Démystifier le Machine Learning à travers AzureML

Pierre LEROY – Clément ROMAC

2017

Global Azure

BOOTCAMP

Speakers

Pierre LEROY

Alternant Data Scientist, Entrepreneur

Co-fondateur d'Enimia
Big Data, Machine Learning

Clément ROMAC

Alternant Machine Learning & IA

Co-directeur d'Enimia
Machine Learning, Deep Learning

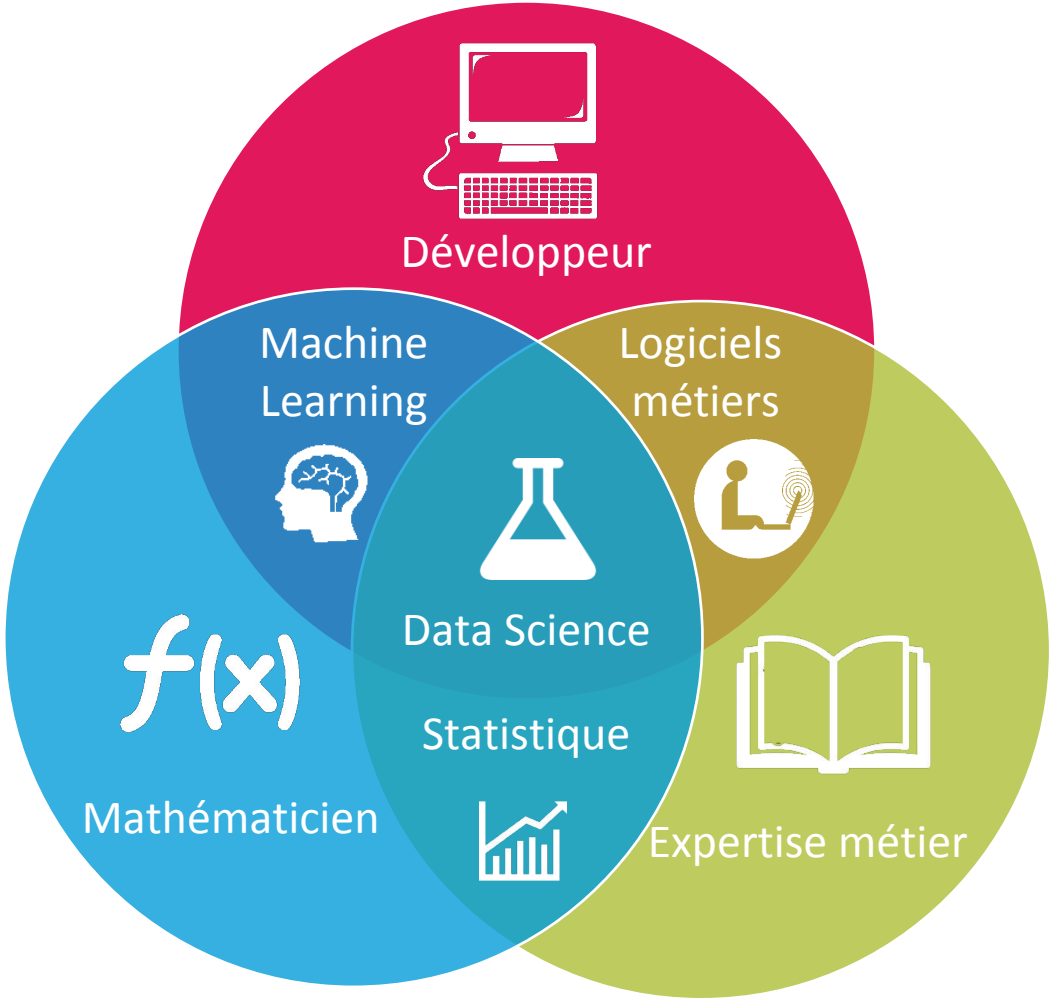
MACHINE LEARNING

DEFINITION DES CONCEPTS DE BASE

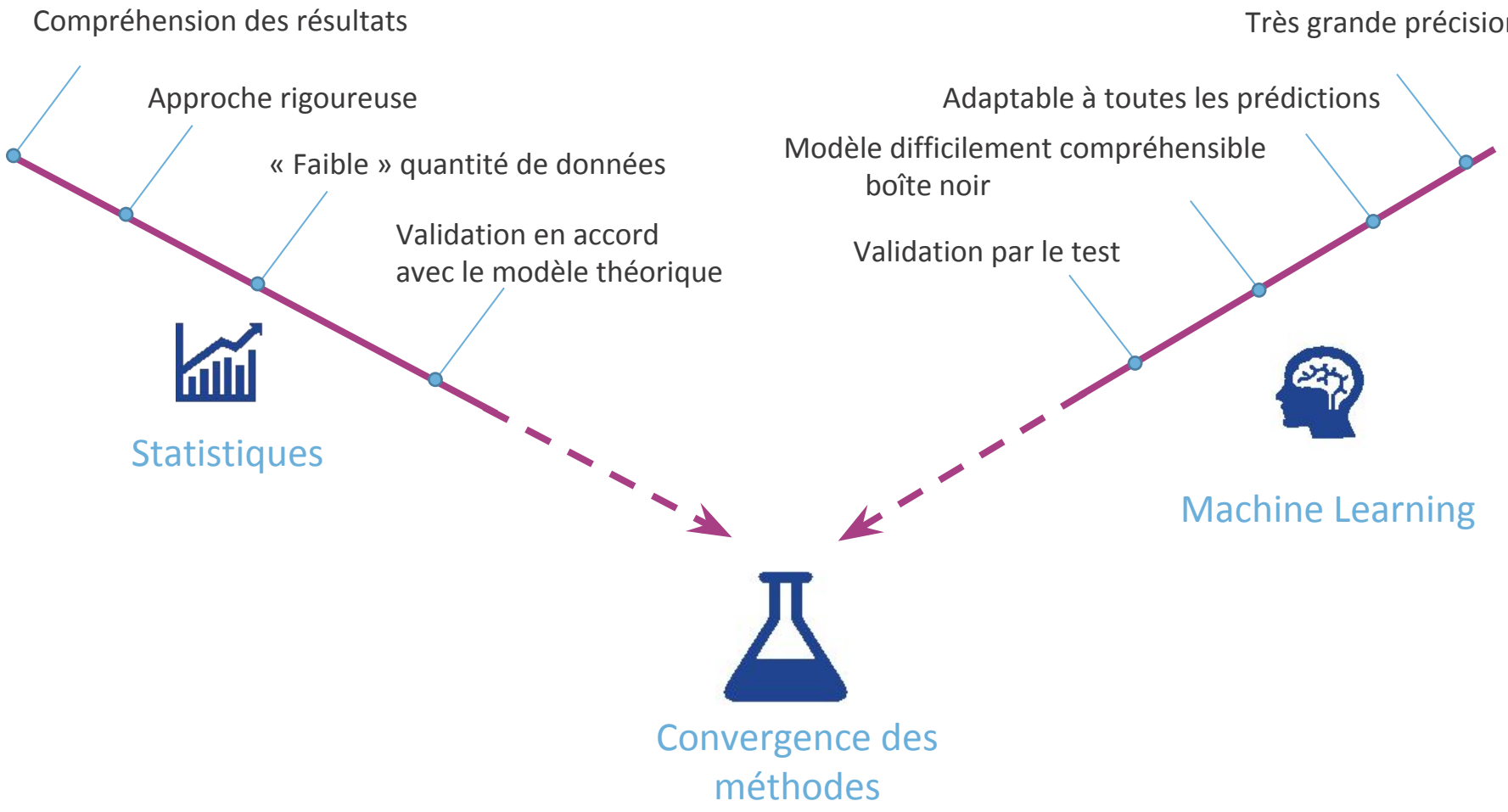
Qu'est ce que le machine learning ?

«*L'apprentissage automatique est la science permettant aux ordinateurs d'accomplir des tâches sans avoir été explicitement programmé dans ce sens*.» Andrew Ng

Qu'est ce que le machine learning ?

