DS Journal of Digital Science and Technology
Volume 2 Issue 3, 1-8, Jul - Sep 2023
ISSN: 2583-5416 / https://doi.org/10.59232/DST-V2I3P101

Original Article

Persistent Homology and Artificial Intelligence Analysis of COVID-19 in Topological Spaces

Benard Okelo[1], Allan Onyango[2]

[1,2]*Department of Pure and Applied Mathematics, JOOUST, Bondo, Kenya.*

[1]bnyaare@yahoo.com

Received: 30 June 2023; Revised: 15 July 2023; Accepted: 02 August 2023; Published: 19 August 2023

Abstract - In this study, we describe aspects of Topological Data Analysis using techniques like Persistent Homology and Simplicial Complex. We then explore Big Data Sets of COVID-19 with extremely large and highly complex properties. We, after that, discuss the characterization and significance of the Hausdorff Spaces an AI in this work. Next, we describe simulations within the Python programming language with applications in TDA.

Keywords - Artificial Intelligence, Machine Learning, Topological Data Analysis, COVID-19, Python.

1. Introduction

The notion of TDA is founded under the ubiquitous theory of persistent homology (see [6, 9, 13, 22-37, 40] and the references therein). Some of the pioneer contributors to TDA include [1, 3, 5], who founded the notion of how features persist as the data is modified. Nevertheless, the genesis of the term TDA expression appears not to have surfaced till contributions by [2, 4]. Thereafter, the authors in [7] became instrumental in the popularization of TDA, establishing the ways topological techniques willremedy challenges encountered while implementing topology to analyze Big Data. In [14], the work put up other developments by observing that Persistent homology is currently one of the more widely known tools from computational topology and Topological Data Analysis. Topology and geometry are tools used to investigate highly composite data [19] by creating a compendium of the data features to uncoverhidden attributes within the dataset. Normally, the dataset of interest is often centered around structures that appear challenging to reveal with traditional methods [8]. The major TDA approach for removing "topological noise" maps the original data to a lower dimensional approximation acquired through a multidimensional assortment. Therefore, Open sets provide an essential approach to understanding the nearness of points without a distance element defined in a topological space [10]. Other inherent mathematical concepts to understand besides topology include continuity, connectedness, and closeness, which embrace nearness. The problem is that no single story is happening in this data. We can therefore say this data has much "noise"! The explosive growth in data [11], voice and video traffic, and ubiquity of social-media content, health records, and many more data sources have been contributing factors to Big Data. It was anticipated that the generated data volume could be 44 zettabytes in 2020, as found in [12]. The vast data volume, with its complexity, has propelled technological advancements [14] realized as well as accelerated increase in bandwidth capacity, processing power, storage capability and transfer velocity [16].

This is partially due to the technological advancement in high-power computing. There is, therefore, an urgent need to establish robust and resilient techniques to process Big Data. BD consists of 5 Vs: Value, Variety, Volume, Veracity and Velocity [15]. The data size to be processed and analyzed constitutes the volume. The speed of growth and usage of this data is the velocity. The varied data formats cum types are the variety.

Veracity involves accuracy plus analysis of the results of the datasets. The richness obtained after processing the dataset is the value. The growing volumes of Voice over IP (VoIP) social-media content [18] underscores the requirement for ways of countering the ambiguity innate in the finite datasets. Presently, roughly eighty percent of datasets remain indeterminate. TDA has lately recorded advances in innumerable directions and application disciplines. The fundamental aim of TDA is to extract multidimensional rich datafeatures based on geometry and topology pre-existing in distributed data-points, as shown in [17, 21, 28]. Connections within the data and topological methods have a close affiliation to neural networks between data pointswhich reveal insight into this united structure.

