



# Guidance of ML Project

(Report submission: due the end of the 16<sup>th</sup> week)



# ML Class

## ➤ Grade system:

Attendance + project + final exam (closed-book).

## ➤ Purpose of project

Practice of the general ML process, including problem definition, data collection, feature extraction, model training, and metric evaluation.



# Problem Definition

## ➡ Problem type

What is your problem (classification / regression / clustering / association rules mining)?

## ➡ Problem description

What is the input and output?



# Data Acquisition

## ➡ Data type

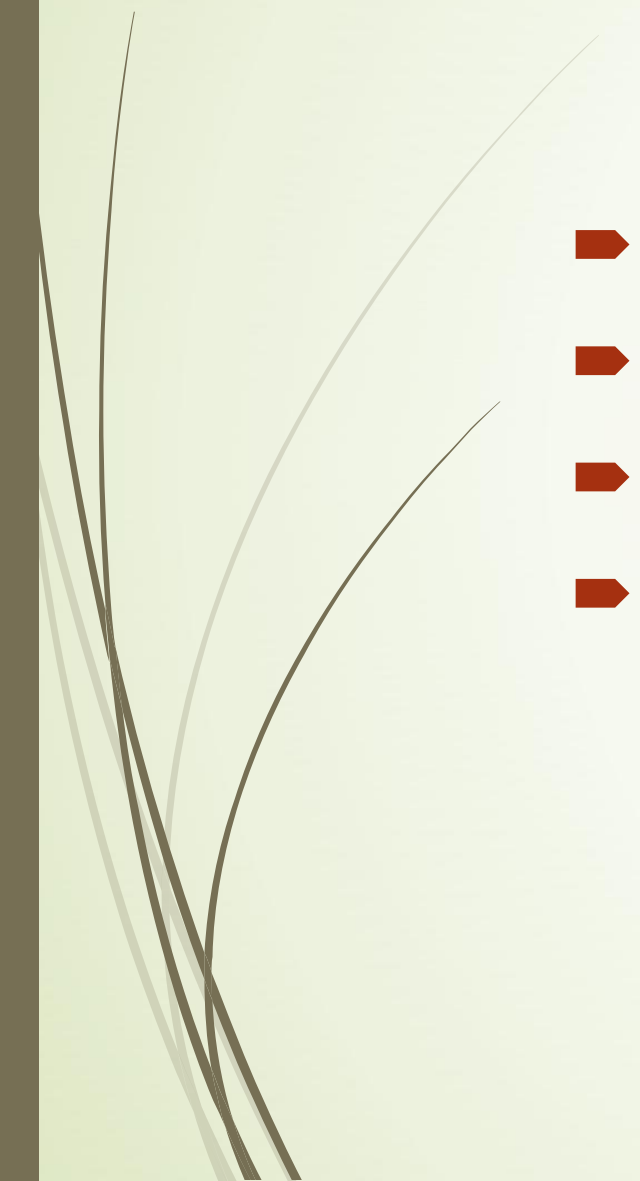
Unlimited. Text/image/video/audio is allowed.

## ➡ Data set

- 1) construct your own data set and state the details of how to collect data (higher score)
- 2) use an existing data set (lower score)



# Feature Extraction

- Which features are extracted from raw data?
  - Any work for data clearing?
  - Any work for dimension reduction?
  - What is the size of the final data set?
- 



# Model Training

- Use at least 2 different machine learning algorithms to train your model and compare their performance
  - training parameters
  - model file or model visualization
  - performance on train/development/test set



# Metric Evaluation

- Which metrics (e.g., accuracy, precision, recall, F1) are used in your project?
- Any work for improving evaluation score?
- Discussion and conclusion