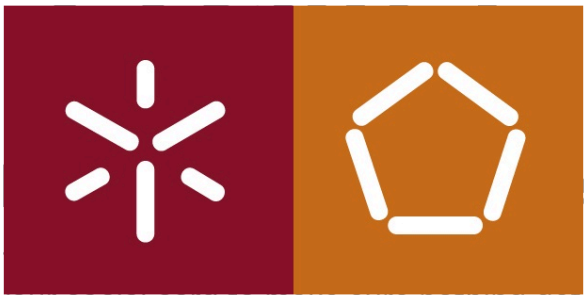


Universidade do Minho

Departamento de Informática

# Sistemas Baseados em Similaridade

## Trabalho Prático Individual 1

Gonçalo Almeida (A84610)

19 de Outubro de 2020

**T1.** Instalar a plataforma Knime.

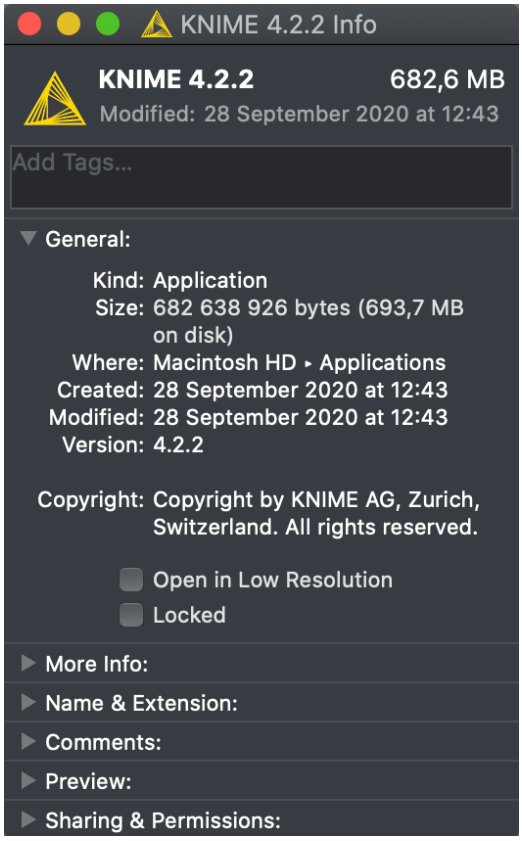



*Figura 1 - Informações da Aplicação Knime*

**T2.** Desenvolver um workflow que, utilizando um nodo reader, faz a correta leitura do dataset disponível em https://bit.ly/3hXCwIG.

