

人工知能アプリケーションの信頼性・安全性

報告書番号：R24JDG20112

利用分野：研究開発

URL：https://www.jss.jaxa.jp/ar/j2024/27140/

● 責任者

清水太郎, 研究開発部門第三研究ユニット

● 問い合わせ先

伊藤光一郎(ito.koichiro@jaxa.jp)

● メンバ

伊藤 光一郎

● 事業概要

JAXA 内のドキュメント自動解析に向けて、画像と言語が同時入力となる VLM (Vision-Language Model) の開発を行い、特に図に対する対話が可能なローカルモデルの開発を進めています。実験では、図認識に関する新たな指示学習方法の検討を行い、その初期成果を人工知能学会と MIRU2025 に投稿しました。

● JAXA スーパーコンピュータを使用する理由と利点

VLLM を構成する LLM の学習に当たっては、多くの VRAM が必要であり、V100 は古い GPU ではあるもののマルチノードによる学習をするために利用した。

● 今年度の成果

本年度は JAXA スーパーコンピュータ (JSS) を活用してローカル LLM のトレーニング・学習を実施した。とりわけ既存ライブラリの活用と、マルチノード学習に向けた実装の確認を行った。LLM の実験は Nvidia Ampere アーキテクチャ以降の GPU でサポートされている Bfloat16 形式での数値表現で行われるが、タスクによっては V100 での Float16 でも学習が可能であった。ただし Overflow することが多く、次世代スパコンに期待する点も多かった。

● 成果の公表

なし

● JSS 利用状況

● 計算情報

プロセス並列手法	MPI
スレッド並列手法	OpenMP
プロセス並列数	1 - 32
1 ケースあたりの経過時間	8 時間

● JSS3 利用量

総資源に占める利用割合※1 (%) : 0.01

内訳

計算資源		
計算システム名	CPU 利用量(コア・時)	資源の利用割合※2 (%)
TOKI-SORA	0.00	0.00
TOKI-ST	1.04	0.00
TOKI-GP	5,568.27	0.09
TOKI-XM	0.00	0.00
TOKI-LM	0.00	0.00
TOKI-TST	1.40	0.00
TOKI-TGP	12,568.75	86.52
TOKI-TLM	0.00	0.00

ファイルシステム資源		
ファイルシステム名	ストレージ割当量(GiB)	資源の利用割合※2 (%)
/home	245.00	0.17
/data 及び/data2	15,310.00	0.07
/ssd	2,510.00	0.13

アーカイバ資源		
アーカイバシステム名	利用量(TiB)	資源の利用割合※2 (%)
J-SPACE	0.00	0.00

※1 総資源に占める利用割合 : 3つの資源(計算,ファイルシステム,アーカイバ)の利用割合の加重平均

※2 資源の利用割合 : 対象資源一年間の総利用量に対する利用割合

● ISV 利用量

ISV ソフトウェア資源		
	利用量(時)	資源の利用割合※2 (%)
ISV ソフトウェア(合計)	0.00	0.00

※2 資源の利用割合：対象資源一年間の総利用量に対する利用割合