9TH EDITION
VISUM 2021
VISION UNDERSTANDING & MACHINE INTELLIGENCE SUMMER SCHOOL

Machine Learning (Part 1)

(**Basic Concepts** + Basic Coding)

Wilson Silva and Tiago Gonçalves

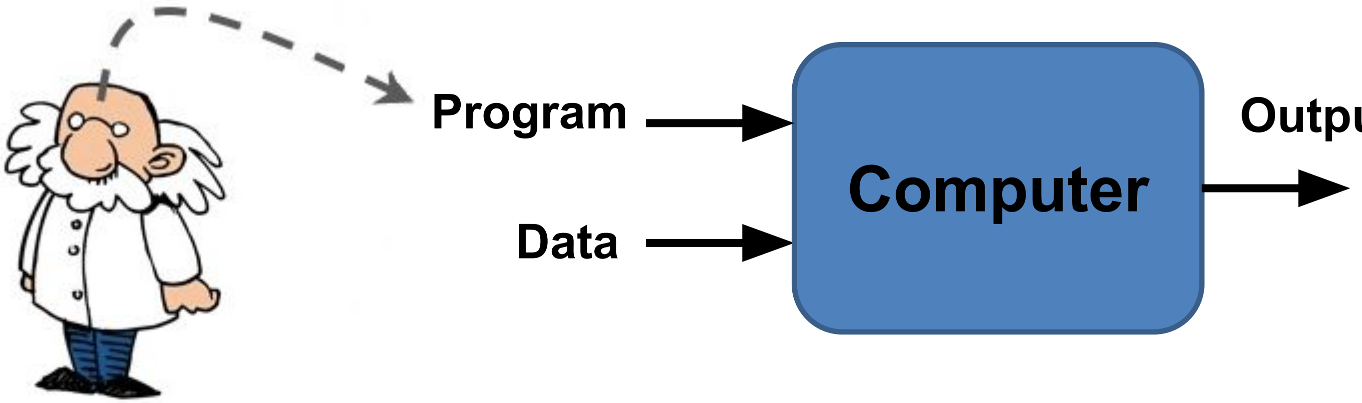
DIGITAL EDITION
JULY 02 - 09

What will we cover?

- What is machine learning?
- What kind of learning settings exist?
- What type of supervised learning problems can you find?
- What type of classifiers are available for you to use?
- How does the learning process work?
- Python basics + Sklearn (the most used toolbox in the realm of machine learning)

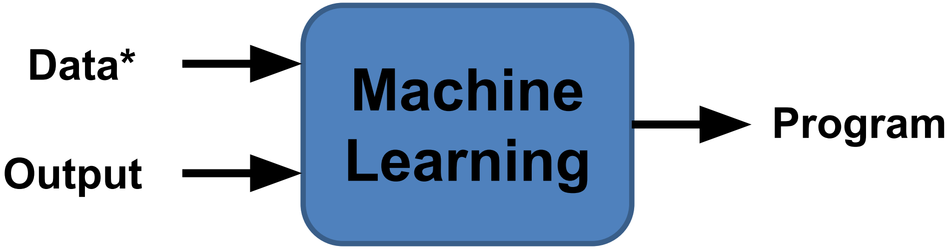
AI (expert based)

- Emulates the decision-making process of a human expert
  - “If-then-else” approach
  - What are the advantages and disadvantages of this approach?

