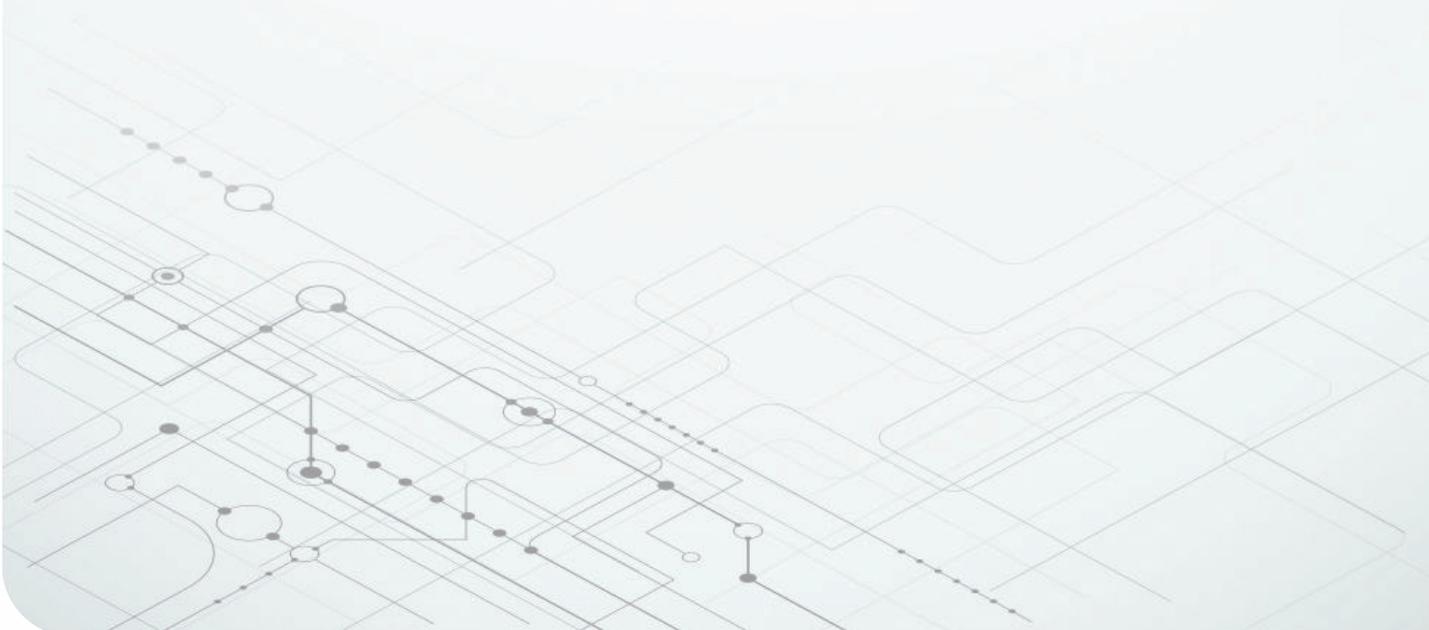


組み込みプロセッサの採用でエッジ側インテリジェンスを強化



Alec May
Product Marketing Engineer
Embedded Processors



概要



1

はじめに

このホワイトペーパーでは、ネットワークのエッジで人工知能 (AI) を活用する目的と利点について説明し、組み込みプロセッサとソフトウェアの進歩がもたらす多様なアプリケーションへの AI の実装をこれまで以上に容易にする方法を説明します。



2

エッジ AI とは？

半導体の世界でエッジ AI がこのような話題になっている理由や、ニュースに出てくる AI との違いをご確認ください。



3

AI はどのようにエッジへと移行しているのか

組み込みハードウェア / ソフトウェア エンジニアが直面してきた課題と、テキサス・インスツルメンツがそれらに対処する方法について解説します。



4

エッジ AI のスケーラビリティ

スケーラビリティと再利用性に関する課題を解決するために、テキサス・インスツルメンツが提供するハードウェアとソフトウェアをご確認ください。

はじめに

AI は、ニッチなテクノロジーから、人々が日常的にやり取りするものへと移行し、エンジニアリングやテクノロジーの分野を超えて成長しています。この傾向により、ほぼすべての業界の企業は、AI を活用して効率性を高め、コストを削減し、製品の機能を向上させる方法を検討するようになりました。広く利用されているクラウドベースの AI ソリューションのアクセシビリティと使いやすさにより、ほとんどの人が AI 向けに設計されたモデルやツールに簡単に取り組むことができます。

