MLJ : Notes

Dainish Jabeen

January 11, 2023

1 General

MLJ is a machine learning toolbox for Julia, that wraps a large number of models and provides great tools such as resampling and evaluation.

Models are structs storing hyperparameters.

Listing 1: Basics

```
1
2       #Split data randomly with seed (rng)
3       y, X = unpack(df,==(:col);rng=123);
4
5       df |> pretty #Output pretty version
6
7       #Load model
8       VAR = @load model pkg=""
9
10      var = VAR() #Default parameters
11
12      #evaluate model with cross validation
13       via error measures
14
15      evaluate(var,X,y,resampling=CV(shuffle=true),measures=[rms])
```

1.1 Types

Each model has an expected variable type needed to train it:

target_scitype(model) provides type needed.

2 Machines

Used to store training outcomes.

Listing 2: Machines

```
1
2         mach = machine(model,X,y)
3
4         #70:30 partition giving an index vector
5         train, test = partition(eachindex(y),0.7)
6
7         fit!(mach,train)
8         yhat = predict(mach,X[text,:]);
9
10        #error rate
11        misclassification_rate(yhat, test)
```