The scenario

Many customers only make an order in an online shop once. There are many reasons why they do not make another order.

Online dealers try to counteract this using appropriate customer loyalty measures. For example, a tried-and-tested method is to hand out vouchers sometime after an order to encourage the customer to make a follow-on purchase. But from an economic perspective sending vouchers to all customers is not a good solution because customers may make a second purchase without the incentive.

The task

Using the existing characteristics of a customer‘s initial order, such as order quantity per type of goods, title and delivery weight, a decision must be made on whether to send a voucher worth € 10.00. The customers who receive a voucher should be those who would not have decided to re-order by themselves.

As shown by empirical results, it should be assumed that the voucher initiates a purchase with an average order value of € 40.00 in 10% of those not making a purchase. If a voucher is sent to a customer that would have re-purchased anyway this results in a € 10.00 loss for the dealer. The aim is therefore to maximize revenue by sending the vouchers to selected customers.

Using the data in the train.txt file a model is to be generated that is then used to decide which customers from the class.txt file will receive a voucher.

Form of the results provided

The results are to be submitted in a csv file with two columns delimited by a comma. The first column should contain client id, the second one the decision about voucher. Few example rows:

3,1

13,0

20,1

23,1

The voucher field must contain a “0” for “no voucher” or a “1” for “send voucher”. Unfortunately results that do not contain all the customer numbers from the data that is to be assessed cannot be graded.

This list must be sent back in the form of a text file (ASCII, no Unicode) with no header.

Assessing the results

If a customer is sent a voucher but would have repurchased anyway, this results in a € 10.00 loss in revenue.

In contrast if a customer who would not have repurchased is sent a voucher this increases revenue on average by € 3.00.

Predicted value		True value
	0	1
0	0	0
1	3	-10

Following this matrix, the task is to maximize the score on the test set computed as:

3 * (number of customers who wouldn’t repurchase without contact, but our model decides to contact them) – 10 * (number of customers who would repurchase without contact and our model decides to contact them)