

PC0117

Introdução a Meta Heurística

Pedro Paes Siniscalchi
Breno de Oliveira Renó

Introdução – Tutorial

01

Manipulação de Dados

...

02

Pré-processamento dos dados

...

03

Modelagem

...

04

Resultados

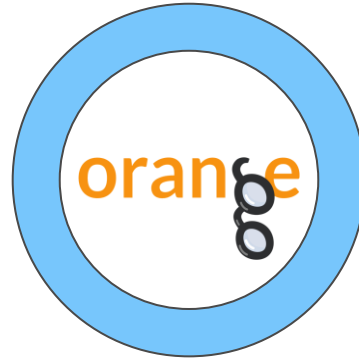
...

05

Agradecimentos

...





Download

O programa é disponibilizado em todas as variações de sistemas operacionais, sendo Windows, Linux e Mac OS. A instalação tem variações para windows sendo arquivo .exe e para linux instalação via terminal.

[Orange Download](#)

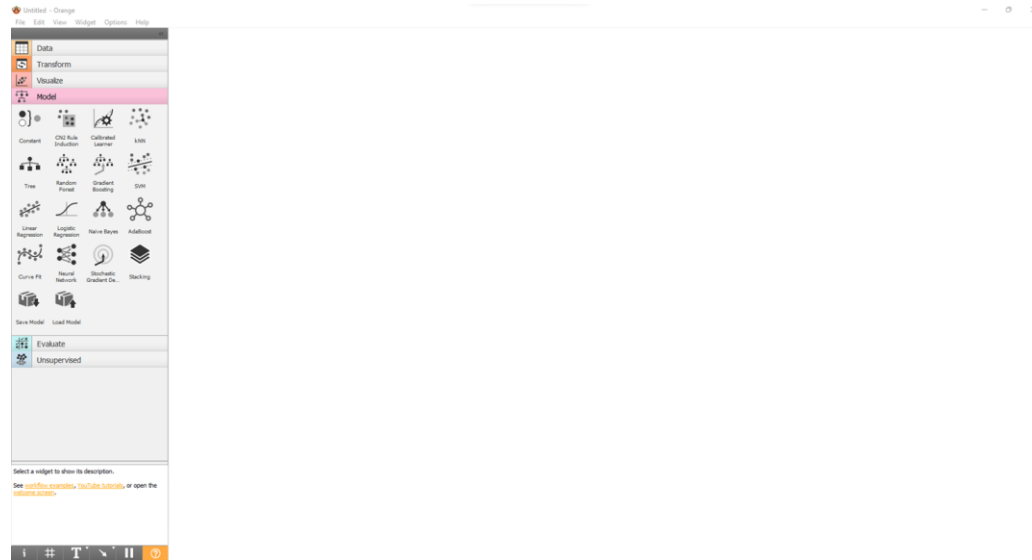
...

Dataset

Para os testes utilizamos o *wine dataset* disponibilizado no [link](#). O dataset é composto por 178 tipos de vinhos distribuídos por classes.

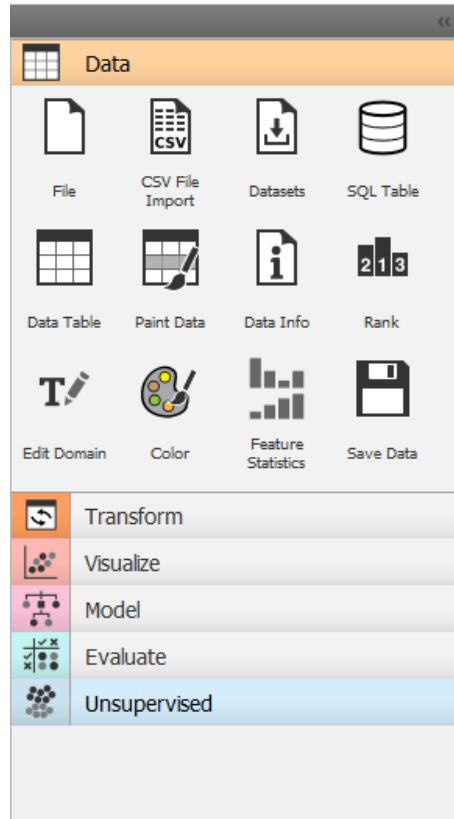


Interface



O programa
apresenta uma
interface amigável
e bem intuitiva.

