

Azure Machine Learning et Titanic : plutôt #TeamJack ou #TeamRose ?

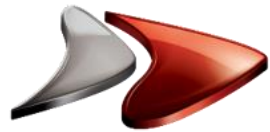
Joël Crest

2019

Global **Azure**
BOOTCAMP



Un autre GRAND merci à nos sponsors locaux !



Projetlys

Exakis Nelite **M**
MagellanPartners

ynov
AIX-EN-PCE

Tentez votre chance

Un livre
2 casques Jabra
10 vouchers à gagner



Faisons connaissance

Joël CREST
Référent Data & Business Intelligence



@JoelCrest



www.bonjourjoel.com



www.linkedin.com/in/joelcrest



www.clubpowerbi.com
<https://www.meetup.com/fr-FR/Club-Power-BI/>


Exakis Nelite 
MagellanPartners

Microsoft
CERTIFIED
Solutions Expert
Data Management and
Analytics

Microsoft
CERTIFIED
Solutions Expert
Business Intelligence



<https://docs.microsoft.com/fr-fr/azure/machine-learning/studio/>

 Microsoft Azure

Présentation

Solutions

Produits

Documentation

Tarification

Formation

Place de marché

Partenaires

Support technique

Blog

plus

Contact de vente: 1-800-867-1389

Azure / Machine Learning / Studio

ModifierPartager

Filtrer par titre

Documentation Machine Learning Studio

Vue d'ensemble

Machine Learning Studio

Comparer nos produits ML

Tutoriels

Exemples

Concepts

Guides pratiques

Informations de référence

Ressources

Documentation Machine Learning Studio

Azure Machine Learning Studio est un outil collaboratif fonctionnant par glisser-déplacer qui permet de générer, tester et déployer des solutions d'analytique prédictive à partir de vos données. Des didacticiels, des vidéos et des exemples de modèle vous montrent comment utiliser Studio pour créer et déployer des modèles Machine Learning.

Didacticiels pas à pas

Consulter des didacticiels étendus qui vous guident dans la création et le déploiement de modèles :

