Université de Montréal

Machine Learning to Analyze SCED Data

Marc J. Lanovaz, Ph.D., BCBA-D
Université de Montréal

April 26th, 2023

Objectives

1. Describe the basic methodology underpinning machine learning research
2. Train simple machine learning models using Python
3. Apply machine learning to single-case datate

What to Expect

- Brief lecture
- Modeling
- Practice, more practice and even more practice
- Remote prompting
- Feedback

Université de Montréal

Artificial Intelligence

"… the study of how to make computers do things at which, at the moment, people are better." (Rich, 1991)

Université de Montréal

Machine Learning

- Subfield of artificial intelligence
- Detect patterns
- Predictions and support decision-making
- Three main types:
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

Dataset

- Samples: Exemplars
- Features: Discriminative stimuli/Salient characteristics
- Class label: Correct response

Dataset – AB Graphs

- 8,000 samples
- Eight features:
  - Mean and standard deviation
  - Intercept and slope
- One class label: Change vs no change

Transforming Data

- Transform continuous feature values to z scores
- (value – mean) / SD

Step 1 – Extract graph data

A	A	A	B	B	B	B	B
7.0	3.2	10.1	1.1	0.2	2.1	2.1	2.1

Expected direction of change: -1 (decrease)
Label: 1 (clear change)

Step 2 – Multiply by -1 (as expected direction of change is to decrease)

A	A	A	B	B	B	B	B
-7.0	-3.2	-10.1	-1.1	-0.2	-2.1	-2.1	-2.1

If the expected direction of change is to increase, skip this step (or multiply by 1).

Step 3 – Normalize the data by subtracting the mean (-3.49) and then dividing the difference by the standard deviation (3.13)

A	A	A	B	B	B	B	B
-1.12	0.09	-2.11	0.76	1.05	0.44	0.44	0.44

Now, the graph data has a mean of 0 and a standard deviation of 1.

Step 4 – Extract eight features from the normalized data

Mean A	Mean B	*SD* - A	*SD* - B	Intercept A	Slope A	Intercept B	Slope B
-1.05	0.63	.90	0.24	-0.55	-0.50	1.25	-0.12

Practice

Manipulating data

Université de Montréal

Algorithms

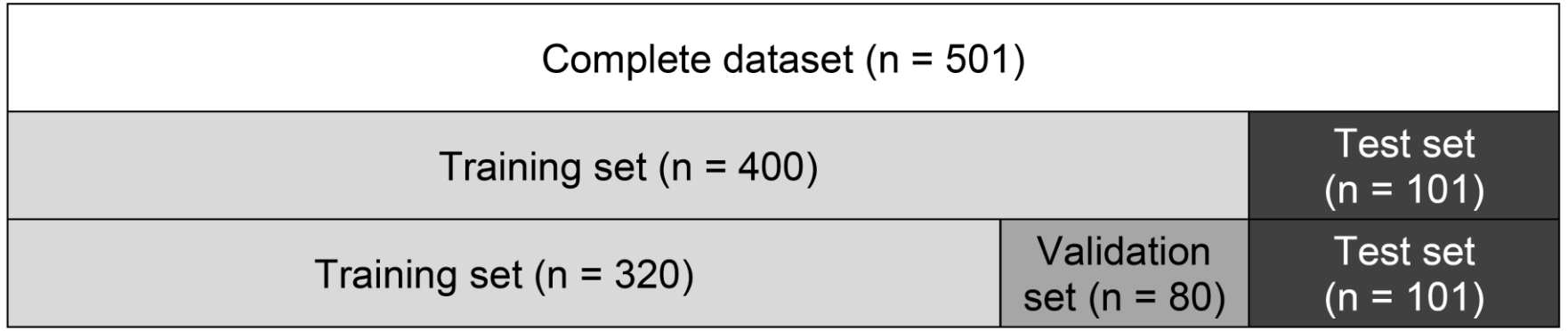
- Algorithm: Teaching method
- Model: Learner
- Predictions: Learner's response
- Hyperparameters: Teaching parameters