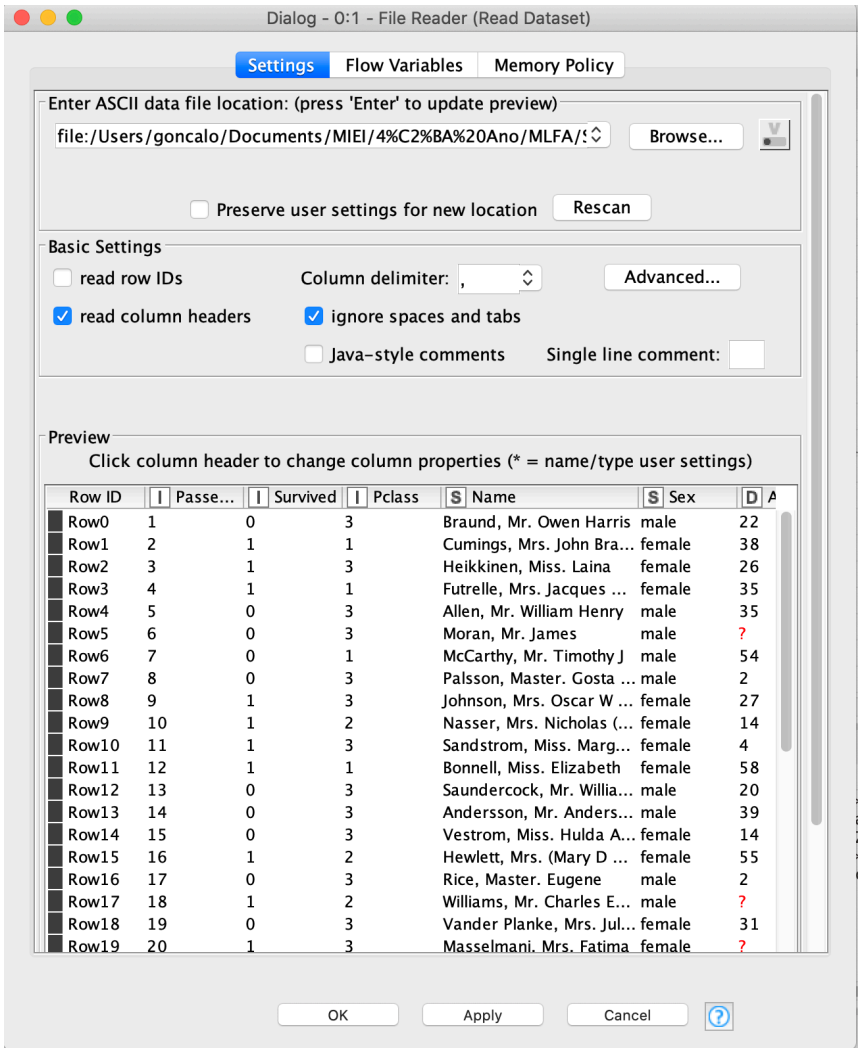
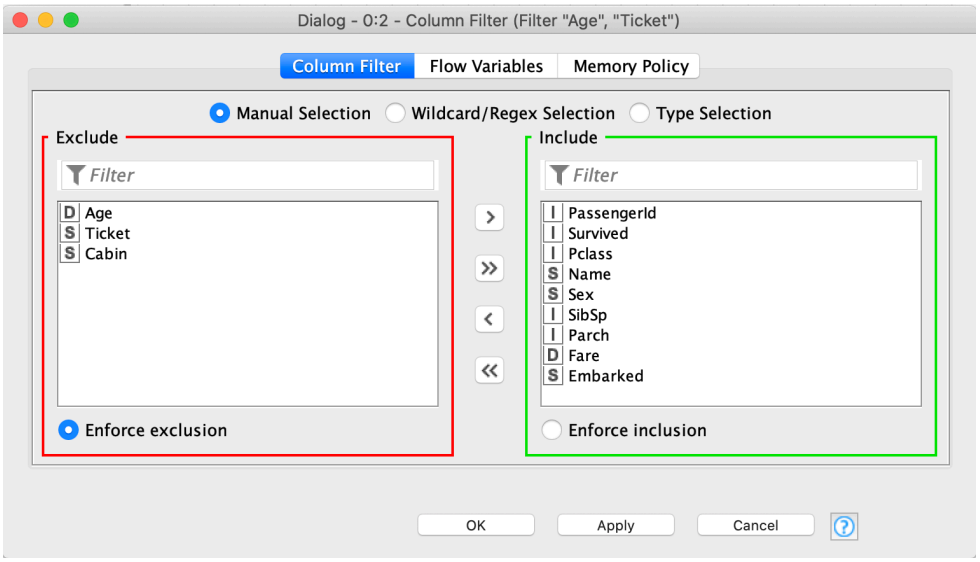


