Artificial Intelligence: *Modeling Human Intelligence with Networks*

Jeová Farias Sales Rocha Neto

`jeova_farias@brown.edu`

Machine Learning and Linear Classifiers

Machine Learning and Classification

Machine Learning

- Studies how to automatically learn to make accurate predictions based on past observations.
- Two kinds kinds of problems in ML:
    - ▶ Supervised Learning: Observations + desired outputs. Ex.: Spam detection.
    - ▶ Usupervised Learning: Only the observations. Ex.: Clustering.
- Our focus in this course: Supervised Learning

Machine Learning and Classification

Machine Learning

- Studies how to automatically learn to make accurate predictions based on past observations.
- Two kinds kinds of problems in ML:
    - ▶ Supervised Learning: Observations + desired outputs. Ex.: Spam detection.
    - ▶ Usupervised Learning: Only the observations. Ex.: Clustering.
- Our focus in this course: Supervised Learning/Classification.

Classification

- Classify example into a given set of categories.

Machine Learning and Classification

Machine Learning

- Studies how to automatically learn to make accurate predictions based on past observations.
- Two kinds kinds of problems in ML:
  - ▸ Supervised Learning: Observations + desired outputs. Ex.: Spam detection.
  - ▸ Usupervised Learning: Only the observations. Ex.: Clustering.
- Our focus in this course: Supervised Learning/Classification.

Classification

- Classify example into a given set of categories.
- What you have: Labeled examples.

Machine Learning and Classification

Machine Learning

- Studies how to automatically learn to make accurate predictions based on past observations.
- Two kinds kinds of problems in ML:
  - ▶ Supervised Learning: Observations + desired outputs. Ex.: Spam detection.
  - ▶ Usupervised Learning: Only the observations. Ex.: Clustering.
- Our focus in this course: Supervised Learning/Classification.

Classification

- Classify example into a given set of categories.
- What you have: Labeled examples.
- What you want to do: Label (classify) the unlabeled data.

Machine Learning and Classification

Machine Learning

- Studies how to automatically learn to make accurate predictions based on past observations.
- Two kinds kinds of problems in ML:
    - Supervised Learning: Observations + desired outputs. Ex.: Spam detection.
    - Usupervised Learning: Only the observations. Ex.: Clustering.
- Our focus in this course: Supervised Learning/Classification.

Classification

- Classify example into a given set of categories.
- What you have: Labeled examples.
- What you want to do: Label (classify) the unlabeled data.

Classification Example

- You are given a bunch of fish!

Classification Example

- You are given a bunch of fish! And someone tells you:
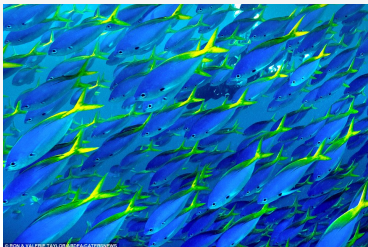
(a) "Look, these ones are *bass*...

Classification Example

- You are given a bunch of fish! And someone tells you:
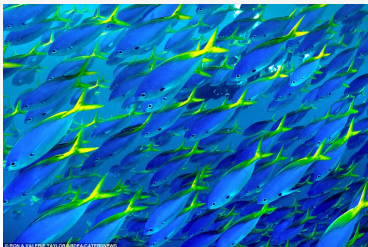
(a) "Look, these ones are *bass*...   *(b)* and these are *tuna*."

- He asks you, "Would you mind tell me which kind of fish Jim is bringing?".
- What would you do?

Classification Example

- Possible steps:
    1. Extract information about the fish you have.

Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info?*

Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info? How about possible info?*

Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info*? *How about possible info?*
  2. Propose a rule based on that information!

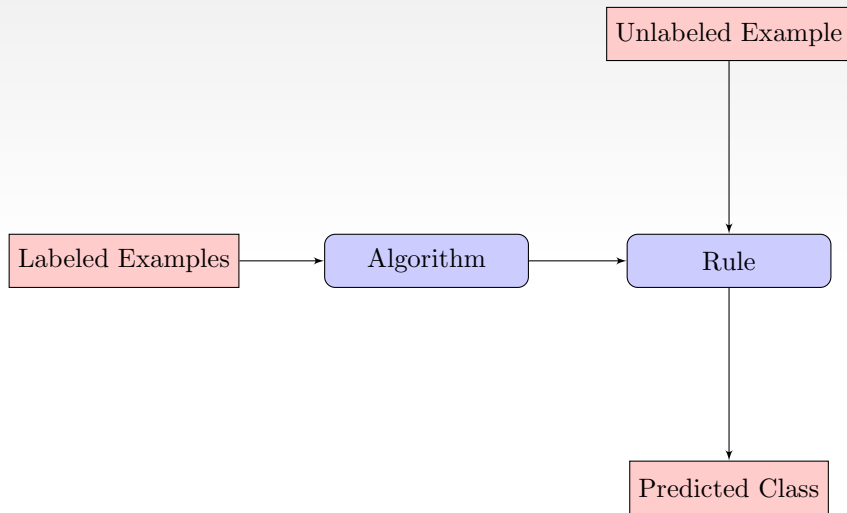# Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info*? *How about possible info?*
  2. Propose a rule based on that information! *Which rule?*

