

ПРИЛОЖЕНИЕ «БАЛАНС ВНИМАНИЯ» ДЛЯ САМОСТОЯТЕЛЬНОГО ПОЛУЧЕНИЯ НАБОРОВ ДАННЫХ ДЛЯ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ БИОЛОГИЧЕСКИХ ПРОЦЕССОВ

Шербина Д.Н.

ФГАОУ ВО «Южный федеральный университет»,

Академия биологии и биотехнологии

E-mail: dnsherbina@sfnedu.ru

При изучении методов математической статистики нужны правдоподобные примеры для демонстрации особенностей разных методов. Студентам естественнонаучного направления не близки примеры с искусственно синтезированными наборами данных, на которых математики зачастую демонстрируют свойства разных алгоритмов классификации и кластерного анализа. Развитые пакеты для научных вычислений (R, seaborn, panda) в своем составе включают популярные наборы данных. В частности, набор 'iris' с измерениями ширины и длины лепестков цветов ириса трех близких видов, близок биологическому сердцу и при этом наглядно демонстрирует сильные и слабые стороны разных подходов к многомерному анализу, изоморфного отображения, визуализации. С такими «знаменитыми» наборами данных следует знакомить студентов на практических работах по статистике, однако, самостоятельные домашние задания лучше всего выполнять на уникальных самостоятельно полученных наборах данных.

Требования к процедуре самостоятельного получения уникальных наборов данных могут быть такие:

1. Личная заинтересованность за счет получения собственных данных.
2. Данные должны иметь серьезную научную интерпретацию, но не требовать специального дорогостоящего оборудования.
3. Возможность многократного повторения процедуры без утраты интерпретационной ценности, или, другими словами, всякий раз при повторении процедуры данные будут новыми.
4. Количество и сложность данных в одном наборе должны быть достаточными для проведения нескольких типов статистического анализа.
5. Получение и хранение данных должны проводиться с помощью открытых ресурсов с сохранением долговременного доступа к своим данным.

Всеми этим требованиям удовлетворяет приложение для экспресс-диагностики индивидуальных свойств внимания, исходно разработанное в

рамках выполнения научно-исследовательской работы <http://balatte.stireac.com/>. Приложение обеспечивает идентификацию пользователя, онлайн тестирование, обработку данных по отдельным записям с наглядными иллюстрацией результатов и возможность экспорта исходных данных в табличном виде. Тест основан на известной психологической парадигме сенсомоторных реакций.

Студентам второго курса было предложено воспользоваться приложением "Баланс внимания" для получения индивидуальных наборов данных. Впоследствии эти наборы данных были использованы для развития практических навыков по удалению выбросов, созданию очищенных сбалансированных таблиц данных, многофакторному вариационному анализу, выделению линейных и нелинейных трендов, подбору характеристик распределений.