

Fine-tuning a pretrained machine learning model to identify declension errors in ICELANDIC

↑TEXT↑

By Isak Grimsson, supervised by Matthew Yee-King

ABSTRACT↑

In this paper we fine-tuned a BERT model for declension error detection (DED). We present our method to create synthetic data to retrain our model, and analyse the results.

MOTIVATION↑

Declension governs how words change their form dependent on the context of the sentence. Latin, a famously declension-happy language, has twenty declension patterns. Icelandic has seventy-three. Making minor declension errors does occasionally occur for native Icelandic speakers, but for non-natives, they are a constant source of frustration and/or embarrassment. At the start of this project and until the addition of Icelandic into ChatGPT-4 there were no tools available for DED.

DATA SET CREATION↑

Originally we had a dataset of 35,000 correctly and 1000 incorrectly declined sentences. But that was not enough data for the model to effectively learn, so we created a synthetic dataset for the model to use, by combining the original dataset with a database of declensions, we algorithmically created multiple incorrect sentences for each correct sentence in our original dataset.

What does fine tuning a pretrained model mean?

A pretrained BERT model has been slowly trained on a lot of data, and has a rich understanding of how words are connected, but it's not very useful, it's trained to predict a missing word, for example:

Input 'This is a [missing word] example sentence'

Output ['simple' 60%, 'clear' 25%, 'confusing' 10%, 'unclear' 5%]

To create a more useful model, we fine-tuned it for DED by using the pretrained model as a starting point and providing it with new data so it could be trained to classify sentences as correct or incorrect.

RESULTS↑

ChatGPT-4's F-1 score = 0.7865 Our Model's F-1 score = 0.9214

		Actual Class	
		Positive	Negative
Predicted Class	Positive	2691	693
	Negative	768	343

		Actual Class	
		Positive	Negative
Predicted Class	Positive	3430	556
	Negative	29	480

CONCLUSIONS AND FUTURE WORK

The DED model has very promising results, convincingly outperforming ChatGPT-4.

Possible future improvements:

- A way to reliably source or generate correct sentences
- Fine-tuning a classification threshold (AUC-ROC curve)
- Making the model available for use online
- Comparing the model's performance with native and non-native Icelandic speakers.