

# Árvores de Decisão @ KNIME

## Aprendizagem e Decisão Inteligentes

Licenciatura em Engenharia Informática/3° ano - 2° semestre Mestrado integrado em Engenharia Informática/4° ano - 2° semestre Cesar Analide, Inês Amorim, Pedro Oliveira



Departamento de Informática

- O fluxo Learner-Predictor em KNIME
- Árvores de Decisão
  - Learner
  - Predictor
  - Scorer
- Experimentação (hands on)

## **Agenda**





**A Machine Learning Pipeline** Universidade do Minho Departamento de Informática 2. Data 3. Data 4. Data 5. Model 6. Candidate Model 8. Performance Problem 7. Model Definition Ingestion Preparation Segregation Training Evaluation Monitoring eployment TEST SET TRAINING SET VALIDATION SET ITERATION

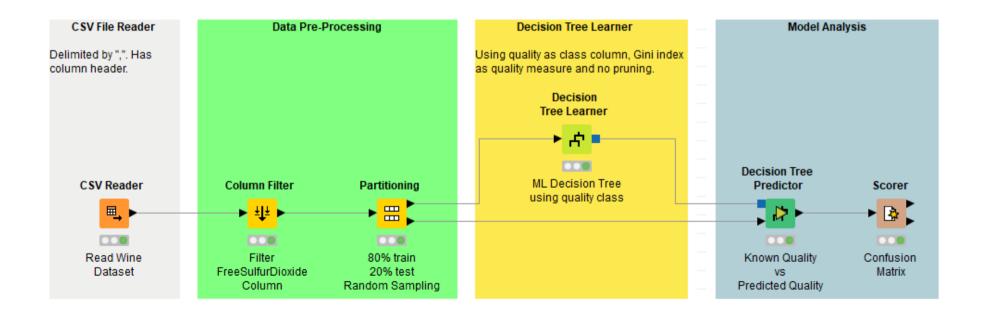
(https://towardsdatascience.com/architecting-a-machine-learning-pipeline-a847f094d1c7)



Departamento de Informática

### Fluxo Learner-Predictor @ KNIME

Em KINME, a criação de modelos de previsão pela aplicação de técnicas de aprendizagem (machine learning), é implementada por nodos «Learner» e «Predictor»:

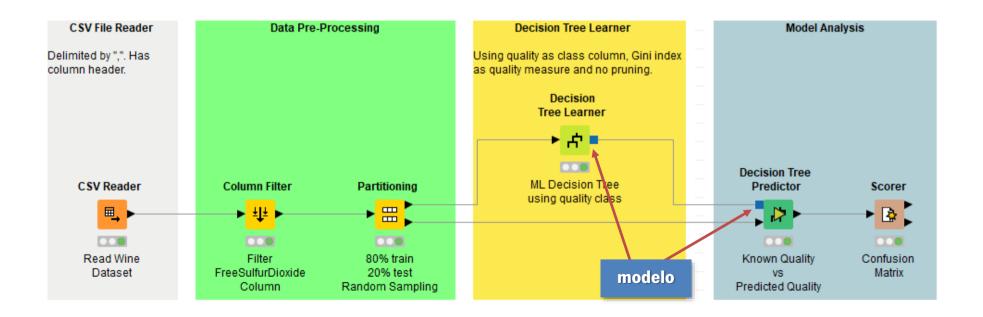




Departamento de Informática

### Fluxo Learner-Predictor @ KNIME

Em KINME, a criação de modelos de previsão pela aplicação de técnicas de aprendizagem (machine learning), é implementada por nodos «Learner» e «Predictor»:

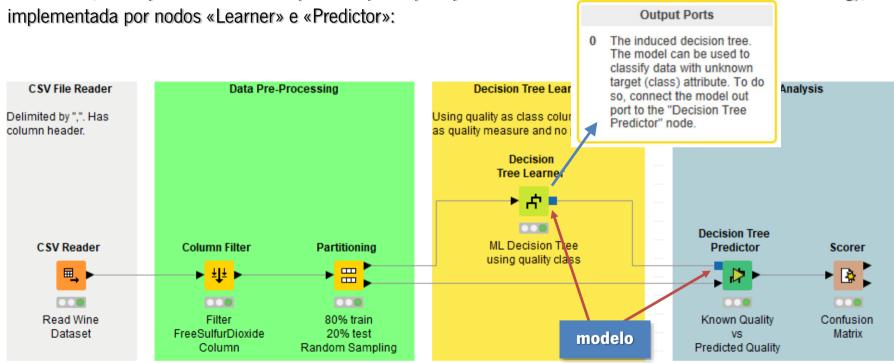




Departamento de Informática

## Fluxo Learner-Predictor @ KNIME

Em KINME, a criação de modelos de previsão pela aplicação de técnicas de anrendizagem (machine learning), é

