



# ÉCOLE MATHÉMATIQUE AFRICAINE : Bases mathématiques de l'intelligence artificielle

## Introduction à l'apprentissage artificiel en santé TP/TD

JEAN-DANIEL ZUCKER



# Machine learning software

---

## Commercial software

- SAS Enterprise Miner ([www.sas.com/technologies/analytics/datamining/miner](http://www.sas.com/technologies/analytics/datamining/miner))
- IBM SPSS Statistics Software ([www-01.ibm.com/software/analytics/spss/products/statistics](http://www-01.ibm.com/software/analytics/spss/products/statistics))
- Salford Systems Data Mining Suite ([www.salford-systems.com](http://www.salford-systems.com))
- Statistica Data Miner ([www.statsoft.com/products/data-mining-solutions](http://www.statsoft.com/products/data-mining-solutions))

## Free Software

- WEKA ([www.cs.waikato.ac.nz/ml/weka](http://www.cs.waikato.ac.nz/ml/weka))
- RapidMiner ([rapid-i.com](http://rapid-i.com))
- Orange ([orange.biolab.si](http://orange.biolab.si))
- Machine Learning & Statistical Learning in R language ([cran.r-project.org/web/views/MachineLearning.html](http://cran.r-project.org/web/views/MachineLearning.html))

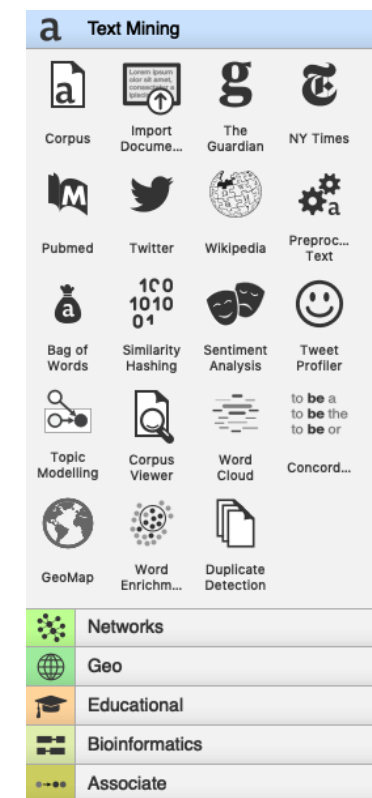
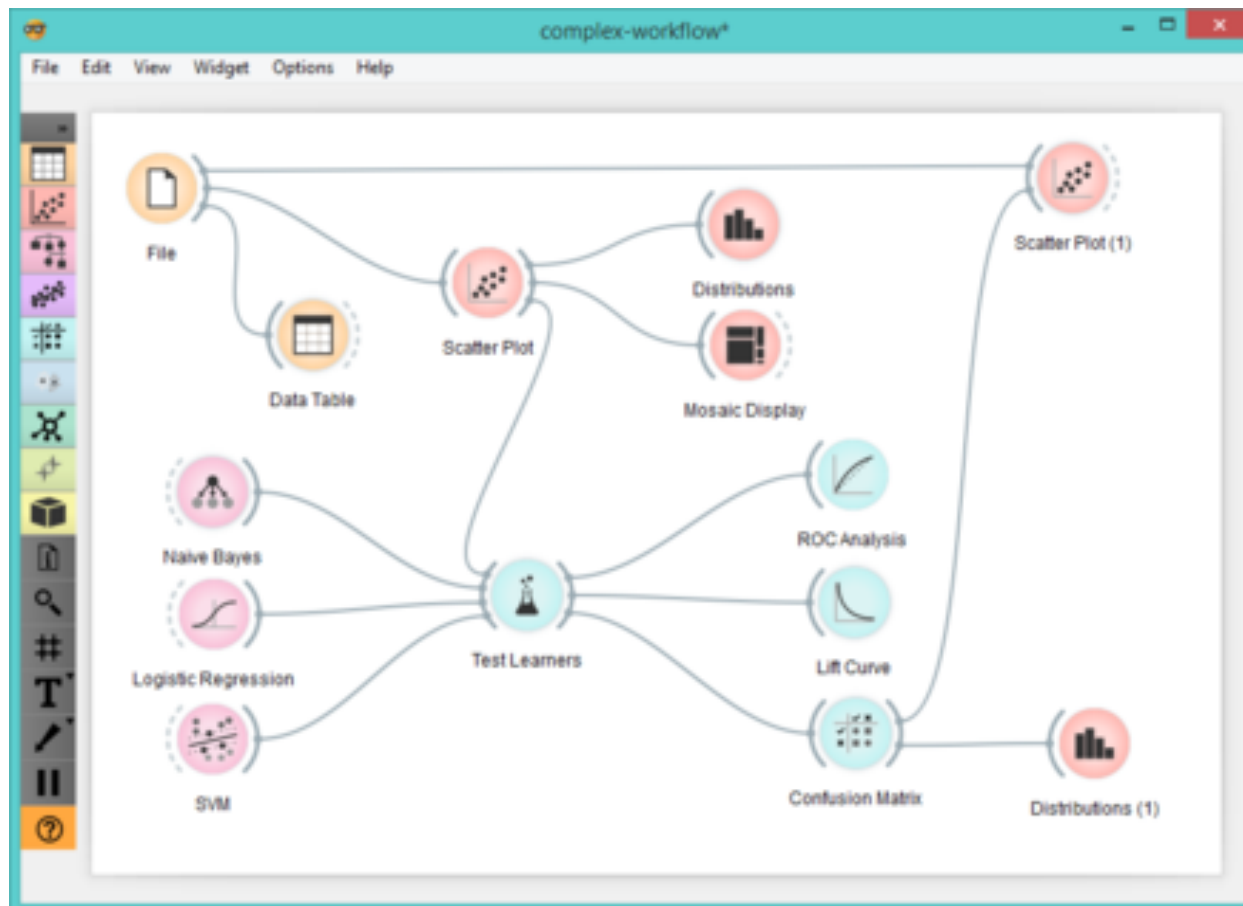
And many others...

