DE VINCI INNOVATION CENTER

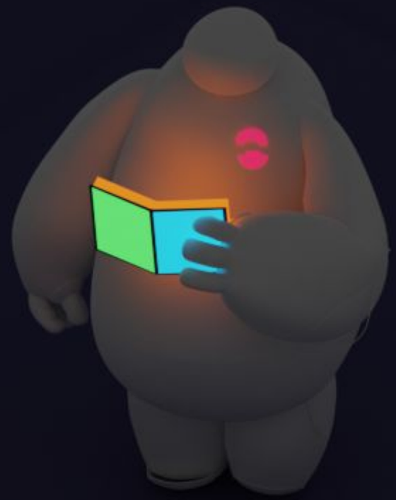
MACHINE LEARNING

Learning from Data

Ylless **HATI**

PHD Student - Computer Science

yliess.hati@**devinci.fr**



LÉONARD DE VINCI PARIS-LA DÉFENSE

ESILV LÉONARD DE VINCI | ÉCOLE D'INGÉNIEURS PARIS-LA DÉFENSE

iiM LÉONARD DE VINCI | INSTITUT DE L'INTERNET ET DU MULTIMEDIA PARIS-LA DÉFENSE

EMLV LÉONARD DE VINCI | ÉCOLE DE MANAGEMENT PARIS-LA DÉFENSE

|00 INTRODUCTION

Pattern Recognition

Study **Correlations** Between **Data Domains**

Tune a **Parametric Model** with **Data**

|00 INTRODUCTION

SUPERVISED

Learn parameters from
Labeled Data

Regression

Classification

UNSUPERVISED

Learn parameters from
Unlabeled Data

Clustering

Latent Space

REINFORCEMENT

Autonomous Agent Learning
from **Experience**

Q-Learning

DE VINCI
INNOVATION
CENTER

01|

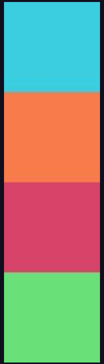
LINEAR ALGEBRA

The Last Space Bender

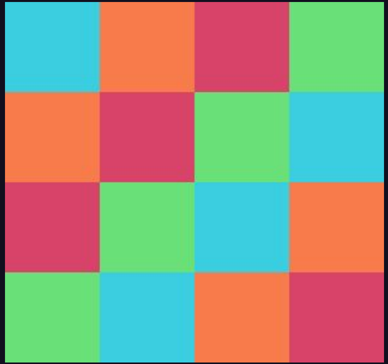
DE VINCI INNOVATION CENTER

|01 LINEAR ALGEBRA

VECTOR



MATRIX



TENSOR

DE VINCI
INNOVATION
CENTER

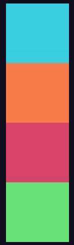
VECTOR

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

LENGTH

$$\|\vec{v}\| = \sqrt{v_x^2 + v_y^2}$$

VECTOR



ADDITION / SUBTRACTION

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

$$\vec{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

$$\vec{v} + \vec{u} = \begin{pmatrix} v_x + u_x \\ v_y + u_y \end{pmatrix}$$

|01 LINEAR ALGEBRA

VECTOR

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

$$\vec{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

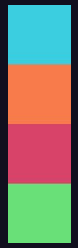
SCALAR (DOT) PRODUCT

$$\vec{v}.\vec{u} = v_x u_x + v_y u_y = |v|\,|u|\,cos\theta$$

DE VINCI INNOVATION CENTER

|01 LINEAR ALGEBRA

VECTOR



$$\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$$

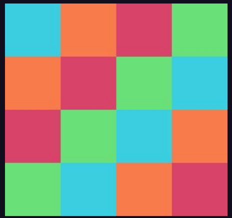
$$\vec{u} = \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}$$

CROSS PRODUCT

$$\vec{v} \times \vec{u} = \begin{pmatrix} v_y u_z - v_z u_y \\ v_z u_x - v_x u_z \\ v_x u_y - v_y u_x \end{pmatrix}$$

DE VINCI INNOVATION CENTER

MATRIX



ADDITION / SUBTRACTION

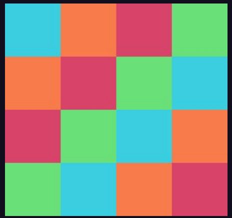
$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$$

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} \end{pmatrix}$$

|01 LINEAR ALGEBRA

MATRIX



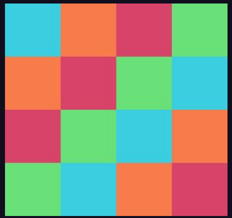
HADAMARD PRODUCT

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$$

$$A \odot B = \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} \end{pmatrix}$$

