BIBLIOGRAPHIC INFORMATION SYSTEM

JOURNAL FULL TITLE: Journal of Biomedical Research & Environmental Sciences

ABBREVIATION (NLM): J Biomed Res Environ Sci **ISSN:** 2766-2276 **WEBSITE:** https://www.jelsciences.com

SCOPE & COVERAGE

- **Sections Covered:** 34 specialized sections spanning 143 topics across Medicine, Biology, Environmental Sciences, and General Science
- Ensures broad interdisciplinary visibility for high-impact research.

PUBLICATION FEATURES

- **Review Process:** Double-blind peer review ensuring transparency and quality
- **Time to Publication:** Rapid 21-day review-to-publication cycle
- **Frequency:** Published monthly
- **Plagiarism Screening:** All submissions checked with iThenticate

INDEXING & RECOGNITION

- **Indexed in:** Google Scholar, IndexCopernicus (ICV 2022: 88.03)
- **DOI:** Registered with CrossRef (10.37871) for long-term discoverability
- **Visibility:** Articles accessible worldwide across universities, research institutions, and libraries

OPEN ACCESS POLICY

- Fully Open Access journal under Creative Commons Attribution 4.0 License (CC BY 4.0)
- Free, unrestricted access to all articles globally

GLOBAL ENGAGEMENT

- **Research Reach:** Welcomes contributions worldwide
- **Managing Entity:** SciRes Literature LLC, USA
- **Language of Publication:** English

SUBMISSION DETAILS

- Manuscripts in Word (.doc/.docx) format accepted

SUBMISSION OPTIONS

- **Online:** https://www.jelsciences.com/submit-your-paper.php
- **Email:** support@jelsciences.com, support@jbresonline.com

Vision: The Journal of Biomedical Research & Environmental Sciences (JBRES) is dedicated to advancing science and technology by providing a global platform for innovation, knowledge exchange, and collaboration. Our vision is to empower researchers and scientists worldwide, offering equal opportunities to share ideas, expand careers, and contribute to discoveries that shape a healthier, sustainable future for humanity.

JOURNAL OF BIOMEDICAL RESEARCH &ENVIRONMENTAL SCIENCES
issn: 2766-2276
SUBJECT AREA(S): COMPUTER SCIENCE
DOI: 10.37871

SHORT COMMUNICATION

Explainability for Graph Spectral Analysis

Mieczyslaw A Klopotek, Slawomir T Wierzchon*, Bartlomiej Starosta, Piotr Borkowski and Dariusz Czerski

Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland

A Few Words About Graph Spectral Analysis

Spectral graph analysis (GSA for short) [1] is a method used to cluster objects when a method is available for determining the degree of affinity for all pairs of objects. It proved to be a very efficient method, but it missed the capability to explain the clustering results. This paper presents our contribution to the field of explainable artificial intelligence when GSA is applied to collections of textual documents. The major strength of our methodology consists of explanations from first principles, contrary to other approaches that describe cluster contents rather than saying why the documents are brought together. GSA was originally developed as a method for approximating cuts in graphs, but was later extended to arbitrary sets of objects. These objects form the vertices of a graph whose edges represent the affinity between pairs of objects. The absence of an edge between a pair of nodes indicates a lack of affinity between the corresponding objects. Although GSA is more involved compared to typical clustering methods, it is more efficient when faced with unknown, non-linear structures. An example here could be text documents, where similarity is calculated as cosine similarity in a certain embedding, such as a term vector space (denoted TVS), word2vec, GloVe, BERT, or other embeddings [2].

Theclusteringisperformedasfollows:Let S beasimilaritymatrix between the objects (with diagonal set to 0), and D be a diagonal matrix with the diagonal elements being sums of corresponding rows of S. One derives some Laplacian of S, e.g. combinatorial Laplacian $L = D - S$, normalized Laplacian $\mathcal{L} = D^{-1/2} L D^{-1/2}$, or some other. Then one performs eigendecomposition of this Laplacian and uses eigenvectors associated with the k lowest eigenvalues as an embedding of the objects , where k is the envisaged number of clusters. . Then one performs k-means clustering in this embedding. The advantage is as follows: clustering of documents

*Corresponding author(s)

Slawomir T Wierzchon, Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland

Email: slawomir.wierzchon@ipipan.waw.pl

DOI: 10.37871/jbres2301

Submitted: 20 May 2026

Accepted: 25 May 2026

Published: 26 May 2026

How to cite this article: Klopotek MA, Wierzchon ST, Starosta B, Borkowski P, Czerski D. Explainability for Graph Spectral Analysis. J Biomed Res Environ Sci. 2026 May 26; 7(5): 3. Doi: 10.37872/jbres2301

in TVS would mean clustering in a multi-thousand-dimensional space, while clustering in the space spanned by these eigenvectors requires only k dimensions.

Explainability - A Grievient Problem of GSA

Although this methodology has proven to be quite effective, it has drawbacks typical of many artificial intelligence algorithms: the result of a clustering is hard to explain, since the embedding in the space of eigenvectors of Laplacians has nothing to do with the content of, for example, the documents being clustered.

