Classification With Randomer Forests

Table of Contents

Randomer forest is a sparse oblique decision forest.

Here we demonstrate how to train, compute out of bag error, and predict on test data.

Load Iris Dataset

```
load fisheriris
X = meas;
Y = cellstr(num2str(grp2idx(species))); %Convert strings of names to strings of nu
classes = unique(Y);
```

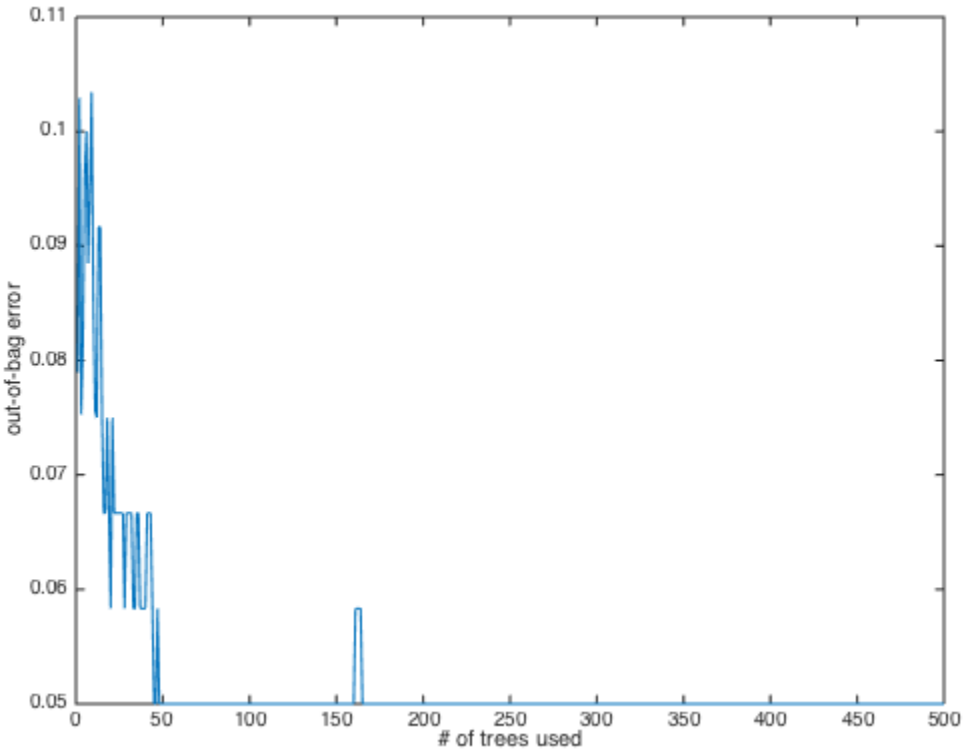
Train Randomer Forest

```
% Use 500 trees, very sparse Rademacher matrix (default) for random projections,
% sample 3 candidate projections at each split node, specify stratified,
% sampling, and connect to two parallel workers

nTrees = 500;
ProjectionMethod = 'sparse';
mtry = 3;
NWorkers = 2;
Stratified = true;

trainIdx = [1:40 51:90 101:140];
testIdx = setdiff(1:150,trainIdx);
Xtrain = X(trainIdx,:);
Ytrain = Y(trainIdx);
Xtest = X(testIdx,:);
Ytest = Y(testIdx);

RerF = rpclassificationforest(nTrees,Xtrain,Ytrain,'sparsemethod',ProjectionMethod
    'nvartosample',mtry,'NWorkers',NWorkers,'Stratified',Stratified);
```

Compute and plot out of bag error vs number of trees

```
oobError = oobpredict(RerF,Xtrain,Ytrain,'every');
plot(1:nTrees,oobError)
xlabel('# of trees used')
```

```
ylabel('out-of-bag error')
```



Predict class labels of test data points and compute test error

```
posteriors = rerf_classprob(RerF,Xtest,'last');
[~,classIdx] = max(posteriors,[],2);
Yhat = classes(classIdx);
testError = sum(strcmp(Ytest,Yhat))/length(Y);
fprintf('test error = %f',testError)
```

```
test error = 0.200000
```

Published with MATLAB® R2014b