Microsoft Azure

Machine Learning solutions with enterprise security and scale

Empowering organizations to build secure, scalable, and equitable ML solutions with Azure Machine Learning

Contents

AI is not the future—it's now. Across industries, leading businesses and organizations are leveraging AI to tackle some of the biggest challenges facing the world. Among top enterprises, 50% report that their companies have adopted AI in at least one business function,[1] and 75% plan on continuing to invest in new AI initiatives over the next six to nine months.[2]

But investing in AI doesn't automatically translate to long-term business success. As organizations progress from small proofs of concept and isolated AI projects to large-scale initiatives, ensuring effective adoption means building solutions that function in real-life scenarios. For enterprise organizations, this means building solutions that can scale to thousands of users without sacrificing cost-efficiency, security, or compliance.

These priorities aren't new for enterprise organizations, but applying them to AI can be especially difficult. At the heart of many AI systems are complex Machine Learning (ML) models that provide the predictive intelligence needed to make smart, automated decisions. Building these models, ensuring their accuracy, and maintaining them means dedicating cross-department teams to the ML lifecycle: training ML models through large datasets, then packaging, validating, deploying, and monitoring them.

"Our goal is to empower organizations to apply AI across the spectrum of their business to engage customers, empower employees, optimize operations, and transform products."

Eric Boyd

Corporate Vice President, Azure AI

[1] https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020

[2] https://blogs.gartner.com/svetlana-sicular/gartners-update-on-ai-investments-in-the-enterprise/

Running the ML lifecycle at scale presents several challenges to enterprise organizations. For one, the vast datasets involved in ML make it difficult to manage costs—especially when models must be continually trained and retrained to maximize accuracy. This is an even bigger problem when stakeholders from a variety of teams don't have clear roles and responsibilities. In such an environment, it can become even harder to meet industry security and regulatory requirements, let alone organizational fairness and ethical standards.

Building and deploying enterprise-ready AI solutions means adopting a strategic approach to ML—one that streamlines execution at scale and helps ensure security, compliance, and AI responsibility. At Microsoft, we're committed to helping reshape software development and enabling next-generation AI capabilities. Through services like Azure Machine Learning, we're helping organizations build enterprise-ready solutions with the tools to maximize efficiency, security, and compliance. In this whitepaper, we will explore the best practices enabled by Azure Machine Learning to help you maximize security and compliance, execute your project at scale, and build responsible, equitable systems.

Put security and compliance at the heart of your ML project

Security and compliance are imperatives for any software development project, but the unique nature of ML makes it especially challenging to meet organizational security standards and industry requirements. Building and operating effective ML models is an intensive process requiring multiple teams, a variety of systems, and complex data infrastructure. With so many teams and systems involved, it can be hard to manage identities and control costs, especially when large-scale data infrastructure complicates responsible data management and monitoring against threats and vulnerabilities.

Since security is one of the biggest reasons ML projects get stuck, it's important to think about your approach to security from the beginning rather than trying to patch it in at the end. A successful security approach means authenticating and authorizing users, securing your network, controlling your data, and enacting strong data governance. So where does one start?

Manage access and identities

Building effective ML models requires input from a variety of teams—but providing too much access increases risk and makes it harder to stay in compliance with industry standards and regulations. Authenticating and authorizing users enables enterprises to limit access to sensitive data and workloads without hampering productivity.

Authentication

Azure Machine Learning helps keep sensitive data and workloads safe by streamlining the authentication of users accessing sensitive systems and automated processes taking part on them. Azure Active Directory (Azure AD) is the primary tool that supports user authentication in Azure Machine Learning and offers three primary authentication workflows:

- **Interactive authentication** enables you to control access to resources like web services on a per-user bases. With interactive authentication, you use your account in Azure Active Directory to either directly authenticate users or provide a token that can be used for authentication.

- **Service principal authentication** supports automated processes that don't involve users, such as authenticating a CI/CD script to train and test a model every time the training code changes. With this form of authentication, you create a service principal account in Azure AD and use it to authenticate the process or get a token.

- **Managed identity authentication** provides an additional layer on top of a service principal, automating their creation and management. Managed identities are becoming the preferred approach to identity management and can be used when running the Azure Machine Learning SDK on an Azure Virtual Machine. This workflow enables the VM to connect to the workspace using the managed identity without storing credentials in Python code or prompting user authentication.

To streamline authentication, you can also configure Azure Machine Learning compute clusters to use a managed identity to access the workspace when training models. This enables seamless communication with services like Azure Container Registry (ACR) via managed keys instead of requiring admin keys for ACR, which can be hard for many enterprises.

