

Ficha de Trabalho – Reconhecimento de Padrões

Docente: Miguel Tavares Coimbra

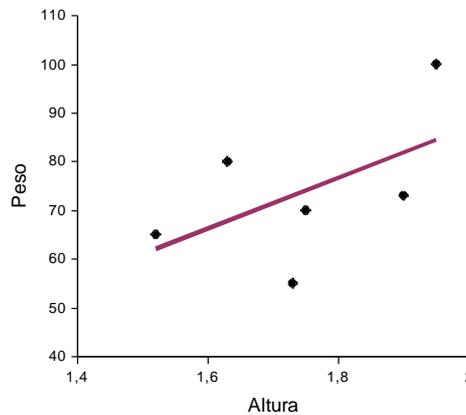
1. Representação do conhecimento.

- a. Considere a seguinte tabela. Usando uma folha de cálculo, faça um gráfico 2D com os valores contidos nesta.

| Indivíduo | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------|------|------|------|------|------|------|
| Peso | 1,75 | 1,90 | 1,52 | 1,63 | 1,95 | 1,73 |
| Altura | 70 | 73 | 65 | 80 | 100 | 55 |

Este gráfico está desenhado na figura em baixo.

- b. Por observação do gráfico, consegue dizer quais destas pessoas podem ser consideradas gordas? E magras?

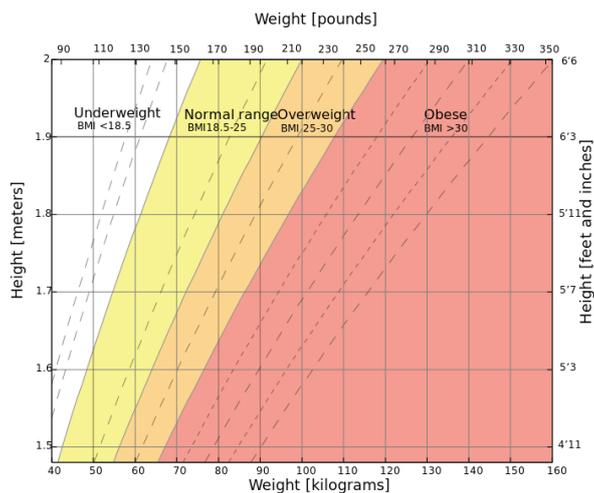


Para resolver esta pergunta é necessário, mesmo que mentalmente, dividir o *espaço de características* em dois sub-espaços: gordo e magro. A linha de regressão linear dá esta pista, propondo uma divisão do espaço em dois. No entanto conseguimos perceber que nem todas as classificações nos inspiram a mesma confiança. Podemos confiar num ponto que está quase sobre a linha de regressão?

- c. Formalize este seu “reconhecimento de um padrão”, criando uma *Regra* (par *condição-acção*) para cada uma destas situações.

Já fazemos isto! Como? Chamamos-lhe ‘índice de massa corporal’ e dividimos o espaço de características em várias ‘classes’. A World Health Organization propõe estas regras de ‘reconhecimento de padrões’ aqui:

http://apps.who.int/bmi/index.jsp?introPage=intro_3.html



(imagem adaptada de: <http://pt.wikipedia.org/>)

- d. Aplique a sua regra aos indivíduos anteriores, preenchendo a seguinte tabela. (Regra A: *Gordo / Não gordo*; Regra B: *Magro / Não magro*).

| Indivíduo | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------|---|---|---|---|---|---|
| Regra A | | | | | | |
| Regra B | | | | | | |

Esta pergunta tem como objectivo percebermos que muitas vezes não chega ter uma regra ‘binária’. Se criarmos uma regra para determinar se uma pessoa é gorda ou não, não estamos a afirmar nada sobre ser magra (pode não ser uma coisa nem outra). Isto quer dizer que teríamos que traçar duas linhas sobre o nosso gráfico da página anterior, se usarmos regras simples. Formalmente estamos a criar dois classificadores diferentes que actuam sobre o mesmo espaço de características, criando assim duas sub-divisões diferentes de espaço. Porque é que isto é importante? Porque não há nenhuma restrição formal que evite que os sub-espacos se sobreponham podendo acontecer algo estranho como uma pessoa ser considerada ao mesmo tempo gorda e magra se não formos cuidadosos.

- e. Considere uma nova tabela. Classifique mentalmente cada pessoa como gorda, magra ou normal. Aplique as regras criadas na alínea c). Os resultados são concordantes?

| Indivíduo | 7 | 8 | 9 | 10 |
|--------------------------|------|------|------|------|
| Peso | 1,72 | 2,05 | 1,67 | 1,82 |
| Altura | 85 | 95 | 65 | 61 |
| Classificação Manual | | | | |
| Classificação Automática | | | | |

Esta pergunta destina-se a perceber as diferenças entre a minha classificação ‘certa’ (assumindo que a manual assim o é) e a criada por uma simples regra. Ou, pensando nos slides da cadeira, a diferença entre o mapa (e os seus detalhes nas fronteiras) de uma regra simples de latitude e longitude.

2. **Reconhecimento estatístico de padrões.**

- a. Considere que criou um algoritmo que segmenta objectos circulares de uma fotografia a preto e branco. O seu objectivo é identificar se estes círculos correspondem a bolas escuras ou claras. Classificando manualmente as várias fotografias, obteve os seguintes resultados:

| | | | | | | | | | | |
|---------------------------|----|----|-----|----|-----|-----|----|-----|----|----|
| Bolas escuras | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Luminosidade Média | 10 | 55 | 152 | 34 | 175 | 101 | 77 | 163 | 44 | 95 |

| | | | | | | | | | | |
|---------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Bolas claras | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Luminosidade Média | 230 | 243 | 180 | 134 | 244 | 153 | 180 | 175 | 220 | 215 |

- b. Calcule a média e o desvio-padrão da distribuição da luminosidade de cada uma das classes.

| Classe | Média | Desvio-Padrão |
|-------------------------|-------|---------------|
| A: Bolas escuras | | |
| B: Bolas claras | | |

- c. Assumindo que cada classe têm uma distribuição Gaussiana, calcule a probabilidade destas novas bolas serem claras ou escuras. Perante estas probabilidades, tome uma decisão: “A bola é escura” ou “A bola é clara”.

$$Gauss_{pdf} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

| | | | | | | | | | | |
|---------------------------|----|-----|-----|-----|-----|-----|-----|-----|----|-----|
| Bolas | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Luminosidade Média | 90 | 145 | 133 | 132 | 146 | 223 | 252 | 204 | 78 | 195 |
| Decisão | | | | | | | | | | |

- d. O resultado correcto está descrito na tabela seguinte. Calcule a precisão do seu classificador. (*C-Clara, E-Escura*)

| | | | | | | | | | | |
|------------------|---|---|---|---|---|---|---|---|---|----|
| Bolas | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Resultado | E | E | C | E | C | C | C | C | E | E |

$$\text{Precisão} = \frac{\text{Nr. Correcto de Class.}}{\text{Nr. Total de Class.}}$$

Estas perguntas são bastante directas sendo apenas necessário aplicar as fórmulas pedidas.

3. **Aprendizagem máquina.**

- a. Para conhecer mais profundamente os métodos discutidos na aula teórica, pode aceder a este excelente repositório online de tutoriais:

<http://www.autonlab.org/tutorials/>