

Regression and Model Trees. Prof. Ruiz, Dept. of Computer Science, WPI.

1. Consider the following dataset:

- a. Calculate the Standard Deviation Reduction (SDR) of Humidity using the split-point 72.5 with respect to the Target attribute.

Humidity	Target
65.0	64.0
70.0	65.0
70.0	69.0
70.0	75.0
75.0	81.0
80.0	68.0
80.0	75.0
85.0	85.0
86.0	83.0
90.0	80.0
90.0	72.0
91.0	71.0
95.0	72.0
96.0	70.0

- b. Calculate the Standard Deviation Reduction (SDR) of Humidity using the split-point 88 with respect to the Target attribute.

2. Consider now the following dataset:

- a. Convert Outlook into a set of binary attributes, as Weka's M5P algorithm would do it.
- b. Calculate the Standard Deviation Reduction (SDR) of one of the resulting binary attributes with respect to the Target.

Outlook	Target
sunny	85.0
sunny	80.0
overcast	83.0
rainy	70.0
rainy	68.0
rainy	65.0
overcast	64.0
sunny	72.0
sunny	69.0
rainy	75.0
sunny	75.0
overcast	72.0
overcast	81.0
rainy	71.0