

# Тема: Прогнозирование погоды с использованием методов машинного обучения

**Команда: МФТИшные котики)**

## **Состав:**

Самаковский Вячеслав  
Ильиных Александр  
Гриднев Константин  
Зайцев Дмитрий  
Алиева Наталья



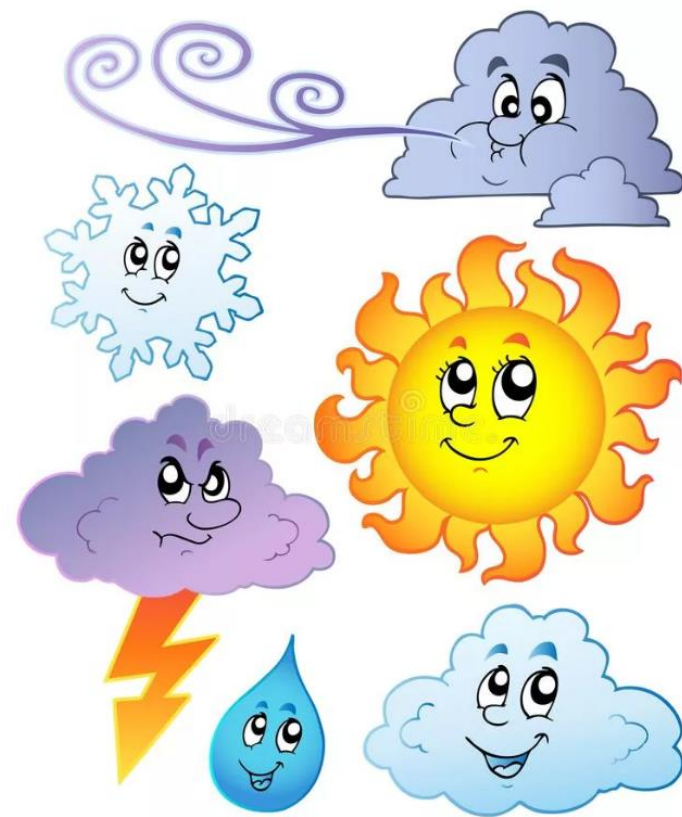
# Бизнес-цель

Прогноз погодных условий на  
основании детекции изображений  
погоды



# ML - задача

- Предметная область: **CV**
- Тип задачи: **классификация** погоды (по фото)
- Таргет: предсказание погоды (снег, дождь, солнечно и т.д.)
- Технология: **Tensorflow, Roboflow, Yolo**



# Принцип работы

## Наша модель

**Tenzorflow**



Датасет: **Kaggle**



## Модели для сравнения

**Roboflow classification**



Датасет: **Kaggle**  
+ аугментация



**Yolo (на Roboflow)**



Датасет: **дефолтный**



Сравнение моделей

# Датасет

- Датасет: **Weather Image Recognition\***
- **6 862 фото – 8 категорий:**  
туман, мороз, молния, дождь, радуга,  
песчаная буря, снег, солнечно



\* <https://www.kaggle.com/datasets/jehanbathena/weather-dataset>

# Preprocessing (для Roboflow)





# Preprocessing (для Roboflow)

## Preprocessing

Auto-Orient: Applied

Resize: Stretch to 640x640

## Augmentations

Outputs per training example: 3

Flip: Horizontal

Crop: 0% Minimum Zoom, 20% Maximum Zoom

Grayscale: Apply to 25% of images

Brightness: Between -25% and +25%

Noise: Up to 5% of pixels

13868 Total Images

[View All Images →](#)



