

Per.5: Don't assume that low-level code is necessarily faster than high-level code





ОНЛАЙН-ОБРАЗОВАНИЕ



ML

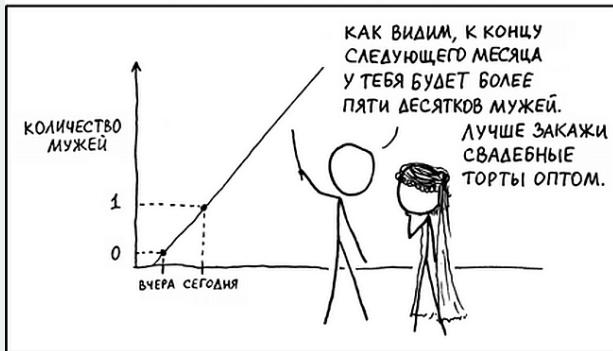
Регрессия

Дмитрий Шебордаев

Ведущий специалист



МОЁ ХОББИ: ЭКСТРАПОЛИРОВАТЬ



Сегодня

- ▶ Преобразование простого текста
- ▶ Бинарная и многоклассовая классификация
- ▶ Кластеризация и тематическое моделирование
- ▶ Регрессия
- ▶ Понижение размерности
- ▶ Рекомендации



Преобразование простого текста

- ▶ How to format my hard disk.
- ▶ Hard disk format problems.



Преобразование простого текста

- ▶ How to format my hard disk.
- ▶ Hard disk format problems.
- ▶ disk format hard how my problems to



Преобразование простого текста

- ▶ How to format my hard disk.
- ▶ Hard disk format problems.
- ▶ disk format hard how my problems to
- ▶ How to format my hard disk.
- ▶ 1 1 1 1 1 0 1



Преобразование простого текста

- ▶ How to format my hard disk.
- ▶ Hard disk format problems.
- ▶ disk format hard how my problems to
- ▶ How to format my hard disk.
- ▶ 1 1 1 1 1 0 1
- ▶ Hard disk format problems.
- ▶ 1 1 1 0 0 1 0



Преобразование простого текста

- ▶ How to format my hard disk.
- ▶ Hard disk format problems.
- ▶ disk format hard how my problems to
- ▶ How to format my hard disk.
- ▶ 1 1 1 1 1 0 1
- ▶ Hard disk format problems.
- ▶ 1 1 1 0 0 1 0
- ▶ униграммы, биграммы, триграммы.



Бинарная и многоклассовая классификация

- ▶ Входит и не входит
- ▶ Входит в один из нескольких
- ▶ Входит и не входит (входит в другой?)



Кластеризация и тематическое моделирование

- ▶ Попадает только в одну группу
- ▶ Попадает разные группы



Регрессия

- ▶ Прогнозирование
- ▶ Линейная
- ▶ Линейка и резинки
- ▶ Мягче или жестче - переобучение
- ▶ За какой период? Внешние факторы



Понижение размерности

- ▶ Garbage in, garbage out (GIGO)
- ▶ Корреляция



Рекомендации

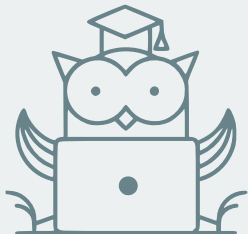
- ▶ Сбор данных
- ▶ Кластеризация документов
- ▶ Кластеризация пользователей
- ▶ Как определить что рекомендовать?
- ▶ Идеальные рекомендации или грязь полезна



ДЗ

- ▶ Набор данных в студию





Спасибо за
внимание!

