

Распознавание образов. Семинар №1

Трофимов Михаил

17 сентября 2014

Инструментарий

Формально ограничений на выбор инструментария нет, но рекомендуется использовать следующий программный стек:

- ▶ Python 2.7+, IPython Notebook
- ▶ NumPy
- ▶ SciPy
- ▶ Pandas
- ▶ Matplotlib
- ▶ scikit-learn (sklearn)

Устанавливается примерно так

```
sudo apt-get install python-setuptools
```

```
sudo apt-get install python-matplotlib
```

```
easy_install ipython numpy scipy pandas scikit-learn
```

Немного подробнее

NumPy

NumPy — это расширение языка Python, добавляющее поддержку больших многомерных массивов и матриц, вместе с большой библиотекой высокоуровневых математических функций для операций с этими массивами. Свободная альтернатива MATLAB.

sklearn

scikit-learn (или `scikits.learn`) - свободная библиотека для машинного обучения. Содержит реализации различных алгоритмов классификации, регрессии и кластеризации. Опирается на использование NumPy/SciPy.

Matplotlib

Matplotlib — библиотека для визуализации данных на языке Python . Matplotlib начинался с подражания графическим командам MATLAB, но является независимым от него проектом.

Где взять примеры?

У каждого пакета есть хорошая документация.
Если не помогло - поисковик / StackOverflow.

Несколько полезных ссылок

<http://habrahabr.ru/post/121031/> - введение в NumPy

<http://habrahabr.ru/post/202090/> - пример использования
Pandas

<http://scikit-learn.org/stable/tutorial/index.html>

http://scikit-learn.org/stable/auto_examples/index.html

Что нужно сделать?

- ▶ Реализовать алгоритм обучения персептрона
- ▶ Реализовать алгоритм обучения логистической регрессии
- ▶ Нарисовать разделяющие поверхности
- ▶ *В каком случае обучение персептрона не сойдется? А у логистической регрессии?
- ▶ *Как с помощью логистической регрессии строить квадратичную разделяющую поверхность?
- ▶ *Как с помощью бинарного классификатора решать задачу классификации с тремя классами?

Как сдавать?

В виде кода(скрипты, ipython notebooks) и/или отчета.
Важно наличие воспроизводимого кода и прикрепленных графиков (если требуется). Сдать можно

- ▶ Лично
- ▶ e-mail: mikhail.trofimov@phystech.edu