



«**Personal\_diary**»

Выполнила: Мосунова Ксения, гр.5040112\40201

17. Dec. 2025.



# **Анализ настроения:** **Личный дневник**

# Постановка задачи

---

**Цель** – создать персональный «эмоциональный дневник», который сам определяет эмоцию по тексту записи, позволяет пользователю вручную её скорректировать и постепенно подстраивает модель под конкретного человека.

## Задачи:

- Создать интерактивный дневник с автоанализом эмоций
- Адаптировать модель под конкретного пользователя
- Обучать модель на основе feedback'ов

# Dataset



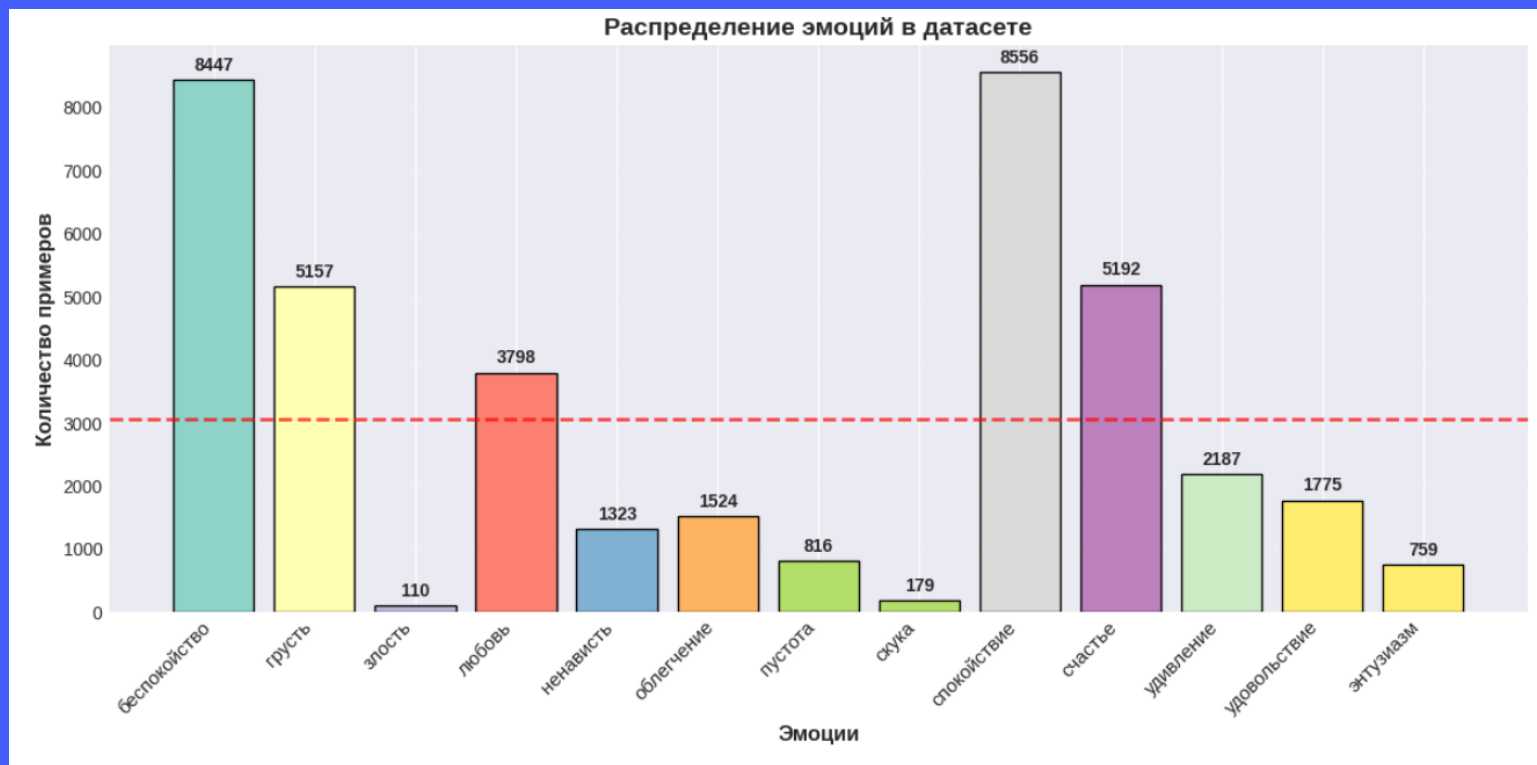
-> 40К англоязычных твитов

-> 13 эмоций (беспокойство, грусть, злость, любовь, ненависть, облегчение, пустота, скука, спокойствие, счастье, удивление, удовольствие, энтузиазм )

