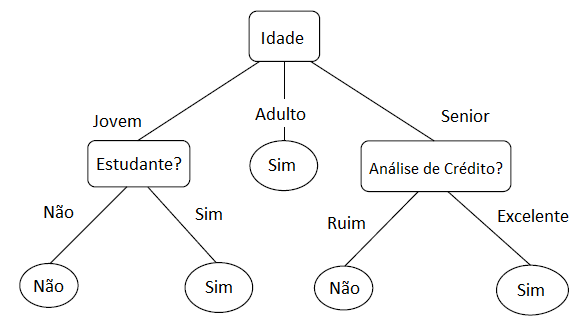
Random Forests é uma melhoria aplicada nos algoritmos de árvore de decisão. Uma aplicação é no kinect do xbox que faz uso de algoritmos de Random forests.

No treinamento da árvore de decisão que é feito vários cálculos matemáticos e definido qual é característica mais importante (aquela que será nodo inicial/raiz), qual sequência entre as características, qual características não são importantes. Em resumo é feito a elaboração da aŕvore de decisão, que nada é mais é que um conjunto de regras.

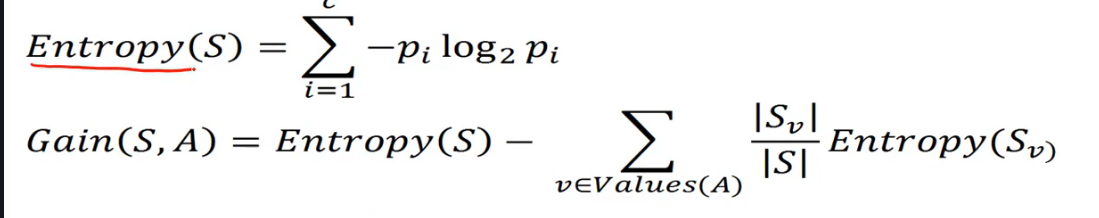
Exemplo:



Regras:

* SE idade é jovem E é estudante vai permitir criar conta.
* SE idade é adulto vai permitir criar conta
* SE idade é Senior E análise de crédito é ruim não vai permitir criar conta.

as duas formulas abaixo são usadas para calcular qual é a caracteristica mais importante.



Na pasta desse módulo tem uma planilha que faz esse cálculo.

overfitting é quando o algoritmo se adapta demais aos dados de um base, logo quando é usado esse algoritmo em outra base pode ter classificações incorretas em grande volume.

A ideia de overfitting seria quando o algoritmo ao inves de aprender ele “decora”.

Para árvore de decisão faz escalonamento geralmente não muda em nada (não aumenta e nem diminui acuracia)

Quanto mais árvore em uma Random Forest maior chance de acontecer um overfitting