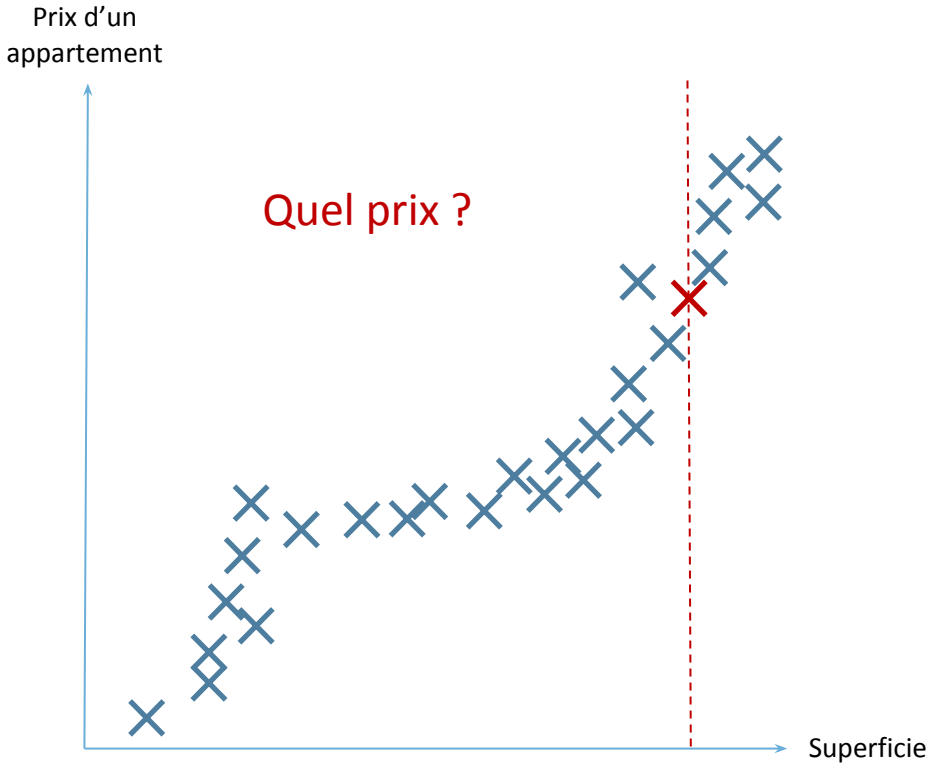


Le machine Learning, des statistiques faites par les développeurs ?

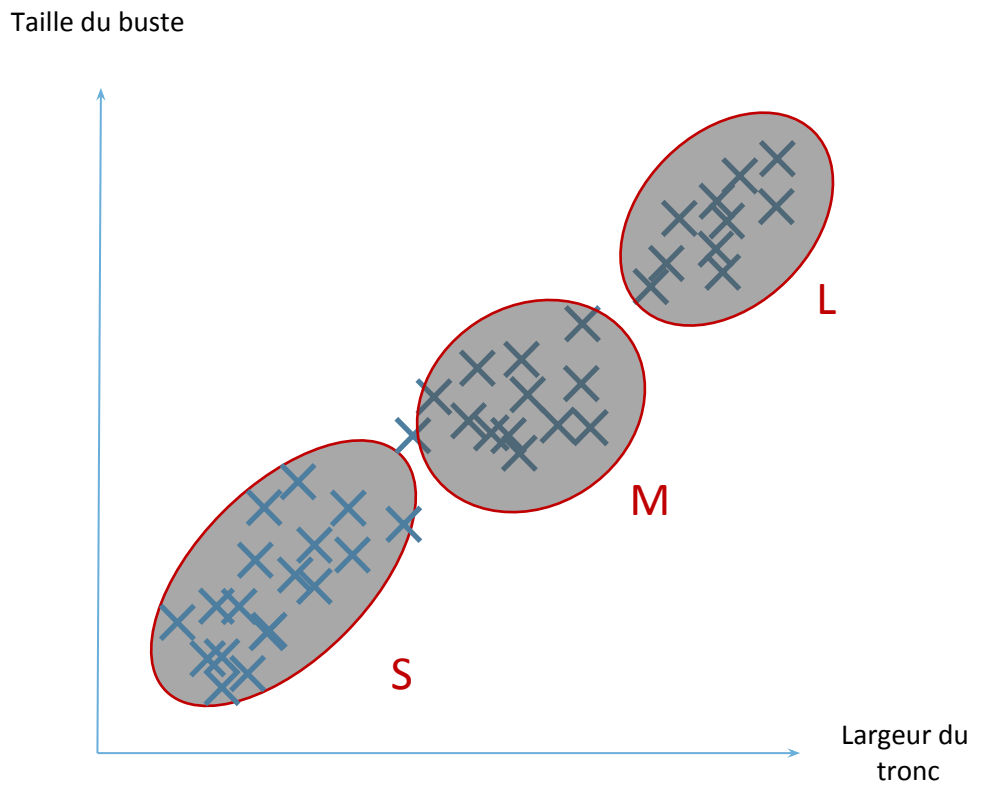


Deux familles d'algorithmes

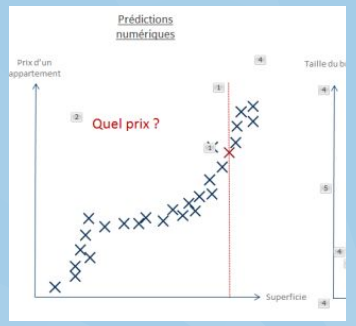
Supervisé
(données labellisées)



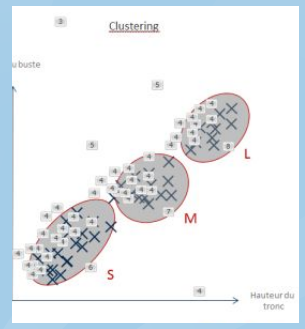
Non Supervisé
(données non labellisées)



