Editorial

We are pleased to present the second issue of the *Journal of Data Processing*, featuring three studies that advance methodological and applied research across diverse data-driven domains.

The inaugural article, **" Modelling EEG Signals as Graphs: A GNN-Based Framework for Eye State Detection with Embedding Space Analysis,"** introduces a graph-centric approach to electroencephalography, representing electrodes as nodes and their functional interdependencies as edges. The authors provide a comprehensive treatment of graph neural network (GNN) architectures, detailing spectral and spatial convolutions, attention mechanisms, and spatiotemporal adaptations. Additionally, the study explores embedding space dynamics to improve the transparency of learned representations, synthesizing recent breakthroughs in graph-based neural decoding while outlining critical challenges and avenues for future research.

The second contribution, **"Evaluating Clustering Strategies for Categorical Microbial Data: From K-Means Limitations to Gower-Based Hierarchical Optimization,"** addresses the complexities of partitioning categorical biological datasets. Analyzing a corpus of approximately 200 bacterial species characterized by taxonomic, ecological, and pathogenic traits, the authors contrast a conventional K-Means pipeline (applied to one-hot encoded features) with a distance-aware hierarchical clustering strategy grounded in Gower dissimilarity. Their findings position the Gower-driven framework as a more robust alternative for microbial classification, demonstrating superior cluster stability, enhanced biological interpretability, and stronger alignment with contemporary ecological taxonomies.

The final paper, **"A Quantitative and Semantic Clustering Framework for High-Risk AI Systems under the EU AI Act,"** proposes a hybrid analytical architecture that fuses semantic encoding, machine learning, and multidimensional risk assessment. To capture risk complexity beyond mere semantic similarity, the authors develop a composite scoring mechanism that integrates impact domains, risk typologies, decision-criticality thresholds, and human-centric consequences. The study concludes by advocating for nuanced, evidence-based regulatory approaches that tailor oversight mechanisms to the specific structural and risk profiles of high-risk artificial intelligence deployments.

Collectively, these contributions reflect significant advancements in data processing methodologies and their interdisciplinary applications. We trust that readers will find them both insightful and impactful.

Editors