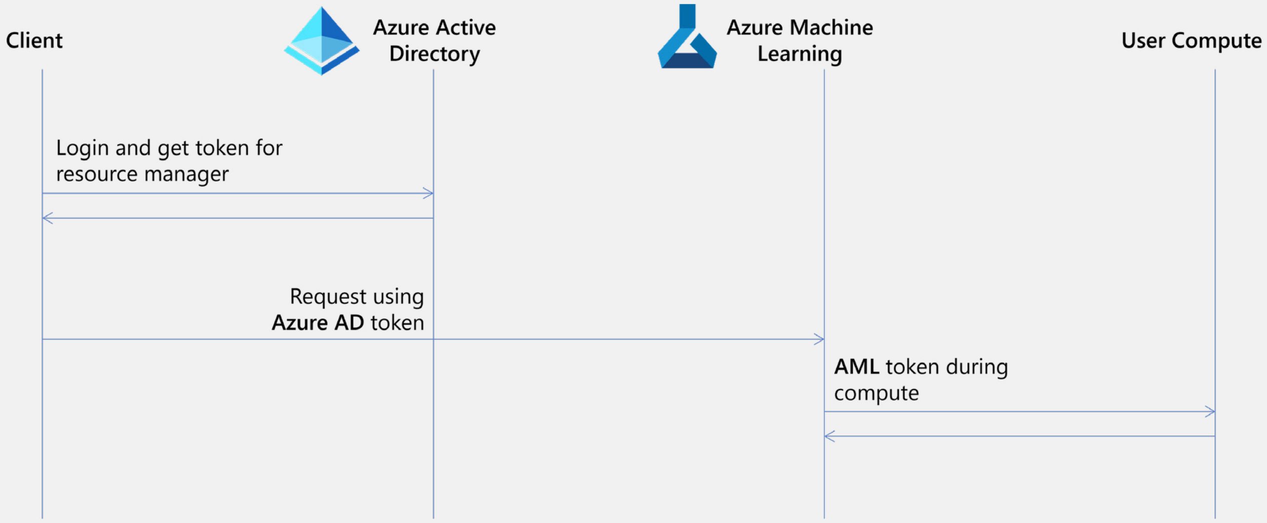
Authorization

Regardless of the authentication type used, Role-Based Access Control (RBAC) is critical to managing the amount of access each user has to mission-critical resources. For example, while solution architects may need granular permissions to all workload resources and assets, other users like business analysts likely only need read access to see insights derived by the ML workload.

Managing access based on identities enables fine-grained control over permissions, and custom roles support further customization to meet your organization's unique needs. With permissions managed by role, you can configure your workspaces to ensure all users have the right permissions to access mission-critical resources without introducing unnecessary risk. Let's take a closer look at how Azure Machine Learning supports RBAC:

- **Built-in roles in Azure Machine Learning:** Azure Machine Learning comes with three built-in roles: owners, contributors, and readers, each with different permission. When creating a new workspace, you can set up role-based access to share workspaces with everyone who needs access while limiting the actions they can take with your data and resources. Azure Machine Learning also includes a new data scientist role that expands on the admin role with specialized access for data scientists. These roles also apply to associated compute targets and data stores, ensuring your resources can only be accessed by authorized users.
- **Custom roles in Azure Machine Learning:** If the built-in roles don't meet your needs, you can create custom roles. These can help you define roles that are best suited for your team structures and organizational policies. You can build custom roles to control operations like creating a compute cluster, submitting a run, registering a datastore, or deploying a model. Custom roles can have read, write, or delete permissions on the various resources of a workspace, such as clusters, datastores, models, and endpoints. You can make the role available at a specific workspace level, a specific resource-group level, or a particular subscription level.

Authorizing user access to compute resources with Azure Active Directory



Isolate networks and monitor against threats

Limiting access to the users who need it is a key first step to securing ML workspaces, but these steps can be rendered ineffective if the network itself isn't secure. ML models require complex infrastructures connected to a variety of services like storage, container registry, and Kubernetes services, and maintaining this infrastructure while protecting sensitive data can put organizations between a rock and a hard place. Thankfully, there are ways to minimize the risks of security threats and vulnerabilities.

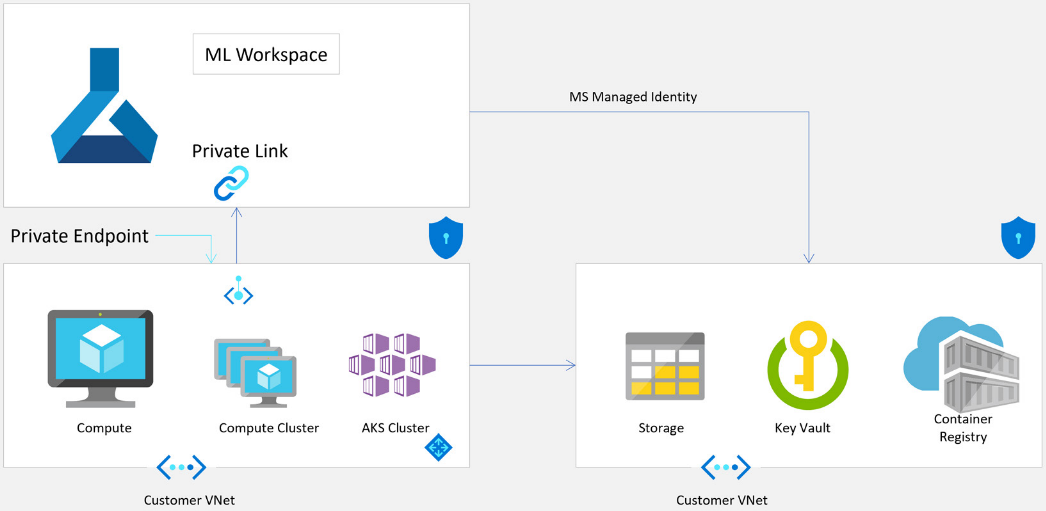
- **Isolate ML resources on a virtual network:** Azure Machine Learning mitigates the risks of security vulnerabilities and threats by giving you the option to move your Azure resources onto an Azure Virtual Network (VNet). This isolates your data from the public internet, making it harder for bad actors to access sensitive data while enabling you to connect out to other mission-critical Azure services.
- **Limit traffic to your VNet with Azure Private Link:** To further protect your ML resources, Azure Private Link enables you to connect to your workspace using a private endpoint—a set of private IP addresses within your virtual network. With Private Link, you can limit access to your workspace to only occur over the private IP addresses, essentially masking your VNet from public traffic. This provides another level of protection against data exfiltration.

