

# Lab-4

## Group-20

Arthur Humblot - s350872, Bekhzod Anvarov - s316063,  
Ghita El Belghiti - s361421

December 10, 2025

- 1 Task 1: Dataset Characterization**
- 2 Task 2: Tokenization**
- 3 Task 3: Model Training**
- 4 Task 3: Model Training**