

Topological Data Analysis Applications in Natural Language Processing: A Survey

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

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

ABSTRACT

The surge of data available on the Internet has driven the adoption of a wide range of computational methods for analyzing and extracting insights from large-scale data. Among these, Machine Learning (ML) has become a central paradigm, offering powerful tools for pattern discovery, prediction, and representation learning across many domains. At the same time, real-world data often exhibit properties such as noise, imbalance, sparsity, limited supervision, and high dimensionality, motivating the use of additional analytical perspectives that can complement standard ML pipelines. One such perspective is Topological Data Analysis (TDA), a statistical framework that focuses on the intrinsic shape and structural organization of data. Rather than replacing ML, TDA offers a complementary lens for characterizing geometric and topological properties that may be difficult to capture with conventional feature-based or purely predictive approaches. This has motivated a growing body of work that integrates TDA into ML workflows, particularly in settings where data structure plays an important role. Despite this promise, TDA has received relatively limited attention in Natural Language Processing (NLP) compared to domains with more overt structural regularities, such as computer vision. Nevertheless, a dedicated community of researchers has explored its use in NLP, leading to **137 papers** that we comprehensively survey in this work. We organize these studies into theoretical and non-theoretical approaches. Theoretical approaches use topology to explain linguistic phenomena, whereas non-theoretical approaches incorporate TDA into ML-based pipelines through a variety of numerical representations. We conclude by discussing the key challenges and open questions that continue to shape this emerging area. Resources and a list of papers are available at: <https://github.com/AdaUchendu/AwesomeTDA4NLP>¹.

¹DISTRIBUTION 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. © 2026 Massachusetts Institute of Technology. Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014

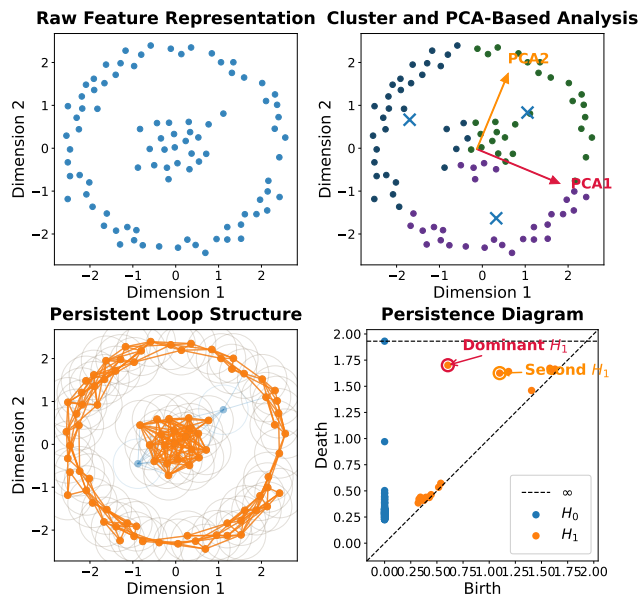


Figure 1: The raw point cloud contains a dense central region together with loop-like outer structure. Conventional ML analysis such as PCA and clustering is strongly influenced by variance and local density, so much of the summary is affected by the central cluster. In contrast, TDA emphasizes global structure. In the persistence diagram, blue points represent *connected component* features and orange points represent *loop* features. For each topological feature, the horizontal coordinate (“birth”) is the scale at which the feature first appears, and the vertical coordinate (“death”) is the scale at which it disappears. Here, the prominent orange points that are further from the diagonal correspond to the persistent loop structures visible in the third panel.

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

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

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 overall 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. Thus, by applying TDA to NLP, we obtain “*topological structures from language*,” which refers not to intrinsic properties of raw text itself, but to the structures that emerge when linguistic data is mapped into high-dimensional embedding spaces. These induced topologies capture relationships among words, sentences, or documents based on their learned representations, rather than any inherent topological features of the text.

TDA is a “collection of powerful tools that can quantify shape and structure in data”² and is inspired by the algebraic topology and geometry mathematical fields. The benefits of TDA are vast, including the ability to extract additional features that are typically not captured by other feature extraction techniques [157, 96, 107]. These features are known as *topological features*. Unsurprisingly, since TDA is used to capture topological features, it has been applied to many tasks where data has distinct graphical structures [107, 63]. These include tasks that have obvious graph-like structures, such as protein classification [33, 82, 158] and drug discovery [2]; to those that are not so obvious, such as diabetes classification [166, 140], image classification [66, 151], and time series analysis [110, 155, 53]. However, since the shape of a text is not apparent, it has not gained as much attention in NLP as it has in the Computer Vision field [66, 151] specifically in the Medical domains [139, 102]. Still, several researchers have found ways to extract unique, global-level features using TDA. Other typical numerical representation techniques in text are unable to extract global-level features, making TDA suitable for the task.

Figure 1 illustrates a toy example that shows the utility of TDA to standard ML analysis. Conventional ML methods often summarize data through variance, centroids, or local grouping, which means dense regions can disproportionately shape the result. In the example, the central cluster absorbs much of the variance, while the outer loop structure plays a less visible role in PCA- and cluster-based summaries. TDA instead tracks topological features such as connected components and loops across multiple scales, allowing it to recover meaningful global structure.

More broadly, TDA aims to answer the central research question - *what is the true shape of data?* We survey **137 papers** that have attempted to find an answer through various approaches. The first application of TDA in NLP was published in 2012 [165], and since then, there have been over 100 papers applying TDA in NLP. There have been a gradual acceleration in the number of published works in TDA applications on various NLP tasks, including ones pertaining to the recent emergence of Large Language Models (LLMs) such as hallucination detection [14, 128], mechanistic interpretability [174, 120], and model efficiency [50, 98]; we project that this trend will continue in the future (Figure 3). Therefore, based on these approaches, we categorize these applications into two - *theoretical* [74, 113] and *non-theoretical* [187, 35] approaches. *Theoretical* approaches involve using TDA to explain linguistic phe-

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

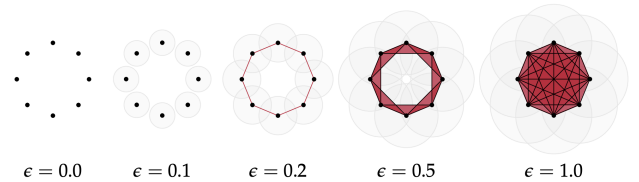


Figure 2: Illustration of the Persistent Homology technique using different radii to find the persistent features [123]. ϵ is the ball diameter.

nomena by 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.

In addition, we observe that theoretical approaches only span 13 papers, while non-theoretical approaches have over 100 papers. Due to the higher number of non-theoretical approaches, we discuss several categories that could be useful for distinguishing applications: numerical representation, tasks, TDA technique, data modality, and learning type. TDA techniques (i.e., Persistent Homology and Mapper), data modality (i.e., text and speech), and learning type (i.e., unsupervised and supervised) are binary, making it difficult to meaningfully grasp distinctness from almost 120 papers. However, with tasks which refer to the problems in which approaches are adopted for; these have seven categories - (1) *classification*, (2) *clustering & topic modeling*, (3) *sentiment & semantic analysis*, (4) *structure & visualization*, (5) *health, social, & scholarly analysis*, (6) *speech processing*, and (7) *model interpretation & analysis*. We observe that classification and model interpretation & analysis are the most popular tasks explored by researchers. In addition, numerical representations leveraged in non-theoretical application include - (1) *TF-IDF*, (2) *Word2Vec*, (3) *GloVe*, (4) *FastText*, (5) *ELMo*, (6) *Transformers*, (7) *Symbolic*, and (8) *Multi-Modal*. We use numerical representation as our main taxonomy for non-theoretical applications because it is the bottleneck for extracting topological features from text.

Finally, we will first discuss the principles behind TDA and the two main techniques employed for TDA feature extraction: *Persistent Homology* and *Mapper*. In addition, we will discuss the selection criteria and taxonomy development of the survey, both the theoretical and non-theoretical approaches, and discuss interesting findings, open problems, and future directions.

2. TOPOLOGICAL DATA ANALYSIS

Topology is defined as “*the study of geometric properties and spatial relations unaffected by the continuous change of shape or size of figures*,” (Oxford Dictionary). TDA is then a collection of powerful techniques that can quantify the shape and structure of data [100]. Two main techniques are used to extract TDA features: *Persistent Homology* and *Mapper*.

2.1 Persistent Homology

Persistent Homology (PH) [38] 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 [100]. Vietoris-Rips complex is a way to build simplicial com-

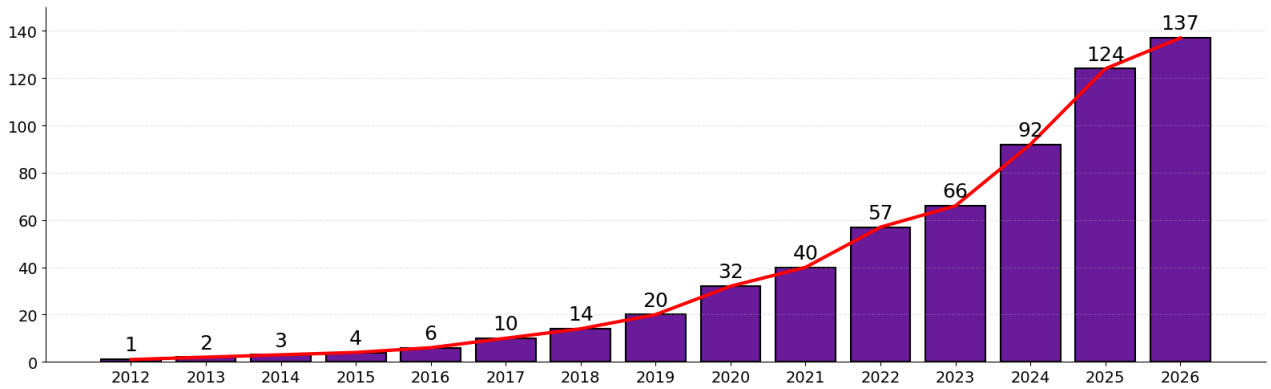


Figure 3: Cumulative number of TDA papers for NLP published from 2012 to April-2026.

plexes 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 *betti numbers* (β_d , d -dimension). Holes in the 0-dimension (β_0) is represented as one vertex, 1-dimension (β_1) is represented as an edge, 2-dimension (β_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. Persistence is defined as the length of time it took a feature to disappear or die ($death - birth$). The *death* is recorded with the radius value at which the points overlap. Lastly, TDA features are typically visualized in a persistence diagram, which is a visual representation of the *birth* (x-axis) vs. *death* (y-axis) features. Other ways of visualizing TDA features include persistence images [1] and barcode plots [54]. Figure 2 illustrates an example of the process of extracting TDA features using PH. In terms of application, PH has been used to extract novel features to complement existing NLP representations and improve various classification performances [35, 157, 170].

Persistent Homology. *This is a TDA technique that studies the deformation of “holes” in different dimensions. Using PH, we can track when features appear and disappear and visualize these features, usually in a persistence diagram. This process allows us to find the true structure of data, typically devoid of noise.*

2.2 Mapper

Mapper is a dimension reduction clustering technique for visualizing TDA-extracted topological structures/signatures. It was proposed by Singh et al. [138] and has been used extensively to visualize topological structures in data to create visually pleasing figures. In addition, Mapper figures have been used to interpret model performance through data probing [23]. The Mapper algorithm works

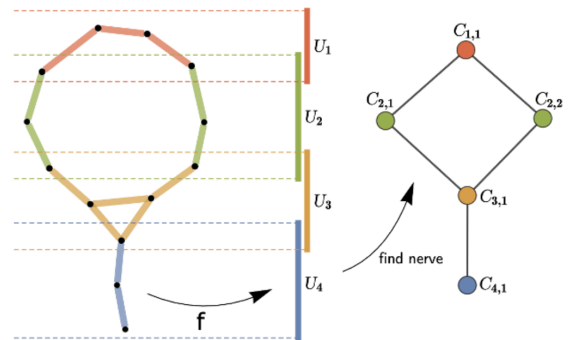


Figure 4: Illustration of Mapper from [101]. The filter function f is a height function, which is a projection onto the y-axis. The cover of the projected space is the four intervals U_i . The Mapper graph on the right is a result of applying the rest of the Mapper algorithm and clustering each preimage in the nearest neighbor.

in four steps³ (Figure 4) following the instructions of [101]: (1) Transform the data 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 [93], UMAP [94], 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. There is an edge between two vertices if the corresponding clusters share common points. Points in the same neighborhood are clustered using a defined clustering technique, such as DBSCAN [42] 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. In this survey, we will discuss how several researchers use Mapper to explain or enhance several phenomena in NLP tasks.

