Прогнозирование качественных переменных в R (24 часа)

Очень часто на практике приходится сталкиваться с прогнозированием качественных переменных, т.е. переменных, принимающих ограниченное количество значений. Курс посвящен обзору современных компьютерных алгоритмов решения данной задачи. Кратко излагаются теоретические основы методов, основной упор делается на практическое применение. Для курса слушателю потребуется доступ к открытому программному обеспечению (среда R и оболочка R-studio). Это может быть компьютерный класс или личный компьютер.

Темы курса:

\* Введение в статистический пакет R

\* Визуализация качественных переменных.

\* Логит и пробит-модели для переменных принимающих бинарные значения. Логит и пробит модели для упорядоченных качественных переменных. Предельные эффекты.

\* Линейный, квадратичный и смешанный дискриминантный анализ.

\* Метод опорных векторов. Кросс-валидация для выбора оптимальных параметров модели.

\* Алгоритм случайного леса (random forest).

\* Использование кривых ROC для сравнения моделей.