University *of Ljubljana*
Faculty *of Computer and Information Science*

Natural language processing course: Analysis and comparison of translation errors and biases in LLMs

Urška Roblek, Tjaša Nadoh

Abstract

The abstract goes here. The abstract goes here.

Keywords

Natual language processing, bias, translation errors, bias in LLMs

Advisors: Aleš Žagar

Introduction

For the purpose of the project, we compiled a dataset of texts containing information about politics for the purpose of finding out how a language model will translate this genre of text and if there might be some political bias in the translations. The LLMs evaluated in this project were ChatGPT, Deepseek and Mistral AI. They were chosen because they are currently popular and widely used and also because they come from three different countries (USA, France and China), which have different cultures as well as legislation. Besides looking at the quality of translations and biases that might be present, we will evaluate if the bias is similar among these three models or if the fact that they are from different countries might affect the bias.

The bias in language might seem insignificant to some, but it can impact how we view things like minority groups or gender roles. Because large language models are made on human language and our texts, they often produce a biased language which is a reflection of our own. As discussed in the article Biases in Large Language Models: Origins, Inventory, and Discussion by Roberto Navigli et al., a language bias occurs in two ways. It starts with the data that is selected to be used to train an LLM, and shows in the output that it produces. Bias can show in gender, religion, age, sexual orientation, race and many more. An example of a gender bias for example would be an LLM using a she when talking about a nurse, but a he when talking about a doctor, even if the text never specifically specifies what gender they are. This usually happens because the data used for training contained most material, where nurses were women and doctors were men, so it leans toward that ”opinion” as well. When training the LLM we should try to make it less biased by choosing different kinds of data, as well as correct any bias that might appear, when fine tuning the model.

Methods

For the purpose of the research we have collected 10 texts from Slovene and English newspapers. The newspapers were picked on the criteria that they were talking about topics commonly found in political conversation, such as global warming, immigration, tariffs, protests etc. The newspapers that published these articles were well-known newspapers for Slovene and English public.

We copied the articles in Word files and put the files in picked language models.

ChatGPT

The first prompt used to translate the texts was simple. ”Translate this text into [preferred language]”. The model then translated the texts according to his own preference and conclusion.

Firstly, we manually analyzed the text. We noticed that texts that came from left-leaning or neutral sources remained more

or less true to the source. The Some examples:
migration - migracije
deportations - deportacije
On the other hand, when analyzing one article that came from a right-leaning source and had quite an opinionated content, the model changed some of the text.
propagandno pozivati - call for (leaves out the propaganda part)
v nasprotju z njimi - *despite it* (makes the text italic)
zaslužni (nanašajoč se na osebe) - ”deserving” (puts it in quotation marks)
srečnežev - recipients

When another very opinionated text was translated, this time from a left-leaning source, it once again leaned towards neutralising the words.
mencavo - hesitantly
z mačeto lomastiti - swing a machete

Equations

You can write equations inline, e.g. $\cos\pi = -1$, $E = m \cdot c^2$ and α, or you can include them as separate objects. The Bayes's rule is stated mathematically as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{1}$$

where A and B are some events. You can also reference it – the equation 1 describes the Bayes's rule.

Lists

We can insert numbered and bullet lists:

1. First item in the list.
2. Second item in the list.
3. Third item in the list.

- First item in the list.
- Second item in the list.
- Third item in the list.

We can use the description environment to define or describe key terms and phrases.

Word What is a word?.

Concept What is a concept?

Idea What is an idea?

Random text

This text is inserted only to make this template look more like a proper report. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam blandit dictum facilisis. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Interdum et malesuada fames ac ante ipsum primis in faucibus. Etiam convallis tellus velit, quis ornare ipsum aliquam id. Maecenas tempus mauris sit amet libero elementum eleifend. Nulla nunc orci, consectetur non consequat ac, consequat non nisl. Aenean vitae dui nec ex fringilla malesuada. Proin elit libero, faucibus eget neque quis, condimentum laoreet urna. Etiam at nunc quis felis pulvinar dignissim. Phasellus turpis turpis, vestibulum eget imperdiet in, molestie eget neque. Curabitur quis ante sed nunc varius dictum non quis nisl. Donec nec lobortis velit. Ut cursus, libero efficitur dictum imperdiet, odio mi fermentum dui, id vulputate metus velit sit amet risus. Nulla vel volutpat elit. Mauris ex erat, pulvinar ac accumsan sit amet, ultrices sit amet turpis.

