



Model architecture — 5gram - - - gpt2 - - - rnn - - - vanilla

Training vocabulary — bllip-lg — bllip-md — bllip-sm — bllip-xs — gptbpe