Interface



Menu lateral com
ferramentas para
os mais variados
tipo de
manipulações de
dados

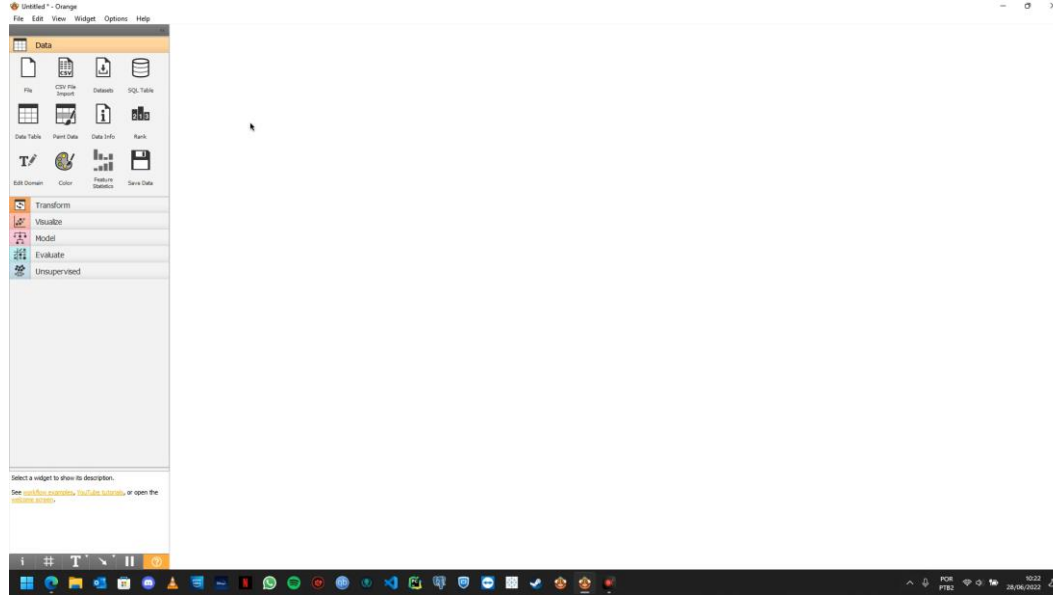


01

Manipulação



Manipulação



Por se tratar de um programa intuitivo, o funcionamento é baseado em drag-and-drop.

Para carregar um *dataset* basta arrastar o ícone file ou *datasets* e seleccionar o que será utilizado.