Phasellus in ligula nunc. Vivamus sem lorem, malesuada sed pretium quis, varius convallis lectus. Quisque in risus nec lectus lobortis gravida non a sem. Quisque et vestibulum sem, vel mollis dolor. Nullam ante ex, scelerisque ac efficitur vel, rhoncus quis lectus. Pellentesque scelerisque efficitur purus in faucibus. Maecenas vestibulum vulputate nisl sed vestibulum. Nullam varius turpis in hendrerit posuere.

Figures

You can insert figures that span over the whole page, or over just a single column. The first one, Figure 1, is an example of a figure that spans only across one of the two columns in the report.

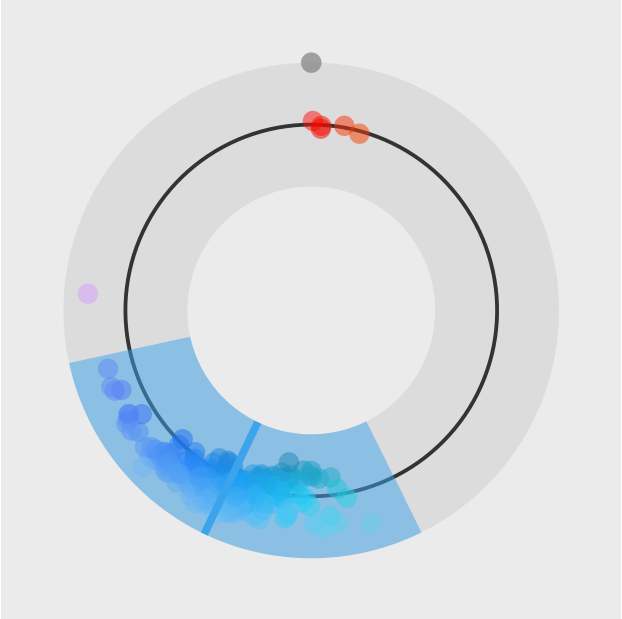

Figure 1. A random visualization. This is an example of a figure that spans only across one of the two columns.

On the other hand, Figure 2 is an example of a figure that spans across the whole page (across both columns) of the report.

Tables

Use the table environment to insert tables.

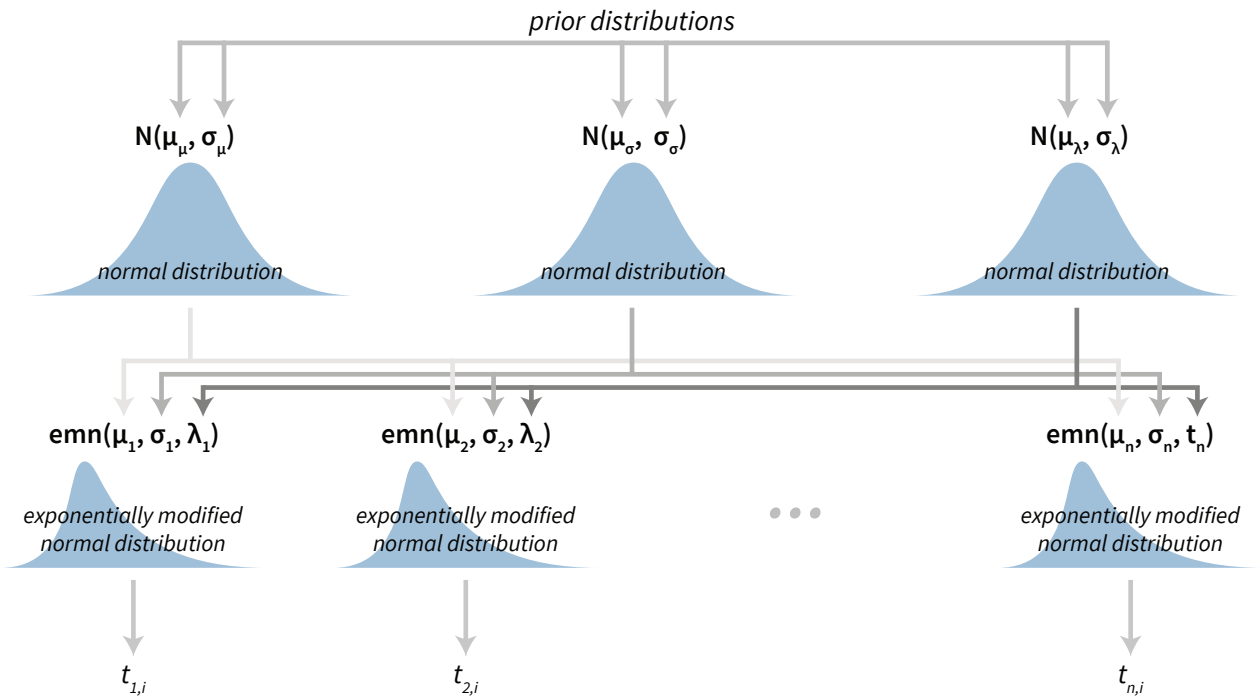

Figure 2. Visualization of a Bayesian hierarchical model. This is an example of a figure that spans the whole width of the report.

