

Разработка программного обеспечения на языке Python

[Обзорная панель](#) ▶ [Мои курсы](#) ▶ [Разработка ПО на языке Python](#) ▶ [Анализ данных и машинное обучение](#) ▶

[Лекция 3. Постановка задачи машинного обучения](#)

Лекция 3. Постановка задачи машинного обучения

Посмотрите видеоуроки и ответьте на контрольные вопросы после лекции

Плюсы и минусы машинного обучения

Реклама 1 из 2

Алгоритмы машинного обучения применимы к задачам из различных предметных областей. Они нашли широкое применение при решении задач в медицине, банковском деле, промышленности и многих других сферах. Рассмотрим основные **преимущества** такого подхода.

Во-вторых – автоматизация. Во время обучения и по мере накопления информации модель способна обнаруживать новые шаблоны. Такие модели могут быть встроены в автоматизированные рабочие процессы. Следующее преимущество – это скорость и реагирование в реальном времени. Модели машинного обучения позволяют получить ответ через доли секунд после поступления новой информации. Четвертое преимущество – это возможность настройки. Многие задачи в предметной области можно решить методами машинного обучения. Модели строятся на основе исторически накопленных данных и допускают настройку под любую систему правил. И, наконец, отметим масштабируемость моделей машинного обучения. Модель легко приспособляется и к увеличивающимся объёмам данных. Некоторые алгоритмы можно использовать для обработки множества данных, используя облачные вычислительные ресурсы.

Очевидно, что алгоритмы машинного обучения обладают рядом **недостатков**. В основе алгоритмов лежат данные. По разным оценкам, специалисты тратят на подготовку данных до восьмидесяти процентов времени, которое затрачено на решение всей задачи. Фактически, **задача специалиста по машинному обучению** – это **сформулировать задачу** так, чтобы к ней можно было **применить алгоритмы анализа данных**. При этом полученные результаты должны иметь практическую ценность и быть измеримыми. Вторая сложность связана с выбором **оптимального набора признаков**. Они должны максимально подробно описывать интересующие свойства объектов. При этом, описание не должно быть избыточным. Есть известная сложность, с которой встречаются те, кто занимается созданием моделей. Так называемое “проклятие данных”. Когда модель идеально работает для обучающей выборки, но для тестовых значений, которые отсутствуют в обучающей выборке, даёт недостоверные результаты. Причина может быть в переобучении.

[Вопросы](#)

ПРЕДЫДУЩИЙ ЭЛЕМЕНТ КУРСА

[◀ Задание 3. Визуальный анализ данных](#)[Перейти на...](#)

СЛЕДУЮЩИЙ ЭЛЕМЕНТ КУРСА

[Анализ данных - Практическое занятие 3 ▶](#)

© 2010-2023 Центр обучающих систем
Сибирского федерального университета, sfu-kras.ru

Разработано на платформе moodle
Beta-version (3.9.1.5.w3)

[Политика конфиденциальности](#)[Соглашение о Персональных данных](#)[Политика допустимого использования](#)

Контакты +7(391) 206-27-05
info-ms@sfu-kras.ru

[Скачать мобильное приложение](#)[Инструкции по работе в системе](#)