Grupo 25 – AMD – AP05

## Exercício 2.b)

Foi feita a análise da classe lenses, e neste caso as frequências vinham de acordo com as frequências nominais que apareceram no dataset, pelo que a soma de todos dava o valor total de entradas.

## Exercício 2.c)

Neste caso optámos por ver a probabilidade da prescrição ser relacionada com o tipo de lene, e foi possível identificar que as pessoas com lentes rígidas tinham mais tendência a sofrer de hipermetropia.

Também foi apresentado as percentagens dentro do grupo que no caso sugerido era de 75% e no tal era de 12.5%.

## Exercício 2.e)

Neste exercício focámo-nos na relação entre tipo de lente e astigmatismo, e conseguimos sugerir com alguma confiança, que pessoas com astigmatismo apenas usam lentes moles, enquanto que quem possui astigmatismo tem de utilizar lentes rígidas.

## Exercício 3.a)

No primeiro teste, os resultados foram 0% para a existência de valores ausentes, isto deve-se a todos os valores apresentados no dataset serem válidos e estarem presentes.

De seguida foi feita uma cópia do ficheiro lenses, no qual retirámos valores de todas os atributos, e a incidência foi retirar mais de idade e prescrição, e apenas um de estigmatismo e de tear\_rate. A classe lenses manteve-se com 0% de valores ausentes pois não retirámos nenhum. No atributo age, retirámos 5 entradas correspondentes a 20.8%, no atributo prescription retirámos 4 valores correspondendo a 16.7% do total, e retirámos 1 entrada tanto a astigmatic como a tear\_rate, que correspondeu em ambos a 4.2%.

Por fim, com o ficheiro “lenses\_with\_missingValues”, faltaram 4 entradas de age, correspondendo a 16.7%, faltou uma entrada de em tear\_rate, correspondendo a 4.2% e faltaram 2 entradas de lenses, correspondendo a 8.3%. Nos restantes atributos não houve valores ausentes.

## Exercício 3.b)

O que a matriz de contingência mostra, é o mesmo que a regra 1R demonstra, excepto que demonstra de forma diferente. O 1R demonstraria a relação como erros/ valor total de entradas, e poderia ser simplificado como fração, enquanto que a matriz de contingência apresenta exclusivamente os valores que estão dentro dos parâmetros.

Por exemplo no caso de relação entre lentes e idade, para a idade prepresbyopic a matriz de contingência apresenta “<pre-presbyopic> hard: 1.0 | none: 2.0 | soft: 2.0 |” enquanto que a regra 1R apresentaria “<pre-presbyopic> hard: 4/5| none: 3/5| soft: 3/5 |”, no entanto acabam por conter a mesma informação.

## Exercício 3.f)

O primeiro gráfico apresenta a matriz de erro, que é apresentada no ficheiro Excel do exercício anterior. Isto apresenta o erro que cada combinação prescrição – lente apresenta. Um dos casos curiosos é hipermetropia – lentes rígidas, que apresenta uma taxa de erro de 1, pelo que de acordo com o dataset podemos sugerir, que uma pessoa que tenha hipermetropia não pode usar lentes rígidas.

Neste exercício a segunda tabela apresenta os mesmo valores para a regra 1R apresentados no slide. Mas isto apenas nas regras apresentadas, que são ('prescription', 'myope', 'hard') e ('prescription', 'hypermetrope', 'none'), sendo que a única diferença é que no slide são apresentados na forma de erro / total, enquanto que aqui são apresentadas na forma de fração.