[ML] Subtracting Data
Problem ID: subtracting

Machine learning models are great, but sometimes they take a long time to train. To fix this, we've decide to delete most of our data! In this problem, your task is to select a subset of the data to remove while minimizing the impact on the performance of the resulting model trained on the data.

Specifically, for this problem you'll be given a training set as input and must output the indexes of k training samples to remove such that the resulting balanced accuracy score[1] when the model is tested on a hidden test set is maximized.

Input

Your program will receive input from standard input.

The first line will contain three space-separated integers n, m, and k.

n lines follow. The i-th line contains m space-separated integers $x_1, x_2, \ldots x_m$ representing the model's training data features of the i-th training data point.

Next, n lines follow. The i-th line contains one integer y_i representing the label for the i-th training data point.

Output

Your program should write to standard output.

Print k lines representing the (0-indexed) indexes of the training data points to remove from the training set.

Constraints

- $10^3 \leq n \leq 10^4$
- $2 \leq m \leq 8$
- $k < n$
- $0 \leq x_i \leq 10^3$
- $0 \leq y_i \leq 19$
- The training and test set are generated from the same distribution.

Scoring

Score is inversely proportional to the change in balanced accuracy score of the model. A baseline random solution has been used to determine the minimum balanced accuracy score to score any points (it is intended that a uniformly random solution scores no points).

Sample Explanation

A simple input and output, as well as the checker used for this problem are attached on the challenge website under the statement for this problem. These are intended to help test solutions locally. The sample input is not part of the real test set.

---

[1]https://scikit-learn.org/stable/modules/generated/sklearn.metrics.balanced_accuracy_score.html