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

Mapper. This is a TDA technique that visualizes the graphical representation of data in order to capture the intrinsic structure. It is very useful for preserving data structure and creating visually pleasing plots, which can be investigated manually to find insights.

3. SURVEY SCOPE

3.1 Selection Criteria

In order to find all NLP papers that applied TDA, we manually searched on Google Scholar using key terms such as *text mining persistent homology*, *language model topological data analysis*, etc., checking related articles of the relevant papers, their cited papers, and different combinations of all three methods. After, obtaining over 60 papers initially, we started creating a taxonomy and categorizing the papers. Initially, we focused on TDA applications in textual data, but as we searched, we found several applications in speech, and collected such papers. Finally, we removed papers that did not apply TDA to text or human speech data. Papers removed, either applied TDA to a graphical representation of reddit social networks, applied non-TDA topological techniques, or applied TDA to non-speech audio data. Using these criteria we selected only papers that fit schema and collected the rest following the same schema.

3.2 Taxonomy Development

Based on the papers selected for the survey, we were able to categorize the applications of these papers into two approaches - *theoretical* and *non-theoretical* applications:

Theoretical applications of TDA in NLP: These focus on understanding, characterizing, or proving properties of language and its representations through the lens of topology. They are less about immediate performance gains and more about insight. This application aims to answer the question - "What do the shapes of embedding spaces tell us about language itself and our models of it?" Example - Analyzing embedding spaces: Using persistent homology to study whether semantic clusters, or syntactic structures, correspond to stable topological features, and finding out what that tells us about language.

Non-theoretical (practical) applications of TDA in NLP: These treat TDA as a tool for solving tasks, regardless of whether deeper linguistic/topological insights are obtained. The emphasis is on utility. This application aims to answer the question - "How can topological summaries directly help with applied NLP tasks?" Example - Feature engineering: Augmenting classifiers for sentiment analysis, topic detection, or authorship attribution with topological signatures.

4. THEORETICAL APPROACHES

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 data. See Figure 5 for an illustration of this pipeline. 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) in language, analyze and visualize the *topology of topic evolution* within texts, and extract the

topological shape of words. See Table 1 for the theoretical approaches and Figure 6 for the flowchart illustrating the taxonomy of theoretical applications of TDA in NLP tasks. In essence, these theoretical topological methods provide a conceptual bridge between linguistic theory and mathematical topology.

4.1 Topological Space

4.1.1 Semantic Topological Space

A **semantic topological space** is a conceptual framework used to represent and analyze the relationships between the 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).

Karlgren et al. [74] 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 [74]. 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 [74]. This suggests that semantic topological space can be better captured with richer and denser data than with sparser data.

In addition, Cavaliere et al. [24] extracts main concepts from the texts by probing the context-aware semantic topological space built with simplicial complexes. Gromov et al. [58] builds a semantic space with bigrams and trigrams of Word2vec embeddings of English and Russian languages to ascertain how distinct the languages are. Next, they use these findings for bot detection. Sakib et al. [126] proposes a metric - *S M Nazmuz Sakib Topological Affix Isometry Index* to measure the structural preservation of the semantic space after the addition of affixes (i.e., -ness, -er, un-). They use persistent homology to create word manifolds to measure how affixes preserve (i.e., isometry) or distorts the meanings and relationships of the words it attaches to. Christianson et al. [29] uses persistent homology to structure the semantic networks from mathematical concepts in college-level linear algebra texts. They find that the networks show strong core-periphery architecture, where concepts are dense and sparse periphery for concepts presented throughout. Their results could inform the optimal design of principles for textbooks.

Furthermore, Wagner et al. [165] 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 [165]. However, in 2012, they found that their technique was unsuccessful due to computational costs, which is a testament to how the field of NLP has improved so that we now have more tractable solutions, such as dimensionality reduction algorithms [94], and TDA packages (Ripser [12], Sklearn-TDA [133], PHAT [13], pytorch-topological⁴), as well as compute resources to efficiently construct topological spaces from large or complex data.

⁴<https://github.com/aidos-lab/pytorch-topological>

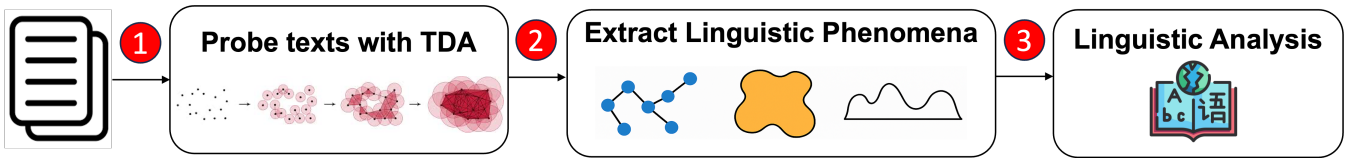


Figure 5: Illustration of the theoretical approaches researchers have employed to (1) probe texts, (2) extract TDA features, (3) use these features to explain or confirm known linguistic phenomena.

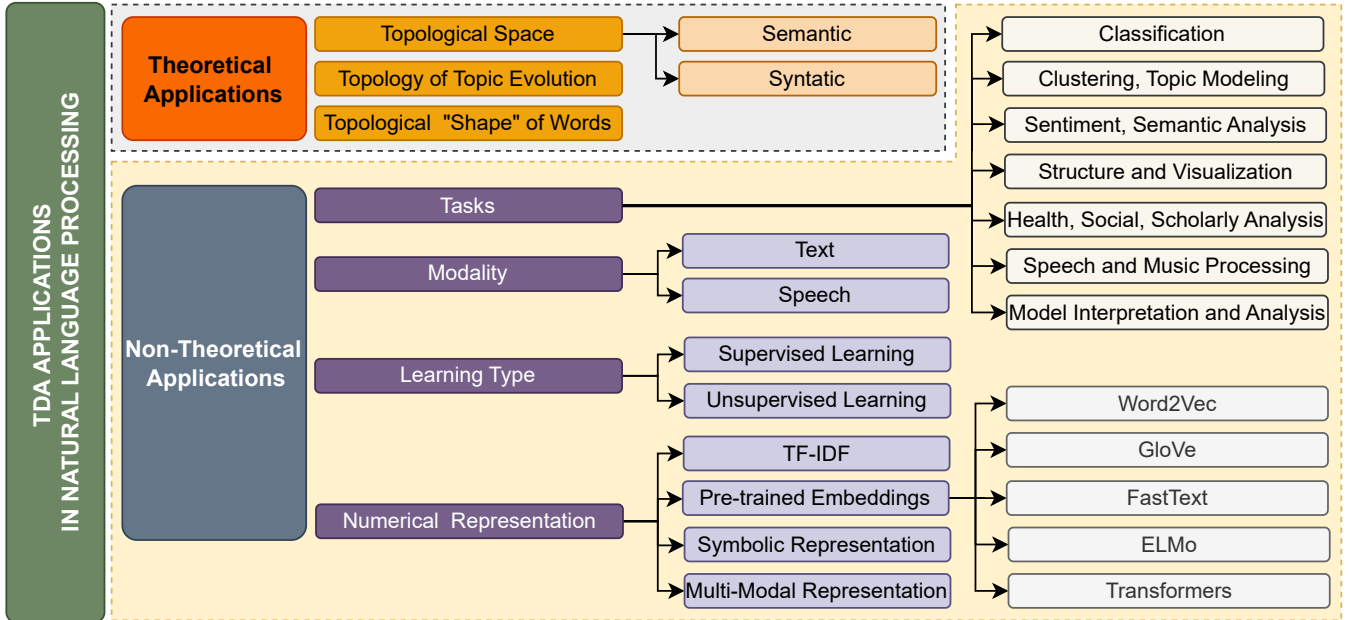


Figure 6: Taxonomy of Topological Data Analysis (TDA) for Natural Language Processing (NLP) Applications

Table 1: Theoretical Applications of TDA in NLP

Name	Category	Task	TDA Technique
[74]	TS-Sem	Identify topical density of the space	Mapper
[24]	TS-Sem	Extracts main concepts from text	Persistent Homology
[165]	TS-Sem	Analyzes document similarities	Persistent Homology
[58]	TS-Sem	Analyzes holes in English and Russian languages	Persistent Homology
[126]	TS-Sem	Measures how word affixes (i.e., -ness) distort or preserve semantic structure	Persistent Homology
[29]	TS-Sem	Creates a topological structure for college-level linear algebra texts	Persistent Homology
[113]	TS-Syn	Analyzes syntactic parameters of different language families	Persistent Homology
[114]	TS-Syn	Explains linguistic structures with homoplasmy phenomena	Persistent Homology
[36]	TSW	Investigates the "shape" of language phylogenies in the Indo-European language family	Persistent Homology
[45]	TSW	Captures grammatical structure expressed by corpus using <i>word manifold</i>	Persistent Homology
[34]	TSW	Analyzes shapes of South American languages: Nuclear-Macro-Jê & Quechuan families	Persistent Homology
[20]	TSW	Analyzes topological similarity between Tifinagh and Phoenician scripts	Persistent Homology
[129]	TTE	Topic evolution within documents	Persistent Homology

Note: TS-Sem = Topological Space (Semantic); TS-Syn = Topological Space (Syntactic); TTE = Topology of Topic Evolution; TSW = Topological "Shape" of Words.

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.*

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. [113] analyzes how syntactic parameters are distributed over different language families, including Indo-

4.1.2 Syntactic Topological Space

European, Niger-Congo, Austronesian, and Afro-Asiatic families. For instance, features in β_0 capture the subdivision into historical, and features in β_1 capture syntactic differences between branches of families of languages, as well as the syntactic influences between them [113]. They investigate the syntactic topological structures of language families, specifically Indo-European, Niger-Congo, Austronesian, and Afro-Asiatic families. Port et al. [113] shows that the three persistent connected components (β_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 [113].

Similarly, Port et al. [114] probes the interpretability of the syntactic topological space by introducing *homoplasmy* phenomena to explain persistent loops. Homoplasmy phenomena in syntax are observed when dissimilar languages exhibit syntactic similarities [114]. Findings reveal that the Indo-European family languages - Czech, Lithuanian, Middle Dutch, and Swiss German have the same homoplasmy phenomena [114] due to the appearance of persistent loops in these languages. This is because Middle Dutch and Swiss German are similar, while Czech and Lithuanian are so different from them, making the homoplasmy phenomenon the most reasonable explanation [114].

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, a novel application of this framework could include the discovery of new linguistic phenomena within syntactic structures.*

4.2 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.

Sami et al. [129] utilizes TDA to visualize the relationship between words in a text block, words in a corpus, and text blocks in a corpus. Text blocks represent a chapter/section in a book, a document in a media corpus, and a webpage in a web corpus [129]. 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. Then, the peripheral nature of the text block and corpus can be visualized using these features. With the Local context features, dimension reduction is achieved by stemming the prefixes and suffixes of words. For the Global context features, word movement is captured, which analyzes topic evolution. 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 predictions by comparing the topology of the topic evolution in the original vs. the perturbed texts.*

4.3 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 et al. [36] 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 [17], 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. [113, 114]’s findings; and (2) TDA methods can successfully capture aspects of the shape of language [36].

Similar to [36, 113, 114], Dong et al. [34] 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 et al. [34] 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 [34].

Fitz et al. [45] 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 [45]. Finally, Bouazzaoui et al. [20] 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 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 words.*

5. NON-THEORETICAL APPROACHES

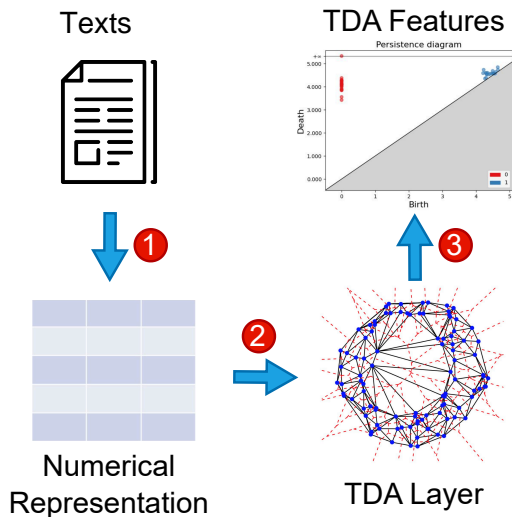


Figure 7: Illustration (inspired by [157]) of the Non-theoretical approach of using TDA as a feature extractor in NLP with three steps: (1)-extracting numerical representations, (2)-reformatting for TDA’s inputs, and (3)-extracting TDA features.

There are several ways to categorize the applied/non-theoretical TDA applications in NLP. These applications can be categorized by *task*, *learning type*, *data modality*, *TDA technique*, and *numerical representation*. We observe that categorizing 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 7 for an illustration of the pipeline for extracting TDA features from numerically represented texts. In addition, 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. Part of Figure 6 illustrates the taxonomy of non-theoretical applications of TDA in NLP tasks. See Table 3 and 4 in Appendix for the list of non-theoretical approaches.

