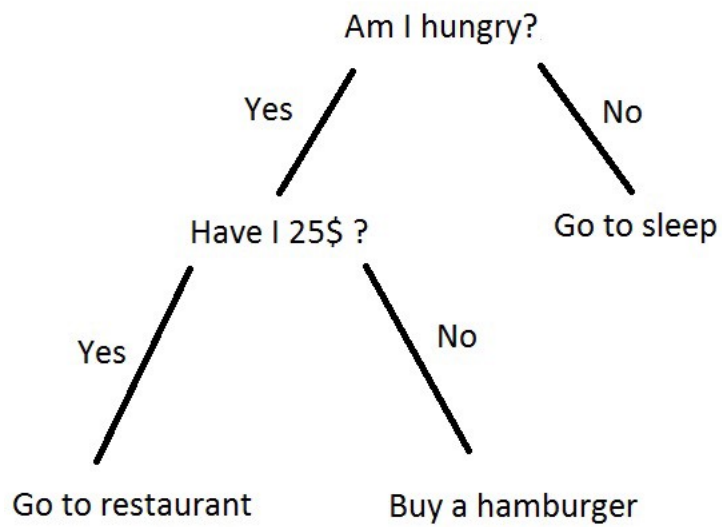


# Trabalho Aprendizagem Automática

João Marques nº39996  
Pedro Claudino nº39870



# 1 Introdução

Na cadeira de Aprendizagem automática foi pedido a realização de uma árvore de decisão com o algoritmo **ID3**. As funções de impureza deste algoritmo serão **entropia**, **gini** e **erro**.

# 2 Desenvolvimento do código

No código encontra-se a *Class* **Node** que tem como intuito criar nodes. Nesta *Class* encontra-se também as funções de adicionar *child nodes* e de obter os *child nodes* de um node.

Na *Class* **DecisionTreeREPrune** encontram-se as funções de criação dos *nodes* (**create\_node**), a função de calculo de entropia(**entropy**), a função **fit** que gera uma árvore baseada em 75 por cento dos fornecidos, a função **info\_gain** que realiza um calculo entre entropias para futuramente ser possível escolher o maior valor de *gain*, a função **count\_unique** que devolve informação sobre um atributo, ou seja, quantas repetições de um certo valor se encontram naquele atributo, a função **buildtree** que é utilizada pela função **fit** na geração da árvore, a função **bestatt** que devolve o *index* do atributo com maior *info gaine* por fim as funções de calculo da exatidão como por exemplo a função **score** que calcula a exatidão com base nos dados de teste (25 por cento) fornecidos.

### 3 Conclusão

A maior atrocidade com que nos deparamos durante esta jornada foi o entendimento sobre como implementar o **fit**. O que causou este problema foi criar as ligações entre os *child nodes* e o *node*. Apesar de ter-mos praticado o *python3* antes da iniciação do trabalho, também encontramos dificuldades na escrita de algumas partes do código o que atrasou muito o seu desenvolvimento. Por outro lado, concluimos que o nosso entendimento sobre árvore de decisão e o nosso nível de conhecimento de *python3* se elevou.