# Обучение модели Roboflow

## Dataset Split

TRAIN SET

87%

12129 Images

VALID SET

8%

1159 Images

TEST SET

4%

580 Images

## weather-classification/1

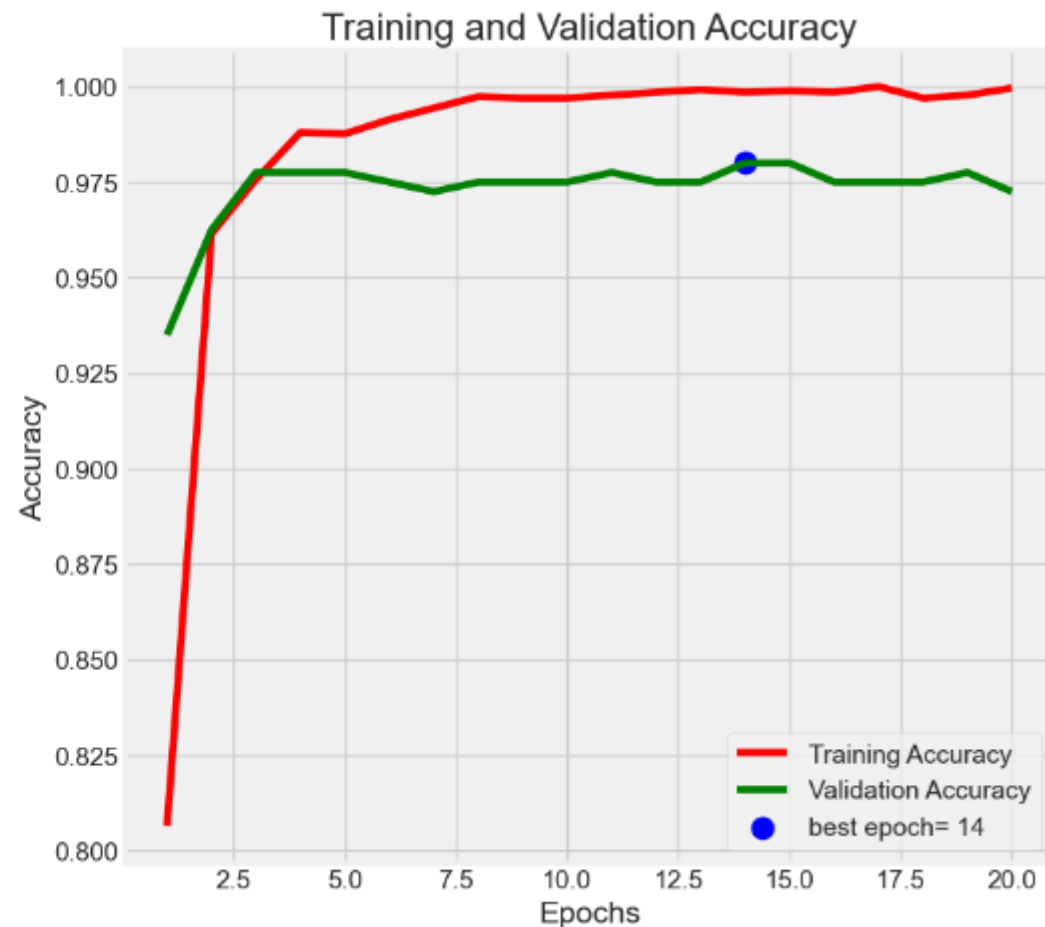
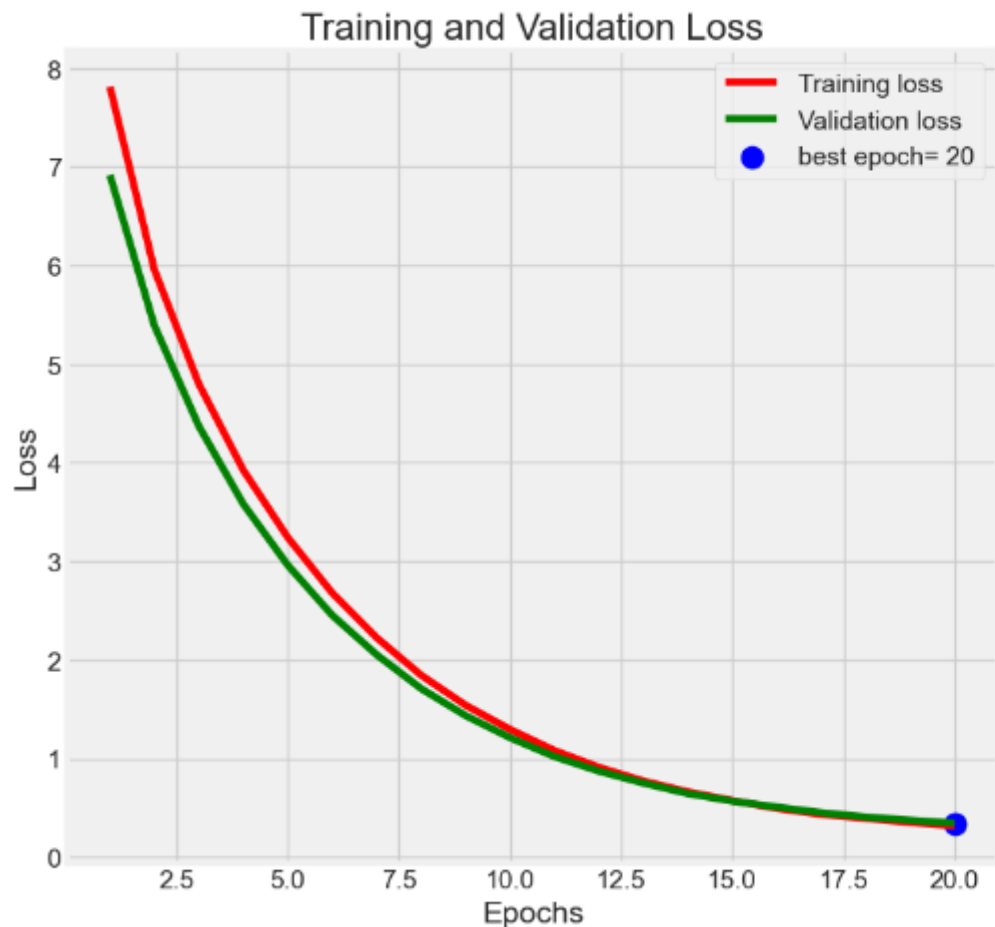
Model Type: Roboflow 2.0 Multi-label Classification

Validation Accuracy ?

