GenAI Code™ Course Syllabus

1 Introduction to Generative AI

The first module introduces the world of generative AI, discussing basic concepts, terminology, and the genealogy of large language models (LLMs). Participants examine key applications of the technology, such as chatbots, image and music generation, and compare open source models and closed commercial solutions. In the practical part, they learn the basics of prompt engineering, constructing effective queries for text and graphical models.

2 GenAI — image

The second module focuses on generating images using diffusion models such as Stable Diffusion. Participants analyze how these models work, create prompts and test them in practice. They work with tools such as Gradio and Streamlit, and experiment with tuning models using Dreambooth. They also learn how to use DALL-E 3 through the OpenAI interface and API. They apply this knowledge to product visualization, interior design and marketing graphics automation projects.

3 GenAI — text

In the third module, participants delve into the topic of text processing by AI models. They learn about the Transformer architecture, explore how LLMs generate content and summaries, and test their application in Q&A systems. Practical tasks include creating tools for automatic document summarization, marketing content generation and personalized language models.

4 GenAI — sound

The fourth module focuses on generating sound, speech and music. Participants analyze how AI models work in this area, and learn how to use tools such as Audiocraft for music creation and Bark/Suno for voice cloning. They also test Text-to-Speech (TTS) systems, such as Parler, and use them to generate realistic voiceovers and automatically produce subtitles. They are working on projects involving the creation of soundtracks, audio effects and synthesized voices.

5 GenAI — video

In the fifth module, participants explore generative video models. They learn to construct animations from text (text-to-video), transform images into video (image-to-video), and apply style transfer (video-to-video). As part of practical tasks, they experiment with the automatic generation of promotional content, product demonstrations and virtual visualizations of 3D spaces.

6 GenAI — multimodal

The sixth module is devoted to multimodal applications of AI, combining text, image, audio and video. Participants test image-to-text, video-to-text, text-to-speech and transcription from image technologies. They analyze ways to integrate different AI models, creating systems for generating visual descriptions, interactive assistants, and applications for automatic presentation processing. They implement projects such as a virtual fitting room, personalized avatars and automatic slide generation.

7 + 8 RAG

The seventh and eighth modules focus on the Retrieval-Augmented Generation (RAG) method to improve the quality of results generated by LLM. Participants analyze problems related to hallucinations and outdated knowledge in models, and learn to use vector databases to improve AI performance. They use these skills to build chatbots tailored to the specific needs of companies and systems for automatic generation of HR documents.

9 Evaluation

The ninth module focuses on the evaluation of generative models. Participants explore metrics used to evaluate LLMs and test tools such as DeepEval, Promptbench and Giskard. They work to assess the quality of generated text, images and multimodal content, optimizing models for specific applications.

10 Monitoring

The tenth module is devoted to monitoring generative AI models. Participants analyze how to ensure the safety and quality of LLM performance using tools such as Llama Guard, Fiddler and Arize Phoenix. They learn to identify potential problems, analyze errors and implement monitoring systems that prevent the generation of unwanted content. They work on projects related to quality control and automatic monitoring of AI models in practical applications.