Créer votre première expérience Studio

Créer et déployer une solution prédictive pour le risque de crédit

Informations de référence

Ligne de commande

Modules PowerShell

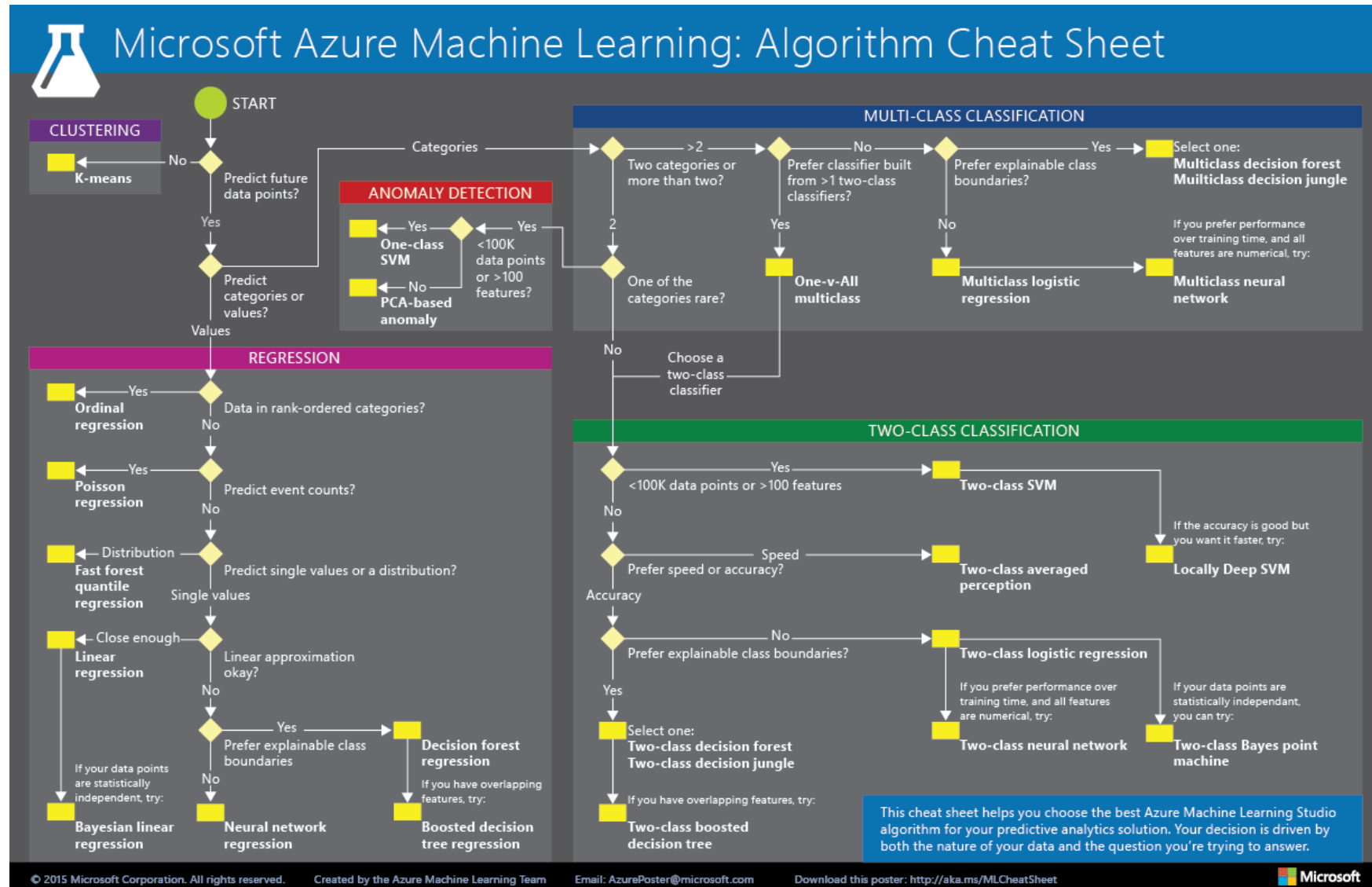
Modules

Algorithmes et informations de référence sur les modules

REST

API de gestion REST

Codes d'erreur de service web



Données

On a besoin d'un échantillon significatif pour entraîner l'algorithme

Algorithmes

On a besoin de choisir l'algorithme correspondant au problème à traiter

Modèle ML

A partir des données et de l'algorithme choisi, on entraîne un modèle

Built-in ML Algorithms

Import Data

Preprocess

Split Data

Train Model

Score Model

A Titanic Probability

Thanks to Kaggle and encyclopedia-titanica for the dataset.

This is the last question of [Problem set 5](#). In this problem you will use real data from the Titanic to calculate conditional probabilities and expectations.



tldr: the ship sinks

On April 15, 1912, the largest passenger liner ever made collided with an iceberg during her maiden voyage. When the Titanic sank it killed 1502 out of 2224 passengers and crew. This sensational tragedy shocked the international community and led to better safety regulations for ships. One of the reasons that the shipwreck resulted in such loss of life was that there were not enough lifeboats for the passengers and crew. Although there was some element of luck involved in surviving the sinking, some groups of people were more likely to survive than others.

The [titanic.csv](#) file contains data for 887 of the real Titanic passengers. Each row represents one person. The columns describe different attributes about the person including whether they survived (*S*), their age (*A*), their passenger-class (*C*), their sex (*G*) and the fare they paid (*X*).

[Question12] Write a program in C, C++, Java or Python that **reads the data file** and finds the answers to the following questions:

Update (May/12): We removed commas from the name field in the dataset to make parsing easier.