Learning types have supervised [40, 83], and unsupervised [144, 18]; *Modality* has text [150, 75], and Speech [131]; and *TDA techniques*, have Persistent Homology [149, 27], and Mapper [64, 40]. For data modality, 90% of applications are concentrated in Text, and for TDA techniques 89% of applications are concentrated in persistent homology. In this survey, we focus on two other broad categories, *Tasks* and *Numerical Representation*, where the connections to prior work are richer.

5.1 Tasks

Tasks in this context are defined as the problem for which a solution is attempted. We categorize these NLP problems that TDA practitioners have attempted into seven categories below. We refer the readers to Figure 8 (left) for the distribution of the number of publications per task.

1. **Classification:** The most popular application is deepfake text detection [89, 154, 80, 79, 157, 168].
2. **Clustering and Topic Modeling:** The most popular application is document clustering and topic modeling [64, 59].

3. **Sentiment and Semantic Analysis:** The most popular applications are linguistic/grammatical acceptability [27, 72], word sense induction & disambiguation [121, 147], and polysemy word classification [73, 136].
4. **Structure and Visualization:** The most popular is using Mapper to visualize model hidden weights [48, 120].
5. **Health, Social, and Scholarly Analysis:** Since this is not a popular application for TDA, the most interesting applications are - prediction of epidemics [110] and categorization of lonely people [39].
6. **Speech Processing:** The most popular applications are studying vocalizations [19, 18].
7. **Model Interpretation and Analysis:** The most popular applications are model probing to reveal behavior in hidden weights [75, 57].

5.2 Numerical Representation

See Figure 8 (right) for the distribution of the number of applications for each numerical representation.

5.2.1 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 Figure 7. For instance, SIFT, a persistent homology-based model with TF-IDF, is developed to differentiate between child and adolescent writings [187]. 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 task*, such as deepfake text detection [89], presidential election speech attribution [68], distinguishing between languages by averaging the persistence landscapes [143], age group categorization of lonely people [39], and movie genre classification [35, 137]. Additionally, Elyasi et al. [40] compares the two popular TDA approaches - Persistent Homology and Mapper to classify Persian poems. Also, using Mapper for the *structure & visualization tasks*, Maadarani et al. [91] explains linguistic properties in poetry writing styles, and Van et al. [159] interprets NLP model behavior. Lastly, we observe applications in the *clustering & topic modeling task* - keyphrase extraction [59], text summarization [78], and twitter topic detection [149]; *sentiment & semantic analysis task* - legal entailment [134], and sentiment analysis of movie reviews [96].

5.2.2 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 [97]. Thus, we observe applications in the *structure & visualization task*, where Haghhigh-Atkhah et al. [61] 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 topologically-enhanced search engine using Mapper [30], measuring distance between the literary style of Spanish poets - Francisco de Quevedo, Luis de Góngora, and Lope de Vega [105], detecting narrative shifts in media discourse [10], distinguishing news articles and poems by detecting plot holes [5], analysis of contradictions within texts [170], and word sense induction and disambiguation [121, 147]. For the *classification task*, researchers detect fraudulent papers [155], and topological loops in logical statements [156].

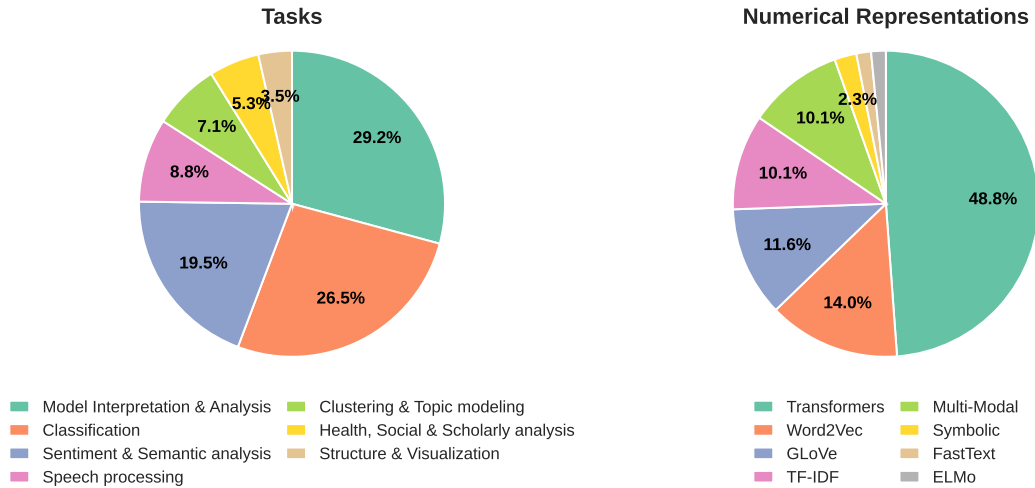


Figure 8: Distribution of the number of publication for each task (Left) and numerical representation type (Right).

Furthermore, we observe applications in the *health, social, and scholarly analysis* task - disease prediction from epidemic curves [110], and Yadav et al. [173] uses doc2vec to create top2vec (i.e., topological features \rightarrow vec) of publication documents to use the holes found with persistent homology to determine missing documents vs. innovative research. Additionally, [64, 169] perform *topic modeling tasks*. Finally, for the *model interpretation and analysis* task, Feng et al. [44] uses both topological and geometrical features to investigate the quality of LLM-enhanced data augmentation, Sun et al. [145] derives the correlation between sentence vectors and their semantics, and Yessenbayev et al. [175, 176] compares the semantic alignment of text and speech embeddings for text-speech pairs.

GloVe Embeddings. GloVe or Global Vectors for Word Representation is another technique for numerically representing texts as embeddings. Topological features extracted from GloVe embeddings have been applied to the following tasks - *classification task*, which include author attribution of novelists [51], fake news detection [32], and deepfake text detection [89, 3]; *sentiment & semantic analysis task*, which include document categorization [52], and capturing circles in circular arguments [156, 181]; *model interpretation and analysis*, where both Haim et al. [62] and Michel et al. [96] compare text representations & embeddings, Spannaus et al. [144] explains model performance, and Zadrozny et al. [180] tests the manifestation of intelligence and understanding in models; and *health, social, and scholarly analysis task*, which include social anxiety detection [21], and keywords extraction of scholarly documents [103].

FastText Embeddings. FastText embeddings are built on the Word2Vec approach by incorporating subword information, improving the representation of rare words, and allowing for embedding out-of-vocabulary words [17]. This type of embedding is not a popular feature extractor that researchers employ to enhance topological features, as only two tasks are attempted - *sentiment & semantic analysis task*, which include polysemy word classification [73, 150, 136], and word sense induction & disambiguation [73]; and *text classification*, where Tymochko et al. [155] detects fraudulent papers.

5.2.3 Pre-trained Contextual Embeddings

Transformer Embeddings. Researchers have evaluated the strength of the TDA features extracted from Transformer-based [161] 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 Transformer 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 task*, specifically for deepfake text detection [154, 80, 168, 60], fake news detection [83], and TEDtalk public speaking ratings classification [31]. Additionally, we observe applications to the *topic modeling task* [22, 65]. Next, for the *model interpretation & analysis task*, Gourgoulia et al. [57] probes LLMs to estimate class separability of text datasets, and Proskura et al. [115] uses topological information from encoder models to select the best models to use when building an ensemble. In addition, Rair et al. [118] uses Mapper to visualize how fine-tuning of models like RoBERTa-Large restructures the embedding space into modular, non-convex regions to align with model predictions. This technique aims to visualize the geometry and topology of when annotators agree and disagree [118]. Furthermore, Arun et al. [6] employs TDA for the *structure & semantic task*, detecting controversial vs. non-controversial political discourse by capturing the shifts in discourse for controversial data. Similarly, Meng et al. [95] performed a *semantic task* of using persistent homology to augment and improve personalized web search. Next, Chandra et al. [25] performs a *health analysis task* using persistent homology to track mental health journeys in online communities. Finally, Rathore et al. [120] performs the *structure & visualization task* in combination with the *model interpretation & analysis task* by visualizing the training process of transformer-based models.

Hidden Weights. *Classification task* include deepfake text detection [157, 122], language translation [7], and code attribution [92]. Next, Garcia et al. [48] explores a combination of the *sentiment & semantic analysis* and *structure & visualization* tasks by using Mapper to visualize polysemous words in the hidden representations of the BERT transformer model. Bensalem et al. [15]

performs a *sentiment analysis* task by using the sentiment scores of original vs. translated texts extracted from a Transformer-based model. They represent these scores in a time series form and use zigzag persistent homology to detect sentiment shift in translated texts. Similarly, Goshev et al. [56] investigates the semantic topology of sentences encoded by transformer embeddings. Zhang et al. [184] uses persistent homology to improve text summarization by capturing the global structure of texts. Alexander et al. [4] combines the *health analysis* and *visualization* tasks to visualize GPT-3’s embeddings of hate speech, misinformation, and psychiatric disorder texts with Mapper. Ruppik et al. [125] performs the *clustering & topic modeling task* through dialogue term extraction.

The rest of the applications attempt the *model interpretation & analysis task*, making it the most popular task. Gardinazzi et al. [50] proposes a novel metric - *persistence similarity* to prune redundant layers in LLMs. Zheng et al. [186] uses the same technique to prune large vision-language models. Balderas et al. [11] proposes Persistent BERT Compression and Explainability (PBCE) to compress BERT by pruning redundant layers. Sun et al. [145] probes the correlation between sentence vectors and their semantics. Garcia et al. [49] performs zero-shot model stitching by employing *topological densification* (i.e., creating a topology-aware loss function). Huang et al. [69] uses a topology-aware loss function to improve prompt tuning. Athreya et al. [8] proposes a framework - HOLE (Homological Observation of Latent Embeddings for Neural Network Interpretability) to visualize the latent space of BERT for Named Entity Recognition (NER) task.

Next, several researchers probe the reasoning processes of LLMs: (1) Tan et al. [146] captures the geometry and topology of reasoning in LLMs using a mathematics examination dataset to capture step-by-step reasoning for solving non-trivial word problems; (2) Li et al. [85] and Zhang et al. [183] probe the reasoning process of the latent space of LLMs when using the chain-of-thought (CoT), tree-of-thought, and graph-of-thought prompts; (3) More et al. [98] proposes Enhanced Dirichlet and Topology Risk (EDTR), which is a novel decoding strategy that leverages persistent homology and Dirichlet-based uncertainty quantification to calculate LLM confidence; (4) Zhang et al. [185] proposes GHS-TDA, a reasoning technique to improve CoT reasoning by constructing a semantically enriched global hypothesis graph using persistent homology to capture global structures and remove redundancies; and (5) Ishimtsev et al. [71] proposes a theory that consciousness might appear when a thinking process loops back on itself with CoT prompting of LLMs. Using this theory, Ishimtsev et al. [71] implements a topology analogy, guided by persistent homology to define: normal reasoning as a *straight path*, self-reflection as a *loop*, and consciousness as a *stable loop around a hole* [71].

Finally, in support of the *model interpretation & analysis task*, several researchers probe the embedding space of LLMs to quantify the topology of the latent space. Fitz et al. [47] measures the topological complexity (known as *perforation*) of the hidden representation of LLMs to understand their topological shapes. Also, Fitz et al. [46] investigates the topological structure of the brain of ChatGPT concerning its notion of fairness. Chauhan et al. [26] proposes a novel scoring metric - *persistence scoring function* which captures the homology of the hidden representations of BERT. Fay et al. [43] investigates the differences in the topological structure of the latent space of adversarial vs. non-adversarial texts in LLMs. Kudriashov et al. [77] probes BERT’s hidden weights on new grammatical features, known as *polypersonality*. Lastly, Yan

et al. [174] builds an *Explainable Mapper* framework that uses two mapper agents to probe the embedding space of language models, and generate readable linguistic explanations using summarization, comparison, and perturbation operations.

Attention Weights. Attention weights extracted from BERT and its variants (e.g., RoBERTa) have been transformed to both directed and undirected graphs, on top of which different TDA features are extracted for the *text classification task* such as deepfake text detection [79], LLM hallucination detection [14], Code-LLM hallucination detection [162], robustness evaluation of TDA features [108], authorship attribution of Japanese texts [127], out-of-distribution detection (OOD) [112, 108], and vulnerability detection in code [141]. Additionally, this framework is applied to the *sentiment & semantic analysis task*, specifically on human linguistic competence (i.e., grammatical acceptability judgment) [27, 116, 108], dialog term extraction [164], and document coherence [72]. Similarly, with the same framework, we observe a *speech processing* application [153]. Finally, researchers attempt the *model interpretation & analysis task*, where Kostenok et al. [75] estimates uncertainty in encoder models; Samaga et al. [128] characterizes the occurrence of hallucination in LLMs; Tsai et al. [152] performs the same experiments to investigate the effects of sandbagging and code-injections in language model latent space; Varadarajan et al. [160] examines why GPT-2 significantly misrepresents gender and race identity categories by identifying which attention heads are responsible for the misclassification of specific identity groups; and Proskura et al. [115] performs dynamic weighting for building ensemble models.

