

Cvičení 6 – Základy práce v prostředí KNIME

- KNIME (Konstanz Information Miner) www.knime.com
- open-source platforma pro analýzu dat

Okna

- **nejdůležitější** – KNIME Explorer, Node repository, Editor

Programy

- jednotlivé úlohy – uzly ([nodes](#))
- propojené uzly – program ([workflow](#))

Data - soubor dataKoupimEV.csv

c_t – koupím si EV? (0 - ano, 1 - nevím, 2 - ne)

$y_{1;t}$ – doba cesty do práce (minuty)

$y_{2;t}$ – náklady na cestu (Kč/km)

$y_{3;t}$ – příjem (Kč/měsíc)

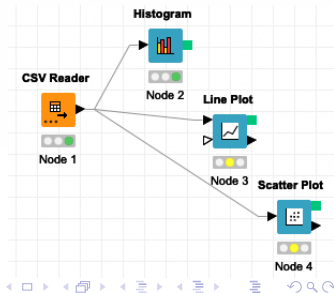
$y_{4;t}$ – věk (roky)

$y_{5;t}$ – dostupnost nabíjecí stanice (1 - ano, 0 - ne)

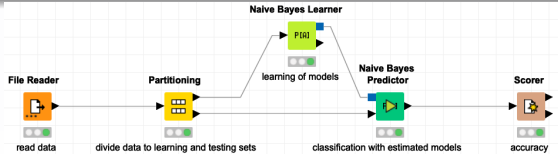
Příklad 1: Jak to funguje?

- ❶ nový workflow – File Reader, Configure, Execute, Results
- ❷ histogram – Node repository/Views nebo napsat “histogram”
 - Configure – sloupec, vygenerovat obrázek, Execute
- ❸ časový graf – Node repository/Views nebo napsat “line plot”
 - Configure – sloupec pro osy, vygenerovat obrázek, Execute
- ❹ x,y graf – Node repository/Views nebo napsat “scatter plot”
 - Configure – sloupec pro osy, vygenerovat obrázek, Execute

- pro nové spuštění uzlu – Reset, Execute
- pro práci doma – Export KNIME Workflow



Samostatná práce - úkol 1: Naivní Bayesův klasifikátor



1. Použijte metodu Naivní Bayes pro klasifikaci dat ([dataKoupimEV.csv](#))
2. Použijte metodu Naivní Bayes pro klasifikaci dat z přednášky ([dataDopravProstredek.csv](#))

File/CSV reader:

- cílová proměnná (Koupě EV) – “string”
- ikonka File/CSV reader, kliknout na sloupec, zvolit “string”
- nebo: Configure/Transformation, “string”
- oddělovač , nebo ; (tabulka s daty)

Partitioning:

relative 80%, draw randomly

Naive Bayes Learner:

zvolit classification column

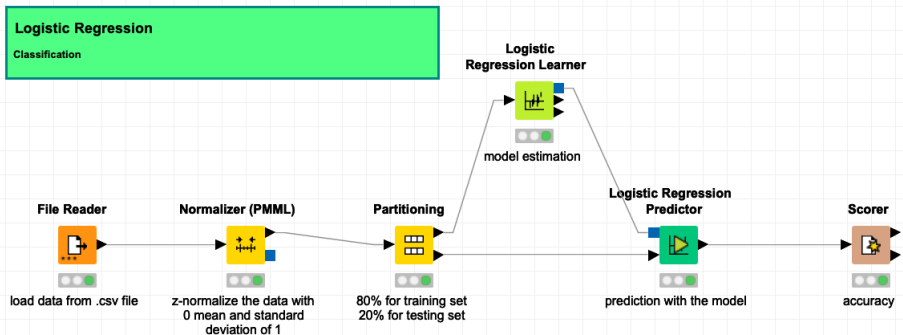
Prediktor:

přejmenovat predikci

Scorer:

zvolit první a druhý sloupec
pro porovnání (Koupě EV a
predikce)

Samostatná práce - úkol 2: Logistická regrese



1. Použijte logistickou regresi pro klasifikaci dat ([dataKoupimEV.csv](#))
2. Použijte logistickou regresi pro klasifikaci dat z přednášky ([dataDopravProstredek.csv](#))
3. Porovnejte přesnost klasifikace použitých metod