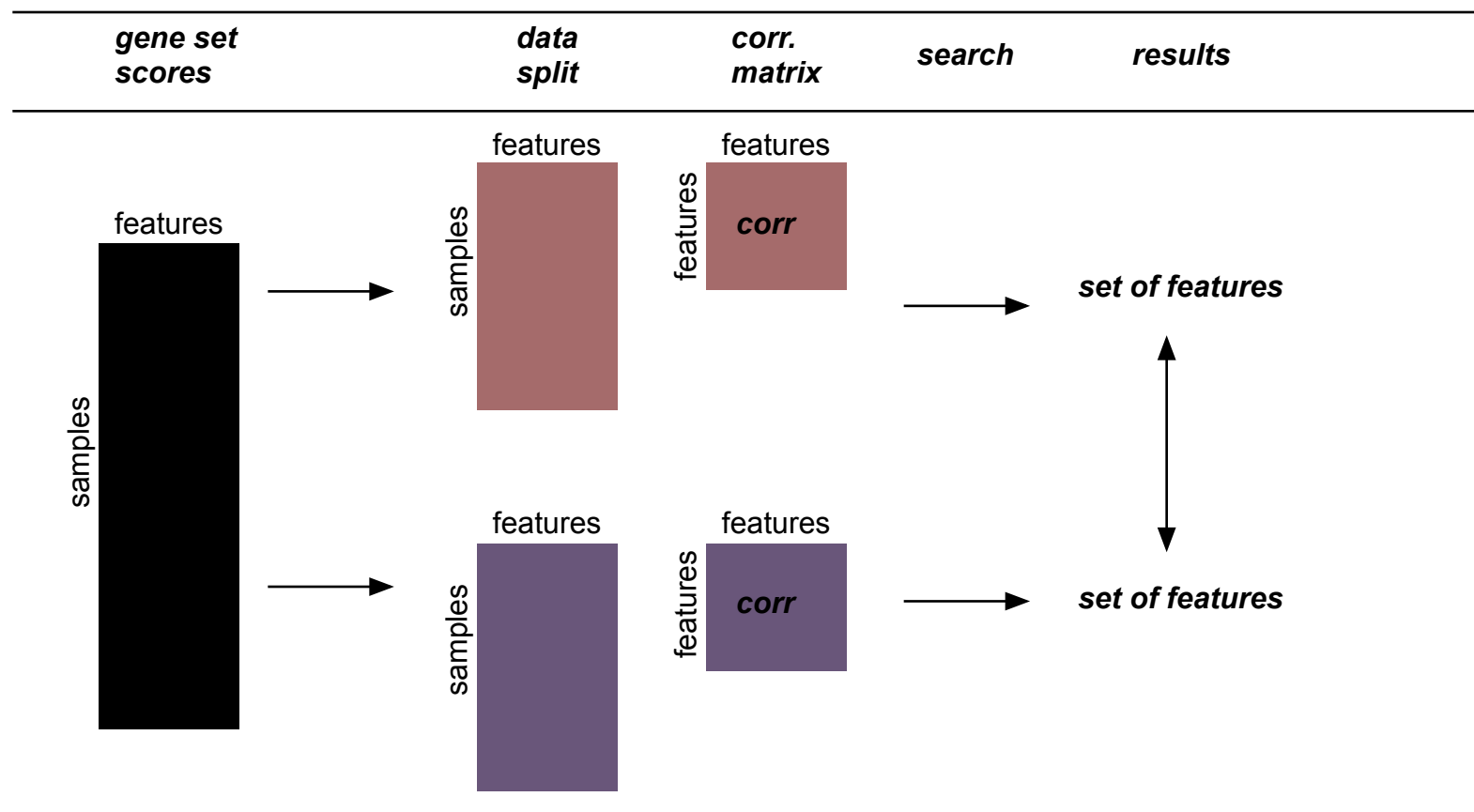


Feature selection via minimization of the average inter-set correlation.
A hill climbing implementation.



Algorithm:

Repeat *Reps* number of times. When done, return *k* items most often selected.

parameters

I : list of all features

D : data matrix, features in columns

M : correlation matrix of data

k : number of desired features

iter : number of iterations

init the result set

resultSet = sample *k* items from *I*

for 1 to *iter*:

x = sample feature that is not already selected,
with probability proportional to distance from result set.

i = sample element in *resultSet*,
probability proportional to average correlation with other result set elements.

try: replace element *i* in *resultSet* with *x*, compute average inter-set correlation
if the inter-set correlation has decreased, then keep *x*, otherwise discard *x*.