

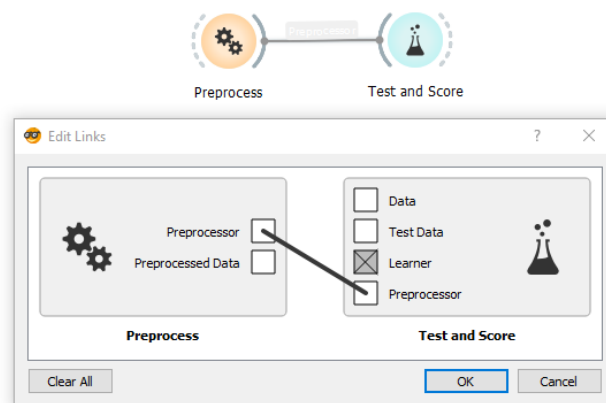
## Homework (Diperbaharui)

### 1. Pemodelan dengan dataset “breast cancer wisconsin”

Anda perlu mempraktekan kembali tutorial diatas untuk pemodelan dengan menggunakan dataset “Breast Cancer wisconsin” lalu **men screenshot canvas beserta hasil pemodelan (AUC, CA, F1, Precision dan Recall) yang telah anda lakukan.**

Data yang digunakan adalah data yang terdapat pada kaggle link berikut <https://www.kaggle.com/uciml/breast-cancer-wisconsin-data>. Untuk cara load data ke dalam aplikasi orange data dapat dilihat pada subbab 1.1.2.

Anda dapat menggunakan widget “*Test and Score*” untuk melakukan pemodelan agar proses sampling data lebih mudah. Anda juga dapat menambahkan widget preprocess untuk mem preprocess data anda sebelum dilakukan proses training dan testing.



Jangan lupa mengklik dua kali garis antara widget Preprocess dan widget Test and Score untuk menghubungkan Preprocessor seperti gambar diatas.

### 2. Pemodelan dengan dataset “Heart Disease” (Nilai Tambah)

Mempraktekan kembali pemodelan sesuai dengan tutorial sebelumnya namun dengan dataset yang berbeda yaitu “Heart Disease” lalu **men screenshot canvas beserta hasil pemodelan (AUC, CA, F1, Precision dan Recall) yang telah anda lakukan.**

Data yang digunakan adalah data yang terdapat pada widget *Datasets* lalu tulis “Heart Disease” pada kolom search seperti dibawah ini :



## Datasets

heart Disease

Title	Size	Instances	Variables	Target	Tags
Heart Disease	23.5 KB	303	14	categorical	biology, medicine

Description

**Heart Disease** (1988), from [UCI ML Repository](#)

This data uses a subset of 14 attributes from the Cleveland database. The 'goal' field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence (values 1,2,3,4) from absence (value 0).

**See Also**

[How to Properly Test Models](#)  
[Minimum Cost of Misclassification](#)

**References**

Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J., Sandhu, S., Guppy, K., Lee, S., & Froelicher, V. (1989). International application of a new probability algorithm for the diagnosis of coronary artery disease. *American Journal of Cardiology*, 64, 304-310.

303