

Машинное обучение

Гейдар Теймурович

НИУ ВШЭ

2025

Машинное обучение (англ., Machine learning) - это наука, изучающая алгоритмы, автоматически улучшающиеся благодаря опыту.

Зачем нужно разбираться в предмете?

Зачем нужно разбираться в предмете?

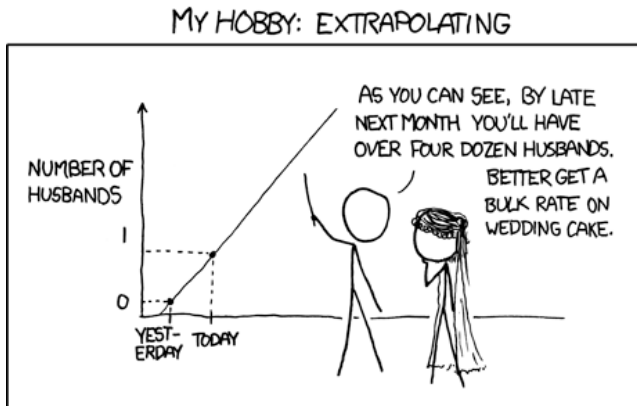


Рис.: <https://xkcd.com/605>

Формула оценки

В курсе предусмотрены домашние задания (в количестве 6 штук) как ключевая форма контроля знаний:

Формула оценки

В курсе предусмотрены домашние задания (в количестве 6 штук) как ключевая форма контроля знаний:

$S = \text{round}(\frac{1}{6} \cdot \sum_{i=1}^6 HA_i)$, где:

- S — оценка за курс по 100-балльной шкале
- HA_i — оценка за домашние задания по 100-балльной шкале

Округление арифметическое.

Формула оценки

В курсе предусмотрены домашние задания (в количестве 6 штук) как ключевая форма контроля знаний:

$S = \text{round}(\frac{1}{6} \cdot \sum_{i=1}^6 HA_i)$, где:

- S — оценка за курс по 100-балльной шкале
- HA_i — оценка за домашние задания по 100-балльной шкале

Округление арифметическое.

Важно!

Для зачета необходимо набрать не менее 60 баллов ($S \geq 60$)

Правила сдачи заданий (нюансы) - I

Для каждой домашней работы обозначаются мягкие и жесткие дедлайны. За каждый день просрочки после мягкого дедлайна снимается **20%** от оценки.

Правила сдачи заданий (нюансы) - I

Для каждой домашней работы обозначаются мягкие и жесткие дедлайны. За каждый день просрочки после мягкого дедлайна снимается **20%** от оценки.

е. г. сдаете домашнюю работу на 3 дня позднее - штраф 60%.

Правила сдачи заданий (нюансы) - I

Для каждой домашней работы обозначаются мягкие и жесткие дедлайны. За каждый день просрочки после мягкого дедлайна снимается **20%** от оценки.

е. г. сдаете домашнюю работу на 3 дня позднее - штраф 60%.

После жёсткого дедлайна работы не принимаются. Даже при опоздании на одну секунду. Сдавайте заранее.

Правила сдачи заданий (нюансы) – I

Для каждой домашней работы обозначаются мягкие и жесткие дедлайны. За каждый день просрочки после мягкого дедлайна снимается **20%** от оценки.

е. г. сдаете домашнюю работу на 3 дня позднее - штраф 60%.

После жёсткого дедлайна работы не принимаются. Даже при опоздании на одну секунду. Сдавайте заранее. Студент может **1 раз** сдать домашнее задание после мягкого дедлайна (но до жёсткого) без штрафов.

При наличии уважительной причины дедлайн по домашнему заданию может быть перенесён на количество дней, равное продолжительности уважительной причины.

Правила сдачи заданий (нюансы) – II

В случае использования больших языковых моделей¹:

- В chunk (ячейку) выше кода, созданного LLM, прикрепляйте промпт, который использовался для генерации.
- Отдельно опишите как подбирали промпты, какие заметили преимущества и недостатки GenAI для данного задания.
- За решение с указанным промптом - штраф 40% для конкретного задания, который может быть пересмотрен в сторону увеличения в следующих случаях:
 - 1 использован ответ LLM без указания промпта (штраф 100%)
 - 2 решение избыточно и, или написано неоптимально (использование magic команд без необходимости, использование циклов в тех случаях, когда операцию можно совершить при помощи инструментов библиотек, etc) (штраф 50%)

¹large language models (LLMs): ChatGPT, GigaChat, Qwen, etc

Правила сдачи заданий (нюансы) – III

- Можно использовать любые свободные источники с **обязательным** указанием ссылки на них.
- Плагиат не допускается. При обнаружении случаев списывания, 0 за работу выставляется всем участникам нарушения, даже если можно установить, кто у кого списал.
- Мы в любом случае оставляем за собой право пригласить студента для защиты своего ДЗ, если заподозрим плагиат.

План курса

- 1 Введение в machine learning;
- 2 Линейные регрессии - I;
- 3 Линейные регрессии - II;
- 4 Регуляризация. Классификация - I;
- 5 Классификация - II;
- 6 Логистическая регрессия, SVM;
- 7 Нелинейная, многоклассовая классификация;
- 8 Снижение размерности;
- 9 Кластеризация;
- 10 Решающие деревья;
- 11 Рекомендательные системы;
- 12 Ансамбли;
- 13 Бустинг;
- 14 Обработка естественного языка.

Python - VHLL (very high-level programming language) язык программирования.

Python - VHLL (very high-level programming language) язык программирования.

Как работать с Python?

- Терминал

Python - VHLL (very high-level programming language) язык программирования.

Как работать с Python?

- Терминал
- VIM (<https://www.vim.org/>)

Python - VHLL (very high-level programming language) язык программирования.

Как работать с Python?

- Терминал
- VIM (<https://www.vim.org/>)
- Anaconda (в т. ч. jupyter notebooks)
(<https://www.anaconda.com/download/>)

QR-код группы в telegram

