




# Flood Prediction

Лундина Светлана,  
Петрова Елизавета



# Цель работы

- 1) **Главная цель:** научиться предсказывать вероятность затопления региона на основе различных факторов
- 2) Исследовать алгоритмы, которые могут быть применимы для решения нашей задачи
- 3) Создать бота для демонстрации работоспособности модели

# Описание исходных данных

**Исходные данные** были взяты с соревнования kaggle: [Regression with a Flood Prediction Dataset](#).

В датасете представлены 20 числовых признаков, которые потенциально оказывают влияние на возникновение потопа.

**Примеры признаков:** интенсивность муссонов, уровень урбанизации, тип дренажной системы и тд.

# Используемые алгоритмы

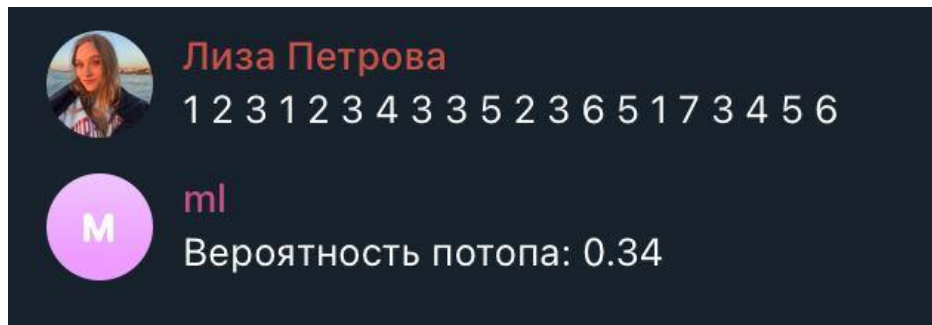
- **Использованные модели:** Множественная регрессия, XGBRegressor, CatBoostRegressor, RandomForestRegressor, SVR
- **Лучшие модели:** CatBoostRegressor и Множественная регрессия (у обеих  $r^2\_score$  примерно 0.84)
- В итоге решили использовать CatBoostRegressor, поскольку при первичном анализе был сделан вывод, что таргет не является сильно линейно зависимым ни от одной из фичей

# Качество моделей

- В качестве финальной модели выбрали **CatBoostRegressor**.
- $r^2\_score = 0.84$
- В качестве метрики взяли  $r^2\_score$ , потому что она является более понятной и интерпретируемой, чем  $mse/rmse$ .

# Результат работы

Результатом является телеграм бот, предсказывающий вероятность потопов на местности по введенным пользователем числовым данным по каждой из фичей.



23:22



< 8439

ml  
bot



No messages here yet...



Message



я

Ну

Как

Й Ц У К Е Н Г Ш Щ З Х

Ф Ы В А П Р О Л Д Ж Э

↑ я ч с м и т ь б ю ↵

123



Пробел

Ввод





Спасибо за внимание!