Classification Example

- Possible steps:
    1. Extract information about the fish you have. *What info*? *How about possible info?*
    2. Propose a rule based on that information! *Which rule?*
    3. (Pray so that the you were given everything you needed...)

Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info*? *How about possible info?*
  2. Propose a rule based on that information! *Which rule?*
  3. (Pray so that the you were given everything you needed...)
  4. Classify the new fish!

Classification Example

- Possible steps:
  1. Extract information about the fish you have. *What info*? *How about possible info?*
  2. Propose a rule based on that information! *Which rule?*
  3. (Pray so that the you were given everything you needed...)
  4. Classify the new fish!

- Then, Jim brings you this:
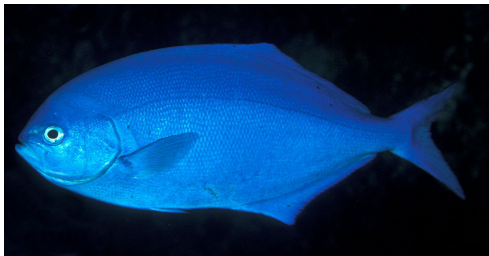
- What do you do?

Classification Pipeline

Unlabeled Example

Labeled Examples → Algorithm → Rule

Predicted Class

Classification Pipeline

Unlabeled Example

Labeled Examples → Algorithm → Rule

How good? ← Predicted Class
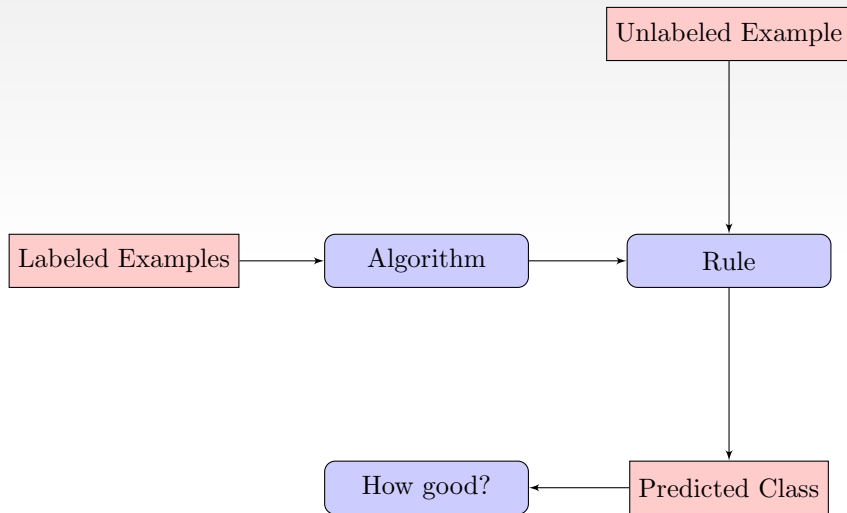
Classification Pipeline
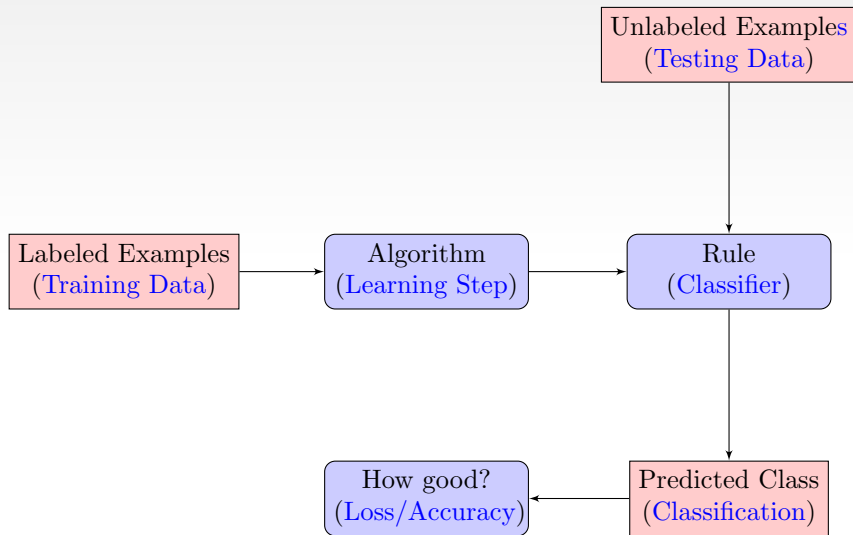
Classification Pipeline
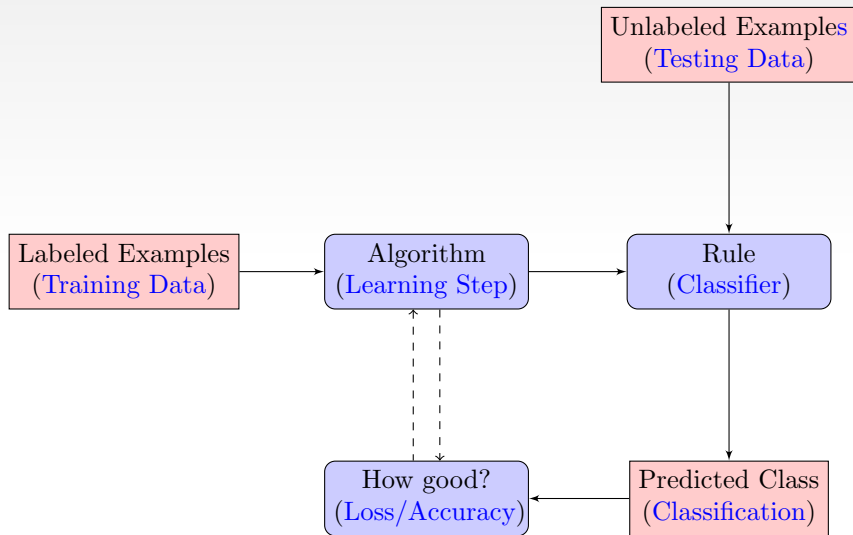
Classification Pipeline