しかし、すべての AI イノベーションがクラウドで起こっているわけではありません。組み込みプロセッサ設計の分野での技術の進歩に伴い、ノート PC や携帯電話などのコンシューマ製品や、ビデオドアベルのようなバッテリー動作アプリケーション、

車載システムのビジョン処理、エネルギーインフラや産業用システム向けのモーターなどに搭載された電子機器で AI の能力が発揮されつつあります。

エッジ AI とは、データの出所に近い場所で AI モデルをローカルで実行する能力であり、電子機器の応答性、効率、信頼性、安全性を強化します。このクラウドからエッジへの移行を可能にする組み込みプロセッサは、デジタル信号処理 (DSP) 向けの特化型コアなどのコンポーネントを統合したもので、AI をエッジ側に取り込むために必要な時間と専門知識を最小限に抑える、使いやすい GUI ベースのツールがサポートしています。

このホワイトペーパーでは、エッジ AI の進化と利点、およびそれを可能にしているハードウェアとソフトウェアの進歩について説明します。

AI とは

AI とは、ある種の知能や推論を示す機械の能力のことです。今日のほとんどの人が AI について考えるとき、彼らはしばしばテキストや画像生成器、またはビデオゲームの仮想対戦相手を想像します。しかし、最も単純なアルゴリズムでさえ、技術的には文字通りの意味での AI の例です。

AI の広範さとその複数の使用事例により、**図 1** に示すように、機械学習やディープラーニングなどのいくつかのサブドメインが生まれています。

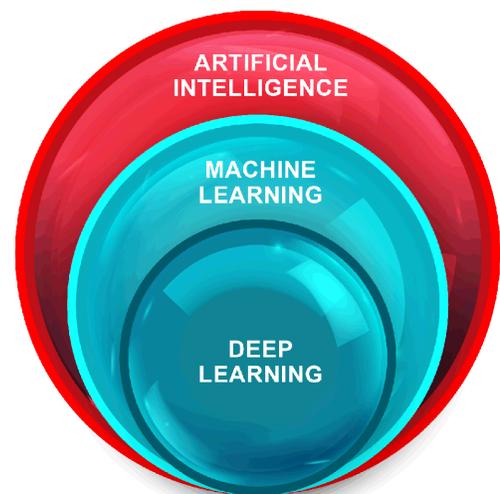


図 1. 異なる AI サブドメイン間の関係。

組み込みアプリケーションで使用する AI の大半は機械学習で、これは、たとえば、画像データの一般的なパターンを分析することにより、歩行者と障害物を認識する、など機械とアルゴリズムが問題を解決する方法を「学習」するサブドメインです。機械学習モデルは、大量のトレーニングデータ、またはすでにラベル付けされているデータを受け取ることで学習することもできます。このトレーニングプロセスにより、機械学習モデルはデータ内のパターンを識別し、将来の推論に使用できます。

機械学習の分野では、非常に複雑な問題を正確に解決できる能力があるため、ディープラーニングが非常に一般的な実装の 1 つになりました。ただし、そのような実装には多くの計算リソースが必要です。ディープラーニングは、人間の脳のニューロンに触発されたデータモデルである多層ニューラルネットワークを使用します。これにより、製品設計者は、自分では

識別できないパターンを認識できるソリューションを作成できます。

エッジ AI とは？

AI とそのサブドメインは、通常、クラウドまたはローカル ハードウェアで処理を実行できます。従来、クラウドベース AI が一般的に使用されてきましたが、これは強力な AI を実行するために必要なコンピューティング能力は、大規模サーバー以外では容易に達成できなかったからです。しかし、組み込みプロセッサの計算能力と電力効率が向上してきた結果、エッジ AI の普及が進んでいます。