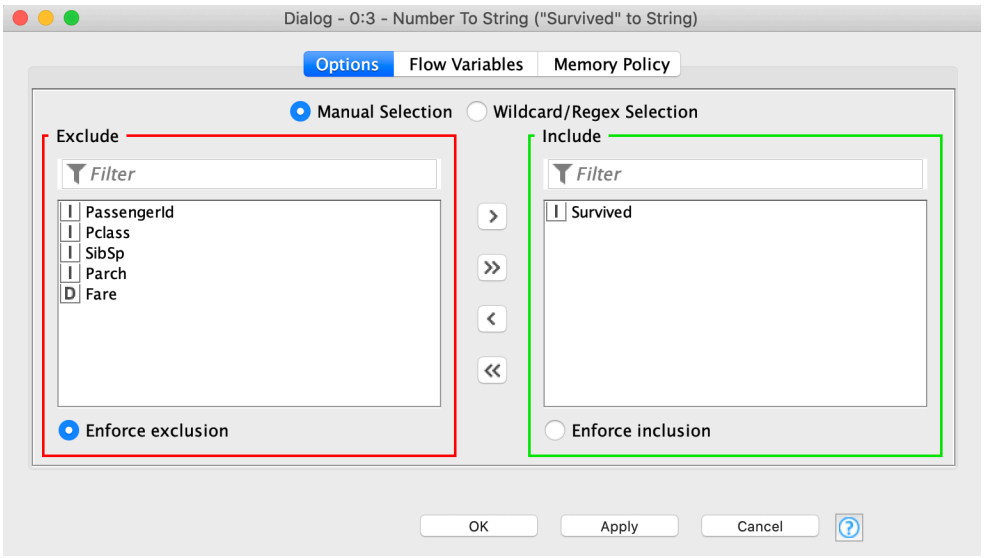

*Figura 2 - Nodo Reader*

**T3.** Utilizar um conjunto de nodos para:

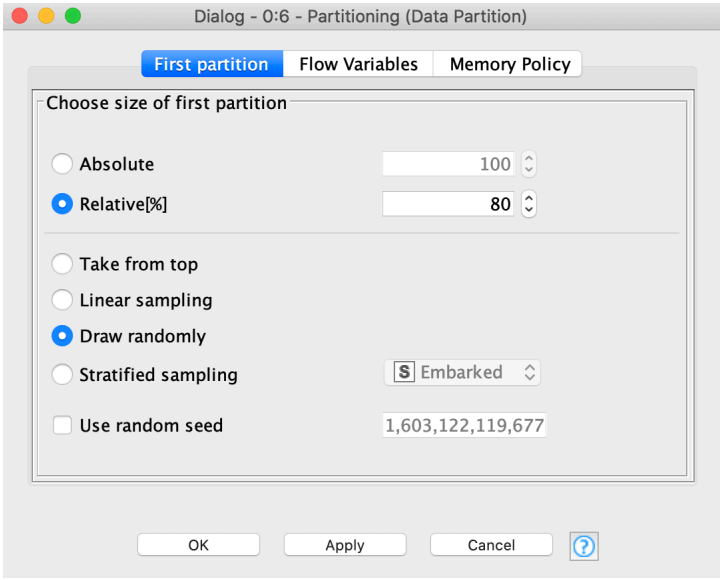
- Filtrar as colunas “Age”, “Ticket” e “Cabin”;



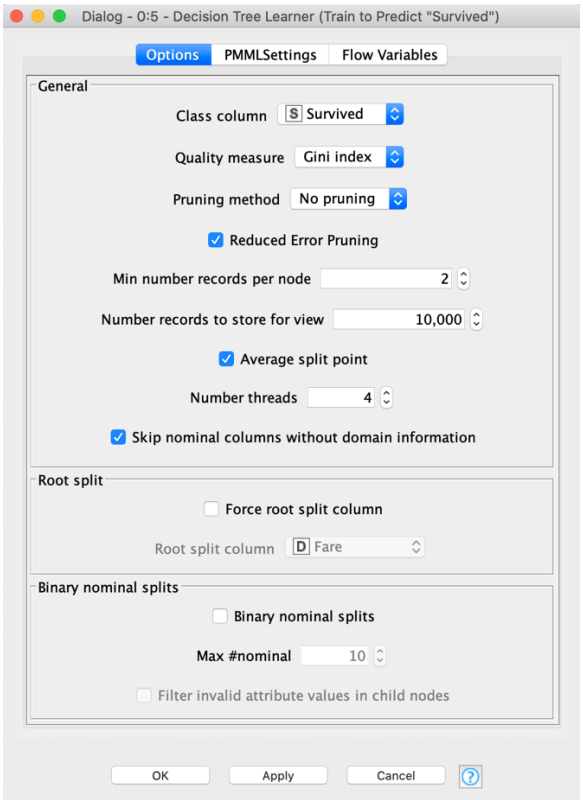

- Fazer o cast da coluna “Survived” para String;

