AZURE MACHINE LEARNING

MICROSOFT AZURE

Mahisha Mudaliar

Sr.No.	Topics	Pg. No.
1.	Prerequisite	3
2.	Azure	4
3.	Azure Machine Learning	5
3.1	Readiness and Security	5
3.2	Azure Machine Learning project workflow	9
3.3	Pros and Cons for Azure Machine Learning	11
3.4	Services offered by Azure Machine Learning	14
3.5	How to train a model?	17
3.6	Costing for Azure Machine Learning	18
3.7	Alternatives of Azure Machine Learning	22
3.8	Steps to reduce the costing of Azure Machine Learning	25

1. Prerequisites:

1. **Azure subscription: -** Azure subscription is necessary.
2. **Azure portal access: -** For this, we must have an account in Azure portal where we can use Azure Machine Learning and other resources.
3. **Python: -** Most Machine Learning work in azure is done with python so we need to install python in our local machine.
4. **Azure Machine Learning SDK: -** We need to download Azure Machine Learning SDK for python which allows us to interact with Azure Machine Learning services and resources.
5. **Development Environment: -** We can work with various developing environments like Visual Studio Code, Jupyter Notebook and Azure Machine Learning Studio.
6. **Knowledge of Machine Learning: -** Having a good understanding of machine learning concepts, algorithms, and data preprocessing is crucial for effectively using Azure Machine Learning to build and deploy machine learning models.

2. What is Machine Learning?

- Machine learning is a subset of artificial intelligence (AI) that focuses on the development of algorithms and statistical models that enable computers to improve their performance on a specific task through experience or data, without being explicitly programmed. In other words, it's a method of teaching computers to learn from data and make predictions or decisions based on that learning.

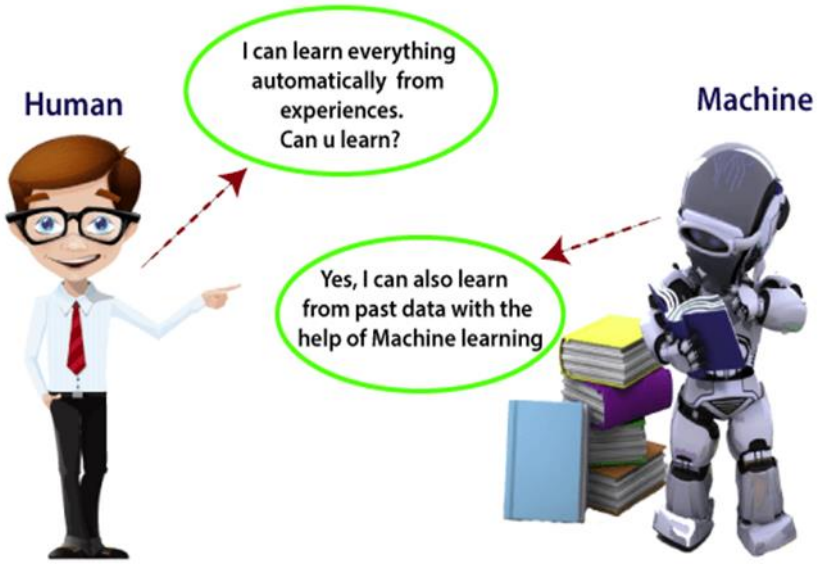


Figure 2.1

3. What is Azure Machine Learning?

- Azure Machine Learning is a cloud-based platform and service provided by Microsoft for building, training, and deploying machine learning models.
- Azure ML gives the full capabilities to design a machine learning solution with Low or No Code, using drag and drop, Designer.
- With the use of Designer and Azure ML, we can add a data source, do data transformation, train the model with different algorithms, as well as do scoring and evaluation without writing any code with publishing options as well.
- The service also interoperates with popular deep learning and reinforcement open-source tools such as PyTorch (ML framework used for apps like computer vison and natural language processing), TensorFlow (library for ML and AI and focus on training and inference (a conclusion reached based on evidence and reasoning) of deep neural networks.), scikit-learn (free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines), and Ray RLlib.
- Individuals and teams deploying MLOPs inside their organization use azure machine learning models into production in a safe and auditable environment.

