ULB

UNIVERSITÉ LIBRE DE BRUXELLES

Implementation of Decision Tree Classifiers
ID3 versus C4.5

Depuydt Antoine
Dansy Efila
Mudura Mircea

May, 2017

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

Introduction

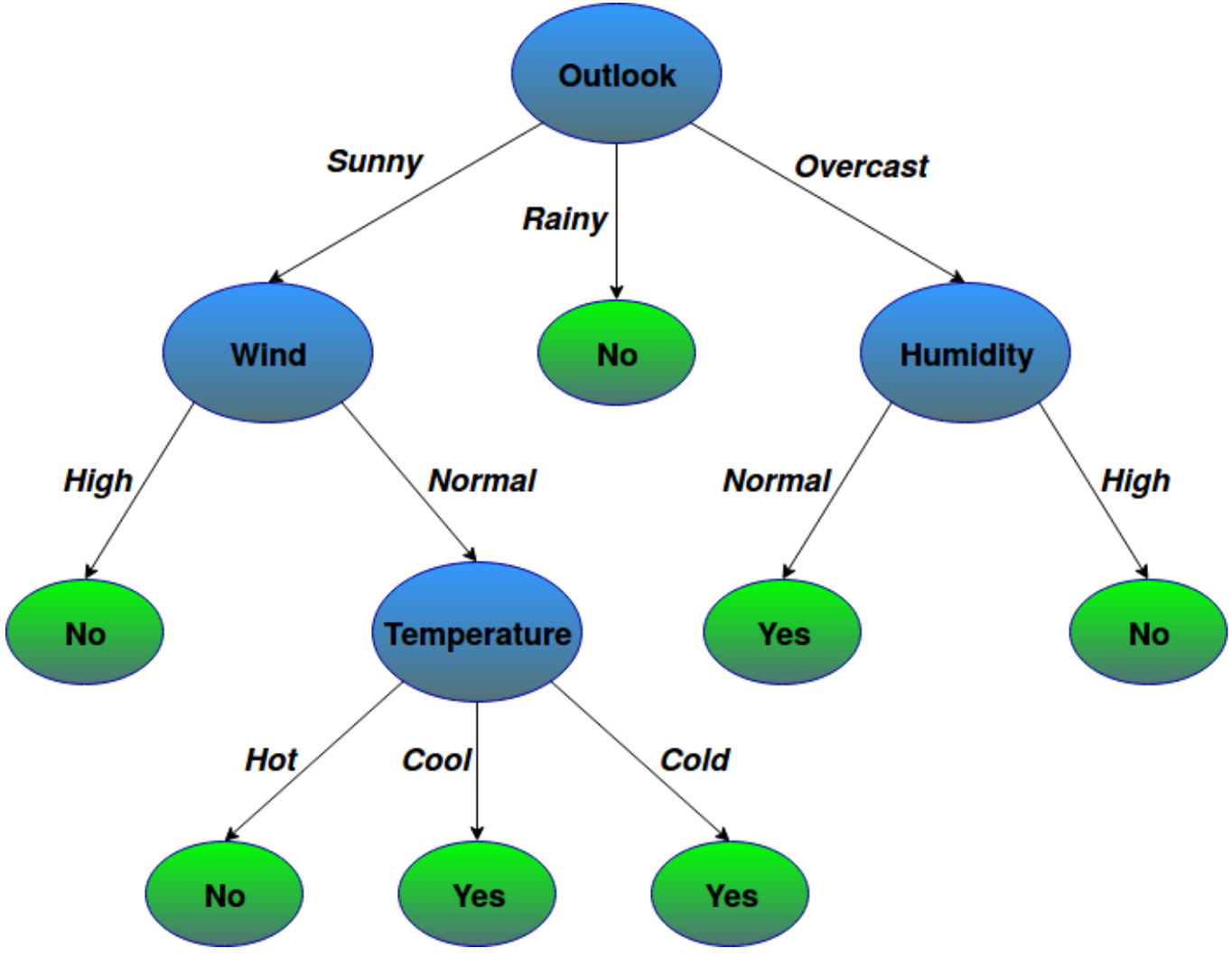
- Data mining: compress, understand and predict
  - Clustering
  - Classification
  - Regression
  - ...
- Techniques to find links
  - Linear Regression
  - Decision Trees
  - Neural Networks
  - ...

Classification

- Well known example: is it a good day to play tennis?
  - Features:
    - Outlook: sunny, overcast, rainy
    - Temperature: hot, cool, cold
    - Wind: high, weak
    - Humidity: high, normal
  - Labels:
    - Yes
    - No

Decision Tree

- Model is easy to understand for users (visual)
- Final result is a tree with decision nodes and leaf nodes



ID3 versus C4.5

- Goal: construct the decision tree using ID3 and C4.5 algorithms
- Objectives: compare ID3 and C4.5 and creating
  - Compare ID3 and C4.5
  - Create an application that classifies any data using both algorithms

ID3

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

C4.5

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

K-fold cross validation

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

Demonstration