BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, Pilani
Pilani Campus
AUGS/ AGSR Division

SECOND SEMESTER 2025-26
COURSE HANDOUT

In addition to part I (General Handout for all courses appended to the Time table), this portion gives further specific details regarding the course.

Course No	**:**	**CS F426**
Course Title	**:**	**Graph Mining**
Instructor-in-Charge	**:**	**Vinti Agarwal**
Course webpage	**:**	**URL**

1. Scope and Objective of the Course:

This course studies managing and mining graphs which are massive and cannot be held in main memory, as the size of the applications is often very large. Classic examples of graphs are the web, social networks, computational biology, communication networks, etc. In some cases the entire graph is not available in the form of continuous stream. Edges are received continuously with time. The course includes the basics of the graphs, static and dynamic graphs, PageRank & random walks, graph or graph node classification, graph clustering, community detection, anomaly detection, frequent sub-graph mining. The course is designed to provide students with an understanding of parallel and streaming graph mining to deal with massive graphs and evolving graphs. The course also aims at providing a holistic view of dealing and mining with graphs.

2. **Text Books**:

T1: Agarwal Charu C. and Wang Haixun, Managing and Mining Graph Data, Springer

T2: Graph Representation Learning by William L. Hamilton

3. Reference Books:

R1: Jure Leskovec, Anand Rajaraman, Jeff Ullman. Mining of Massive Datasets. Book 2nd edition. Cambridge University Press.

R2: Chakrabarti, D. and Faloutsos, C., 2012. Graph mining: laws, tools, and case studies. Synthesis Lectures on Data Mining and Knowledge Discovery, 7(1), pp.1-207.

R3: Nagiza F. Samatova, William Hendrix, John Jenkins, Kanchana Padmanabhan, Arpan Chakraborty. Practical Graph Mining with R. CRC Press

R4: Ma Y., Tang J. Deep learning on Graphs. Cambridge University Press 2021

4. Course Plan:

Module No.		**Lecture Session**	**Reference**	**Learning outcomes**
M1: Graph Basics Static and dynamic graphs	Week 1-2	Course Overview	T1 Ch1 T2 Ch2	To understand graphs and their applications
		Eigenvalues and Eigenvectors computation, PageRank, HITS, Neumann Kernels, (Co-citation graph etc.)	T1 Ch 2 & Class notes	To analyze graphs by computing centrality measures and kernels.

	Week 3	Graph Laplacian, its types, Laplacian Eigenmaps, spectral clustering vs PCA	T2 Ch2	To use graph Laplacian–based frameworks to learn low-dimensional representations. To learn how to propagate features and labels in graphs, and different approaches to label diffusion
	Week 4	Graph diffusion, regularization framework, Feature/label diffusion kernels, Label spreading vs Label propagation	R1 Ch 5	
	Week 5	Similarity graphs: kNN, Gaussian, etc.	Class notes	To construct similarity graphs using distance-based methods To learn how to detect communities in a graph
		Graph clustering: Betweenness Centrality, Girvan- Newman Algorithm	T1 Ch 9 R1Ch10	
M2: Graph Mining	Week 6	Machine Learning on Graphs	T1 Ch13	To evaluate machine learning models on graphs by selecting appropriate loss functions and performance metrics for classification and clustering tasks
		Performance evaluation: Classification(basic & class specific metrics, macro/micro; precision, recall, F1-score); clustering (using prior knowledge, without prior knowledge)	Class notes	
		Loss functions (MSE, cross entropy, CCE, max margin/hinge loss, Supervised contrastive loss)	Class notes	
M3: Deep learning for graphs	Week 7	Neighborhood Reconstruction Methods: Factorization-based approaches, Random walk-based approaches, Limitations of shallow embeddings	Class Notes T2 Ch3	To learn how to solve a graph mining problems with deep learning
	Week 8	Multi-relational decoders: RESCAL, TransE, TransH, DisMult, ComplEx, comparison of their Representational abilities	T2 Ch4	To compare and select an appropriate multi-relational decoder model based on the relational patterns
	Week 9	Graph Neural Network (GNNs): Permutation invariance and equivariance, Neural Message Passing (section 5.1), Neighborhood Aggregation: Neighbourhood normalization, set aggregation, neighbourhood attention	T2 Ch5	Understand the core principles underlying Graph Neural Networks, including how they propagate and aggregate information over graph structures
	Week 10-11	Spectral graph learning: MPNN Message Passing Neural Networks), GCN(Graph Convolutional Networks), GAT(Graph Attention Networks)	Class notes	To understand and compare major graph neural network paradigms (both spectral and spatial)
		Spatial vs spectral graph learning methods: GRAPHSAGE		