QUIZZ



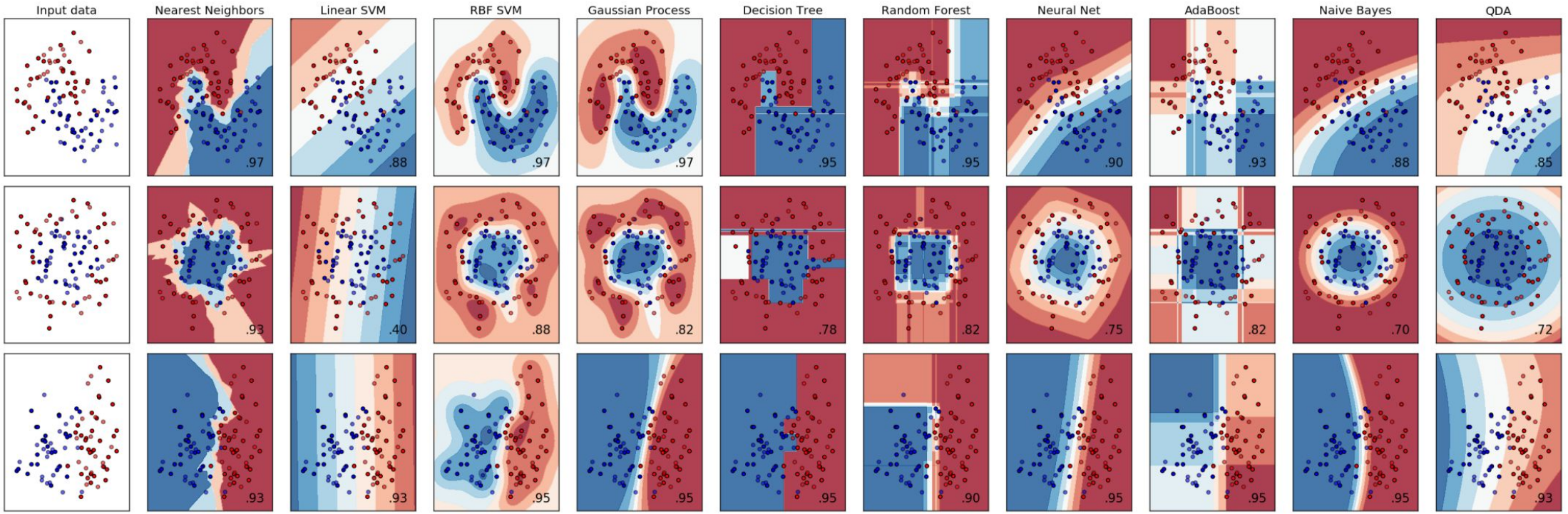
Supervisé



Non supervisé

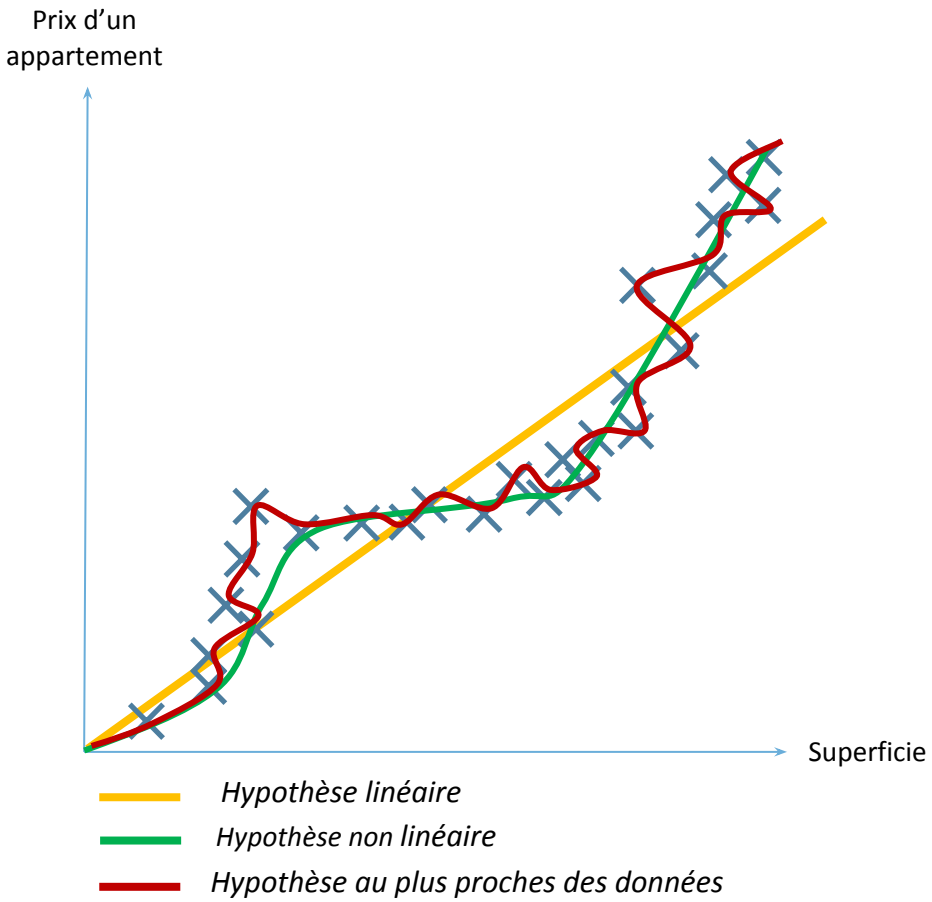
	Supervisé	Non supervisé
Connaître l'âge d'un utilisateur	✓	
Détecter d'anomalies		✓
Prévoir si un programme est malveillant		✓
Reconnaître un visage sur une image	✓	
Prévoir le temps de rentabilité d'une action en bourse	✓	

Comment choisir le « bon » algorithme ?

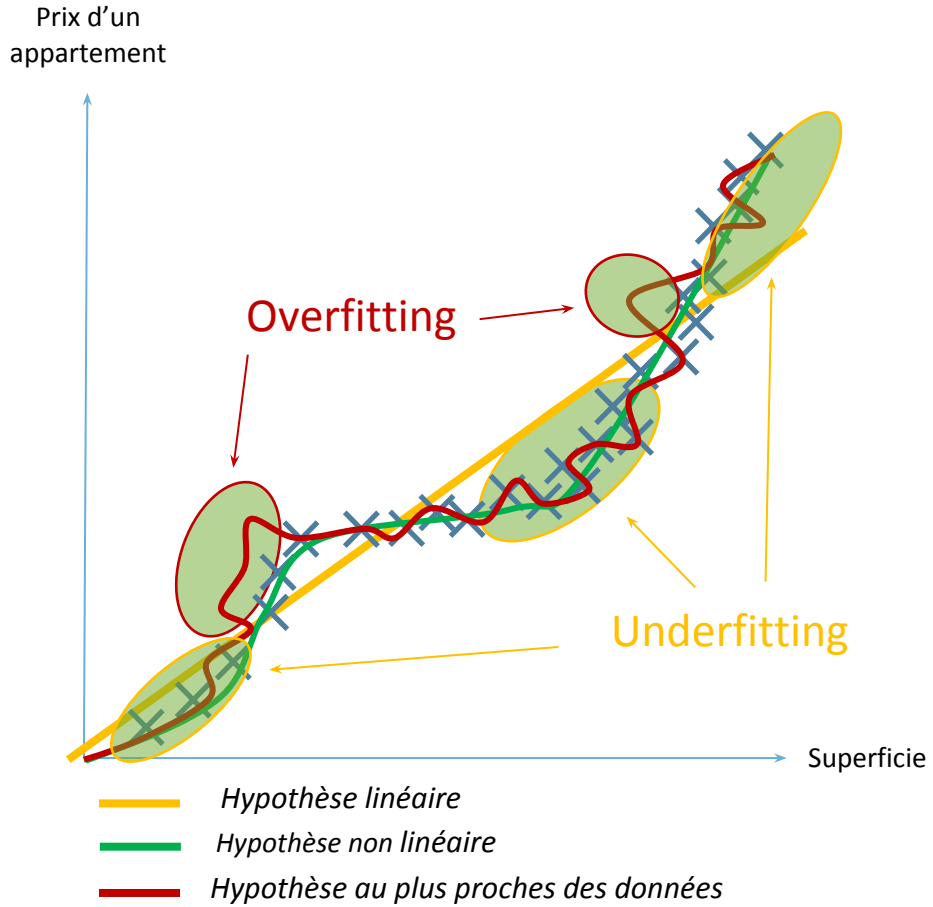


Quelles métriques pour valider mon algorithme ?