Between VNets and Azure Private Link, Azure Machine Learning provides you with the capabilities to ensure end-to-end private IP deployment for your ML environments. Best of all, you don't have to worry about setting these up on your own—Azure ML has a dedicated ARM template within Azure Quickstart templates to automatically deploy secure configurations.

Control your data

Effective data management is key not only to protecting against security threats, but also ensuring compliance with industry regulations. ML models require as much data as possible to

Using VNets and Private Link to isolate ML workspace from public network



maximize model accuracy, so organizations must be prepared to manage the privacy, encryption, and confidentiality of large volumes of data coming from a variety of sources—whether it's at rest or in transit.

Encrypting sensitive data at rest

Azure Machine Learning includes a special tag for workspaces that hold High Business Impact (HBI) information, such as personal customer data. If your workspace contains sensitive data, you can tag it with the hbi_workspace flag. This flag controls the amount of data Microsoft collects for diagnostic purposes and enables additional encryption in Microsoft-managed environments. By tagging your workspace with the hbi_workspace flag, your workspace automatically:

1. Encrypts the local disk in your Azure Machine Learning compute cluster and cleans up your local disk between runs.
2. Securely passes credentials for your storage account, container registry, and SSH account from the execution layer to your compute clusters using your key vault.
3. Enables IP filtering to ensure the underlying batch pools cannot be called by any external services other than AzureMachineLearningService.

Encrypting metadata at rest

It is important to note that Azure Machine Learning does store some metadata, although this metadata does not include sensitive information like trained models or results—that data is strictly stored and controlled by you. Instead, metadata includes operational information such as metrics and telemetry from training and is stored in an Azure Cosmos DB instance associated with a Microsoft subscription managed by Azure Machine Learning.

By default, the Azure Cosmos DB is encrypted at rest with Microsoft-managed keys. If you plan on storing your data—such as run history information—outside of the multi-tenant Cosmos DB instance hosted within the Microsoft-owned subscription, you might consider reconfiguring the Azure Cosmos DB to be managed by your own Customer Managed Keys (CMK) instead. To do so, create a dedicated Cosmos DB instance for use with your workspace.

Encrypting data in transit

When you're ready to use your data, Azure Machine Learning uses TLS 1.2 protocols to secure internal communication between various Azure Machine Learning microservices. All Azure Storage access also occurs over a secure channel, and all external calls made to the scoring endpoint happen over TLS as well.

Policies and monitoring

As important as identity, network, and data management are to security, effective policy and monitoring are crucial to ensuring your security initiatives are working as intended. Maintaining compliance with regulatory standards requires clear policies, and ensuring all stakeholders have the insights they need means providing extensive observability across complex computing environments.

Let's look at how Azure Machine Learning makes it easy to establish comprehensive, compliant policies and provide all your stakeholders with the insights they need—all from a single pane of glass:

- **Streamline security policy with Azure Policy and Azure Security Center:** Azure Machine Learning integrates with Azure Policy and the Azure Security Center, which provide Microsoft created and managed definitions related to different compliance standards. There are several policy options with Network Security and Data Protection that can be applied via Azure Policy. For example, an organization can ensure that there are no ML workspaces created without a Private Link configured.
- **Enable cross-environment observability with Azure Monitor:** Azure Monitor collects and aggregates data from a variety of sources into a common data platform where it can be used for analysis, visualization, and alerting. It provides a consistent experience on top of data from multiple sources, which gives you deep insights across all your monitored resources and even with data from other services that store their data in Azure Monitor. It can provide you the ability to capture both Metrics and workspace activity logs, storing data for up to 90 days.

Execute efficient ML at an enterprise scale

Security may be an imperative, but it’s not the reason organizations starts an ML project. Enterprise organizations want to see real business value from their ML investment. But implementing a ML project and seeing tangible benefits from it are two different things—considering the complexity of ML projects, it’s easy for costs and timelines to bloat if your organization isn’t careful.

This is especially true considering that the sheer

Policies and governance

With Azure Policy in Azure ML, you can manage many policy and governance features, including:

Customer-managed keys

Audit or enforce whether the workspace must use a customer-managed key.

Private Link

Audit or enforce whether workspaces use a private endpoint to communicate with a virtual network.

Private Endpoint

Configure the Azure Virtual Network subnet where the private endpoint should be created.

