

決定木 その1





決定木とは

慶應義塾大学理工学部
櫻井彰人

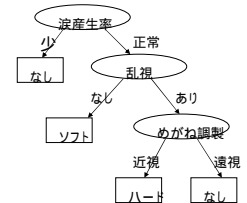
木とは？

- 根がある
- 根から枝分かれしていく
 - 枝の先では、くっつかない
- 枝の先端は葉である(現実ではない)
- どうでもいいことですが、「連理の枝」って知っていますか？



木とは

-  であって
 - どの節(点)も繋がっている
-  がないもの
- 根から枝分かれしたとき
 - 連理がある  がある
 - 連理がない  がない

決定木



どんなものか

- 木
- 木の節(ノード)に 
- 木の枝(エッジ)に 
- ただし、葉(これも節の一つ)には、ある特別な「属性値」

どう使うか

- 「属性 = 属性値」の組(決定表の一行)に対し、
- 根からスタートして、
- 自分の属性値に従って、枝をたどり、葉に至ると
- 葉には行うべき行動とか属するクラスの名称が書かれている

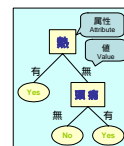
属性と属性値

- 属性: 「表」で言えば、縦の欄
 - うるさく言えば、確率変数
 - 人間でいえば、身長、体重、年齢、生年月日、...
- 属性の名: 「表」で言えば、縦の欄の名称
 - 確率変数の名称
 - 「身長」、「体重」、...
- 属性値: 「表」に入っている値
 - 確率変数がとる値
 - 身長であれば、180cm というような連続数値や、「高い」「普通」「低い」と離散値(カテゴリ値)

宿題

- 決定木に対応する決定表を作ってください。

どういう意味だと思いますか？



これなら分りやすいか？

