Чек-лист выбора AI-модели

Не выбирайте модель только по leaderboard. Проверьте её на своём workload.

1. Capability

☐ Какие 10–30 реальных задач должна решать модель?

☐ Нужны ли coding, vision, long context, computer use или tools?

☐ Какой уровень ошибок допустим?

2. Economics

☐ Сколько input/output tokens проходит в месяц?

☐ Есть ли caching, batch или повышенная цена long context?

☐ Какова стоимость с учётом повторных запросов и tools?

3. Production

☐ Какие latency и rate limits нужны?

☐ Какова data policy?

☐ Нужен managed API или собственный deployment?

4. Eval

☐ Одинаковые inputs для всех моделей.

☐ Считать factual errors и edit burden.

☐ Повторить тест после крупного обновления модели.