

ML Basic

Практика

Подготовка датасета из сырых данных



Проверить, идет ли запись

Меня хорошо видно && слышно?



Ставим "+", если все хорошо
"-", если есть проблемы



Тема вебинара

ML Basic

Практика

Подготовка датасета из сырых данных



Николай Осипов

MLOps Engineer @ Kadam.net

Опыт: 4+ лет в области Data Science

Направления: MLOps, DS инфраструктура, табличные данные, CV

Telegram: [@NickOsipov](https://t.me/NickOsipov)



Правила вебинара



Активно
участвуем



Off-topic обсуждаем
в учебной группе в tg



Задаем вопросы в чат



Вопросы вижу в чате,
могу ответить не сразу

Условные обозначения



Индивидуально



Время, необходимое
на активность



Пишем в чат



Говорим голосом



Документ



Ответьте себе или
задайте вопрос

Маршрут вебинара

Цели и смысл

Практика

Рефлексия

```
feature_engineering.py

if __name__ == "__main__":
    df = load_data()
    df = preprocessing(df)
    result = transformation(df)
    report = get_report(result)
    save_data(result)
    save_report(report)
```

Цели вебинара

К концу занятия вы сможете

1. Загружать и предобрабатывать данные
2. Проводить feature engineering с осмысленными признаками
3. Работать с данными и готовить датасет с примером из реальной задачи

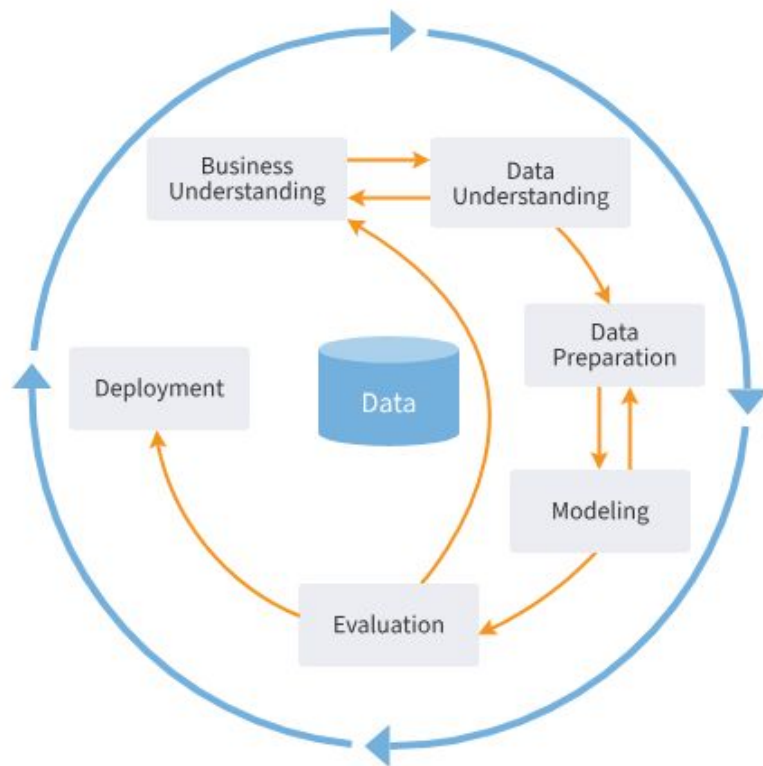
Смысл

Для чего нам это нужно

1. Для наработки навыка подготовки датасетов для своих моделей
-

Практика

Где мы находимся



Понимание бизнес-целей
(Business Understanding)

Понимание данных
(Data Understanding)

Подготовка данных
(Data Preparation)

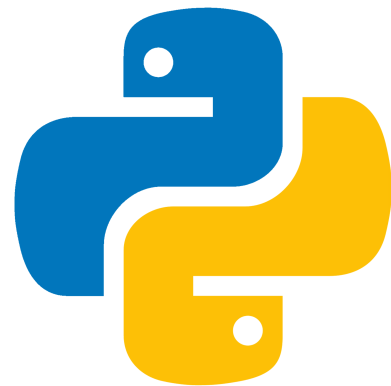
Моделирование
(Modeling)

Оценка
(Evaluation)

Внедрение
(Deployment)

План практики

1. Загрузим сырые данные
2. Проведем небольшой разведочный анализ
3. Подготовим новые признаки
4. Проанализируем результаты
5. Сохраним датасет



Рефлексия

Цели вебинара

К концу занятия вы сможете

1. Познакомиться с основными магическими методами;
2. Узнать основные магические методы для арифметических операций;

Рефлексия



Отправьте в чат эмодзи, который отражает ваше настроение после занятия.



Продолжите высказывание: теперь я умею...



Что планируете использовать в работе?

Следующий вебинар



6 мая 2025

Продвинутый ООП, исключения



Ссылка на вебинар
будет в ЛК за 15 минут



Материалы
к занятию в ЛК —
можно изучать



Обязательный материал
обозначен красной
лентой



**Заполните, пожалуйста,
опрос о занятии
по ссылке в чате**

Спасибо за внимание!

Приходите на следующие вебинары



Николай Осипов

Telegram: [@NickOsipov](https://t.me/NickOsipov)