Department: Mathematics and Computer Sciences
Division: Computer Sciences
Level and Major: Graduate

Course Title: Special Topics in Data Mining
Number of Credits: 3
Prerequisite: Permission Group
Lecturer:

Course Description: Data Science projects life cycle and tools EDA Explore individual features, pairs and groups, clean up, missing values, leakages, feature evaluation Necessary python libraries: Pandas, NumPy, Numba / Dask Visualization Modeling (I assume you know the basics) Feature preprocessing and generation, Model selection, Model evaluation, Recommender systems Special Topics Modern Approaches in NLP Transfer Learning (in NLP or Vision) Spatial Data Science Auto-ML Time-series Active Learning

Course Goals and Objectives:

Course Topics:

- Data Science projects life cycle
- Data Analysis Basics (EDA)
- NumPy, Pandas, Dask
- Jupyter lab
- Visualization (Matplotlib, Seaborn)
- Data Clean up
- Feature Evaluation
- Feature preprocessing and generation
- Model evaluation and confusion matrix
- Recommender Systems
- Modern Approaches in NLP
- Transfer Learning
- Spatial Data Science
- Active Learning

Reading Resources:

Evaluation: