Filling the Gap: Decoding of Word Embeddings for Generation of Coherent New Words

Safa AlSaidi, Amandine Decker, Stephanie Monteiro

M2 – Software Project

Reminder of our aim

- Apply decoder to the regression task (solving analogies)

$$A : B :: C : X \xrightarrow{X=?} A : B :: C : D$$
$$e.g.\ \ star : stars :: cat : X \rightarrow cats$$

- Current output: vectors ($\neq$ word)
- Aim: transform these vectors into words

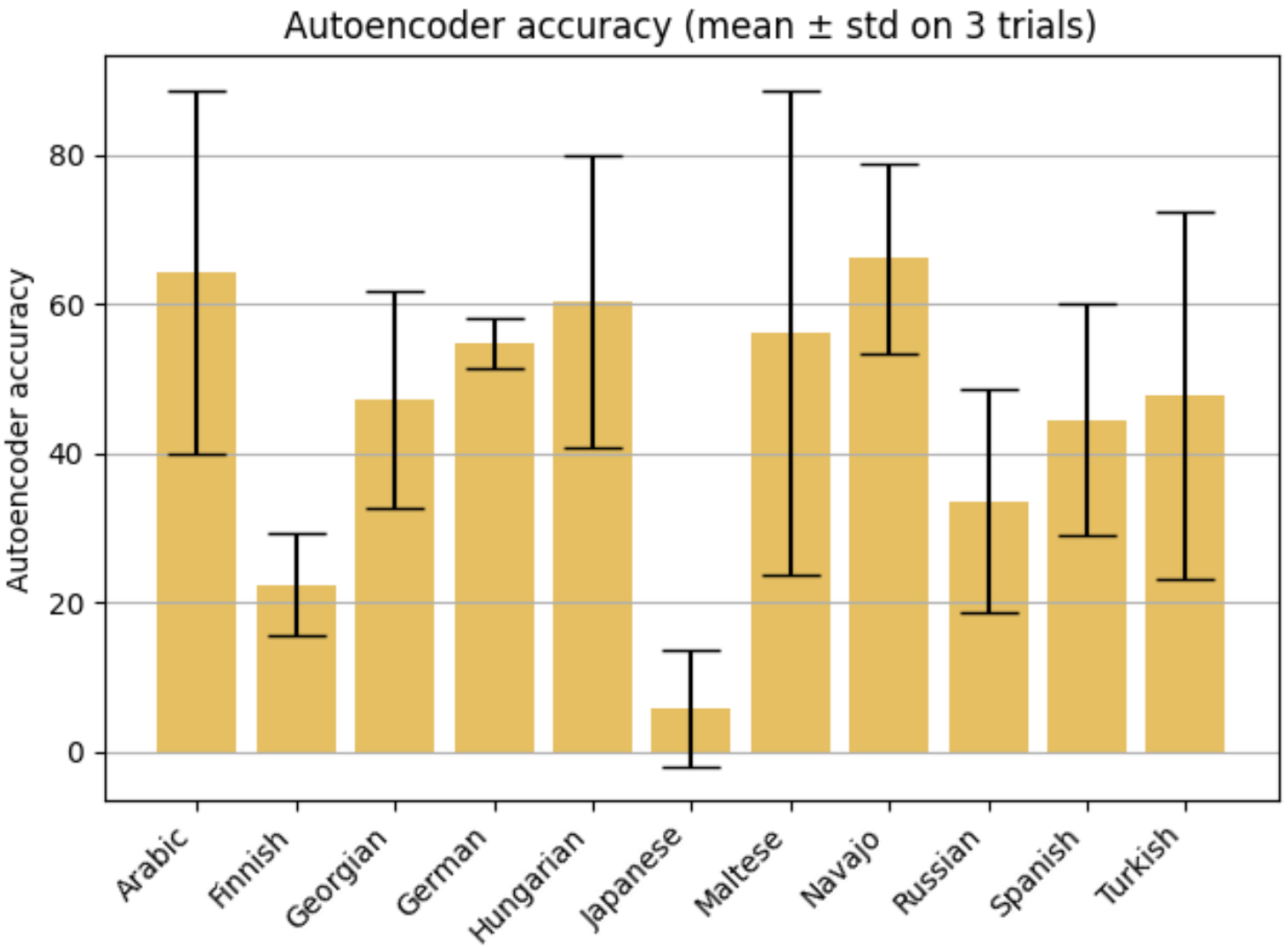
What we managed to do?

