

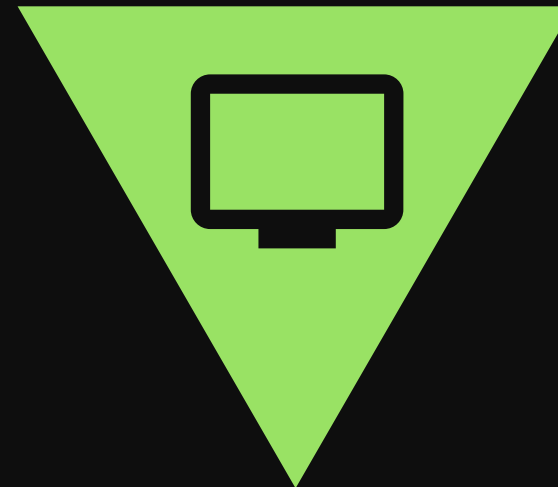
# DEVIAN

**DEV IT SIMPLE**

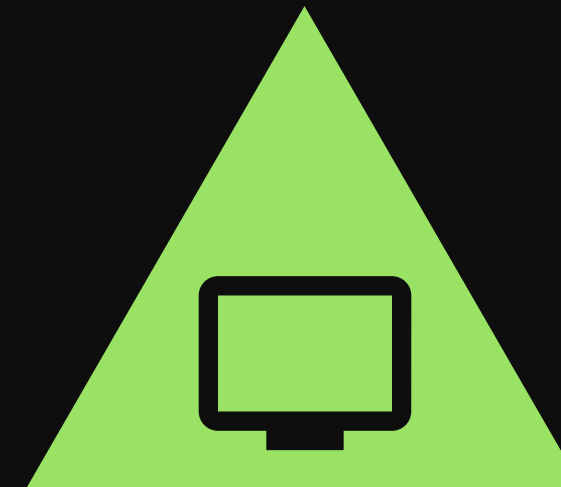
# ПОСТАНОВКА ЗАДАЧИ



*Android-  
приложение*



*Сервис обработки  
и хранения  
отзывов*



*Сервис  
классификации  
отзывов*



# ОБРАБОТКА СЫРЫХ ДАННЫХ



## Предобработка

- Токенизация
- Лемматизация
- Удаление стоп-слов

## Создание выборки

- Эвристические правила для каждого класса



# ЭВРИСТИЧЕСКИЕ ПРАВИЛА

*Проблема отсутствия разметки*



- Для каждого класса составлен словарь слов/биграмм, наличие которых может указывать на принадлежность к данному классу
- На основе эвристик была создана обучающая выборка



# ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ



*Модель для предсказания  
команды*

*Комплементарный наивный  
байесовский классификатор*

**87.1%**

Accuracy CV5 на отложенной выборке



*Модель для предсказания  
тональности и типа отзыва*

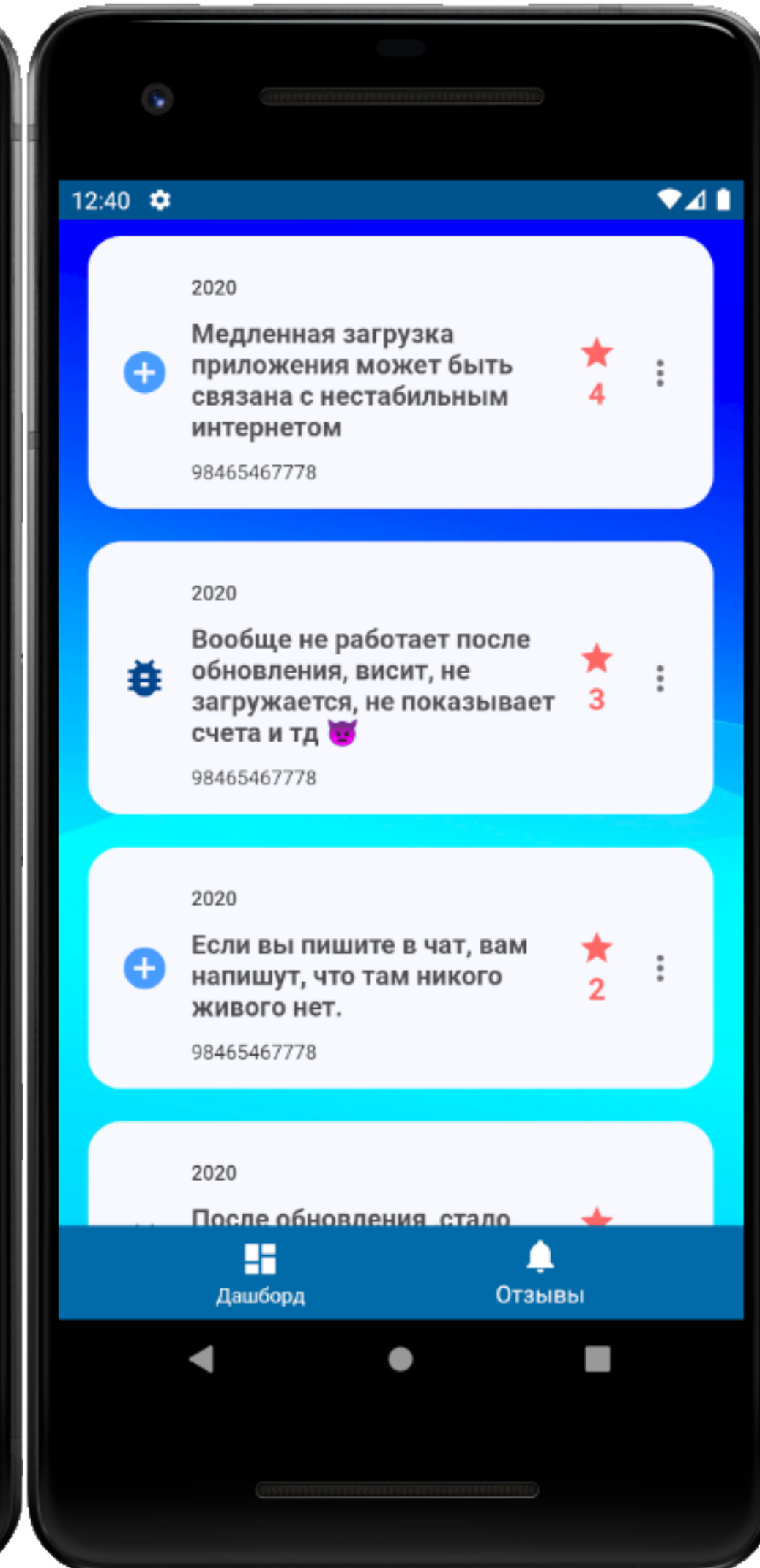
*Комплементарный наивный  
байесовский классификатор*

**92.6%**

Accuracy CV5 на отложенной выборке

# НАШЕ РЕШЕНИЕ

- Просмотр отзывов
- Общая статистика по отзывам
- Ответ на отзыв
- Возможность указать неправильную классификацию



# КОМАНДА DEVIAN

КРЕАТИВНЫЕ ГЕНИИ



***Михаил Притугин***

ML / Python-разработчик  
21 год



***Максим Аминов***

Android-разработчик  
21 год

MAKE IT TOGETHER

