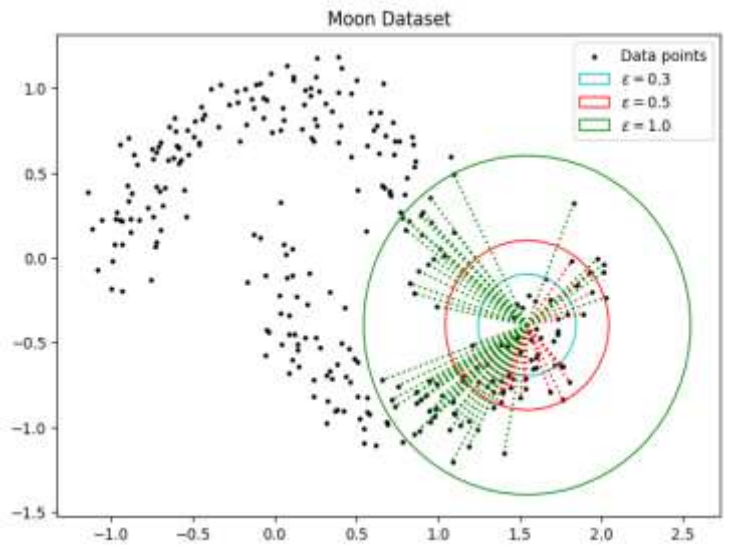
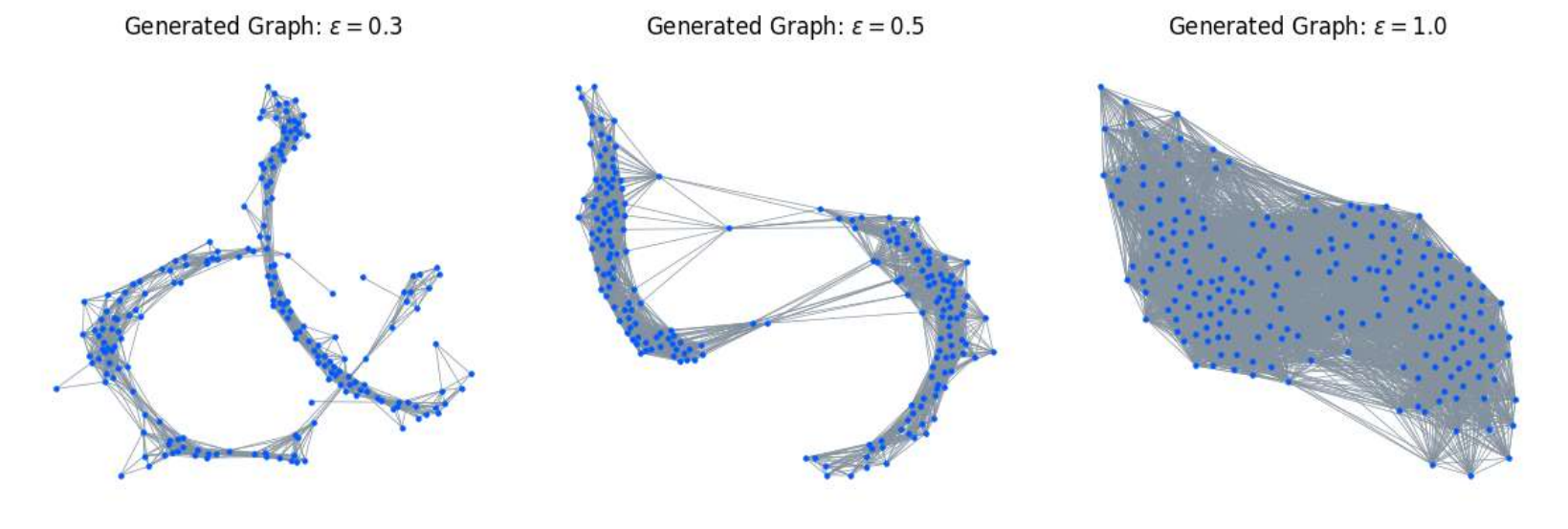
Generate Graph Structure from Vector Data