Private DNS zone

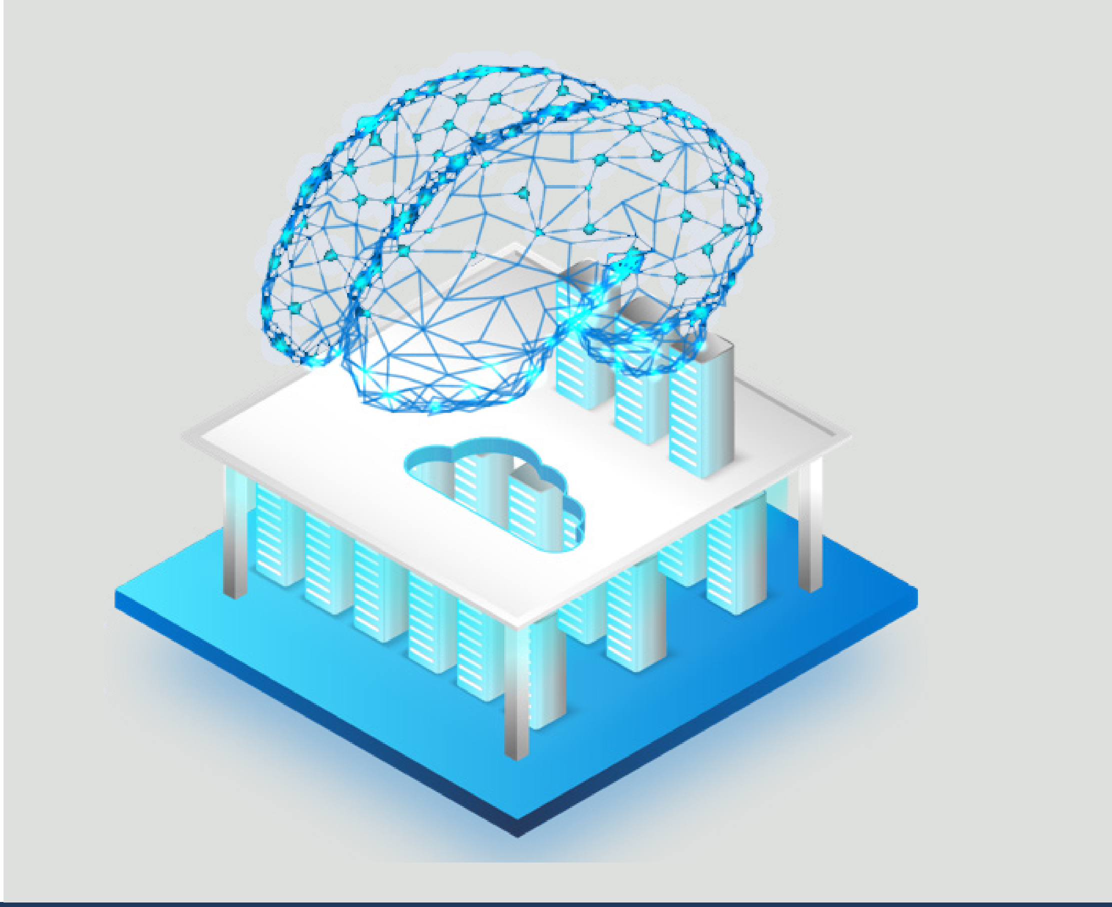
Configure the private DNS zone to use for the private link.

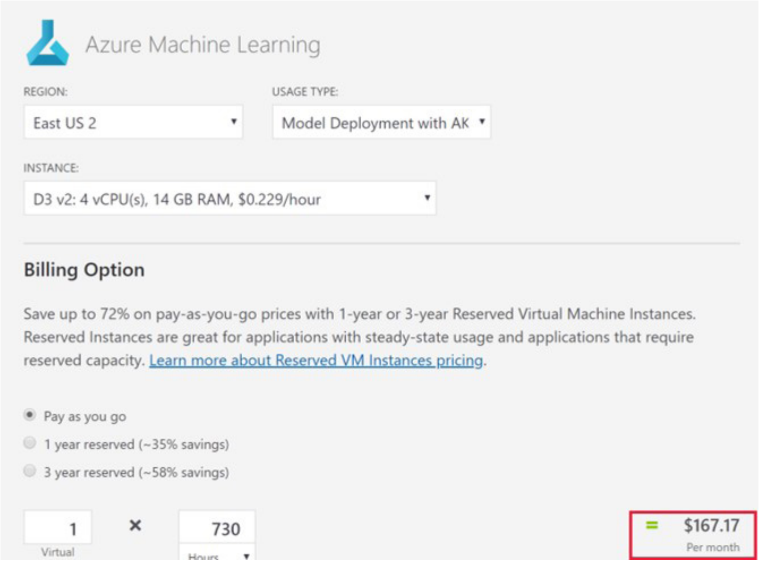
amount of data available to enterprise organizations. ML projects normally require large volumes of data, but enterprise-scale ML necessitates even more compute power—up to thousands of nodes and hundreds of high-powered GPUs. Running so much compute requires cutting-edge infrastructure to train and deploy effective ML models, which means leveraging the latest CPU and GPU compute power running in multi-node clusters to crunch through large volumes of data.

To avoid incurring large costs due to high compute consumption, enterprise organizations need ways to optimize their ML compute usage—by managing and controlling their costs, implementing hybrid and multi-cloud environments, leveraging flexible, open-source resources, and automating the ML lifecycle. Let's look at how Azure Machine Learning helps organizations scale and optimize their ML resources.

Manage and control your costs

One of the biggest concerns for enterprise organizations building a new ML model is uncertainties around costs. No matter how well thought-out your AI initiative is, building the right data model takes time. Increased time often means added cost, especially if there are no controls in place





Azure Machine Learning

REGION: East US 2

USAGE TYPE: Model Deployment with AK

INSTANCE: D3 v2: 4 vCPU(s), 14 GB RAM, $0.229/hour

Billing Option

Save up to 72% on pay-as-you-go prices with 1-year or 3-year Reserved Virtual Machine Instances. Reserved Instances are great for applications with steady-state usage and applications that require reserved capacity. Learn more about Reserved VM Instances pricing.

