Руководство пользователя к программе для построения модели множественной регрессии с Web реализацией пользовательского интерфейса

Разработчики:

Подчезерцев А.Е.

Солодянкин А.А.

Оглавление

[Описание решаемой задачи 3](#_Toc515392011)

[Структура исходных файлов 4](#_Toc515392012)

[Описание интерфейса программы 5](#_Toc515392013)

[Структура каталогов приложения 8](#_Toc515392014)

[Программные требования 9](#_Toc515392015)

[Аппаратные требования 10](#_Toc515392016)

# Описание решаемой задачи

Данная программа предназначена для вычисления параметров регрессии для произвольного набора данных – датасета.

# Структура исходных файлов

Исходные данные, загружаемые в программу, могут иметь любое имя и расширение и находиться в любом месте на компьютере. В первой строчке должны находиться названия столбцов, указывающих какие данные хранятся в нем. Названия не должны содержать пробелов и символов разделителей. В последующих строчках находятся сами данные, сепированные разделителем. В каждой строчке количество данных должно совпадать с общим числом столбцов. Формат разделителя для дробных чисел – точка. Символ разделитель данных выбирает пользователь, рекомендуется ставить точку с запятой. Формат конца строки – любой – Unix, Mac или Windows, кодировка - Unicode.

После загрузки файла через форму данные помещаются в файл Work/Data/.temp, после чего файл проверяется и преобразуется с форматом разделителя запятая, конец строки – Windows и сохраняется с именем равным количеству секунд, прошедших с начала эпохи Unix, и расширением “.csv”.

# Описание интерфейса программы

Для запуска приложения необходимо запустить файл run.py из папки Scripts. После инициализации приложения откроется браузер по адресу <http://127.0.0.1>:8080 и отобразится начальное окно (Рисунок 1).

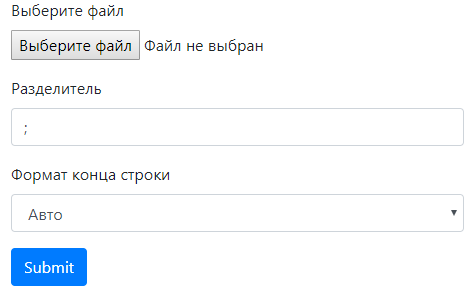


Рисунок 1. Форма загрузки датасета

Пользователю необходимо выбрать исходный файл, удовлетворяющий требованиям, выбрать разделитель и формат конца строки, после чего нажать кнопку Submit для загрузки файла на сервер. В случае успеха пользователь будет переадресован на страницу обработки (Рисунок 2), иначе произойдет перезагрузка.

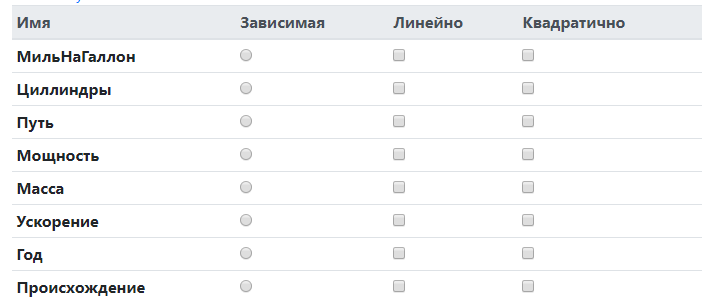


Рисунок 2. Форма обработки данных

На данной форме необходимо выбрать зависимую переменную, относительно которой будут производиться вычисления, и зависимые с параметром вхождения – линейно или квадратично (Рисунок 3) и нажать кнопку обработать для вычисления.

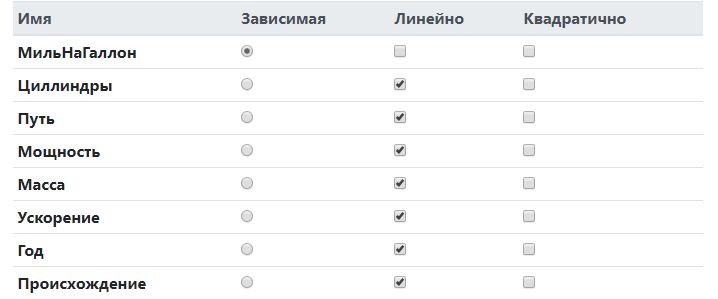


Рисунок 3. Параметры регрессии

В случае успеха будет выведена информация с результатами регрессии (Рисунок 4) или ошибкой (Рисунок 5).

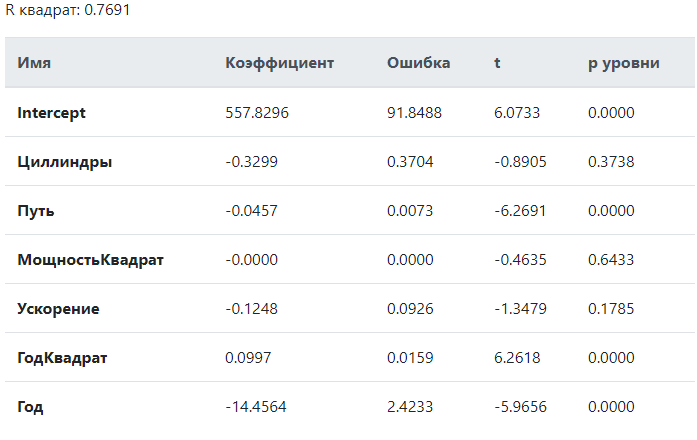


Рисунок 4. Результат регрессии



Рисунок 5. Сообщение об ошибке

# Структура каталогов приложения

Основной каталог решения – Work, внутри которого находятся следующие подкаталоги:

* Data – предобработанные датасеты;
* Graphics – шаблоны страниц;
* Library – библиотека универсальных функций;
* Notes – руководства пользователя и разработчика;
* Output – выходные данные приложения;
* Scripts – скрипты для запуска приложения.

# Программные требования

Для корректной работы приложения необходимо установить следующие библиотеки:

* Django – web framework;
* NumPy – математические функции;
* Pandas – обработка и анализ данных;
* Statsmodels – статистические исследования;

# Аппаратные требования

Любая система (32- или 64- битная/ARM и другие), на которую можно установить Python 3.5+ и необходимые библиотеки (https://www.python.org/downloads/)