ELMo Embeddings. ELMo embeddings are a type of word representation that captures both the meaning of words and their usage in context [109]. Similar to other embeddings, ELMo has also been leveraged to extract topological features. Tymochko et al. [155] performs *text classification* to detect fraudulent papers by examining their titles and abstracts. This is done in comparison of other embeddings (Word2Vec, GloVe, FastText, and Frequency Time Series) to determine the best embeddings to extract strong topological features. Similarly, Alimpiev et al. [5] compares using ELMo, GloVe, Word2Vec, and BERT embeddings to perform semantic analysis on news articles and poems.

5.2.4 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)*, *principles of speech sound coding (PSSC)*, and one-hot encoding, of which topological features are then extracted.

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 Ukrainian tongue twisters [179, 76]. These applications attempt two tasks,

S1: Captures local and global structure Models both neighborhood-level and corpus-level linguistic structure.	L1: High computational cost Can be expensive on large corpora and high-dimensional constructions.
S2: Robustness to noise Persistent homology filters small noises while preserving salient structure.	L2: Interpretability challenges Topological summaries are often less intuitive for NLP practitioners.
S3: Effective in low-resource settings Remains informative with limited, noisy, or sparsely labeled data.	L3: Limited software support Many TDA toolkits require adaptation for text and language data.
S4: Reveals complex relationships Uncovers structural, semantic patterns missed by feature-based methods.	L4: Limited pipeline integration TDA is not yet commonly supported in standard NLP workflows.
S5: Reduced manual feature design Structural signals that complement engineered or statistical features.	L5: Limited adoption Few large-scale NLP applications and industrial deployments exist.
S6: Compatible with embeddings Applies naturally to word, sentence, and contextual embeddings.	L6: Requires specialized expertise Effective use often requires knowledge of both topology and NLP.
S7: Geometric insight into language Characterizes syntactic, semantic structure from a geometry perspective.	L7: Lack of standardized evaluation No widely accepted benchmark exists for TDA-based NLP methods.

Figure 9: Strengths (S) and limitations (L) of applying TDA in NLP.

Category*	Method	Description
Linear	PCA	Reduces dimensionality while preserving variance through orthogonal transformations.
Linear	MDS	Projects high-dimensional data into lower dimensions by preserving pairwise distances.
Nonlinear	t-SNE	Projects high-dimensional data into 2D or 3D while preserving local relationships.
Nonlinear	UMAP	Similar to t-SNE, but faster and often better at preserving global structure.
Nonlinear	Isomap, LLE	Captures intrinsic structure in high-dimensional data using graph-based techniques.
Neural	Autoencoders	Learn compressed data representations through encoding and decoding.
Neural	Geometric Deep Learning	Applies neural networks to non-Euclidean spaces such as graphs and manifolds.
Graph	Spectral Clustering, GNNs	Uses graphs to model relationships and structure within data.
Clustering	DBSCAN, K-Means, HDBSCAN	Groups similar data points based on distance or density.
Kernel	SVM, Kernel PCA	Uses non-linear mappings to extract complex structures in data.
Geometric	Delaunay Triangulation, Convex Hull	Uses geometric techniques to extract data characteristics by analyzing spatial boundaries.

(*) **Linear**: Linear Projection; **Nonlinear**: Nonlinear Projection; **Neural**: Neural Network-based Methods; **Graph**: Graph-based Methods; **Clustering**: Clustering Methods; **Kernel**: Kernel Methods; **Geometric**: Geometric Methods.

Table 2: Alternatives to Topological Data Analysis in NLP

sentiment & semantic analysis and *speech processing*, respectively. Yurchuk et al. [179] uses the PLC to create word embeddings for Ukrainian tongue twisters and extract topological features from such embeddings with persistent homology. This is to distinguish tongue twister from a simple narrative sentence using support vector machine and decision tree classifiers. Similarly, Kovaliuk et al. [76] uses PSSC for classifying spoken Ukrainian tongue twisters. Additionally, Escolar et al. [41] uses one-hot encoding to represent cooking recipes numerically. Using these symbolic representations of recipes, they use the persistent homology’s concept of holes to create new recipes, which were implemented and confirmed to be acceptable by a sensory evaluation study [41].

5.2.5 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 most popular task performed by researchers is *speech processing*, where applications include - studying human vowels and infant vocalizations [19, 18], speech recognition [178, 81], emotion recognition from audio speech [55, 119] and audio in videos [104], depression detection from audio clips [148], recognizing voiced and voiceless consonants in speech [188].

Next, we observe several applications with Vision-Language Models (VLMs), which also attempt the *model interpretation & analysis task* - integrating Representation Topology Divergence (RTD) with

the loss function to align the topological structures of image and text representations during tuning [67]; implementing a topological approach to align the image and text latent manifolds in VLMs [182, 117, 172]; assessing the adversarial robustness of VLMs by measuring topological consistency [163]; assessing the topological alignment in VLMs between images and multi-lingual text (i.e., French, Spanish, Russian, etc.) [177]; and multimodal recommendations improvements [9].

6. BENEFITS AND CHALLENGES OF TDA

The use of TDA in NLP presents both clear opportunities and important challenges, many of which also extend to adjacent topological approaches. Figure 9 summarizes the main strengths and limitations of applying TDA across NLP tasks. Among its most salient strengths is its effectiveness in low-resource settings, which makes it particularly attractive for noisy, limited, and heterogeneous data. Although TDA has been shown to be a powerful mathematical technique by the numerous applications discussed above, there are alternatives to TDA (Table 2). While these alternatives have shown great utility in a variety of tasks, such as dimension reduction [16], TDA is the only technique that can extract not only local but global features. More broadly, TDA provides a principled way of capturing local and global textual structure, uncovering complex relationships, and maintaining robustness to noise. At the same time, its wider adoption remains constrained by several practical barriers, including high computational cost, limited interpretability, relatively immature software support, and the need for specialized ex-

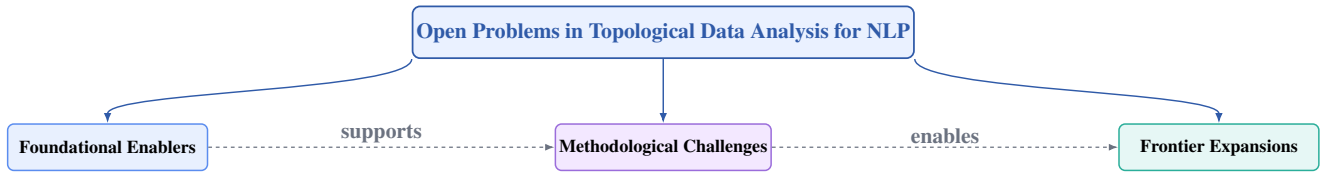


Figure 10: A high-level taxonomy of open problems in applying topological data analysis to NLP, organized into foundational enablers, methodological challenges, and frontier expansions.

pertise.

However, while persistence diagrams are not intuitively interpretable, in the context of NLP, 0D features (connected components) corresponds to clusters of semantically similar documents (topics), while 1D features (loops) can capture transitional or overlapping themes where documents form continuums rather than discrete groups. These structures help identify stable topic groupings, semantic relationships, and potential outliers, with robustness to noise. Compared to PCA or t-SNE, persistence diagrams provide complementary insights. PCA captures linear variance and may miss nonlinear structure, while t-SNE emphasizes local clustering but can distort global relationships. In contrast, persistence diagrams analyze the data in its original space and quantify both local and global structure across scales.

These strengths are further reflected in a broader body of work on related topological approaches in NLP. Beyond Persistent Homology and Mapper, researchers have explored methods grounded in simplicial homology, morse theory, homotopy, and other notions of connectedness and shape to extract structural information from language data. Although these methods are often less formalized than Persistent Homology, they similarly demonstrate the value of topological thinking for modeling semantic structure, discourse organization, stylistic variation, and model behavior. For example, topological notions such as the connected component dimension (β_0) have been used to assess document coherence through semantic connectedness [28], while more recent work has extended these ideas to dialogue semantics [130] and to topological semantic spaces for studying sociolinguistic phenomena in LLMs [70]. Other studies treat text as an object with shape, using topology to characterize stylistic and linguistic patterns [90] or to quantify semantic differences between real and fake news through thematic complexity and connectedness [135]. More recently, topological ideas have also appeared in safety-oriented LLM research, where homotopy-based methods are used to obfuscate malicious prompts or optimize jailbreak attacks [84, 167].

Taken together, these studies suggest that the appeal of TDA in NLP is part of a broader methodological advantage of topological approaches: their ability to reveal structural regularities in language that are often difficult to capture through conventional feature-based or purely predictive methods alone. Similarly, limitations of TDA, such as computational expense, interpretability challenges, and limited software, highlight the practical obstacles that must be addressed before topological methods can see broader adoption in NLP. While ongoing advances in algorithms, software, and computing infrastructure offer reasons for optimism, these strengths and challenges together point to several promising directions for future work, which we discuss in Section 7.

7. OPEN PROBLEMS

Although TDA has shown growing promise in NLP, many challenges remain before its full potential can be realized. In this section, we organize the main open problems and future directions into three broad categories: foundational enablers, methodological challenges, and frontier expansions. Together, these categories highlight how researchers can better leverage the strengths of TDA while addressing its current limitations and risks (Figure 10).

7.1 Foundational Enablers

LLM-Assisted TDA Code Generation. 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, 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. [87] proposes using 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 [87]. Similarly, experts can develop specialized code generators, such as fine-tuning models like Code-Llama [124] on TDA concepts. In addition, Li et al. [86] proposes to benchmark LLM as topological thinkers by giving them tasks to implement persistent homology on. The idea is to create *LLM4PH*, an LLM for persistent homology. Therefore, we observe that we are closer to creating a dedicated LLM for TDA code generation which could significantly lower the barrier to entry, encouraging the NLP community to explore TDA applications more innovatively.

Topology-Linguistics Alignment. TDA can come across as not intuitive. Even though it has the potential to be applied to interpret different behaviors of modern NLP models, there is still a need for theoretical approaches that better tie TDA features to linguistic phenomena but intuitively. For example, Draganov et al. [36] investigates the shape of words and their embeddings in Indo-European languages and find similar conclusions to Port et al. [113, 114], which investigate the syntactic topological space of such languages. They find that TDA features represent historical facts, such that languages clustered closely together are similar or influenced by each other. These applications show how TDA can be used to reveal and confirm linguistic phenomena. However, we currently observe only 13 theoretical TDA works in NLP, compared to over 100 non-theoretical ones. Thus, we need more theoretical TDA approaches, as it is impractical to achieve the depth and understanding of performance from a topological perspective without further investigations.

7.2 Methodological Challenges

Interpretability of TDA Results. 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 [132], where the structure is clearly apparent. Consequently, the interpretation of TDA features 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 extracted from the model; and (2) explain model performance by using TDA to probe the prediction space or data. Either task requires 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 β_0 , and β_1 's representations to different properties of natural texts such as coherency, and writing style. This can be done either through visualizing the latent space of a model [120, 174, 128] or probing the latent space of models with TDA [144, 171, 142, 174].

TDA for Understanding NLP Model Behavior. Existing studies suggest that TDA is promising for interpreting the behavior of NLP models; however, current efforts remain scattered across tasks, representations, and model settings. In particular, TDA has been used to probe neural representations and make language models less opaque by characterizing the geometry and topology of their learned hidden space. For example, existing work spans representation geometry and alignment across modalities, including text embeddings, sentence semantics, text–speech alignment, adversarial versus non-adversarial latent spaces, and image–text alignment [62, 145, 175, 176, 43, 163, 177, 152]. Other studies use TDA to probe hidden states and model weights, revealing structural properties of language model latent spaces and uncovering grammatical or representational patterns [26, 77, 47]. TDA has also been applied to training and optimization dynamics, model-level explanation and diagnostics, model efficiency & safety, reliability, and reasoning-related behavior [120, 49, 69, 98, 185, 144, 57, 174, 115, 50, 11, 180, 46, 75, 128, 160, 44, 118, 146, 85, 71]. Beyond these works, how can TDA move beyond case-specific analysis to become a principled and generalizable framework for interpreting complex model behavior in NLP is of great value for future work.