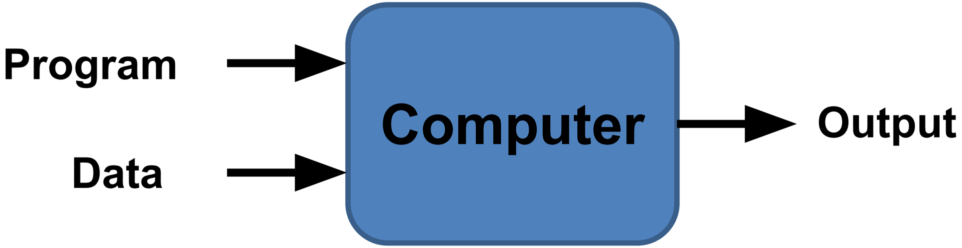


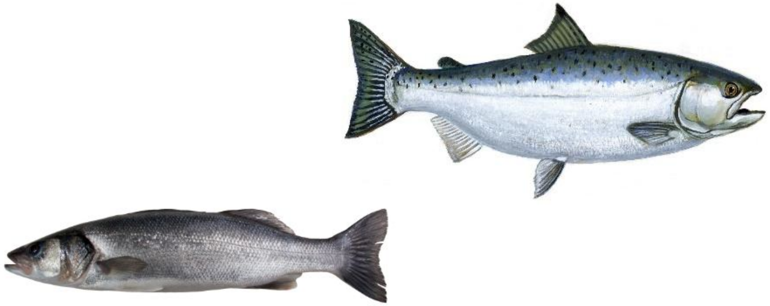
AI (machine learning based)

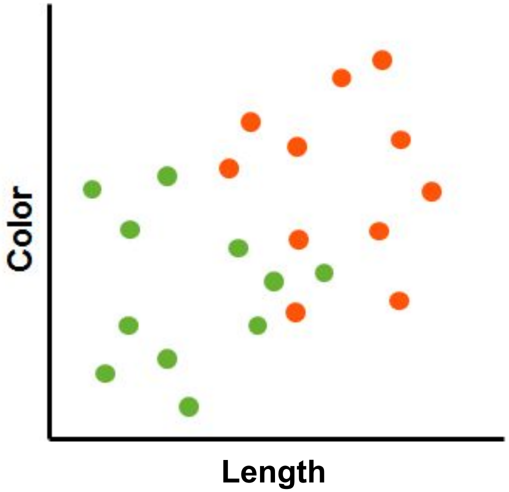
Step 1:



Step 2:

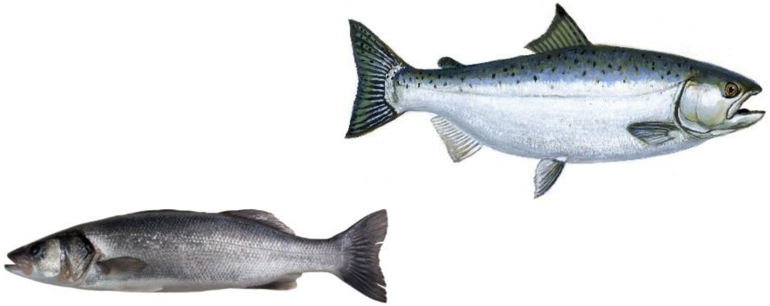




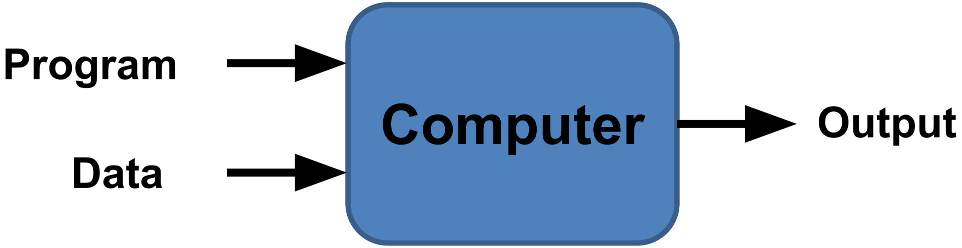


AI (deep learning based)

Step 1:



Step 2:

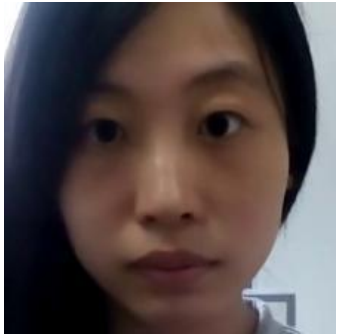


Taxonomy of Learning Settings

- <u>Supervised Learning</u>

  - In supervised learning, the training data you feed to the algorithm includes the desired solutions, called labels

