LABORATORIO DI:

METODI E MODELLI MATEMATICI IN PYTHON

A CURA DI: ANTONIO MIRARCHI & GIUSEPPE TROTTA

https://www.labmetodiemodelli.it/

Dove eravamo rimasti?

Manipolare I dati con
Pandas e MatPlotLib

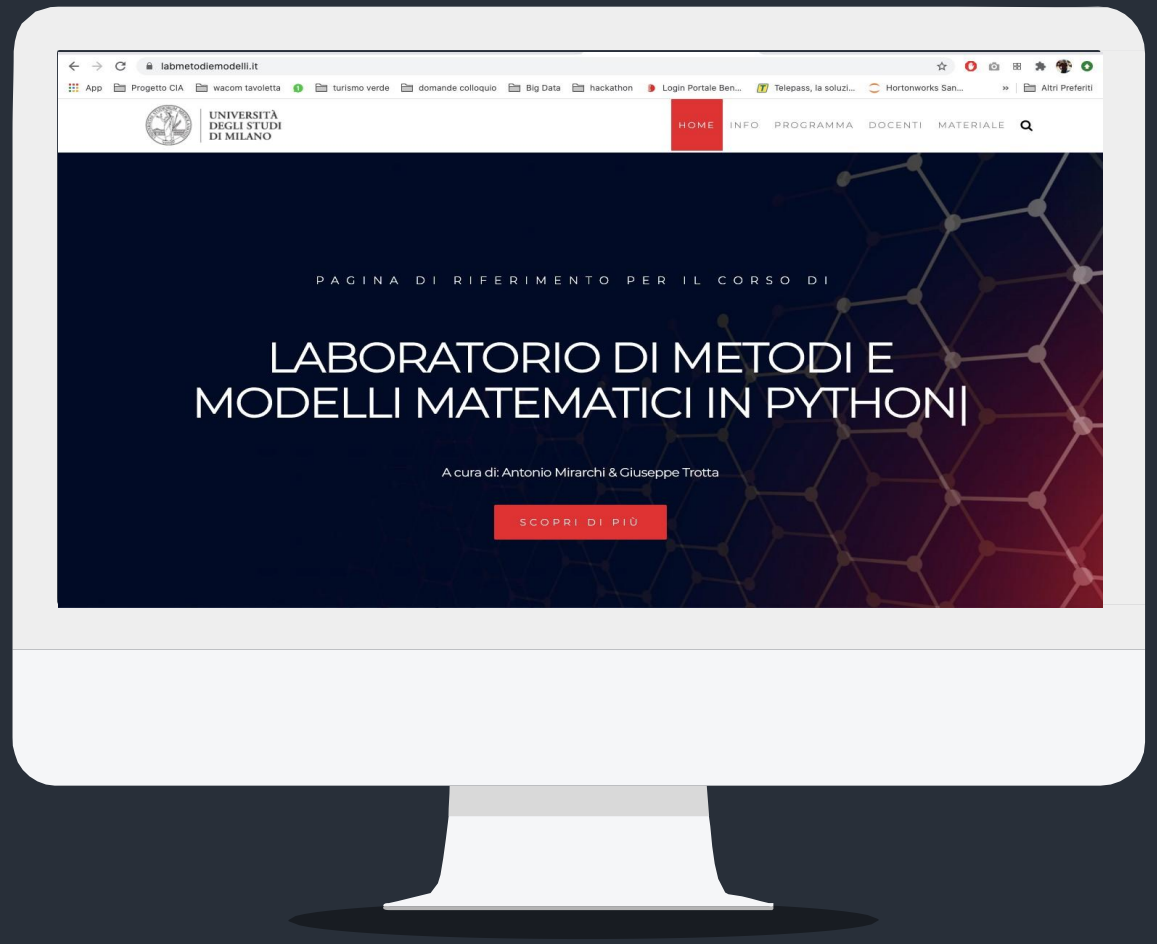
https://www.labmetodiemodelli.it/

Let's Code!