Improved TDA Feature Extraction and Representation Selection. Unlike some other data modalities that possess an intrinsic geometric structure, texts acquire their “shape” only through their numerical representations. In other words, the topology we observe is not an inherent property of the text itself, but of the embedding method used to encode it. A corpus represented with TF–IDF will therefore exhibit a geometry characteristic of sparse lexical weighting, whereas Transformer-based pre-trained embeddings induce a different shape driven by contextual semantic structure. These differences are partly explained by the distinct linguistic features each representation captures - lexical frequency in the case of TF–IDF, and richer semantic and syntactic information in contextual embeddings. However, the resulting geometric and topological variations are often unintuitive. As a result, it becomes difficult to determine which numerical representation is most appropriate for extracting meaningful TDA features for a given task. To address this challenge, we must develop principled methods for selecting representations that align with the objectives of the analysis. This includes exploring new forms of numerical encoding, such as *symbolic representations* that capture the diversity of textual phenomena across tasks. Equally important is the development of improved strategies for leveraging existing embeddings in ways that enhance the stability, interpretability, and task-relevance of the extracted topological features.

7.3 Frontier Expansions

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 [108, 26, 43, 163]. For instance, Perez et al. [108] shows that their topologically-augmented BERT model is more robust than the vanilla BERT model when tested against perturbations generated by TextAttack [99]. Chauhan et al. [26] also show 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. Recently, we have observed interesting applications that use topology to track the latent space of LLMs before and after adversarial perturbations are introduced [43, 163]. These studies highlight the emerging application of topology in robustness evaluations, which will benefit the NLP, ML, and security communities.

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 5.1, we can see that TDA has been applied to seven non-theoretical NLP tasks. While many of the tasks are interesting, especially the speech processing and health applications, there are still nuanced niche fields that could benefit from TDA. One glaring application is on multi-lingual tasks [62, 49, 177]. Due to the benefits of TDA, which include performing robustly on heterogeneous, imbalanced, and noisy data, its application to multi-lingual tasks is necessary. Other applications include: *Topology-aware neural networks*, *Topological interpretability*, *semantic and syntactic structural analysis*, *forensic authorship*, *multi-modal (e.g., VLMs) analysis*, etc.

Topological Deep Learning for NLP. Due to the benefits of TDA and deep learning, a new niche field is born - Topological Deep Learning (TDL) as “the collection of ideas and methods related to the use of topological concepts in deep learning” [107]. Initially, TDL is described as an ensemble of topological features extracted by TDA techniques such as persistent homology and deep learning features. In this setting, TDL is a traditional deep learning model with extra features (i.e., TDA-extracted features). However, as the field has advanced, a new definition for TDL has emerged - “the collection of ideas and methods related to the use of topological concepts in deep learning” [107]. 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) [37, 106], 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 require 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 [88, 111, 85, 79]. Nevertheless, more research is still needed to validate such an approach.

8. CONCLUSION

Our world is currently experiencing an explosion of data and an 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 limita-

tions. 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 137 papers, which we comprehensively surveyed in this paper. Finally, we conclude that while the applications of TDA in NLP have improved greatly since 2012, there is still room for improvement, specifically in reducing the barrier to entry for non-TDA experts to apply it to their NLP tasks.

9. ETHICAL STATEMENT

This survey highlights emerging applications of Topological Data Analysis (TDA) in Natural Language Processing. While our primary goal is to synthesize existing work, we recognize that several use cases carry important ethical considerations and dual-use risks. Topological methods can inadvertently expose latent sensitive attributes (e.g., dialect, health cues, authorship), enabling re-identification or profiling even when data is anonymized. Applications in speech, emotion, and health-related domains further raise fairness, consent, and equity concerns, particularly for minority groups and low-resource languages. Therefore, we emphasize the need for bias and robustness audits, careful data governance and licensing, and privacy-preserving mechanisms when sharing derived features. Responsible release practices, such as restricting code that enables circumvention, conducting red-team evaluations, and requiring IRB or ethics review for clinical or surveillance-adjacent uses, are essential. Finally, given the computational demands of TDA pipelines, their environmental impact should be considered as well.

10. ACKNOWLEDGMENTS

The authors thank Dr. Charlie Dagli for his encouragement, reading the paper drafts, and providing invaluable recommendations.

References

- [1] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18(8):1–35, 2017.
- [2] M. Alagappan, D. Jiang, N. Denko, and A. C. Koong. A multimodal data analysis approach for targeted drug discovery involving topological data analysis (tda). In *Tumor Microenvironment: Study Protocols*, pages 253–268. Springer, 2016.
- [3] L. Alanís-López, J. P. B. Lafarga, L. R. G. Sánchez, E. I. L. Otañez, A. Ramirez-Cabello, and A. Ucan-Puc. Detection of ai generated texts using deep learning and topological data analysis. In *Mexican Congress on Artificial Intelligence*, pages 69–82. Springer, 2025.
- [4] A. W. Alexander and H. Wang. Topological data mapping of online hate speech, misinformation, and general mental health: A large language model based study. *PLOS Digital Health*, 4(7):e0000935, 2025.
- [5] E. Alimpiev and V. Myers. Plot holes and text topology. *Stanford CS224N Custom Project*, 2020.
- [6] A. Arun, K. K. Chandra, A. Sinha, B. Velayutham, J. Arora, M. Jain, and P. Kumaraguru. Topo goes political: Tda-based controversy detection in imbalanced reddit political data. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 2616–2625, 2025.
- [7] E. Asriani, I. Muchtadi-Alamsyah, and A. Purwarianti. Topological data analysis for transformer nmt: Exploring the use of cohomology-based persistence landscapes as a representation of global context. In *2025 12th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*, pages 1–6, 2025.
- [8] S. M. Athreya and P. Rosen. Hole: Homological observation of latent embeddings for neural network interpretability. *arXiv preprint arXiv:2512.07988*, 2025.
- [9] K. Bachiri, A. Yahyaouy, M. Malek, and N. Rogovschi. Topological data analysis and graph-based learning for multimodal recommendation. *IEEE Access*, 2025.
- [10] M. M. Bailey and M. I. Heiligman. Detecting narrative shifts through persistent structures: A topological analysis of media discourse. *arXiv preprint arXiv:2506.14836*, 2025.
- [11] L. Balderas, M. Lastra, and J. M. Benítez. A green ai methodology based on persistent homology for compressing bert. *Applied Sciences*, 15(1):390, 2025.
- [12] U. Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, 2021.
- [13] U. Bauer, M. Kerber, J. Reininghaus, and H. Wagner. Phat–persistent homology algorithms toolbox. *Journal of symbolic computation*, 78:76–90, 2017.
- [14] A. Bazarova, A. Yugay, A. Shulga, A. Ermilova, A. Volodichev, K. Polev, J. Belikova, R. Parchiev, D. Simakov, M. Savchenko, et al. Hallucination detection in llms via topological divergence on attention graphs. *arXiv preprint arXiv:2504.10063*, 2025.
- [15] A. Bensalem, M. A. Bensalem, and M. A. Chadli. Detecting token-level sentiment change in text translation through zig zag persistent homology. <https://hal.science/hal-05139630/>, 2025.
- [16] J. A. D. Binnie, P. Dłotko, J. Harvey, J. Malinowski, and K. M. Yim. A survey of dimension estimation methods, 2025.
- [17] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146, 2017.
- [18] G. Bonafos, C. Bourot, P. Pudlo, J.-M. Freyermuth, L. Reboul, S. Tronçon, and A. Rey. Dirichlet process mixture model based on topologically augmented signal representation for clustering infant vocalizations. In *Proc. Interspeech 2024*, pages 3605–3609, 2024.
- [19] G. Bonafos, J.-M. Freyermuth, P. Pudlo, S. Tronçon, and A. Rey. Topological data analysis of human vowels: Persistent homologies across representation spaces. *arXiv preprint arXiv:2310.06508*, 2023.

- [20] H. Bouazzaoui, M. A. Elomary, and M. I. Mamouni. 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, 2021.
- [21] M. Byers. *The Hidden Shape of Data: Topological Data Analysis for Anxiety Detection in Text*. PhD thesis, Texas State University, 2021.
- [22] C. Byrne, D. Horak, K. Moilanen, and A. Mabona. Topic modeling with topological data analysis. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 11514–11533, 2022.
- [23] G. Carlsson. Topological methods for data modelling. *Nature Reviews Physics*, 2(12):697–708, 2020.
- [24] D. Cavaliere, S. Senatore, and V. Loia. Context-aware profiling of concepts from a semantic topological space. *Knowledge-Based Systems*, 130:102–115, 2017.
- [25] J. Chandra, S. K. Navneet, and Y. Zhang. The topology of recovery: Using persistent homology to map individual mental health journeys in online communities. *arXiv preprint arXiv:2602.23886*, 2026.
- [26] J. Chauhan and M. Kaul. Bertops: Studying bert representations under a topological lens. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022.
- [27] D. Cherniavskii, E. Tulchinskii, V. Mikhailov, I. Proskurina, L. Kushnareva, E. Artemova, S. Barannikov, I. Piontkovskaya, D. Piontkovski, and E. Burnaev. Acceptability judgements via examining the topology of attention maps. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 88–107, 2022.
- [28] I.-J. Chiang. Discover the semantic topology in high-dimensional data. *Expert Systems with Applications*, 33(1):256–262, 2007.
- [29] N. H. Christianson, A. Sizemore Blevins, and D. S. Bassett. Architecture and evolution of semantic networks in mathematics texts. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2239), 2020.
- [30] F. Cornell. An explainable topological search engine with giotto-tda. In *gt-da-challenge-2020*, 2020.
- [31] S. Das, S. A. Haque, and M. I. Tanveer. Persistence homology of tedtalk: Do sentence embeddings have a topological shape? *arXiv preprint arXiv:2103.14131*, 2021.
- [32] R. Deng and F. Duzhin. 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, 2022.
- [33] T. K. Dey and S. Mandal. Protein classification with improved topological data analysis. In *18th International Workshop on Algorithms in Bioinformatics (WABI 2018)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2018.
- [34] R. Dong. Linguistics from a topological viewpoint. *arXiv preprint arXiv:2403.15440*, 2024.
- [35] P. Doshi and W. Zadrozny. 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, 2018.
- [36] O. Draganov and S. Skiena. 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, 2024.
- [37] S. Ebli, M. Defferrard, and G. Spreemann. Simplicial neural networks. In *TDA & Beyond*, 2020.
- [38] H. Edelsbrunner, J. Harer, et al. Persistent homology—a survey. *Contemporary mathematics*, 453(26):257–282, 2008.
- [39] B. Effah. *Topological data analysis of open-ended responses*. PhD thesis, University of Cape Coast, 2017.
- [40] N. Elyasi and M. H. Moghadam. An introduction to a new text classification and visualization for natural language processing using topological data analysis. *arXiv preprint arXiv:1906.01726*, 2019.
- [41] E. G. Escolar, Y. Shimada, and M. Yuasa. A topological analysis of the space of recipes. *International Journal of Gastronomy and Food Science*, 39:101088, 2025.
- [42] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.
- [43] A. Fay, I. García-Redondo, Q. Wang, H. Dubossarsky, and A. Monod. The shape of adversarial influence: Characterizing LLM latent spaces with persistent homology. In *The Fourteenth International Conference on Learning Representations*, 2026.
- [44] S. J. Feng, E. M. Lai, and W. Li. Geometry of textual data augmentation: Insights from large language models. *Electronics*, 13(18):3781, 2024.
- [45] S. Fitz. The shape of words-topological structure in natural language data. In *Topological, Algebraic and Geometric Learning Workshops 2022*, pages 116–123. PMLR, 2022.
- [46] S. Fitz. Do large gpt models discover moral dimensions in language representations? a topological study of sentence embeddings. *arXiv preprint arXiv:2309.09397*, 2023.
- [47] S. Fitz, P. Romero, and J. J. Schneider. Hidden holes - topological aspects of language models. In *NeurIPS 2024 Workshop on Symmetry and Geometry in Neural Representations*, 2024.
- [48] J. S. Garcia. *Applications of topological data analysis to natural language processing and computer vision*. PhD thesis, Colorado State University, 2022.
- [49] A. García-Castellanos, G. L. Marchetti, D. Kragic, and M. Scolamiero. Relative representations: Topological and geometric perspectives. In *UniReps: 2nd Edition of the Workshop on Unifying Representations in Neural Models*, 2024.