# Introduction to Orange3



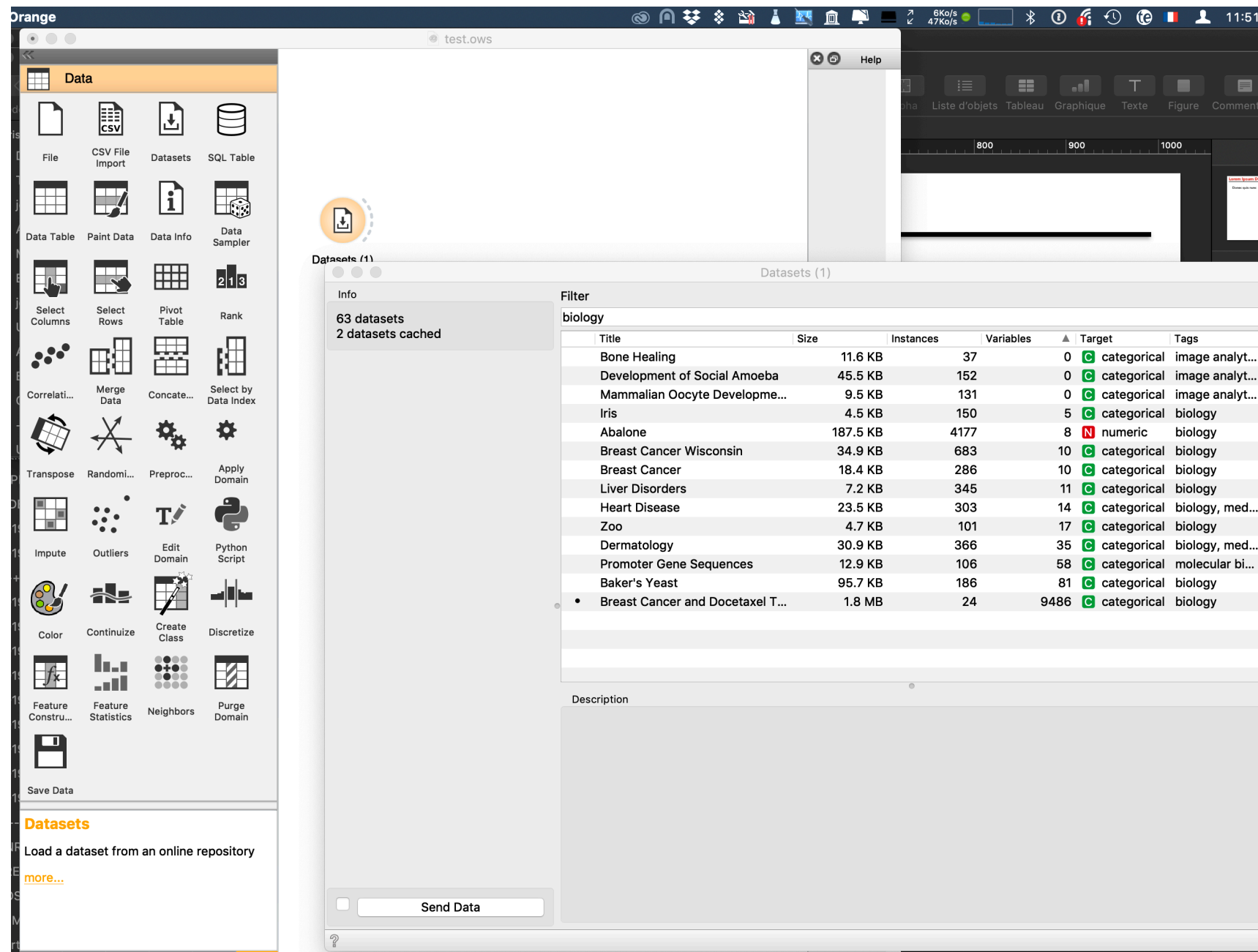
[https://en.wikipedia.org/wiki/Orange\\_\(software\)](https://en.wikipedia.org/wiki/Orange_(software))

