

Курсы по машинному обучению

от команды предиктивной аналитики myTracker

- 12 лекций (4 блока по 3 лекции)
- 4 домашних задания (2 Kaggle и 2 Google Forms)
- 12 практических занятий
- 3 классных лектора



Антон Лоскутов



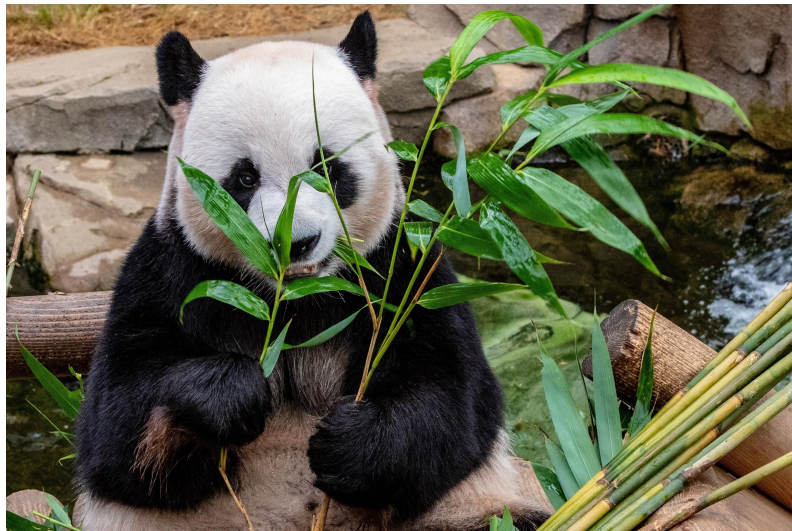
Артем Хапкин



Иван Мажаров



- a. Базовые навыки работы с табличными данными
- b. Основные функции Pandas, примеры использования
- c. Практика на знакомство с данными



- a. Копирование в pandas, оптимальное использование функций, оптимизация памяти
- b. Работа с датами
- c. Высокопроизводительные типы данных



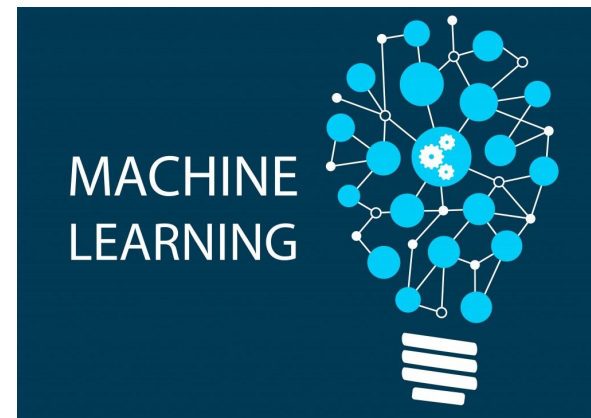
- a. Основные функции построения графиков в pandas, matplotlib, seaborn, plotly
- b. Визуальный анализ данных, объяснения графиков
- c. Практика на построение графиков и анализ результатов