● Pay as you go
○ 1 year reserved (~35% savings)
○ 3 year reserved (~58% savings)

1 Virtual × 730 Hours = $167.17 Per month

Customers can use the Azure pricing calculator to estimate the compute costs of their ML project and set realistic expectations before getting started.

to limit compute consumption on individual workspaces.

To realize real value from your ML projects, it's important to keep costs under control and streamline the ML lifecycle—maximizing time and resource utilization wherever possible to see faster time to value. Azure Machine Learning includes a multitude of fail-safes to protect against bloated costs and maximize the efficiency of your ML investment. Some best practices include:

- **Estimate costs before getting started with Azure Machine Learning:** The Azure pricing calculator is a helpful tool for estimating costs before you begin your ML initiative. Through the Azure pricing calculator, you can see the likely costs of individual Azure Machine Learning instances based on your usage type and billing option. This helps you set expectations and ensure you have the resources necessary to successfully complete your ML initiative.
- **Implement cost budgets and controls:** Once you've validated that you have the resources needed to realize your ML initiative, it's important to keep tabs on your compute utilization to ensure projects stay on budget. You can create budgets to manage costs and alerts that automatically notify stakeholders of spending anomalies and overspending risks. Alerts are based on spending compared to budget and cost thresholds. Budgets and alerts are created for Azure subscriptions and resource groups, so they're useful as part of an overall cost monitoring strategy.
- **Reduce start-up time:** Azure Machine Learning helps you jumpstart your ML experience to see immediate productivity. Compute instances make it easy to get started with Azure Machine Learning development and provide management and enterprise readiness capabilities for IT administrators. Setup scripts enable you to automatically customize and configure the compute instance at provisioning time, and Azure Resource Manager templates streamline the creation of resources as a single coordinated operation.
- **Plan, manage, and share resources across**

projects: As an organization grows its number of machine learning use cases and teams, it requires increased operating maturity from IT and Finance, as well as coordination between individual machine learning teams to ensure efficient operations. Company-scale capacity and quota management become important to address scarceness of compute resources and overcome management overhead. Azure allows you to set limits for quota allocation at both a subscription and workspace level. Restricting who can manage quota through Azure Role-Based Access Control (RBAC) can help ensure predictable resource utilization and costs.

- **Manage costs with autoscaling compute:** Batch endpoints enable you to run asynchronous batch inference jobs to increase the efficiency your compute efficiency. Compute resources are automatically provisioned when the job starts and automatically de-allocated as the job completes. You only pay for the compute you use, and you can override compute resource settings (like instance count) and manage advanced settings (like mini-batch size, error threshold, and so on) for each individual batch inference job to speed up execution while keeping costs low.

Achieving ML at scale with Azure Machine Learning

agl

"We're setting up all the infrastructure we need to work efficiently at scale, reuse code, and have a library of things that we can build out, rinse, and repeat. With Azure Machine Learning, we're increasing speed-to-value while reducing cost-to-value."

Sarah Dods

Head of Advanced Analytics, agl

Hybrid and multi-cloud ML

Another way of managing large-scale ML deployments while controlling costs is to implement a hybrid or multi-cloud environment. There are a multitude of reasons why enterprises would want to leverage this infrastructure: some may still be on their journey to the cloud and don't have all their data moved to the cloud, while others may need to keep sensitive data on-premises or on a sovereign or private cloud for regulatory purposes. Others may want to leverage existing third-party or on-prem data or compute infrastructure, and even more may want to implement a "cloud burst" scenario, in which they start with on-prem compute resources and scale into the cloud as compute needs grow.

With so many use cases tied to them, it should come as no surprise that hybrid and multi-cloud deployments are becoming more and more common. Still, deploying across multiple environments comes with its own challenges and considerations, requiring flexible, streamlined deployment and management options.

Deploy a hybrid or multi-cloud environment with Kubernetes

When deploying a hybrid environment, there are several key choices to consider when building the solution architecture—particularly around compute options and deployment topologies. Whatever compute and deployment topologies you choose, they must be available across all your target environments.

Kubernetes is an ideal compute option for hybrid and multi-cloud deployments since it's available for deployment on-premises or as a service on all public clouds. Kubernetes clusters support both CPU and GPU compute options, enabling ML workloads to utilize the right kind of compute. Not only is Kubernetes available across deployment environments, it also enables you to leverage a well-understood, widely adopted deployment topology with containers.

Azure Kubernetes Service is ideal for high-scale production deployments, providing fast response times, autoscaling, logging, model data collection, authentication, and hardware acceleration options such as GPU and Field-Programmable Gate Arrays (FPGA).

Simplify management across hybrid and multi-cloud environments

Managing the deployment and lifecycle of your hybrid or multi-cloud infrastructure means enabling seamless interaction across on-premises and public cloud infrastructures—whether Azure, Amazon Web Services, or Google Cloud Platform. Seamless infrastructure management enables you to monitor the health of the deployment and assist with automation, tracking, security, and more, all from a centralized management console.

