

Домашка «Титаник»

Нужно зарегистрироваться на сайте www.kaggle.com и принять участие в конкурсе «Titanic: Machine Learning from Disaster». Крайний срок сдачи отчёта: в ночь с 14 на 15 апреля 2013 года.

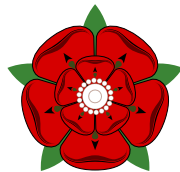


Версия Белой Розы

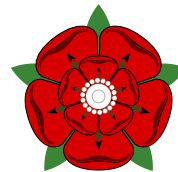


1. Домашнее задание можно делать в одиночку или группой из двух человек.
2. А можно всё-таки группой из трёх человек? Нет :)
3. Письменный отчёт должен содержать как-минимум:
 - (a) Логин группы
 - (b) Графический анализ имеющихся данных
 - (c) Результаты оценивания logit и probit моделей
 - (d) Графический анализ logit и probit моделей
 - (e) «Если бы я был пассажиром Титаника, то я спасся бы с вероятностью...».

С помощью logit и probit моделей необходимо построить 95%-ый доверительный интервал для вероятности спасения каждого из участников группы, сдающей домашку. Пол и возраст взять фактические, а остальные объясняющие переменные — по своему желанию.



Версия Белой Розы



1. Домашнее задание можно делать только в одиночку :)
2. Нет, нельзя :)
3. Письменный отчёт должен содержать как-минимум:
 - (a) Логин
 - (b) Графический анализ имеющихся данных
 - (c) Результаты оценивания logit и probit моделей
 - (d) Прогнозирование с использованием Random Forest
 - (e) Прогнозирование с использованием метода опорных векторов (SVM)
 - (f) Графический анализ оценённых моделей
 - (g) «Если бы я был пассажиром Титаника, то я спасся бы с вероятностью...».

С помощью логит и пробит моделей необходимо построить 95%-ый доверительный интервал для своей вероятности спасения. Для Random Forest требуется только точечная оценка вероятности спасения. Пол и возраст взять фактические, а остальные объясняющие переменные — по своему желанию.