Testing Models

- Training set
- Testing set

Holdout Cross-Validation

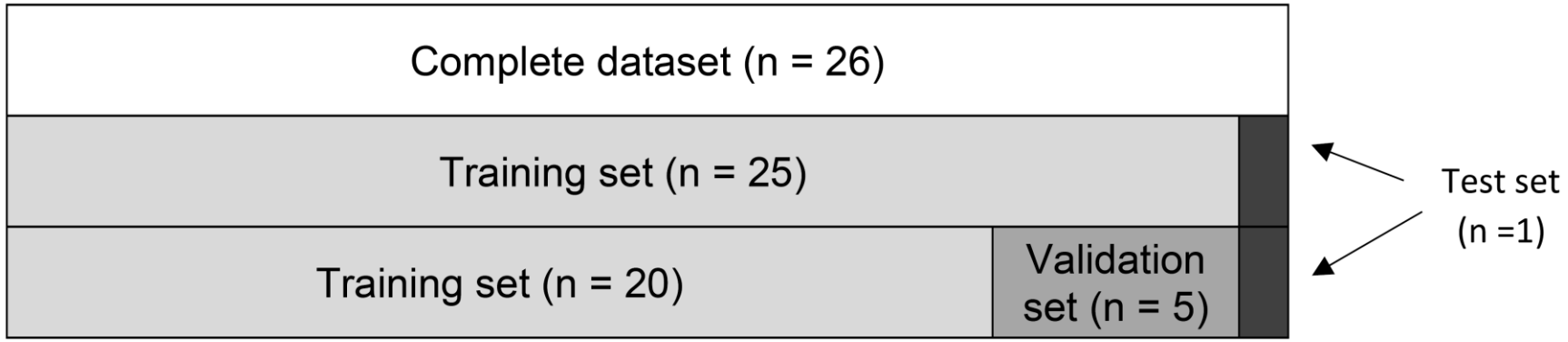
Holdout Cross-Validation



Note. This process runs only once with each set being sampled randomly without replacement.