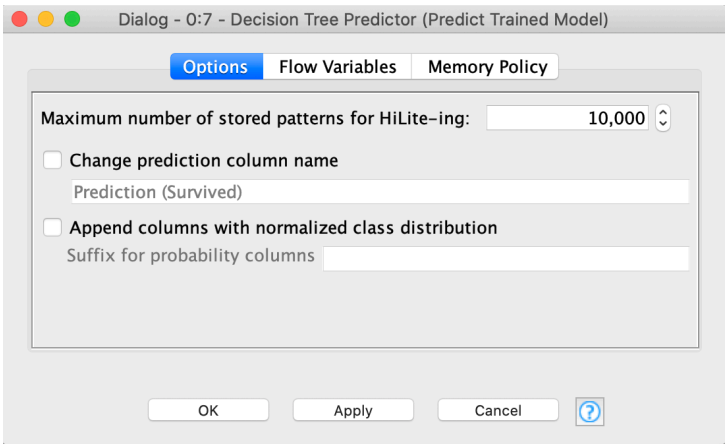


- Particionar os dados, de forma aleatória, utilizando 80% para aprendizagem e 20% para teste;

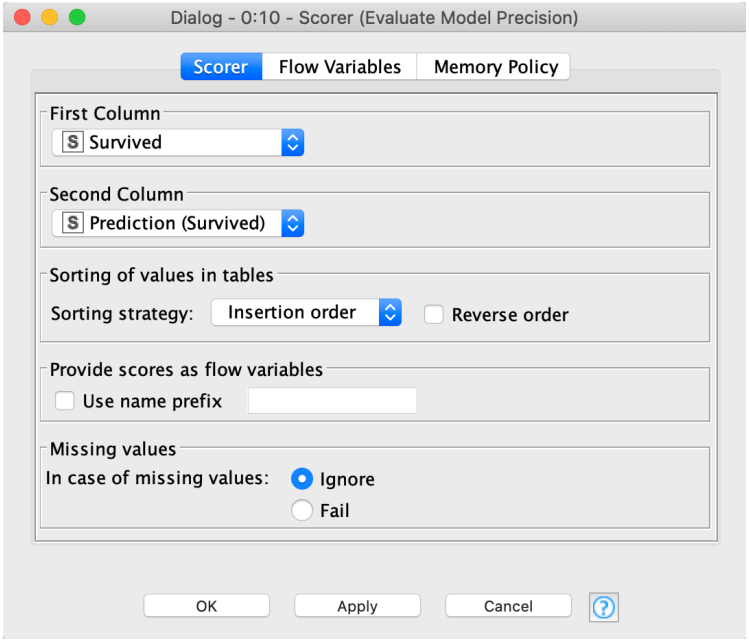



- Aplicar um nodo Decision Tree Learner para treinar uma Árvore de Decisão e um Decision Tree Predictor para obter previsões utilizando o modelo treinado;






- Avaliar a precisão (accuracy) do modelo utilizando o nodo Scorer e a respetiva matriz de confusão.



**T4**. Experimentar várias combinações de parâmetros no nodo *Decision Tree Learner* e documentar as performances obtidas.