According to [20], mostcommonly, every other form of TDA [23-24] revolves around the steps given by: Firstly, the data sample is presumed as datapoints which are finite quantified as a metric space $\mathbf{R}^d$ [39]. Worth to mention is that the metric choice may be vital to guarantee remarkable topological and geometric featuresof the data set [31]. Secondly, a mathematical structure is computed on top of the dataset to tap more from the fundamental concepts of geometry and topology. This is mostly cases in SC or a convention of SCs that depicts the high dimensional data structures at varied degrees [36]. Thirdly, from these high dimensional data structures built on atop the data set, topological or geometric information is derived [29]. The shape of the data from which we extract the topological/geometrical high-dimensional features can either be crude structural summaries or relevant approximations that need further approaches like persistent homology and visualization [26]. The extracted topological and geometric information gives rise to insightful features and descriptors in the data, which, when injected into further analysis and machine learning procedures, reveal very rich results and significant meaning that can be used in other disciplineslike medicine [23], biology [27] and astrophysics [29], just to mention a few [25].

2. Materials and Methods

In this section, we present the techniques, materials, and tools that we used to achieve the objectives of this work. These include; Big Data Sets, Separation criterion of Hausdorff Spaces, AI and ML techniques, and the development of algorithms and simulations using Python programming language [38]. Big Data is described as hugely large, highly composite datasets to be analyzed by traditional techniques but might reveal structure, patterns, relationships, and shapes when computational analysis methods are involved.

3. Results and Discussion

3.1. TDA and Big Data Sets

The principle idea behind TDA involves applying techniques to recognise shapes and patterns within data. The finite TDPs within a Euclidean space become the driving force for Big Data to be considered in TDA. TDA perceives the point cloud data as a discrete cluster of points within a compact topological space infinitely composed of many points. When neighboring data points are “connected” to reveal geometry atop the dataset, this reveals rich topological features. The idea of distance closeness among TDPs is the backbone that TDA exploits to qualify data sets as metric spaces. The driving force behind Topological Data Analysis has to construct higher-dimensional structures by linking pairs of TDPs by edges and by $(k + 1) - tuple$ of surrounding TDPs. This drives us to a concept called simplicial complexes, which makes it easy to recognize emerging topological characteristics, including points, lines, holes, cycles, and voids.

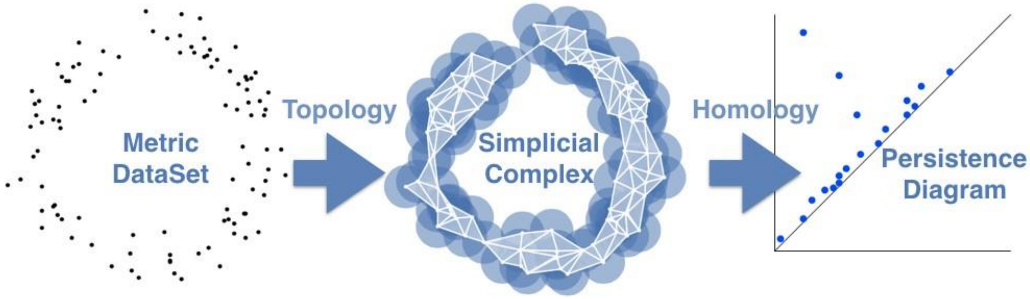

Fig. 1 Encoding point cloud persistent topological features (left) by approximation of the space by simplicial complexes (middle) into a persistence diagram (right) [20]

Figure 1 shows point cloud persistent topological features (left) by approximation of the space by simplicial complexes (middle) into a persistence diagram which shows clustering when the distance between the points is reduced.

3.2. Persistent Homology

Given that $u \geq v$, then non-empty sets $A_u \subseteq A_v$ such that u and v are distinct points in the set A. Moving from u to v, the components of the non-empty set A_u may merge as new ones are born, which have higher chances of merging with each other or with the existent components of A_u. Consequently, these components may change their topology with holes and other structures forming and disappearing. This is a perfect demonstration of persistent homology. The 'persistence' ideology is a result of the changing of the level u with no change in the homology untila critical point f of level u is reached, which means that the topology of the excursion sets 'persists' (remains static), between the varying heightsof critical points. This concept of Persistent Homology (PH) persists further; every time two components merge, we treat the firstof these to have appeared as though its existence continues beyond the point of merging. Moving forward, a more promising illustration of persistent homology is through barcodes. Suppose $dim(M) = N$, and given a smooth of f, if A_u is nonempty, then $dim(A_u)$ seamlessly becomes N. The barcode for the excursion set f becomes the collection of $N + 1$ graphs, with each homology group having one. A bar in the $k - th$ graph, beginning at u_1 and terminating at $u_2 (u_1 \geq u_2)$, reveals $H_k(A_u)$ that emerged at u_1 and vanished at u_2.