Input



Label:
Real (Bona-fide)

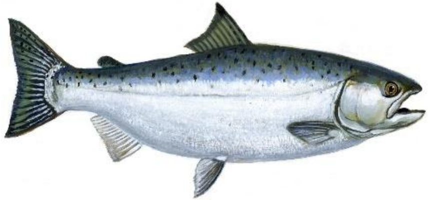
Input

Label:
Fake (Attack)

Taxonomy of Learning Settings

- <u>Unsupervised Learning</u>

  - In unsupervised learning, as you might guess, the training data is unlabelled. The system tries to learn without a teacher

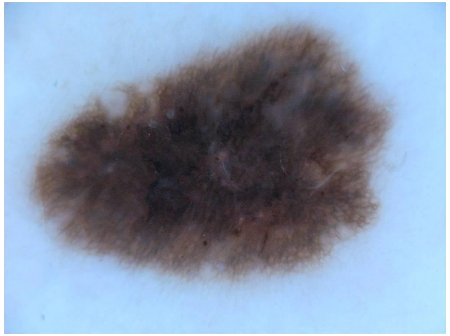


Taxonomy of Learning Settings

- Semi-supervised Learning

  - Some algorithms can deal with partially labelled training data, usually a lot of unlabelled data and a little bit of labelled data
  - Most semi-supervised learning algorithms are combinations of unsupervised and supervised algorithms

Input 1



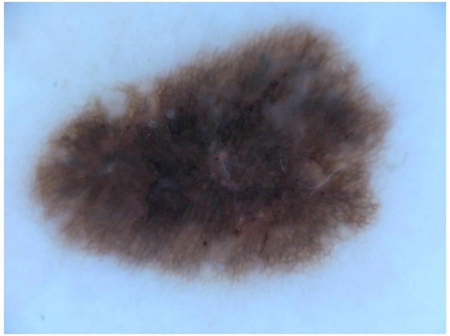
Output 1

Input 2

Taxonomy of Learning Settings

- <u>Weakly-supervised Learning</u>

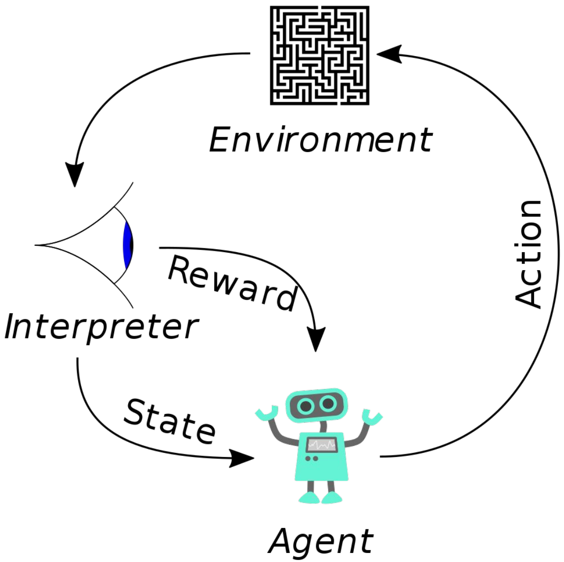
  - When we have noisy, limited or imprecise sources that are used to provide supervision

Input	**Output (we wanted)**	**Output (we have)**
		Class: Benign

Taxonomy of Learning Settings

- <u>Reinforcement Learning</u>

  - Reinforcement learning is a different beast

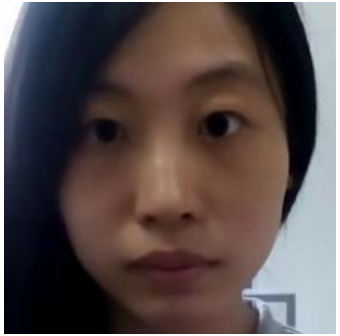


Supervised Learning

- <u>Classification (binary)</u>

  - System capable of distinguishing between just two classes

Input



Label:
Real (Bona-fide)