図 2 は、データの受信と処理、およびクラウドベースのリソースとのやり取りに関して、エッジ AI とクラウド AI がどのように異なっているかを示しています。

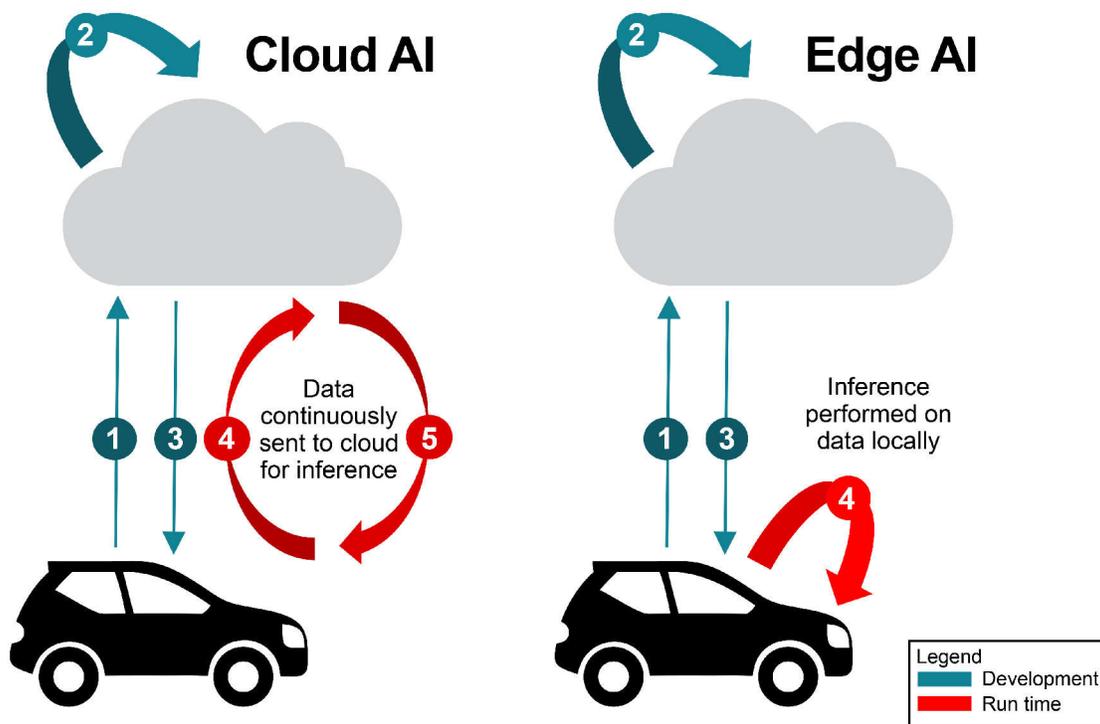


図 2. クラウドベース AI とエッジ AI の比較。

エッジ AI は多くの場合、開発中のモデルトレーニングにクラウドまたはデスクトップのリソースを使用します。モデルを組み込みデバイスに展開すると、新しいデータについて、独立してローカルでモデルの推論と決定を行うことができます。

最近まで、AI の特に有用な事例では、平均的なコンシューマエレクトロニクスが提供できる処理能力を上回る処理能力を必要としていました。これが、機械学習モデルのトレーニングや実装が、多くの場合クラウドベースのリソースで行われてきた理由です。クラウドベースの実装はハードウェアへの投資を最小限に抑えることで利便性を提供しますが、AI の採用も制限します。クラウドベースの AI 実装は、クラウドアクセス、つまり、ネットワーク接続のないアプリケーションでは一切、使用できません。さらに、エッジ AI は、クラウドベース AI と比較して、セキュリティ、安全性、応答性の向上にも貢献します。

半導体の進歩と AI ツールチェーンの改良に伴い、組み込みプロセッサとマイコン (MCU) に AI ソリューションを直接実装することで、AI をエッジに移行することができるようになりました。AI をエッジ側に導入すると、データを供給するセンサの近くで計算と AI の推論を実施することになりますが、電子機器が収集するセンサ データの量が継続的に増加していることを考えるとこのことは重要です。クラウドとの間での大量のデータの送受信はコストがかかり複雑で単一障害点となる可能性があるため、データ量が増加するとクラウドのみのリソースは実用的ではなくなります。

ネットワーク エッジで AI モデルを実行すると、たとえば衝突検出のため、車内に搭載しているカメラセンサなど、センサデータに基づく推論や意思決定のレイテンシを通常、短縮できます。エッジ AI 機能を活用すると、車両はクラウドからの推論を待たずに、推論をより高速に実行し、外部刺激にリアルタイムで応答できるようになります。

ネットワーク接続への依存度の低減など、エッジ AI にはクラウドベース AI と比較した場合、その他にもいくつかの利点があります。エッジ AI は、クラウドへのアクセスが不可能なアプリケーションで使用でき、ネットワークの停止によって引き起こ