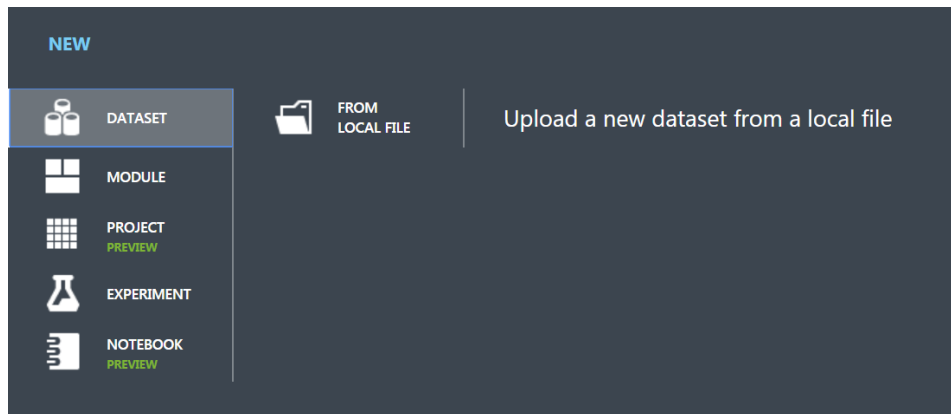
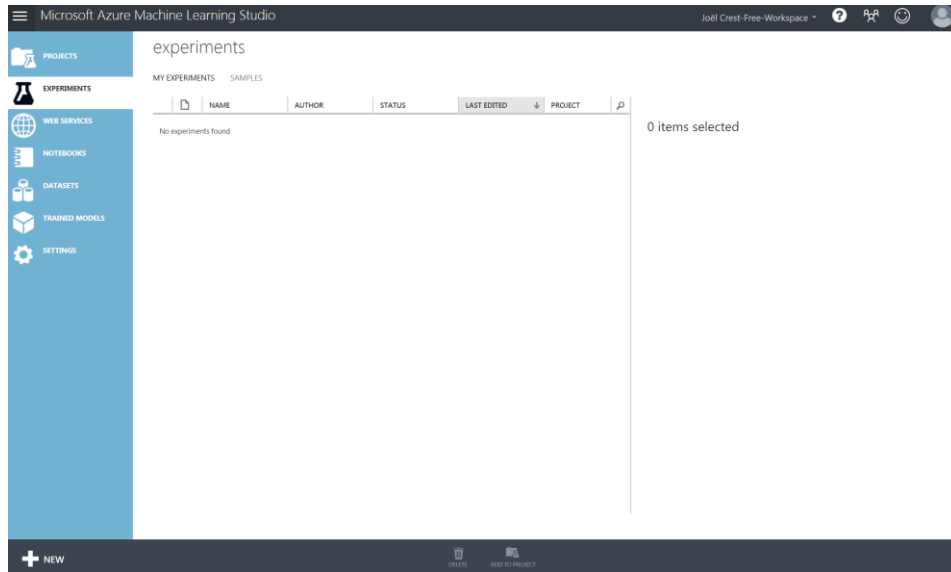
Titanic Dataset



Téléchargez le fichier

<https://studio.azureml.net/>

New + Dataset



✓ Upload of the dataset 'titanic.csv' has completed.