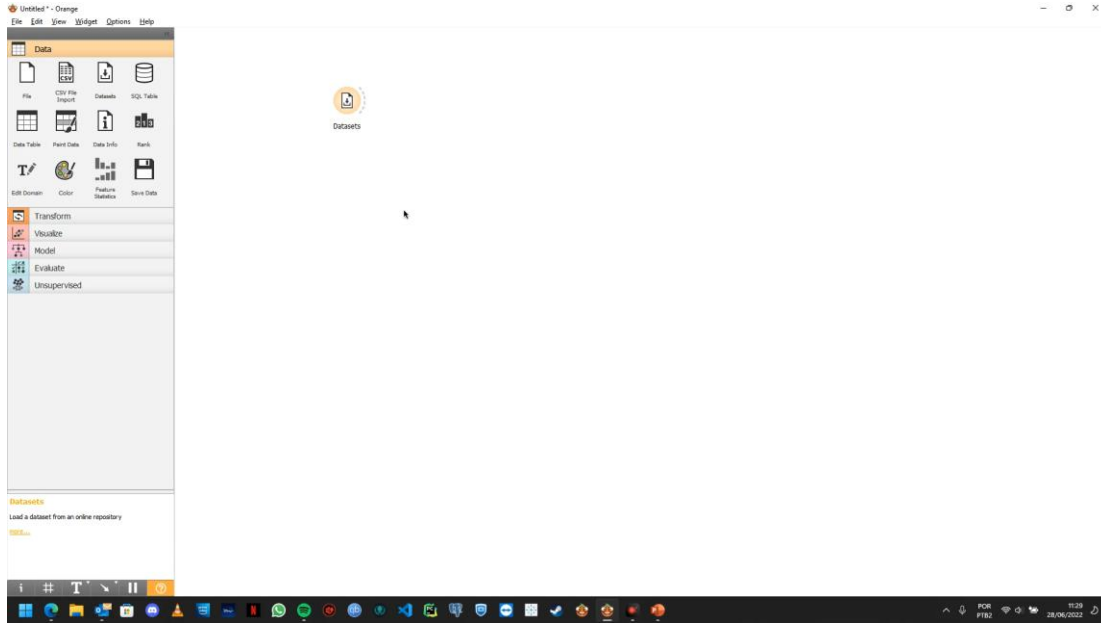


02

Processamento



Manipulação e Modelagem

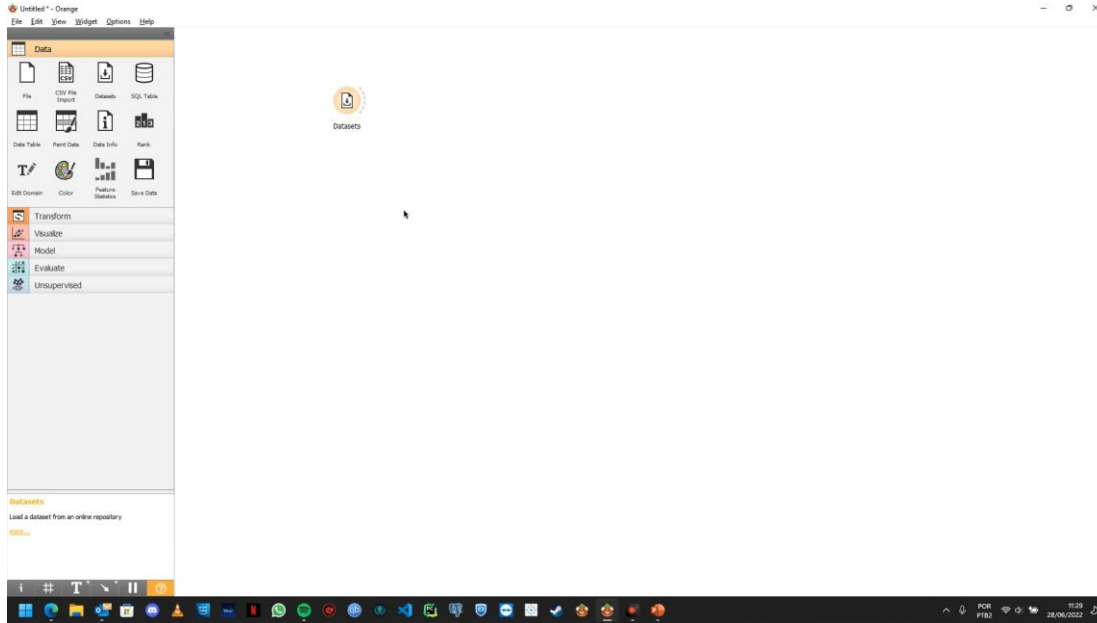


Para as manipulações é necessário inserir um menu de testes e a rede neural. Ao linkar os processamentos é necessário se atentar que a saída dos dados é sempre para direita e entrada sempre a esquerda.

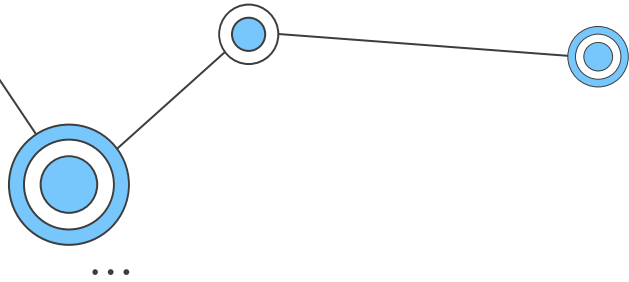
03

Modelagem

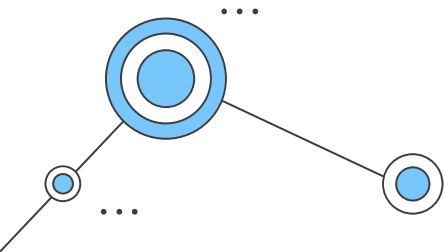
Manipulação e Modelagem



Após as inserções de testes e rede neural, deve-se desmonstrar os resultados. A escolha de mostragem fica a critério do projeto trabalhado. Nesta desmontração utilizamos a curva ROC e a matriz de confusão.