される潜在的なダウンタイムを最小限に抑えることができます。また、クラウドベースの AI にはネットワーク接続が必要なため、アクセスのサービス料が繰り返し発生する可能性があります。消費者向け製品を設計する際のビジネスモデルとしては困難な場合があります。

AI はどのようにエッジへと移行しているのか

組み込みプロセッサや使いやすさに配慮した Web ベース ソフトウェアの進歩により、より多くの設計者が自らの設計の中でエッジ AI 機能を具体化できるようになる様子を見ていきます。

エッジ AI のハードウェア革新: これまでは、組み込みプロセッサの処理と消費電力の制約に加え、社内のプログラミングに関する高度な専門知識とリソースが原因で、エッジ AI の幅広い採用が制限されていました。AI 計算のパフォーマンス要件を満たすことができる組み込みデバイスは、多くの場合サイズが大きく、多くの電力を消費して大量の熱を発生しました。

近年、エッジ AI を実現するために必要な計算処理の高速化を可能にする、専用のハードウェアソリューションが登場しています。ただし、これらのハードウェアソリューションにはいくつかのトレードオフがあり、エッジアプリケーションでの幅広い採用を妨げていました。グラフィックス処理ユニット (GPU)、フィールドプログラマブル ゲートアレイ (FPGA)、特定用途向け集積回路 (ASIC) などの専用ハードウェア ソリューションは素晴らしい成果を達成できてきましたが、通常は消費電力 (特に GPU や FPGA の場合) や柔軟性の欠如 (ASIC の場合) によって制限がありました。

エッジ AI 機能を統合したより新しい組み込みプロセッサは、この電力の制限やコストに関する検討事項に対処します。これらのデバイスの多くには、ニューラル プロセッシング ユニット (NPU) やデジタル信号プロセッサ (DSP) などの内蔵部品が搭載されています。

図 3 に、モーターとソーラー アークの障害検出用に NPU を内蔵したテキサス・インスツルメンツの C2000™ マイコンの例を示します。

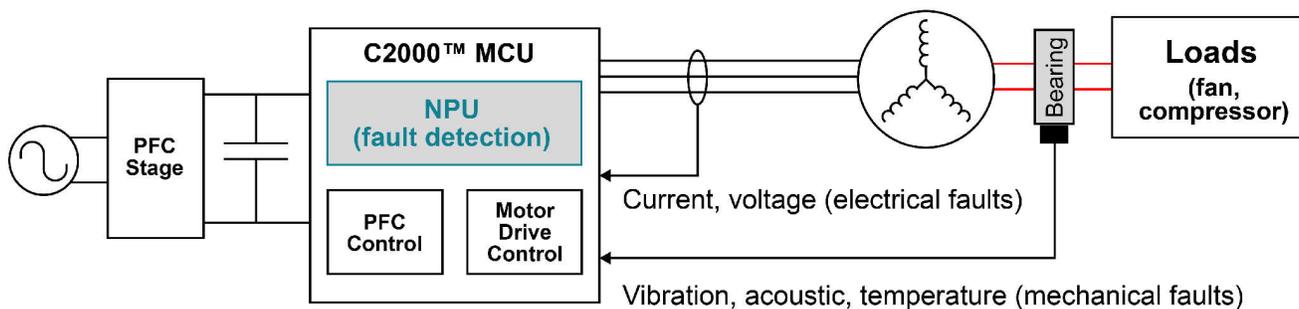


図 3. リアルタイム制御システム内のエッジ AI 対応障害監視ソリューション。

NPU は、組み込みプロセッサの設計に統合できる専用ハードウェアアクセラレータの一種であり、必要な計算レベルを実現できます。たとえば、[TMS320F28P550SJ](#) など、テキサス・インスツルメンツの TMS320F28P55x シリーズのリアルタイム C2000 マイコンは、モーター ベアリングとソーラー アーク障害検出用のオンチップ NPU を搭載しており、これらのプロセスをリアルタイム モーター制御に使用されるメイン CPU から解放します。障害を早期に検出すると、ダウンタイムや修理コストの削減と同時に、安全性の向上にも役立ちます。モーター障害検出機能は、暖房・換気・空調 (HVAC)、ファクトリオートメーション、ロボット、電気自動車などのアプリケーションにとって不可欠です。これらのアプリケーションは、さまざまなセンサから集計された時系列データを使用して、高精度の推論を行い、システム内で異常を迅速に検出することができます。技術記事『[エッジ AI 対応 MCU を用いたリアルタイム制御システムでのシステム障害検知の最適化](#)』をご覧ください。エッジ AI 対応 C2000 マイコンを使用してモーターとアークの故障の検出を簡素化する方法を確認できます。

