Data challenge

Bankable customers

Table des matières

[Etude des données 2](#_Toc30613446)

[Données déterministes 2](#_Toc30613447)

[Ethique 2](#_Toc30613448)

[Filtrage 2](#_Toc30613449)

[Partitionning 2](#_Toc30613450)

[Choix de conception 2](#_Toc30613451)

[Un arbre de décision 2](#_Toc30613452)

[Logistique 2](#_Toc30613453)

[Optimisation 2](#_Toc30613454)

[Le nœud « Parameter Optimization Loop Start » 2](#_Toc30613455)

[Le nœud « Parameter Optimization Loop End » 2](#_Toc30613456)

[Tests 3](#_Toc30613457)

[Tests on KNN classifier 3](#_Toc30613458)

[Schema 3](#_Toc30613459)

[Tests on decisionnal tree 3](#_Toc30613460)

[Tree without transformation 3](#_Toc30613461)

# Etude des données

## Données déterministes

## Ethique

## Filtrage

## Partitionning

# Choix de conception

## Un arbre de décision

## Logistique

(tout seul dans le groupe)

## Optimisation

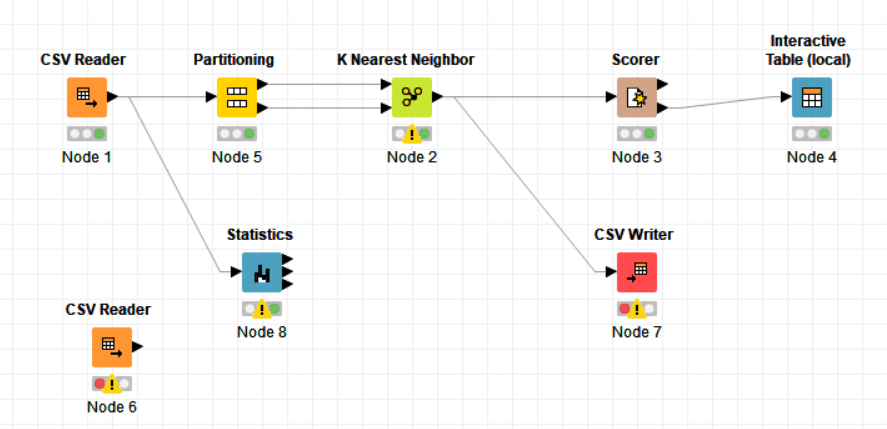
### Le nœud « Parameter Optimization Loop Start »

### Le nœud « Parameter Optimization Loop End »

# Tests

## Tests on KNN classifier

### Schema

**** Test 1

**Classifier :** K Nearest Neighbor

**Partitionning :** 70% for training and 30% for testing

**Params :** 2 neighbors precision

**Column predicted :** income

**Accuracy :** 0.785

**Coherence :** 0.264

## Tests on decisionnal tree

### Tree without transformation

Test 1

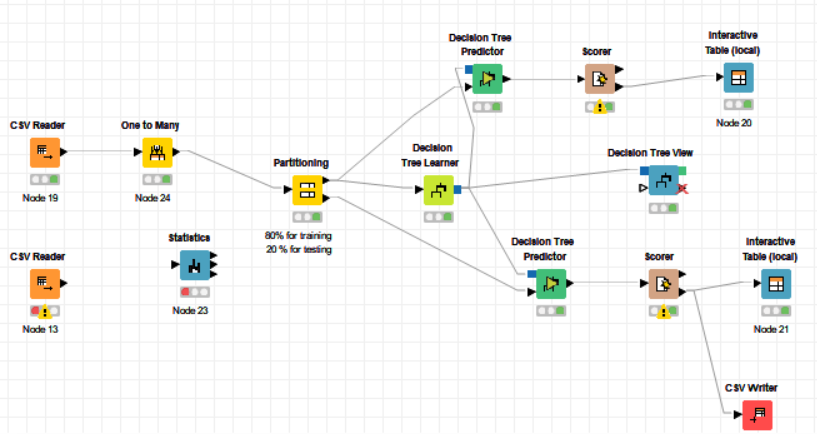
**Classifier :** Decision Tree

**Partitionning :** 80% for training and 20% for testing

**Column predicted :** income

**Accuracy :** 0.857

**Coherence :** 0.588

 Test 1

**Classifier :** Decision Tree

**Partitionning :** 80% for training and 20% for testing

**Transforming :** One to many on « sex » column

**Column predicted :** income

**Accuracy :** 0.857

**Coherence :** 0.588