

職務経歴書

2024 年 07 月 09 日現在

氏名 戴 雄介

☎ +1-217-305-0416 | @ xdail2@illinois.edu | [in linkedin.com/in/xiongjie-dai](https://www.linkedin.com/in/xiongjie-dai) | github.com/Xiongjiedai

■職務要約

フリーランスとして活動中の人工知能/機械学習エンジニア。様々なハードウェアプラットフォーム上での大規模言語モデル (LLM) の推論最適化に専念し、パフォーマンス評価のための革新的なクラウドベース戦略を適用。llama.cpp ツールキットを使用し、AI アプリケーションでの GPU ベンチマークの理解と適用を向上させ、業界から高い評価を受ける (GitHub で 600 以上のスターを獲得)。また、イリノイ大学アーバナ・シャンペーン校にて大学院コースアシスタントとして勤務し、教員と協力して R、Python、SQL を使用した魅力的な機械学習コースの内容とコーディング評価を設計。学問的な問題を解決し、学生の成績評価に貢献。

教育背景においては、統計学の修士 (GPA 3.84/4.00)、経済学と応用数学の二重学位 (GPA 3.93/4.25) を取得。厳格な定量スキルと数学、統計学の知識を活かし、データ駆動型の問題解決に強みを持つ。多言語能力 (英語、中国語) を生かし、国際的な環境でも活躍。

■職務経歴

2022 年 09 月～2022 年 12 月 イリノイ大学アーバナ・シャンペーン校		
事業内容：大学院コースアシスタント		契約社員 として勤務
2022 年 09 月 ～2022 年 12 月	本社 ・ 教員と協力して、R、Python、SQL で魅力的な機械学習のコース内容と革新的なコーディング評価を設計しました。 ・ 週に一度のオフィスアワーで学生の成績を評価し、学問的な問題を解決しました。コースに関する苦情は一切ありませんでした。	
2023 年 07 月～現在 フリーランス		
事業内容：人工知能/機械学習エンジニア		正社員 として勤務
2023 年 07 月 ～現在	スキル: Git, Shell, RunPod, LLM, GenAI ・ LLaMA モデルの包括的なテストを実施し、llama.cpp ツールキットを使用して NVIDIA の GPU と Apple Silicon を比較し、パフォーマンス評価のための革新的なクラウドベース戦略を適応させ、デバイス上でのローカル大規模言語モデル (LLM) の概念実証を確立しました。これにより、業界からの賞賛を受け、GitHub で 600 以上のスターを獲得しました。 ・ 様々なモデルサイズの推論速度を文書化・分析し、効率的な LLM 推論のための最適な設定に関する実用的な洞察を提供しました。	

■プロジェクト

2022 年 05 月 イリノイ大学アーバナ・シャンペーン校	
事業内容：CNN を用いたデータ拡張の効果の説明	
	<div>スキル: Python, PyTorch, Matplotlib, Gradio</div> <div><ul style="list-style-type: none">・ さまざまなシードで三つの異なる畳み込みニューラルネットワーク（CNN）モデルを開発・評価し、モデル性能への拡張技術の影響を特定することに焦点を当てました。二項の猫犬分類チャレンジに焦点を当てました。・ 正則化（ドロップアウト、L1、L2）を適用して過学習を避け、二項クロスエントロピーをコスト関数として、SGD を最適化アルゴリズムとして使用しました。データ拡張を反転させることで、最高の CNN モデルの正確性は 59.6% から 74.8% に向上しました。</div>
2022 年 12 月 イリノイ大学アーバナ・シャンペーン校	
事業内容：アマゾンレビューおよび薬データセットの感情分析	

