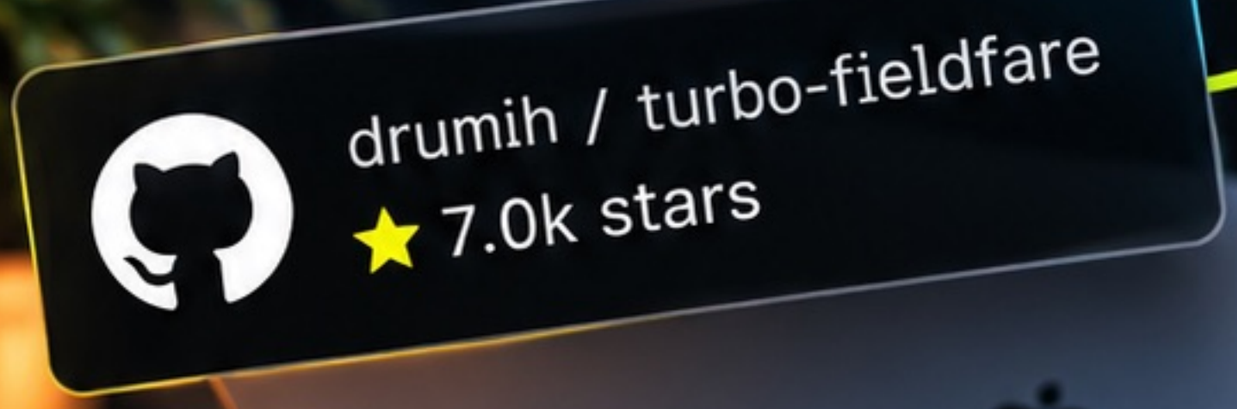
AI МОДЕЛЬ С 26 МЛРД ПАРАМЕТРОВ, КОТОРАЯ РАБОТАЕТ НА 2 ГБ ОЗУ.

Запускается на MacBook Air
с 8 ГБ ОЗУ.
Не нужен облачный сервис,
не нужен API ключ.



26 МЛРД ПАРАМЕТРОВ НА НОУТБУКЕ

ВСЕГО 8 ГБ ОЗУ.

РЕАЛЬНЫЕ ИЗМЕРЕНИЯ, А НЕ ОБЕЩАНИЯ.

drumih / turbo-fieldfare

Type / to search

Code | Issues | Pull requests | Actions | Security | Insights

README.md

Measured decode speed

Machine	Memory	Measured decode
8 GB M2 MacBook Air	**8 GB**	**5.1–6.3 tok/s**
M5 Pro	24 GB	31–35 tok/s

MacBook Air

M1

Текстовая генерация работает даже **на M1.**

ШАГ 1

СОБИРАЕМ ВСЁ В 3 КОМАНДЫ

Клонируем репозиторий и компилируем **Swift runtime**. Этого достаточно, чтобы поднять всю сборку.