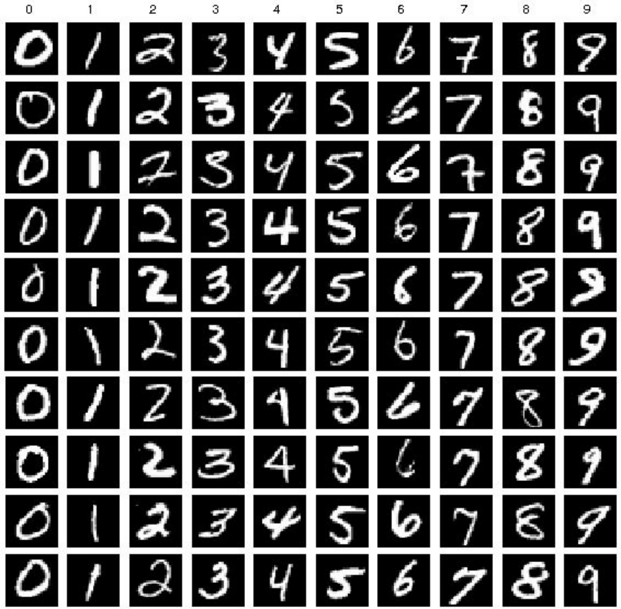
Input

Label:
Fake (Attack)

Supervised Learning

- <u>Classification (multi-class)</u>

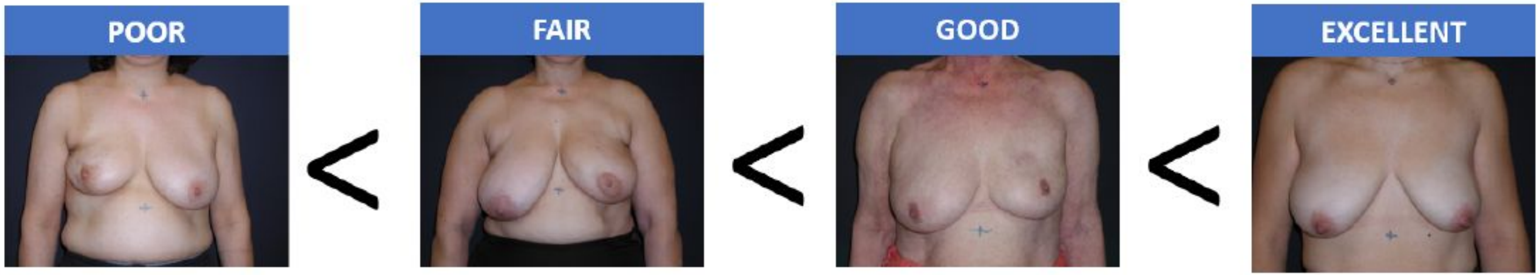
  - System capable of distinguishing between more than two classes



Supervised Learning

- <u>Classification (ordinal)</u>

  - System capable of distinguishing between more than two classes
  - Classes have an inherent natural order



Supervised Learning

- <u>Regression</u>

  - Predicting house price
    - Output: price (a scalar)
    - Inputs: size, orientation, distance to key services, etc.