-> Вид: {"sentiment": "label", "content": "text"}

# Подготовка:

- > Перевод датасета на русский язык
- > Удаление повторяющихся строк
- > Удаление спецсимволов
- > Приведение к нижнему регистру для классических моделей
- > Генерация записей для эмоций с небольшим количеством записей



## Промпт для генерации твитов необходимой эмоции

```
prompts = [  
    f"Напиши короткое сообщение от первого лица, в котором однозначно чувствуется эмоция {label}.",  
    f"Напиши пример твита, где человек явно испытывает {label}.",  
    f"Придумай одно предложение, передающее сильную эмоцию {label}.",  
    f"Напиши короткую фразу, в которой ясно читать эмоцию {label}.",  
    f"Напиши дневниковую запись на одну-две строки о том, как человек ощущает {label}.",  
]
```

label = [злость, ненависть, облегчение, пустота, скука, удивление,  
удовольствие, энтузиазм]

Модель: ai-forever/rugpt3large\_based\_on\_gpt2

# Построенные модели

01

Наивный Байесовский классификатор

02

Логистическая регрессия

03

Метод опорных векторов

04

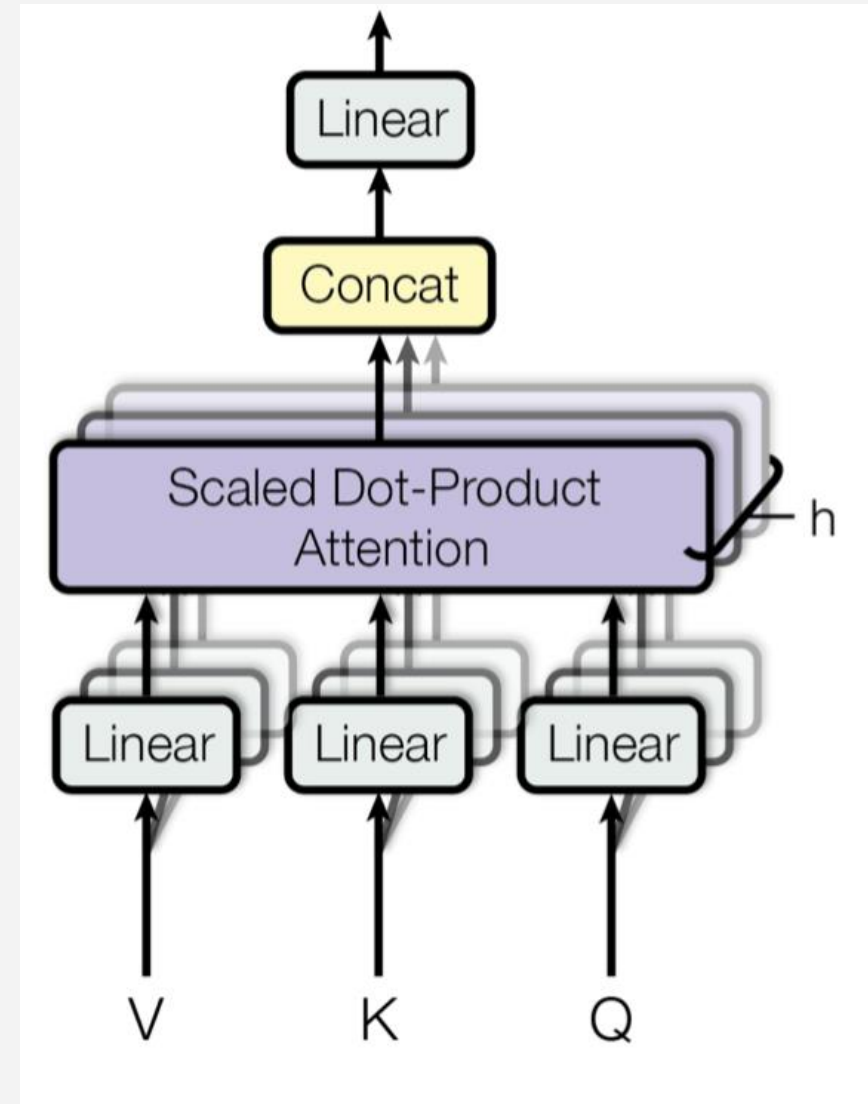
Случайный лес

**F1 score  $\approx$  54%**

# ruBERT

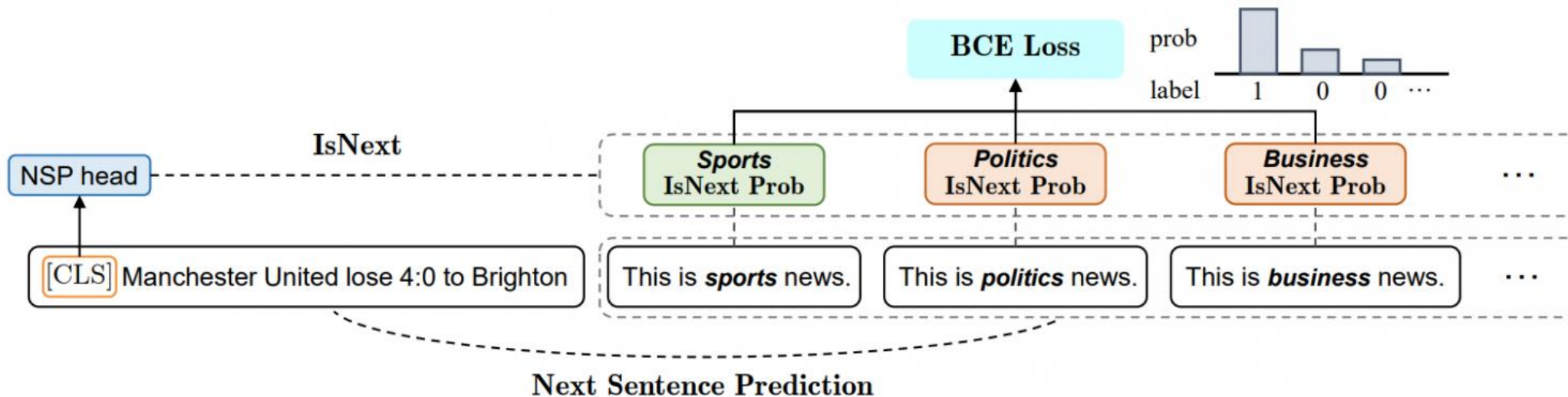
Модель BERT основана на **трансформере** — нейросетевой архитектуре, которая использует **механизм внимания** для анализа и понимания текста.

**F1 score = 60%**



# ruBERT

Модель BERT основана на **трансформере** — нейросетевой архитектуре, которая использует **механизм внимания** для анализа и понимания текста.