- [50] Y. Gardinazzi, K. Viswanathan, G. Panerai, A. Cazzaniga, M. Biagetti, et al. Persistent topological features in large language models. In *Forty-second International Conference on Machine Learning*, 2025.
- [51] S. Gholizadeh, A. Seyeditabari, and W. Zadrozny. Topological signature of 19th century novelists: Persistent homology in text mining. *big data and cognitive computing*, 2(4):33, 2018.
- [52] S. Gholizadeh, A. Seyeditabari, and W. Zadrozny. A novel method of extracting topological features from word embeddings. *arXiv preprint arXiv:2003.13074*, 2020.
- [53] S. Gholizadeh and W. Zadrozny. A short survey of topological data analysis in time series and systems analysis. *arXiv preprint arXiv:1809.10745*, 2018.
- [54] R. Ghrist. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75, 2008.
- [55] R. Gonzalez-Diaz, E. Paluzo-Hidalgo, and J. F. Quesada. 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, 2019.
- [56] I. Goshev, P. Sekuloski, I. Chorbev, D. Kitanovski, and V. D. Ristovska. Topology as a lens for semantic organization in transformer embeddings. *INTERNATIONAL SCIENTIFIC JOURNAL "MATHEMATICAL MODELING"*, 2025.
- [57] K. Gourgoulis, N. Ghalyan, M. Labonne, S. Moran, J. Sabelja, et al. Estimating class separability of text embeddings with persistent homology. *Transactions on Machine Learning Research*, 2024.
- [58] V. A. Gromov, Q. N. Dang, and A. S. Erbolova. A language and its holes: The first-order homology of the large-scale geometrical structure of a natural language. *Complexity*, 2025(1):9659172, 2025.
- [59] H. Guan, W. Tang, H. Krim, J. Keiser, A. Rindos, and R. Sazdanovic. A topological collapse for document summarization. In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2016.
- [60] A. Guilinger, E. Best, and V. Awasthi. Topological data analysis for distinguishing human written and ai generated abstracts. *preprints.org*, 2025.
- [61] P. Haghhighatkah, A. Fokkens, P. Sommerauer, B. Speckmann, and K. Verbeek. Story trees: Representing documents using topological persistence. In *Proceedings of the Thirteenth LREC 2022*, pages 2413–2429, 2022.
- [62] S. Haim Meir and O. Bobrowski. Unsupervised geometric and topological approaches for cross-lingual sentence representation and comparison. In S. Gella, H. He, B. P. Majumder, B. Can, E. Giunchiglia, S. Cahyawijaya, S. Min, M. Mozes, X. L. Li, I. Augenstein, A. Rogers, K. Cho, E. Grefenstette, L. Rimell, and C. Dyer, editors, *Proceedings of the 7th Workshop on Representation Learning for NLP*, pages 173–183, Dublin, Ireland, May 2022. Association for Computational Linguistics.
- [63] F. Hensel, M. Moor, and B. Rieck. A survey of topological machine learning methods. *Frontiers in Artificial Intelligence*, 4:681108, 2021.
- [64] W. J. Holmes. Topological analysis of averaged sentence embeddings. Master’s thesis, Wright State University, 2020.
- [65] M. D. Hopp, V. Labatut, A. Amalvy, R. Dufour, H. Stone, H. K. Jach, and K. Murayama. Persistent homology of topic networks for the prediction of reader curiosity. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 28121–28132, 2025.
- [66] M. Horn, E. De Brouwer, M. Moor, Y. Moreau, B. Rieck, and K. Borgwardt. Topological graph neural networks. In *International Conference on Learning Representations*, 2022.
- [67] D. Huang. Topology-aware clip few-shot learning. *arXiv preprint arXiv:2505.01694*, 2025.
- [68] J. Huang. A tda approach of analyzing election speeches with nlp techniques. *Journal of Computing Sciences in Colleges*, 38(3):215–215, 2022.
- [69] Z. Huang, C. Zhang, H. Bian, S. Zhang, C.-l. A. Tai, J. Zhang, C. Qin, J. Qu, Y. Ye, Y. Yang, et al. Optimizing soft prompt tuning via structural evolution. *arXiv preprint arXiv:2602.16500*, 2026.
- [70] I. Ionescu, J. Leung, Y. Siglidis, et al. Generative topolinguistics: Bidirectional interfaces for emergent language topologies. *Antikythera: Journal for the Philosophy of Planetary Computation*, 2025.
- [71] T. Ishimtsev. The hole-in-the-bagel theory of consciousness—a phenomenological study and topodynamic model for emergent coherence in llms. <https://zenodo.org/records/18346699>, 2026.
- [72] S. Jain, R. Singhal, S. Krishna, Y. K. Singla, and R. R. Shah. Beyond words: A topological exploration of coherence in text documents. In *The Second Tiny Papers Track at ICLR 2024*, 2024.
- [73] A. Jakubowski, M. Gasic, and M. Zibrowius. Topology of word embeddings: Singularities reflect polysemy. In I. Gurevych, M. Apidianaki, and M. Faruqui, editors, *Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics*, pages 103–113, Barcelona, Spain (Online), Dec. 2020. Association for Computational Linguistics.
- [74] J. Karlgren, M. Bohman, A. Ekgren, G. Isheden, E. Kullmann, and D. Nilsson. Semantic topology. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1939–1942, 2014.
- [75] E. Kostenok, D. Cherniavskii, and A. Zaytsev. Uncertainty estimation of transformers’ predictions via topological analysis of the attention matrices. *arXiv preprint arXiv:2308.11295*, 2023.
- [76] T. Kovaliuk, I. Yurchuk, and O. Gurnik. Topological structure of ukrainian tongue twisters based on speech sound analysis. *6th International Workshop on Modern Data Science Technologies*, 2024.

- [77] S. Kudriashov, V. Zykova, A. Stepanova, J. Raskind, and E. Klyshinsky. The more polypersonal the better—a short look on space geometry of fine-tuned layers. In *International Conference on Neuroinformatics*, pages 13–22. Springer, 2024.
- [78] A. Kumar and A. Sarkar. Extractive text summarization using topological features. In *International Workshop on Combinatorial Image Analysis*, pages 105–121. Springer, 2022.
- [79] L. Kushnareva, D. Cherniavskii, V. Mikhailov, E. Artemova, S. Barannikov, A. Bernstein, I. Piontkovskaya, D. Piontkovski, and E. Burnaev. Artificial text detection via examining the topology of attention maps. In *Proceedings of the 2021 EMNLP*, pages 635–649, 2021.
- [80] L. Kushnareva, T. Gaintseva, G. Magai, S. Barannikov, D. Abulkhanov, K. Kuznetsov, E. Tulchinskii, I. Piontkovskaya, and S. Nikolenko. Ai-generated text boundary detection with roft. In *1st Conference on Language Modeling (COLM)*, volume 2024, 2024.
- [81] C. Laméris. *Topological Featurization of Speech Data for Speech Recognition*. PhD thesis, University of Groningen, 2024.
- [82] Z. Lamine, M. I. Mamouni, and M. W. Mansouri. A topological data analysis of the protein structure. *International Journal of Analysis and Applications*, 21:136–136, 2023.
- [83] R. Lavery, A. Jurek-Loughrey, and L. Bai. 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, 2024.
- [84] L. Lazo, H. Jelodar, and R. Razavi-Far. Llm security and safety: Insights from homotopy-inspired prompt obfuscation. *arXiv preprint arXiv:2601.14528*, 2026.
- [85] C. Li, C. Zhang, Y. Lu, S. Chen, X. Wang, J. Zhang, Z. Wang, Z. Jin, K. Liu, S.-H. Bae, et al. Understanding chain-of-thought in large language models via topological data analysis. *arXiv preprint arXiv:2512.19135*, 2025.
- [86] H. Li, H. Wan, Y. Huang, Y. Chen, Y. Gel, and H. Jiang. Large language models as topological thinkers: A benchmark on graph persistent homology, 2026.
- [87] J. Liu, L. Shen, and G.-W. Wei. Chatgpt for computational topology. *Foundations of data science (Springfield, Mo.)*, 6(2):221, 2024.
- [88] X. Liu, Z. Zhang, Y. Wang, H. Pu, Y. Lan, and C. Shen. 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, 2023.
- [89] B. Løvlie. Text classification via topological data analysis. Master’s thesis, Norwegian University of Science and Technology (NTNU), 2023.
- [90] X. Luong, M. Juillard, S. Mellet, and D. Longrée. Trees and after: The concept of text topology. *Literary and Linguistic Computing*, 22(2):167–186, 2007.
- [91] L. Maadarani and S. G. Hajra. The shape of poems. In *Fall Poster Forum*, 2020.
- [92] K. MacDonald, A. Ruparelia, B. Rogers, A. Bovell, and B. Stone. Binary malware attribution using llm embeddings and topological data analysis. *Conference on Applied Machine Learning for Information Security*, 2024.
- [93] A. Maćkiewicz and W. Ratajczak. Principal components analysis (pca). *Computers & Geosciences*, 19(3):303–342, 1993.
- [94] L. McInnes, J. Healy, N. Saul, and L. Großberger. Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29):861, 2018.
- [95] Y. Meng, R.-H. Li, H. Qin, X. Wu, H. Duan, Y. Lu, and G. Wang. Encoding group interests with persistent homology for personalized search. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54(9):5606–5616, 2024.
- [96] P. Michel, A. Ravichander, and S. Rijhwani. 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, 2017.
- [97] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26, 2013.
- [98] A. More, A. Zhang, N. Bonilla, A. Vivekan, K. Zhu, P. Sharafoleslami, and M. Chaudhary. Optimizing chain-of-thought confidence via topological and dirichlet risk analysis. *arXiv preprint arXiv:2511.06437*, 2025.
- [99] J. Morris, E. Lifland, J. Y. Yoo, J. Grigsby, D. Jin, and Y. Qi. 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, 2020.
- [100] E. Munch. A user’s guide to topological data analysis. *Journal of Learning Analytics*, 4(2):47–61, 2017.
- [101] J. Murugan and D. Robertson. An introduction to topological data analysis for physicists: From lgm to frbs. *arXiv preprint arXiv:1904.11044*, 2019.
- [102] J. L. Nielson, J. Paquette, A. W. Liu, C. F. Guandique, C. A. Tovar, T. Inoue, K.-A. Irvine, J. C. Gensel, J. Kloke, T. C. Petrossian, et al. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. *Nature communications*, 6(1):8581, 2015.
- [103] L. Novak. Network and topological analysis of scholarly metadata: A platform to model and predict collaboration. Master’s thesis, Purdue University, 2019.
- [104] E. Paluzo-Hidalgo, R. Gonzalez-Diaz, and G. Aguirre-Carrazana. Emotion recognition in talking-face videos using persistent entropy and neural networks. *Electronic Research Archive*, 30(2):644–660, 2022.
- [105] E. Paluzo Hidalgo, R. González Díaz, and M. Á. Gutiérrez Naranjo. Summary and distance between sets of texts based on topological data analysis. *arXiv preprint arXiv:1912.09253*, 2019.