98.7%



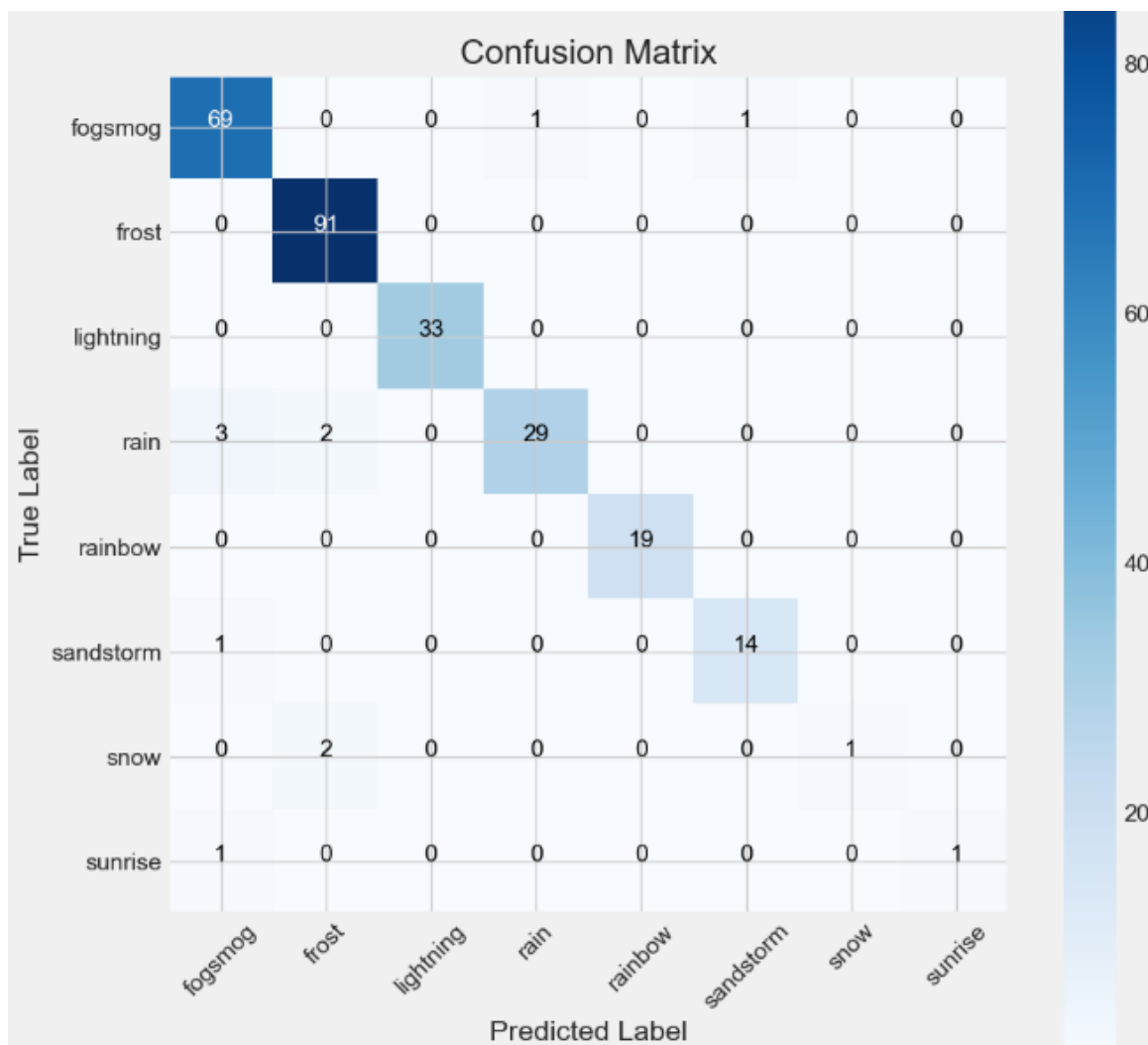
# Производительность модели Tensorflow



# Оценка модели Tensorflow

```
4/4 [=====] - 6s 1s/step - loss: 0.2667 - accuracy: 1.0000
4/4 [=====] - 6s 1s/step - loss: 0.3607 - accuracy: 0.9727
4/4 [=====] - 6s 1s/step - loss: 0.3989 - accuracy: 0.9590
Train Loss: 0.26669108867645264
Train Accuracy: 1.0
-----
Validation Loss: 0.36070716381073
Validation Accuracy: 0.97265625
-----
Test Loss: 0.39886474609375
Test Accuracy: 0.9589552283287048
```