1. Build the autoencoder
2. Applied it on the classification task
3. Obtained a primary set of results
4. Started analyzing what was faulty in our decoder

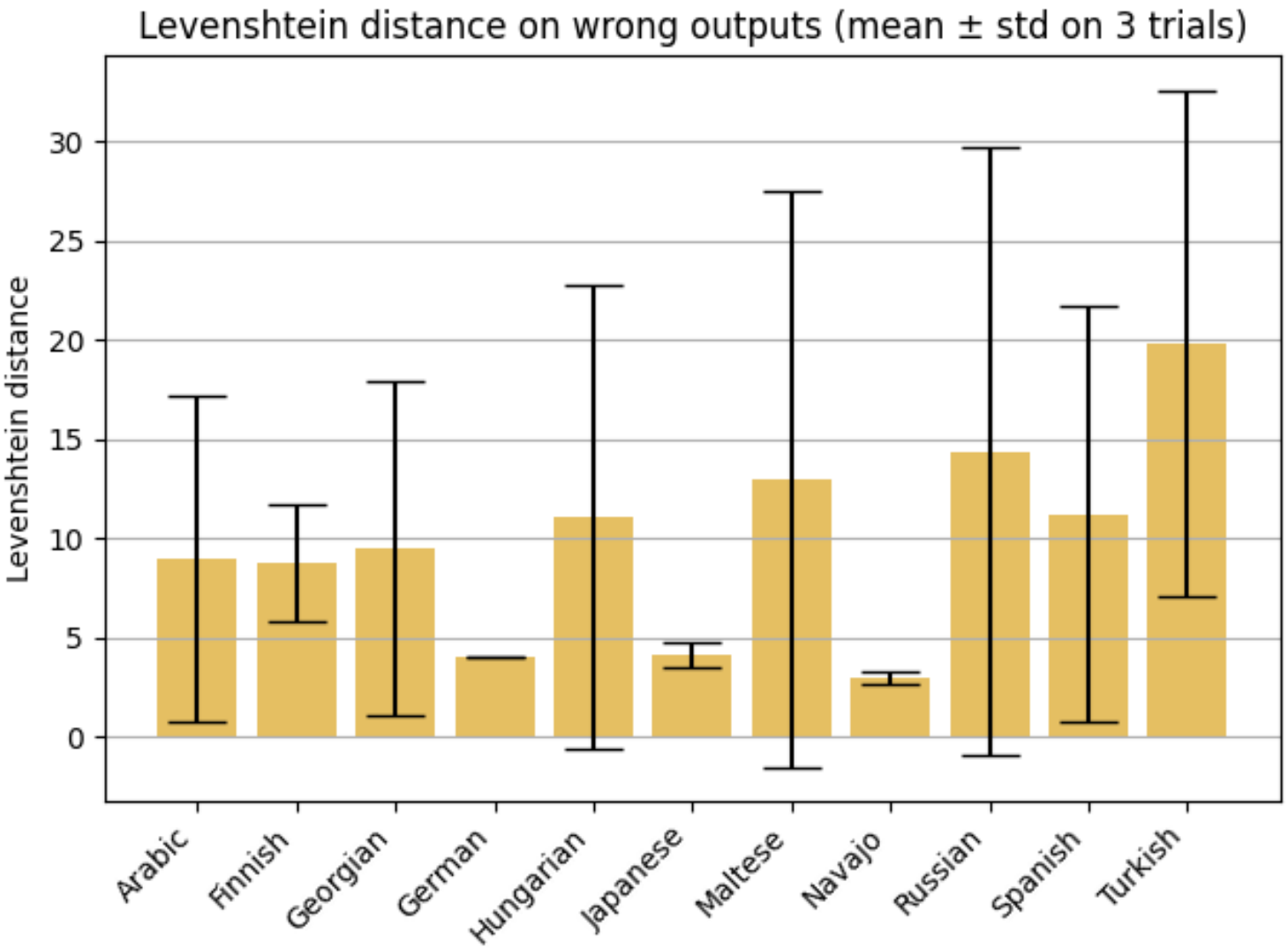
Autoencoder

1. Encoder
   - CNN based model
   - filters of sizes 2 to 6 $\rightarrow$ encode subwords ?
2. Decoder
   - 1 layer GRU
   - hidden layer input = word embedding
   - cells input = last decoded character's ID (start with 'BOS' ID)

