

Midterm

Question types

- Short answer questions
- Programming questions (NOT)
- Multi-choice questions
- True/False questions

Scope

- Introduction
- Data
 - Distance calculation
 - Sampling
 - Exploration (I will not ask you to write program; but you may be given some plots and I will ask you questions related to those plots)
- Classification
 - Decision trees (Gini index, entropy, information gain)
 - NOT: KNN, Bayesian Classifier, Logistic Regression, SVM
- Classification issues & concepts
 - Non-balanced datasets
 - Performance measurements: Contingence matrix, accuracy, F1, Precision, recall, ROC, AUC

How to take the exam

- Come to class
- Take it online (Lockdown browser + camera)
- Calculator
- Not cell phone
- 1 A4 page cheat sheet
- Blank paper
- Pencil/pen
- 1hr 15 minutes