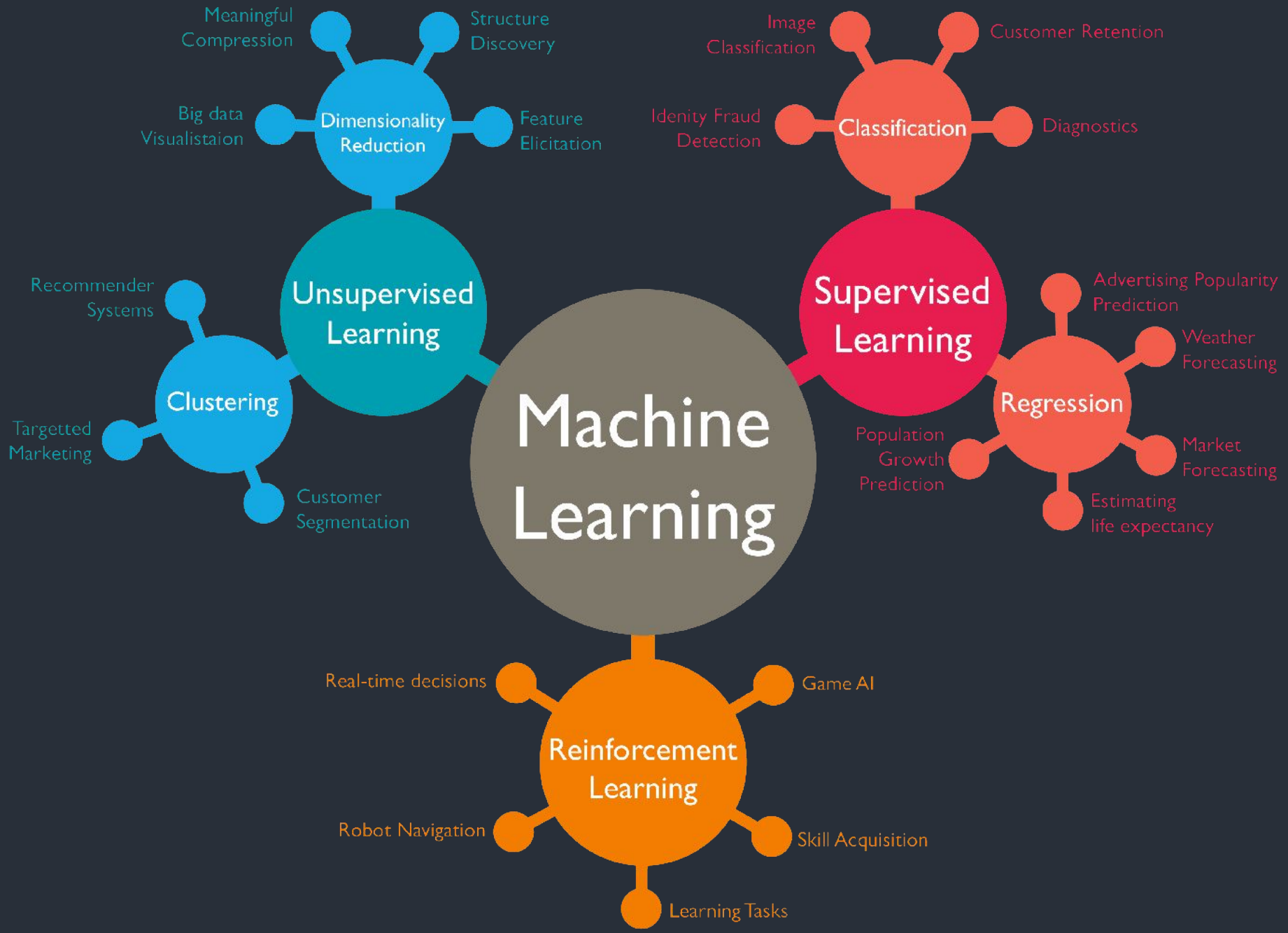
https://www.labmetodiemodelli.it/

LEZIONE 6 – I Modelli Predittivi in Python

1. Machine Learning
2. Costruire un Modello Predittivo
3. Regressione



https://www.labmetodiemodelli.it/



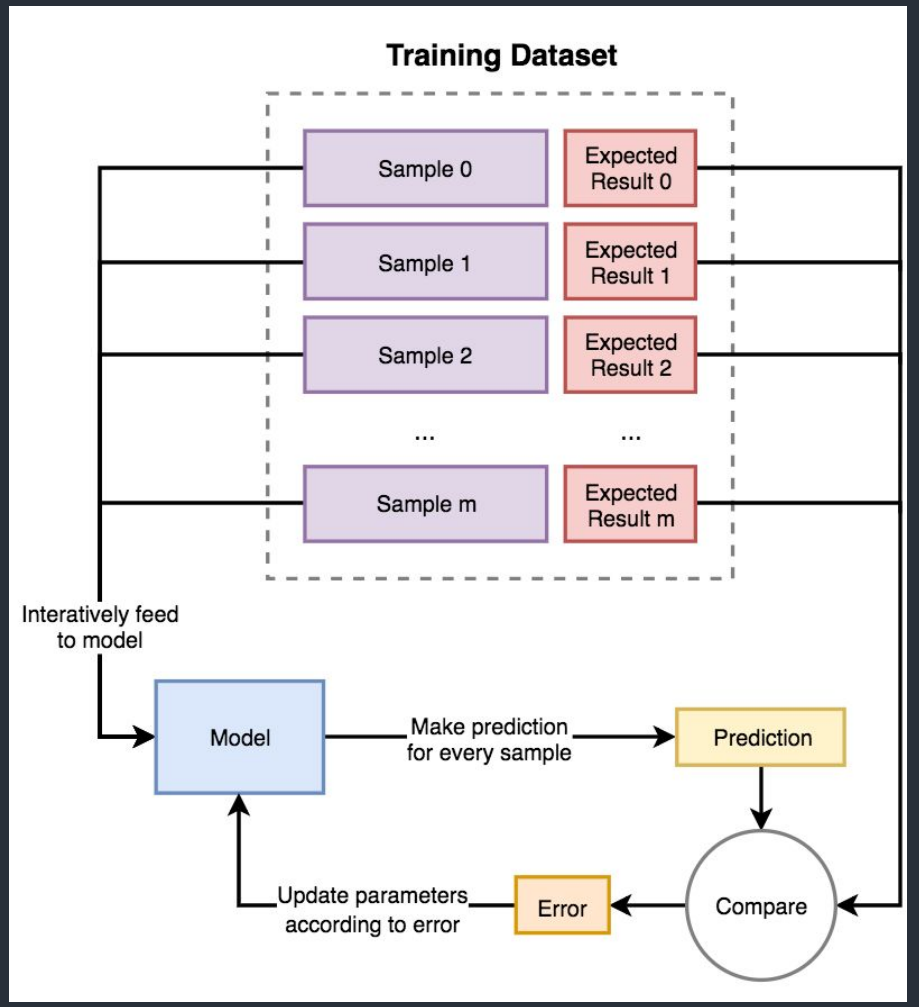
https://www.labmetodiemodelli.it/

Supervised Learning

In training process, samples are being iteratively fed to the model. For every sample, the model uses the current state of parameters and returns a prediction. Prediction is compared to label, and the difference is called an error. **The error is a feedback for the model of what went wrong and how to update itself in order to decrease the error in future predictions.** This means that model will change the values of its parameters according to the algorithm based on which it was created.

Supervised Learning models are **trying to find parameter values that will allow them to perform well on historical data**. Then they are **used for making predictions on unknown data**, that was not a part of training dataset.