- [106] E. Paluzo-Hidalgo, R. Gonzalez-Diaz, and M. A. Gutierrez-Naranjo. Trainable and explainable simplicial map neural networks. *Information Sciences*, 667:120474, 2024.
- [107] T. Papamarkou, T. Birdal, M. M. Bronstein, G. E. Carlsson, J. Curry, Y. Gao, M. Hajij, R. Kwitt, P. Lio, P. Di Lorenzo, et al. Position: Topological deep learning is the new frontier for relational learning. In *Forty-first International Conference on Machine Learning*, 2024.
- [108] I. Perez and R. Reinauer. The topological bert: Transforming attention into topology for natural language processing. *arXiv preprint arXiv:2206.15195*, 2022.
- [109] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. 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, 2018.
- [110] G. Petri and A. Leita. Prediction of disease type from topological features of time series. In *gtda-challenge-2020*, 2020.
- [111] H. T. Phan, N. T. Nguyen, and D. Hwang. Fake news detection: A survey of graph neural network methods. *Applied Soft Computing*, 139:110235, 2023.
- [112] A. Pollano, A. Chaudhuri, and A. Simmons. Detecting out-of-distribution text using topological features of transformer-based language models. *The IJCAI-2024 AISafety Workshop*, 2024.
- [113] A. Port, I. Gheorghita, D. Guth, J. M. Clark, C. Liang, S. Dasu, and M. Marcolli. Persistent topology of syntax. *Mathematics in Computer Science*, 12(1):33–50, 2018.
- [114] A. Port, T. Karidi, and M. Marcolli. Topological analysis of syntactic structures. *Mathematics in Computer Science*, 16(1):2, 2022.
- [115] P. Proskura and A. Zaytsev. 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, 2024.
- [116] I. Proskurina, E. Artemova, and I. Piontkovskaya. 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, 2023.
- [117] A. A. Rahim et al. Topological perspectives on optimal multimodal embedding spaces. *arXiv preprint arXiv:2405.18867*, 2024.
- [118] N. Rair, A. Goupil, V. Vrabie, and E. Chochoy. When annotators disagree, topology explains: Mapper, a topological tool for exploring text embedding geometry and ambiguity. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 8468–8491, 2025.
- [119] M. B. Raskin. The shape of emotion in speech: A topological data analysis of speech emotion recognition. *IEEE Access*, 14:52964–52978, 2026.
- [120] A. Rathore, Y. Zhou, V. Srikumar, and B. Wang. Topobert: Exploring the topology of fine-tuned word representations. *Information Visualization*, 22(3):186–208, 2023.
- [121] M. Rawson, S. Dooley, M. Bharadwaj, and R. Choudhary. Topological data analysis for word sense disambiguation. *arXiv preprint arXiv:2203.00565*, 2022.
- [122] R. Rejimoan. Detection of machine-generated text by integrating roberta embeddings with topological features. *International Journal of Applied Mathematics*, 38(6s):17–30, 2025.
- [123] B. Rieck. Topological data analysis for machine learning. *ECML-PKDD 2020 (Tutorial)*, 2020.
- [124] B. Roziere, J. Gehring, F. Gloeckle, S. Sootla, I. Gat, X. E. Tan, Y. Adi, J. Liu, R. Sauvestre, T. Remez, et al. Code llama: Open foundation models for code. *arXiv preprint arXiv:2308.12950*, 2023.
- [125] B. M. Ruppik, M. Heck, C. van Niekerk, R. Vukovic, H.-c. Lin, S. Feng, M. Zibrowius, and M. Gašić. 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*, 2024.
- [126] S. M. N. Sakib. S M Nazmuz Sakib Affix Isometry Index, 2025. researchgate.net.
- [127] W. Sakurai, M. Asano, D. Imoto, M. Honma, and K. Kurosawa. 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, 2025.
- [128] S. N. Samaga, G. G. Arroyo, and T. K. Dey. Halluzig: Hallucination detection using zigzag persistence. *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics*, 2026.
- [129] I. R. Sami and K. Farrahi. A simplified topological representation of text for local and global context. In *Proceedings of the 25th ACM International Conference on Multimedia*, pages 1451–1456, 2017.
- [130] A. B. Santacana. A topological semantics of dialogue: Nerve structures and logical extraction. *arXiv preprint arXiv:2506.00615*, 2025.
- [131] L. Sassone, M. Manetti, M. G. Bergomi, and M. Ferri. *Bridging Topological Persistence and Machine Learning for Music Information Retrieval*. PhD thesis, Sapienza – University of Rome, 2022.
- [132] N. Saul and D. L. Arendt. Machine learning explanations with topological data analysis. In *Demo at the Workshop on Visualization for AI Explainability (VISxAI)*, 2018.
- [133] N. Saul and C. Tralie. Scikit-tda: Topological data analysis for python. URL <https://doi.org/10.5281/zenodo.2533369>, 2019.
- [134] K. Savle, W. Zadrozny, and M. Lee. Topological data analysis for discourse semantics? In *Proceedings of the 13th International Conference on Computational Semantics-Student Papers*, pages 34–43, 2019.

- [135] B. Scalvini and A. Mashaghi. Semantic topology: a new perspective for communication style characterization. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 9223–9233, 2025.
- [136] D. Shehu. An analysis of the effect of polysemy on the topology of the latent manifold. Master’s thesis, Eindhoven University of Technology, 2024.
- [137] K. Shin. Genre classification: A topological data analysis approach. *kevin-shin.com*, 2019.
- [138] G. Singh, F. Mémoli, G. E. Carlsson, et al. Topological methods for the analysis of high dimensional data sets and 3d object recognition. *PBG@ Eurographics*, 2:091–100, 2007.
- [139] Y. Singh, C. M. Farrelly, Q. A. Hathaway, T. Leiner, J. Jagtap, G. E. Carlsson, and B. J. Erickson. Topological data analysis in medical imaging: current state of the art. *Insights into Imaging*, 14(1):58, 2023.
- [140] Y. Skaf and R. Laubenbacher. Topological data analysis in biomedicine: A review. *Journal of Biomedical Informatics*, 130:104082, 2022.
- [141] P. Snopov and A. N. Golubinskiy. Vulnerability detection via topological analysis of attention maps. *arXiv preprint arXiv:2410.03470*, 2024.
- [142] P. Solunke, V. Guardieiro, J. Rulff, P. Xenopoulos, G. Y.-Y. Chan, B. Barr, L. G. Nonato, and C. Silva. Mountaineer: Topology-driven visual analytics for comparing local explanations. *IEEE Transactions on Visualization and Computer Graphics*, 2024.
- [143] B. Sovdat. Text mining via homology. Master’s thesis, UNIVERSITY OF LJUBLJANA, 2016.
- [144] A. Spannaus, H. A. Hanson, G. Tourassi, and L. Penberthy. Topological interpretability for deep learning. In *Proceedings of the Platform for Advanced Scientific Computing Conference*, pages 1–11, 2024.
- [145] T. Sun and B. Nelson. Topological interpretations of gpt-3. *arXiv preprint arXiv:2308.03565*, 2023.
- [146] X. W. Tan, N. Tan, G. Lee, and S. Kok. The shape of reasoning: Topological analysis of reasoning traces in large language models. *arXiv preprint arXiv:2510.20665*, 2025.
- [147] T. Temčinas. Local homology of word embeddings. *arXiv preprint arXiv:1810.10136*, 2018.
- [148] M. Tlachac, A. Sargent, E. Toto, R. Paffenroth, and E. Rundensteiner. 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, 2020.
- [149] P. Torres-Tramón, H. Hromic, and B. R. Heravi. Topic detection in twitter using topology data analysis. *Current Trends in Web Engineering*, pages 186–197, 2015.
- [150] J. F. Triki. Analysis of word embeddings: A clustering and topological approach. Master’s thesis, The University of Bergen, 2021.
- [151] I. Trofimov, D. Cherniavskii, E. Tulchinskii, N. Balabin, E. Burnaev, and S. Barannikov. Learning topology-preserving data representations. In *ICLR 2023 International Conference on Learning Representations*, 2023.
- [152] A. Tsai, S. Subramanian, C. Liu, K. Lopez, L. Zinn-Brooks, A. Shulz, and A. Uchendu. The shape of vulnerability: How adversarial perturbations reshape the topology of language model latent spaces. In *ACL 2026 Student Research Workshop*, 2026.
- [153] E. Tulchinskii, K. Kuznetsov, D. Cherniavskii, S. Barannikov, S. Nikolenko, and E. Burnaev. Topological data analysis for speech processing. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pages 311–315, 2023.
- [154] E. Tulchinskii, K. Kuznetsov, L. Kushnareva, D. Cherniavskii, S. Nikolenko, E. Burnaev, S. Barannikov, and I. Piontkovskaya. Intrinsic dimension estimation for robust detection of ai-generated texts. *Advances in Neural Information Processing Systems*, 36, 2024.
- [155] S. Tymochko, J. Chaput, T. Doster, E. Purvine, J. Warley, and T. Emerson. 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, 2021.
- [156] S. Tymochko, Z. New, L. Bynum, E. Purvine, T. Doster, J. Chaput, and T. Emerson. Argumentative topology: Finding loop (holes) in logic. *arXiv preprint arXiv:2011.08952*, 2020.
- [157] A. Uchendu, T. Le, and D. Lee. Topformer: Topology-aware authorship attribution of deepfake texts with diverse writing styles. *ECAI 2024*, 2024.
- [158] L. Valeriani, D. Doimo, F. Cuturello, A. Laio, A. Ansuini, and A. Cazzaniga. The geometry of hidden representations of large transformer models. *Advances in Neural Information Processing Systems*, 36, 2024.
- [159] H. J. van Veen. Novel topological shapes of model interpretability. In *TDA and Beyond at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, 2020.
- [160] K. Varadarajan and T. Songdechakraiwut. Augmenting bias detection in llms using topological data analysis, 2025.
- [161] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [162] D. Voronkova, I. Trofimov, A. Dmitriev, E. Tulchinskii, E. Burnaev, and S. Barannikov. Topology of attention detects hallucinations in code LLMs, 2026.
- [163] M. Vu, G. Zollicoffer, H. Mai, B. Nebgen, B. Alexandrov, and M. Bhattarai. Topological signatures of adversaries in multimodal alignments. In *Forty-second International Conference on Machine Learning*, 2025.

- [164] R. Vukovic, M. Heck, B. Ruppik, C. van Niekerk, M. Zibrowius, and M. Gasic. 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, 2022.
- [165] H. Wagner, P. Dłotko, and M. Mrozek. 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, 2012.
- [166] M. Wamil, A. Hassaine, S. Rao, Y. Li, M. Mamouei, D. Canoy, M. Nazarzadeh, Z. Bidel, E. Copland, K. Rahimi, et al. Stratification of diabetes in the context of comorbidities, using representation learning and topological data analysis. *Scientific Reports*, 13(1):11478, 2023.
- [167] Z. Wang, D. Anshumaan, A. Hooda, Y. Chen, and S. Jha. Functional homotopy: Smoothing discrete optimization via continuous parameters for llm jailbreak attacks. In *The Thirteenth International Conference on Learning Representations*, 2025.
- [168] D. Wei, M. Mao, X. Fang, and M. Chau. Short-phd: Detecting short llm-generated text with topological data analysis after off-topic content insertion. In *2nd Conference on Language Modeling (COLM)*, volume 2025, 2025.
- [169] M. Wright and X. Zheng. Topological data analysis on simple english wikipedia articles. *The PUMP Journal of Undergraduate Research*, 3:308–328, 2020.
- [170] X. Wu, X. Niu, and R. Rahman. Topological analysis of contradictions in text. In *Proceedings of the 45th International ACM SIGIR*, pages 2478–2483, 2022.
- [171] P. Xenopoulos, G. Chan, H. Doraiswamy, L. G. Nonato, B. Barr, and C. Silva. Gale: Globally assessing local explanations. In *Topological, Algebraic and Geometric Learning Workshops 2022*, pages 322–331. PMLR, 2022.
- [172] M. Xu, Q. Hu, X. Hu, S. Abousamra, X. Yu, W. Lyu, K. Qi, D. Samaras, and C. Chen. Topo-r1: Detecting topological anomalies via vision-language models. *arXiv preprint arXiv:2603.13054*, 2026.
- [173] H. Yadav, T. B. Smith, P. Bubenik, and C. McCarty. What is missing from this picture? persistent homology and mixup barcodes as a means of investigating negative embedding space. *arXiv preprint arXiv:2510.14327*, 2025.
- [174] X. Yan, R. Sevastjanova, S. van der Ben, M. El-Assady, and B. Wang. Explainable mapper: Charting llm embedding spaces using perturbation-based explanation and verification agents, 2025.
- [175] Z. Yessenbayev and Z. Kozhirkbayev. Comparison of word embeddings of unaligned audio and text data using persistent homology. In *International Conference on Speech and Computer*, pages 700–711. Springer, 2022.
- [176] Z. Yessenbayev and Z. Kozhirkbayev. 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, 2024.
- [177] J. You, K. Dasol, and J.-H. Jung. Topological alignment of shared vision-language embedding space. In *UniReps: 3rd Edition of the Workshop on Unifying Representations in Neural Models*, 2025.
- [178] Z. Yu, P. Feng, Q. Qu, H. Zhang, and Y. Zhu. Topological deep learning for speech data, 2025.
- [179] I. Yurchuk and O. Gurnik. Tongue twisters detection in ukrainian by using tda. In *CEUR Workshop Proceedings*, pages 163–172, 2023.
- [180] W. W. Zadrozny. Abstraction, reasoning and deep learning: A study of the "look and say" sequence. *arXiv preprint arXiv:2109.12755*, 2021.
- [181] W. W. Zadrozny. A note on argumentative topology: Circularity and syllogisms as unsolved problems. *arXiv preprint arXiv:2102.03874*, 2021.
- [182] H. Zhang, L. Zhang, Y. Zhang, and Z. Mao. Homology consistency constrained efficient tuning for vision-language models. *Advances in Neural Information Processing Systems*, 37:93011–93032, 2024.
- [183] J. Zhang, Q. Sun, C. Zhang, X. Wang, Z. Huang, Y. Zhou, P. Zheng, C.-I. A. Tai, S.-H. Bae, Z. Ma, et al. Tda-rc: Task-driven alignment for knowledge-based reasoning chains in large language models. *arXiv preprint arXiv:2604.04942*, 2026.
- [184] J. Zhang, C. Zhang, S. Chen, Y. Liu, C. Li, Q. Sun, S. Yuan, F. D. Puspitasari, D. Han, G. Wang, et al. Text summarization via global structure awareness. *arXiv preprint arXiv:2602.09821*, 2026.
- [185] J. Zhang, C. Zhang, S. Chen, X. Wang, Z. Huang, P. Zheng, S. Yuan, S. Zheng, Q. Sun, J. Zou, L.-H. LEE, and Y. Yang. Learning global hypothesis space for enhancing synergistic reasoning chain. In *The Fourteenth International Conference on Learning Representations*, 2026.
- [186] P. Zheng, C. Zhang, Y. Wen, W. Liu, Q. Sun, J. Mo, J. Zhang, J. Lee, T.-H. Kim, K. Liu, et al. Topology-aware layer pruning for large vision-language models. *arXiv preprint arXiv:2604.16502*, 2026.
- [187] X. Zhu. Persistent homology: An introduction and a new text representation for natural language processing. In *IJ-CAI*, pages 1953–1959, 2013.
- [188] Y. Zhu, P. Feng, S. Yi, Q. Qu, and Z. Yu. Topology-enhanced machine learning for consonant recognition. *Research Square*, 2024.