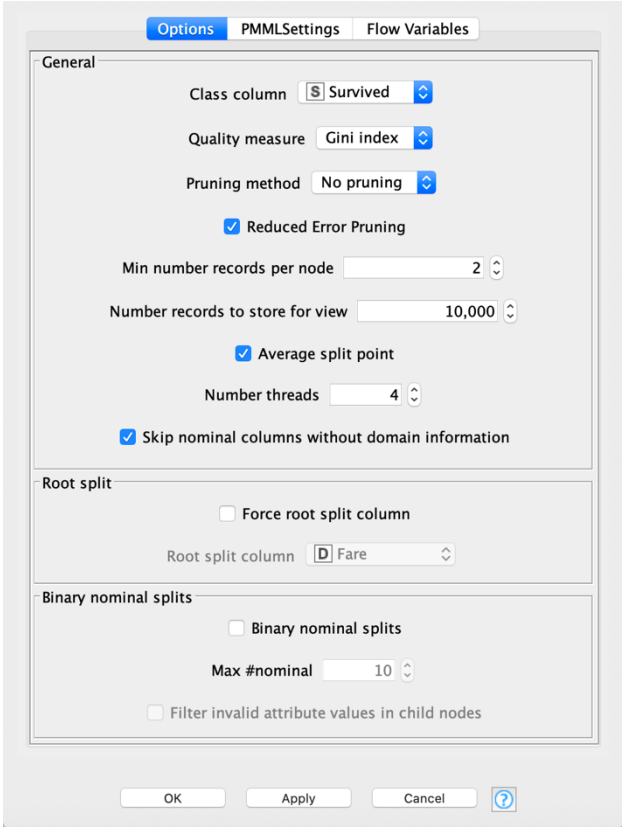



*Figura 3 - Configuração 1*

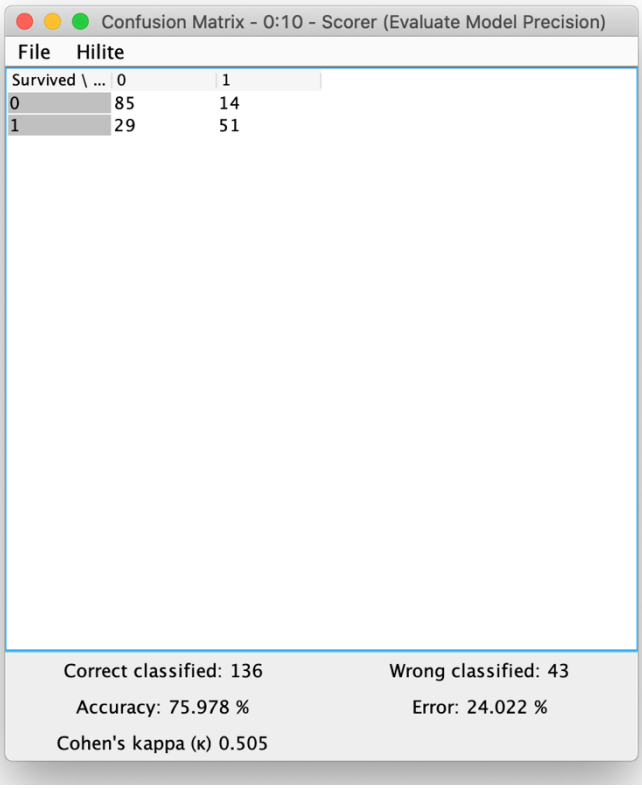



*Figura 4 - Matriz de Confusão 1*

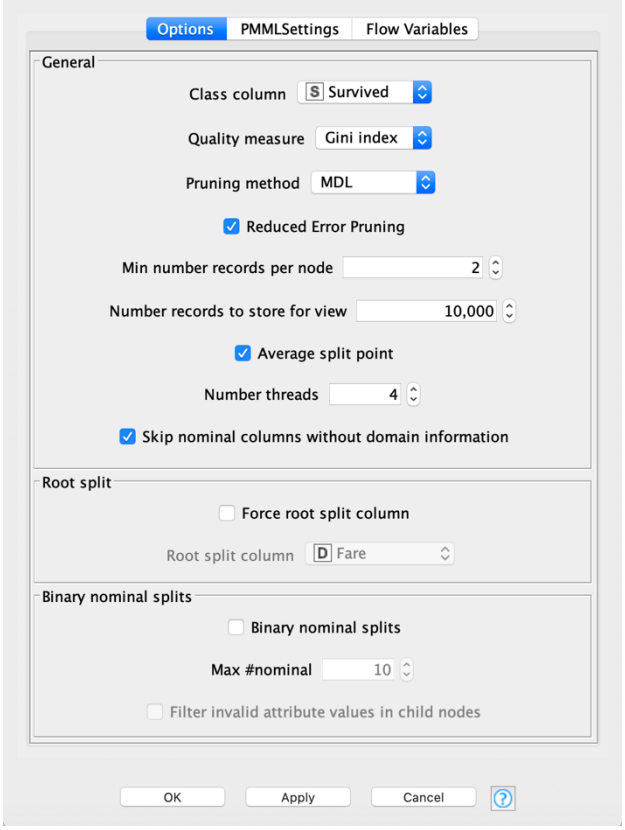