Azure Machine Learning makes this possible with Azure Arc, which helps simplify the management and deployment of your hybrid or multi-cloud projects. Azure Arc standardizes visibility, operations, and compliance across a wide range of resources and locations by extending the Azure control plane. Since it can ship and manage solutions from Azure to any Kubernetes distributions in any location, it enables you to securely build and train machine learning models anywhere. Azure Arc is useful for both administrators managing the ML infrastructure and data scientists utilizing resources for distributed model training:

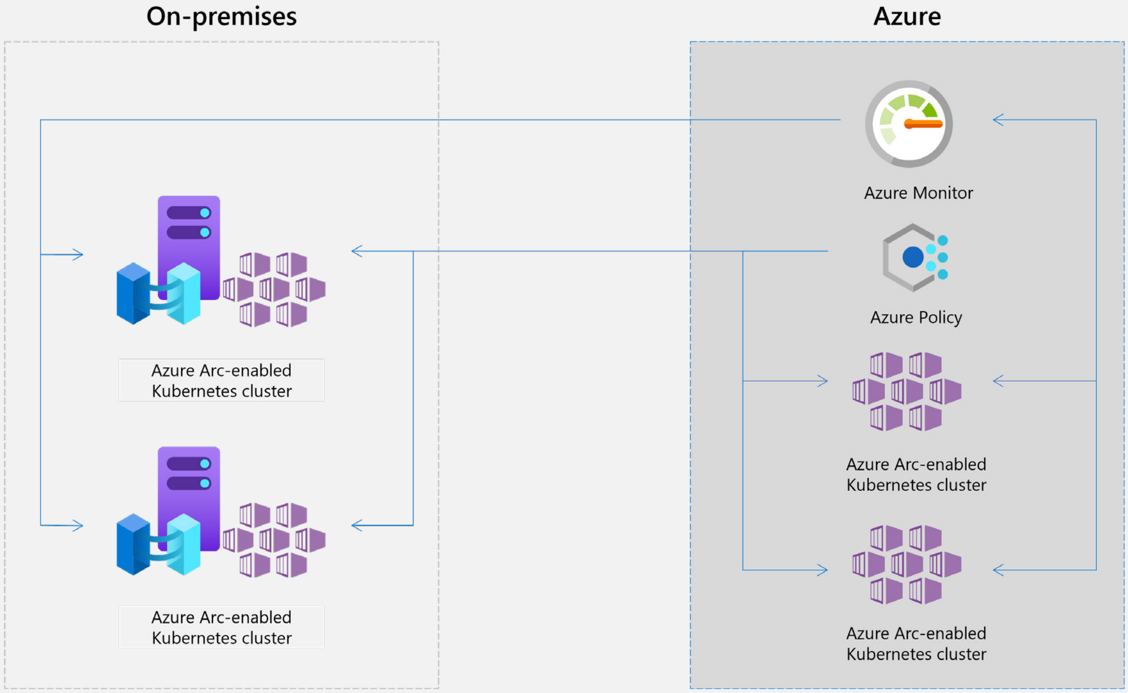
- **Azure Arc for administrators:** Administrators can ensure a seamless setup of the agent on any OSS Kubernetes cluster such as Azure Kubernetes Services, RedHat OpenShift, and more. This makes it easy to set up clusters and run cloud-native tools like GitOps. Once the agent is successfully deployed, administrators can either grant data scientists access to the entire cluster or just a slice. Administrators also get a transparent, flexible view to oversee the cluster's lifecycle and security, manage autoscaling, and upgrade to newer Kubernetes versions.
- **Azure Arc for data scientists:** Azure Arc ensures that data scientists can focus on ML processes rather than managing the resources they runs on. Data scientists don't need to know anything about Kubernetes to use it as a compute target—Azure Arc simply marks the Kubernetes cluster as an "Attached Compute" option that data scientists can select when choosing their training compute target. This enables data scientists to leverage familiar tools like Jupyter notebooks, TensorFlow, and PyTorch with Kubernetes clusters.

Open and interoperable ML

Another way enterprise organizations can get the most from their ML investment is to leverage open, interoperable tools and systems. ML has already matured in large part thanks to investments from the open-source community, and as the technology continues to evolve, this influence is unlikely to go away. It's important for enterprise organizations investing in ML to choose services and solutions that are open and interoperable or risk suffering vendor lock-ins and missing out on integrations that can help maximize efficiency.

Azure Machine Learning is built around a commitment to the open-source community, leveraging integrations with powerful tools, languages, libraries, and frameworks, while promoting the promise of openness and interoperability.

Supporting hybrid ML deployments at scale with Azure Kubernetes Services and Azure Arc


On-premises
Azure
Azure Monitor
Azure Policy
Azure Arc-enabled Kubernetes cluster
Azure Arc-enabled Kubernetes cluster
Azure Arc-enabled Kubernetes cluster
Azure Arc-enabled Kubernetes cluster

Choice of tools and languages

Data scientists and developers often have strong preferences for different developer tools, whether that's interactive notebooks like Jupyter, or feature-rich IDEs like Visual Studio Code. These users are equally opinionated about their preferred language for ML—R or Python.

Azure Machine Learning empowers data scientists and developers to use the tools of their choice by seamlessly integrating with Visual Studio Code, Jupyter, and RStudio:

- **Visual Studio Code integrations:** Data scientists and developers can securely utilize compute instances as execution environments from Visual Studio Code or open notebooks in Visual Studio Code from the Azure Machine Learning Studio. Additionally, native integration with GitHub and Azure DevOps enable more collaborative workflows.
- **JupyterLab and Jupyter Notebooks:** Just like with Visual Studio Code, Azure Machine Learning compute instances come pre-installed with Jupyter and JupyterLab, enabling data scientists to author, train, and deploy models in a fully integrated notebook experience—directly from the ML workspace.
- **RStudio:** Azure Machine Learning resources work with R to provide a scalable environment for training and deploying a model. The RStudio Server Open Source Edition is configured and installed on the Azure Machine Learning compute instance for R users to get started quickly and build models faster.

