

Bootcamp Natural Language Processing & Large Language Models

Program Directors

Dr. Aygul Zagidullina
aygul.zagidullina@hslu.ch

Dr. Elena Nazarenko
elena.nazarenko@hslu.ch

Website

hslu.ch/bootcamp-nlp

Overview

This practical bootcamp offers you an opportunity to immerse yourself in the world of natural language processing and to gain practical experience over the course of a week. Over six intensive days, you will have access to expert guidance and support and learn everything from the basic concepts to the latest developments and innovation in the field.

Natural Language Processing (NLP)

This training introduces the foundational concepts of NLP, including tokenization, lemmatization, and part-of-speech tagging. Participants will explore core modeling techniques such as vector representations and probabilistic models. The course also provides an overview of modern transformer-based language models, setting the stage for advanced NLP applications.

LLM Techniques and Applications

Participants will also develop Large Language Model (LLM) skills. The course covers the fundamentals of prompt engineering and explores retrieval-augmented generation (RAG) to enhance LLM contexts. In the framework of hands-on projects, attendees will build LLM-powered agents capable of executing tasks, solving problems, and integrating with various tools.

In addition, the course covers strategies for optimizing LLMs, including fine-tuning embeddings and developing domain-specific RAG chat models. Key topics such as model evaluation, bias mitigation, fairness in NLP, security considerations, and basic MLOps practices are also addressed to ensure responsible and effective deployment of LLMs.

Aims

Participants finish the bootcamp with an in-depth understanding of NLP and LLM concepts, methods and tools, reinforced by practical, hands-on experience. The bootcamp modules build on each other and offer both foundational knowledge and advanced technical skills to apply across a variety of domains. This empowers participants to tackle challenges in data science, AI, and beyond with confidence, and to apply their expertise with a view to driving innovation in their respective fields.

Methodology

This certificate program is delivered in the shape of an intense bootcamp, with focus on day-to-day practical projects. Participants should be ready to get actively involved in hands-on exercises and projects and deepen their understanding and consolidate their skills in the process.

We offer a dynamic learning environment that transcends the traditional classroom setting. This to ensure that our participants develop a profound understanding of Natural Language Processing and the skills to apply it in practice.

Module 1: Foundational Principles of NLP

- Introduction to NLP and preprocessing techniques
- Advanced NLP techniques
- Representation and embeddings
- Probabilistic and sequence models

Module 2: Transformer Models and Fundamentals of LLMs

- Introduction to transformer models
- Transfer learning
- Applying transformers for NLP tasks (chatbots, classification, summarization)
- Introduction to LLMs and prompt engineering

Module 3: RAG and Language Agents

- Introduction to vector databases
- RAG (LlamaIndex/LangChain)
- RAG variations (GraphRAG, CAG)
- Semantic search and document retrieval
- Language agents

Module 4: Parameter-Efficient Fine-Tuning and Model Evaluation in LLMs

- Introduction to PEFT (LoRA, QLoRA)
- Full fine-tuning vs. PEFT
- Fine-tuning (full and PEFT) vs. RAG (what to use when)
- Evaluation metrics for LLMs and RAG
- Bias in LLMs
- Responsible AI

Module 5: MLOps for NLP and Deploying Models to Production

- Introduction to MLOps for NLP and LLMs
- Deploying models for production
- Cloud vs. on-premise deployment strategies
- Testing strategies for AI applications

Module 6: Real-World Applications for NLP and LLMs

- Real-World Applications for NLP and LLMs
- Industry case studies
- Hands-on applied NLP tasks
- Emerging trends and challenges associated with LLMs