

Curso: Mestrado em Engenharia Biomédica
U.C.: Aprendizagem e Extração do Conhecimento

Plano de Aula 05	
Docente:	Hugo Peixoto José Machado
Tema:	RapidMiner: Revisão e Continuação da Aula Anterior
Ano Letivo:	2023-2024 – 1º Semestre
Duração da aula:	2 horas

1. Exercício 1

Importe o dataset happyscore_income.csv e prepare o processo de Data Mining da forma que considerar mais correta:

[1] Execute os processos de Business e Data Understanding para perceber qual o processo de DM que deverá ser aplicado. Deverá colocar:

- [a] 1 frase que caracterize o objetivo do negócio.
- [b] 1 frase que caracterize o objetivo do processo de DM.
- [c] 1 tabela com o resumo dos atributos encontrados (tipo, min, max, med, nulos, etc).

[2] Crie o modelo de DM no RapidMiner e justifique a utilização de todos os operadores de preparação de dados que incluiu no seu modelo.

[3] Descreva as principais conclusões a que chegou.

2. Exercício 2

Importe o dataset insurance.csv e prepare o processo de Data Mining da forma que considerar mais correta:

[1] Execute os processos de Business e Data Understanding para perceber qual o processo de DM que deverá ser aplicado. Deverá colocar:

- [a] 1 frase que caracterize o objetivo do negócio.
- [b] 1 frase que caracterize o objetivo do processo de DM.
- [c] 1 tabela com o resumo dos atributos encontrados (tipo, min, max, med, nulos, etc).

[2] Utilizando a visualização do RapidMiner avalie os diferentes gráficos e execute uma descrição do que for encontrando. Por exemplo:

- [a] Relação entre o Sexo e o Prémio do Seguro
- [b] Relação entre o número de dependentes e o Prémio do Seguro
- [c] Relação entre se é ou não fumador e o Prémio do Seguro
- [d] Relação entre a localização e o Prémio do Seguro

[2] Crie o modelo de DM no RapidMiner e justifique a utilização de todos os operadores de preparação de dados que incluiu no seu modelo.

[3] Descreva as principais conclusões a que chegou.