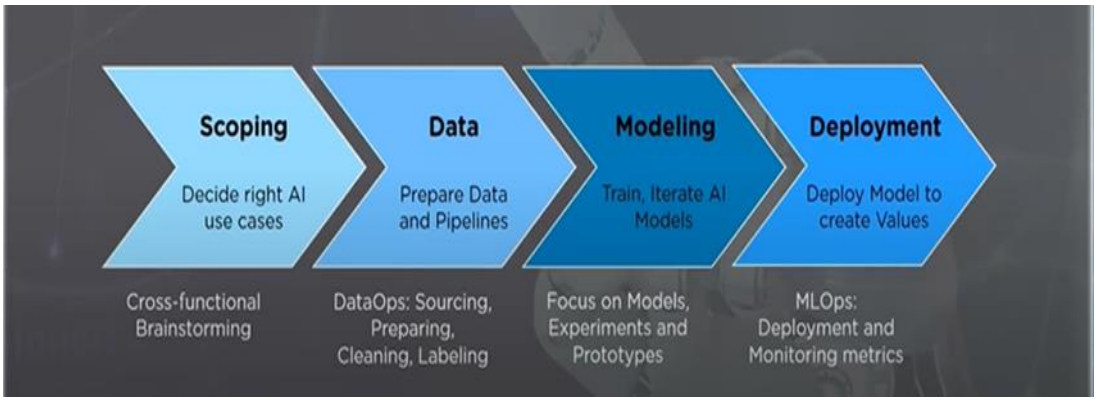
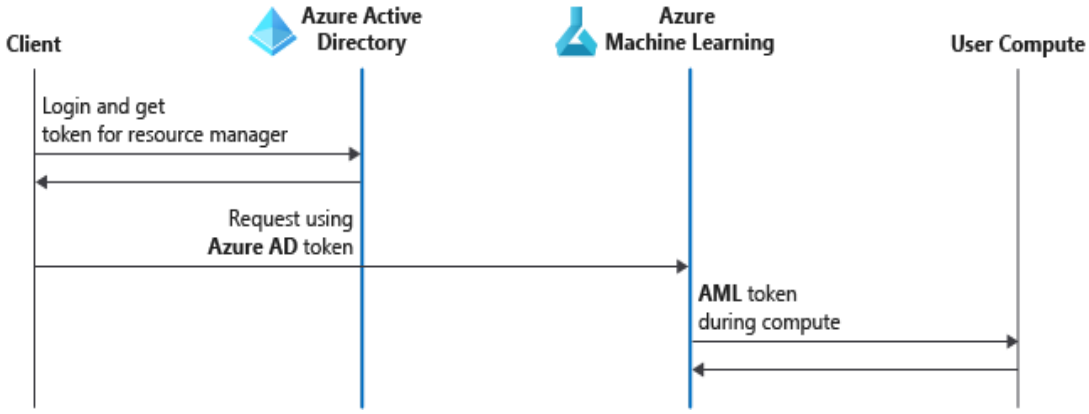


Figure 3.1

3.1 Readiness and Security:

- Azure Machine Learning works with Azure cloud platform to make ML projects more secure.
- This is how Azure Machine Learning Authentication works:



- The following security integrations are available:

1. **Azure Virtual Networks (VNets):**
    - Azure Virtual Networks are basic building blocks in Microsoft Azure providing the foundation for creating an isolated and secure network environment for Azure cloud.
    - Some of the properties of VNets are:
        i. **Isolation:** VNets allow us to isolate our Azure resources from one another, creating a network boundary that prevents unauthorized access.
        ii. **Subnets:** Within a VNet, we can create one or more subnets. Subnets are logical subdivisions of the VNet and allow us to segment resources within the VNet for better organization and network control.
        iii. **DDos Protection:** Azure VNets benefit from Distributed Denial of Service (DDoS) protection by default, helping protect your network from common DDoS attacks.

iv. **User-Defined Routes:** Azure VNets support user-defined routes, which allow us to customize routing to control traffic flow within the network.

v. **Integration with Azure Active Directory (Azure AD):** VNets can be integrated with Azure AD for user and service principal authentication, enhancing security and access control.