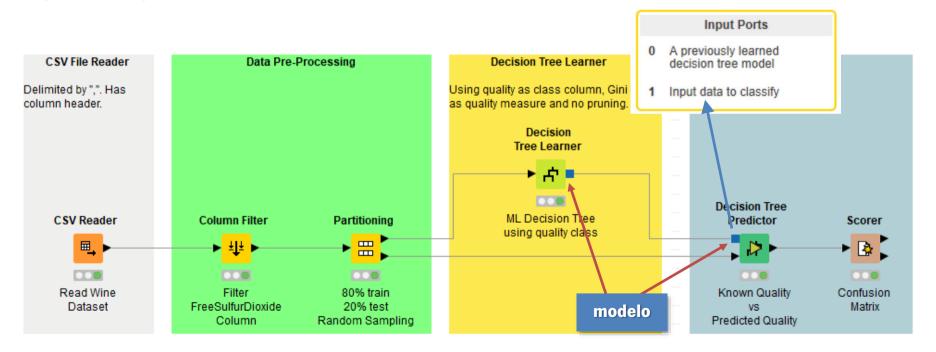




Departamento de Informática

## Fluxo Learner-Predictor @ KNIME

Em KINME, a criação de modelos de previsão pela aplicação de técnicas de aprendizagem (machine learning), é
implementada por nodos «Learner» e «Predictor»:

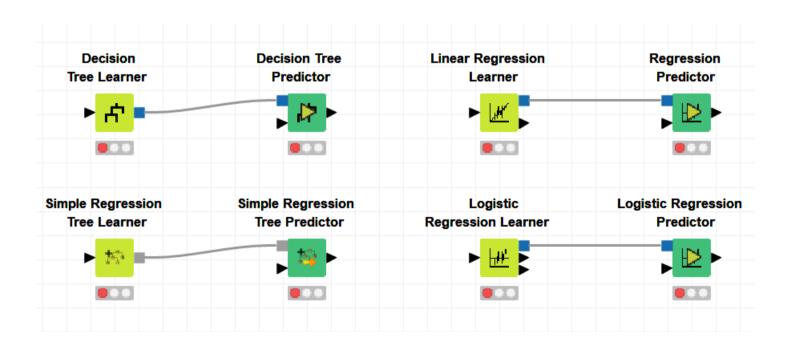




Departamento de Informática

# Nodos Learner-Predictor @ KNIME (exemplos)

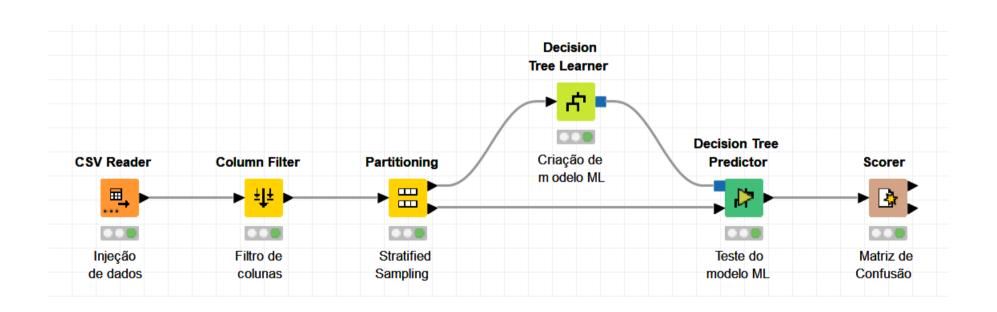
• Alguns exemplos de nodos «Learner» e «Predictor», para várias técnicas de aprendizagem:





Departamento de Informática

## Árvores de Decisão





#### Árvores de Decisão ▲ Dialog - 6:4 - Decision Tree Learner (Criação de) Universi Departan Options PMMLSettings Flow Variables Job Manager Selection General Class column S class Quality measure Gairmand Pruning method No pruning ~ Reduced Error Pruning Decision Min number records per node 2 💠 Tree Learner 10 000 🖨 Number records to store for view Average split point 8 🜲 Number threads **Decision Tree** Skip nominal columns without domain information Criação de CS Predictor Scorer -Root split m odelo ML Force root split column Root split column D petal\_width ~ Binary nominal splits Binary nominal splits Teste do Matriz de modelo ML Confusão d Max #nominal 10 💠 Filter invalid attribute values in child nodes

?

Cancel

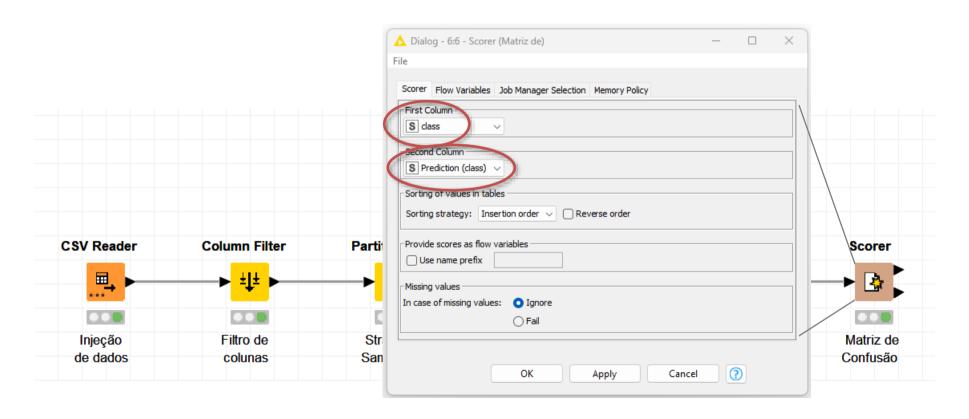
OK

Apply



Departamento de Informática

## Árvores de Decisão







Departamento de Informática

