

100B/1Tパラメータからなる大規模マルチモーダル基盤モデルの構築

実施者 株式会社Preferred Elements

事業概要 日本語性能に優れ、言語・画像・音声に対応したマルチモーダル100Bモデルの開発および1Tパラメータの言語モデルの事前学習の検証を行う。

事業成果

データ構築
学習手法の確立

高効率な
大規模学習の実現

マルチモーダル化
に向けた技術開発

- 実施内容**
- ① 100B モデルの事前学習・追加学習
 - ② 指示学習
 - ③ 画像モーダル向けの事前学習・追加学習
 - ④ 音声モーダル向け追加学習
 - ⑤ 1Tモデルの事前学習の検証

- 開発・検証する基盤モデル**
- 100Bマルチモーダルモデル
 - 言語・画像・音声に対応
 - 一部タスクで世界最高レベルの性能
 - 1T言語モデル
 - グローバルレベルでも大規模なモデルの学習を検証



社会実装イメージ

一部モデル・ノウハウ等の公開・提供



ソースコード
(フィンチューニング用など)



モデル
事前学習済
100Bモデル
ウェイト

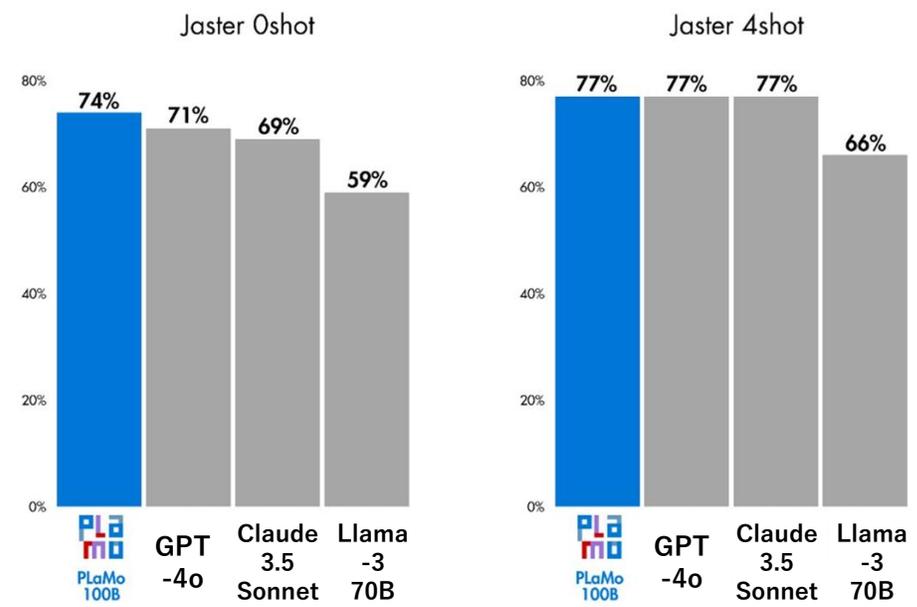


開発ノウハウ
マルチモーダル
化及び1Tモデル
学習

- 自社ビジネスとしての展開**
- 基盤モデルのAPI提供
 - 基盤モデルのライセンス提供
 - 基盤モデル及びAPIに付帯するビジネス（エンジニアリング・コンサルティング）

国際的競争力のある大規模言語モデルを
日本でゼロから構築できる技術基盤を確立

日本語性能を評価するベンチマークJaster(*) 0shotおよび4shotにて、
GPT、Claude、Llamaを上回り**最高性能を記録**



(*) 日本語性能を評価する標準ベンチマークで、GENIAC特有のカテゴリセット（NLI, QA, RC, MC, MR, FAの6カテゴリ）で評価。GENIACでは、Weights & Biases社のllm-leaderboard/g-leaderboardで主な基盤モデルをベンチマーク評価を実施。評価スコアは、ベンチマークテストに全問正解すると1点満点となる。