

## 4006 - Remedial Forecast

#### Latin America - Mexico and Central America - 2007/2008

The Academic Development Department of the University of Hidalgo pretends to implement a remedial program to reduce the fail index in the Mathematics V course. With this purpose, the Department does a future performance forecast of the students with base on historical data. This data record includes the Id number (IN), average (A), number of study hours per week (SH), number of class hours per week (CH), and the result indicator obtained (R); when a student fails R = 0, otherwise R = 1. Students who obtain a failing forecast must take a remedial workshop on extra class hours.

The next table is an example of the historical results of students who have taken a Mathematics V courses before. The number of historical results is *HR*.

The characteristic vector of any student is represented by:

$$S_{IN} = \langle A_{IN}, SH_{IN}, CH_{IN} \rangle$$

, for example:

$$S_{735} = < 8, 13, 25 > , S_{677} = < 6, 10, 22 >$$

The Manhattan distance is given by:

$$Md_{(Si, Sj)} = |A_i - A_j| + |SH_i - SH_j| + |CH_i - CH_j|$$

Here is an example:

$$Md_{(8677,8512)} = |6 - 8| + |10 - 15| + |22 - 20| = 9$$

The next procedure is used in forecasts, with base on statistics, and the possibility that a new student in a mathematics V course could fail.

#### Step 1.

Measure the Manhattan distance between the characteristic vector of the student and the characteristic vectors of all the students in the historical table (without considering the result parameter R).

#### Step 2.

Sort the students table based on their distance; in case of tie, take the order as it appears in the original table. Once sorted, take the first k (an odd integer number), 1 - k - HR/2 to determine the value of

the more frequent indicator result and assign it to the student who the forecast is being done to, as seen in the example for the NC = 300 student

Step1.

Like k = 3, the indicator result forecasted is 1.

### Input

The Input to this problem will consist of a (non-empty) series of up to 50 data sets. Each data set will be formatted according to the following description, and there will be a blank space separating data sets.

The first data set contains historical data of up to 100 students who have taken this course before. The first number in each line is an integer for the Id Number, followed by three integers corresponding to the characteristic vector of that student.

The following data sets contain an integer in the first line, corresponding to the k value and in the lines to follow a table with the data of new students to whom the forecast has to be done.

### **Output**

The output consists of a list of pairs of integers that correspond to the students Id Number and the forecasted result, separated by a blank space.

## Sample Input

```
214 8 13 25 1
315 7 10 28 0
550 9 20 22 1
120 6 10 22 0
335 8 11 20 0
220 10 20 22 1
450 7 15 21 0
180 10 14 20 1
250 7 12 23 0
300 8 16 23 0
3
300 8 20 20
320 7 15 24
5
340 8 15 25
365 10 25 20
440 9 30 20
```

# **Sample Output**

```
300 1
320 0
340 0
365 1
440 1
```

Mexico and Central America 2007-2008