

Введение в машинное обучение

ВВОДНАЯ ЛЕКЦИЯ

Дьяконов А.Г.

**Московский государственный университет
имени М.В. Ломоносова (Москва, Россия)**

Лектор

Дьяконов Александр Геннадьевич

**профессор кафедры
«математических методов прогнозирования»**



Для кого

Бакалавры младших курсов (1-2) ВМК МГУ

Планы

1 поток	2 поток		3 поток
		ММП	
если пока не слушали Техносферу, курсеру и т.п.			будет курс МОиАД

Специфика

изучаем «с нуля»

Что будет – общее введение

Объяснение области

Наука о данных

Анализ данных

Машинное обучение

Большие данные

Искусственный интеллект

...

Примеры задач

Классификация

Регрессия

Кластеризация

Рекомендации

Поиск аномалий

...

Объяснение терминов

Переобучение

Недообучение

Сложность алгоритмов

Регуляризация

Модель

Параметры

Гиперпараметры

Скользящий контроль

Ансамблирование

Правило Байеса

Разделяющая поверхность

Метаалгоритм

...

Что будет – инструментарий



Язык программирования Python

<https://www.python.org/>



**Библиотека для матричных
вычислений и линейной алгебры**

<http://www.numpy.org/>



Библиотека для научных вычислений

<https://www.scipy.org/>



Библиотека для визуализации

<https://matplotlib.org/>



Библиотека для машинного обучения



<http://scikit-learn.org/>

Совет по инструментарию



**научный дистрибутив
Anaconda Python
от Continuum**

<https://www.anaconda.com/download/>

Python 3.6 version *	Python 2.7 version *
 Download	 Download
64-Bit Graphical Installer (631 MB) ?	64-Bit Graphical Installer (564 MB) ?
32-Bit Graphical Installer (506 MB)	32-Bit Graphical Installer (443 MB)

Совет по инструментарию



Python, R, Julia, Scala, F#

<http://jupyter.org/>



<https://www.jetbrains.com/pycharm/>

эволюция IPython Notebook

**для создания и обмена
«ноутбуками»:**

- код
- полнотекстовые комментарии
- уравнения
- визуализация

**интегрированная среда
разработки для языка
программирования Python**

Basic Numerical Integration: the Trapezoid Rule

A simple illustration of the trapezoid rule for definite integration:

$$\int_a^b f(x) dx \approx \frac{1}{2} \sum_{k=1}^N (x_k - x_{k-1}) (f(x_k) + f(x_{k-1})).$$

First, we define a simple function and sample it between 0 and 10 at 200 points

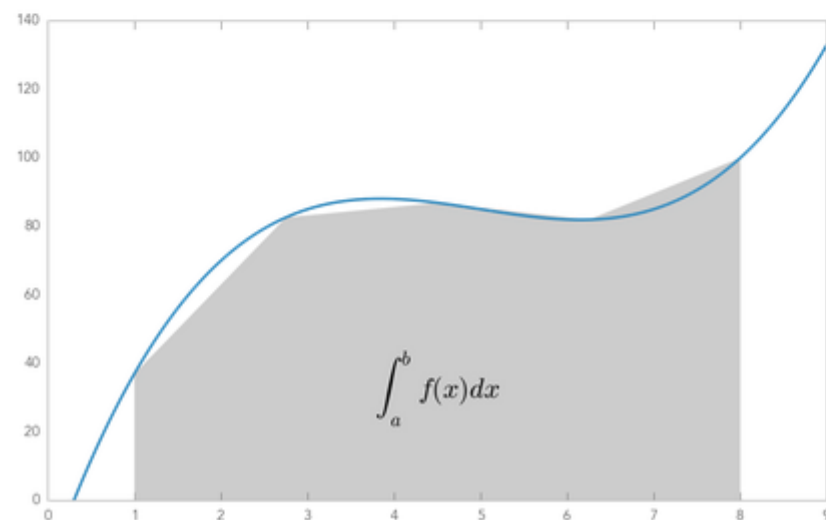
```
In [1]: %matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
```

```
In [2]: def f(x):
        return (x-3)*(x-5)*(x-7)+85

x = np.linspace(0, 10, 200)
y = f(x)
```

Choose a region to integrate over and take only a few points in that region

```
In [4]: plt.plot(x, y, lw=2)
plt.axis([0, 9, 0, 140])
plt.fill_between(xint, 0, yint, facecolor='gray', alpha=0.4)
plt.text(0.5*(a+b), 30, r"$\int_a^b f(x)dx$", horizontalalignment='center', fontsize=20);
```



Цели

Дать представление о машинном обучении

Сделать «практический курс»

Реклама деятельности кафедры

Подготовить к поступлению на ММП
перечисленный инструментарий – основной в бакалавриате

Важно: не способ попасть на кафедру

Лекции

раньше – 1 раз в 2 недели

(не сильно отвлекают, есть время на самостоятельную подготовку)

сейчас – 1 раз в неделю

пятница, 18:30

Советы на будущее

Походить на разные спецкурсы / спецсеминары

Если понравится, продолжить изучение ML

Можно уже начинать учить Питон...
смотрите все приведённые ссылки

Есть кафедры, специализирующиеся в ML

Есть программы обучения

<https://ozonmasters.ru/>

Техносфера

ШАД

MADE

Что надо в ближайшее время

Проверить запись на курс

Канал в телеграмме

ссылка выслана всем зарегистрированным

Экзамен

оценка по сдаче заданий и тестов

страницы курса

[http://www.machinelearning.ru/wiki/index.php?title=Введение в машинное обучение](http://www.machinelearning.ru/wiki/index.php?title=Введение_в_машинное_обучение)

<https://github.com/Dyakonov/IML>