Table 1. Table of grades.

Name		
First name	Last Name	Grade
John	Doe	7.5
Jane	Doe	10
Mike	Smith	8

Code examples

You can also insert short code examples. You can specify them manually, or insert a whole file with code. Please avoid inserting long code snippets, advisors will have access to your repositories and can take a look at your code there. If necessary, you can use this technique to insert code (or pseudo code) of short algorithms that are crucial for the understanding of the manuscript.

Listing 1. Insert code directly from a file.

```
import os
import time
import random

fruits = ["apple", "banana", "cherry"]
for x in fruits:
  print(x)
```

Listing 2. Write the code you want to insert.

```
import(dplyr)
import(ggplot)

ggplot(diamonds,
           aes(x=carat, y=price, color=cut)) +
  geom_point() +
  geom_smooth()
```

Results

ChatGPT

Upon manually analyzing the translations that came from ChatGPT, the texts were translated well. But when analyzing the texts that came from very opinionated sources, be it left or right leaning, it tended to leave out or change some of the words, which made the texts lose some of their original tones.

More random text

This text is inserted only to make this template look more like a proper report. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam blandit dictum facilisis. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Interdum et malesuada fames ac ante ipsum primis in faucibus. Etiam convallis tellus velit, quis ornare ipsum aliquam id. Maecenas tempus mauris sit amet libero elementum eleifend. Nulla nunc orci, consectetur non consequat ac, consequat non nisl. Aenean vitae dui nec ex fringilla malesuada. Proin elit libero, faucibus eget neque quis, condimentum laoreet urna. Etiam at nunc quis felis pulvinar dignissim. Phasellus turpis turpis, vestibulum eget imperdiet in, molestie eget neque. Curabitur quis ante sed nunc varius dictum non quis nisl. Donec nec lobortis velit. Ut

cursus, libero efficitur dictum imperdiet, odio mi fermentum dui, id vulputate metus velit sit amet risus. Nulla vel volutpat elit. Mauris ex erat, pulvinar ac accumsan sit amet, ultrices sit amet turpis.

Phasellus in ligula nunc. Vivamus sem lorem, malesuada sed pretium quis, varius convallis lectus. Quisque in risus nec lectus lobortis gravida non a sem. Quisque et vestibulum sem, vel mollis dolor. Nullam ante ex, scelerisque ac efficitur vel, rhoncus quis lectus. Pellentesque scelerisque efficitur purus in faucibus. Maecenas vestibulum vulputate nisl sed vestibulum. Nullam varius turpis in hendrerit posuere.

Nulla rhoncus tortor eget ipsum commodo lacinia sit amet eu urna. Cras maximus leo mauris, ac congue eros sollicitudin ac. Integer vel erat varius, scelerisque orci eu, tristique purus. Proin id leo quis ante pharetra suscipit et non magna. Morbi in volutpat erat. Vivamus sit amet libero eu lacus pulvinar pharetra sed at felis. Vivamus non nibh a orci viverra rhoncus sit amet ullamcorper sem. Ut nec tempor dui. Aliquam convallis vitae nisi ac volutpat. Nam accumsan, erat eget faucibus commodo, ligula dui cursus nisi, at laoreet odio augue id eros. Curabitur quis tellus eget nunc ornare auctor.

Discussion

Use the Discussion section to objectively evaluate your work, do not just put praise on everything you did, be critical and exposes flaws and weaknesses of your solution. You can also explain what you would do differently if you would be able to start again and what upgrades could be done on the project in the future.

Acknowledgments

Here you can thank other persons (advisors, colleagues ...) that contributed to the successful completion of your project.

References

Roberto Navigli, Simone Conia, and Björn Ross. 2023. Biases in Large Language Models: Origins, Inventory, and Discussion. J. Data and Information Quality 15, 2, Article 10 (June 2023), 21 pages. https://doi.org/10.1145/3597307

data used for translation and bias analysis: https://www.theguardian.com/us-news/2025/mar/20/trump-executive-order-education-department
https://arxiv.org/pdf/2412.04782