1. Classification
2. Regression



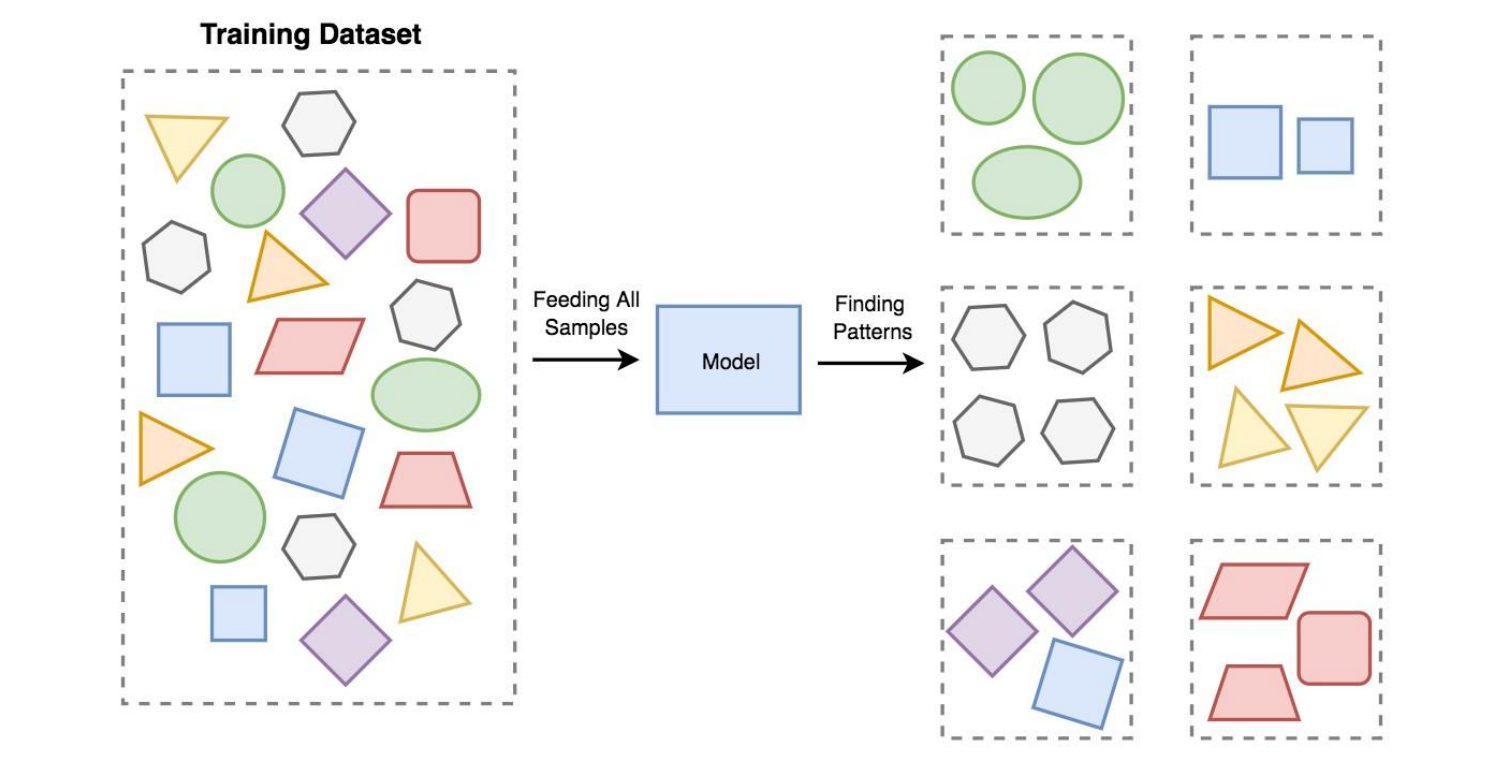
https://www.labmetodiemodelli.it/

Unsupervised Learning

Group of algorithms that try to draw inferences from non-labeled data (without reference to known or labeled outcomes). In Unsupervised Learning, there are no correct answers.