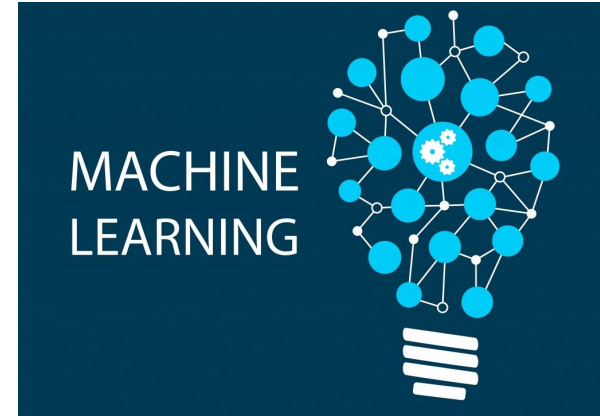
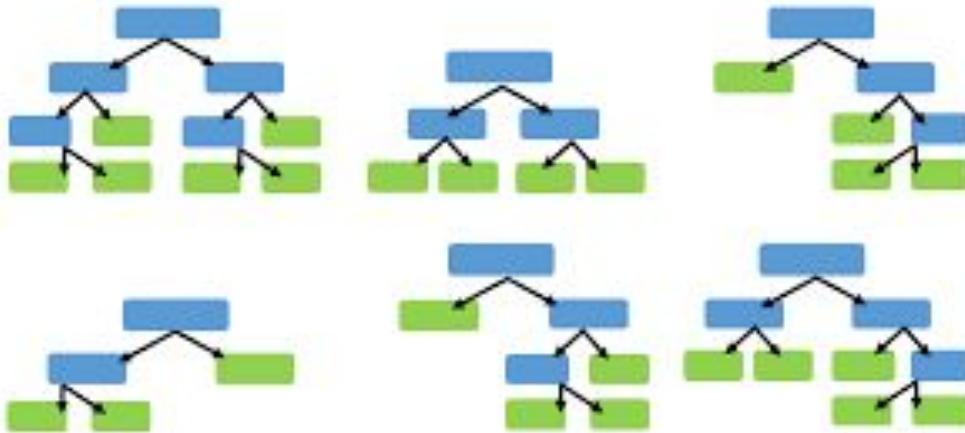
Соревнование на небольшой датасете для проверки ML навыков.



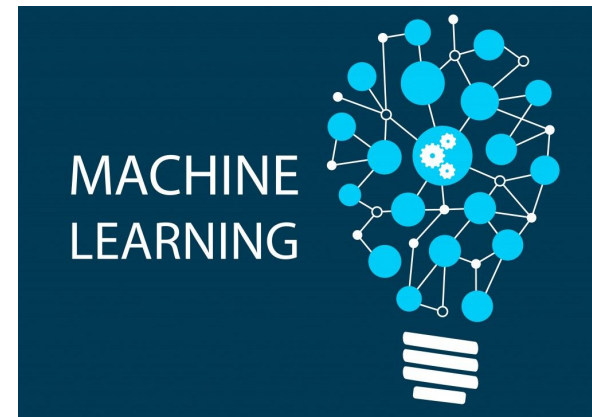
Небольшое задание для проверки навыков работы с pandas/numpy



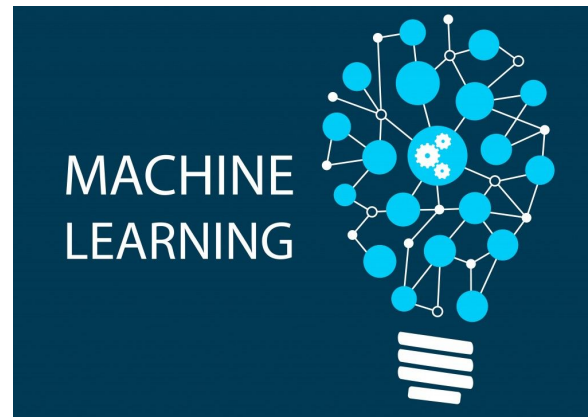
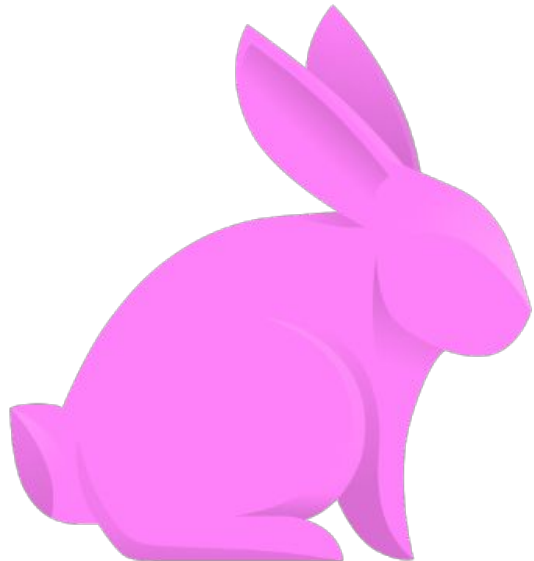
- a. Основы машинного обучения
- b. Деревья решений, энтропия
- c. Практика на знакомство с библиотекой Scikit-learn, знакомство с основными методами API библиотеки