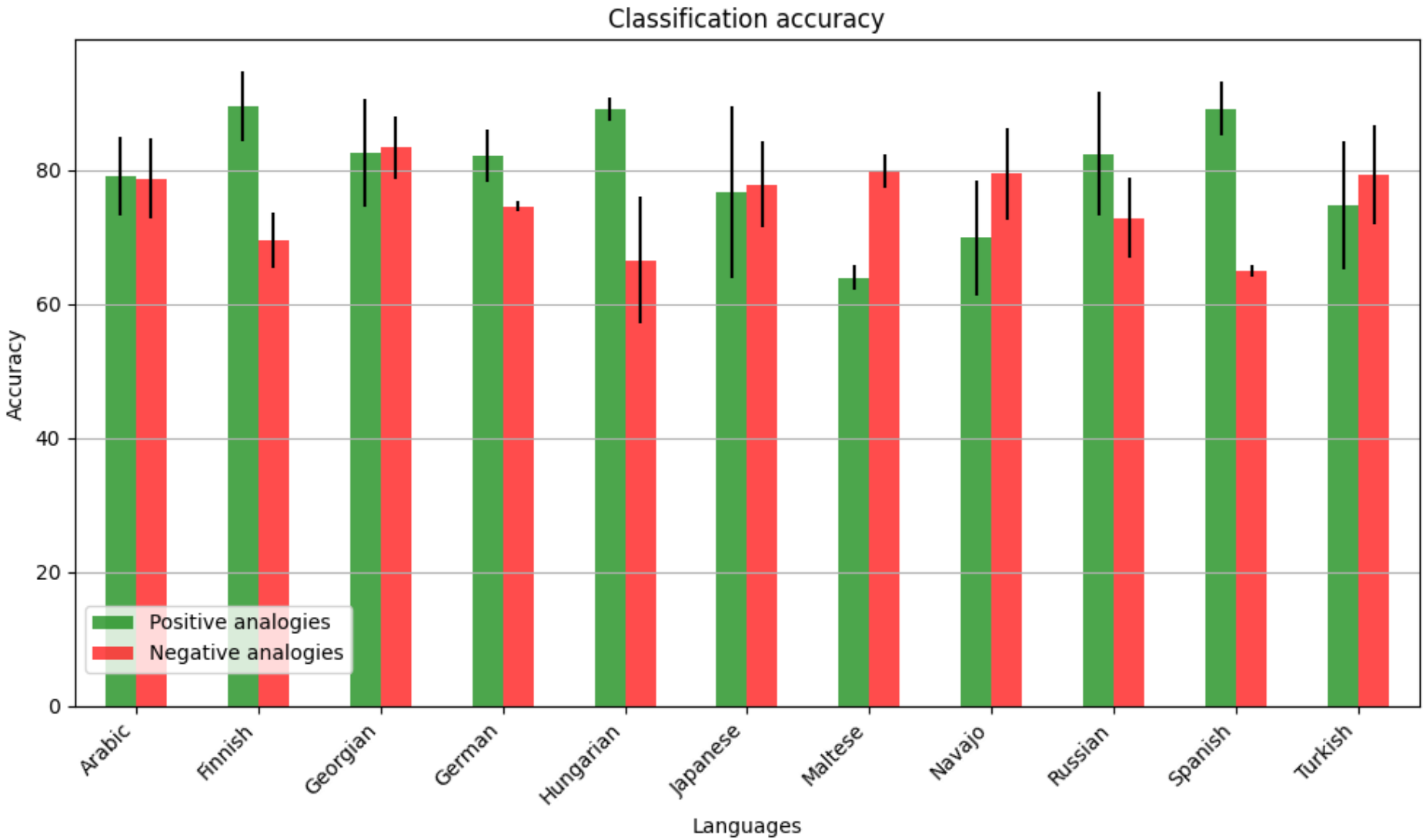
Autoencoder accuracy (input word = output word)



Levenshtein distance when output word != input word



Accuracy for the classification task



Discussion

Some modifications we tried:

- bidirectional: outputs empty words
- concatenate the embeddings to the input cells (= last character ID embedding + word embedding): random outputs of max_len

What went wrong with the decoder

1. Built files per language with only incorrect words
2. Evaluated manually by fluent/native speakers
3. Result: obtained some leads that would help us investigate how the model produces the words

Comments and leads to investigate

Summary of the comments:

- prefixes, inflections, and suffixes: exist in the languages;
- incorrect words: variability among the languages (real words: German around 65%, Russian none, Spanish 15%, Arabic 10%)

Leads to investigate:

- how the model decodes the central part of a word
- the ratio of words that exist and don't exist per language
- force the model to work with the embeddings (seem to prioritise intern coherence)