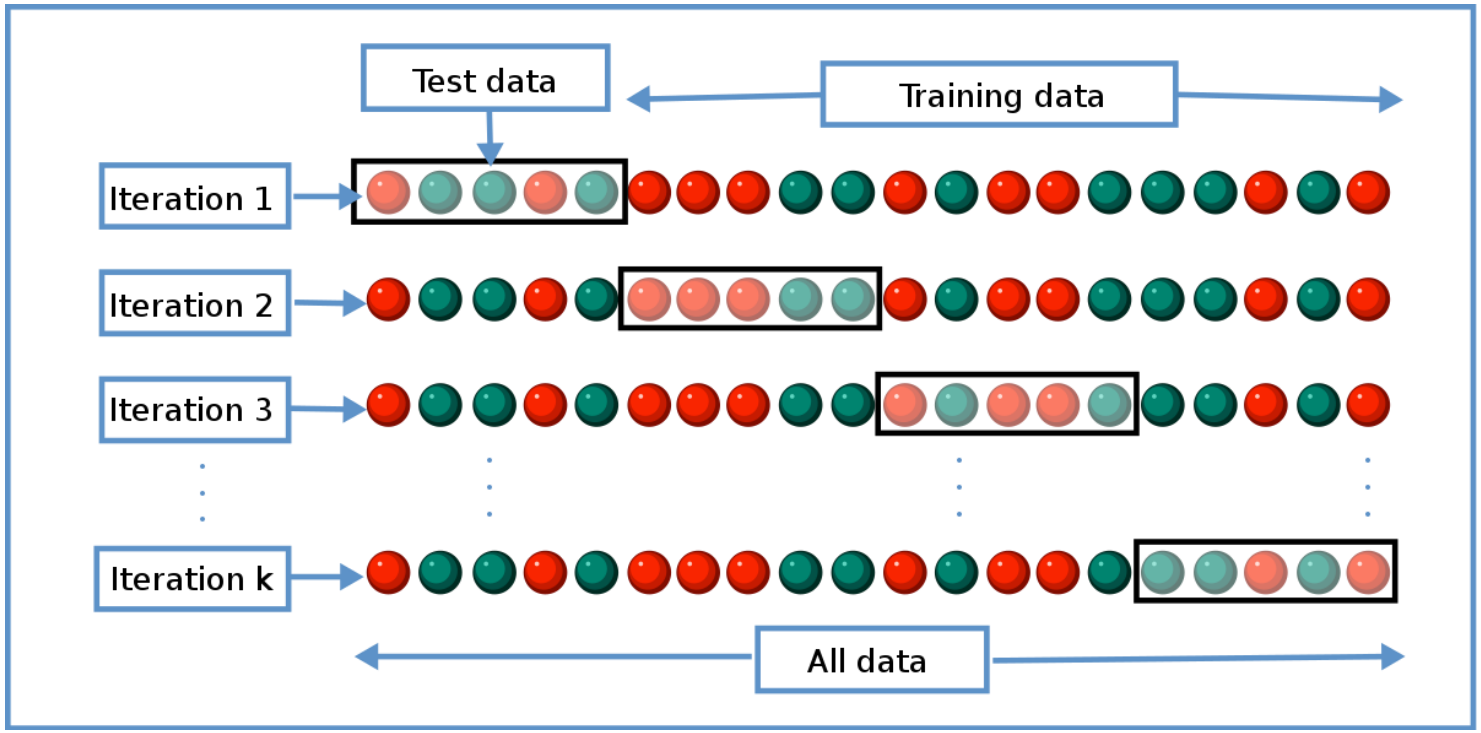
Leave-One Out Cross-Validation

Leave-One Out Cross-Validation



Note. This process runs 26 times so that each sample is in the test set once.

K-Fold Cross-Validation



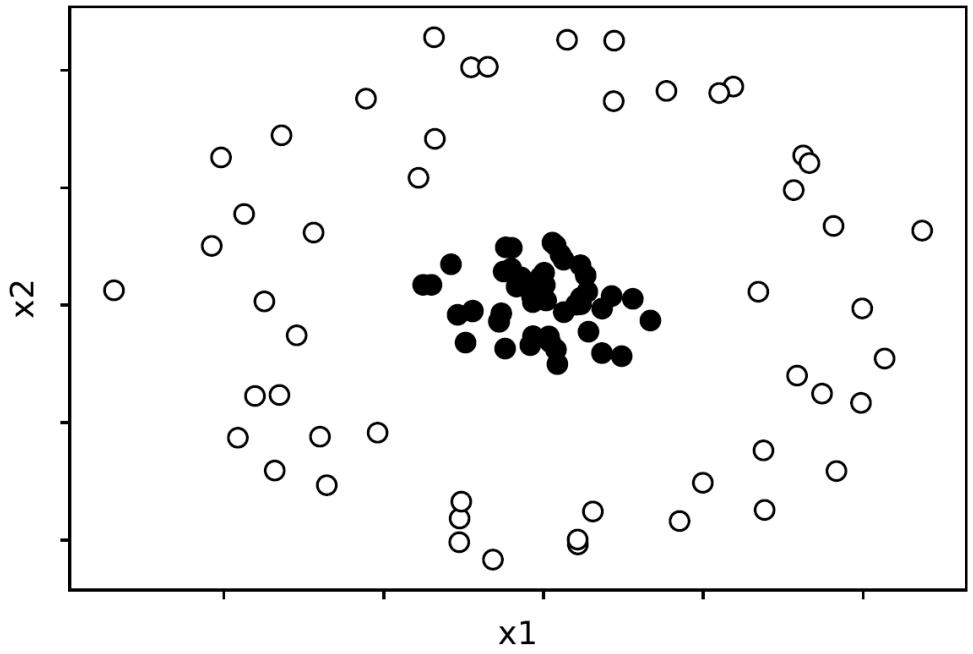
Source: https://en.wikipedia.org/wiki/Cross-validation_(statistics)

