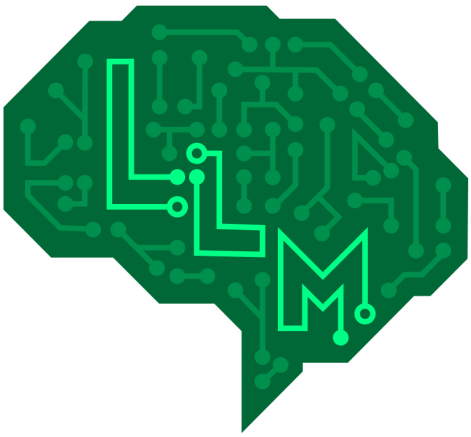
P01

We are developing an LLM that is small, strong in Japanese language, and customizable

Large language model “tsuzumi”

IOWN Pick Up

NTT version Large Language Models



Background

LLMs have attracted attention for their advanced and flexible language processing capabilities, but the operational cost of running models is a challenge. The huge amount of computational resources required to train models also makes customization difficult.

Summary

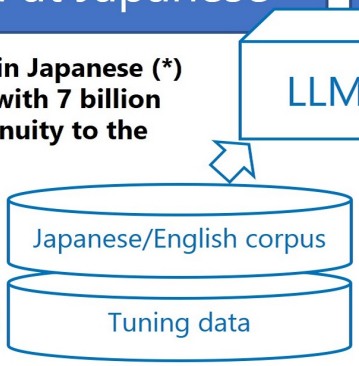
We are developing a small, low-power LLM. It excels at understanding the Japanese language through the use of large volumes of high-quality data in Japanese. Flexible customization is possible by learning additional data specific to particular industries, organizations, or individuals.

Feature 1: Small and good at Japanese

Achieved top-level accuracy in Japanese (*) with a version of our model with 7 billion parameters by applying ingenuity to the training data

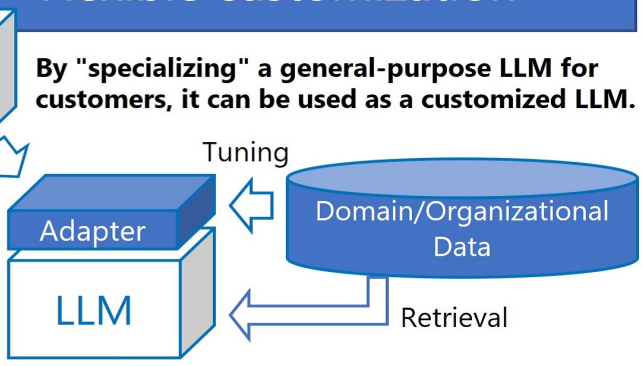
Over a trillion texts

Accumulation of many years of research and development and newly created data



Feature 2: Flexible customization

By "specializing" a general-purpose LLM for customers, it can be used as a customized LLM.



(*) Evaluated with Rakuda, Japanese Vicuna QA, JGLUE

Features

- We achieved top-level accuracy in Japanese task evaluation with a version of our model with 7 billion parameters
- We developed a unique model using data and knowledge accumulated through more than 40 years of natural language processing research at NTT
- We use adapters that can be trained with additional knowledge, be specialized to specific industries and organizations, and can flexibly customize the LLM

Future_benefits

Our LLM is specialized for business domains and users, to achieve a "sustainable world" through collaboration and mutual growth between AI and humans.

Exhibiting Company

NIPPON TELEGRAPH AND TELEPHONE CORPORATION

Contact

rdforum-exhibition@ml.ntt.com