



UNIVERSIDADE DO MINHO

DEPARTAMENTO DE INFORMÁTICA

DESCOBERTA DE CONHECIMENTO

Ficha 4

8 de Março de 2019

Francisco Oliveira (A78416)



Conteúdo

1	Parte1	2
1.1	Exercício 1	2
1.2	Exercício 2	2
1.3	Exercício 3	2
1.3.1	A	2
1.4	Exercício 4	2
1.5	Exercício 5	2
2	Parte2	2
2.1	Exercício 2	2
2.2	Exercício 4	3

1 Parte1

1.1 Exercício 1

Uma limitação do modelo de correlação é a definição de relações falsas.

Embora possa ser calculada uma correlação entre dois atributos de um certo conjunto de dados, pode não existir uma razão lógica para explicar que a alteração de um atributo influencia os valores do outro atributo. Pode ser uma simples coincidência, ou outra explicação que o modelo não fornece.

1.2 Exercício 2

Os coeficientes tentam encontrar o valor a variabilidade compartilhada entre atributos, através de uma fórmula matemática. Os coeficientes de correlação apenas tentam medir a proximidade (força) da interação entre os atributos.

1.3 Exercício 3

Correlação positiva é quando ambos os atributos aumentam, ou diminuem, ora seja alteração no mesmo sentido.

Correlação negativa é o oposto, ora seja, quando um atributo aumenta e o outro diminui.

1.3.1 A

Se *diminuem* á mesma taxa então é um correlação positiva, pois ambos alteram no mesmo sentido

1.4 Exercício 4

A força de relação é medida comparando os valores de um parâmetro com outro, e calculando um coeficiente.

O limite é entre -1 e 1.

1.5 Exercício 5

Atributos que poderiam ser incluídos no dataset, e que poderão influenciar o uso de óleo são: certificado energético da casa e localização da casa.

2 Parte2

2.1 Exercício 2

Informação Relevante:

Este conjunto de dados dizem respeito ao consumo de combustível do ciclo urbano em "miles per gallon"(mpg).

Número de Instâncias: 398

Número de Atributos: 9 incluindo o atributo class

Informações sobre Atributos dos Veículos:

1. mpg: contínuo (miles per gallon)
2. cilindros: discreto com vários valores
3. deslocamento: contínuo

4. cavalos de potência: contínuo
5. peso: contínuo
6. aceleração: contínua
7. ano do modelo: discreto com vários valores
8. origem: discreto com vários valores
9. nome do carro: string (único para cada instância)

Valores Atributos Ausentes: a potência do cavalo tem 6 valores indefinidos

2.2 Exercício 4

Podemos ver que o atributo *MPG* tem uma correlação negativa com os atributos *Cylinders* e *Weight*, o que significa que quanto maior o número de cilindros e maior for o peso do veículo, menor será o número de milhas que fará por 'galon' (mpg), ou seja, maior será o consumo de combustível.

Attribut...	mpg	cylinders	displac...	horsep...	weight	acceler...	model y...	origin
mpg	1	-0.778	-0.806	-0.779	-0.835	0.421	0.568	0.566
cylinders	-0.778	1	0.951	0.843	0.896	-0.505	-0.349	-0.563
displace...	-0.806	0.951	1	0.897	0.933	-0.544	-0.370	-0.609
horsepo...	-0.779	0.843	0.897	1	0.865	-0.689	-0.416	-0.455
weight	-0.835	0.896	0.933	0.865	1	-0.417	-0.307	-0.581
accelerat...	0.421	-0.505	-0.544	-0.689	-0.417	1	0.288	0.206
model ye...	0.568	-0.349	-0.370	-0.416	-0.307	0.288	1	0.181
origin	0.566	-0.563	-0.609	-0.455	-0.581	0.206	0.181	1

Figura 1: Matriz de correlação