Icon

Description automatically generated

Universidade do Minho

Departamento de Informática

Sistemas Baseados em Similaridade

Trabalho Prático Individual 1

Gonçalo Almeida (A84610)

28 de Outubro de 2020

**T1.** Carregar, no Knime, ambos os datasets. Utilizar um nodo Joiner para agregar, por “area code” e “phone”, os dados provenientes das duas readers. Transformar o atributo Churn em nominal;

A picture containing flower, bird, tree

Description automatically generated

Figura - Carregar os datasets

Graphical user interface, text, application, email

Description automatically generated

Figura - Nodo Joiner

Graphical user interface, application

Description automatically generated

Figura - Transformar o atributo Churn em nominal

**T2.** Particionar os dados de forma estratificada (pela feature “Churn”), utilizando 75% para aprendizagem e 25% para teste. Aplicar um Decision Tree Learner e um Decision Tree Predictor. Avaliar a precisão (accuracy) do modelo utilizando o nodo Scorer e a respetiva matriz de confusão;

Graphical user interface, text, application

Description automatically generated

Figura 4 - Nodo Reader

Figura - Partição dos dados

Graphical user interface, application

Description automatically generated

Figura - Decision Tree Learner

Graphical user interface, text, application

Description automatically generated

Figura - Decision Tree Predictor

Graphical user interface, text, application

Description automatically generated

Figura - Scorer

Graphical user interface, application, table

Description automatically generated

Figura - Matriz de Confusão

**T3.** Remover, iterativamente, features do dataset e reavaliar a performance dos modelos candidatos. Descrever os resultados obtidos;

Graphical user interface, application

Description automatically generated

Table

Description automatically generated

Graphical user interface, application

Description automatically generated

Table

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application, table

Description automatically generated

**T4.** Seguir as práticas de bons-hábitos na construção de workflows;

Diagram

Description automatically generated

Figura - Workflow

**T5.** Utilizar o output de um nodo Decision Tree Learner para criar uma imagem de uma

Árvore de Decisão e guardar essa imagem no ambiente de trabalho;

Graphical user interface, text, application, email

Description automatically generated

Figura - Decision Tree To Image

Graphical user interface, text, application

Description automatically generated

Figura - Image Writer

**T6.** Explorar os datasets originais, procurando extrair informação relevante dos dados.