2. **Azure Key vault**: Azure Key Vault is a cloud service for securely storing and accessing secrets. A secret is anything that we want to tightly control access to, such as API keys, passwords, certificates, or cryptographic keys.

➢ Some important functionalities of Azure key vaults are:

i. **Secrets Management:** Azure Key Vault allows us to securely store and manage sensitive information, such as API keys, connection strings, and passwords, as secrets. These secrets are stored in a highly secure manner.

ii. **Key Management:** Key Vault can generate and manage cryptographic keys used for encryption and decryption. It supports various key types, including RSA, AES, etc. Key rotation and versioning are also supported.

iii. **Integration with Azure Services:** Key Vault seamlessly integrates with various Azure services, such as Azure Virtual Machines, Azure Functions, Azure Logic Apps, and Azure App Service, making it easy to use keys and secrets securely in your applications.

3. **Azure container registry:** Azure Container Registry (ACR) is a fully managed container registry service provided by Microsoft Azure. It's designed to help developers and organizations securely store, manage, and deploy container images, making it a fundamental part of Azure's container ecosystem.

➢ We can store following things in ACR:

1. **Docker Container Images:** designed to store Docker container images. These images contain the application code, runtime,

libraries, and dependencies required to run an application in a containerized environment.

2. **Linux and Windows Images:** ACR can store both Linux-based and Windows-based container images, enabling us to manage applications for various operating systems.
3. **Custom Images:** We can store custom-built container images that are created from your application code and configuration files. These images can be tailored to your specific requirements.
4. Frameworks and Runtimes: Container images can include specific programming frameworks (e.g., Node.js, .NET Core) and runtimes (e.g., Java, Python) required for your application.
5. Official and Community Images: We can store official Docker images from trusted sources, as well as community-contributed images available on the Docker Hub or other container registries.

And many more.

3.2 Azure Machine Learning project workflow:

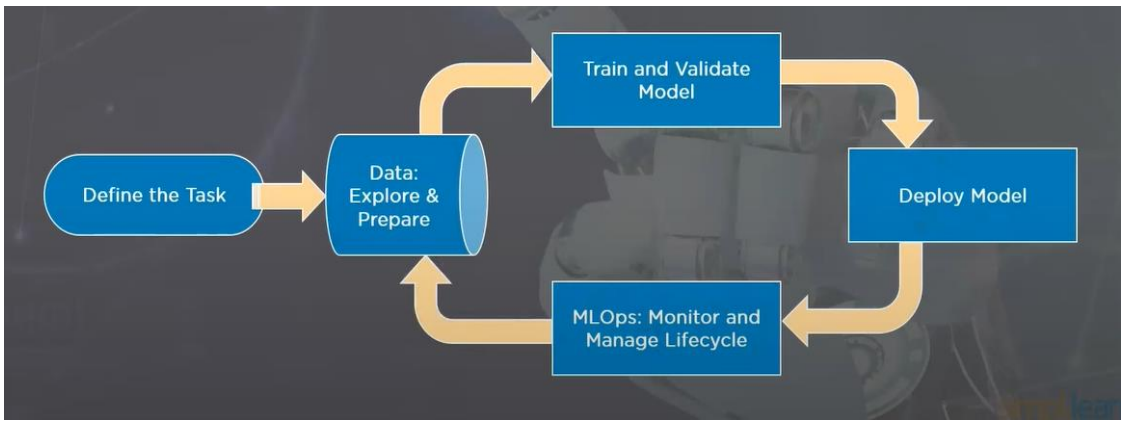


Figure 3.2.1

1. **Define the Problem:**
    - Clearly define the business problem or task we want to solve with machine learning.
    - Specify the goals and objectives of the project.

2. **Data Collection and Preparation:**
    - Collect and gather the necessary data for your project.
    - Clean, preprocess, and transform the data as required.
    - Split the data into training, validation, and test sets.

3. **Create an Azure Machine Learning Workspace:**
    - Set up an Azure Machine Learning workspace, which serves as the centralized hub for your project.

4. **Experimentation:**
    - Create and manage experiments to explore different algorithms and model architectures.
    - Use Jupyter Notebooks, Azure ML Studio, or other tools to develop and test your models.