Classification (Learning tools)

$$y = f(\mathbf{x})$$

- Training: given a training set of labelled examples {(x1, y1),...,(xn, yn)}, estimate the prediction function f by minimizing the prediction error on the training set

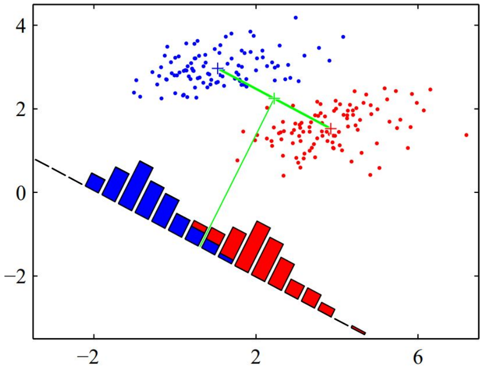
- Testing: apply f to a never seen before test example x and output the predicted value y = f(x)

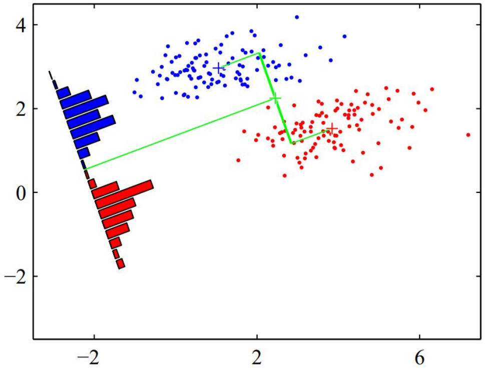
Classification (Learning tools)

- <u>Discriminant function</u>

- No computation of posterior probabilities (probability of a certain class given the data).
- Directly map each x onto a class label

Tools: Fisher’s Linear Discriminant, SVM, etc.





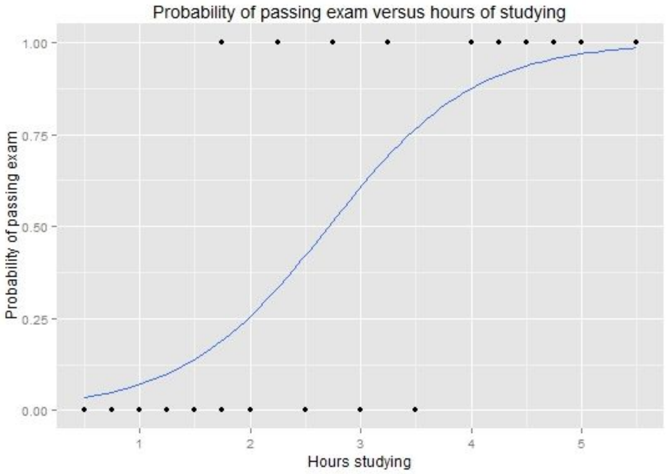
Classification (Learning tools)

- <u>Probabilistic Discriminative Models</u>

- Computation of posterior probabilities: p(C = k | x)
- Model posterior probabilities directly

Tools: Logistic Regression

$$p = \frac{1}{1 + e^{-\omega_0 + \omega_1 x}}$$



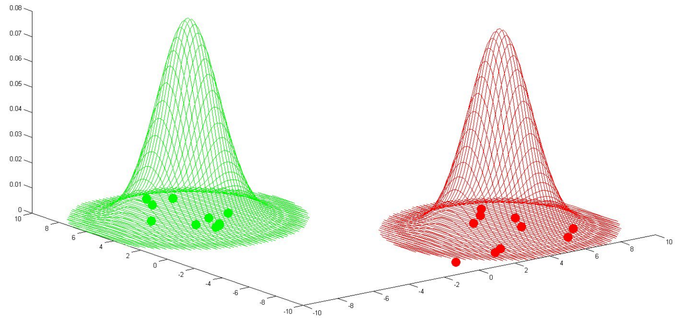
Classification (Learning tools)

- <u>Probabilistic Generative Models</u>

- Model class priors, p(C = k), and class-conditional densities, p(x | C = k)
- Use Bayes theorem to compute posterior probability, p(C = k | x)

