CORSEN AI Academy

Grok in Production

Course syllabus

Version 2026-11

Overview

Deploy Grok to production confidently: rate limits, intelligent caching, tokenization, version migration, and best practices.

Learning objectives

- Diagnose a rate limit exceedance and implement exponential backoff on HTTP 429
- Calculate usage cost based on the four token types and the tokenize endpoint
- Interpret cost fields in API responses to track actual spending
- Increase the automatic cache hit rate to improve latency and cost in production
- Migrate an integration from a legacy model to a pinned version without functional regression

Audience and prerequisites

Developers, engineers, and integrators looking to leverage Grok via API in their applications.

Confirmed experience in application development and deployment to production.

Duration and format

Estimated duration : 1 h 30 min. 100% online self-paced course, accessible 24/7 from a browser. Progress is saved, resume anytime. Estimated duration: 1 h 30 min.

Content

- Rate Limits (4 lessons)
- Tokens (4 lessons)
- Cache and Performance (3 lessons)
- Migration (3 lessons)

Assessment

Self-assessment based on the Key takeaways concluding each lesson, and direct practical application of examples on your own examples. A final quiz of 5 corrected questions at the end of the last lesson validates acquired knowledge.

Accessibility

Self-paced course accessible to people with disabilities: text content compatible with screen readers, keyboard navigation, and compliant contrasts. For any adaptation needs, email hello@corsen.ai.

Price and recognition

Free

Certificate of completion issued by Corsen AI.

Access time

Immediate after account creation and email confirmation.

Corsen AI — Paris, France — SIRET 10297040700015
academy@corsen.ai — https://academy.corsen.ai