Données connues en base de données

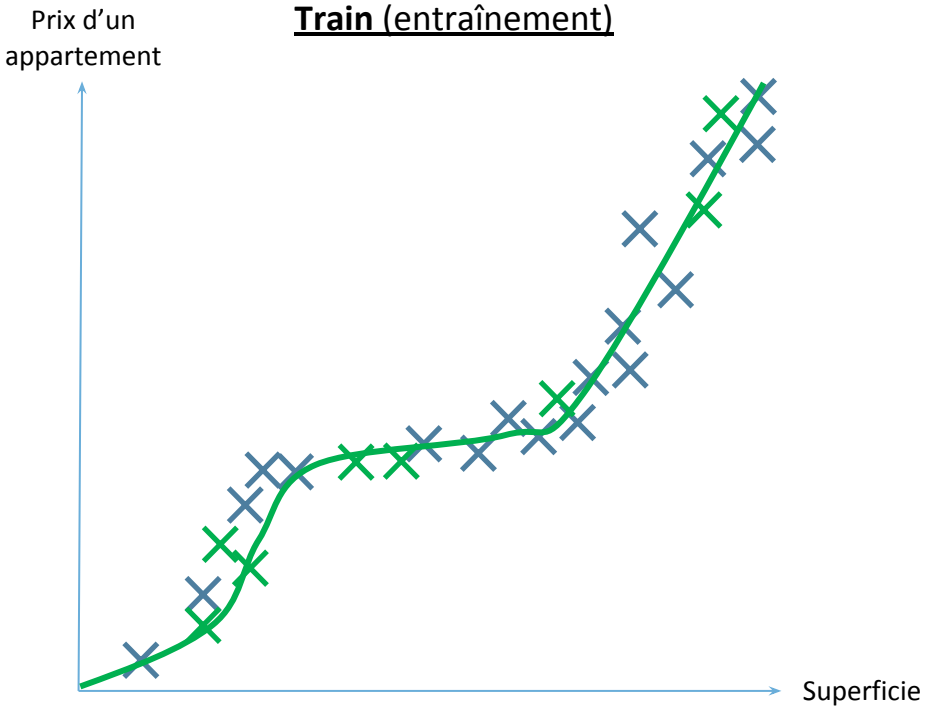


Nouvelles données entrantes

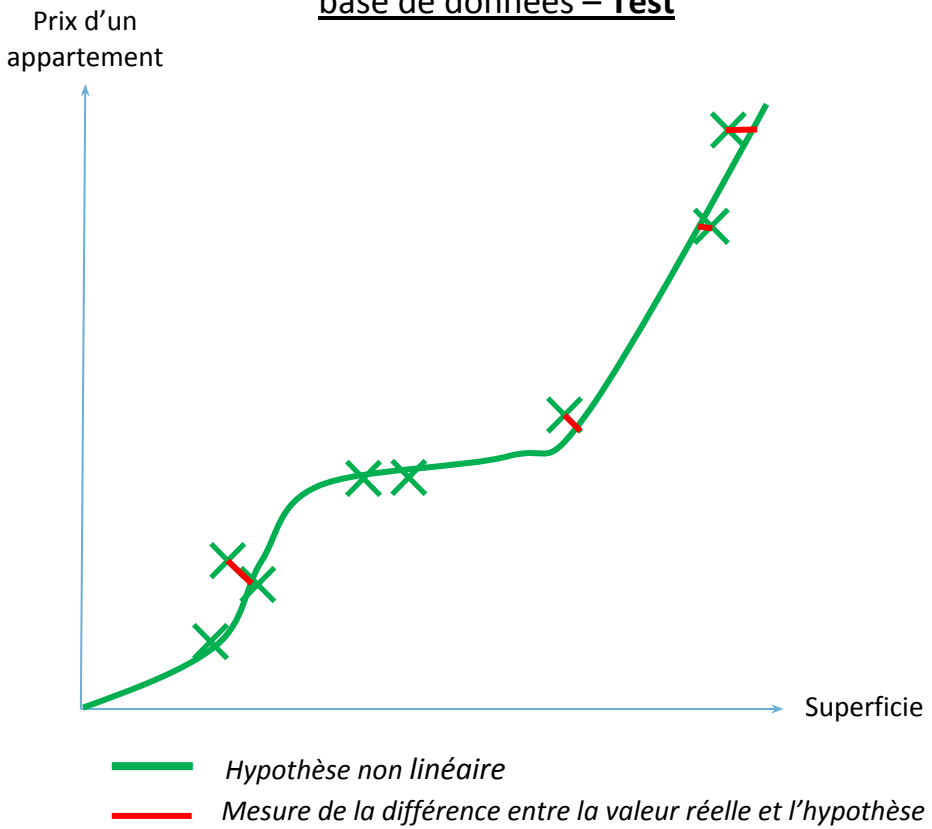


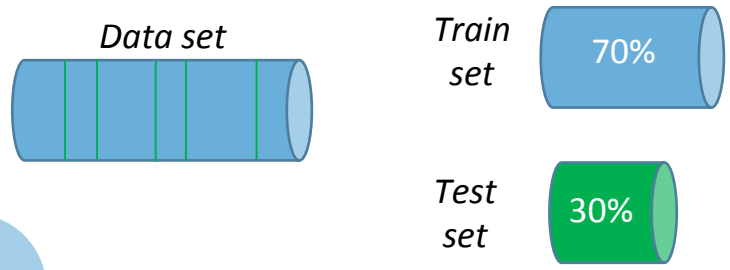
Vérifier son algorithme

Données connues en base de données – **Train** (entraînement)

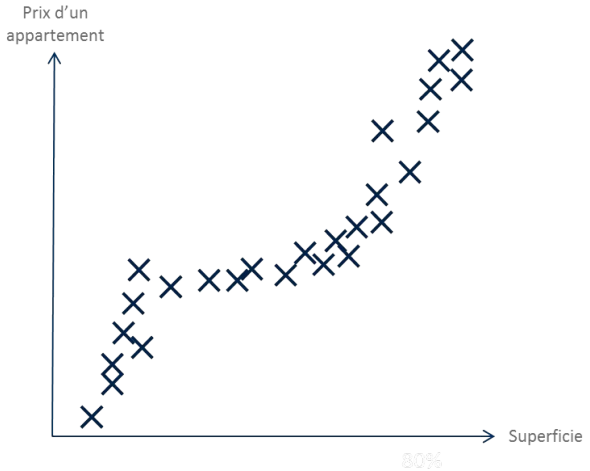


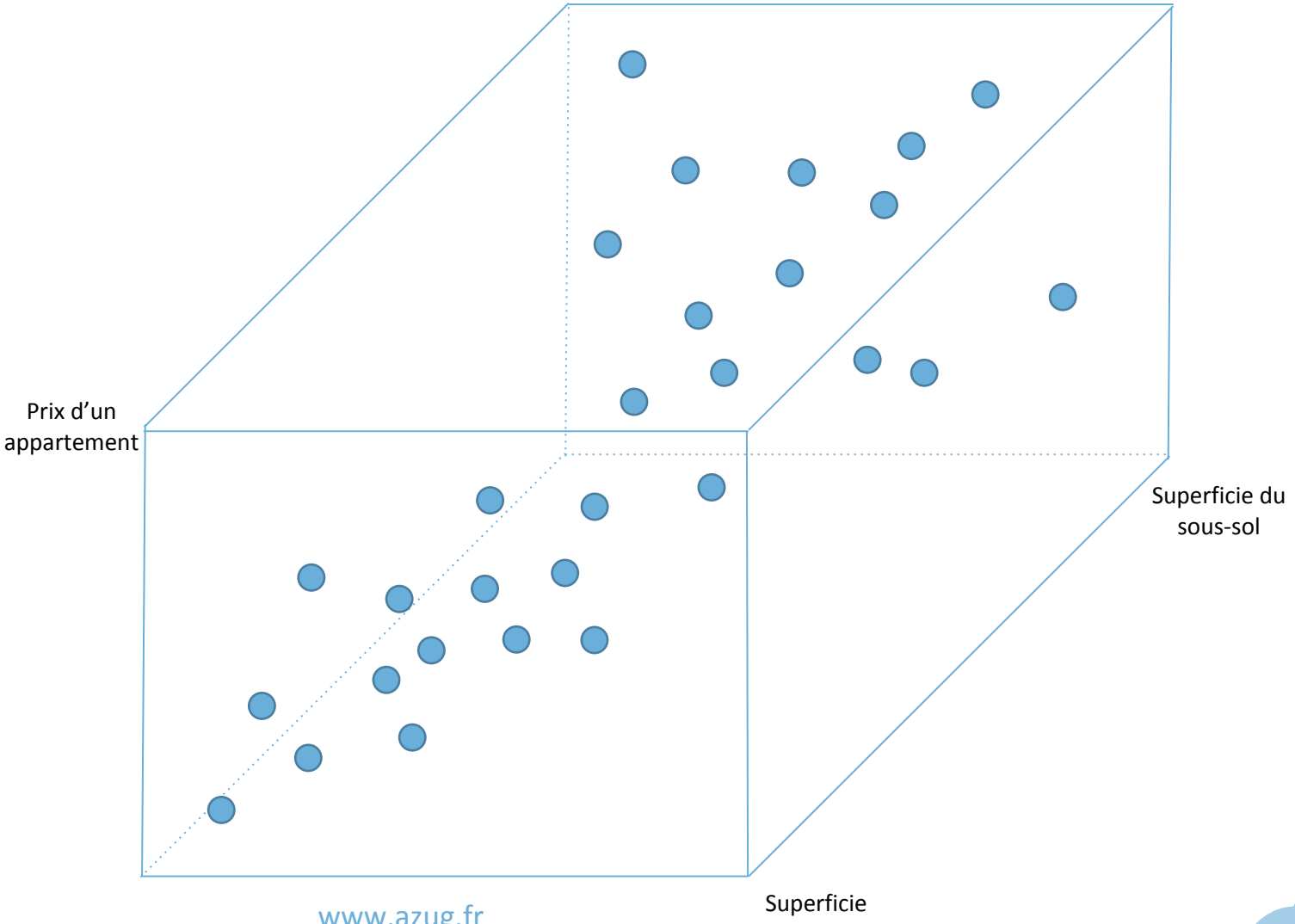
Données connues en base de données – **Test**





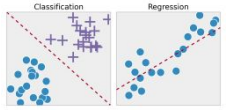
Plusieurs caractéristiques





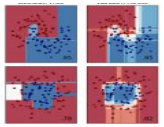
Récapitulatif

Nécessite des données

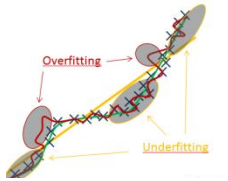


Adaptable à chaque problématique :
Supervisé ou Non Supervisé

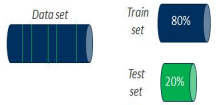
Plus ou moins une boîte noir



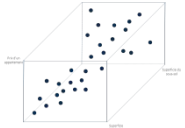
Il existe plusieurs algorithmes qui sont plus ou moins efficaces en fonction de la problématique