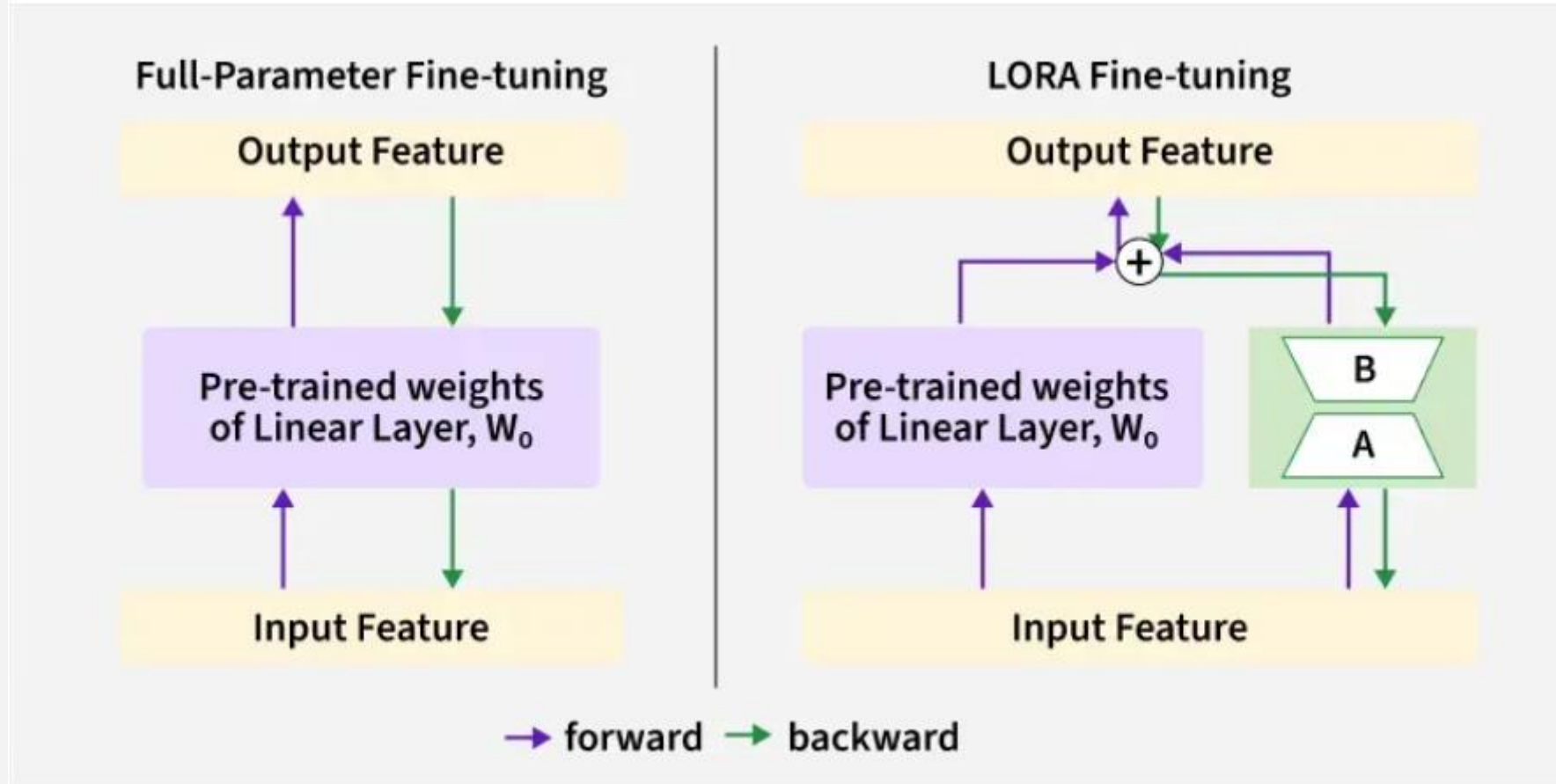


**F1 score = 60%**