Tools: Bayes

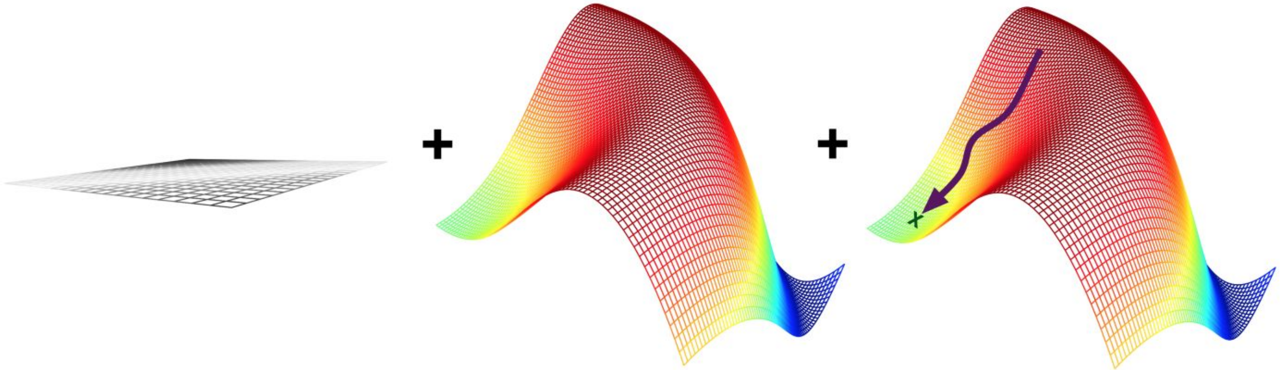
$$p(C = k|x) = \frac{p(x \mid C = k)\, p(C = k)}{p(x)}$$



Common steps in all approaches

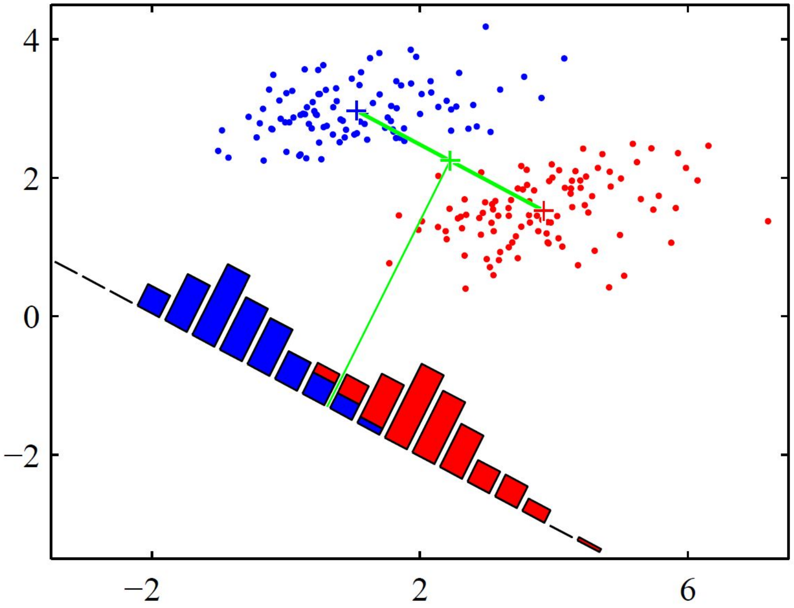
- <u>The learning of a model from the data entails:</u>