New + Experiment



NEW

- DATASET
- MODULE
- PROJECT
PREVIEW
- EXPERIMENT**
- NOTEBOOK
PREVIEW

Search experiment templates

Microsoft Samples

Blank Experiment

Experiment Tutorial

Sample 1: Download dataset from UCI: Adult 2 class dataset

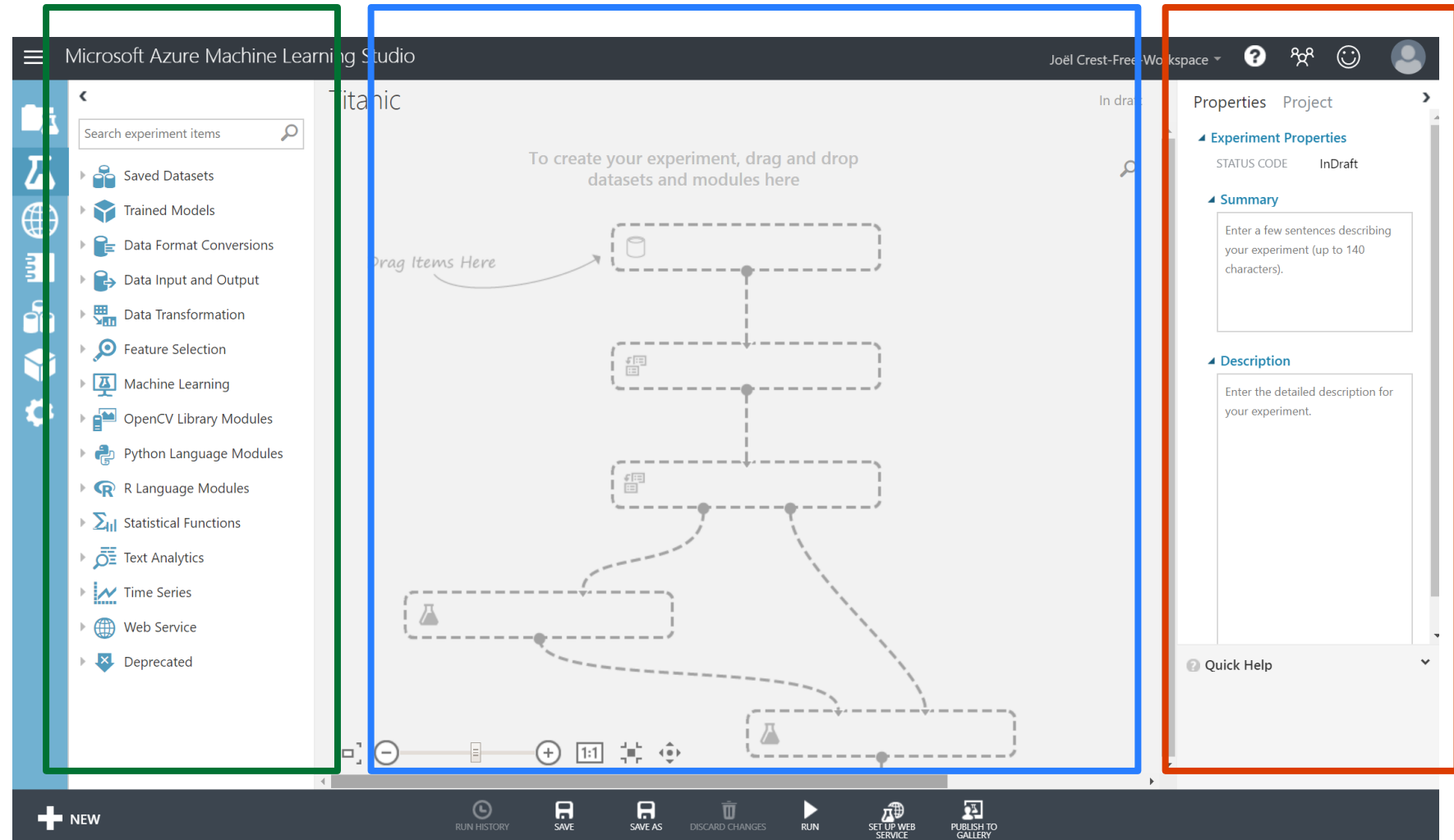
Blank Experiment



Expérience ML

Propriétés

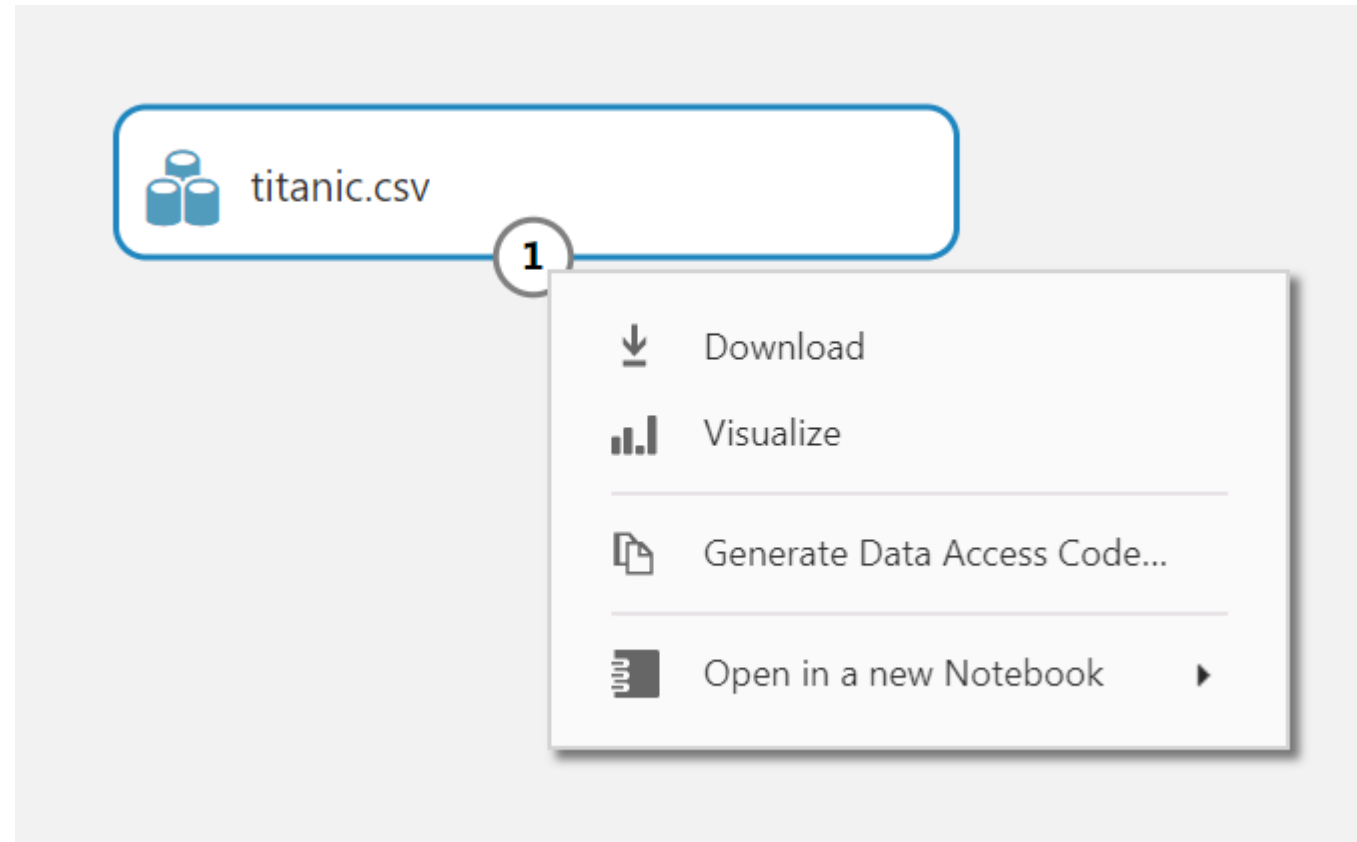
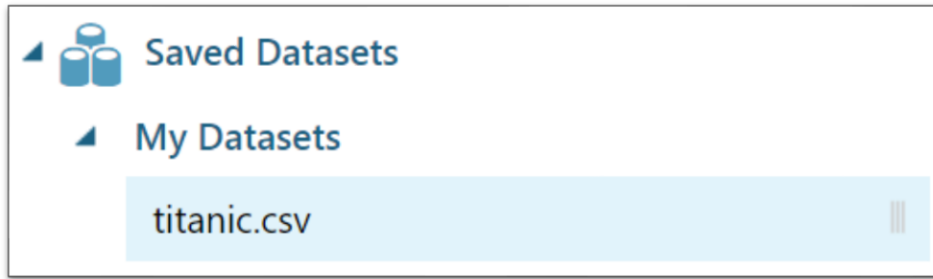
Menu items



The screenshot displays the Microsoft Azure Machine Learning Studio interface. The top bar shows the title "Microsoft Azure Machine Learning Studio" and the user "Joël Crest-Free Workspace". The interface is divided into three main sections, each highlighted with a colored border:

- Menu items (Green border):** Located on the left, it contains a search bar "Search experiment items" and a list of categories with expandable sub-items:
 - ▶ Saved Datasets
 - ▶ Trained Models
 - ▶ Data Format Conversions
 - ▶ Data Input and Output
 - ▶ Data Transformation
 - ▶ Feature Selection
 - ▶ Machine Learning
 - ▶ OpenCV Library Modules
 - ▶ Python Language Modules
 - ▶ R Language Modules
 - ▶ Statistical Functions
 - ▶ Text Analytics
 - ▶ Time Series
 - ▶ Web Service
 - ▶ Deprecated
- Experience ML (Blue border):** The central workspace, titled "Titanic", contains a diagram with dashed boxes and arrows representing a workflow. A text box at the top says "To create your experiment, drag and drop datasets and modules here". A label "Drag Items Here" with an arrow points to the first dashed box. The diagram shows a sequence of steps: a data source, followed by two transformation steps, and finally a model training step. A feedback loop arrow connects the output of the model training step back to the input of the first transformation step. The bottom of this section has a toolbar with icons for zooming and other actions.
- Properties (Orange border):** Located on the right, it contains a "Properties" tab and a "Project" tab. The "Properties" tab is active and shows:
 - Experiment Properties:** STATUS CODE: InDraft
 - Summary:** A text area with the prompt "Enter a few sentences describing your experiment (up to 140 characters)."
 - Description:** A text area with the prompt "Enter the detailed description for your experiment."
 - Quick Help:** A section with a question mark icon and a dropdown arrow.