APPENDIX

Table 3: Non-theoretical applications of TDA in NLP. For the TDA techniques - **PH**: Persistent Homology and **M**: Mapper. Task categories - **cl**: classification, **C & TM**: clustering & topic modeling, **S & SA**: sentiment & semantic analysis, **S & V**: structure & visualization, **H,S,& SA**: health, social, & scholarly analysis, **S**: speech processing, **MI & A**: model interpretation & analysis (**PART I**)

Name	Task	Problem	Numerical Representation	Learning Type	Modality	TDA
SIFT [187]	cl	child vs. adolescent writing detection	TF-IDF	Supervised	Text	PH
[89]	cl	deepfake text detection	TF-IDF, GloVe	Supervised	Text	PH
[68]	cl	election speech feature extraction	TF-IDF	Supervised	Text	PH
[35, 137]	cl	movie genre	TF-IDF	Supervised	Text	PH
[143]	cl	distinguishing between languages	TF-IDF	Supervised	Text	PH
[134]	S & SA	legal document entailment	TF-IDF	Supervised	Text	PH
[96]	S & SA	clustering and sentiment analysis	TF-IDF, GloVe	Supervised	Text	PH
DoCollapse [59]	C & TM	keyphrase extraction	TF-IDF	Unsupervised	Text	PH
TOPOL [149]	C & TM	Twitter topic detection	TF-IDF	Supervised	Text	PH
[78]	C & TM	text summarization	TF-IDF	Unsupervised	Text	PH
[159]	cl	Propaganda tweet	TF-IDF	Unsupervised	Text	M
[40]	cl	classify Persian poems	TF-IDF	Supervised	Text	PH & M
[39]	cl	age group categorization of lonely people	TF-IDF	Supervised	Text	PH & M
[91]	cl	nursery rhyme classification from different continents - Australia, Asia, Africa, Europe, and North America	TF-IDF	Supervised	Text	PH
[61]	S & V	building document structure	Pre-trained (Word2Vec)	Unsupervised	Text	PH
BERT+TDA [170]	S & SA	contradiction detection	Pre-trained (Word2Vec)	Supervised	Text	PH
[175]	MI & A	speaker recognition & text processing	Pre-trained (Word2Vec)	Unsupervised	Speech	PH
[176]	MI & A	speaker recognition & text processing	Pre-trained (Word2Vec)	Unsupervised	Speech	PH
[30]	S & SA	building a topological search engine	Pre-trained (Word2Vec)	Unsupervised	Text	M
[64]	C & TM	document clustering and topic modeling tasks	Pre-trained (Word2Vec)	Supervised	Text	M
[121]	S & SA	word sense induction and disambiguation	Pre-trained (Word2Vec)	Unsupervised	Text	PH
[147]	S & SA	word sense induction and disambiguation	Pre-trained (Word2Vec, GloVe)	Unsupervised	Text	PH
[44]	MI & A	Geometry of textual data augmentation	Pre-trained (Word2Vec)	Supervised	Text	PH
[155]	cl	fraudulent paper detection	Pre-trained (Word2Vec, GloVe, EIMo)	Supervised	Text	PH
[110]	H,S,& SA	disease epidemic prediction	Pre-trained (Word2Vec)	Supervised	Text	PH
[173]	H,S,& SA	publication analysis	Pre-trained (Word2Vec)	Unsupervised	Text	PH
[156]	cl	finding topological loops in logical statements	Pre-trained (Word2Vec, GloVe)	Unsupervised	Text	PH
[169]	C & TM	distinguish subsets in data	Pre-trained (Word2Vec)	Unsupervised	Text	PH
[105]	S & SA	measuring the distance between the literary style of Spanish poets	Pre-trained (Word2Vec)	Supervised	Text	PH
[10]	S & SA	detecting narrative shifts	Pre-trained (Word2Vec)	Unsupervised	Text	PH
[5]	S & SA	distinguishing news articles & poems	Pre-trained (Word2Vec, GloVe)	Unsupervised	Text	PH
[51]	cl	extract the topological signatures of novelists	Pre-trained (GloVe)	Supervised	Text	PH
TIES [52]	S & SA	document categorization & sentiment analysis	Pre-trained (GloVe)	Unsupervised	Text	PH
[144]	MI & A	phenotype prediction and news group categorization	Pre-trained (GloVe)	Unsupervised	Text	M
[181]	S & SA	finding topological loops in logical statements	Pre-trained (GloVe)	Unsupervised	Text	PH
[21]	H,S,& SA	social anxiety detection	Pre-trained (GloVe)	Supervised	Text	PH
[62]	MI & A	compare cross-lingual sentence representations	Pre-trained (GloVe)	Unsupervised	Text	PH
[3]	cl	deepfake text detection	Pre-trained (GloVe)	Supervised	Text	PH
[180]	MI & A	investigates the manifestations of intelligence and understanding in neural networks	Pre-trained (GloVe)	Supervised	Text	PH
[32]	cl	fake news detection	Pre-trained (GloVe, BERT)	Supervised	Text	PH
[105]	H,S,& SA	analyzing scholarly network	Pre-trained (GloVe, BERT)	Unsupervised	Text	PH
[73]	S & SA	(1) polysemy word classification, and (2) word sense induction & disambiguation	Pre-trained (FastText)	Unsupervised	Text	PH
[150, 136]	S & SA	polysemy word	Pre-trained (FastText)	Supervised	Text	PH
PHD [154]	cl	deepfake text detection	Pre-trained (Transformers - CLS)	Supervised	Text	PH
Short-PHD [168]	cl	deepfake text detection (for short-text)	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[60]	cl	deepfake text detection (for academic abstracts)	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[80]	cl	deepfake text detection	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[57]	MI & A	class separability estimation	Pre-trained (Transformers - CLS)	Unsupervised	Text	PH
TopoBERT [120]	S & V and MI & A	visually analyzing the fine-tuning process of a Transformer-based model	Pre-trained (Transformers - CLS)	Unsupervised	Text	M
[83]	cl	fake news detection	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[31]	cl	classification of public speaking ratings from TED talks	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[25]	H,S,& SA and S & V	tracking individual mental health journeys	Pre-trained (Transformers - cls)	Supervised	Text	M
[22, 65]	C & TM	topic modeling	Pre-trained (Transformers - CLS)	Unsupervised	Text	M
TOPFORMER [157]	cl	deepfake text detection	Pre-trained (Transformers - Hidden)	Supervised	Text	PH
TDA-BERTa [122]	cl	deepfake text detection	Pre-trained (Transformers - Hidden)	Supervised	Text	PH
[7]	cl	Portuguese-English language translation	Pre-trained (Transformers - Hidden)	Supervised	Text	PH

Table 4: Non-theoretical applications of TDA in NLP. For the TDA techniques - **PH**: Persistent Homology and **M**: Mapper. Task categories - **cl**: classification, **C & TM**: clustering & topic modeling, **S & SA**: sentiment & semantic analysis, **S & V**: structure & visualization, **H,S,& SA**: health, social, & scholarly analysis, **S**: speech processing, **MI & A**: model interpretation & analysis (**PART II**)

Name	Task	Problem	Numerical Representation	Learning Type	Modality	TDA
[48]	S & SA	polysemy word	Pre-trained (Transformers - Hidden)	Supervised	Text	M
[15]	S & SA	polysemy word	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[56]	S & SA	semantic analysis of embeddings	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[184]	S & SA	text summarization	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[4]	H,S,& SA and S & V	Hate speech, Misinformation & Psychiatric disorder	Pre-trained (Transformers - Hidden)	Supervised	Text	M
PBCE [11]	MI & A	Model compression	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
Persistent Similarity [50]	MI & A	Probing layers in LLMs	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[145]	MI & A	Correlation between sentence vectors	Pre-trained (Word2Vec, Transformers - Hidden)	Unsupervised	Text	PH
Persistence Scoring Function [26]	MI & A	captures the homology of the high-dimensional hidden representations	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
Topological Denisification [49]	MI & A	zero-shot model stitching	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
TSLoss [69]	MI & A	topology loss function for prompt tuning	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
HOLE [8]	MI & A	Homological Observation of Latent Embeddings for Neural Network Interpretability	Pre-trained (Transformers - Hidden)	Supervised	Text	PH
[46]	MI & A	topology of fairness	Pre-trained (Transformers - Hidden)	Unsupervised	Text	M
[43]	MI & A	adversarial vs. non-adversarial text representation in LLMs	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[146]	MI & A	shape of reasoning process of LLMs	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[85, 183]	MI & A	reasoning process of LLMs using chain-of-thought prompts	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
EDTR [98]	MI & A	measures LLM confidence	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
GHS-TDA [185]	MI & A	improving reasoning process of chain-of-thought	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[71]	MI & A	tracing reasoning process of chain-of-thought	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[125]	C & TM	dialogue term extraction	Pre-trained (Transformers - Hidden)	Supervised	Text	PH
[77]	MI & A	polypersonality	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[174]	MI & A	explainability of latent space	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[47]	MI & A	topological complexity of LLM hidden space	Pre-trained (Transformers - Hidden)	Unsupervised	Text	PH
[14, 162]	cl	LLM Hallucination detection	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[92]	cl	Code attribution	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[79]	cl	deepfake text detection	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[27, 116, 72]	S & SA	grammatical acceptability judgment	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[75]	MI & A	Uncertainty estimation of model predictions	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[128]	MI & A	hallucination detection in LLMs	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[152]	MI & A	investigation of adversarial influence in the latent space	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[160]	MI & A	bias detection	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[108]	cl	spam detection, grammatical acceptability judgment, and movie sentiment analysis	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[127]	cl	Authorship attribution of Japanese texts	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[112]	cl	out-of-distribution detection	Pre-trained (Transformers - Attention, CLS)	Supervised	Text	PH
[115]	MI & A	estimation of weights for ensembles of classification models	Pre-trained (Transformers - Attention, CLS)	Supervised	Text	PH
[118]	MI & A	visualizes fine-tuning process to see when annotators disagree	Pre-trained (Transformers - Attention, CLS)	Supervised	Text	PH
[6]	S & V and S & SA	controversial vs. non-controversial political discourse detection	Pre-trained (Transformers - CLS)	Supervised	Text	PH
PHPS [95]	S & V and S & SA	personalized web search	Pre-trained (Transformers - CLS)	Supervised	Text	PH
[141]	cl	vulnerability detection in code	Pre-trained (Transformers - Attention)	Supervised	Text	PH
TopoHuBERT [153]	S	speaker recognition	Pre-trained (Transformers - Attention)	Supervised	Speech	PH
[164]	S & SA	dialogue term extraction	Pre-trained (Transformers - Attention)	Supervised	Text	PH
[179]	S & SA	Ukrainian tongue twisters	Symbolic Representations	Supervised	Text	PH
[76]	S	Ukrainian tongue twisters	Symbolic Representations	Supervised	Speech	PH
[41]	H,S,& SA	recipe discovery	Symbolic Representations	Unsupervised	Text	PH
[19]	S	human vowel	Multi-Modal	Supervised	Speech	PH
[18]	S	infant vocalization	Multi-Modal	Unsupervised	Speech	PH
[178, 81]	S	speech recognition	Multi-Modal	Unsupervised	Speech	PH
[55, 104, 119]	S	emotion recognition	Multi-Modal	Supervised	Speech	PH
[148]	S	depression detection	Multi-Modal	Supervised	Speech	PH
[188]	S	consonants recognition	Multi-Modal	Supervised	Speech	PH
[163]	MI & A	multi-modal adversarial robustness assessment	Multi-Modal	Supervised	Text	PH
[67, 182, 117, 172]	MI & A	aligning the topological structures of image and text representations during tuning	Multi-Modal	Supervised	Text	PH
[9]	MI & A	aligning the topological structures of image and text representations during tuning	Multi-Modal	Supervised	Text	PH
[177]	MI & A	multi-modal multi-lingual topological alignment	Multi-Modal	Supervised	Text	PH