Models based on this type of algorithms can be used for discovering unknown data patterns and data structure itself

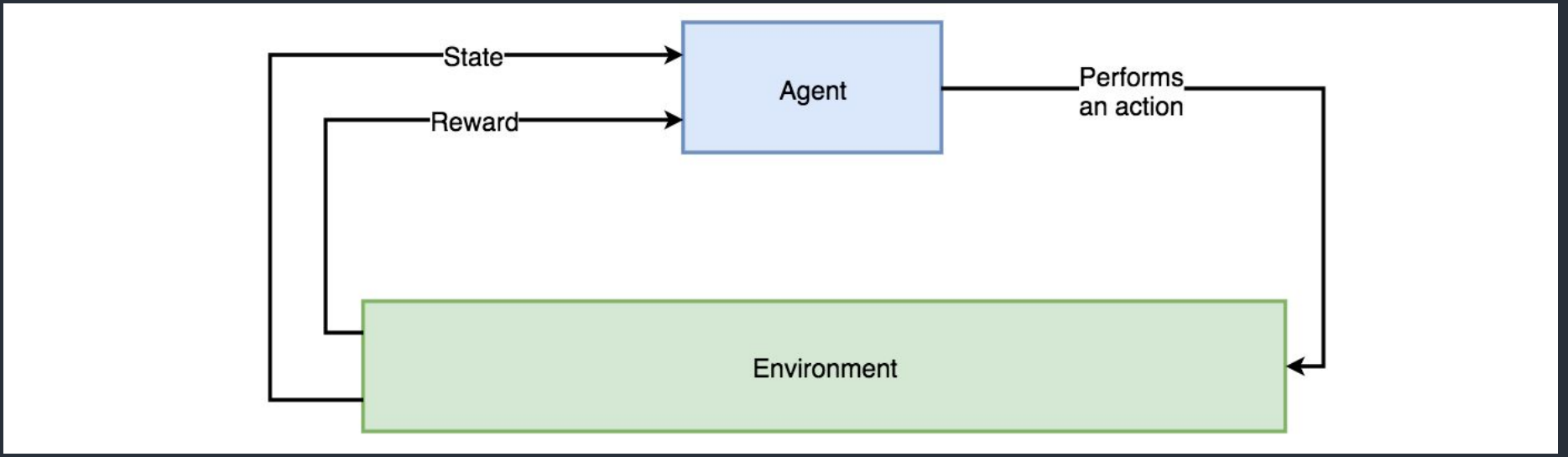
1. Pattern recognition and data clustering
2. Reducing data dimensionality



https://www.labmetodiemodelli.it/

Reinforcement Learning

Training of an agent is a process of trial and error. It needs to find itself in various situations and get punished every time it takes the wrong action in order to learn. The goal of optimization can be set in many ways depending on Reinforcement Learning approach e.g. based on Value Function, Gradient Policy or Environment Model.



https://www.labmetodiemodelli.it/

APPRENDIMENTO SUPERVISIONATO

https://www.labmetodiemodelli.it/

Assumption

REGRESSIONE LINEARE

REGRESSIONE POLINOMIALE

REGRESSIONE LOGISTICA

REGRESSIONE LINEARE BAYESIANA

https://www.labmetodiemodelli.it/

Build a Predictive Model

01 Import Libraries

02 Import Dataset

03 Cleaning Datasets – Null Value, Remove column etc.

04 Pre-visualizations and Correlation

05 Training a model (Train-Test_Split)

https://www.labmetodiemodelli.it/

Linear Regression

Regression attempts to predict one dependent variable (usually denoted by Y) and a series of other changing variables (known as independent variables, usually denoted by X).

Linear Regression is a way of predicting a response Y on the basis of a single predictor variable X. It is assumed that there is approximately a linear relationship between X and Y. Mathematically, we can represent this relationship as:

$$Y \approx \beta_0 + \beta_1 X + \varepsilon$$

where β_0 and β_1 are two unknown constants that represent intercept and slope terms in the linear model and ε is the error in the estimation.