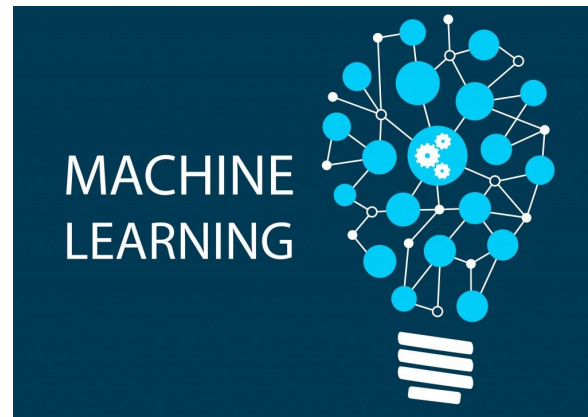
- a. Композиция алгоритмов, bagging, bootstrap
- b. Случайный лес
- c. Практика на применение случайного леса и оценку важности признаков



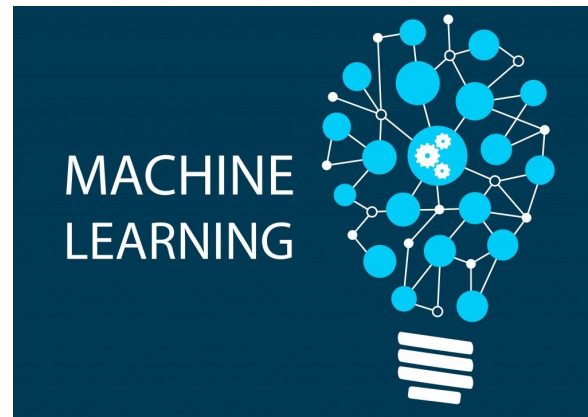
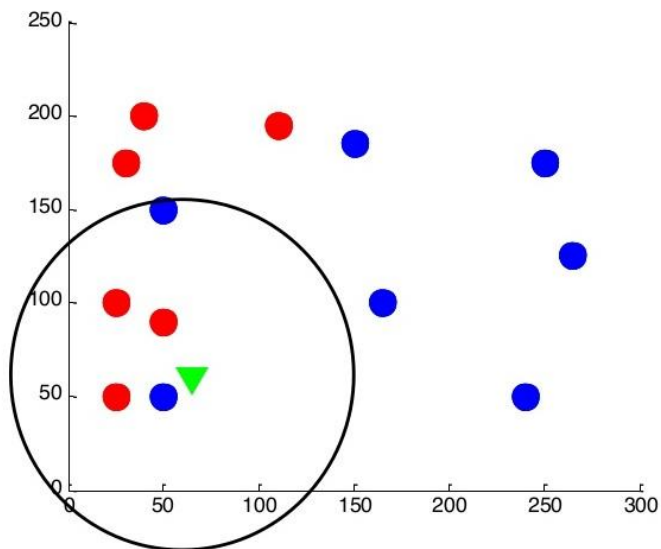
- a. Линейные модели классификации
- b. Кросс-валидация
- c. Практика на логистическую регрессию



- a. Онлайн-обучение
- b. Обучение на гигабайтах
- c. Vowpal Wabbit
- d. Основы работы с текстами



- a. Обучение без учителя
- b. Principal Component Analysis
- c. Кластеризация



- a. Метод ближайших соседей
- b. Потенциальные методы
- c. Методы Парзенковского окна

