Meeting — Summary LLM

Bas Donkers, Dennis Fok, Finn-Ole Höner

May 7, 2024

Updates

Improved the optimization: `.generate()` is the wrong method, need `.forward()` $\rightarrow$ wrong conditionals! Now maximizes the likelihood and is faster!

- Space looks different now; Does maximizing the likelihood collapse the encoding space?
- Maybe even further performance gains by using `past_key_values`?
- Idea: LLM judge to assess grammar / quality of slogans?
- (MLFlow)

MPhil: Need to start working towards the paper now

Admin:

- ☐ There seem to be issues with the access to the data for me now
- ☐ SLURM application
- ☐ Set appointment for pre-defense
- ☐ Find third assessor
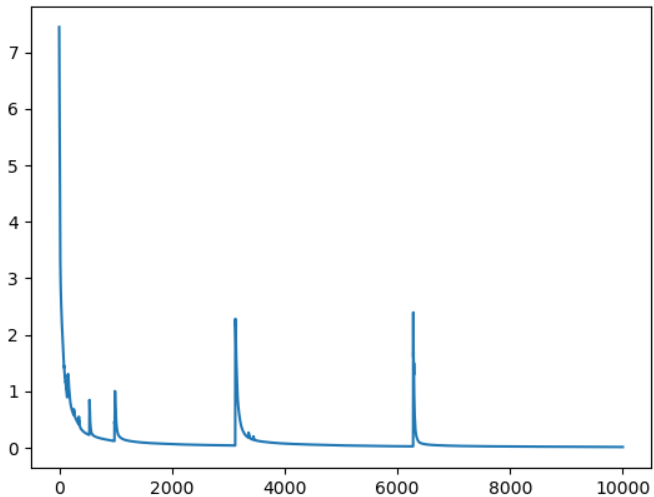- ☐ Complete data management survey

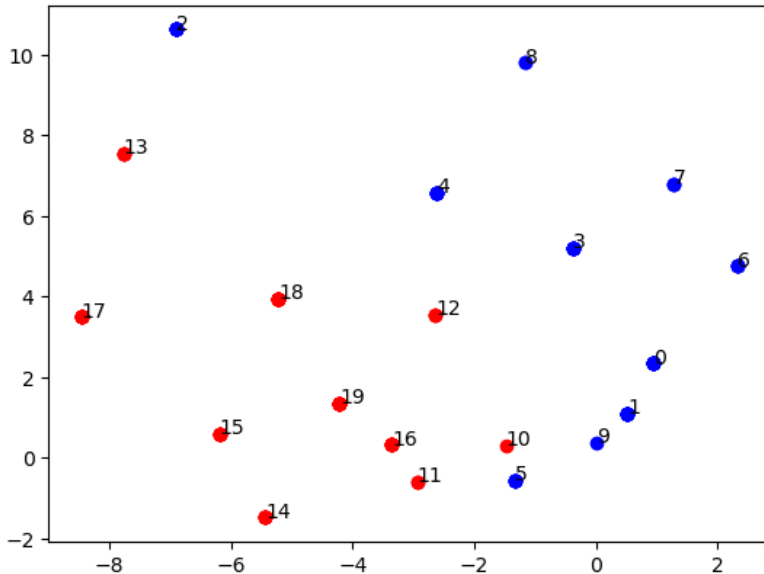**Figure 1:** Maxmimum Likelihood (nLL)
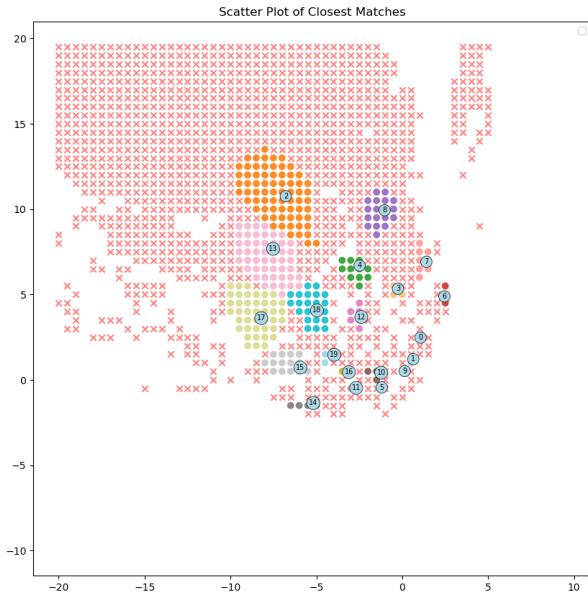
Figure 2: Encoded data points

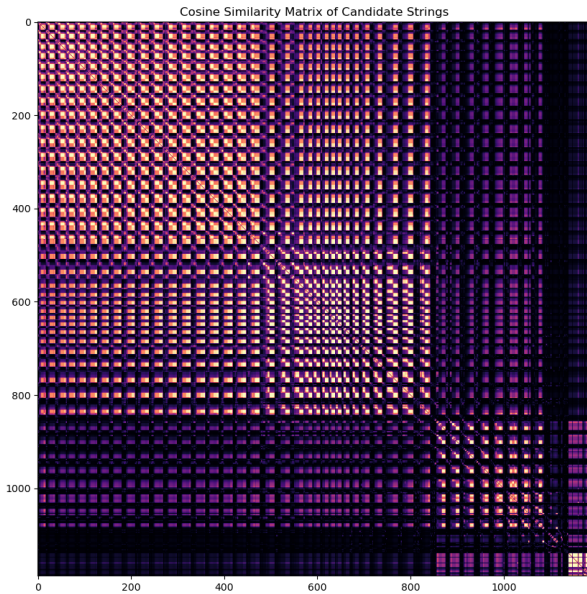**Figure 3:** Exploring the encoding space with generation, x are candidates

Figure 4: Cosine similarity between the candidate points