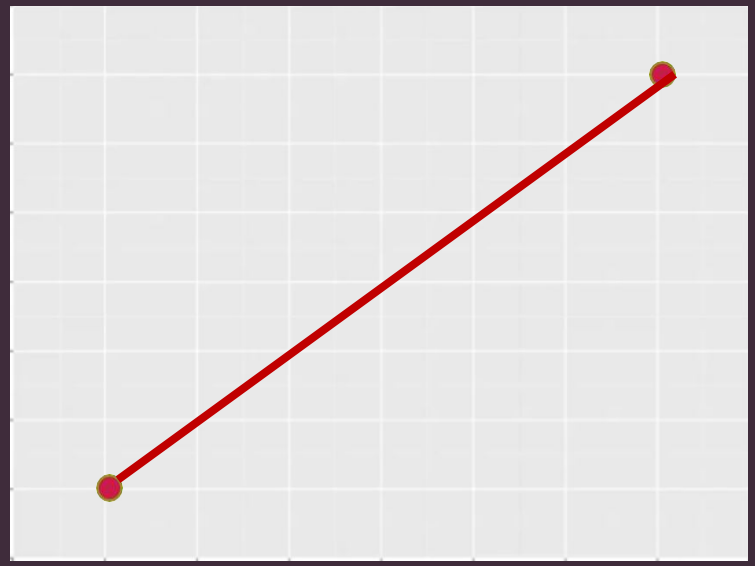
https://www.labmetodiemodelli.it/

Linear Regression

Let's take the simplest possible example. Calculate the regression with only two data points.

Here we have 2 data points represented by two black points. All we are trying to do when we calculate our regression line is draw a line that is as close to every point as possible.

Here, we have a perfectly fitted line because we only have two points.

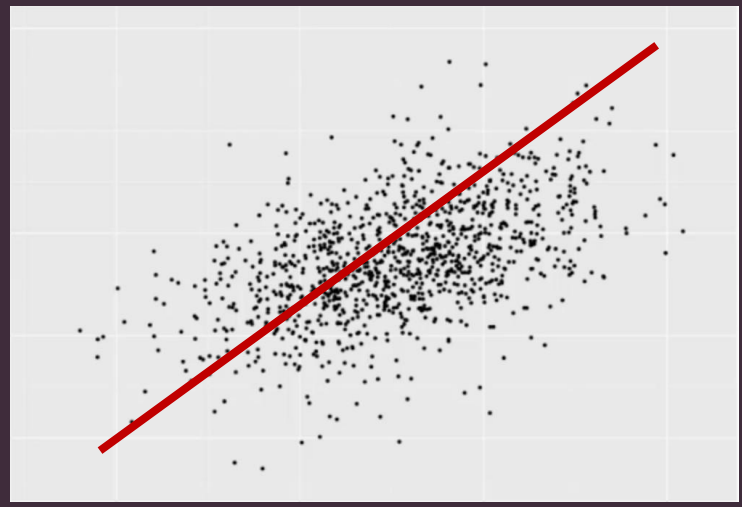


https://www.labmetodiemodelli.it/

Linear Regression

What's up with multiple points?

By applying linear regression we can take multiple X's and predict the corresponding Y values. This is depicted in the plot ->



Our goal with linear regression is to minimise the vertical distance between all the data points and our line.

