Part six is Using LIWC to Compare Linguistic Features of AI vs Human Answers

This is my task completion process

Background and Motivation

This task aims to analyze the language differences between two platforms Gemini and Stack Overflow

The Goal is Determine which answer is more rational, and suitable for knowledge transfer

We will Use the LIWC tool to analyze three types of language features: Cognitive Processing, Positive Emotion, Negative Emotion

To Compare Gemini and Stack Overflow in these three aspects

Methodology Overview

First, I cleaned the data and extracted the hit words using LIWC dictionaries.

Here are my dictionaries

It helps LIWC count how many words match each category.

Proportion calculation

Next, we carry out Proportion calculation. The process is as follows.

Count the number of words belonging to each type of keyword in each answer text, calculate their proportion in the total vocabulary, and output the percentage.

That is, the proportion of positive, negative and cognitive keywords matched in the answers of Stack Overflow and Gemini respectively in the total number of answer words.

Output example

This is an example of a line taken from the output file.

This includes keyword matching lists and percentage data for human and AI answer texts.

Average value calculation & visualization

Then perform Average value calculation & visualization

Calculate the average of the proportion of each column of Stack Overflow and ai. And present the data in a bar chart to make the data more readable.

We can see that Gemini uses more cognitive processing words.

Stack Overflow uses slightly more emotional words.

T-test

But we don't know whether the difference in this result is significant or whether it is a result of chance. Therefore, we use the t-test to verify the conclusion.

The results of the t-test are shown in the figure. According to the value, the difference is not accidental and is significant.

Conclusion and reflection

Finally, draw conclusions and reflections.

Through language feature analysis, the differences in cognitive and emotional expressions between AI and human responses were revealed.

The results show that AI is more rational and more suitable for knowledge transfer, supporting the overall goal of the project.

The LIWC tool has a good reference value in academic analysis.

In the future, it can combine more data and introduce texts from multiple languages ​​and fields for more extensive analysis.

The above is the entire content of our group report.