Reviewer: Elshan Naghizade

Submitter: Eljan Mahammadli

Project: Name Generation with Autoregressive Character-level Language Modeling

OVERALL THOUGHTS

The project has a conspicuous business appeal as a backend service for a web application for an AI-based tool for parents to generate names for their newborns based on the provided keywords. Moreover, despite the abundance of similar services in the English-speaking world, an app targeted towards the Azerbaijani audience taking into account the Azerbaijan-specific naming traditions does hold a potential value.

The technical part of Eljan's proposal is somewhat feasible with the caveat being that the models would be trained on relatively small Azerbaijani datasets scraped off of only a small number of baby name portals.

However, there seems to be no adequate way to systematically validate the efficiency of the proposed project since subjective company or baby names can only be evaluated manually in an ad-hoc manner.

STRONG SIDES

- ✓ Will be based on modern NLP models even though starting with just simple N-Grams.
- ✓ The APIs required for the development of the system are readily available.

TO BE IMPROVED

- ✓ Due to the time limitations (11 weeks), the models can only be trained on small datasets, therefore, putting enormous strain on the quality of the data collected
- ✓ Addressing the issue of the generation of "un-original" names is complicated as there are no clear metrics of originality (except if the project is to be directed towards businesses looking for a company name where the focus lands on copyrights

SUGGESTIONS

- ✓ Manually assure the quality of the dataset since the web-scraped ones usually are of sub-par quality.
- ✓ The only way to validate the platform is to launch a small web service where a user can leave a star-based numeric grade for the generated name. (Take into account that it requires the project to be finished way earlier than the deadline to leave time for such manual validations).