Работает на
Apple Silicon Mac
даже с 8 ГБ ОЗУ

macOS 26
с Metal 4

Swift 6.2
или новее

Нужно всего
14.3 GB
свободного места

```
bash

nikhil@macbook ~ % git clone
https://github.com/drumih/turbo-fieldfare.git

Cloning into 'turbo-fieldfare'...

nikhil@macbook ~ % cd turbo-fieldfare

nikhil@macbook turbo-fieldfare %
swift build -c release

Building for production...

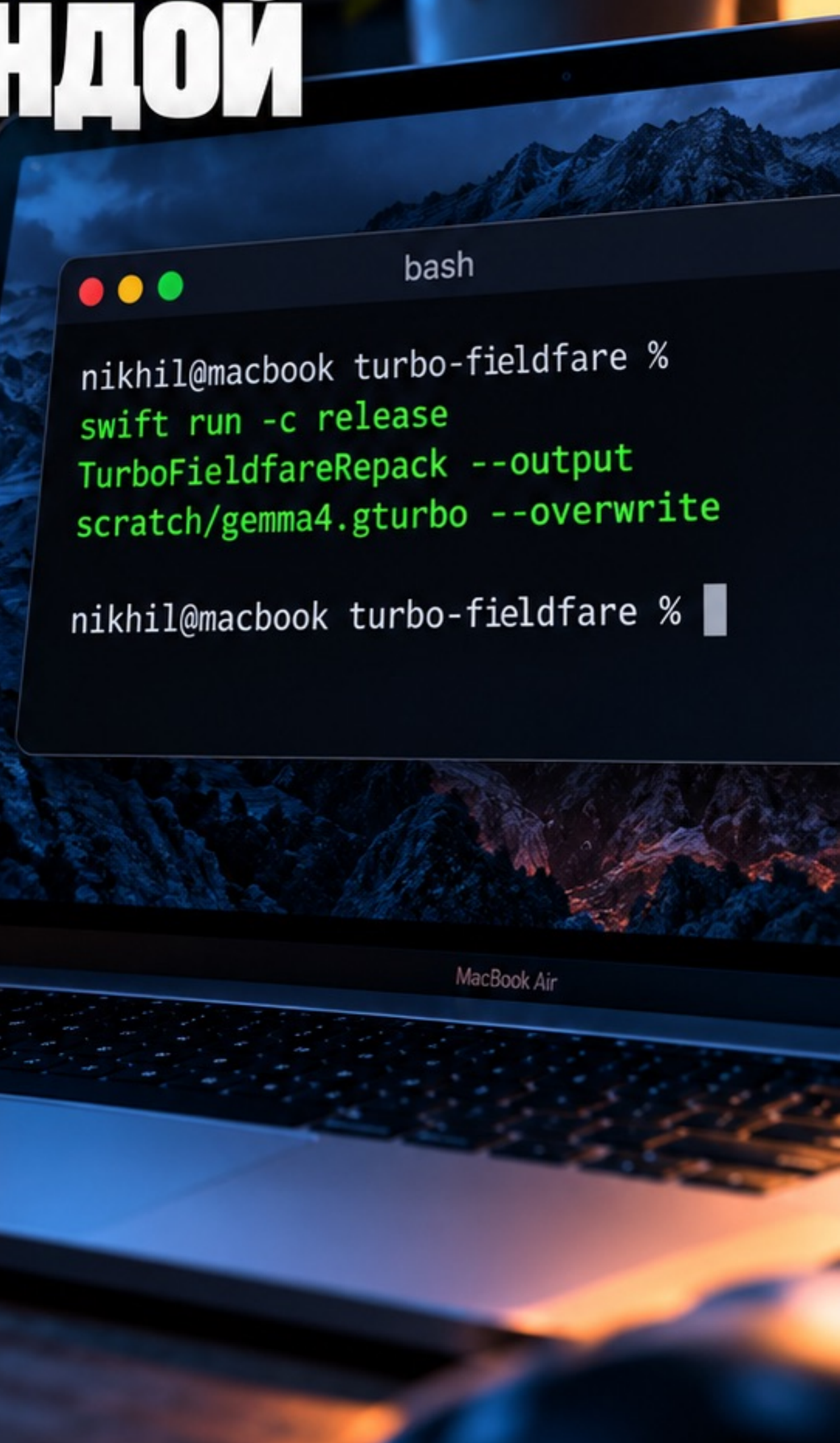
nikhil@macbook turbo-fieldfare %
```

ШАГ 2

СКАЧИВАЕМ МОДЕЛЬ ОДНОЙ КОМАНДОЙ

Одна команда скачивает **Gemma 4** и добавляет её в **runtime**.

- Стримит **14.3 GB** checkpoint
- **Не хранит** полную копию на диске
- Опциональный image pack **1.1 GB**
- Интернет нужен только **один раз**



ШАГ 3

ПРЕВРАЩАЕМ МОДЕЛЬ В API

Запускаем локальный сервер и подключаем свои приложения **на Mac.**

- Работает на **localhost**
- Всё остаётся **на вашем Mac**
- Порт **8080**, путь **/v1**
- Можно подключать **через любое приложение** или чат на Mac



ИЗ 26 МЛРД ПАРАМЕТРОВ АКТИВНЫ ТОЛЬКО 4 МЛРД

Это смесь экспертов (MoE). Роутер выбирает **8 экспертов на каждый токен**, а остальные остаются на SSD.



26B	**Всего параметров** Полный размер модели
3.88B	**Активно на каждый токен** Используется только 4B из 26B
1.35 GB	**Ядро модели в RAM** Эта часть всегда в памяти
16	**Слотов экспертов в кэше** Остальные ждут на SSD, при необходимости подгружаются

ЗАПУСТИЛИ БЫ ВЫ МОДЕЛЬ В 26В НА 8 ГБ ОЗУ?

Сохраните этот пост
и напишите в комментариях,
на каком **Mac** вы бы её запустили.