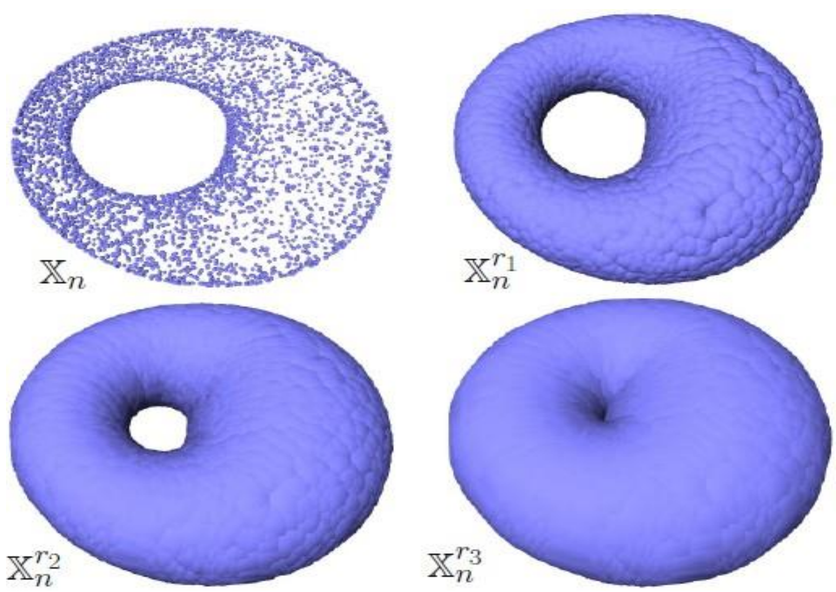

Fig. 2 The point cloud instance X_n derived from the torus surface (R^3 in the top left), which leads to varying radii $r_1 < r_2 < r_3$ [15]

Figure 2 displays a more impressive illustration with a three-dimensional view. It is imperative to indicate that, unlike a Two- Dimensional (2D) space, our comprehension is served with immense visual information from the barcodes since the increase in the N-dimension parameter space projects more clarity. As a result, therefore, it becomes uncomplicated to observe barcodes with six sets of bars for the six persistent homologies [28].

Big Data sets can be described as hugely large, highly composite datasets to be analyzed by traditional techniques, but might reveal structure, patterns, relationships, and shapes when computational analysis methods are involved [22]. BD is found almost everywhere, with many connections and meaning hidden within this complex data. Precisely, five "Vs" (Value, Veracity, Volume, Velocity and Variety) are used to sum up the description of BD. Given the traditional software tools, every attempt to store, extract, search, share, analyze or process this massive and complex data set has achieved little success. Despite this failure, different levels of society, including businesses, education, health, transport, research and many more, continue to persistently and desperately sort insights from Big Data analysis to enhance their efficiency and performance. The human ability to visualize Big Data sets is not keeping pace with its exploding natureof production. Therefore, it is urgent to develop advanced and proficient analysis methods to manage this kind of data. One such method is the application of geometrical and topological tools. Geometry involves the study of distance functions, which works very well with large finite data sets. The mathematical concepts formulated by mathematicians that unite topological and geometric techniques often revolve around point clouds, basically finite point sets equipped with a distance function. An important objective inBD analysis is to understand how the data is organized on a large scale hence retrieving qualitative information about the data. For instance, given a data set of diabetic patients, insightfully distinguishing the two distinct forms of the disease is firstly important. Given a very large data set X, it may prove challenging to apply a clustering algorithm on top of the data set.

