

# Tutorial Proposal for FLINS-ISKE2026 Conference

**Title:** Deep Clustering: Techniques and Future Perspectives

**Duration:** 3 Hours

**Organizers:** Bo Peng<sup>1</sup>, Jie Lu<sup>1</sup>

**Presenter:** Bo Peng

<sup>1</sup> *University of Technology Sydney, Australia*

*Email:* [Bo.Peng-7@student.uts.edu.au](mailto:Bo.Peng-7@student.uts.edu.au); [Jie.Lu@uts.edu.au](mailto:Jie.Lu@uts.edu.au);

## 1 Introduction

Clustering is a fundamental problem in machine learning that aims to group data instances into several clusters such that instances within the same cluster share similar semantics, while those in different clusters are dissimilar. By organizing data in this way, clustering can reveal the inherent semantic structure underlying the data, thereby supporting more effective decision-making.

In the early stages of clustering research, various classical techniques—such as K-means, spectral clustering, and DBSCAN—were developed to partition instances based on similarity measures or density patterns in raw or linearly transformed data. Although these shallow clustering methods are theoretically well-founded and highly interpretable, they struggle to handle the complex, high-dimensional, and nonlinear data commonly encountered in real-world applications. Deep clustering addresses these limitations by leveraging the powerful representation learning capabilities of neural networks to extract more informative features for clustering.

Motivated by the growing interest in deep clustering, the primary goal of this tutorial is to provide a comprehensive review of deep clustering methods by categorizing them according to prior knowledge. We observe that innovations in prior knowledge generally follow two major trends: (i) from mining to constructing prior knowledge, and (ii) from internal to external sources of prior information. In addition, we present a benchmark evaluation on five widely used datasets and analyze the performance of methods incorporating different types of priors. Finally, we will also introduce some potential challenges for future research. We do hope this tutorial will offer new insights and inspire future research in the deep clustering community.

In summary, the objective and motivation of this tutorial is three-fold:

- present milestones in modern deep clustering
- review a broad selection of previous works in different categories
- uncover future directions of deep clustering.

## 2 Tutorial Outline and Schedule

The outline and schedule of this 3-hour tutorial is given as follows:

1. **Opening, Background, and Key Insights** (20 minutes)
  - Tutorial Goals and Real-world Motivations.
  - Problem Definition of Deep Clustering
2. **Deep Clustering with Priors**
  - Deep Clustering with Structure Prior (20 minutes)
  - Deep Clustering with Distribution Prior (20 minutes)
  - Deep Clustering with Augmentation Invariance (20 minutes)
  - Deep Clustering with Neighborhood Consistency (20 minutes)
  - Deep Clustering with Pseudo-labeling (20 minutes)
  - Deep Clustering with External Knowledge (20 minutes)
3. **Conclusion and Future Trends** (20 minutes)

- Fairness, Generalizability, Interpretability, and Robustness
- Non-parametric deep clustering
- Deep Clustering for Long-tailed/Fine-grained Data

4. Q&A (20 minutes)

### 3 Teaching Methods

We promote the tutorial as follows:

- We plan to accompany our tutorial with intuitive examples.
- We will provide the participants with specific time slots to ask their questions.
- We will release all related materials (e.g., presentation slides, survey paper, references, open-source data and code, etc.) in advance.

### 4 Target Audience

Clustering is a broad and foundational topic in machine learning. As such, our potential audience includes academic researchers and industry practitioners working in machine learning, data mining, computer vision, and natural language processing. We assume that the audience has a general background in probability and statistics, along with a working knowledge of deep neural networks.