<p>スキル: R markdown, Word2Vec</p> <ul style="list-style-type: none"> ・ Amazon および Drug Review データセットに対する感情分析を進め、四つのクラシックな埋め込みおよびNLP 方法 (BoW, Word2Vec, GloVe, fastText) を比較しました。FastText が最高のパフォーマンスを達成しました (Amazon での正確性 86.49%, Drug での正確性 78.69%)。 ・ テキストの正規化およびベクトル化のためのデータ前処理技術を使用しました。一般的な単語をフィルタリングするために2 つの単純な t 検定を使用しました。分類のためにナイーブベイズおよびランダムフォレストアルゴリズムを活用しました。
2022 年 12 月 イリノイ大学アーバナ・シャンペーン校
事業内容: スネークゲームの強化学習
<p>スキル: Python, Numpy</p> <ul style="list-style-type: none"> ・ ステート、アクション、リワードのマルコフ決定プロセス (MDP) フレームワークを定義し、時間差 (TD) Q 学習アルゴリズムを使用して AI エージェントがスネークゲームをマスターするようにしました。 ・ 新しい状態を探索することと既知の状態を利用することの間にバランスをとる探索ポリシーを実装しました。時間の経過とともに Q 値更新プロセスを最適化するために減衰学習率を使用しました。 ・ モデルの堅牢性と適応性を確保するために厳格なテストフェーズを通じてエージェントのパフォーマンスを分析しました。
2023 年 05 月 イリノイ大学アーバナ・シャンペーン校
事業内容: ニューラル機械翻訳: RNN およびトランスフォーマーモデル
<p>スキル: Python, PyTorch</p> <ul style="list-style-type: none"> ・ PyTorch を使用してスペイン語から英語へのテキスト翻訳に焦点を当てたニューラル機械翻訳 (NMT) システムの開発をリードしました。 ・ GRU とアテンションメカニズムを使用して RNN を設計・実装し、エンコーダーおよびデコーダーコンポーネントを構築しました。トランスフォーマーモデルの翻訳品質を向上させるために、マルチヘッドアテンションと位置エンコーディングを使用しました。 ・ RNN モデルの BLEU-4 スコア 0.058、トランスフォーマーモデルの BLEU-4 スコア 0.059 を達成しました。

■スキル

Git	git clone、git push、git add、git commit、git checkout、git pull を熟知
Shell	shell script の作成が可能
Python	PyTorch、TensorFlow、NumPy、Pandas、Matplotlib、Gradio を熟知
R	ggplot2、Tidyverse などのデータサイエンス基本操作を熟知
SQL	data pipeline の構築を熟知

■資格

中国 C1 カテゴリー運転免許	2019 年 10 月取得
アメリカ Class D 運転免許	2023 年 08 月取得

■自己PR

<機械学習プロジェクトのリーダーシップとイノベーション>

フリーランスとして活動中の機械学習エンジニアであり、大規模言語モデル (LLM) の推論最適化とパフォーマンス評価のためのクラウドベース戦略を導入しました。このプロジェクトで、llama.cpp ツールキットを使用し、NVIDIA GPU と Apple Silicon の比較分析を行い、デバイス上での LLM の概念実証を成功させました。この成果は業界から高く評価され、GitHub で 600 以上のスターを獲得しました。これらの経験から、機械学習とデータサイエンスの最前線で活躍し、新たな技術革新をリードする能力を持っています。

<データ駆動型問題解決の専門知識>

イリノイ大学アーバナ・シャンペーン校で統計学の修士号を取得し、高度な定量分析とデータ管理スキルを身につけました。学術的な背景を活かし、データサイエンスプロジェクトでの深い洞察と戦略的アプローチを提供します。R、Python、SQL を駆使し、教育コースの設計から実践的なデータ分析まで、幅広く対応可能です。

<多言語能力と国際的な視点>

ビジネスレベルの英語、基礎レベルの日本語、そして母国語である中国語を流暢に操ることで、国際的なプロジェクトや多様な文化背景を持つチームでのコミュニケーションにも強みを持っています。これにより、世界各国のクライアントやプロジェクトに柔軟に対応し、グローバルな視点からの問題解決を実現します。

以上