

# Statistical Learning

## - Course Syllabus -

Marco Zanotti

2021-2022

### Lecture 0 - Requirements: Tidyverse

#### **Tidyverse:**

- Core Tidyverse
- Wrangle (tidyr & dplyr)
- Program (purrr)

#### **R Code Evaluation Methods:**

- Standard Evaluation
- Non-Standard Evaluation
- Tidy Evaluation

#### **Pipe Operator:**

- Basic Piping
- Argument Placeholder
- Re-using Placeholder for Attributes

### Lecture 1: Tidymodels

#### **Recipe:**

- feature engineering
- steps
- recipes

#### **Parsnip:**

- engines
- modelling
- workflows

#### **Rsample:**

- sampling
- validation
- cross-validation

### Lecture 2: Machine Learning, Regression & Classification

#### **Regression Algorithms:**

- Linear Regression
- Ridge, LASSO, Elastic Net
- MARS

- SVM
- KNN
- Bagging
- Random Forest
- XGBoost, LightGBM, CatBoost
- Cubist
- Multi Layer Perceptron

#### **Classification Algorithms:**

- Naive Bayes
- Logistic Regression
- Ridge, LASSO, Elastic Net
- SVM
- KNN
- CART
- Bagging
- Random Forest
- XGBoost, LightGBM, CatBoost
- Cubist
- Multi Layer Perceptron

### **Lecture 3: Hyperparameter Tuning**

#### **Tune:**

- tuning
- grid searches
- validation

### **Lecture 4: Ensemble Learning & Stacking**

#### **Stacks:**

- ensembling
- stacking
- elastic net stacking
- simple ensembles

### **Lecture 5: Automatic Machine Learning**

#### **H2O:**

- H2O modelling
- AutoML
- Tidymodels integration, h2oparsnip

### **Lecture 6: Deep Neural Networks**

#### **Tensorflow**

#### **Keras:**

- network structure
- activation functions
- backpropagation
- training

- evaluation
- tuning