

Задание «Модельная задача»

1. Скачиваете задачу по адресу
(номер вашей задачи соответствует номеру в списке группы на сайте
<http://www.machinelearning.ru>)

<http://AlexanderDyakov.narod.ru/mat317n1train.mat>
<http://AlexanderDyakov.narod.ru/mat317n2train.mat>
<http://AlexanderDyakov.narod.ru/mat317n3train.mat>
<http://AlexanderDyakov.narod.ru/mat317n4train.mat>
<http://AlexanderDyakov.narod.ru/mat317n5train.mat>
<http://AlexanderDyakov.narod.ru/mat317n6train.mat>
<http://AlexanderDyakov.narod.ru/mat317n7train.mat>
<http://AlexanderDyakov.narod.ru/mat317n8train.mat>
<http://AlexanderDyakov.narod.ru/mat317n9train.mat>
<http://AlexanderDyakov.narod.ru/mat317n10train.mat>
<http://AlexanderDyakov.narod.ru/mat317n11train.mat>
<http://AlexanderDyakov.narod.ru/mat317n12train.mat>
<http://AlexanderDyakov.narod.ru/mat317n13train.mat>
<http://AlexanderDyakov.narod.ru/mat317n14train.mat>
<http://AlexanderDyakov.narod.ru/mat317n15train.mat>
<http://AlexanderDyakov.narod.ru/mat317n16train.mat>
<http://AlexanderDyakov.narod.ru/mat317n17train.mat>
<http://AlexanderDyakov.narod.ru/mat317n18train.mat>
<http://AlexanderDyakov.narod.ru/mat317n19train.mat>
<http://AlexanderDyakov.narod.ru/mat317n20train.mat>
<http://AlexanderDyakov.narod.ru/mat317n21train.mat>
<http://AlexanderDyakov.narod.ru/mat317n22train.mat>
<http://AlexanderDyakov.narod.ru/mat317n23train.mat>
<http://AlexanderDyakov.narod.ru/mat317n24train.mat>

2. В MATLAB-файле содержится признаковая матрица X (2000 объектов и 50 признаков) и вектор классификации Y . В задаче два класса (+1 и -1), если объекту поставлена метка 0, то его класс не известен и необходимо его установить.
3. Решаете задачу как можно точнее (мера качества – процент правильных ответов). Предоставляете ответ – вектор классификации Y_2 (по сути, это вектор Y , в котором нули заменены на восстановленные Вами значения). Также предоставляете признаки, от которых, по Вашему мнению, зависит ответ (в задаче есть фиктивные признаки – нужно их отгадать). Для оценки качества определения существенных признаков используется F-мера.

К 7 мая (срок ориентировочный ± 1 неделя! Будем уточнять 25 апреля) нужно подготовить презентацию на тему, как Вы решали задачу и какие результаты получили (время выступления ~ 5 минут). Доклады будут сделаны перед группой и оценены преподавателем. Сами ответы (классификация и перечень существенных признаков) будут собираться позднее (форматы ответов также будут уточняться).