One may otherwise opt to cluster subsamples from X. The confronting question is always whether sampling a cluster is enough proof of representation of the whole data set. An alternate approach is to construct two clusters from the whole dataset, assuming confidence in their consistency. Hence, taking the subsamples X_1 and X_2, including their union $X_1 \cup X_2$. A clustering scheme is then applied to each of the sets separately by denoting the three sets X_1, X_2, $X_1 \cup X_2$ by sets of clusters $C(X_1)$, $C(X_2)$, and $C(X_1 \cup X_2)$, respectively. Suppose the clustering scheme on the data sets induced maps of the collection of clusters. Suppose these clusters in C(X1) and C(X2) in C(X1 ∪ X2) consistently correspond under the maps. In that case, it is enough to deduce that subsample clustering corresponds to the clustering on the entire data set X. Elsewhere,given a varying Big Data set X, clusters can appear, disappear, merge or even split into distinct clusters.

Functoriality can be instrumental in studying this analysis behaviour. For all $t_0 < t_1$, we represent TDS withinclusters t_0 and t_1 as $X[t_0, t_1]$. For all $t_0 < t_1 < t_2 < t_3$, the point cloud data set results in diagram 2.3.2. If functoriality is applied in the clustering scheme, a correspondent diagram 2.3.3 of the data set is obtained. As revealed in the illustration, this set will contain an over-time clustering behaviour. Illustration 2.3.4 corresponds to a unit cluster at an initial time t_0, which breaks into two collections within the interval $[t_1, t_2]$that finally merge back to the interval $[t_2, t_3]$. Despite the development of fresh BD analysis techniques for complex BDS, shape identification andinterpretation have increasingly proven more challenging to visualize. Because of the existence of more structure to be mined from this BDS thanthe traditional ones can output, a remarkable new method of "shape" identification of these BDS is TDA. Several methods exist to construct shapes from point cloud data. One of these methods is described herein: We encircle every TDP within a "ball" whose radius is ε centered within the TDP. While ε size increases, the cluster no longer looks like isolated points but gradually gains shape. As it gets larger, an irregular unit component emerges. This technique is therefore used to generate a SC, which begins with vertices as TDPs. Wherever an intersection occurs between any two balls, an edge is inserted in the middle, while three edges bounded face is added as an intersection among any 3 balls occurs. As this process continues, a high-dimensional face with n+1 intersections is created. This is referred to as a *Čech* complex.

3.3. The Algorithm and Simulations of COVID-19 Cases

High-dimensional gets transformed into lower-dimensional graphical forms when passed through the Mapper algorithm's four stages. The 1st and 2nd proportions of the three-dimensional data computed the graph shape shown in Figure 2. During the last stage, adjacent bins with common data points (i.e. with nonempty intersections) get connected together into a simple skeleton summary of the original dataset. Compared to the standard dimensionality reduction techniques, the Mapper algorithm yields better results as it leverages the techniques from both clustering and dimensionality reduction of the initial highly featured space. As a demerit, standard DR techniques (DR) extrapolate TDPs to a minimal reduction state, from whence examination is done. Therefore, as a consequence, in as much as an accurate prediction is a motivation, this method always results in the loss of some information from the dataset due to interpretability. Comparatively, Mapper takes advantage of the lens from the lower dimension to optimize the data in the initial highly dimensional space. The datasets used in this study are downloaded from Kaggle.com, an open-source community of data scientists, ML, and a huge published Repository of BD sets. In our coding, the world is represented by a Hausdorff space X and subspaces of H represent regions or countries unless otherwise stated. These subspaces represent Big Data sets. We specifically focus on a collection of time series COVID-19 datasets of cases reported from all countries of all the six world continents (Africa, Europe, Asia, North America, South America and Oceania) starting February 24*th*, 2020, until June 28*th*, 2021. The dataset contained 98,904 rows and 60 columns. The huge volume of this dataset qualifies it as a candidate for a Bid Data set. The next section describes the connectedness of our data set.