ビジョン処理などの認識ベースのアプリケーションでは、マイクロプロセッサはビジョン パイプラインに必要なコンポーネントと、必要とされる AI の性能を満たすことができる NPU を内蔵する必要があります。[AM67A](#) や [TDA4VM](#) のようなテキサス・インスツルメンツの分析プロセッサは、ディープラーニングアクセラレータを内蔵しており、AI の性能と消費電力を最適化できます。これらのディープラーニングアクセラレータは、それぞれ C7x DSP で構成されています。このアクセラレータは、マトリクス乗算アクセラレータを内蔵しており、ビジョン処理で一般的に使用されるニューラル ネットワークの処理性能を大幅に向上します。このアーキテクチャは、テキサス・インスツルメンツの DSP 技術を活用し、AI の計算を均質なコンピュ

ーティング エlement に集中化することで、効率的なデータ移動を実現し、消費電力を最適化します。

組み込みプロセッサをできるだけ簡単に設計できるように、テキサス・インスツルメンツのような半導体メーカーは、システムオンモジュールのような「すぐに導入できる」ハードウェアコンポーネントを開発する [サードパーティーパートナー](#) とも密接に協力しています。これらのハードウェアソリューションは、特定の組み込みプロセッサとコア コンポーネントを 1 枚の PCB でサポートするように設計されています。

エッジ AI ソフトウェアの革新: ハードウェアの観点では、組み込みデバイスでの効率的な AI 計算の進歩に加えて、オープンソースコミュニティや半導体メーカーも、プログラミングに関する専門知識を最小限に抑えながら、AI モデルのテストと導入を容易にしています。AI をよりユーザーフレンドリーにし、場合によっては GUI ベースにすることで、AI の専門知識を強化するために追加のリソースやトレーニングに投資する必要性を減らすことができます。

AI モデルに精通している設計者にとって、PyTorch や TensorFlow のようなオープンソースプラットフォームは、各組み込みプラットフォームの詳細の多くを抽象化することで、エッジ AI ソリューションの開発プロセスを簡素化しやすくなります。これらのツールを使用して開発したモデルを、互換性のある組み込みデバイスにインポートすることができます。

テキサス・インスツルメンツのデバイス上で AI モデルのテストと導入をさらに効率化することを希望する開発者は、Web ベースのツールで構成された**テキサス・インスツルメンツの Edge AI Studio**を利用することができます。Edge AI Studio は、テキサス・インスツルメンツのリモートハードウェアとグラフィカル ユーザー インターフェイスを使用したテキサス・インスツルメンツの組み込みデバイス上で、エッジ AI アプリケーションの開発の簡素化と迅速化を意図して設計されています。このツールには、Model Composer、Model Analyzer、モデル選択ツール、Model Maker が含まれているため、設計者は評価ボードに物理的に接続することなく、モデルとその性能を迅速に評価できます。

エッジ AI のスケーラビリティ

組み込みマイコンまたはマイクロプロセッサを使用して製品を開発する場合、製品が長期的にどのように進化し、スケーリングするかを考慮することは常に重要です。エンジニアは、1 つのマイクロプロセッサを使用したソリューションの開発に数か月を費やし、その後、開発した製品をより高性能のプロセッサに更新するとき最初から開発をやり直さなくてはならない、というようなことはしたくありません。

これらの組み込みデバイスを開発する半導体メーカーは、機能、性能、コストの点でスケーラビリティの高い製品ラインアップを開発するとともに、AI 向けの各種組み込みプロセッサの間でシームレスな移行戦略を保証する必要があります。そうすれば、開発者が複数のデバイス間でデータや成果物を再利用し、開発作業をできるだけシンプルにすることができます。エッジ AI も例外ではありません。たとえば、家庭用ロボットを製作する設計者は、サラウンド ビジョン向けに 3 台のカメラを搭載したハイエンド版と、フロント カメラが 1 台のみのエントリー レベル バージョンの両方を製造したいと考える場合があります。ビジョン プロセッサのスケーラブルな製品ラインアップにより、ハイエンド モデルからエントリー レベルのモデルへのソフトウェアの移植が可能になり、両方の製品の作製に必要なリソースの量を最小限に抑えることができます。スケーラビリティにより、開発者は製品の進化に応じて、研究開発への投資

