Homology
(Triki, 2021; Shehu, 2024)	polysemy word classification	Pre-trained (FastText)	Supervised	Text	Persistent Homology
PHD (Tulchinskii et al., 2024)	deepfake text detection	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
Short-PHD (Wei et al., 2025)	deepfake text detection (for short-text)	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
(Guilinger et al., 2025)	deepfake text detection (for academic abstracts)	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
(Kushnareva et al., 2024)	deepfake text detection	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
(Gourgoulas et al., 2024)	class separability estimation	Pre-trained (Transformers - CLS)	Unsupervised	Text	Persistent Homology
TopoBERT (Rathore et al., 2023)	visually analyzing the fine-tuning process of a Transformer-based model	Pre-trained (Transformers - CLS)	Unsupervised	Text	Mapper
(Lavery et al., 2024)	fake news detection	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
(Das et al., 2021)	classification of public speaking ratings from TED talks	Pre-trained (Transformers - CLS)	Supervised	Text	Persistent Homology
(Byrne et al., 2022)	topic modeling	Pre-trained (Transformers - CLS)	Unsupervised	Text	Mapper
TOPFORMER (Uchendu et al., 2024)	deepfake text detection	Pre-trained (Transformers - Hidden)	Supervised	Text	Persistent Homology
(Garcia, 2022)	polysemy word classification	Pre-trained (Transformers - Hidden)	Supervised	Text	Mapper
(Alexander and Wang, 2023)	Hate speech, Misinformation & Psychiatric disorder classification	Pre-trained (Transformers - Hidden)	Supervised	Text	Mapper
Persistent BERT Compression and Explainability (PBCE) (Balderas et al., 2025)	Model compression	Pre-trained (Transformers - Hidden)	Unsupervised	Text	Persistent Homology
Persistent Similarity (Gardinazzi et al., 2024)	Probing layers in LLMs	Pre-trained (Transformers - Hidden)	Unsupervised	Text	Persistent Homology
(Sun and Nelson, 2023)	Correlation between sentence vectors	Pre-trained (Word2Vec, Transformers - Hidden)	Unsupervised	Text	Persistent Homology
Persistence Scoring Function (Chauhan and Kaul, 2022)	captures the homology of the high-dimensional hidden representations	Pre-trained (Transformers - Hidden)	Unsupervised	Text	Persistent Homology
Topological Densification (García-Castellanos et al., 2024)	zero-shot model stitching	Pre-trained (Transformers - Hidden)	Unsupervised	Text	Persistent Homology
(Ruppik et al., 2024)	dialogue term extraction	Pre-trained (Transformers - Hidden)	Supervised	Text	Persistent Homology
(Kudriashov et al., 2024)	polypersonality	Pre-trained (Transformers - Hidden)	Unsupervised	Text	Persistent Homology
(Bazarova et al., 2025)	LLM Hallucination detection	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Kushnareva et al., 2021)	deepfake text detection	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Cherniavskii et al., 2022; Proskurina et al., 2023; Jain et al., 2024)	grammatical acceptability judgment	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Kostenok et al., 2023)	Uncertainty estimation of model predictions	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Perez and Reinauer, 2022)	spam detection, grammatical acceptability judgment, and movie sentiment analysis	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Sakurai et al., 2025)	Authorship attribution of Japanese texts	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Pollano et al., 2024)	out-of-distribution detection	Pre-trained (Transformers - Attention, CLS)	Supervised	Text	Persistent Homology
(Proskura and Zaytsev, 2024)	estimation of weights for ensembles of classification models	Pre-trained (Transformers - Attention, CLS)	Supervised	Text	Persistent Homology
(Snopov and Golubinskiy, 2024)	vulnerability detection in code	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
TopoHuBERT (Tulchinskii et al., 2023)	speaker recognition	Pre-trained (Transformers - Attention)	Supervised	Speech	Persistent Homology
(Vukovic et al., 2022)	dialogue term extraction	Pre-trained (Transformers - Attention)	Supervised	Text	Persistent Homology
(Yurchuk and Gurnik, 2023)	Ukrainian tongue twisters classification	Symbolic Representations	Supervised	Text	Persistent Homology
(Kovalyuk et al., 2024)	Ukrainian tongue twisters classification	Symbolic Representations	Supervised	Speech	Persistent Homology
(Bonafos et al., 2023)	human vowel classification	Multi-Modal	Supervised	Speech	Persistent Homology
(Bonafos et al., 2024)	infant vocalization classification	Multi-Modal	Unsupervised	Speech	Persistent Homology
(Gonzalez-Diaz et al., 2019; Paluzo-Hidalgo et al., 2022)	emotion recognition	Multi-Modal	Supervised	Speech	Persistent Homology
(Tlachac et al., 2020)	depression detection	Multi-Modal	Supervised	Speech	Persistent Homology
(Zhu et al., 2024)	consonants recognition	Multi-Modal	Supervised	Speech	Persistent Homology
(Sassone et al., 2022)	music genre classification	Multi-Modal	Supervised	Speech	Persistent Homology
(Vu et al., 2025)	multi-modal adversarial robustness assessment	Multi-Modal	Supervised	Image & Text	Persistent Homology
(Bergomi, 2015)	music classification	Multi-Modal	Supervised	Speech	Persistent Homology

Table 5: Non-theoretical applications of TDA in NLP