

AI296

Develop, Test, and Run Granite Family LLMs with Red Hat Enterprise Linux AI

Course Description

An introduction to fine-tuning, serving, and using generative AI models with Red Hat Enterprise Linux AI.

Develop, Test, and Run Granite Family LLMs with Red Hat Enterprise Linux AI (AI296) provides students with the fundamental knowledge about how to fine-tune, serve, and use generative AI models tailored to specific business needs. This course helps students build core skills to train models in a secure and private environment.

This course is based on Red Hat Enterprise Linux AI 1.4.

Recommended training

- A basic understanding of AI & ML is recommended but not required.
- Familiarity with the Linux command line.

Course Outline

1. **Generative AI Fundamentals: Capabilities, Challenges, Models, and Techniques**
2. **Define Generative AI, its importance, benefits, and challenges. Learn to select suitable GenAI and LLM models and techniques, considering their capabilities, use cases, and limitations.**
3. **Granite Models For Enterprise Generative AI**
4. **Describe and understand Granite models, and their capabilities to solve enterprise use cases, compared to larger, closed models.**
5. **Training Large Language Models with Red Hat Enterprise Linux AI**
6. **Enable technical and non-technical stakeholders to collaborate in the creation, training, and deployment of large language models (LLMs) optimized for specific business needs by using Red Hat Enterprise Linux AI and IBM Granite.**
7. **Deploying Trained Models with Red Hat Enterprise Linux AI**

8. **Deploy, serve, and operate generative AI models trained with Red Hat Enterprise Linux AI.**