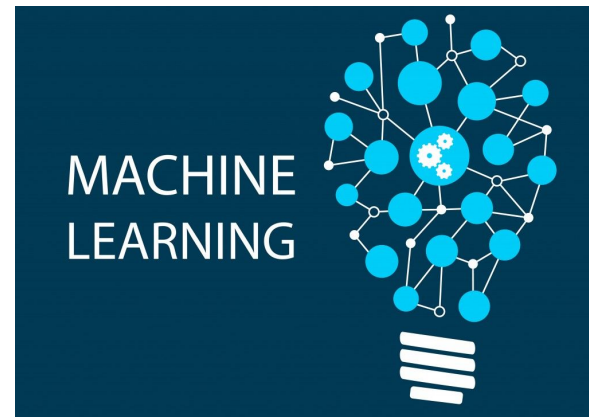
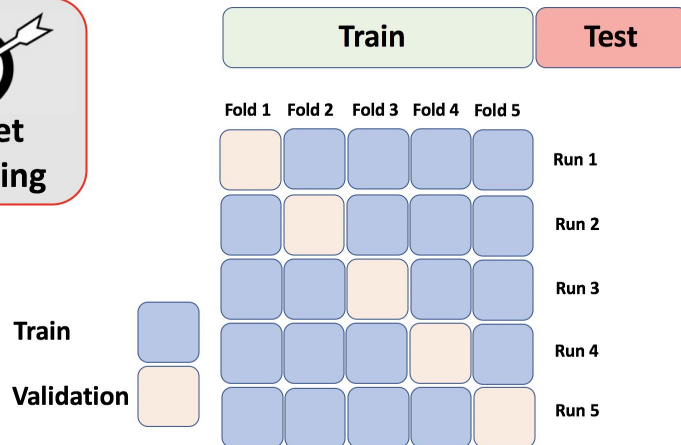


Более серьезное соревнование, в котором надо применить все накопленные навыки

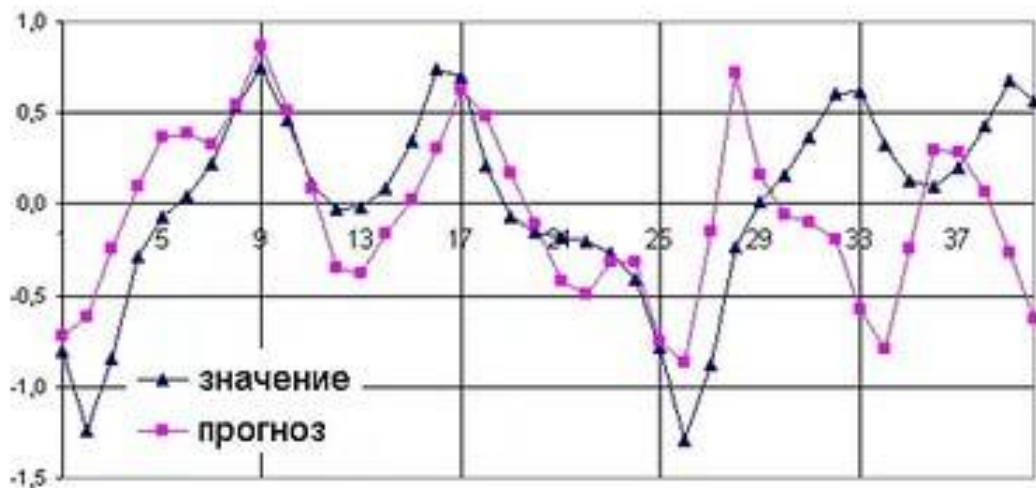


Небольшое задание по метрическим алгоритмам

Target-encoding



- a. OH-кодирование
- b. Bag Of Words
- c. Kernel Trick



MACHINE
LEARNING



- a. Параметрическая регрессия
- b. Непараметрическая регрессия
- c. Facebook Prophet



Yandex
CatBoost

Microsoft
LightGBM

dmlc
XGBoost

- a. Современный взгляд на бустинг
- b. Теоретические основы
- c. Лучшие на сегодня реализации