MATRIX



MATRIX MULTIPLICATION

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$$

$$A \cdot B = \begin{pmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} & a_{1,1}b_{1,2} + a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} & a_{2,1}b_{1,2} + a_{2,2}b_{2,2} \end{pmatrix}$$

|01 LINEAR ALGEBRA

MATRIX

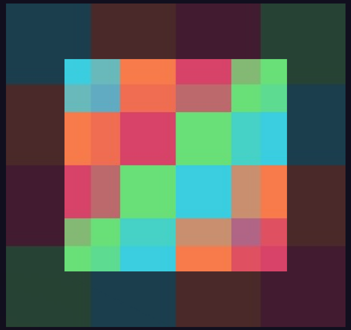
EYE

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

DE VINCI
INNOVATION
CENTER

|01 LINEAR ALGEBRA

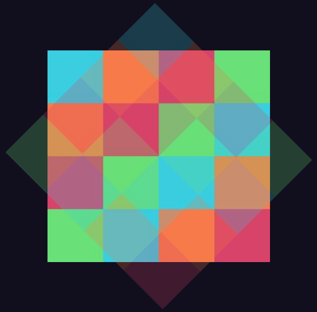
MATRIX



SCALE

$$\begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

MATRIX



ROTATION

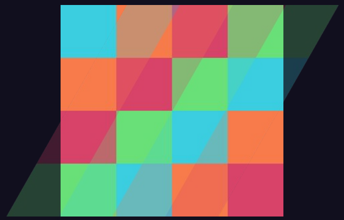
$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ cos(\theta) & -sin(\theta) & 0 & 0 \\ sin(\theta) & cos(\theta) & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} cos(\theta) & 0 & sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ -sin(\theta) & 0 & cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_z(\theta) = \begin{pmatrix} cos(\theta) & -sin(\theta) & 0 & 0 \\ sin(\theta) & cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

DE VINCI INNOVATION CENTER

MATRIX



SHEAR

$$\begin{pmatrix} 1 & 0 & \lambda & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

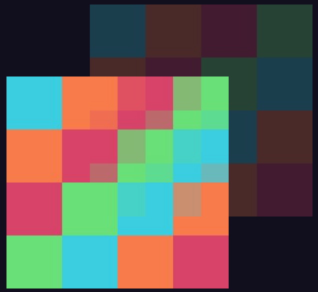
MATRIX

REFLECTION

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

|01 LINEAR ALGEBRA

MATRIX



TRANSLATION

$$\begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

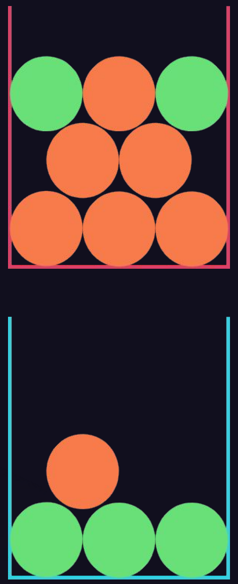
DE VINCI INNOVATION CENTER

|02

PROBABILITIES

Roll the Dice

DE VINCI INNOVATION CENTER



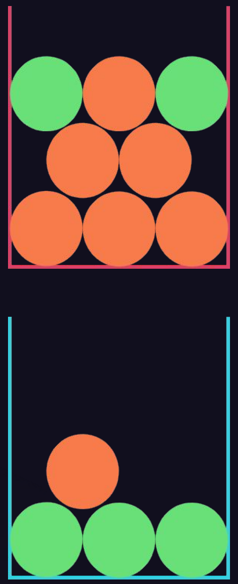
RANDOM VARIABLES

$$B = \begin{cases} r & \text{if Box is red} \\ b & \text{if Box is blue} \end{cases}$$

$$F = \begin{cases} a & \text{if Fruit is an apple} \\ o & \text{if Fruit is an orange} \end{cases}$$

DE VINCI INNOVATION CENTER

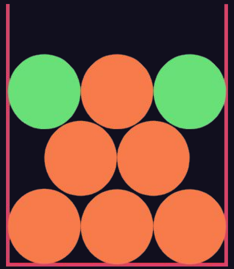
PROBABILITIES

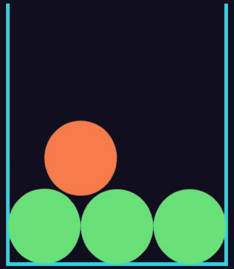


$$p(B = r) = \frac{\#\text{ red boxes picked}}{\#\text{ total boxes picked}} \in [0; 1]$$

