In my pursuit of another degree, I am taking a course COSC478, “Applications of Generative AI using Large Language Models (LLM)”. The course description can be seen [here](https://courseinfo.canterbury.ac.nz/GetCourseDetails.aspx?course=COSC478).

Just so you know, ChatGPT is an LLM.

I have an assignment to come up with a way of assessing an LLM’s ability to translate English to a second language. For my efforts, I have chosen Dutch as the second language. I now need an expert in this second language to translate 4 sentences. I am asking you to act as my expert. 😊

I am to compare your translation with what I obtain from the LLM. The comparison is made with a method prescribed in the class. I will be assessed on the way that I manually and programmatically implement this comparison. My task for you, as my expert, is to translate the 4 sentences prescribed in the table that follows. Please fill in the second column of the table with your Dutch translation on each of the four lines.

Thank you for your consideration!

David Ewing

+64276427681 – TEL

|  |  |
| --- | --- |
| **English sentence to be translated to Dutch** | **Your translation (Dutch)** |
| A transformer architecture is a deep learning model architecture that uses self-attention mechanisms to process input data. | Transformer architectuur is een diep leren architectuur model dat een aandachtsmechanisme gebruikt om invoergegevens te verwerken. |
| Tokenisation is the process of converting text into smaller units that can be processed by a large language model. | Tokenisatie is het proces om tekst te converteren naar kleinere eenheden die verwerkt kunnen worden door een groot taalmodel. |
| Both cats and dogs make wonderful and loyal companions for life. | Zowel katten als honden zijn fantastische en loyale metgezellen voor het leven. |
| Cats love to nap around the house and dogs like chasing sticks. | Katten houden ervan om dutjes te doen in huis en honden vinden het leuk om achter stokken aan te rennen. |

1:

Transformer architectuur is een diep leren architectuur model dat een aandachtsmechanisme gebruikt om invoergegevens te verwerken.

2:

Tokenisatie is het proces om tekst te converteren naar kleinere eenheden die verwerkt kunnen worden door een groot taalmodel.

Or

Tokenisatie is het proces om tekst te converteren naar kleinere eenheden die door een groot taalmodel verwerkt kunnen worden.

(When converting is necessary for processing by LLM, we use the word ‘zodat’ between ‘eenheden’ and ‘die’.

But if that was your intention, I bet you used the sentence  ‘so that they can be processed’ …

3:

Zowel katten als honden zijn fantastische en loyale metgezellen voor het leven.

4:

Katten houden ervan om dutjes te doen in huis en honden vinden het leuk om achter stokken aan te rennen.