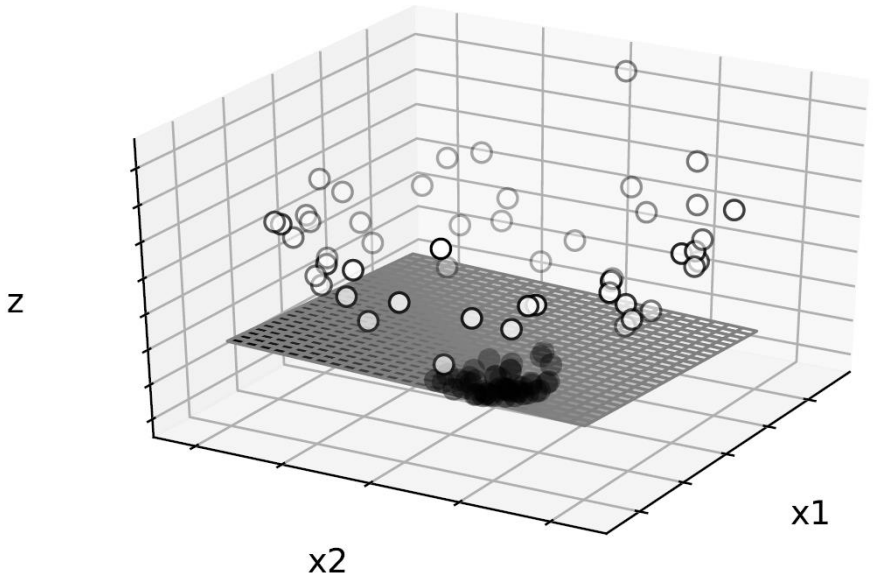
Practice

Splitting the data

Université de Montréal

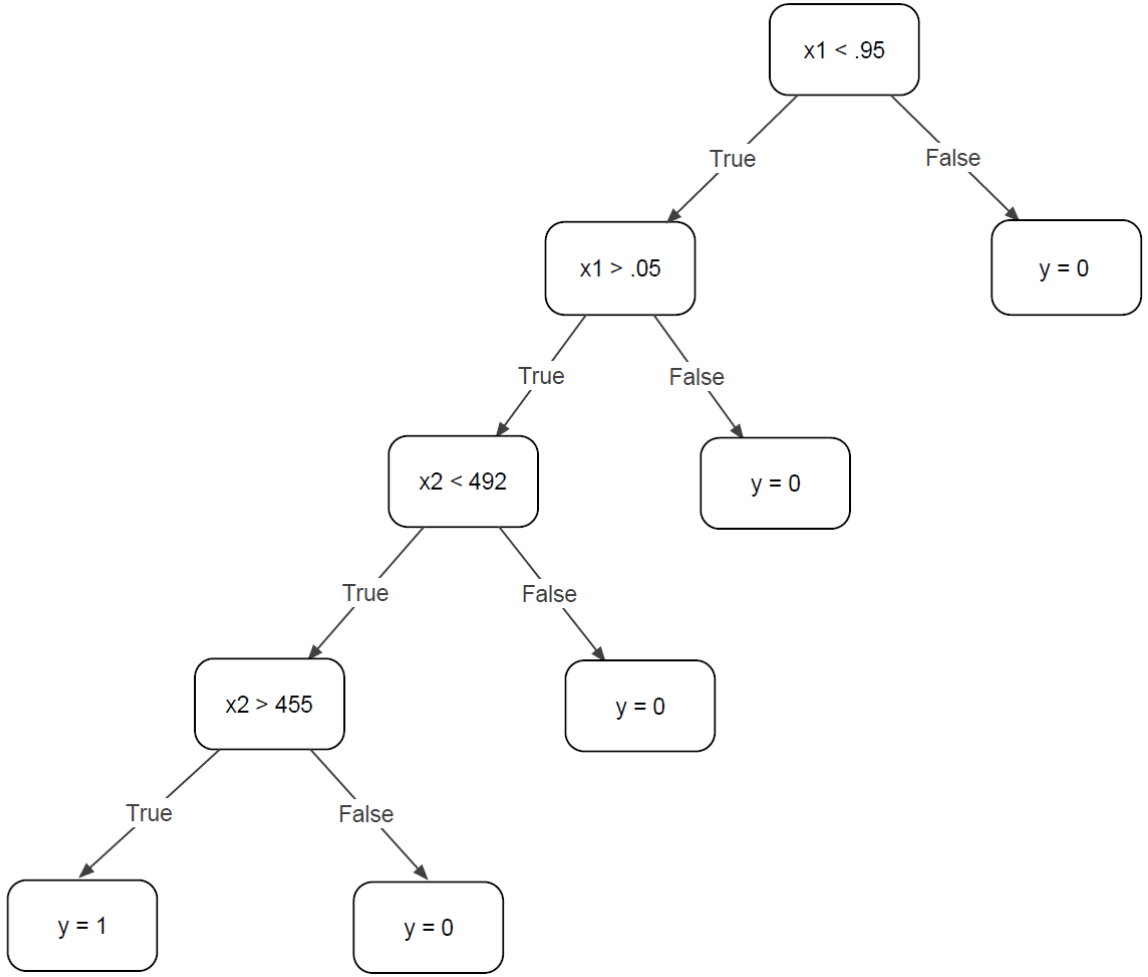
Support Vector Classifier





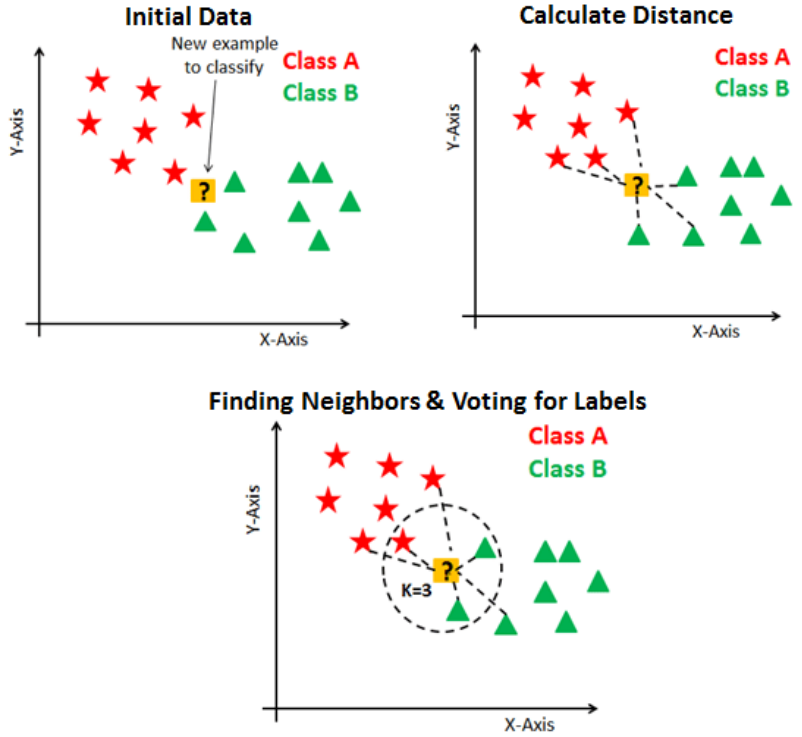
Université de Montréal

Random Forest



Université de Montréal

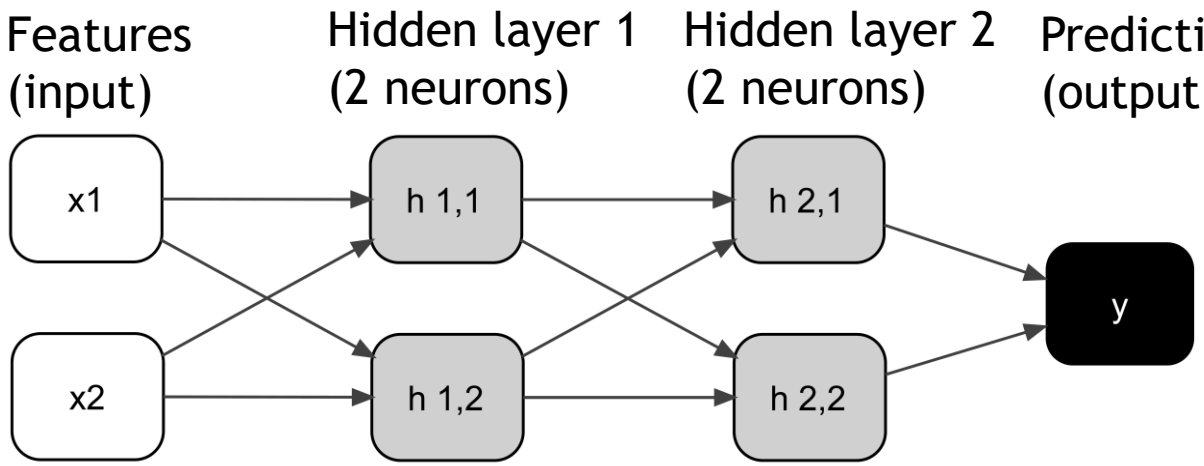
K-NEAREST NEIGHBORS



Source: https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn

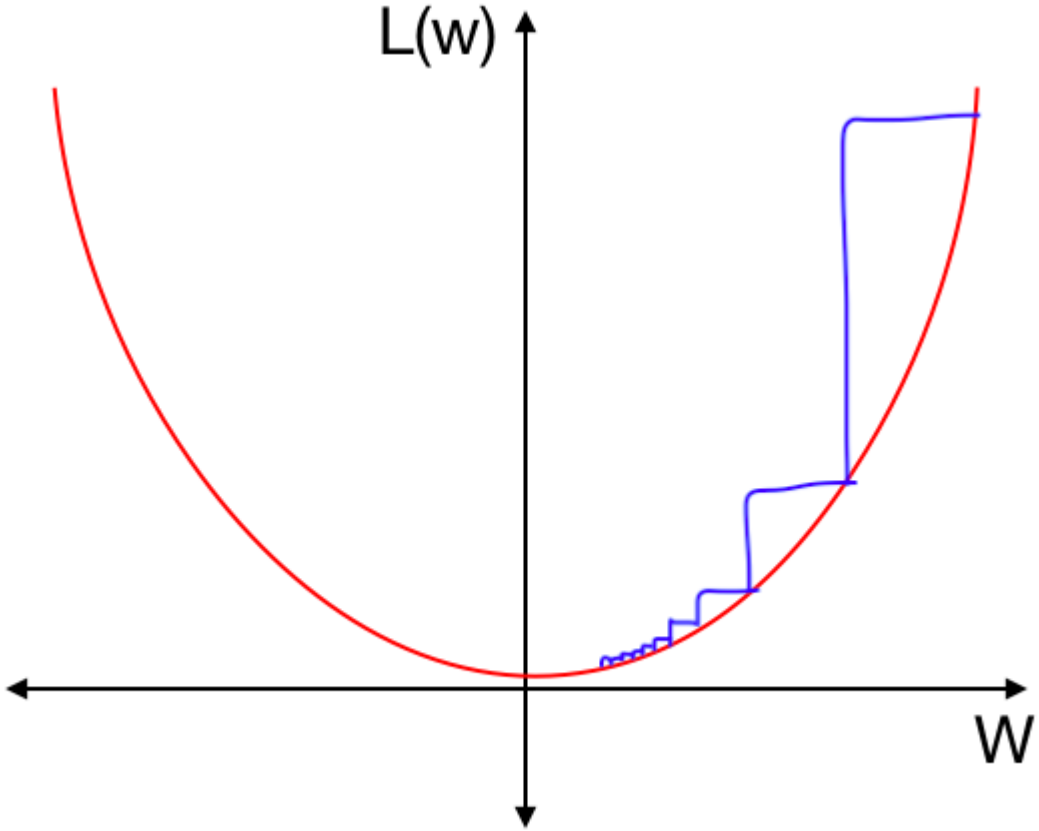
Université de Montréal

Dense Neural Network



Yes or no?

Backpropagation of the Errors



Source: https://towardsdatascience.com/https-medium-com-reina-wang-tw-stochastic-gradient-descent-with-restarts-5f511975163

Practice

Training a model

Université de Montréal

Measures

- Outcomes
  - Accuracy
  - Type I error rate
  - Power
- Comparisons
  - Human observers
  - Best practice

Practice

Computing outcome measures

Université de Montréal

Hyperparameter Tuning

- Number of trees
- Function parameters
- Learning rate
- Epochs
- Number of neighbors
- Loss function

Hyperparameter Tuning

- Epochs
- Number of neurons
- Number of layers
- Activation function
- Loss function
- Dropout

Practice

Hyperparameter tuning

Université de Montréal

Other Considerations

Université de Montréal

Selecting an Algorithm

- Number of samples
- Type of data
- Need to explain
- Stability
- Tuning requirements

Feature Selection

- Filter-based methods
- Wrapper-based methods
- Embedded methods

Benefits

- Reliable decision-making
- Non linear relationships
- Unaffected by common MOs
- Improvements in accuracy

Université de Montréal

Drawbacks

- Black box problem
- “Stupid” decisions
- Number of samples required
- Overtraining and overfitting

Review Objectives

1. Describe the basic methodology underpinning machine learning research
2. Train simple machine learning models using Python
3. Apply machine learning to single-case datate

Université de Montréal

Questions/Comments?

marc.lanovaz@umontreal.ca