Given $H_1, H_2 \subseteq H$, such that $H_1 \cap H_2 = \emptyset$, then collections generated by H_1 and H_2 might be equated to non-empty intersections, and these Hausdorff properties come in handy in building a SC. As a result, this result leads to a multiresolution" map of the TDS. The connectedness of points in a topological space means the nonexistence of separation between the points in the topological space. Let **X** be our COVID-19 data set throughout this study. We apply a non-linear fit on X by reducing the dimensionality while our data set's geometric structure, shape and connectivity remain preserved.

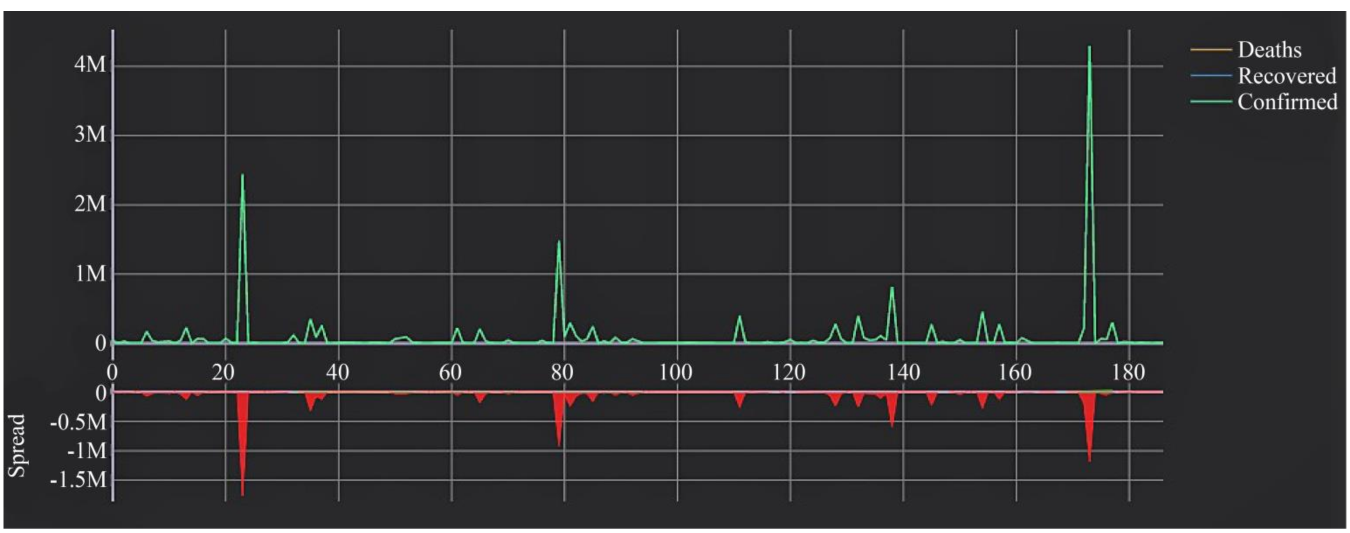

Fig. 3 Line plot visualization of dataset H revealing the total confirmed Covid-19 infection cases globally

After Exploratory Data Analysis (EDA) and dimensionality reduction on our COVID-19 data set, denoted by H, we produce a few line graph visualizations to reveal the relationship between the total confirmed initially, recovered and deaths cases across the six world continents. A visualization of the total confirmed COVID-19 cases in the world is revealed ata glance through the line graph in Figure 3.

Figure 4 gives a line graph visualization of the total recovered cases in the world. A visualization of both the totalconfirmed, recovered and death cases of COVID-19 in the world is also depicted here. In addition, after performing Exploratory Data Analysis and dimensionality reduction on our COVID-19 data set denoted by H, we compute surface plot visualizations to reveal the topological structure and graphical properties from our data set H. We use Python libraries to produce surface plot visualizations that reveal the structured shape between the total confirmed, total recovered and total deaths cases across the six world continents. Each TDP is enclosed around a ball at the center at a point with a radius ε. While ε increases in size, the cloud gradually ceases to appear as secluded points but slowly gains shape. As the output increases in size, a single shapeless blob emerges. This event eventually generates a SC, which starts with vertices as TDPs. An edge is inserted betweenthe two balls as they intersect. Meanwhile, a face bounded by the three edges is added as 3 balls intersect.