2 Phénomènes : Overfitting et Underfitting



Séparation du jeu de données en 2 parties :
Entraînement et Test

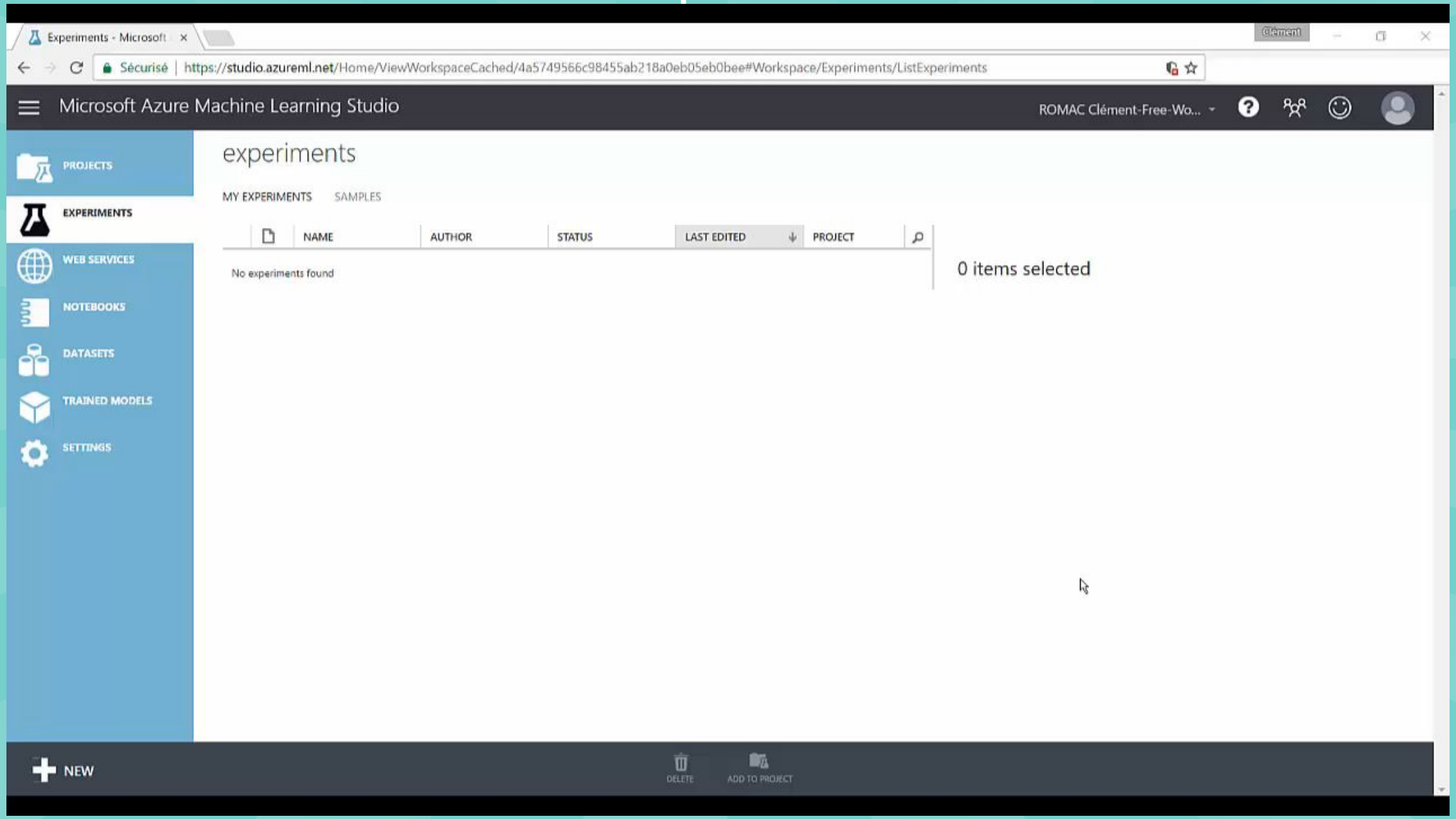


Modèle en x dimension(s), où
x = nombre de paramètres

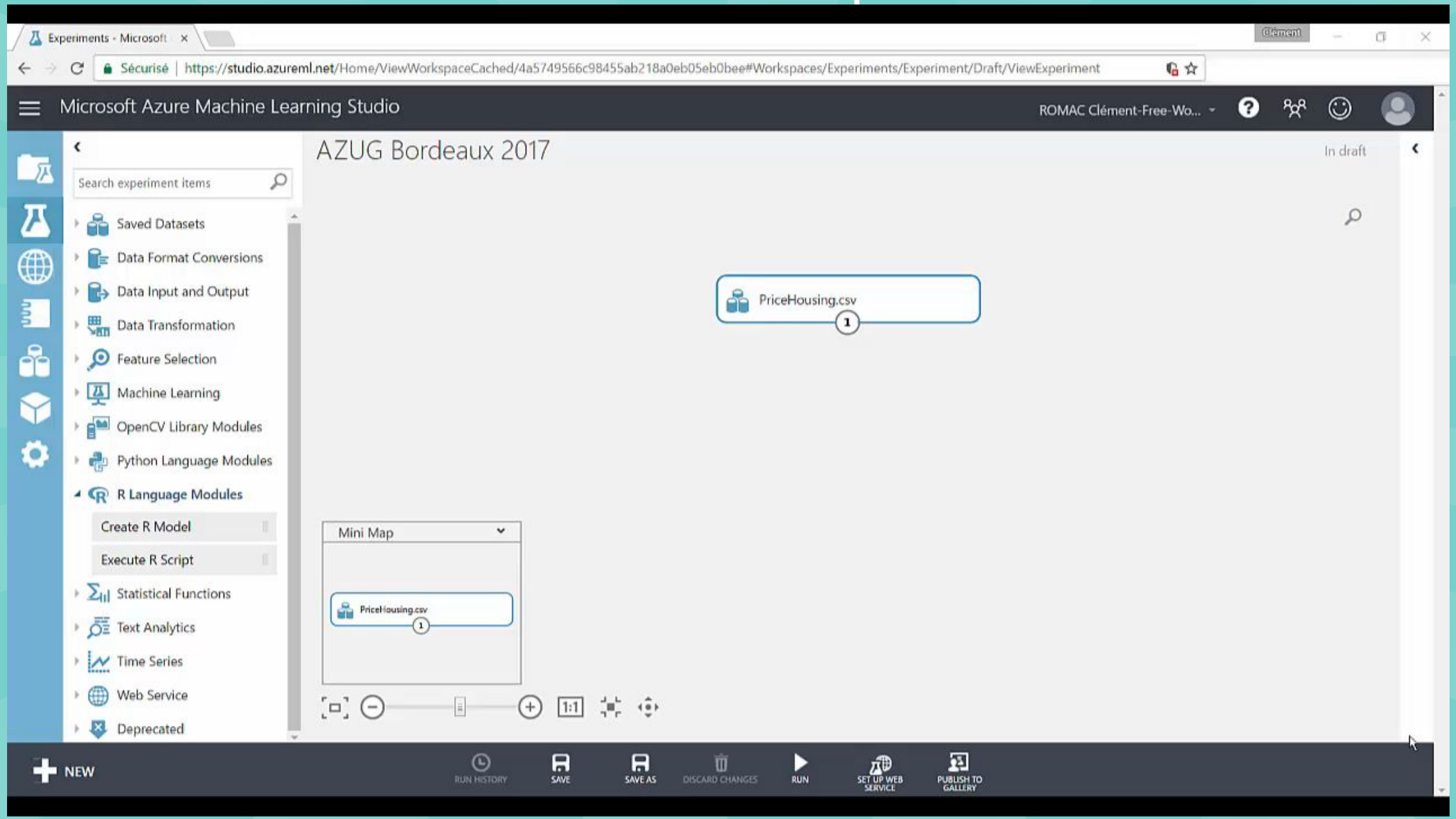
AZURE ML

PREDIRE LE PRIX D'UN APPARTEMENT

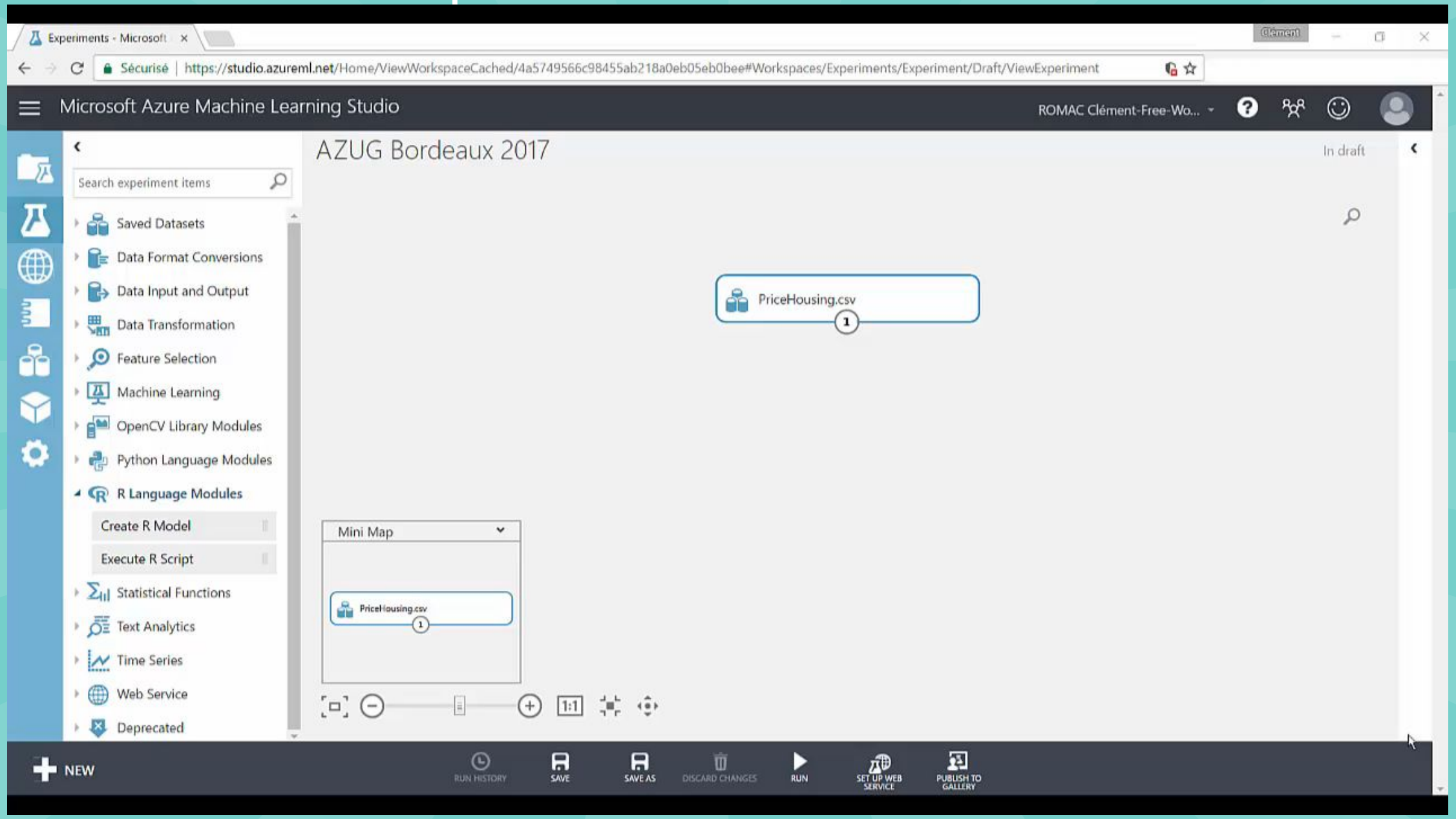
Un rapide tour



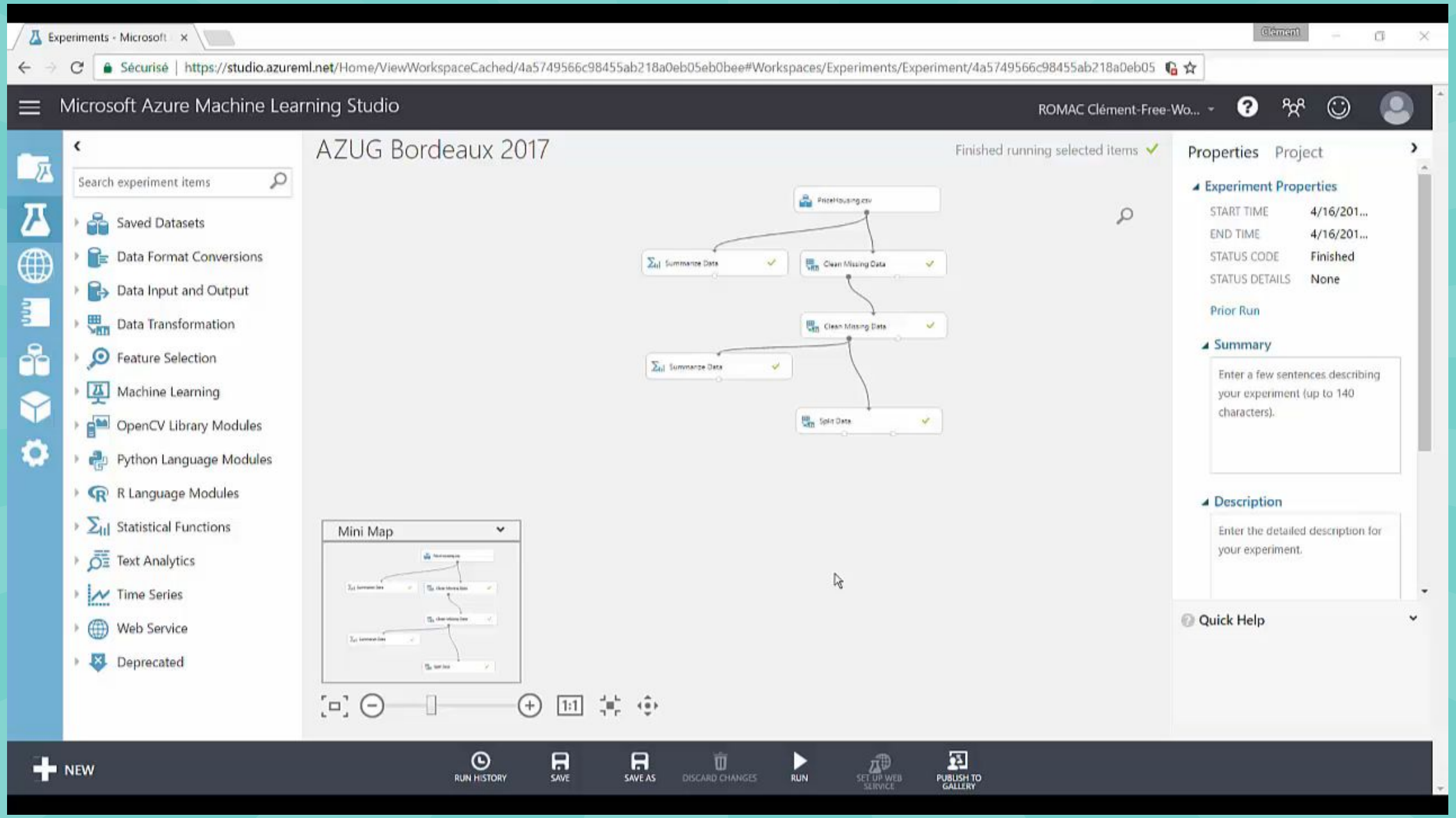
Création d'une expérience



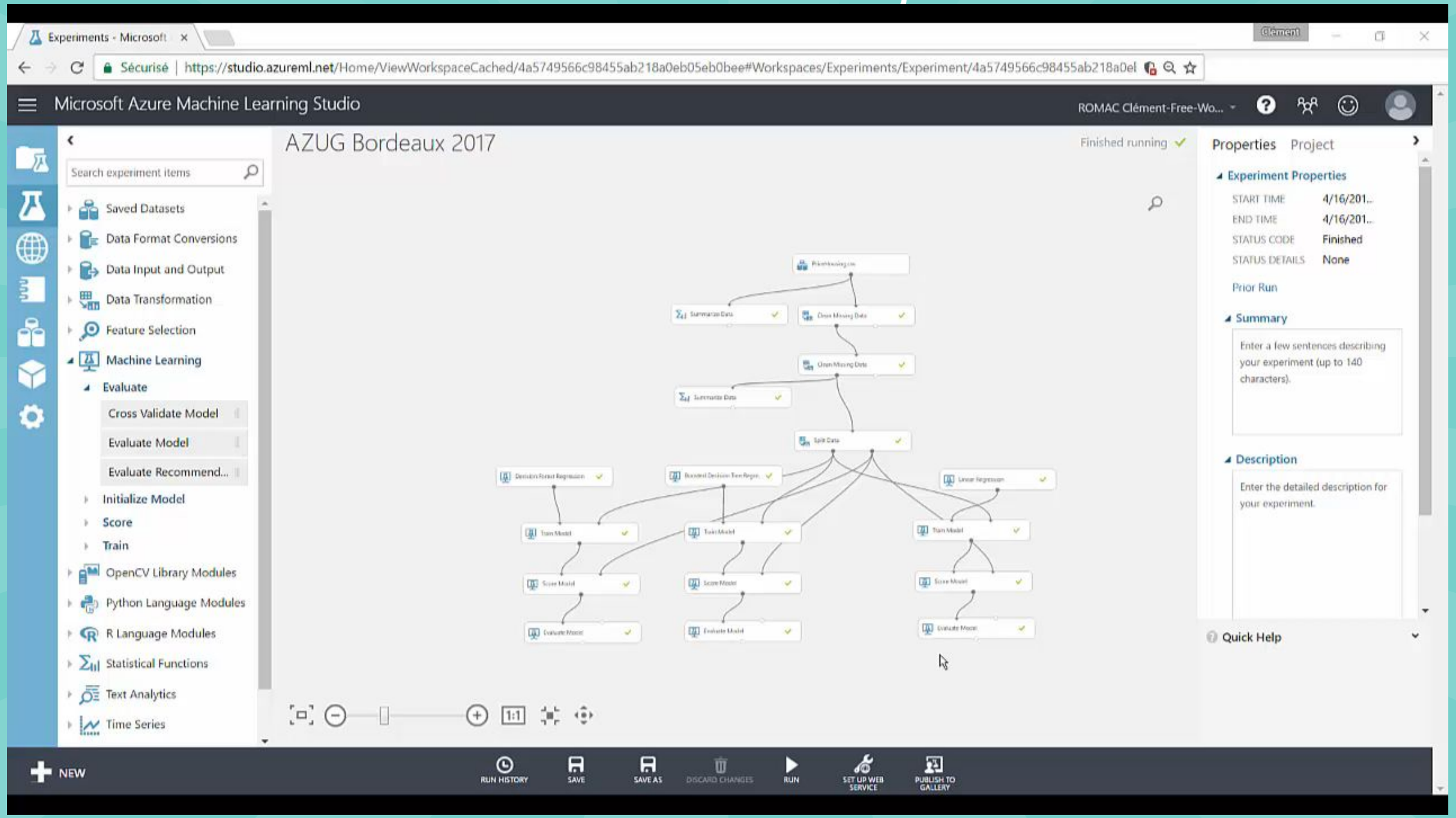
Exploration des données



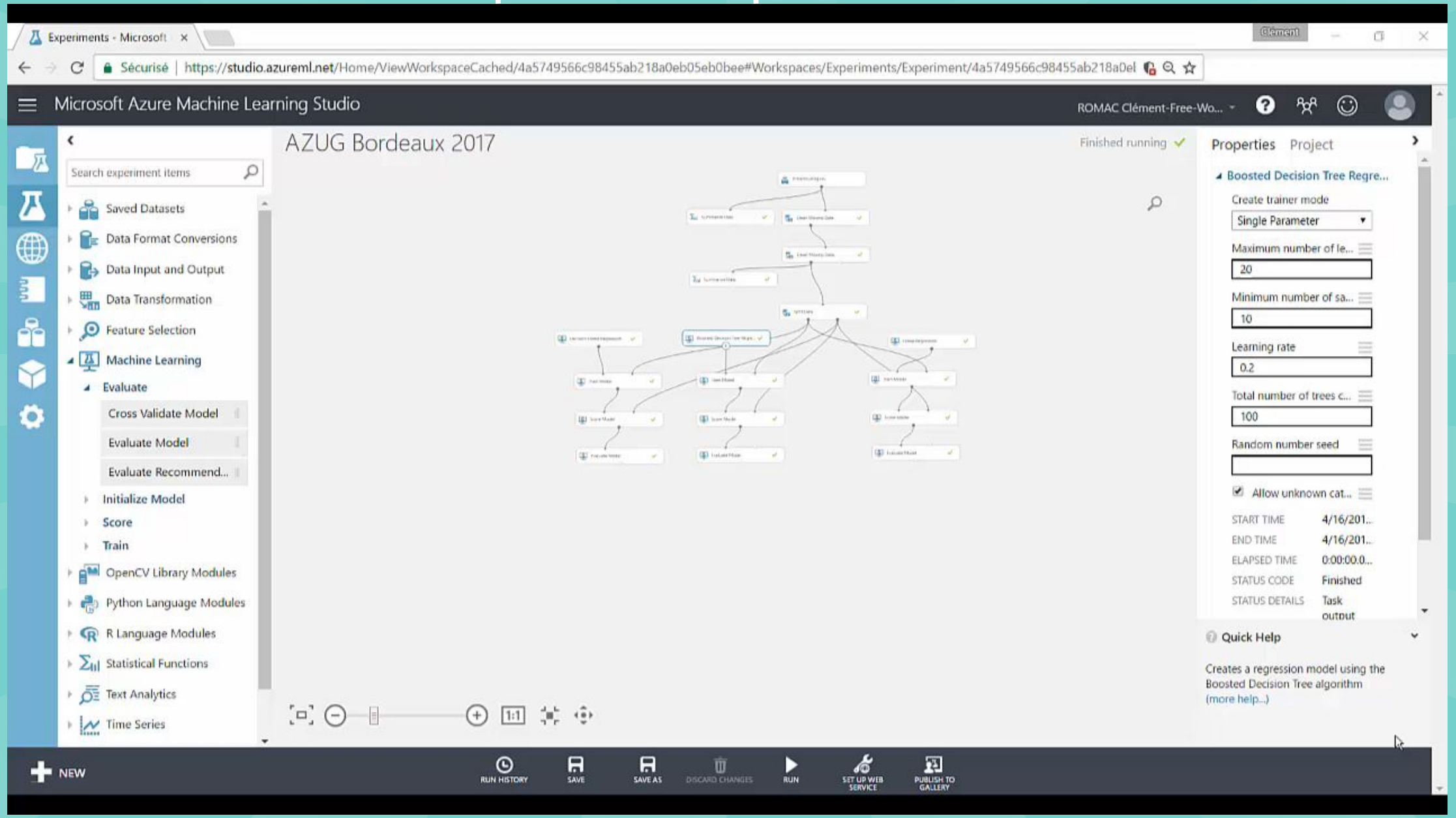
Entrainement d'un modèle