This week for us!

- **Data** *(for testing and for training)* – Points in $\mathbb{R}^2$.
- **Algorithm** – Perceptron (tomorrow!)
- **Classification** – In two classes (tomorrow), more than two (Thursday).
- **Classifier** – Decision boudaries given by lines in $\mathbb{R}^2$.
- **Loss Function** – Zero-one loss.

Linear Classifiers

- Our problem: m points in the plane, i.e., points $p^i = [x_1^i, x_2^i]$.

Linear Classifiers

- Our problem: m points in the plane, i.e., points $p^i = [x_1^i, x_2^i]$.
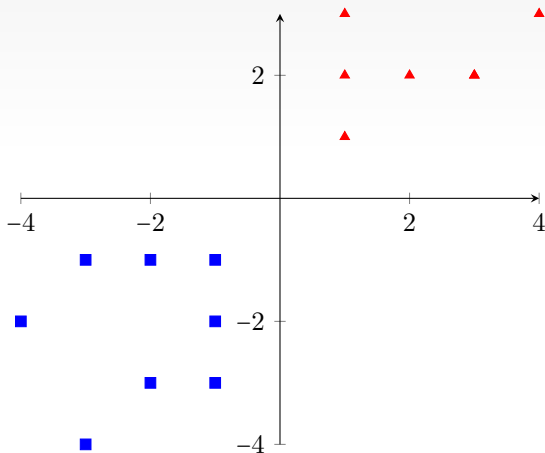
Linear Classifiers

