

# Big Data Analytics Algorithms: Exercises 2x10%

## Big Data analytics Recommendation Systems

Assoc. Prof. Dr.techn. Dang Tran Khanh (Đặng Trần Khánh)

[khanh@hufi.edu.vn](mailto:khanh@hufi.edu.vn)

# Outline

- Frameworks
  - Spark MLib, H2O
- Algorithms
  - Clustering
  - Classification
  - Regression
  - Exercise (2x10% in class):
    - Case Study: Classifying Handwritten Digits (cf. 11.5 [1]) → choose “Digit Classification” with H2O or Spark
      - Datasets: <https://www.kaggle.com/competitions/digit-recognizer/data>
    - Case Study: Genome Data Analysis (cf. 11.6 [1])
- Recommendation Systems
  - Presentation: G4.3
- Reading:
  - [1] chapter 11
  - Internet

