Deploying Small Language Models (LFWS307)

Course Duration: 24 Hours

TOC

Monitoring with Entrenar

Kubernetes Deployment

Capstone: Multi-Target Deployment

Course Summary