AI テクノロジーの導入に関する 4 つのキーポイント

人工知能 (AI) テクノロジーを使い始めたい、あるいはその活用範囲を広げたいと考えているビジネスリーダーたちには、共通の疑問や懸念事項があります。たとえば、安定性を維持しリスクを抑えつつ、運用効率を高めてビジネスを成長させるにはどのように AI を使用すればよいか、などです。ここでは、AI 導入のプロセスの進度にかかわらず、考慮するべき 4 つのポイントをご紹介します。

1　柔軟なプラットフォームを選ぶ

AI の導入に取り掛かったばかりであっても、ユースケースを拡張してより広く AI を活用していこうとしている段階であっても、AI デプロイメントの成熟をサポートできるスケーラブルで柔軟なプラットフォームを選ぶことが重要です。

AI 導入の初期段階にある場合は、チームが小さなワークロードを構築して実行できる、安定して柔軟、かつオープンなインフラストラクチャが必要です。社内で AI の使用が拡大し、ワークロードが増えるのにあわせて大規模にモデルをトレーニングできるプラットフォームを使いましょう。

プラットフォームが安定していれば、データサイエンティストや開発者は顧客に真の価値を提供するモデルを構築でき、顧客は使用しているプラットフォームが信頼でき、セキュリティを重視しているという確信を深めることができます。

すでに AI の機能を使いこなしている組織の場合は、AI の可能性をフルに引き出せるようサポートするテクノロジーが必要です。AI を使用する際には、場合によっては GPU やその他のハードウェア・アクセラレータに対する大きな投資が必要になります。

スケーラブルなオープン・プラットフォームは次のことができます。

- AI で駆動するアプリケーションをより短時間で構築およびモダナイズするための優れた基盤となる
- AI で駆動するアプリケーションをより迅速にデプロイする
- 急速に変化し続ける AI 環境において競争力を獲得するのに役立つ

適切なインフラストラクチャがあれば、将来どのような AI イノベーションが出現しても受け入れることができます。

2　ハイブリッド AI に備える

生成 AI の登場により、コンピューティングリソースの要求は増大し、処理はネットワークのエッジへと移動しました。生成 AI のユースケースでは、クラウドリソースとオンプレミスリソースを組み合わせることでメリットを得られます。ただし、一部のワークロード (特にコンプライアンスやデータグラビティに関する課題が大きい場合) はオンプレミスのみで実行する必要があります。ハイブリッド AI では、データが生成される場所の近くを含め、最も適切な場所でワークロードを実行することができます。

ハイブリッド AI には次のようなメリットがあります。

- **柔軟性：**実験的なワークロードやプロダクション・ワークロードをそれぞれ最も適切な環境で実行することで、コストを削減できます。
- **提供のスピードアップ：**小規模ワークロードの推論をエッジで行うことで、ユーザーに素早く情報を提供できます。
- **リスクの低減：**AI アプリケーションをエッジで実行することで、大量の情報をコアに転送する必要が少なくなり、データの保護が強化されます。

ハイブリッド AI を使用すると、IT チームはコアデータセンターやパブリッククラウドでモデルをトレーニングしてエンドポイントに配信し、有益なインテリジェンスをより迅速に提供させることができます。大きな処理能力を必要とするワークロードは中央クラウドとエッジの全体で分散処理し、処理能力の必要性と正確な提案の迅速な提供を両立させることができます。

3　最初のモデルの先を考える

IT チームは 1 つの AI モデルを最小の努力でプロダクション稼働させることができるかもしれません。しかし、成功はスケーリングの能力に依存することが少なくありません。ビジネス価値を最大限に高めるには、そのモデルの複数のバージョンを、場合によっては 1 日に数百回デプロイできなければなりません。

MLOps プロセスを導入すれば、デプロイを簡単にスケーリングするのに役立ちます。MLOps は DevOps とよく似たプロセスで、開発者、運用マネージャー、データサイエンティストが協力し合って AI ベースのアプリケーションの開発を加速させ、デプロイを成功させることを可能にします。MLOps を使用すると、アプリケーション開発の拡張がより簡単になるので、より短時間でより多くのサービスを顧客に届けることができます。

MLOps はコスト効率の高い効率的な開発方法になり得ます。複数チームが共通のプラットフォームで共同作業することで効率性を高め、時間やコストの節約につなげることができます。これは、組織の純利益に直接影響します。

4　セキュリティを重視する

サイバー攻撃は受けた場合の被害額が大きく、ビジネスエグゼクティブは機密データを保護しリスクを緩和するための方法をより強く求めるようになっています。生成 AI を使用すると機密データがさらされる可能性が高まる場合があるため、その緊急性はさらに高まります。

以下のような脅威を回避するためには、強力なセキュリティポスチャが極めて重要です。

- **データポイズニング：**誤った、あるいは悪意のあるデータを AI モデルに注入し、不正確な情報や詐欺的な提案が提供されるようにします。
- **モデル窃取：**機械学習モデルをリバースエンジニアリングし、コピーを使用してデータを抽出します。
- **バックドア攻撃：**AI モデルの中に悪意のあるコードを潜ませ、攻撃者が情報を抜き取れるようにします。

これらをはじめとするさまざまな戦術は組織を経済的および法的なリスクにさらし、ビジネスに対する大きな脅威となる可能性があります。AI に関連する脆弱性を最小限に抑えるには、セキュリティパッチやリリースの処理に関して追跡できる実績を持ち、きちんとしたプロセスを確立している信頼できるアプリケーション・プラットフォーム・プロバイダーからの支援を活用するとよいでしょう。

さらに詳しく

AI 戦略を加速しましょう。Red Hat の AI ソリューションをご覧ください。

Red Hat について

Red Hat は、受賞歴のあるサポート、トレーニング、コンサルティングサービスをお客様に提供し、複数の環境にわたる標準化、クラウドネイティブ・アプリケーションの開発、複雑な環境の統合、自動化、セキュリティ保護、運用管理を支援します。

アジア太平洋
+65 6490 4200
apac@redhat.com

オーストラリア
1 800 733 428

インド
+91 22 3987 8888

インドネシア
001 803 440 224

日本
03 4590 7472

韓国
080 708 0880

マレーシア
1 800 812 678

ニュージーランド
0800 450 503

シンガポール
800 448 1430

中国
800 810 2100

香港
800 901 222

台湾
0800 666 052

fb.com/RedHatJapan
twitter.com/RedHatJapan
linkedin.com/company/red-hat

jp.redhat.com