# Матрица ошибок Tensorflow



	precision	recall	f1-score	support
fogsmog	0.93	0.97	0.95	71
frost	0.96	1.00	0.98	91
lightning	1.00	1.00	1.00	33
rain	0.97	0.85	0.91	34
rainbow	1.00	1.00	1.00	19
sandstorm	0.93	0.93	0.93	15
snow	1.00	0.33	0.50	3
sunrise	1.00	0.50	0.67	2
accuracy			0.96	268
macro avg	0.97	0.82	0.87	268
weighted avg	0.96	0.96	0.96	268

# Сравнение моделей

## Roboflow собственная модель vs дефолтная модель

```
from roboflow import Roboflow

# Наша модель
rf_cats = Roboflow(api_key="yAV8c8VxFzT7RXVpeS3a")
project_cats = rf_cats.workspace().project("weather-classification")
model_cats = project_cats.version(1).model

# Модель из интернета
rf = Roboflow(api_key="yAV8c8VxFzT7RXVpeS3a")
project = rf.workspace().project("weather-classification-w5xug")
model = project.version(8).model
```

# Сравнение моделей: предсказание

*# Наша модель*

```
print(model_cats.predict("test_data/11.jpg").json())  
print(model_cats.predict("test_data/1688761400_kartin-papik-pro-p-kartinki-vechernii-dozhd-52.jpg").json())  
print(model_cats.predict("test_data/1830.jpg").json())
```

*# Модель из интернета*

```
print(model.predict("test_data/11.jpg").json())  
print(model.predict("test_data/1688761400_kartin-papik-pro-p-kartinki-vechernii-dozhd-52.jpg").json())  
print(model.predict("test_data/1830.jpg").json())
```

# Результаты: сравнение моделей по трем фото

Изображение/ признак	Дождь			Дождь с радугой			Молния		
									
	Tensorflow	Roboflow	Yolo	Tensorflow	Roboflow	Yolo	Tensorflow	Roboflow	Yolo
Наличие признака в модели	+	+	?	+	+	?	+	+	?
Результат	rain: -	rain: confidence 0.94	rain: confidence 0.99	rain: -	rain: confidence 0.42 rainbow: confidence 0.47	rain: confidence 0.99	lightning: -	lightning: confidence 0.92	shine: confidence 0.98

Validation Accuracy: tensorflow = 0.97, roboflow = 0.98

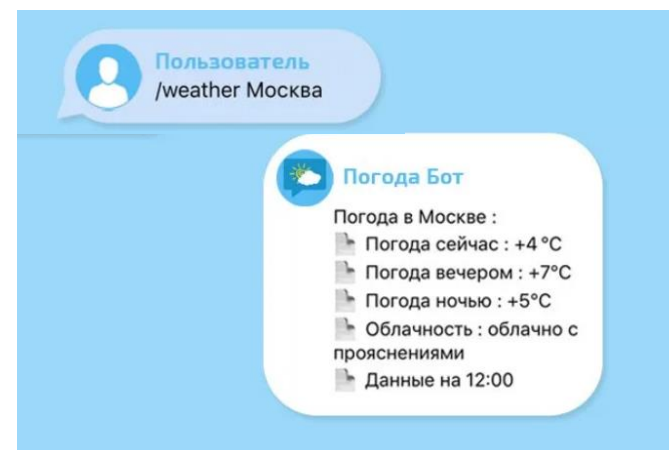
# Вывод

- Модель Roboflow, обученная на аугментированном датасете, показала более высокие результаты: она правильно предсказывает класс, но при этом уступает по confidence.
- Модель Tensorflow, обученная на неаугментированном датасете, показывает более низкую предсказательную способность, т.к. в одно случае угадала только один из признаков.
- Модель Yolo обучена на другом датасете, в котором содержится меньше классов, поэтому модель ошибается, но демонстрирует самый высокий confidence.



# Перспективы развития модели

- В зависимости от детекции реальных погодных условий, записанных с камер видеонаблюдения города, определять вероятность аварий (на сколько опасно сегодня садиться за руль)
- Пользователь пишет город проживания, система находит онлайн камеры видеонаблюдения, считывает серию скринов погоды, и пользователь получает ответ какая погода



Для вас старалась команда МФТИшных котиков)  
Спасибо за внимание!