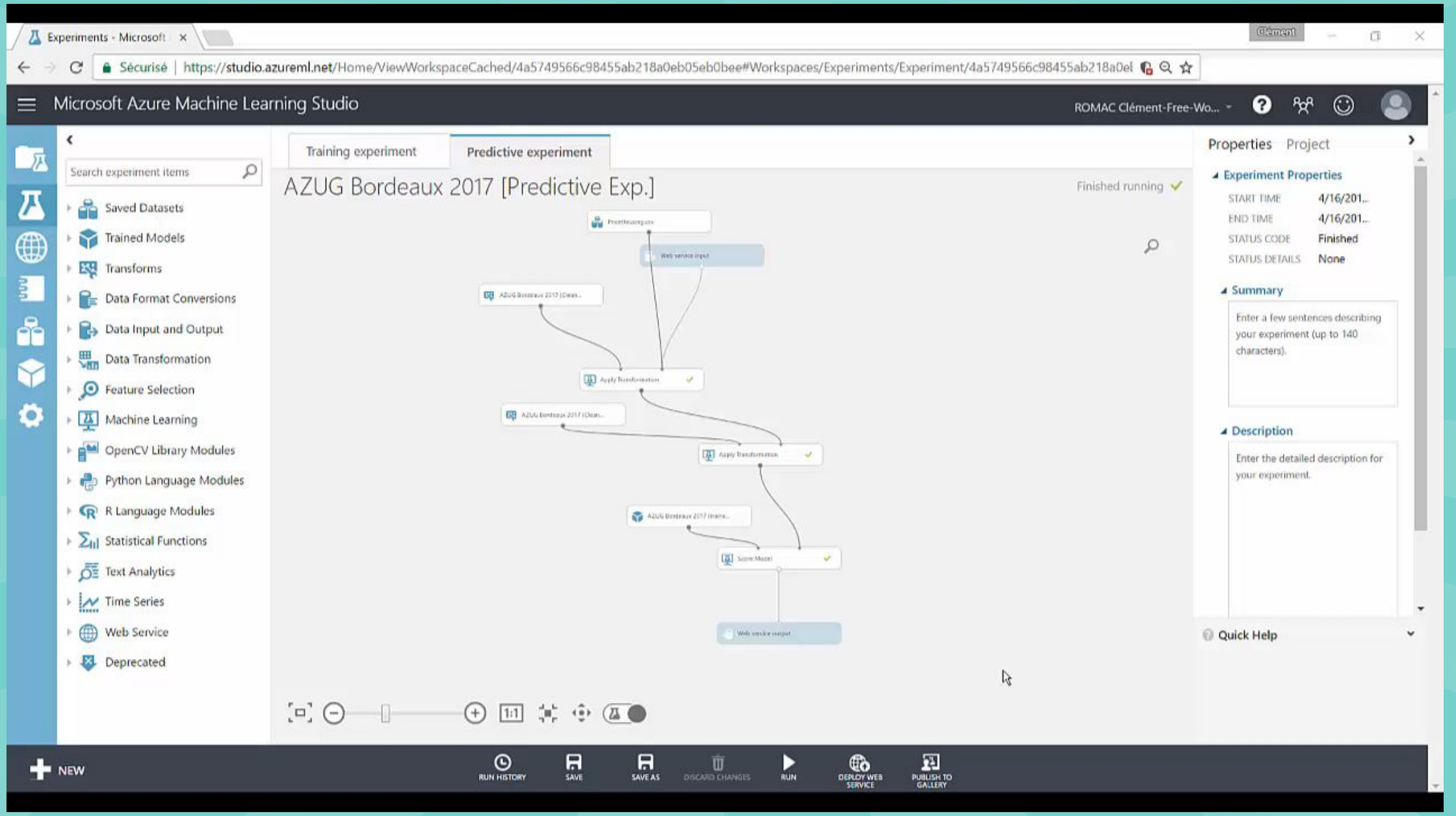
Choix du meilleur algorithme



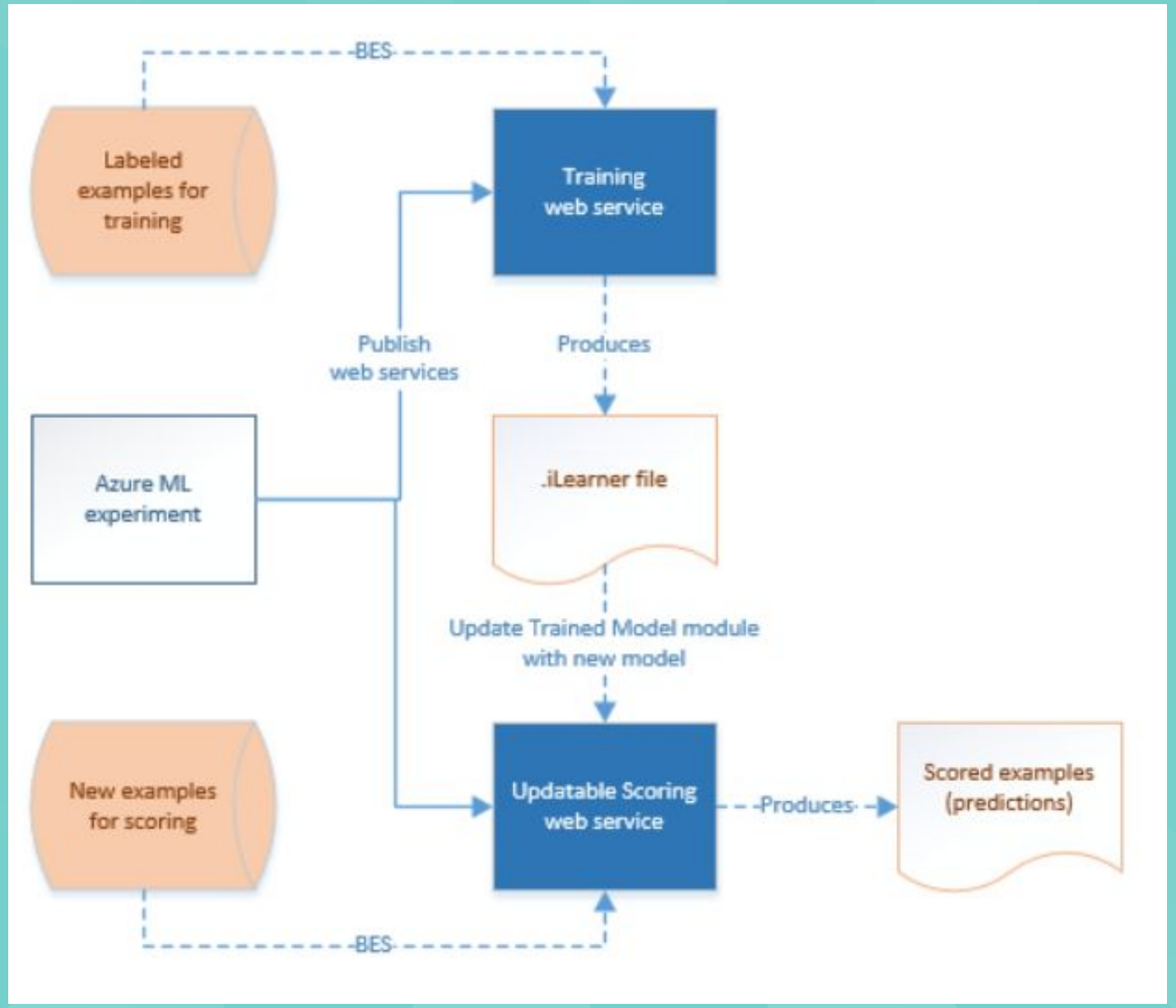
Expérience prédictive



Aller plus loin



Aller plus loin



Vision globale d'un projet de Machine Learning

Équipe métier :
Création d'une *intuition* par l'équipe technique

Équipes métiers :
recueil au près des différentes équipes métiers (différentes sources) et création *d'un entrepôt de donnée*

Data cleaner :
Utilisation d'un ETL pour nettoyer; transformer, enrichir les données. Préparation des données à une modélisation