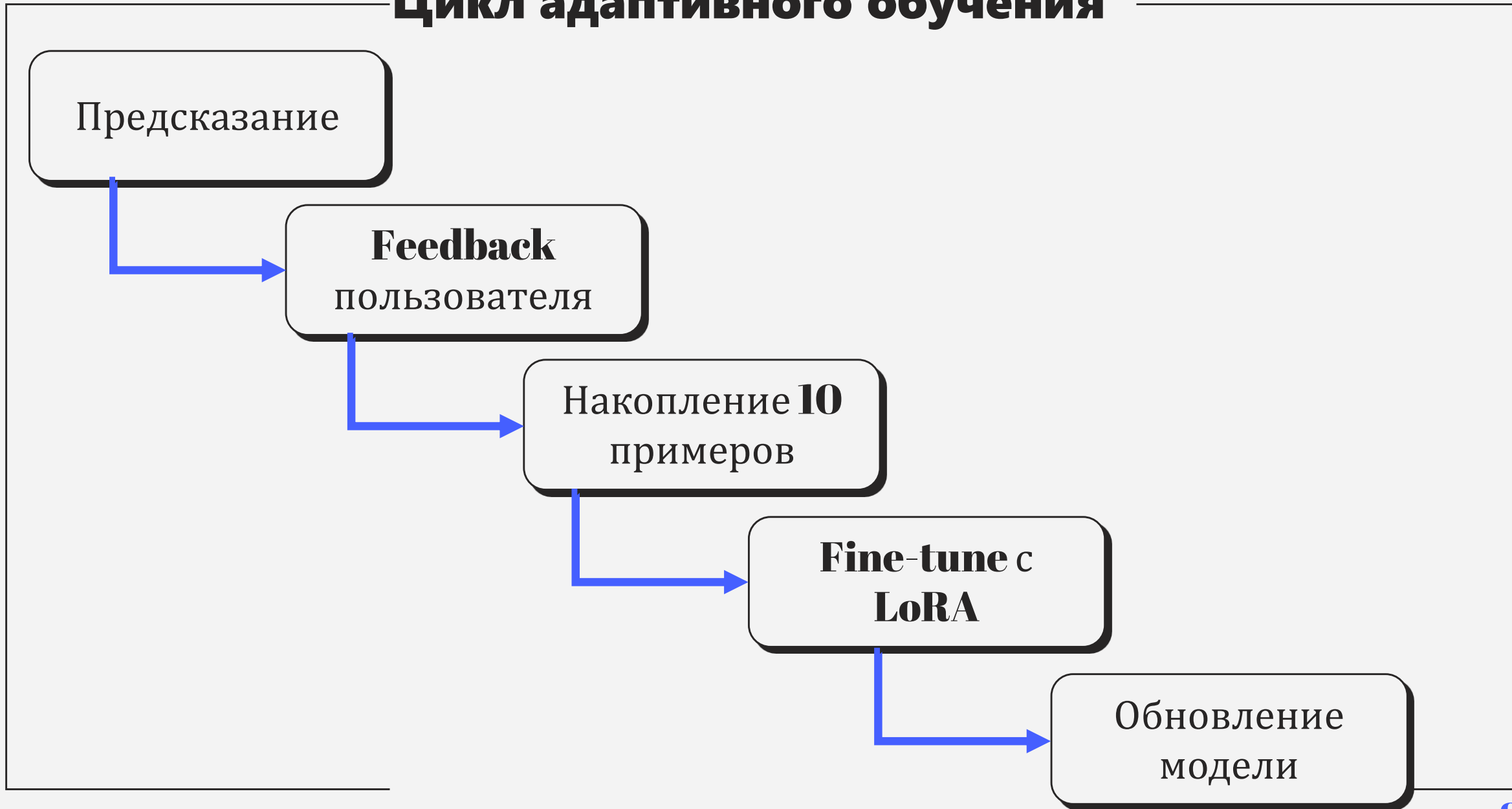


# LoRA

–метод «параметрически-эффективного дообучения»