をあるプラットフォームから別のプラットフォームへと移行することもできます。

まとめ

エッジ AI はまだ比較的新しいものですが、私たちの日常生活を再構築する可能性、特に、ほとんどすべての用途でより高い応答性とパフォーマンスを実現する能力が注目されつつあります。低消費電力でコスト効率の優れた組み込みプロセッサや、直観的なソフトウェアとモデル トレーニング ツールの進歩により、どのような経験レベルの設計者にとっても導入障壁はこれまでになく低くなっています。これが、エッジ AI デバイスや、私たちが日々使っており、信頼している電子機器の中での動作を管理したりデータを収集したりする、重要なコンポーネント (たとえばセンサーや電力供給、コネクティビティ向けの半導体など) の各世代で継続することが期待できます。

その他の資料

- テキサス・インスツルメンツのエッジ AI プロセッシングの製品ラインアップと設計リソースを <https://www.ti.com/technologies/edge-ai.html> でご確認ください。
- テキサス・インスツルメンツのエッジ AI ビジョンプロセッサについては、ホワイトペーパー『[高集積プロセッサを搭載した効率的なエッジ AI システムの設計](#)』をご覧ください。
- スマートホームやリテール エッジ AI のビジョンベース アプリケーション向け設計の主な課題については、技術記事『[ビデオドアベルやスマートリテールの設計でビジョン プロセッサを活用し、エッジ AI 機能を拡張する方法](#)』をご覧ください。

重要なお知らせ:ここに記載されているテキサス・インスツルメンツ社および子会社の製品およびサービスの購入には、TI の販売に関する標準の使用許諾契約への同意が必要です。お客様には、ご注文の前に、TI 製品とサービスに関する完全な最新情報のご入手をお勧め致します。TI は、アプリケーションに対する援助、お客様のアプリケーションまたは製品の設計、ソフトウェアのパフォーマンス、または特許の侵害に対して一切責任を負いません。ここに記載されている他の会社の製品またはサービスに関する情報は、TI による同意、保証、または承認を意図するものではありません。

すべての商標は、それぞれの所有者に帰属します。

重要なお知らせと免責事項

TI は、技術データと信頼性データ（データシートを含みます）、設計リソース（リファレンス・デザインを含みます）、アプリケーションや設計に関する各種アドバイス、Web ツール、安全性情報、その他のリソースを、欠陥が存在する可能性のある「現状のまま」提供しており、商品性および特定目的に対する適合性の黙示保証、第三者の知的財産権の非侵害保証を含むいかなる保証も、明示的または黙示的にかかわらず拒否します。

これらのリソースは、TI 製品を使用する設計の経験を積んだ開発者への提供を意図したものです。(1) お客様のアプリケーションに適した TI 製品の選定、(2) お客様のアプリケーションの設計、検証、試験、(3) お客様のアプリケーションに該当する各種規格や、その他のあらゆる安全性、セキュリティ、規制、または他の要件への確実な適合に関する責任を、お客様のみが単独で負うものとし、

上記の各種リソースは、予告なく変更される可能性があります。これらのリソースは、リソースで説明されている TI 製品を使用するアプリケーションの開発の目的でのみ、TI はその使用をお客様に許諾します。これらのリソースに関して、他の目的で複製することや掲載することは禁止されています。TI や第三者の知的財産権のライセンスが付与されている訳ではありません。お客様は、これらのリソースを自身で使用した結果発生するあらゆる申し立て、損害、費用、損失、責任について、TI およびその代理人を完全に補償するものとし、TI は一切の責任を拒否します。

TI の製品は、[TI の販売条件](#)、または [ti.com](#) やかかる TI 製品の関連資料などのいずれかを通じて提供する適用可能な条項の下で提供されています。TI がこれらのリソースを提供することは、適用される TI の保証または他の保証の放棄の拡大や変更を意味するものではありません。

お客様がいかなる追加条項または代替条項を提案した場合でも、TI はそれらに異議を唱え、拒否します。

郵送先住所：Texas Instruments, Post Office Box 655303, Dallas, Texas 75265
Copyright © 2024, Texas Instruments Incorporated