Data scientist :
utilisation d'algorithmes de Machine Learning pour réaliser une prédiction

Data scientist :
phase d'amélioration avec combinaisons d'algorithmes et techniques de validations croisées

Merci à nos sponsors

PLATINUM

METSYS

Microsoft

[sfeir]

LOCAUX

afib groupe aginius

formation PCL

PARTENAIRES MEDIA

eni Editions

programmez! le magazine des développeurs

2017 Global Azure BOOTCAMP

Sponsors internationaux

psgility

Cloudionix

JET BRAINS

Microsoft

myget

SentryOne
Monitor • Diagnose • Optimize

SERVICEBUS 360

Nous suivre

AZUGFR
Azure User Group France

Facebook
facebook.com/groups/azugfr/

LinkedIn
Linkedin.com/inspirasign

Twitter
twitter.com/AZUGFR

Web
www.azug.fr

Meetup
meetup.com/AZUG-FR/

Nous suivre

LinkedIn

Linkedin.com/in/pierreleroyfr/

LinkedIn

Linkedin.com/in/clement-romac

Twitter

twitter.com/ClementRomac

Enimia

Web

http://enimia.com/


L'enregistrement est ouvert !
http://www.spsevents.org/city/Monaco/Monaco2017
3 juin 2017
#SPSMonaco
Monaco
SHAREPOINT SATURDAY
aOS
MVP
Microsoft®
Most Valuable Professional


Merci
d'être venus
A bientôt !