How we plan to proceed

1. Proceed with two languages: Arabic and German
2. Query a dictionary to know the ratio of Exist/DNE words

Future work

- ~~22 Nov - Regression model + decoder / Variational auto-encoder~~
- 10 Dec - Qualitative analysis / ~~Multilingual model~~
- 14 Jan - Application docker & webpage
- 3 Feb - Report

Future work

- ~~22 Nov - Regression model + decoder / Variational auto-encoder~~
- 10 Dec - Qualitative analysis / Regression model + decoder
- 14 Jan - Application docker & webpage
- 3 Feb - Report

شكرا جزيلا

Thank you

Merci

អរគុណ

Obrigado

Our decoder structure

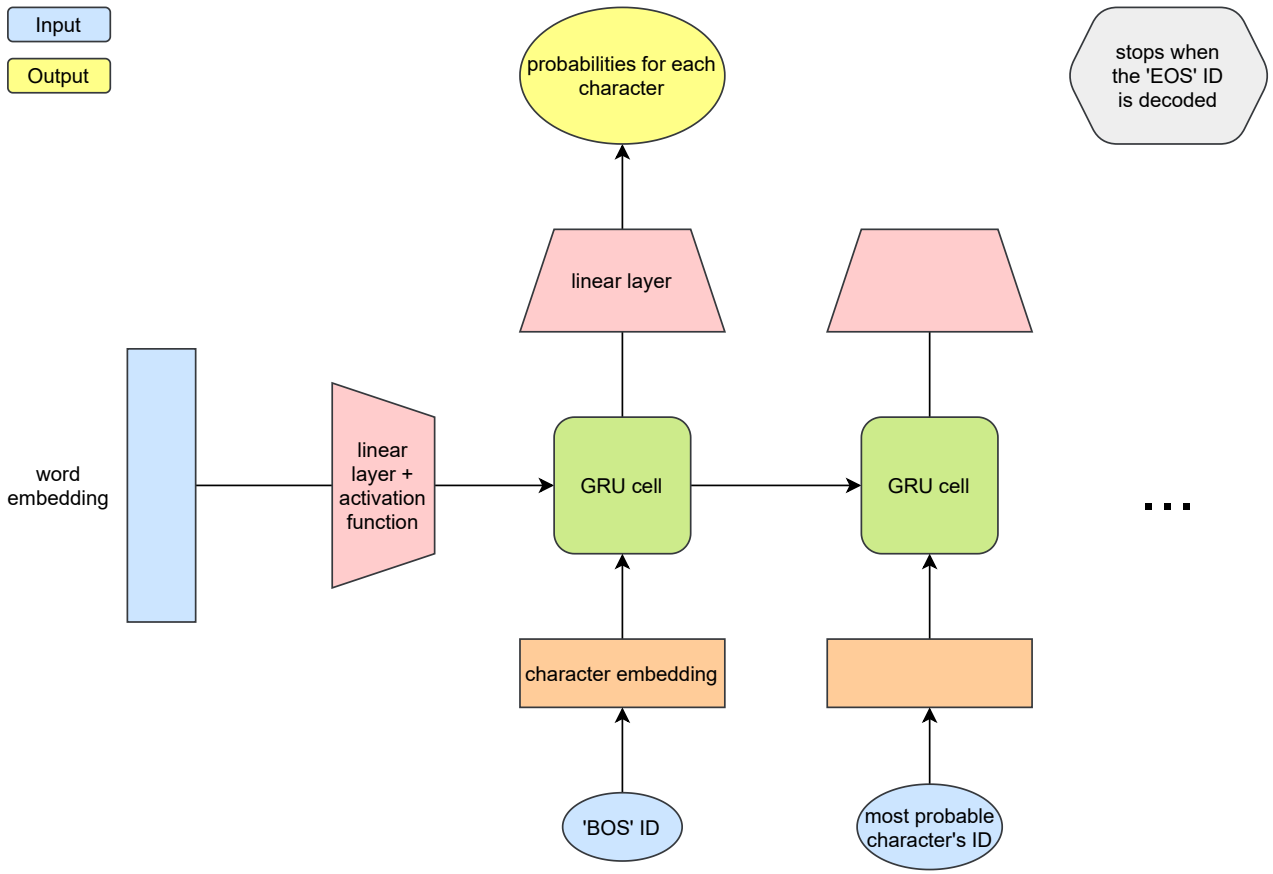


Figure: Our GRU based decoder

Inspired by this blogpost `https://rajatvd.github.io/Generating-Words-From-Embeddings/`