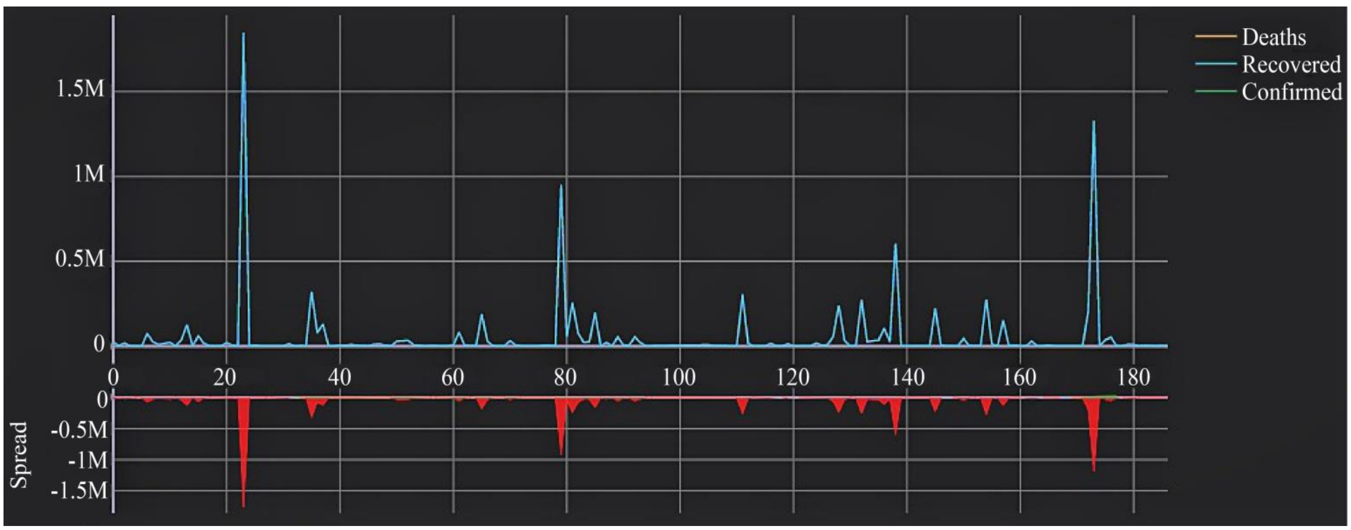

Fig. 4 Line plot visualization of dataset H revealing the total recovered cases globally

4. Conclusion

In summary, we have described aspects of Topological Data Analysis using techniques like Persistent Homology and Simplicial Complex. We then explored Big Data sets of COVID-19 with extremely large and highly complex properties. After that, we discussed the characterization and significance of the Hausdorff Spaces an AI in this work. Lastly, we have described simulations within the Python programming language with applications in TDA.

References

[1] Ashish Dure, "The Effects of Human–AI in IoT," *Trans AI*, vol. 5, pp. 42-97, 2021.

[2] Gerald Beer, "Upper Semicontinuous Functions and the Stone Approximation Theorem," *Journal of Approximation Theory*, vol. 34, pp. 1-11, 1982. [Google Scholar] [Publisher Link]

[3] Y. Chen, Y. Cho, and L. Yang, "Note on the Results with Lower Semicontinuity in Topological Spaces," *Bulletin of the Korean Mathematical Society*, vol. 39, pp. 535-541, 2002. [Google Scholar] [Publisher Link]

[4] J. David, and S. Duke, *An Introduction to Hausdorff Spaces*, 4th ed., John Wiley and Sons, Inc., 2011.

[5] B. Hantoute, "On AI and Optimization," *SIAM Journal on Optimization*, vol. 13, pp. 84-93, 2022.