$$p(B = r) + p(B = b) = 1$$

JOINT & CONDITIONAL





Sum Rule

Joint probabilties $p(B = r, F = a)$

Marginal probabilty $p(B = r) = p(B = r, F = a) + p(B = r, F = o)$

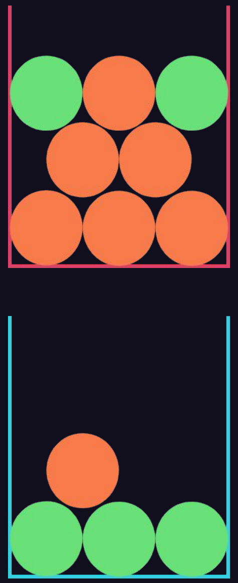
Product Rule

Conditional probabilties $p(B = r | F = o)$

$$p(B = r, F = o) = p(F = o | B = r)\; p(B = r)$$

|02 PROBABILITIES

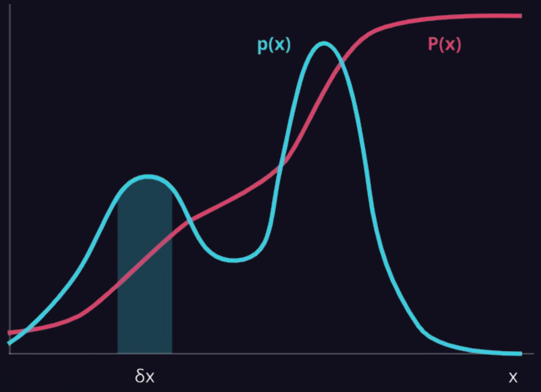
BAYES THEOREM



$$p(B=r|F=o) = \frac{p(F{=}o|B{=}r)p(B{=}r)}{p(F{=}o)}$$

$$p(B=r|F=o) = \frac{p(F{=}o|B{=}r)p(B{=}r)}{p(F{=}o|B{=}r)p(B{=}r)+p(F{=}o|B{=}b)p(B{=}b)}$$

|02 PROBABILITIES

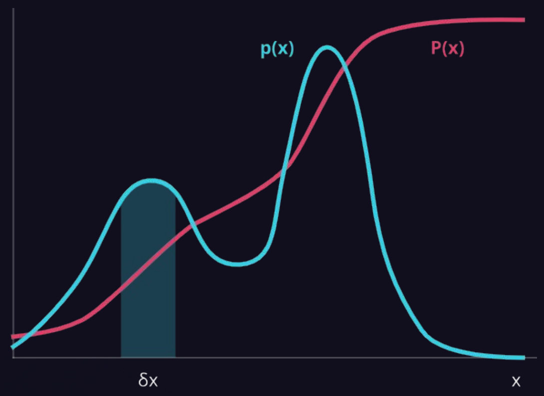


DENSITY

$$p(x \in (a,b)) = \int_a^b p(x)dx$$

$$\int_{-\infty}^{+\infty} p(x)dx = 1$$

$$p(x) \geq 0$$



EXPECTATION & COVARIANCE

$$\mathbb{E}[f] = \int p(x)f(x)dx$$

$$var[f] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2]$$

$$var[f] = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$$

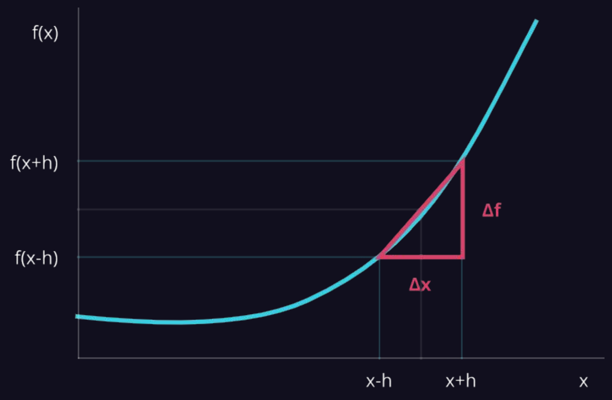
$$cov[x, y] = \mathbb{E}_{x,y}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

03|

OPTIMIZATION

One Method to Rule them All

DE VINCI INNOVATION CENTER



DERIVATIVES

$$f'(x) = \lim_{h \to 0} \frac{\Delta f}{\Delta x} = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h}$$