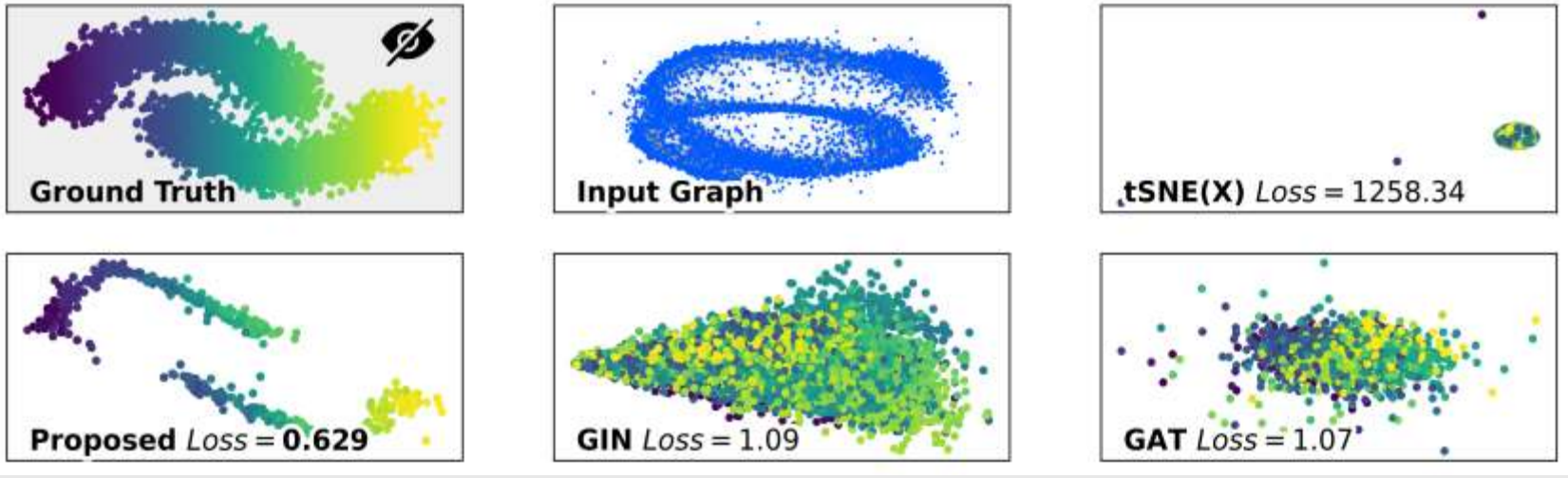


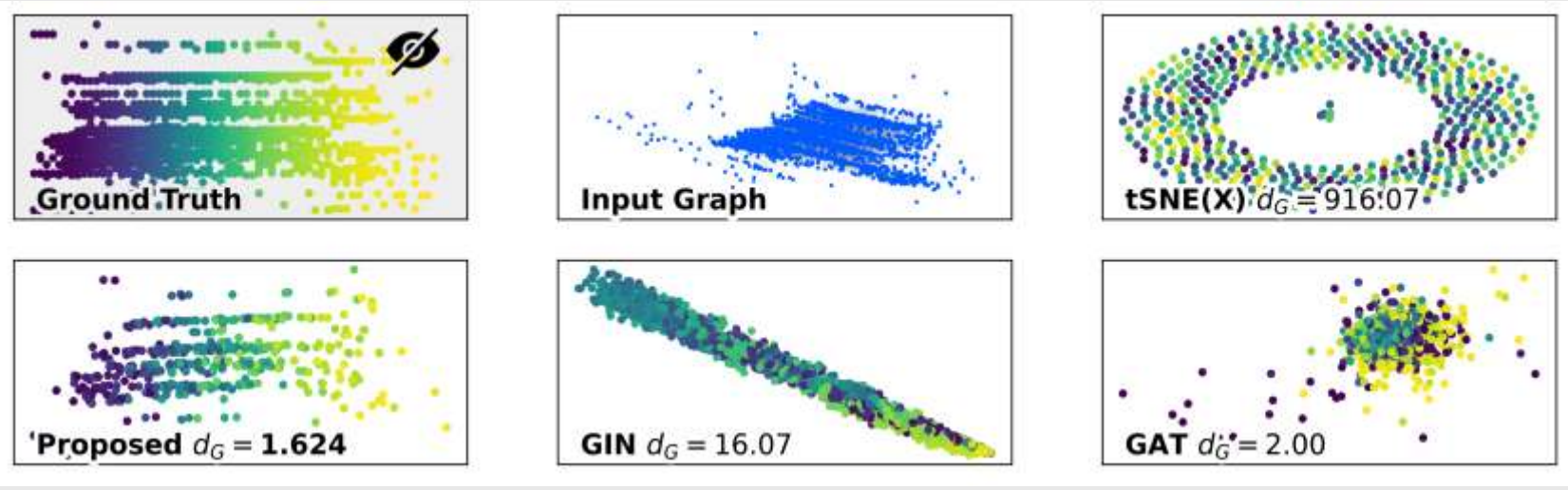
https://github.com/ChengxiPan/INFSCI-2415/blob/main/main.ipynb

The above two images describe the progress of a Clustering Algorithm called $\varepsilon - boom$, which generate different graph structures with the increase of ε.

- The left image shows the original distribution of Moon Dataset and the basic idea of generating graph using $\varepsilon - boom$.
  - In the shown case, $\varepsilon = [0.3, 0.5, 1.0]$, respectively.
  - Draw a circle (with radius = ε) around each node, and build edges to other nodes that fall within this circle. With the increase of ε, the number of generated edges also increase.
- The right image illustrates generated graph by different ε.
  - Nodes have few connections when ε is small, meaning it can only capture local information.
  - When ε grows to a limit, the spring_layout of graph resembles the original distribution.
  - It can be predicted that connections will be built between every 2 nodes when ε is extremely large.

Reveal the original distribution from Graph





The left two images describe the whole process of two steps of my work on two datasets, with the first called moon and the second called adult:

↖: the original distribution of the dataset
↑ : the generated graph using $\varepsilon - boom$
↗: the result of t-SNE
↙: the result of my reconstruction method
↓ : the result of Graph Isomorphism Network
↘: thhe result of Graph Attention Network

In this case, the loss function applied is called d_G, which enforces the embeddings output by the model to align to the target embedding space while preserving geometric properties

It is notable that the proposed reconstruction method is efficient in feature recovery.

https://github.com/ChengxiPan/INFSCI-2415

Presented by: Chengxi Pan [INFSCI 2415]