# Lab 06

1. Increase the N value from 20 (original value) to 200 with multiple N values in between and observe the change of graph density and degree distribution

* The graph density for the random graph with 200 nodes and 400 edges is approximately \*\*0.0201\*\*. This indicates a relatively sparse graph, meaning that only about 2% of the possible edges between nodes are present.
* In terms of degree distribution, the degree sequence shows that some nodes have as many as 11 connections, while others have none, reflecting the random nature of the graph. If you increase the number of nodes (N), the density will likely decrease further unless you increase the number of edges (E) accordingly.

1. Explain

2.1 differences between supervised learning, self-supervised learning and semi-supervised learning methods

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Supervised Learning** | **Self-Supervised Learning** | **Semi-Supervised Learning** |
| Label Availability | Requires labeled data | Generates labels from data | Uses both labeled and unlabeled data |
| Data Requirements | Large labeled dataset | Unlabeled dataset | Small labeled + large unlabeled dataset |
| Example | Image classification | Predicting missing words in a sentence | Sentiment analysis with few labeled reviews |

2.2 Explain the differences between transductive learning and inductive learning

|  |  |  |
| --- | --- | --- |
| **Category** | **Transductive Learning** | **Inductive Learning** |
| Generalization | Does not generalize | Generalizes to unseen data |
| Test Data | Known during training | Unseen during training |
| Example | Graph-based models | Typical supervised learning models |

1. Explain
   1. Increase the number of epochs from 50 to 500 and observe the change in validation accuracy

* Validation Accuracy Increases , But after 300 it doesn’t change
  1. Experiment without self-loops added to GCNConv() layers in the GCN() model and detail the model accuracy increase/decrease
* decrease in accuracy compared to the default model with self-loops.
  1. Increase the number of GCNConv() layers in the GCN() model upto 8 layers from original 3 layers.
* accuracy decreases after adding too many layers, even with the best hyper-parameters.
  + 1. In\_channels and out\_channels in GCNConv() can be considered as hyper-parameters and you can use the best performing values you find.
* Accuracy improves, but increases training time
  + 1. Add skip connections between some of the GCNConv() layers and try to see if that can improve the model performance
* improved performance compared to the deep GCN without skip connections

1. Explain the differences between Message Passing GNN, graph convolution network (GCN), graph attention network (GAT) and GraphSAGE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Key Concept** | **Aggregation Method** | **Attention** | **Use Case** |
| **Message Passing GNN** | General framework for GNNs | Custom aggregation of neighboring nodes | No | General framework, adaptable for any GNN tasks |
| **GCN** | Convolution for graphs | Weighted sum of neighboring node features | No | Simple tasks like node classification |
| **GAT** | Attention-based GNN | Weighted sum with attention on neighbors | Yes | Complex tasks with variable neighbor importance |
| **GraphSAGE** | Efficient, inductive GNN | Sample and aggregate from fixed-size neighbors | No | Large graphs, inductive tasks like recommendations |