5. **Data Splitting:**
    - Split the dataset into training, validation, and test sets. The training set is used to train the machine learning model, the validation set is used for model tuning, and the test set is used for final model evaluation.

6. **Model Training:**
    - Train machine learning models using Azure ML's automated machine learning capabilities or custom scripts.
    - Track and log model training runs, including hyperparameters and metrics.

7. **Model Evaluation:**
    - Evaluate model performance using metrics like accuracy, precision, etc.
    - Use tools like Azure ML Designer or Jupyter Notebooks to visualize and analyze results.

8. **Model Optimization:**
    - Optimize model hyperparameters and architecture based on evaluation results.

9. **Model Deployment:**
    - Deploy the trained model as a web service or container to make predictions.
    - Azure ML supports various deployment options, including Azure Kubernetes Service (AKS) and Azure Functions.

10. **Monitoring and Maintenance:**
    - Implement monitoring and logging for your deployed model to track its performance and detect issues. Azure Application Insights can be helpful for this purpose.

11. Feedback Loop:

- Continuously gather feedback from users and stakeholders to improve and refine your machine learning solution.

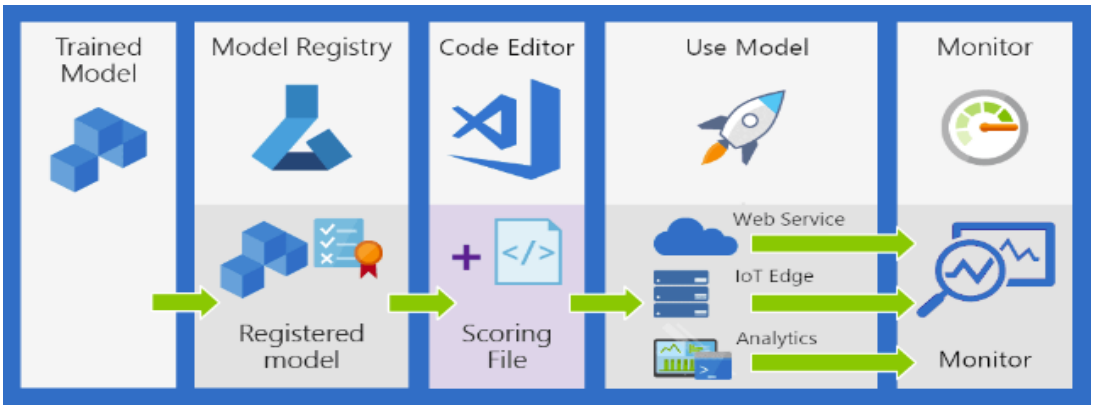


Figure 3.2.2

3.3 Pros and cons for Azure Machine Learning:

- **Pros:**
  - ✓ **Scalability:** Azure ML leverages the scalability of Microsoft Azure, allowing we to scale your machine learning workloads up or down based on demand. This is especially valuable for handling large datasets and complex computations.
  - ✓ **Ease of use:** Azure ML provides a user-friendly interface and a set of tools that make it accessible to data scientists, developers, and analysts with varying levels of expertise in machine learning.
  - ✓ **Integration with Azure Services:** Azure ML seamlessly integrates with other Azure services, including data storage, databases, and Azure Kubernetes Service (AKS), making it easy to create end-to-end machine learning pipelines.
  - ✓ **Automated Machine Learning (AutoML):** Azure ML includes AutoML capabilities that automate many aspects of the machine learning process, making it easier to build high-performing models with minimal manual effort.
  - ✓ **Support for Different Frameworks:** Azure ML supports popular machine learning frameworks like TensorFlow, PyTorch, and scikit-learn, as well as custom code, giving we flexibility in model development.
  - ✓ **Security and Compliance:** Azure ML provides robust security features, including identity and access management, encryption, and compliance certifications which are crucial for handling sensitive data.