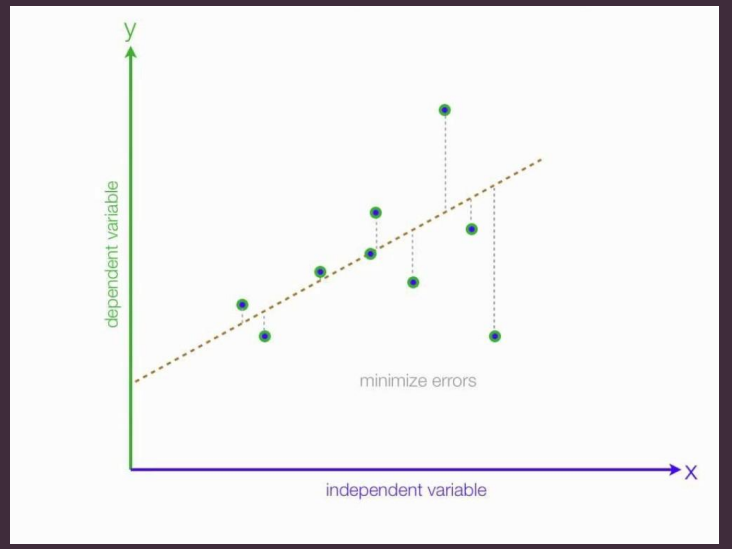
https://www.labmetodiemodelli.it/

Linear Regression

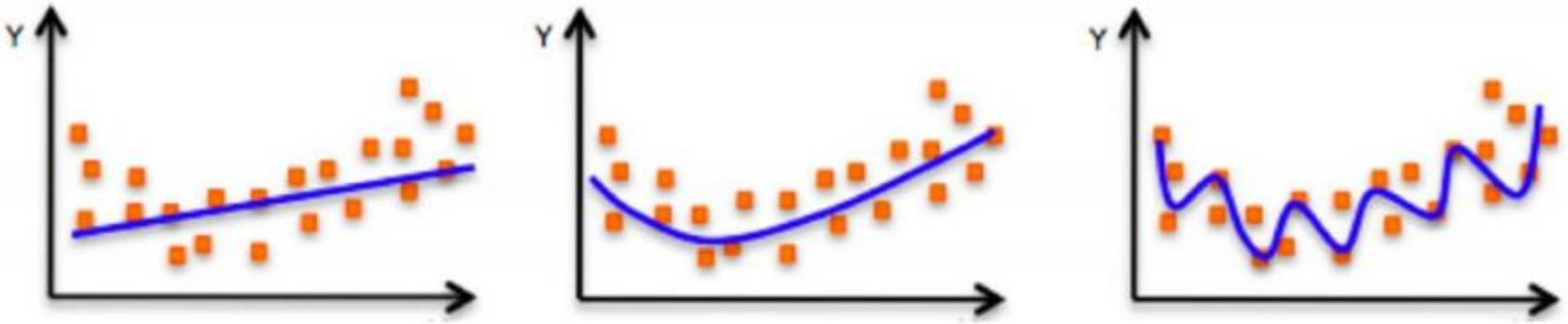
What's up with multiple points?

$$Y = a X + b$$

to find the best fit line, which minimizes error(the sum of the square of the distance between points and the line). The distance between the points and line are taken and each of them is squared to get rid of negative values and then the values are summed which gives the error which needs to be minimized.



https://www.labmetodiemodelli.it/



https://www.labmetodiemodelli.it/

Let's Code!

https://www.labmetodiemodelli.it/

Multiple Linear Regression

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.

The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable.

In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

$$y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_p x_{ip}+\epsilon$$

https://www.labmetodiemodelli.it/

Multiple Linear Regression

$$y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_p x_{ip}+\epsilon$$

where, for i=n observations:
y_i=dependent variable
x_i=expanatory variables
β_0=y-intercept (constant term)
β_p=slope coefficients for each explanatory variable
ϵ=the model's error term (also known as the residuals)

https://www.labmetodiemodelli.it/

Let's Code!

https://www.labmetodiemodelli.it/

Prossimi Appuntamenti

26 NOV	Costruire Modelli Predittivi – P2 + Test Intermedio
03 DIC	Costruire Modelli Predittivi – P3
10 DIC	Costruire Modelli Predittivi – P4

http://www.labmetodiemodelli.it/