New Jersey Institute of Technology

Digital Commons @ NJIT

Computer Science Syllabi

NJIT Syllabi

Fall 2023

CS 634: Data Mining

Usman Roshan

Follow this and additional works at: https://digitalcommons.njit.edu/cs-syllabi

Textbooks:
None required
Grading: 20% mid-term, 80% course project
Course Overview: This course will focus on methods to find patterns on data. We will start with unsupervised methods like k-means clustering, other clustering methods, and neural network representation learning methods. We will then look at text mining, recommender systems, self-supervised learning, and supervised methods.

Course plan:

Topic	Date	Notes
Introduction		
Unsupervised learning - k-means clustering and other clustering methods		
Running k-means in Python scikit-learn		
Data visualization with principal component analysis		
Supervised learning - linear models and support vector machines		
Representation learning - neural networks and autoencoders		
Text mining - word and document representations, regular expressions		
Recommender systems - vector-space modeling, document and image similarity search, searching in representation space		
Self-supervised learning - learning text and image representations		
Project presentations		