M4: Recent advances in Graph Learning	Week 12	GNNs for Link Prediction, Recommendation System Graph agnostic vs graph aware models	Class notes	To apply graph neural networks to different tasks
	Week 13-15	Advanced topics: Graph Transformers, Graph Autoencoders, LLMs and Graphs, Graph foundational models. Spotlight papers discussion from KDD, ECML, LoG, AAAI, NeurIPS, etc.	Class notes	How Graph Transformers extend attention-based models to graph-structured data, and compare their advantages and limitations relative to message-passing GNNs. different integration paradigms of LLMs and GNNs, examining when LLMs serve as representation enhancers versus when they function as direct predictors in graph-based learning tasks.

Lab Plans:

NOTE: Lab sessions will be conducted using the Python programming language; therefore, prior knowledge of Python programming is a prerequisite. Preparatory materials for each laboratory session will be made available in advance to ensure that students are adequately prepared prior to attending the labs.

The tentative lab details are given below:

Lab 1	Introduction to Python and Networkx
Lab 2	PageRank algorithm without and with teleportation
Lab 3	Dimensionality reduction using PCA, Laplacian eigenmaps, and Spectral clustering
Lab 4	Label propagation & Label spreading
Lab 5	ML model training on graph data with CE, SCL losses, etc.
Lab 6	Graph Encoder-decoder model -1 (Matrix factorization, Random Walk)
Lab 7	Graph Encoder-decoder model -2 (RESCAL)
Lab 8	Vanilla GCN implementation
Lab 9	GAT model implementation
Lab 10	GraphSAGE model implementation
Lab 11	Graph Transformers/Graph autoencoders NGCF
Lab 12	Graphs and LLM in the RAG pipeline

Course Prerequisites: A basic knowledge of probability theory, data mining, and machine learning is sufficient. Formal certification in these courses is not required. But you must have sufficient knowledge of these domains to better understand the concepts of graph mining.

5. **Evaluation Scheme:**

Details of the course-project assessment

Component	Duration	Weightage (%)	Date & Time	Mode
Quiz	40 min	5%	23-02-2026 (11:00am)	Closed Book
Weekly Lab (2hrs)	2 hours	5%	Sat 3:00-5:00PM	IPC - 6116
Paper presentation + participation	30 min for each student	5%	Feb 26-March 03 2026	
Course project evaluation C1	30 min per group	5 %	Feb 26-March 03 2026	In-person
Mid sem exam	90 min	20 %	Mar 12, 2026	
Course project evaluation C2	45 min per group	10%	March 18-March 24	In-person
Course project evaluation C3	45 min per group	15%	April16-April 25	In-person
Comprehensive Examination	3 Hours	25%	11/05/2026 FN	(Partly Open book)
Course project evaluation C4	30 min per group	10%	During comprehensive exams(as per students' availability)	In-person

Project Progress Assessment (Tentative plan)

- **Assessment** 1 (Feb): Literature review and proposed methodology **5%**
- **Assessment** 2 (Mar): Methodology implementation (Results & Code(Github)) **10%**
- **Assessment** 3 (April-beginning): Final Report writing(Overleaf ACL format) +Viva **15%**
- **Assessment** 4 (April-end): Feedback incorporation and refinements **10%**

6. Chamber Consultation Hour: Once a week, with prior appointment through email.

7. Notices: All the notices/communications concerning this course will be through the **Nalanda LMS**.

8. Make-up Policy:

Prior Permission is a must, and Make-up shall be granted only in genuine cases based on the individual's need and circumstances.

Instructor-in-charge

CS F420