

# Классификация парнокопытных

Кейс от Министерства природных  
ресурсов и экологии Российской  
Федерации

Команда:  
ikanam\_chipi\_chipi



# ikanam\_chipi\_chipi



**Аделя Сабирова**  
Data engineer, designer



**Максим Ляра**  
Team lead, Data scientist  
TG: @maxlyara1



**Станислав Палатов**  
Data scientist

# Проблематика

## Качество снимков

Ночные/черно-  
белые/некачественные  
снимки

01

02

## Сбалансированность

Оптимизация под слабые  
компьютеры

## Автономность

Возможность работы без  
подключения к сети  
«Интернет»

03

04

## Ошибки в данных

Удаление снимков, ошибочно  
попавших в данные/классы

## Наше решение

**3,7%**

### Неизвестный класс

Именно такой процент загруженных нами фотографий модель не смогла определить

**94%**

### Точность

Процент верно определенных классов изображений на валидационной выборке

**93%**

### F1 score

Это метрика на валидации

# Шаги решения



## Предобработка данных

Удалили крайне некачественные снимки, ошибки в данных



## Аугментация данных

Для обучения добавили больше ночных, ч/б снимков



## Сбалансированная модель

Выбрали нейросеть с наилучшим соотношением точность/скорость



## Конкретная задача

Дообучили модель на данных для задачи классификации подвидов оленей



## Понятный интерфейс

Создали интерфейс, позволяющий загружать папки/файлы



## Требуемый функционал

Добавили функцию распределения классифицированных изображений в папки

# Демонстрация решения

[https://github.com/maxlyara1/deer\\_types\\_classifier\\_hackathon](https://github.com/maxlyara1/deer_types_classifier_hackathon)

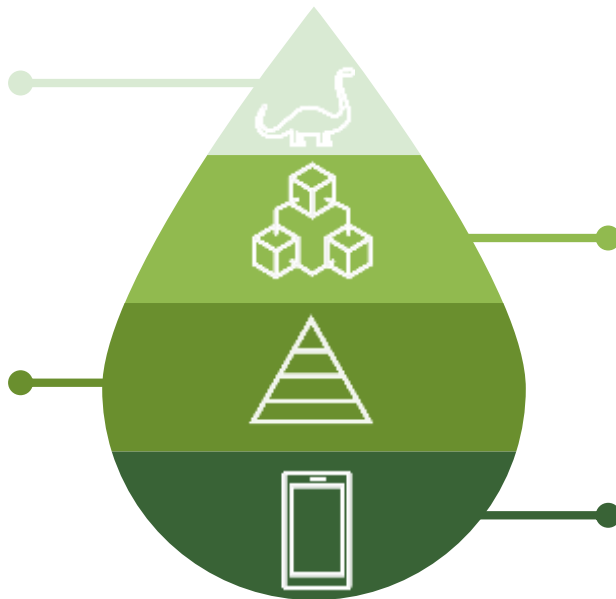
# Масштабируемость

## Адаптируемость

Возможность  
классификации других  
животных, птиц

## Вариативность

Выбор модели в зависимости  
от необходимого соотношения  
точность/скорость



## Big Data

Поиск новых массивов данных  
для улучшения качества  
модели

## Мобильность

Создание приложения  
для мобильных устройств

5,3 млн параметров  
20,5 МБ – размер модели  
Признана эталонной моделью в  
классе сверточных нейронных  
сетей  
Предобучена на датасете  
ImageNet с 1,5 млн изображений

## **Model EfficientNet-B0**

Удобный для написания кода  
Подходит для нашей задачи  
Возможность сделать  
интуитивно понятный интерфейс

## **Streamlit**

# Используемые технологии