*Figura 5 - Configuração 2*

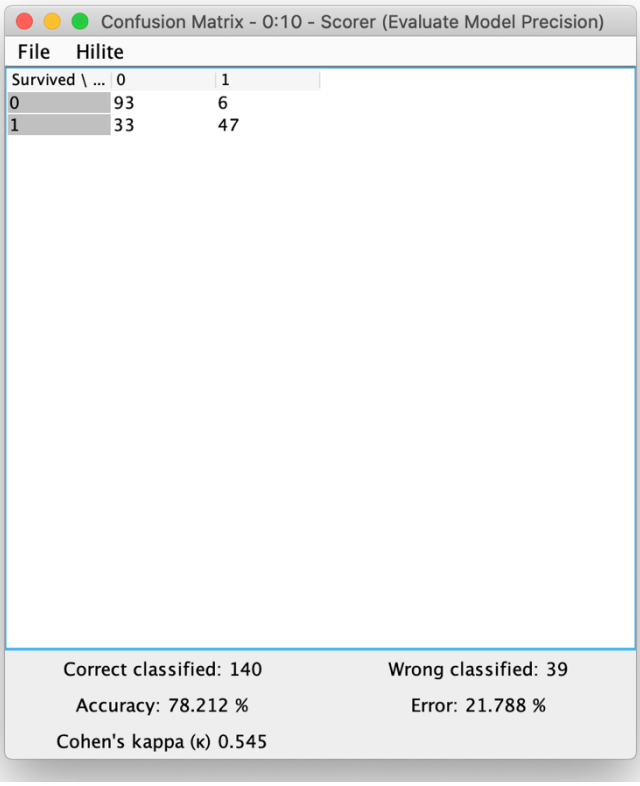



*Figura 6 - Matriz de Confusão 1*

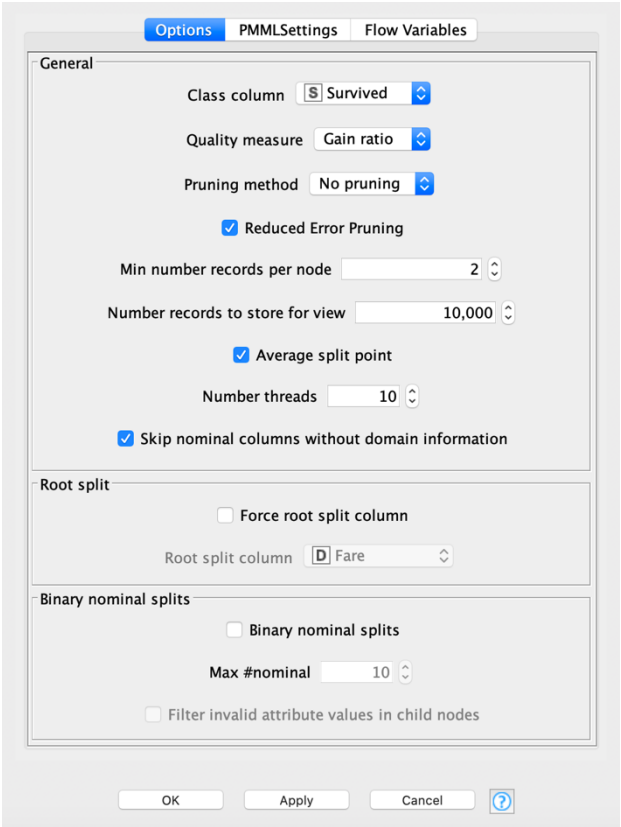



*Figura 7 - Configuração 3*

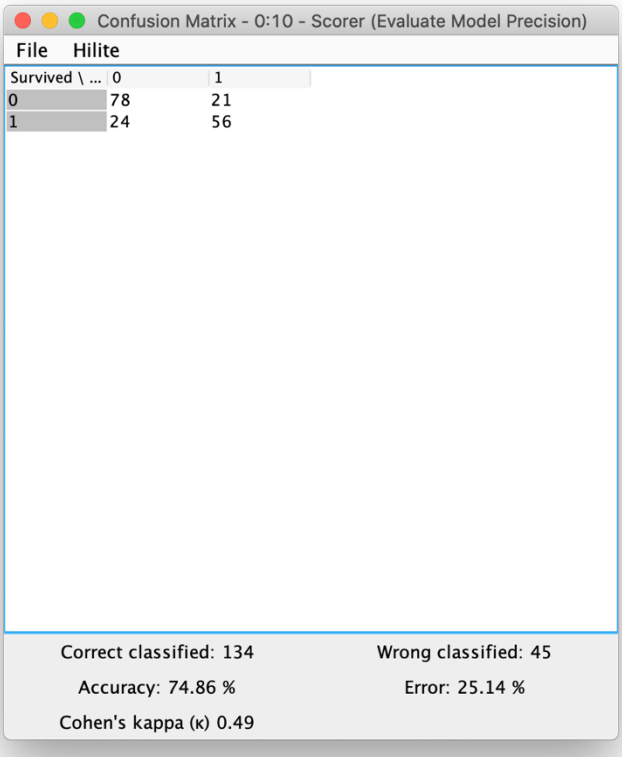