- **Cons:**
  - **Cost:** Using Azure ML can become expensive, especially for large-scale projects or when working with extensive computational resources. It's important to monitor and optimize costs.
  - **Vendor Lock-In**: Using Azure ML may lead to vendor lock-in, where your machine learning workloads are tightly integrated with Azure services, making it challenging to migrate to other cloud providers or on-premises environments.
  - **Complexity:** For advanced machine learning tasks and complex workflows, configuring and managing Azure ML resources can become complicated.
  - **Limited Offline Mode:** Some tasks and features in Azure ML may require a continuous internet connection, limiting the ability to work in offline or air-gapped environments.
  - **Dependency on Azure Updates:** Azure ML relies on Azure updates and changes, which can lead to occasional adjustments or modifications required in your machine learning workflows.

3.4 Services offered by Azure Machine Learning:

1. **Azure Machine Learning Studio:**
   Azure Machine Learning Studio is a collaborative web-based environment for data scientists and analysts. It offers a low-code/no-code interface for building machine learning experiments and pipelines.
2. **Azure Machine Learning Workspaces:**
   Workspaces serve as centralized hubs for machine learning projects, allowing us to organize resources, datasets, experiments, and models. Workspaces can be used to collaborate with team members.
3. **Azure Machine Learning Compute:**
   Azure ML Compute provides scalable, cloud-based compute resources for training machine learning models. It supports various VM types and GPU acceleration.
4. **Azure Machine Learning Datasets:**
   Datasets allow us to store, manage, and version your data in Azure Machine Learning. We can create reusable datasets and use them in experiments.
5. **Azure Machine Learning Automated Machine Learning (AutoML):**
   AutoML automates the process of selecting, training, and tuning machine learning models. It helps users build high-quality models with minimal manual effort.
6. **Azure Machine Learning Designer:**
   The Designer is a visual interface for creating machine learning workflows. It allows users to drag and drop modules to build experiments without writing code.

7. **Azure Machine Learning Deployment Services:**
   Deployment Services allow us to deploy machine learning models as web services or endpoints in Azure. We can deploy to Azure Kubernetes Service (AKS), Azure Functions, or other platforms.
8. **Azure Machine Learning Pipelines:**
   Pipelines enable us to automate and orchestrate end-to-end machine learning workflows. They consist of a series of steps, including data preparation, training, and deployment.
9. **Azure Machine Learning CLI and SDKs:**
   Azure ML offers command-line tools (CLI) and software development kits (SDKs) for Python and R to interact with machine learning resources programmatically.
10. **Azure Machine Learning Python SDK:**
    The Python SDK allows developers and data scientists to programmatically create, manage, and interact with Azure Machine Learning resources and services.
11. **Azure Machine Learning R SDK:**
    Similar to the Python SDK, the R SDK provides R-language support for Azure Machine Learning.
12. **Azure Machine Learning Interpretability:**
    Interpretability tools help us understand and explain the behavior and predictions of machine learning models.
13. **Azure Machine Learning Data Prep SDK:**
    Data Prep SDK is used for data preparation and transformation tasks, allowing us to clean and preprocess data before training models.
14. **Azure Machine Learning Model Interpretability Toolkit (MIT):**
    MIT is a set of interpretability libraries that enable us to explain model decisions and understand feature importance.
15. **Azure Machine Learning Automated ML for Databricks:**
    This integration allows us to use AutoML capabilities within Azure Databricks notebooks.

16. **Azure Machine Learning Synapse Integration:**

Azure Synapse Analytics can be integrated with Azure Machine Learning for advanced analytics and machine learning within a unified platform.