Many explanation methodologies rely on applying a classification algorithm based on attributes of the original objects (e.g. ID3, C4.5, CART). In our opinion, this does not apply to document collections. After all, it is unnatural for them for at least two reasons: (1) too few words would be taken into account, (2) a statement that something belongs to a cluster because it does not contain a word is a poor argument. Besides, in text clustering, the set of most significant words is used as a cluster description, not word presence. Last but not least, such an approach is a descriptive one and does not state the reason why a document belongs to a cluster.

Our Solution

Therefore, we have chosen a different pathway. Recall that if one clusters texts in TVS via k-means, then one would obtain the causal explanation of the clusters as follows: Compute the cluster center vector in TVS from the documents belonging to that cluster. So each term is assigned a value, being the coordinate of this vector in TVS. Take the top n he highest coordinates values. This is the causal description of the cluster. Now we apply the following trick: We would get in this way cluster descriptions of clusters obtained via any lustering algorithm if we could prove that this alternative algorithm delivers the same clusters as the k-means in TVS would do. This is what we actually apply to get explainable GSA clustering [3]. We have created a two-stage proof for such an equivalence. We invented a kernel matrix K which has the property that clustering in the space of its eigenvectors is exactly equivalent to clustering via k-means in TVS on the one hand, and approximately equivalent to combinatorial Laplacian-based GSA clustering. Let A be a matrix of the form: $A = \mathbf{1}\mathbf{1}^T - I - S$, where I the identity matrix, and $\mathbf{1}$ is the (column) vector consisting of ones, both of appropriate dimensions. The matrix A is nonnegative and has a diagonal equal to zero, so it may be considered as a kind of (squared) pseudo-distance, needed by the Gower embedding method [4]. Let K be a doubly centered matrix of order n defined as follows $K = -\frac{1}{2}(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)A(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)$. Now, from the conceptual point of view, we can cluster the data, each time using k-means, in the TVS space, obtaining a clustering $\mathcal{C}_{TVS}$, or in the space spanned by eigenvectors of K matrix obtaining a clustering $\mathcal{C}_K$, or in the GSA space obtaining a clustering $\mathcal{C}_{GSA}$, each time of course using the proper representation. As we have proven, the clusterings of documents $\mathcal{C}_{TVS}$, $\mathcal{C}_K$ and $\mathcal{C}_{GSA}$ are (nearly) identical. This serves as a bridge to the following procedure: group the text data using GSA, obtaining $\mathcal{C}_{GSA}$, and determine cluster membership explanation using TVS, by assuming that $\mathcal{C}_{GSA} = \mathcal{C}_{TVS}$. Clustering in GSA is much simpler than in TVS, as the dimensionality is orders of magnitude lower. If we take for granted that $\mathcal{C}_{GSA} = \mathcal{C}_{TVS}$, then for each cluster in $\mathcal{C}_{TVS}$ we can determine its cluster center by just computing the document embedding vector means of cluster members, which is much easier than clustering. With the computed cluster centers, the explanation methodology of TVS is applicable straightforwardly. We have extended this approach by creating an appropriate kernel atrix bridging conceptually clustering based on the normalized Laplacian and explanation via weighted cluster in TVS.

If the documents, however, are embedded not in TVS, but rather in word2vec, GloVe [5] or BERT [6] like space, we need a modification of this approach, replacing the simplistic identification of most important terms via decreasing sorting the cluster center vector by computing scalar products of the cluster center vector and individual term embedding vectors and sorting these products. The remaining procedure turns out to be the same as above. Hereby, we had to handle the issue of negative similarities between documents. Our approach was a geometrical one [7].

In summary, with our proposals, the application of GSA can be extended to domains where it was rejected due to missing explainability and issues regarding negativity of document similarity. Some issues remain unresolved, in particular GSA, due to the usage of a squared similarity matrix, is highly memory-consuming when the document collections are high (hundreds of thousands of documents or more). There exist efforts to overcome this limitation, but they constitute new challenges to explainability with which we intend to deal in the future research.

References

1. Wierzchoń ST, Kłopotek MA. Modern Algorithms of Cluster Analysis. Studies in Big Data. Vol. 34. Springer Verlag; 2018.
2. Cao H. Recent advances in text embedding: A comprehensive review of top-performing methods on the MTEB benchmark. CoRR. 2024;abs/2406.01607v1. doi:10.48550/arXiv.2406.01607
3. Starosta B, Kłopotek MA, Wierzchoń ST, Czerski D, Sydow M, Borkowski P. Explainable graph spectral clustering of text documents. PLoS One. 2025;20(2):e0313238. doi:10.1371/journal.pone.0313238
4. Borg I, Groenen PJF. Modern Multidimensional Scaling: Theory and Applications. 2nd ed. Springer Series in Statistics. New York (NY): Springer; 2005. doi:10.1007/0-387-28981-X
5. Kłopotek MA, Wierzchoń ST, Starosta B, Borkowski P, Czerski D, Laskowski E. Explainable graph spectral clustering for glove-like text embeddings. 2025.
6. Kłopotek MA, Wierzchoń ST, Starosta B, Borkowski P, Czerski D. Towards explainable graph spectral clustering for BERT embeddings. J Autom Mob Robot Intell Syst. 2026;20(1):53-65.
7. Kłopotek MA, Wierzchoń ST, Starosta B, Czerski D, Borkowski P. A method for handling negative similarities in explainable graph spectral clustering of text documents – extended version. 2025.