A quantidade de iterações é feita de forma empírica pois pode ocorrer um *underfitting* ou um *overfitting*. Para os testes foram utilizadas 200 iterações.

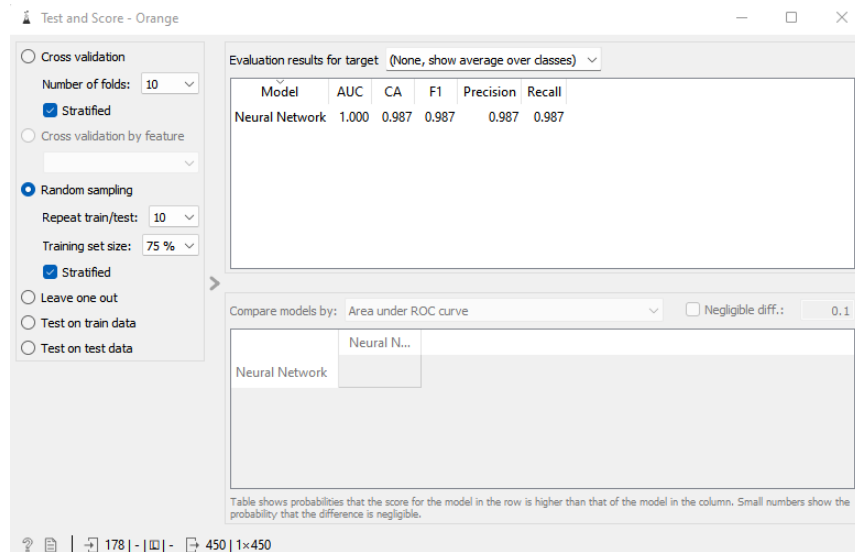


Configuração de Teste

Por default a rede em *Random Sampling* (utilização de forma aleatória):

- Quantidades de repetições
- Aproveitamento dos dados para treino

A porcentagem de aproveitamento refere-se a quantidade de dados que serão utilizados para treino e teste, a quantidade é empírica e pode ser variada de acordo com o projeto.



Parâmetros dos Testes



Acurácia

Indica uma performance geral do modelo. Dentre todas as classificações, quantas o modelo classificou corretamente.



Precisão

Dentre todas as classificações de classe do Positivo que o modelo fez, quantas estão corretas.



Recall

Dentre todas as situações de classe Positivo como valor esperado, quantas estão corretas.



F1 - Score

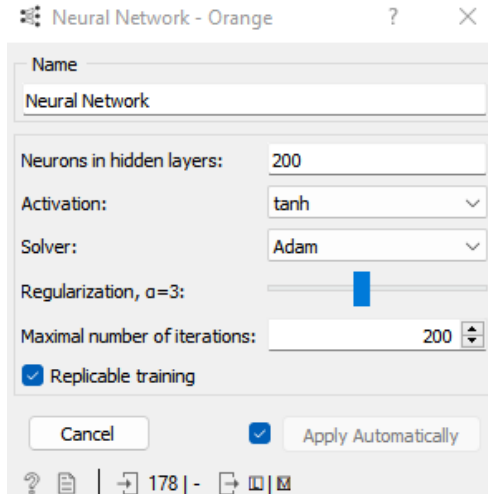
Média harmônica entre precisão e recall.