Sélectionnez votre jeu de données



Data Transformation

Manipulation

Select Columns in Dataset



titanic.csv



Select Columns in Dataset



1

Filtrez les colonnes

Properties Project

Select Columns in Dataset

Select columns

Selected columns:

Launch the selector tool to make a selection

Launch column selector

Select columns

BY NAME

WITH RULES

AVAILABLE COLUMNS

All Types search columns



Name
Fare

2 columns available

SELECTED COLUMNS

All Types search columns

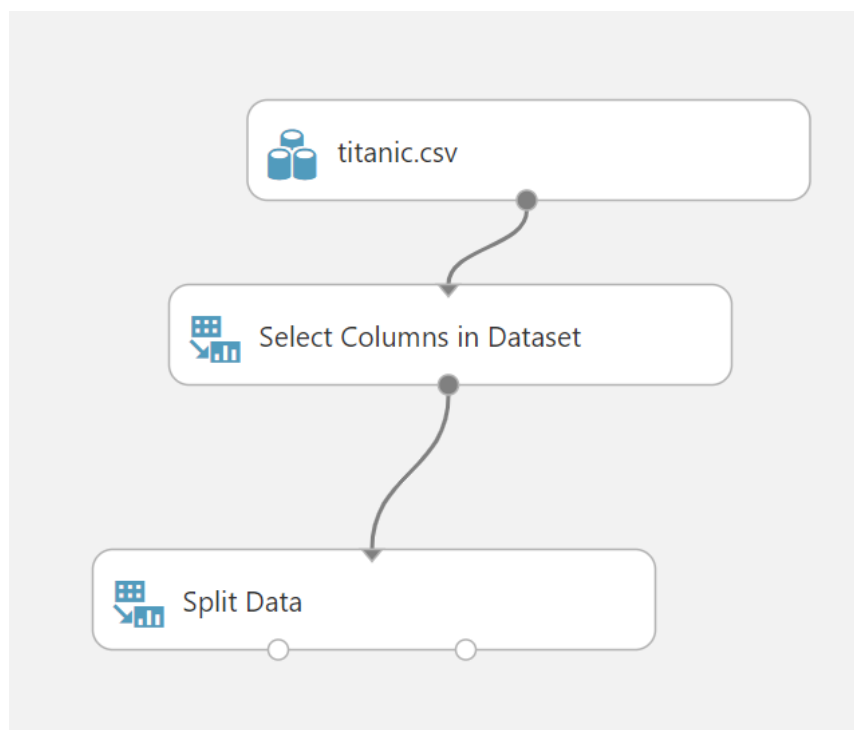
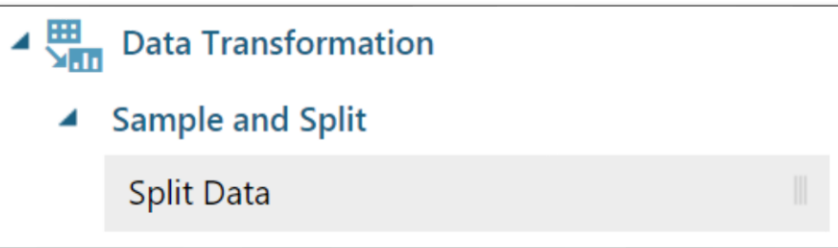


Survived
Pclass
Sex
Age
Siblings/Spouses Aboard
Parents/Children Aboard

6 columns selected



Splittez le jeu de données



Properties Project

▲ Split Data

Splitting mode

Split Rows ▼

Fraction of rows in the first... ≡

0.8

☒ Randomized split ≡


Random seed ≡

0

Stratified split

False ▼

Sélectionnez l'algorithme et entraînez le modèle

 Machine Learning

- Initialize Model
 - Classification
 - Two-Class Boosted Decision Tree

 Machine Learning

- Train
 - Train Model

Train Model

Label column

Selected columns:

