



Phystech@DataScience

Вводный блок, лекция 1

4.02.2023



План лекции

Организационные моменты

Основные задачи анализа данных

Примеры прикладных задач



План курса

Весенний семестр

- ▶ Вводный блок – 3 недели
 - ▶ Python и библиотеки (numpy, matplotlib, pandas)
 - ▶ Повторение теории вероятностей
 - ▶ Базовый анализ данных
- ▶ Прикладная статистика – 5 недель
- ▶ Машинное обучение – 5 недель

Осенний семестр

- ▶ Расширенный блок математической статистики
- ▶ Классическое машинное обучение (новые главы) и глубокое обучение

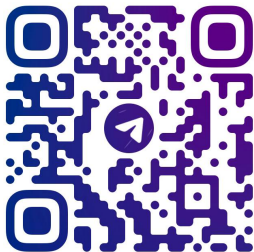


Формат занятий



Сдвоенная пара, теория перемешана с практикой.
Обязательное наличие ноутбука на парах.

Telegram-бот
@miptstats_pds_bot
Код регистрации: ROC_AUC=0.9345



Сайт команды miptstats.github.io

Почта mipt.stats@yandex.ru



Хочу зачесть этот курс

Указывать в заявлении код **РУП 31 001**

Факультатив

1. Пройти отбор*
2. Получить оценку за курс
3. **При желании** взять отрывной и записать оценку за курс в свой диплом

Индивидуальный план

1. Пройти отбор*
2. Договориться со своей физтех-школой и кафедрой о замене предмета
3. Подписать все документы



Правила выставления оценки

Правила:

1. если $L \leq 25\%$,
то $O = 1 + 3 * T$.
2. если $L > 25\%$ и ($B < 50\%$ или $C < 20\%$),
то $O = 3 + 4 * L$;
3. если $L > 25\%$ и $B \geq 50\%$ и $C \geq 20\%$,
то $O = \max(3 + 4 * L, 3 + B + 7 * LC)$.

Обозначения:

- ▶ L — доля выполнения легких заданий (их немного);
- ▶ C — доля выполнения сложных заданий (их много);
- ▶ T — доля выполнения тестов;
- ▶ LC — доля выполнения всех заданий (кроме тестов);
- ▶ B — доля правильных ответов на вопросы в боте на занятии;
- ▶ O — итоговая оценка, округляется вверх.



Правила выставления оценки

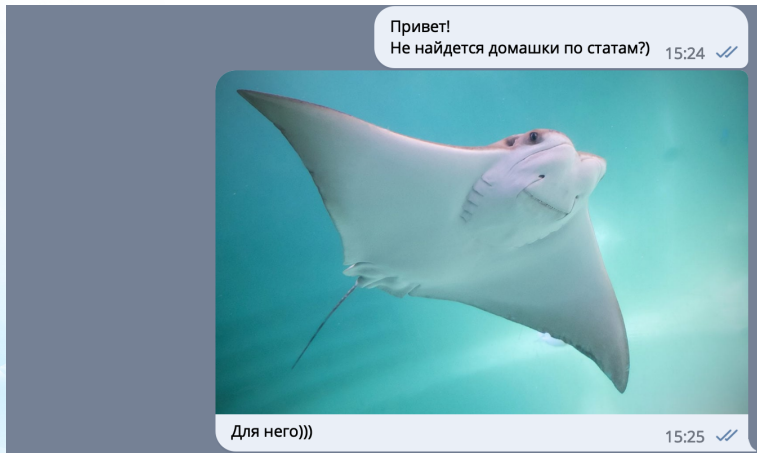
Правила:

1. если $L \leq 25\%$,
то $O = 1 + 3 * T$.
2. если $L > 25\%$ и ($B < 50\%$ или $C < 20\%$),
то $O = 3 + 4 * L$;
3. если $L > 25\%$ и $B \geq 50\%$ и $C \geq 20\%$,
то $O = \max(3 + 4 * L, 3 + B + 7 * LC)$.

Обозначения:

- ▶ L — доля выполнения легких заданий;
- ▶ C — доля выполнения сложных заданий;
- ▶ T — доля выполнения тестов;
- ▶ LC — доля выполнения всех заданий (кроме тестов);
- ▶ B — доля правильных ответов на вопросы в боте на занятии;
- ▶ O — итоговая оценка, округляется вверх.

Только если не был замечен...



Замечен скат - устный экзамен в конце семестра для всех участников :)



Правила комфорта

- ▶ Постарайтесь задавать вопросы на занятии в тот момент, когда это актуально, не перебивая на полуслове.
Другой вопрос лучше задать в перерыве или после занятия.
- ▶ Цените труд проверяющих :)
В каком из случаев проверяющему больше захочется пойти навстречу автору вопроса?
 - ▶ *"Объясните вашу претензию, почему вы мне сняли баллы, я же все сделал, я не согласен"*
 - ▶ *"Добрый день! По такой-то задаче вы написали ..., но я считаю ..., потому что ..., и у меня в работе написано ..."*



Вопросы?