Interoperability with popular libraries and frameworks

Of course, tools are just one piece of the puzzle. For data science teams to be truly productive, they need the latest data processing and machine learning frameworks and libraries. These comprise a wide-ranging list that includes PyTorch, TensorFlow, scikit-learn, pandas, koalas, and many more.

Not only does Azure Machine Learning come pre-installed with many of the most in-demand frameworks and libraries for use at scale, it also provides two dev/test environment deployment options to maximize data scientist productivity:

- **Azure Machine Learning Compute Instances** provide an integrated workstation for data scientists that comes pre-loaded with PyPI, Conda, deep learning, and ONNX packages.

Languages

Python, R, .Net

Frameworks

Apache Spark, TensorFlow, PyTorch, etc.

OSS Contributions

ONNX, InterpretML, Fairearn, SmartNoise, Error Analysis toolkit

Ecosystem Support

Azure Databricks, MLflow, Apache Spark, Dask, Kubernetes

- **Data Science Virtual Machines (DSVM)** offer a more flexible experience for data scientists with different OS, hardware environment, and compute engine needs. DSVMs come in both Windows and Linux versions, enable login via RDP or SSH, and come with pre-installed with the latest frameworks and libraries. DSVMs even come with developer tools, databases, docker, and compute engines like Apache Spark or Apache Drill.

Contributing back and driving innovation within open source

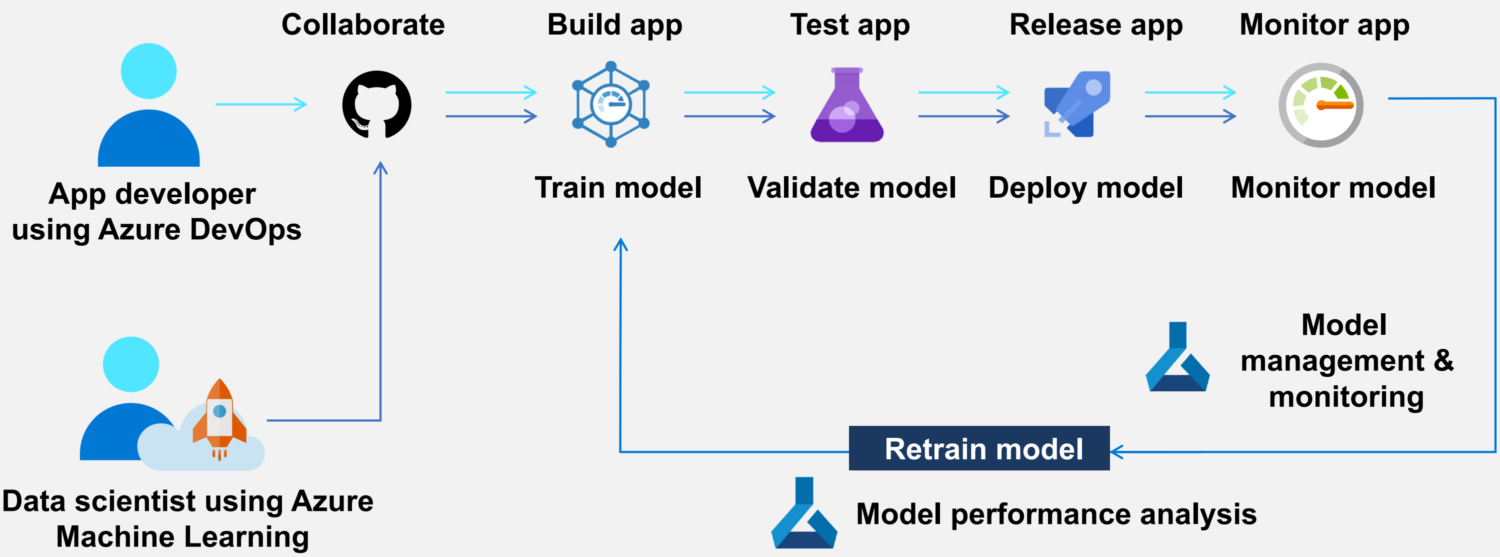
Microsoft is committed to empowering open-source ML and being a good Samaritan within the community. Not only do we collaborate on OSS projects like ONNX, PyTorch, and MLflow, we've also made Responsible ML toolkits like InterpretML, Fairlearn, and Error Analysis available to the broader ML community.

Automate and accelerate the machine learning lifecycle

Creating a machine learning model—taking it from raw data to a deployed model—involves multiple teams and roles working to train and test, package, validate, deploy, monitor, and then retrain the model. Managing the entire lifecycle at scale is complicated and requires large amounts of time and resources. Without repeatable processes, data scientists must reinvent the wheel each time they create and deploy a new model, and siloed teams impede workflow alignment and collaboration—increasing the likelihood of model and code versioning issues that can impact the usefulness of the model.

Just as with traditional software development, organizations can automate and accelerate

MLOps workflow


Collaborate
Build app
Test app
Release app
Monitor app
App developer using Azure DevOps
Train model
Validate model
Deploy model
Monitor model
Model management & monitoring
Retrain model
Model performance analysis
Data scientist using Azure Machine Learning

the machine learning lifecycle using DevOps processes. But although many of the stages are the same, the uniqueness of ML model creation, deployment, and monitoring presents additional complexity and considerations:

- **Code and dataset management:** The source code of the model training process has limited value if it is not also accompanied by the dataset or datasets that were used to create the trained model.
- **Auditability:** It can be difficult to ensure that models meet regulatory standards and performance thresholds over time.
- **Traceability:** It can also be difficult to trace the result of an inference created in a production environment all the way back to the source code and training data sets used to build a model.
- **Explainability:** Black box models make it difficult to understand how the model works.
- **Quality assurance:** Extensive quality checks on both trained models—for interoperability, fairness, and accuracy—and deployed models—for data drift and performance issues—can be exceptionally challenging.