Amostragens

01

Rede Neural

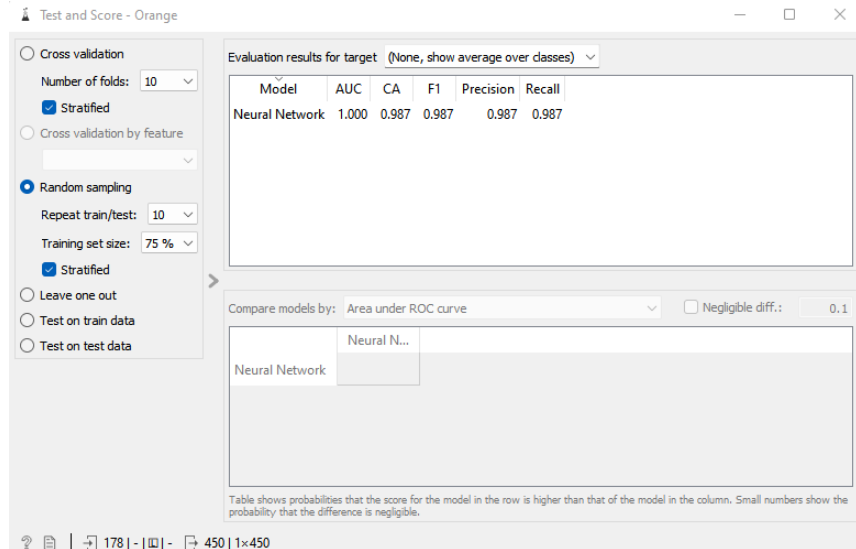
Neurônios (camada oculta) – 200
Teste de Ativação – Identity, Logistic, tanh e ReLu
Regularization (tamanho do neurônio) – 3
Número de Iterações – 200



02

Testes

Área abaixo da curva ROC (AUC) – 1.000
Acurácia (CA) – 0.987
F1 – 0.987
Precision – 0.987
Recall – 0.987



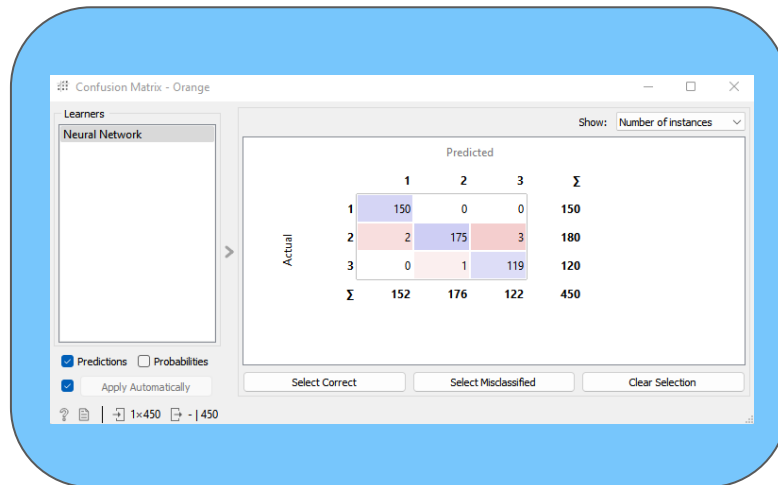


04

Resultados

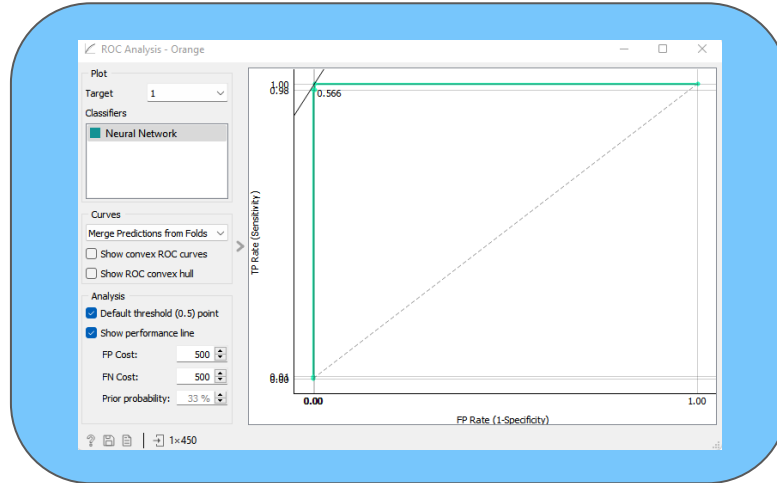


Matriz de Confusão



Como primeiro resultado, a matriz de confusão demonstra a quantidade de instâncias criadas e sua combinação com outras classes. Devido as repetições, os dados foram proporcionalmente aumentados assim gerando uma variabilidade.

Curva ROC



Como segundo resultado escolhemos a curva ROC que demonstra a quantidade de verdadeiros positivos em relação ao FP Rate (taxa de sucessão e aproveitamento). Quanto maior a ocupação da área, maior será a taxa de sucesso e consequentemente melhor o aproveitamento.

Obrigado!