First order Derivative

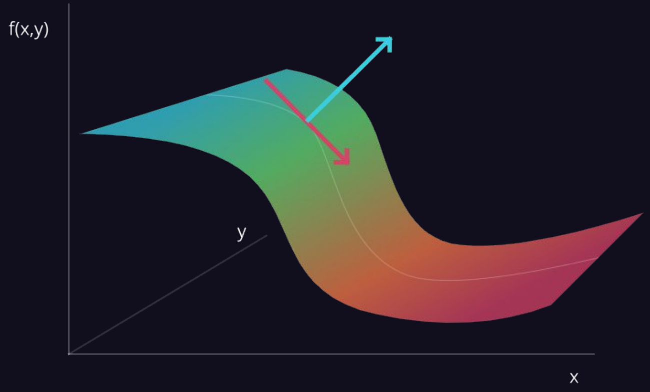
Direction of the Slope

Second order Derivative

Rate of Changes in the Slope

DE VINCI INNOVATION CENTER

|03 OPTIMIZATION

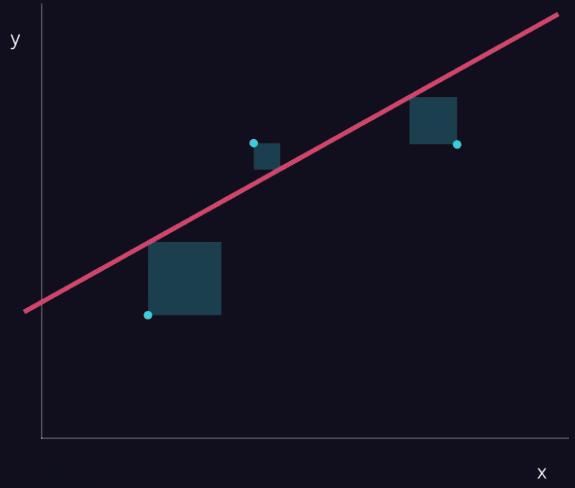


GRADIENTS

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix}$$

DE VINCI INNOVATION CENTER

|03 OPTIMIZATION



LEAST SQUARES

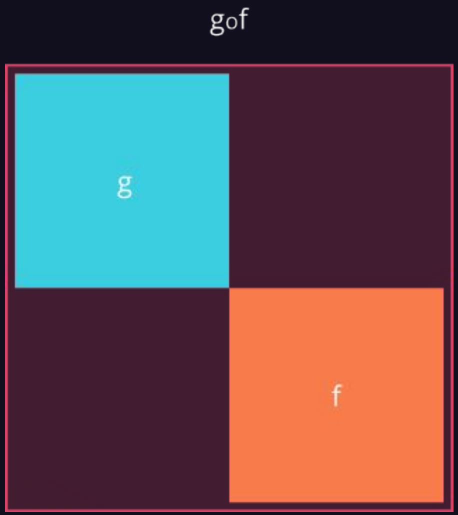
Objective Function

Minimize Squared Distances
from Expected Value

$$\hat{y} = ax + b$$

$$\sum_i (y_i - \hat{y}_i)^2 = 0$$

DE VINCI
INNOVATION
CENTER

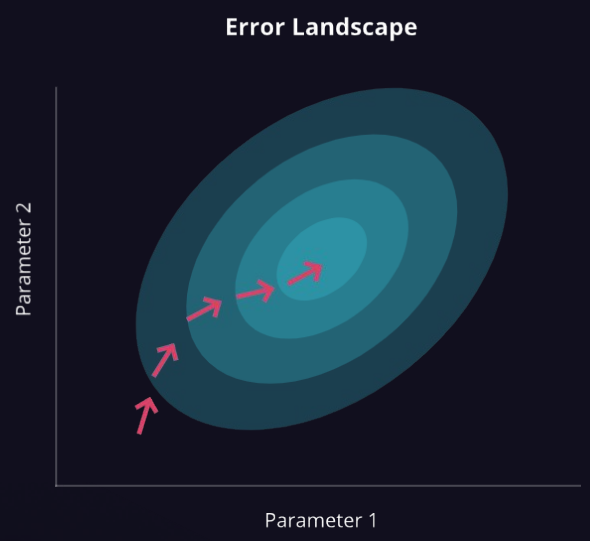


CHAIN RULE

$$h = g \circ f$$

$$h'(x) = g'(f(x))f'(x)$$

|03 OPTIMIZATION

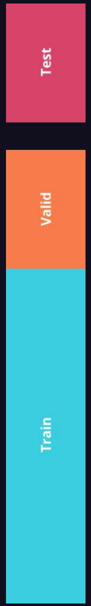


GRADIENT DESCENT

Steps

1) **Forward Propagate** Through the Chain
2) Compute Output **Error**
3) **Backpropagate** Error Through the Chain
4) **Update** Weights w/ Learning Rate
5) **Repeat** Until Convergence Threshold

DE VINCI INNOVATION CENTER



DATASET SPLIT

Training Set

Samples used to Fit/Train the Model

Validation Set

Samples used to provide an Unbiased Evaluation of the Model Becomes Biased during Training

Testing Set

Samples used to provide an Unbiased Evaluation of the Model after Training

DE VINCI INNOVATION CENTER

CROSS-VALIDATION

Iteration 1	Valid	Train	Train	Train
Iteration 2	Train	Valid	Train	Train
Iteration 3	Train	Train	Valid	Train
Iteration 4	Train	Train	Train	Valid
	Fold 1	Fold 2	Fold 3	Fold 4

Method

1) Split Dataset into **k-Folds**
2) **Train on k-1 Folds** and Validate w/ Last
3) **Repeat** k-times
4) Use **Ensemble** Method for **Inference** or **Retrain** on all Dataset