[6] K. John, *AI and Topological Spaces*, John Wiley and Sons, New York, 2020.

[7] Frans Van Gool, "Lower Semicontinuous Functions with Values in a Continuous Lattice," *Commentationes Mathematicae Universitatis Carolina,* vol. 33, no. 3, pp. 505-523, 1992. [Google Scholar] [Publisher Link]

[8] E. Hernndez, and R. Lopez, "A New Notion of Semi-Continuity of Vector Functions and Its Properties," *Journal of Optimization,* vol. 39, pp. 1831-1846, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jeremiah Ratican, and James Hutson, "The Six Emotional Dimension (6DE) Model: A Multidimensional Approach to Analyzing Human Emotions and Unlocking the Potential of Emotionally Intelligent Artificial Intelligence (AI) via Large Language Models (LLM)," *DS Journal of Artificial Intelligence and Robotics,* vol. 1, no. 1, pp. 44-51, 2023. [Google Scholar] [Publisher Link]

[10] Juan M. Górriz et al., "Artificial Intelligence within the Interplay between Natural and Artificial Computation: Advances in Data Science, Trends and Applications," *Neurocomputing*, vol. 410, pp. 237-270, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jun-Ho Huh, and Yeong-Soek Seo, "Understanding Edge Computing: Engineering Evolution with Artificial Intelligence," *IEEE Access,* vol. 7, pp. 164229-164245, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] L. Jord, *TDA and Convexity,* University of Toronto, 2014.

[13] F. Kurey, *Introductory Topology with Applications*, John Wiley and Sons, 2016.

[14] P. Kumlin, *A Note on Topological Spaces and L_p-Spaces*, Functional Analysis Lecture Notes, Chalmers, 2003.

[15] Andrew J. Kurdila, and Michael Zabarankin, *Convex Functional Analysis, Systems and Control, Foundations and Applications,* Birkhauser Verlag, Basel, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[16] D. Nuredi, "Topological Data Sets and Functions," *Ukrainian Mathematical Journal,* vol. 9, pp. 54-63, 2019.

[17] Frédéric Chazal, and Bertrand Michel, "An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists," *Frontiers in Artificial Intelligence*, vol. 4, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Arnold Mashud Abukari, and Akhmad Daniel Sembiring, "Enhancing Social Media Experience through Voice over Internet Protocol (VOIP) Using Asterisk and PHP," *International Journal of Information Technology*, vol. 6, no. 2, pp. 1-8, 2020. [Google Scholar] [Publisher Link]

[19] Xing-Wei Xu et al., "Improved Fish Migration Optimization with the Opposition Learning Based on Elimination Principle for Cluster Head Selection," *Wireless Networks*, vol. 28, no. 3, pp. 1017-1038, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Aliaa F. Raslan et al., "An Improved Sunflower Optimization Algorithm for Cluster Head Selection in the Internet of Things," *IEEE Access,* vol. 9, pp. 156171-156186, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Rambabu Bandi, Venugopal Reddy Ananthula, and Sengathir Janakiraman, "Self-Adapting Differential Search Strategies Improved Artificial Bee Colony Algorithm-Based Cluster Head Selection Scheme for WSNs," *Wireless Personal Communications,* vol. 121, no. 3, pp. 2251-2272, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Farzad Kiani, Amir Seyyedabbasi, and Sajjad Nematzadeh, "Improving the Performance of Hierarchical Wireless Sensor Networks Using the Metaheuristic Algorithms: Efficient Cluster Head Selection," *Sensor Review*, vol. 41, no. 4, pp. 368-381, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] K.C. Avinash Khatri et al., "Genetic Algorithm Based Techno-Economic Optimization of an Isolated Hybrid Energy System," *ICTACT Journal on Microelectronics,* vol. 8, no. 4, pp. 1447-1450, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Mandli Rami Reddy et al., "Energy-Efficient Cluster Head Selection in Wireless Sensor Networks Using an Improved Grey Wolf Optimization Algorithm," *Computers*, vol. 12, no. 2, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Nirmal Adhikari, J. Logeshwaran, and T. Kiruthiga, "The Artificially Intelligent Switching Framework for Terminal Access Provides Smart Routing in Modern Computer Networks," *BOHR International Journal of Smart Computing and Information Technology*, vol. 3, no. 1, pp. 45-50, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] A. Vaniprabha et al., "Examination of the Effects of Long-Term COVID-19 Impacts on Patients with Neurological Disabilities Using a Neuro Machine Learning Model," *BOHR International Journal of Neurology and Neuroscience*, vol. 1, no. 1, pp. 21-28, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Bhairo Singh Rajawat et al., "Improved Election of Cluster Head Using CH-PSO for Different Scenarios in VANET," *Communication, Networks and Computing: First International Conference*, vol. 839, pp. 110-120, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[28] B. Gopi, J. Logeshwaran, and T. Kiruthiga, "An Innovation in the Development of a Mobile Radio Model for A Dual-Band Transceiver in Wireless Cellular Communication," *BOHR International Journal of Computational Intelligence and Communication Network*, vol. 1, no. 1, pp. 27-32, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] J. Logeshwaran et al., "The Role of Integrated Structured Cabling System (ISCS) for Reliable Bandwidth Optimization in High-Speed Communication Network," *ICTACT Journal on Communication Technology*, vol. 13, no. 1, pp. 2635–2639, 2022. [Google Scholar] [Publisher Link]

