Use of the Matlab function "sequentialfs" to perform feature selection

1 `inmodel = sequentialfs(fun,X,y)`

This function selects a subset of features from the data matrix *X* that best predict the data in *y* by sequentially selecting features until there is no improvement in prediction. More precisely, starting from an empty feature set, `sequentialfs` creates candidate feature subsets by sequentially adding each of the features not yet selected. For each candidate feature subset, `sequentialfs` performs 10-fold cross-validation by repeatedly calling `fun` with different training subsets of *X* and *y* (`XTRAIN` and `ytrain`) and test subsets of *X* and *y* (`XTEST` and `ytest`), as follows:

```
criterion = fun (XTRAIN, ytrain, XTEST, ytest)
```

`XTRAIN` and `XTEST` contain the data taken from the columns of *X* that correspond to the current candidate feature set.

`inmodel`=logical vector indicating which features are finally chosen

`fun`= function handle to a function that defines the criterion used to select features. 'fun' must return a scalar value.

X = data matrix (rows are observations, columns are variables or features)

y = column vector of response values or class labels for each observation in X

Observation: X and y must have the same number of rows.

2 `[inmodel, history] = sequentialfs(fun, X,y)`

This function returns information on which feature is chosen at each step. `history` is a scalar structure with the following fields:

`Crit`: a vector containing the criterion values computed at each step

`In`: a logical matrix in which row *i* indicates the features selected at step *i*

Example

```
x=randn(100,10);
t=randn(100,1);
opt=statset('display','iter');
[fs,history]=sequentialfs(@myfun,x,t,'cv','none','opt', opt,
'nfeatures', 3);


function perf = myfun(x,t)

%create a network
hiddenLayerSize=3;
net=fitnet(hiddenLayerSize);

xx=x'; tt=t';

% train the network
[net,tr]=train(net,xx,tt);

% test the network
y=net(xx);
perf=perform(net,tt,y);

end
```

Meaning of the parameters of sequentialfs:

- parameter 'cv': specifies the method used to compute the criterion for each candidate feature subset. "cv" means Cross-validation. When the value of 'cv' is 'none', `sequentialfs` calls the function `my_fun` without separating test and training sets;
- parameter 'opt': allows to set the possible options; in particular, when the value of 'opt' is opt, sequentialfs displays, for each step of the feature selection process, the feature (i.e., column) added and the value of the criterion. In order to achieve this, it is necessary to insert the following instruction before calling sequentialfs:

  ```
  opt=statset('display','iter');
  ```

- parameter 'nfeatures': specifies the number of features at which `sequentialfs` should stop. The parameter `fs` returned by `sequentialfs` includes exactly these features. The default value is empty, indicating that `sequentialfs` should stop when a local minimum of the criterion is found.