
```
$data = shift;
$strainlabels = shift;

#####
##Read data and labels from files
#####
open(IN, $data);
@data = ();
$i=0;
while( eof(IN) != 1){
    $l=<IN>;
    @a=split(/\s+/, $l);
    for(my $j=0; $j<@a; $j++){
        $data[$i][$j] = $a[$j];
    }
    $data[$i][scalar(@a)] = 1;
    $i++;
}
close(IN);

$rows = scalar(@data);
$ref = $data[0];
$cols = scalar(@$ref);

open(IN, $strainlabels);
@class = ();
@class_size = ();
while( eof(IN) != 1){
    $l=<IN>;
    @a=split(/\s+/, $l);
    $class[$a[1]] = $a[0];
    $class_size[$a[0]] = $class_size[$a[0]] + 1;
    if($class[$a[1]] == 0) { $class[$a[1]] = -1; }
}

#####
##Initialize w
#####
@w = ();
for(my $j=0; $j<$cols; $j++){
    $w[$j] = .02 * rand(1) - .01;
# $w[$j] = .0002 * rand(1) - .0001;
}
```

```

#####
##Gradient descent iteration
#####
$eta = .0001;
#$eta = .000001;
#$eta = .000000001;

for(my $k=0; $k<100000; $k++){

  #####Compute dellf#####
  @dellf = ();
  for(my $i=0; $i<$rows; $i++){
    if(defined($class[$i])){
      $dp = &dot_product(\@w, $data[$i]);
      for(my $j=0; $j<$cols; $j++){
        $dellf[$j] += ($class[$i]-$dp)*$data[$i][$j];
      }
    }
  }

  #####Update w#####
  for(my $j=0; $j<$cols; $j++){
    $w[$j] = $w[$j] + $eta*$dellf[$j];
  }

  $error = 0;
  #####Compute error#####
  for (my $i=0;$i < $rows;$i++){
    if(defined($class[$i])){
      $error += ($class[$i] - &dot_product(\@w, $data[$i]))**2;
    }
  }
  print "error = $error\n";
}

print "w = ";
$normw = 0;
for(my $j=0; $j<$cols-1; $j++){
  $normw += $w[$j]**2;
  print "$w[$j] ";
}
print "\n";
$normw = sqrt($normw);
print "||w||=$normw\n";
$d_origin = $w[scalar(@w)-1]/$normw;
print "distance to origin = $d_origin\n";

```

```

#####
#Prediction
#####
open(OUT, ">p_output");
for (my $i = 0 ; $i < $rows; $i++){
    if(!defined($class[$i])){
        $dp = &dot_product(\@w, $data[$i]);
        if($dp>0){
            print OUT "1 $i\n";
        }
        else{
            print OUT "0 $i\n";
        }
    }
}
close(OUT);

#####
##Sub routines
#####
sub dot_product{

    my $refw = $_[0];
    my $refx = $_[1];
    my $dp = 0;
    for(my $j=0; $j<$cols; $j++){
        $dp += $$refw[$j] * $$refx[$j];
    }
    return $dp;
}

```