[30] Adhirath Kapoor et al., "Ransomware Detection, Avoidance, and Mitigation Scheme: A Review and Future Directions," *Sustainability*, vol. 14, no. 1, pp. pp. 1-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Eugenio Montefusco, "Lower Semi-Continuity of Functionals via the Concentration-Compactness Principle," *Journal of Mathematical Analysis and Applications,* vol. 263, pp. 264-276, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[32] J. Moreau, *Convexity and Duality in Functional Analysis and Optimization,* Academis Press, New York, 1966.

[33] A.A. Offia, "On Convex Optimization in Hilbert Spaces," *International Journal of Mathematics and Statistics Invention*, vol. 8, no. 4, pp. 7-9, 2020. [Google Scholar] [Publisher Link]

[34] Pawan Budhwar et al., "Artificial Intelligence–Challenges and Opportunities for International HRM: A Review and Research Agenda," *The International Journal of Human Resource Management,* vol. 33, no. 6, pp. 1065-1097, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[35] P. Rajadurai, "Machine Learning-based Secure Cloud-IoT Monitoring System for Wireless Communications," *DS Journal of Artificial Intelligence and Robotics,* vol. 1, no. 1, pp. 37-43, 2023. [Publisher Link]

[36] Jason Lively, James Hutson, and Elizabeth Melick, "Integrating AI-Generative Tools in Web Design Education: Enhancing Student Aesthetic and Creative Copy Capabilities using Image and Text-Based AI Generators," *DS Journal of Artificial Intelligence and Robotics,* vol. 1, no. 1, pp. 37-43, 2023. [Google Scholar] [Publisher Link]

[37] Scott Varagona, "*Inverse Limits with Upper Semicontinuous Bonding Functions and Indecomposability*," Thesis, University of Hauston, 2011. [Google Scholar] [Publisher Link]

[38] Jean-Philippe Vial, "Strong Convexity of Set and Functions and TDA*,*" *Journal of Mathematical Economics,* vol. 9, pp. 187-205, 1982. [Publisher Link]

[39] Zili Wu, "Uniform Convergence Theorems Motivated by Dini's Theorem for a Sequence of Functions," *Journal of Mathematical Analysis,* vol. 11, no. 6, pp. 27-36, 2020. [Google Scholar] [Publisher Link]

[40] Hsinchun Chen, Roger H.L. Chiang, and Veda C. Storey, "Business Intelligence and Analytics: from Big Data to Big Impact," *Management Information Systems Research Center,* vol. 36, no. 4, pp. 1165-1188, 2012. [CrossRef] [Google Scholar] [Publisher Link]