

Performance Engineer

Performance Engineer

募集職種

人材紹介会社

マイケル・ページ・インターナショナル・ジャパン株式会社

求人ID

1595366

業種

ソフトウェア

雇用形態

正社員

勤務地

東京都 23区

給与

700万円 ~ 1500万円

更新日

2026年06月08日 10:00

応募必要条件

キャリアレベル

中途経験者レベル

英語レベル

無し

日本語レベル

ネイティブ

最終学歴

大学卒：学士号

現在のビザ

日本での就労許可は必要ありません

募集要項

本ポジションは、生成AIモデルの研究開発におけるワークロード全体の性能分析・改善を担うエンジニア職です。GPUや分散環境を活用し、大規模言語モデルの性能向上と開発効率化を推進します。

企業情報

クライアントは、生成AIおよび大規模言語モデルの研究開発を主軸とするテクノロジー企業です。日本語に特化したAIモデルの開発と社会実装を通じて、次世代のデジタル基盤構築を目指しています。

職務内容

- ・ MLOpsにおける各種ワークロード（学習・蒸留・推論・評価）の性能分析および最適化
- ・ GPUリソースやストレージの割当管理および効率利用の推進
- ・ 新規手法・ツールの調査、評価、導入によるMLOps改善
- ・ 大規模計算処理における課題特定および改善施策の実行

条件・待遇

- ・フレックス勤務制度
- ・社会保険完備（健康保険・厚生年金・雇用保険など）
- ・各種手当（通勤手当、残業手当など）
- ・福利厚生サービス、財形貯蓄、団体保険
- ・確定拠出年金制度（401K）
- ・完全週休2日制（土日祝）、年末年始休暇、有給休暇

To apply online please click the 'Apply' button below. For a confidential discussion about this role please contact Yuki Ishihara on +81 3 6627 6071.

スキル・資格

- ・大規模計算処理のワークロードに関する知識
- ・計算機科学に関する深い理解
- ・PythonおよびC++での開発・解析経験
- ・機械学習または深層学習の知識・開発経験
- ・専門外の人にも分かりやすく説明できるコミュニケーション力

会社説明

クライアントは、生成AIおよび大規模言語モデルの研究開発を主軸とするテクノロジー企業です。日本語に特化したAIモデルの開発と社会実装を通じて、次世代のデジタル基盤構築を目指しています。