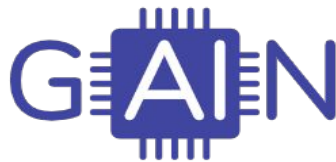




Quick Start with NLP and Hugging Face for Georgian

Beso Mikaberidze (MIMC)



This work was partially supported by the European Union under Horizon Europe project "GAIN" (no.101078950).



2 Agenda

1. What is NLP?
2. Key NLP Tasks
3. Real-Life Applications
4. Impact on Everyday Life
5. Introduction to Hugging Face
6. Overview of the Hands-On Workshop



3 What is NLP?

Natural Language Processing, or NLP, is a field at the intersection of:

1. Computer Science,
2. Artificial Intelligence, and
3. Linguistics.

It focuses on enabling computers to understand, interpret, and generate human language meaningfully.



4 Key NLP Tasks

- **Sentiment Analysis** is a text classification task that determines the emotional tone of the text.
- **Named Entity Recognition** is the token classification task that identifies person, organization, or location names in the text.
- **Part-of-Speech Tagging** is the token classification task that assigns grammatical categories to each token.
- **Machine Translation** is a sequence-to-sequence task that translates text from one language to another.
- **Question Answering** is a language understanding (sometimes also generation) task that builds systems capable of answering questions posed in natural language.



5 Real-Life Applications

NLP is everywhere in our daily lives.

- **Voice Assistants** - It's in the voice assistants like Siri and Alexa that we talk to.
- **Chatbots** - It's in the chatbots on customer service websites.
- **Translation** - It's in services like Google Translate.
- **Spam** - It even filters the spam in our email.



6 Impact on Everyday Life

- Communication: Breaking language barriers.
- Accessibility: Assisting those with disabilities.
- Business: Analyzing customer feedback.
- Education: Enhancing language learning tools.



7 NLP and Georgian Language

- Overcoming the Challenges of Low-Resource Languages
- The Vital Role of NLP Tools in Preserving Georgian
- Unlocking New Opportunities for Technological Growth



8 Introduction to Hugging Face

Hugging Face is an open-source provider of NLP technologies.

It offers a library of pre-trained models and tools that make it easier to build NLP applications.



9 Why Use Hugging Face?

- Extensive model repository
- User-friendly APIs
- Strong community support
- Supports multiple languages, including Georgian



10 Overview of the Hands-On Workshop

Hugging Face workshop about text classification
Adapted on Georgian Sentiment Analysis

Steps:

- Load and process a datasets
- Fine-tune, evaluate and test a pre-trained model on your data
- Push/pull a model to/from the Hugging Face Hub
- Infer a model using the Hugging Face pipeline



11 Resources and Materials

Workshops:

- [Georgian Sentiment Analysis](#) (Original source: [Hugging Face O2 text classification](#))
- [Hugging Face workshops](#)

Datasets:

- [Georgian Sentiment Analysis](#) (Source [Paper](#))
- [IMDB Movie reviews](#)

Hugging Face docs:

- [NLP Course](#)
- [Transformers](#)

Hugging Face videos

- [Datasets overview](#)
- [Tokenizers Overview](#)
- [Word-based tokenizers](#)
- [Character-based tokenizers](#)
- [Subword-based tokenizers](#)
- [The Trainer API](#)



Let's Get Started!