Homework #3: Getting Started with Machine Learning Tool (100 points)

In this assignment, your task is to *download*, *install*, and *experiment* with a machine learning / data mining tool of your choice. Examples tools include:

- KNIME[1]: https://www.knime.org/
- WEKA: http://www.cs.waikato.ac.nz/ml/weka/
- Orange: http://orange.biolab.si/
- RapidMiner: https://rapidminer.com/

With regard to experimenting with your chosen tool, you have **two** deliverables:

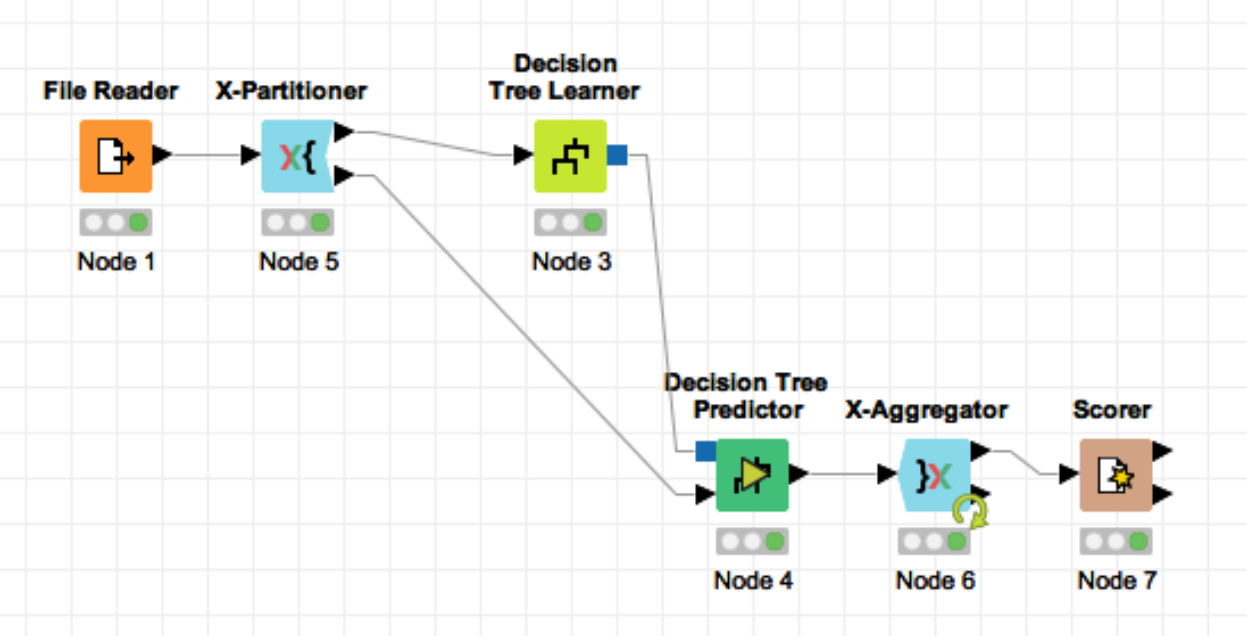
1. Replicate the general machine learning workflow (on page 2) that uses a *decision tree* classifier to classify the canonical "Iris" dataset. When you are finished, take a **screen shot** of your workflow and **document** the classification results (e.g., accuracy, precision, recall, etc.,).

2. Create a pattern recognition workflow to experiment with the "Census Income" dataset from homework #1. Take a **screen shot** of your workflow and **document** the classification results.

    a. What type of classifier did you use?
    b. What features did you use? Why?
    c. What accuracy did you reach? recall? precision?

To clarify, you must submit a single document showing the completed deliverables.

[1] I am most proficient with KNIME; I used it throughout my PhD. KNIME also has a WEKA extension.


File Reader
X-Partitioner
Decision Tree Learner
Node 1
Node 5
Node 3
Decision Tree Predictor
X-Aggregator
Scorer
Node 4
Node 6
Node 7