Launch the selector tool to make a selection

Launch column selector

Select a single column

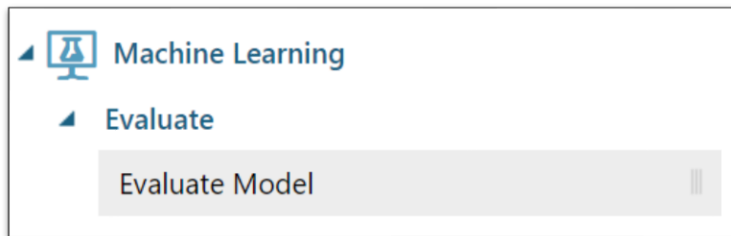
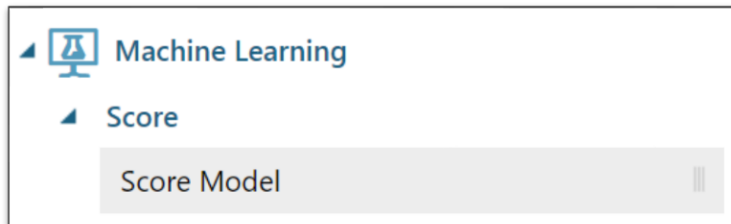
BY NAME

WITH RULES

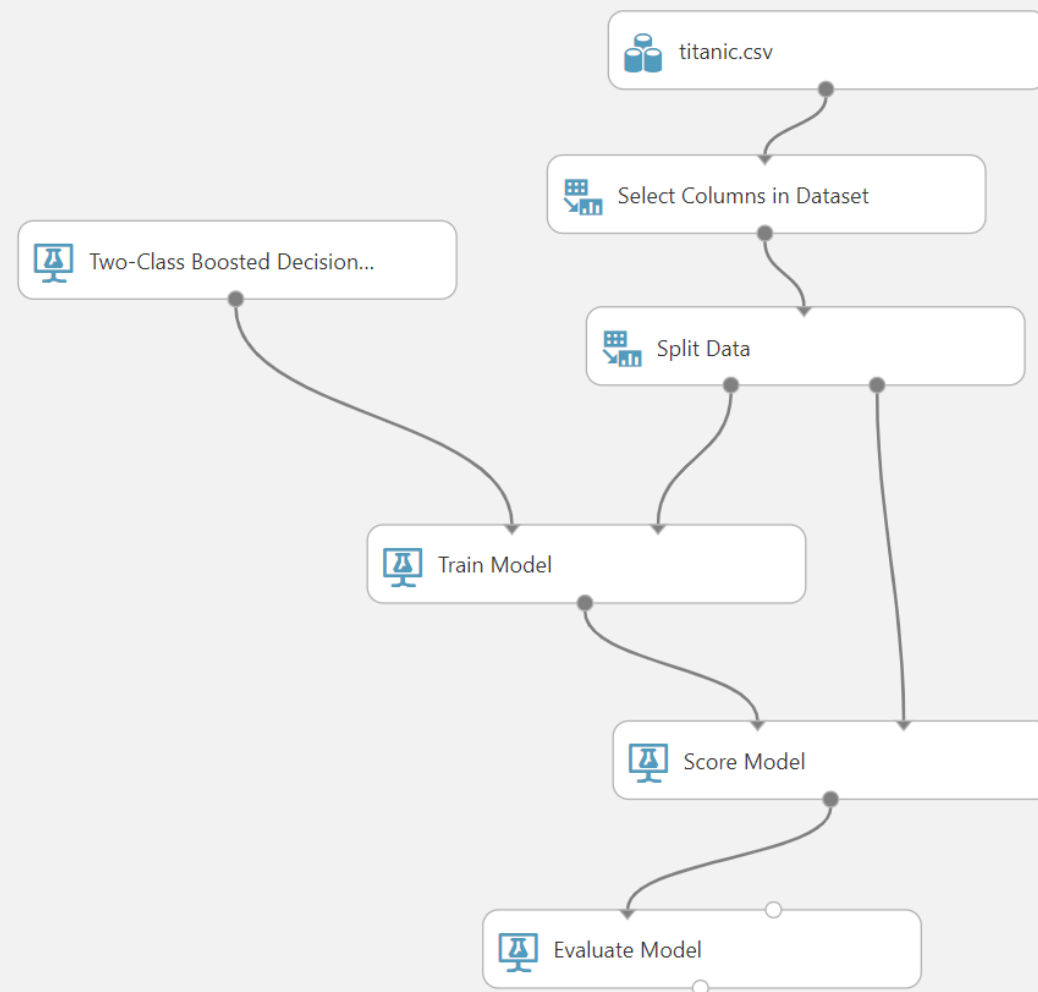
Include ▼ column names ▼

Survived ✕

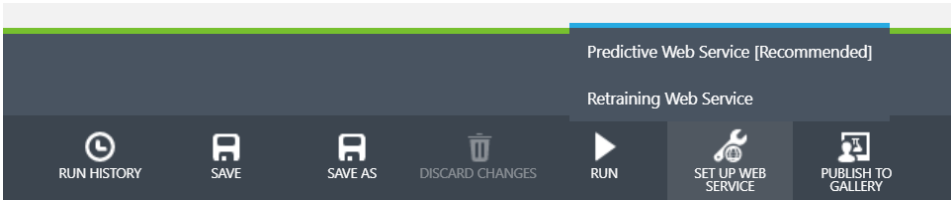
Scorez puis évaluez votre modèle



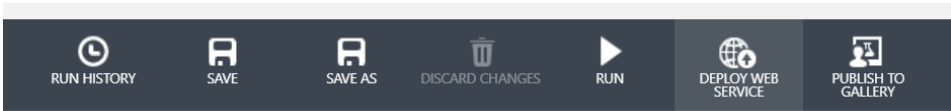
Titanic



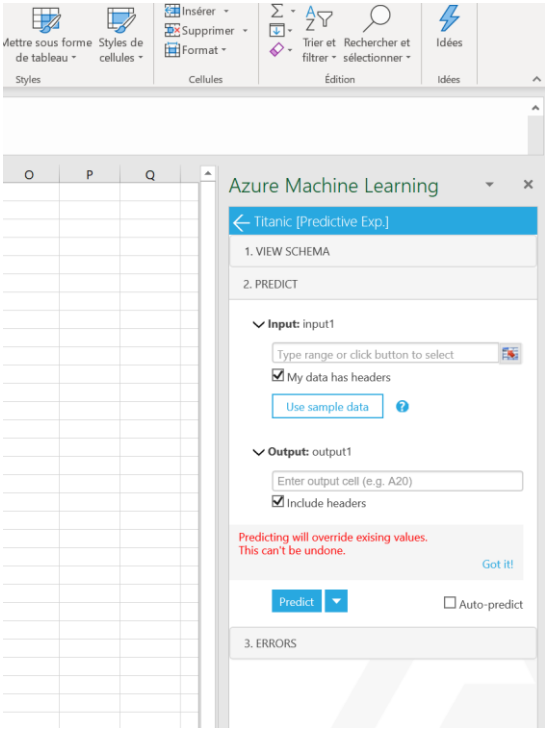
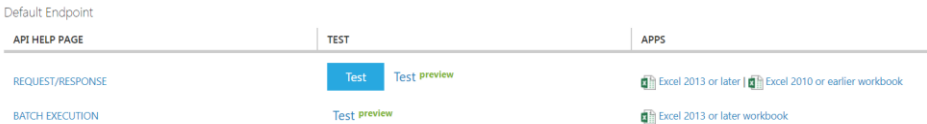
Configurez le web service



Déployez le web service



Testez le web service dans Excel



Moins de slides, plus de code !

Apportez votre propre laptop.
Nous fournissons les pass AZURE

Plusieurs
lots à
gagner



Bâtiment Centraix – 2^{ème} étage,
2 Rue le Corbusier

Inscription gratuite
Service de restauration assuré sur place

2019
Global Azure
BOOTCAMP

Une journée d'ateliers et de formation gratuite
par les experts des communautés MICROSOFT
Samedi 27 avril à Aix en Provence

9h	Accueil – café, viennoiseries
9h30	Présentation des Communautés aOS, CMD et Club Power BI
10h05	Monitorer efficacement vos ressources Azure et On Prem – Hands On Lab Jean-François BERENGUER
10h50	Pause
11h10	Infrastructure as Code (IAC) : simplifiez vous la vie (et l'IT) Rémy BOVI
12h00	Azure Advanced Threat Protection : sécurisez votre environnement Azure et On Prem Seyfallah TAGREROUT
12h45	Déjeuner
13h30	Azure Machine Learning et Titanic : plutôt #TeamJack ou #TeamRose ? Joël CREST
14h20	Le voyage d'une application en mode DevOps du pipeline d'intégration vers un cluster Kubernetes – Hands On Lab Thomas HIDROT
15h10	Pause
15h30	Stockage hybride dans Azure : étendez vos serveurs de fichiers Windows avec Azure File Sync ! Jean-François APREA
16h15	Tirage au sort et remise des lots