

Carregar e usar modelo do Orange em Python

Paulo Ferrari

Para carregar um modelo treinado no Orange Data Mining e usá-lo para fazer previsões em Python, você pode seguir os passos abaixo. O software Orange permite a exportação de modelos treinados para o formato pickle (.pkcls), que é compatível com Python.

1. Treinamento e exportação do modelo no Orange:

Treine o modelo no Orange e use o widget "Save Model" para salvar o modelo treinado em um arquivo .pkcls.

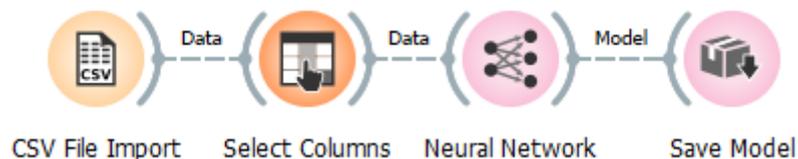


Figura 1. Salvar modelo no Orange

Documentação dos widgets: <https://orangedatamining.com/widget-catalog>

2. Carregar o modelo em Python:

O seguinte trecho de código mostra como carregar o modelo e obter os nomes dos atributos (`domain`) do seu dataset e ver em qual ordem ficaram representados no modelo. Também mostra como obter as classes do atributo target.

```
import pickle

with open('caminho/para/modelo.pkcls', 'rb') as file:
    model = pickle.load(file)

domain = model.domain

classes = model.domain.class_var.values

print(f'domain: {domain} - len: {len(domain)}')
print(f'classes: {classes} - len: {len(classes)}')
```

Obs: A lib pickle é nativa do Python, não precisa instalar, basta fazer o import.

Documentação da lib pickle: <https://docs.python.org/3/library/pickle.html>

3. Usar o modelo para fazer previsões:

Instale a lib do Orange com o seguinte comando.

```
pip install Orange3
```

O seguinte trecho de código mostra como carregar os dados e fazer previsões, assumindo que os dados estão num arquivo .csv com o mesmo formato do dataset usado para treinamento do modelo. Esse arquivo .csv deve conter o cabeçalho com os nomes dos atributos (incluindo o atributo target) para que o modelo consiga mapear cada atributo corretamente na previsão. Por isso, dar print na variável `domain` para saber a ordem dos atributos pode ser útil caso não queira incluir o cabeçalho ou utilize dados em outro formato.

```
import Orange

# Load your data to make predictions. Assume you have a data
file 'data.csv' to predict

data = Orange.data.Table('caminho/para/dados_previsao.csv')

# Make predictions

predictions = model(data)

# Print predictions

for prediction in predictions:

    p = int(prediction)

    c = classes[p]

    print(f'Prediction: {p} - Class: {c}')
```

Documentação da lib Orange3: <https://pypi.org/project/Orange3>

Se eventuais warnings aparecer você pode resolvê-los ou suprimir sua impressão no terminal com o seguinte comando.

```
import warnings

warnings.filterwarnings("ignore") # Suppress all warnings
```