

В данном задании вам необходимо предсказать с помощью алгоритмов из пакета Orange, выжил ли пассажир “Титаника”, в зависимости от информации о нём. Данные находятся в кубике “Datasets” во вкладке с названием “Titanic”.

Один из основных показателей качества решения задачи классификации — доля правильных ответов (на английском обычно называется accuracy). Как следует из названия, этот показатель равен доле объектов, на которых модель выдала ответ, совпадающий с истинным. В нашей задаче эта метрика будет показывать, для какой доли пассажиров из выборки модель угадала, выжил он или нет.

В качестве алгоритма машинного обучения используйте логистическую регрессию (Logistic Regression) и дерево решений (Tree). Качество полученного алгоритма необходимо измерить на обучающей выборке (Test on train data) через долю верных ответов (в таблице кубика Test & Score она записана в столбце CA).

Подсказка: у алгоритмов существуют настраиваемые параметры - их использование поможет получить лучшее качество.

Ниже напишите краткий отчёт о результатах. Запишите, какие модели вы пробовали обучать и какие значения параметров в нихставляли. Для каждого варианта укажите, какое значение доли верных ответов у вас получилось. Также для каждого варианта сделайте скриншот окна Orange и прикрепите его к отчёту. В конце отчёта укажите лучшее значение качества, которое удалось достичь.

Подсказка: у алгоритмов существуют настраиваемые параметры -- их использование поможет получить лучшее качество.