To address these challenges with machine learning, organizations need an approach that brings the agility of DevOps to the ML lifecycle. We call this approach MLOps. To learn more about MLOps best practices with Azure Machine Learning, read the MLOps whitepaper.

Build responsible, equitable ML models

Managing your ML models in a way that maximizes security and efficiency is crucial to ensuring that your ML project doesn't incur more risk than it's worth. But your ML model doesn't just exist in a vacuum—the insights it generates fuel AI systems that are increasingly embedded in people's everyday lives. If your model is improperly designed, trained, or validated, it risks delivering biased insights that can cause real-world problems for your users.

Mitigating risk with ML systems means more than protecting data and privacy—it means enacting a responsible AI approach that promotes positive social impact and preserves human dignity. At Microsoft, we've developed a series of data governance pillars that ladder up to our company-wide AI principles: understand, protect, and control. Each of these principals are supported by tools and methods that can help ensure responsible ML.

Understand

As ML systems become deeply integrated into our daily business processes, we need to be able to trust that they are accurate and fair. Transparency is critical to establishing that trust. Organizations need processes and tools to ensure that data scientists can tune models to meet ethical standards and end-users can understand why certain results are generated. Simply put, when AI systems are used to help make decisions that impact people's lives, people must understand how those decisions were made. Tools like InterpretML, Fairlearn, and Error Analysis help ensure transparency by enabling model interoperability, fairness, and accuracy.

Protect

Privacy and security are key pillars of trust. They require especially close attention in machine learning projects because data is the lifeblood of ML models, and data may contain Personally Identifiable Information (PII). Models are trained using large quantities of sensitive data, then while they're operating they make inferences about people. Organizations need to consider how to protect confidential information while still creating models that are accurate and valuable. They also need to comply with data protection and privacy laws such as Europe's GDPR or the U.S.'s HIPAA. SmartNoise and Microsoft SEAL help safeguard private data by obfuscating an individual's sensitive information through differential privacy and homomorphic encryption.

Control

Enterprise organizations must ensure their ML solutions operate reliably and consistently. Control and accountability are essential for upholding reliability and accuracy. Control starts in the training phase, where ML systems should be exposed to unusual circumstances and rigorously tested to prevent performance failures down the line. After deployment, it's equally important for organizations to properly maintain ML systems throughout their lifespan. Maintenance includes continual monitoring to spot issues and periodic retraining to ensure models remain relevant and accurate. Throughout the entire lifecycle, thorough documentation and audit trails keep people accountable for how their systems operate and help organizations meet regulatory requirements.

As complex as the enterprise ML lifecycle may be, it's important to never lose sight of the end user who will be impacted by your model. By pairing these principles with the features and tools in Azure Machine Learning, your organization can work towards delivering results that deliver the greatest benefit to society while providing transformative business value. Learn more about how your organization can implement responsible AI with the responsible AI whitepaper.

Azure Machine Learning
Responsible AI

Microsoft's holistic approach to responsible AI

1. **Principles:** We've established guiding principles for how AI should be developed, used, and maintained.
2. **Practices:** We've also established a system for internal oversight that provides guardrails for first and third-party AI solutions. Our AI governance teams are tasked with developing policies and guidance, documenting best practices, helping address issues that arise, and educating employees about responsible AI.
3. **Tools:** Our data scientists and developers use tools and resources that make it easier spot and mitigate potential issues when training and maintaining ML models.

Conclusion

Trying to start an AI or ML project can be a daunting task—particularly at the scale of an enterprise organization. Focusing on security, execution, and responsibility is imperative to realizing a successful ML project without ballooning costs or taking on unnecessary risk.

Microsoft is committed to helping enterprise organizations realize their AI and ML goals at scale. Azure Machine Learning is built on our own best practices developed through years of experimentation to help you get started quickly and accelerate your time to market with all of the tools and controls that make it easier to keep costs low and protect against risk. To learn more, visit the Azure Machine Learning page.

Additional resources

eBook

Mastering Azure Machine Learning

Explore this free eBook from Packt for hands-on guidance, real examples, and executable code on Azure ML.

eBook

Principles of Data Science

Get a comprehensive beginner's guide to statistical techniques and theory.

Paper

Four Real-Life Machine Learning Use Cases

Dive into four practical end-to-end machine-learning use cases on Azure Databricks.

Demo

NLP Recipe

See how to leverage recent advances in NLP algorithms, neural architectures, and distributed ML systems.

Sample Code

Recommenders Recipe

Get examples and best practices for building recommendation systems.

Sample Code

Computer Vision Recipe

See examples and best practices for building computer vision systems.

Microsoft

This whitepaper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT. The descriptions of other companies in this document, if any, are provided only as a convenience to you. Microsoft cannot guarantee their accuracy, and the companies and products may change over time. Also, the descriptions are intended as brief highlights to aid understanding, rather than as thorough coverage. This document is provided "as is." Information and views expressed in this document, including URL and other Internet websitereferences, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Microsoft product.