**Orange** est une boîte à outils de visualisation de données, d'apprentissage machine et de data mining open-source. Il dispose d'une interface de programmation visuelle pour l'analyse explorative des données et la visualisation interactive des données, et peut également être utilisé comme une bibliothèque Python.



# L'environnement Orange3

## TP1: familiarisation avec Orange 3 et exploration des datasets médicaux



The screenshot displays the Orange3 data mining environment. On the left is the 'Data' widget panel with various icons for loading and processing data. The main window shows the 'Datasets (1)' widget, which lists 63 datasets. The 'biology' filter is selected, showing a table of datasets with columns for Title, Size, Instances, Variables, Target, and Tags. The 'Send Data' button is visible at the bottom of the widget panel.

Title	Size	Instances	Variables	Target	Tags
Bone Healing	11.6 KB	37	0	C categorical	image analyt...
Development of Social Amoeba	45.5 KB	152	0	C categorical	image analyt...
Mammalian Oocyte Developme...	9.5 KB	131	0	C categorical	image analyt...
Iris	4.5 KB	150	5	C categorical	biology
Abalone	187.5 KB	4177	8	N numeric	biology
Breast Cancer Wisconsin	34.9 KB	683	10	C categorical	biology
Breast Cancer	18.4 KB	286	10	C categorical	biology
Liver Disorders	7.2 KB	345	11	C categorical	biology
Heart Disease	23.5 KB	303	14	C categorical	biology, med...
Zoo	4.7 KB	101	17	C categorical	biology
Dermatology	30.9 KB	366	35	C categorical	biology, med...
Promoter Gene Sequences	12.9 KB	106	58	C categorical	molecular bi...
Baker's Yeast	95.7 KB	186	81	C categorical	biology
Breast Cancer and Docetaxel T...	1.8 MB	24	9486	C categorical	biology

# Overfitting and regularization orange3

---

## TP2: mini tutoriel sur la régularisation

<https://blog.biolab.si/tag/regularization/>

nécessite l'addson Education d'Orange3

## TP3: faire le tutoriel

[https://docs.biolab.si/2/\\_downloads/bio-tutorial.pdf](https://docs.biolab.si/2/_downloads/bio-tutorial.pdf)

# Autres liens

---

## Excellents cours Machine Learning

[http://eric.univ-lyon2.fr/~ricco/cours/supports\\_data\\_mining.html](http://eric.univ-lyon2.fr/~ricco/cours/supports_data_mining.html)

## Excellent cours Machine Learning en santé

<https://mlhc19mit.github.io/>

## Dalex un package pour l'interpretabilité des modèles

<https://pbiecek.github.io/DALEX/>

[https://rawgit.com/pbiecek/DALEX\\_docs/master/vignettes/DALEX\\_and\\_keras.html](https://rawgit.com/pbiecek/DALEX_docs/master/vignettes/DALEX_and_keras.html)