- Now, what is the class of this point? Blue square or red triangle?
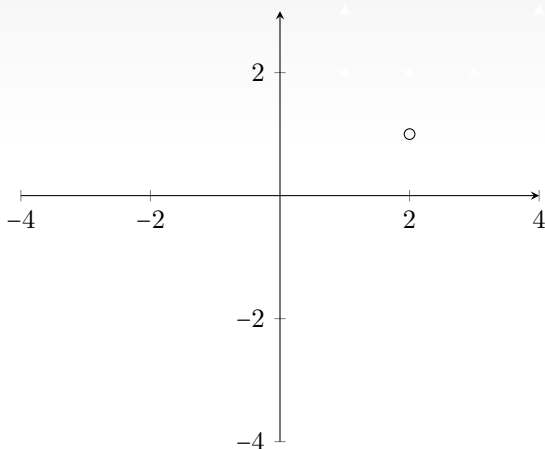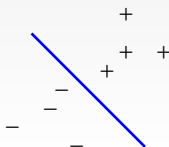
The Neuron Model!
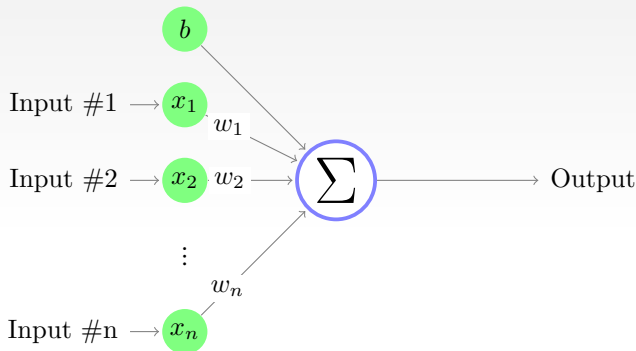
$$x = [x_1, x_2, \ldots, x_n], \quad w = [w_1, w_2, \ldots, w_m]$$

$$f_{act}(x \cdot w + b) = \text{sign}\,(x \cdot w + b) = \begin{cases} +1, & x \cdot w + b > 0 \\ -1, & x \cdot w + b < 0 \end{cases}$$
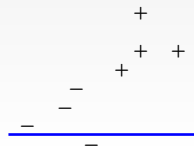
The Neuron Model!
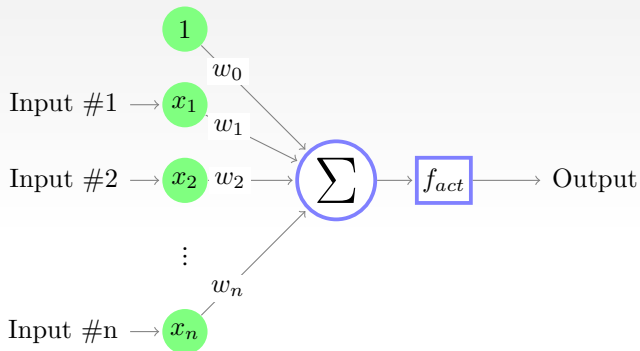
$$x = [1, x_1, x_2, \ldots, x_n], \quad w = [w_0, w_1, w_2, \ldots, w_m]$$

$$f_{act}(x \cdot w) = \text{sign}(x \cdot w) = \begin{cases} +1, & x \cdot w > 0 \\ -1, & x \cdot w < 0 \end{cases}$$

Loss Function

- Which of the following is the best?
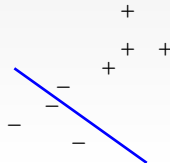
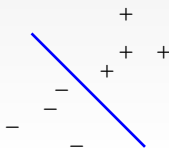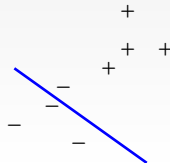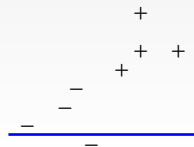(a)   (b)   (c)

- Why?

Loss Function

■ Which of the following is the best?
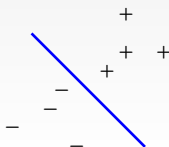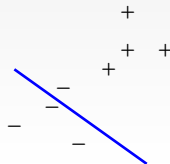

(a)   (b)   (c)

■ Why? Which one is the second best?

Loss Function

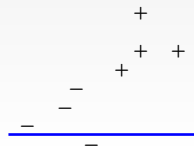- Which of the following is the best?

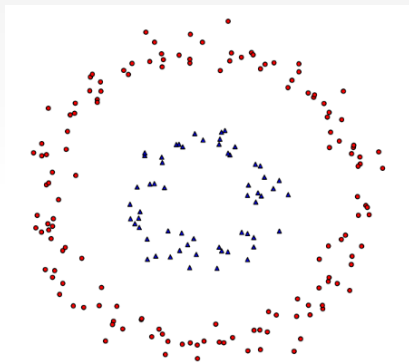*(a)*   *(b)*   *(c)*

- Why? Which one is the second best?
- Zero-one loss: # Number of misclassified examples.

$$\mathcal{L}_w(X) = \sum_{i=1}^{n} [f_w(x^i) \neq y_i]$$

$$[\cdot] : \{\text{False}, \text{True}\} \rightarrow \{0, 1\}$$

Non-Linearly separable data

- What is decision boundary of the this:

- Could I use straight lines to separate it??
- What is Non-Linearly separable dataset?

Python Notebooks!