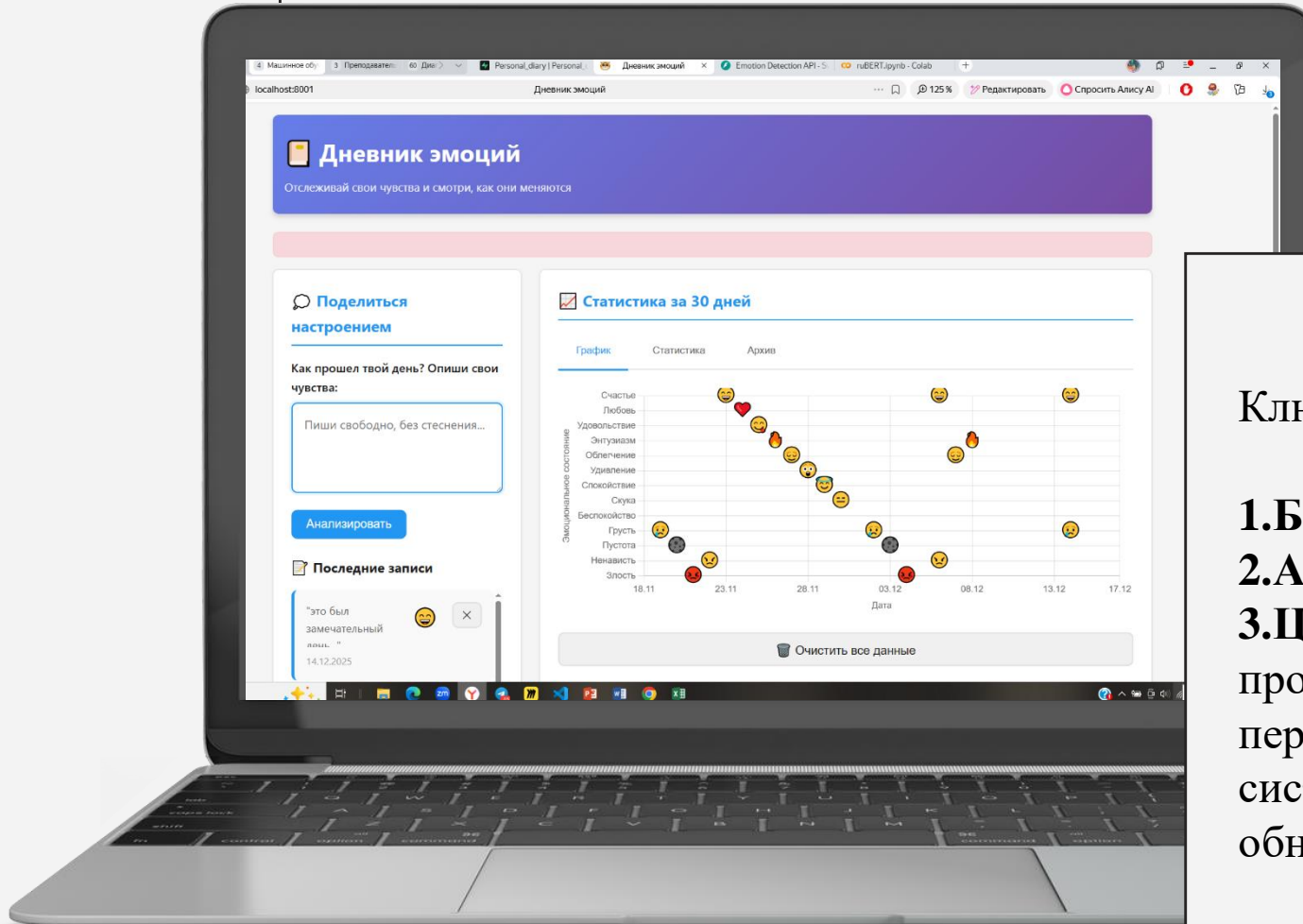


# Цикл адаптивного обучения



## Итоги:

Создана полнофункциональная система адаптивного распознавания эмоций, которая объединяет современные методы глубокого обучения (BERT, LoRA) с практическим применением в персональном дневнике.



### Ключевые компоненты решения

1. Базовая модель (ruBERT)
2. Адаптивный слой (LoRA)
3. Цикл обучения — автоматический процесс накопления исправлений и переобучения: за каждые 10 фидбеков система выполняет fine-tune и обновляет модель.

# Источники:

1. Данные: <https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text>
2. Статья «Анализ тональности текста методами машинного обучения», портал CyberLeninka: <https://cyberleninka.ru/>
3. Hugging Face Transformers
4. Документация по моделям BERT, токенизации и fine-tuning. Ссылка: <https://huggingface.co/docs/transformers>
5. «Модели BERT для машинного обучения: гайд для начинающих» SkillFactory / Habr, 2024. Ссылка: <https://habr.com/ru/companies/skillfactory/articles/862130/>
6. «Как дообучать огромные модели с максимальным сохранением знаний» Habr, 2023. Ссылка: <https://habr.com/ru/articles/781988/>

**Thanks !**