- Model representation
- Evaluation
- Optimization



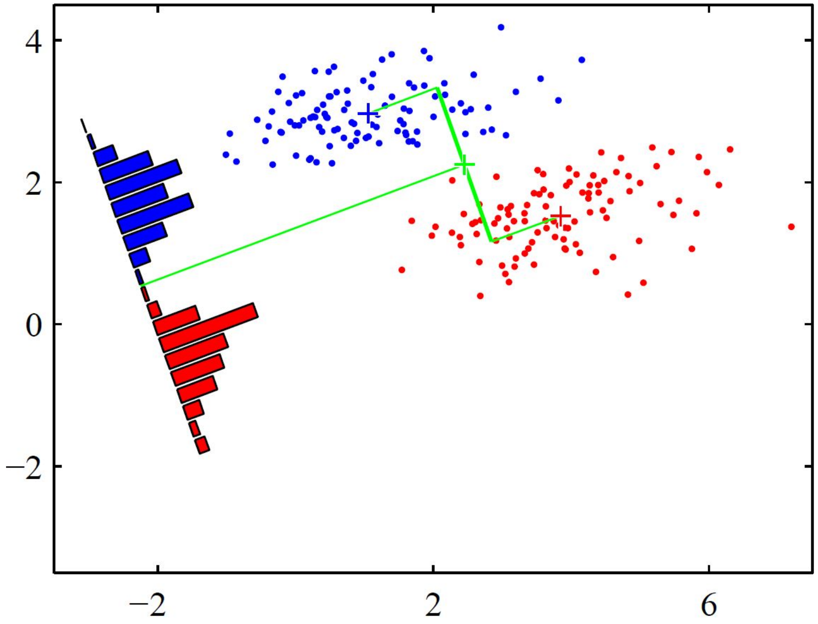
Design of a classifier (example)

- Use the (hyper-) plane orthogonal to the line joining the means
  - Project the data in the direction given by the line joining the class means



Design of a classifier (example)

- Project the data in the direction that maximizes the ratio between class variance to within class variance (Fisher)



Design of a classifier (example)

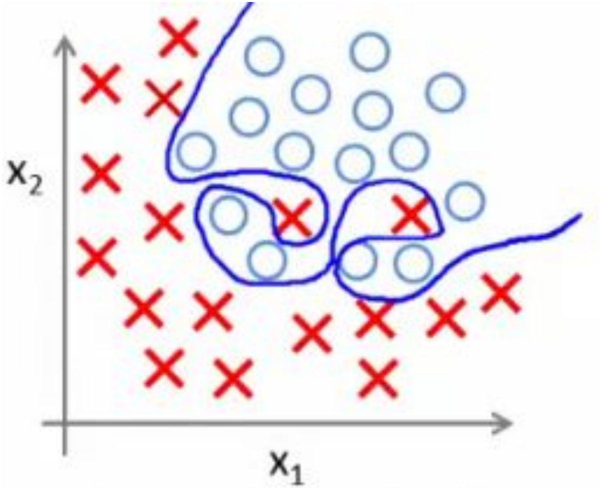
- Fisher's linear discriminant
  - Model representation: class of linear models
  - Evaluation: find the direction **w** that maximizes J(**w**)

$$J(\omega) = \frac{(m_2 - m_1)^2}{{s_1}^2 + {s_2}^2} \qquad J(\omega) = \frac{\omega^T S_B \omega}{\omega^T S_W \omega}$$

  - Optimization:

$$\omega^* \infty {S_W}^{-1} (m_2 - m_1)$$

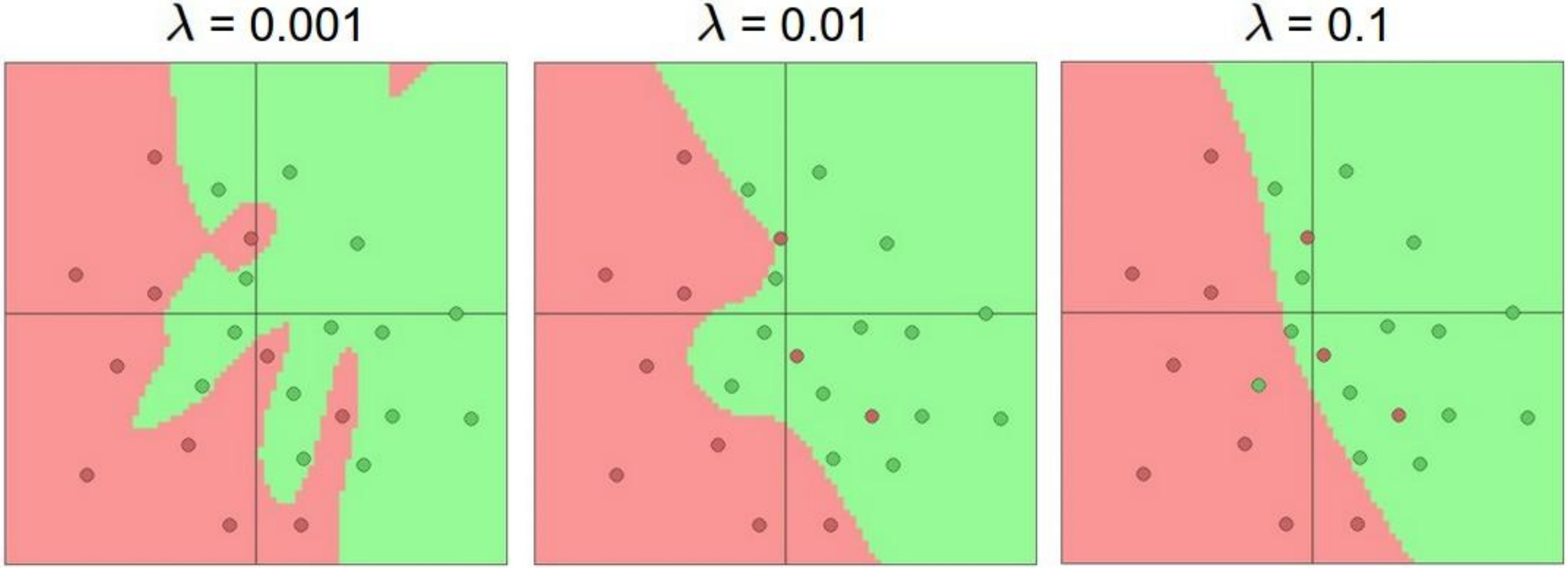
Avoid overfitting and data memorization



$$g(\theta_0 + \theta_1 x_1 + \theta_2 x_1^2 \\ +\theta_3 x_1^2 x_2 + \theta_4 x_1^2 x_2^2 \\ +\theta_5 x_1^2 x_2^3 + \theta_6 x_1^3 x_2 + \dots$$

Avoid overfitting and data memorization

- <u>Evaluation:</u>
  - Minimize (error in data) + λ (model complexity)



9TH EDITION

VISUM 2021

VISION UNDERSTANDING & MACHINE INTELLIGENCE SUMMER SCHOOL

Machine Learning (Part 2)

(Basic Concepts + **Basic Coding**)

https://colab.research.google.com/drive/1ZXYpsBx6y74LULIdLLgwcFNdHRkoixtf?usp=sharing

Wilson Silva and Tiago Gonçalves

DIGITAL EDITION

JULY 02 - 09

References

- Jaime S. Cardoso. Machine Learning Classes. Doctoral Program in Electrical and Computer Engineering.
- Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- Pedro Domingos, “A few useful things to know about machine learning”, CACM, 2012
- Medium post about Pedro Domingos’ paper: https://medium.com/@devnag/machine-learning-representation-evaluation-optimization-fc7b26b38fdb
- H. Li, W. Li, H. Cao, S. Wang, F. Huang, and A. C. Kot, “Unsupervised domain adaptation for face anti-spoofing,” IEEE Transactions on Information Forensics and Security, vol. 13, no. 7, pp. 1794–1809, 2018. (Face PAD dataset)
- Codella N, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza S, Kalloo A, Liopyris K, Mishra N, Kittler H, Halpern A. "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)". arXiv: 1710.05006 [cs.CV] Available: https://arxiv.org/abs/1710.05006 (Skin lesion images)
- Cardoso, J.S.; Cardoso, M.J. Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. Artif. Intell. Med. 2007, 40, 115–126. doi:10.1016/j.artmed.2007.02.007. (Aesthetic evaluation)