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

FIRST/SECOND SEMESTER 2022-23

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
TAs : Aditya Sharma, Harsh Mahajan

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 are often very large. Classic examples of graphs are web, social networks, computational biology, communication networking 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: William L. Hamilton. Graph Representation Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning

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.

4. Course Plan:

Module No.	Lecture Session	Reference
M1: Graph Basics Static and dynamic graphs	Lecture 1: Course Overview Graph preliminaries	T1 Ch1
	Lectures 2-5: PageRank, Triangle computation, Personalized Random walk	

	Lectures 5-6: Graph laplacian and regularization framework	
	Lectures 7-10: Applications of Graph laplacian: Spectral clustering, graph diffusion, Random walk	T1 Ch 2 & Class notes Class Notes
	Lectures 11-12: Case Studies: Feature/label smoothing	R1 Ch 5 T1 Ch 8
M2: Graph Mining	Lectures 13-16: Kernel methods on graphs, diffusion kernels, Multiscale Laplacian Graph Kernels.	
	Lecture 14-18: Node Classification Link Based methods Incremental methods Semi-supervised learning Bipartite Graphs SimRank Link Prediction: network proximity measures, recommender systems Graph classification	T1 Ch 11 Class Notes
	Lecture 19-22: Graph Clustering Clustering in Bipartite graphs Hierarchical Clustering Community Detection Graph Partitioning Spectral methods Graph Similarity and Alignment Iterative similarity methods Edge similarity scores Network Similarity Methods Network Alignment	T1 Ch 9
	Lecture 25-27: Sub-graph mining Sub-graph Enumeration Frequent Subgraph Mining Apriori-based Approach Pattern-growth Approach Mining Significant Graph Patterns Branch-and-Bound Approach	T1 Ch 12
	Lectures 28-30: Graph Summarization	
	Case Study	
M3:	Lectures 31-34:	T1 Ch13

Streaming Graphs	Streaming Graphs Stream Model for massive graphs Graph streams	Class Notes
M4: Deep learning for graphs	Lectures 35-42: Deep learning for graphs Graph Convolutional Networks (GCN), Multi-relational Graph Convolutional Networks (m-RGCN), etc.	

Course Prerequisites: Basic knowledge of probability theory, data mining and machine learning. 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:

Component	Duration	Weightage (%)	Date & Time	Nature of component (Close Book/ Open Book)
Mid-Semester Test	90 Min.	30	02/11/22 (Tentative)	Closed Book
Lab practice + 1 group project	(2 hrs per week + 1 month)	30 (15 %+ 15 %)	To be announced	To be announced
Comprehensive Examination	3 Hours	40	22/12/2022 (Tentative)	Closed Book

6. Chamber Consultation Hour: Once a week. Tuesday from 10:30 AM – 11:30 AM.

7. Notices: All the notices/communication concerning this course will be through the **CANVAS platform** only. You are requested to check this periodically. E-mail will be used as and when required.

8. Make-up Policy:

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

9. Note (if any):

- Assignment(s) (programming/reading) will be given to the students. This will immensely help the students in gaining a better understanding of the subject.

Instructor-in-charge

CS F426