



ANÁLISIS DE DATOS

Ricardo Aler Mur

SELFASSESSMENT OF DECISION TREES WITH QUESTIONS AND ANSWERS

- 1) What does the entropy measure?**
- 2) What is more likely to overfit, a tree with with many nodes or a tree with few nodes?**
- 3) Why the entropy (or gain information or information gain) does not work well when an attribute has many values?**
- 4) What is a simple way to learn rules from trees?**
- 5) What types of trees for regression do exist?**
- 6) How is it possible to build regression models with categorical variables?**