*Figura 8 - Matriz de Confusão 3*

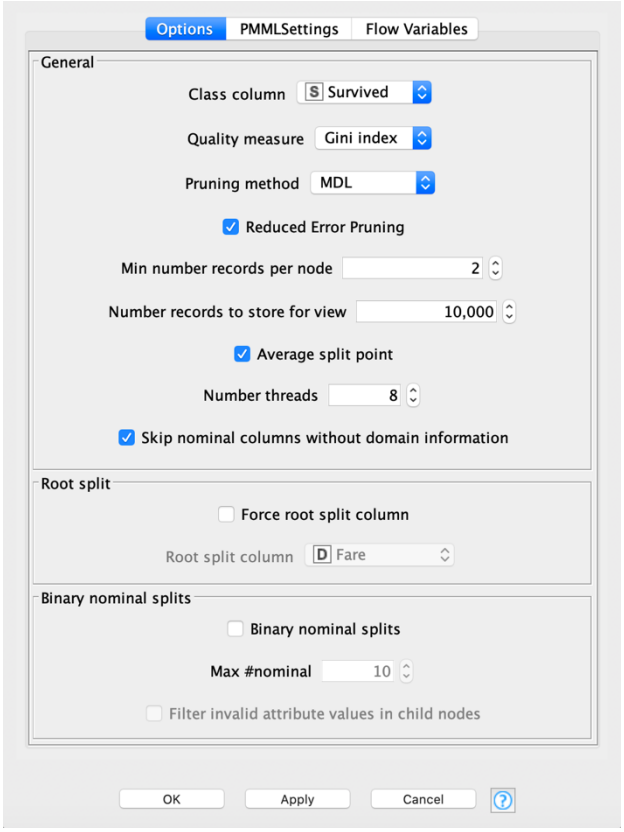



*Figura 9 - Configuração 4*

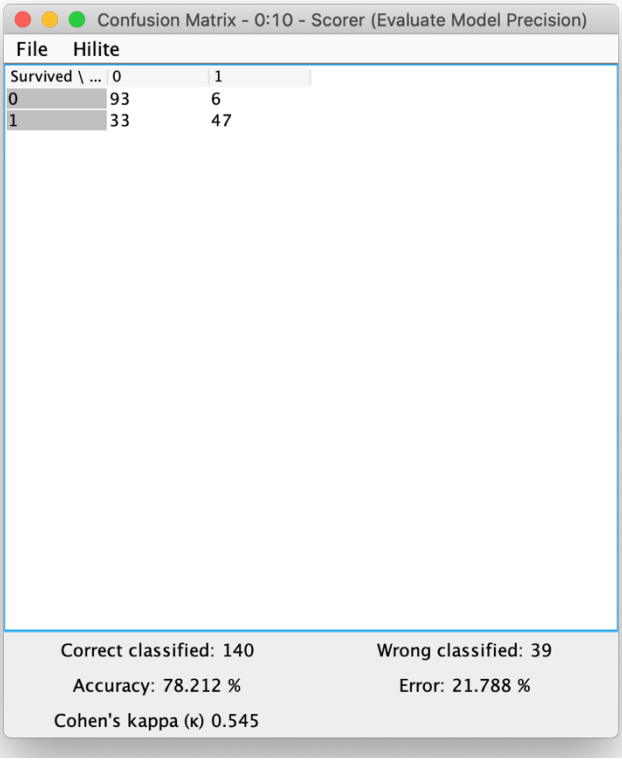



*Figura 10 - Matriz de Confusão 4*