# RECSM Summer School: Machine Learning for Social Sciences

## Session 1.1: Course Overview

Reto Wüest

Department of Political Science and International Relations
University of Geneva

# Outline

1. Session 1
2. Session 2
3. Session 3
4. General Information

# Session 1

| No. | Topic |
|---|---|
| 1.1 | Introductions and course overview |
| 1.2 | General introduction to machine learning (prediction and inference, supervised and unsupervised learning) |
| 1.3 | Assessing model accuracy (overfitting, bias-variance trade-off, cross-validation) |
| | Break |
| 1.4 | Shrinkage methods I: ridge regression |
| 1.5 | Shrinkage methods II: the lasso |
| 1.6 | Application of ridge regression and the lasso |

# Session 2

| No. | Topic |
|---|---|
| 2.1 | Introduction to classification and regression trees |
| 2.2 | Advantages and disadvantages of trees |
| 2.3 | Bagging, random forests |
| | Break |
| 2.4 | Boosting |
| 2.5 | Application I: classification and regression trees |
| 2.6 | Application II: bagging, random forests, boosting |

# Session 3

| No. | Topic |
|---|---|
| 3.1 | Introduction to unsupervised learning |
| 3.2 | Principal components analysis (PCA) |
| 3.3 | $K$-means clustering |
| | Break |
| 3.4 | Hierarchical clustering |
| 3.5 | Application I: PCA |
| 3.6 | Application II: clustering methods |

# General Information

- All course materials are available at:
  `http://retowuest.net/recsm-2018/`