

Inteligência Artificial

2º Projeto

Grupo 71

Bernardo Rosa - 88077

João Santos Jorge - 88079

Introdução:

Neste relatório pretendemos analisar a solução utilizada para a realização do segundo projeto da unidade curricular de Inteligência Artificial do primeiro semestre do ano letivo de 2020/2021.

O problema apresentado “decision tree learning” tem por objetivo retornar uma árvore de decisão a partir de um conjunto de exemplos e respetivas classificações.

As árvores de decisão têm grandes vantagens como:

- Facilidade de interpretação e visualização
- Custo de utilização da árvore é logarítmico de acordo com o número de dados
- Capaz de utilizar tanto tipos numéricos como outros tipos de dados (no projeto foi utilizado apenas valores “booleanos”)
- (...)

Descrição da solução do projecto:

A nossa implementação da função “create_decision_tree” foi baseada na abordagem do livro da disciplina, representada na imagem abaixo.

```
function DECISION-TREE-LEARNING(examples, attributes, parent_examples) returns  
a tree  
  
  if examples is empty then return PLURALITY-VALUE(parent_examples)  
  else if all examples have the same classification then return the classification  
  else if attributes is empty then return PLURALITY-VALUE(examples)  
  else  
     $A \leftarrow \operatorname{argmax}_{a \in \text{attributes}} \text{IMPORTANCE}(a, \text{examples})$   
    tree  $\leftarrow$  a new decision tree with root test A  
    for each value  $v_k$  of A do  
       $\text{exs} \leftarrow \{e : e \in \text{examples} \text{ and } e.A = v_k\}$   
      subtree  $\leftarrow$  DECISION-TREE-LEARNING(exs, attributes – A, examples)  
      add a branch to tree with label ( $A = v_k$ ) and subtree subtree  
  return tree
```

Solução mais pequena:

Para encontrar a solução mais pequena foi utilizado no cálculo da entropia a fórmula dada nas aulas, apresentada abaixo.

$$GI(A) = I\left(\frac{p}{n+p}, \frac{n}{n+p}\right) - \sum_i \frac{p_i + n_i}{p+n} I\left(\frac{p_{A=i}}{n_{A=i} + p_{A=i}}, \frac{n_{A=i}}{n_{A=i} + p_{A=i}}\right)$$

E para casos iguais de entropia foi escolhido aleatoriamente um desses valores. Daí, os resultados dos testes variarem de execução para execução, podendo assim obter-se árvores de maior ou menor comprimento.

```
tree no noise 19 /26 short 1 /2
tree noise 3 /4 good in test 0 /4
points 5.8 + reduced tree 1.5 + noise 0.0

Expected grade: 7.3 (/14)+hidden tests (4pt)+report (2pt)
joao@joao-VirtualBox:~/Documents/tecnico/3ano/IA/projecto2/IA2021proj2alunos$
```

```
tree no noise 21 /26 short 1 /2
tree noise 3 /4 good in test 0 /4
points 6.5 + reduced tree 1.5 + noise 0.0

Expected grade: 8.0 (/14)+hidden tests (4pt)+report (2pt)
joao@joao-VirtualBox:~/Documents/tecnico/3ano/IA/projecto2/IA2021proj2alunos$
```

Bibliografia:

“Artificial Intelligence: A Modern Approach”, Russel Stuart and Norvig Peter

<https://scikit-learn.org/stable/modules/tree.html#tree>