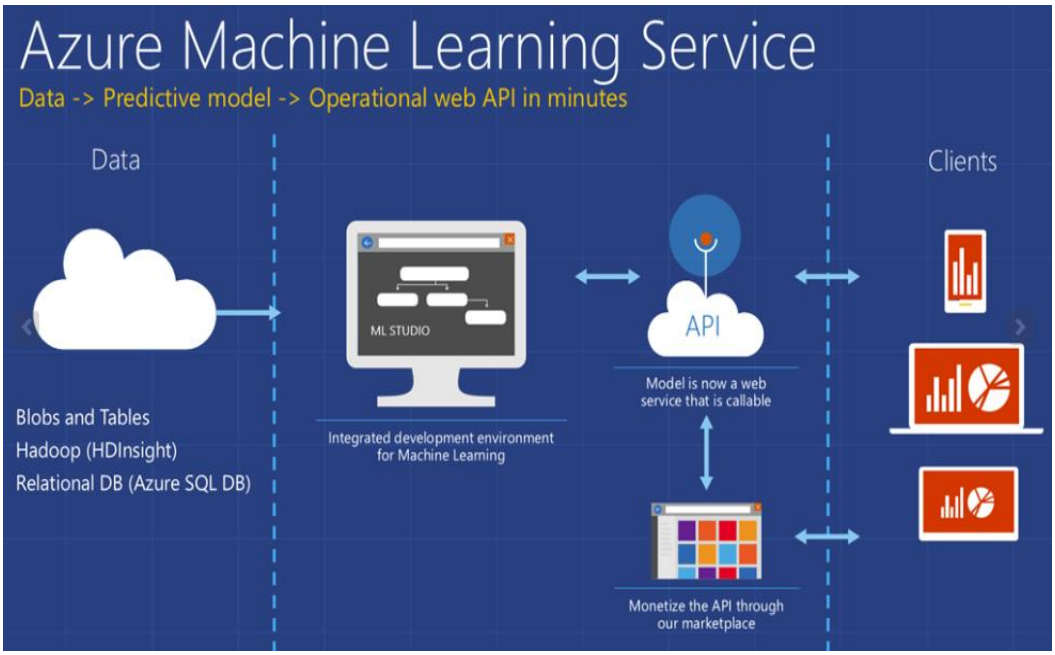


Figure 3.4.1

3.5 How to train a model?

- To train a model in Azure Machine Learning, we need to follow some steps as follows:

1. **Set Up Your Azure Machine Learning Workspace:**

- Create an Azure Machine Learning workspace in your Azure subscription. Access the workspace through the Azure portal or Azure Machine Learning Studio.

2. **Prepare Your Data:**

   Upload and store your dataset in Azure Blob Storage, Azure Data Lake Storage, or any other supported Azure data source.

   Ensure your data is properly cleaned, transformed, and preprocessed for training.

3. **Create a Machine Learning Experiment:**

   In Azure Machine Learning, we can create a new experiment to manage your training process.

   Use the visual interface or code-based approach, depending on your preference and expertise.

4. **Select a Compute Target:**

   Choose the compute target for training your model. Azure Machine Learning supports various compute options, including local compute, Azure Machine Learning Compute, and Azure Kubernetes Service (AKS).

5. **Define Your Training Script:**

   Write the Python or R script that defines your machine learning model and training process. This script should include the data loading, data preprocessing, model creation, and training code.

6. **Specify Training Parameters:**

   Configure training parameters such as hyperparameters, batch size, and training epochs. We can use Azure Machine Learning's hyperparameter tuning capabilities to optimize these settings.

7. **Submit the Training Job:**

   Submit your training job to Azure Machine Learning. The platform will take care of provisioning the compute resources, setting up the environment, and running your training script.

8. **Monitor Training Progress:**

   While the training job is running, we can monitor its progress using Azure Machine Learning's tracking capabilities. This includes viewing metrics, logs, and visualizations.

9. **Evaluate Model Performance:**

   After training, assess the performance of your model using validation datasets or cross-validation techniques.

   Azure Machine Learning provides tools for model evaluation and comparison.

10. **Register the Trained Model:**

    If your model meets your performance criteria, we can register it in your Azure Machine Learning workspace. This allows us to version and manage the model.

11. **Deploy the Model:**

    Once your model is registered, we can deploy it as a web service or endpoint using Azure Machine Learning Deployment Services. This makes your model accessible for predictions.

12. **Test and Monitor the Deployed Model:**

    After deployment, test the model's predictions and monitor its performance in a production environment.

    Azure Machine Learning provides monitoring and logging capabilities to track model behavior.

13. **Automate Model Retraining:**

    Consider setting up automated pipelines in Azure Machine Learning to retrain and deploy your model at regular intervals or in response to changes in data.

14. **Scale as Needed:**

    Azure Machine Learning allows us to scale your compute resources up or down based on demand. Adjust the compute configuration as your workloads evolve.

3.6 Costing for Azure Machine Learning:

- Costing required for Azure Machine Learning depends on various aspects and parameters:

- This is how the pricing calculator for Virtual Machine looks like:

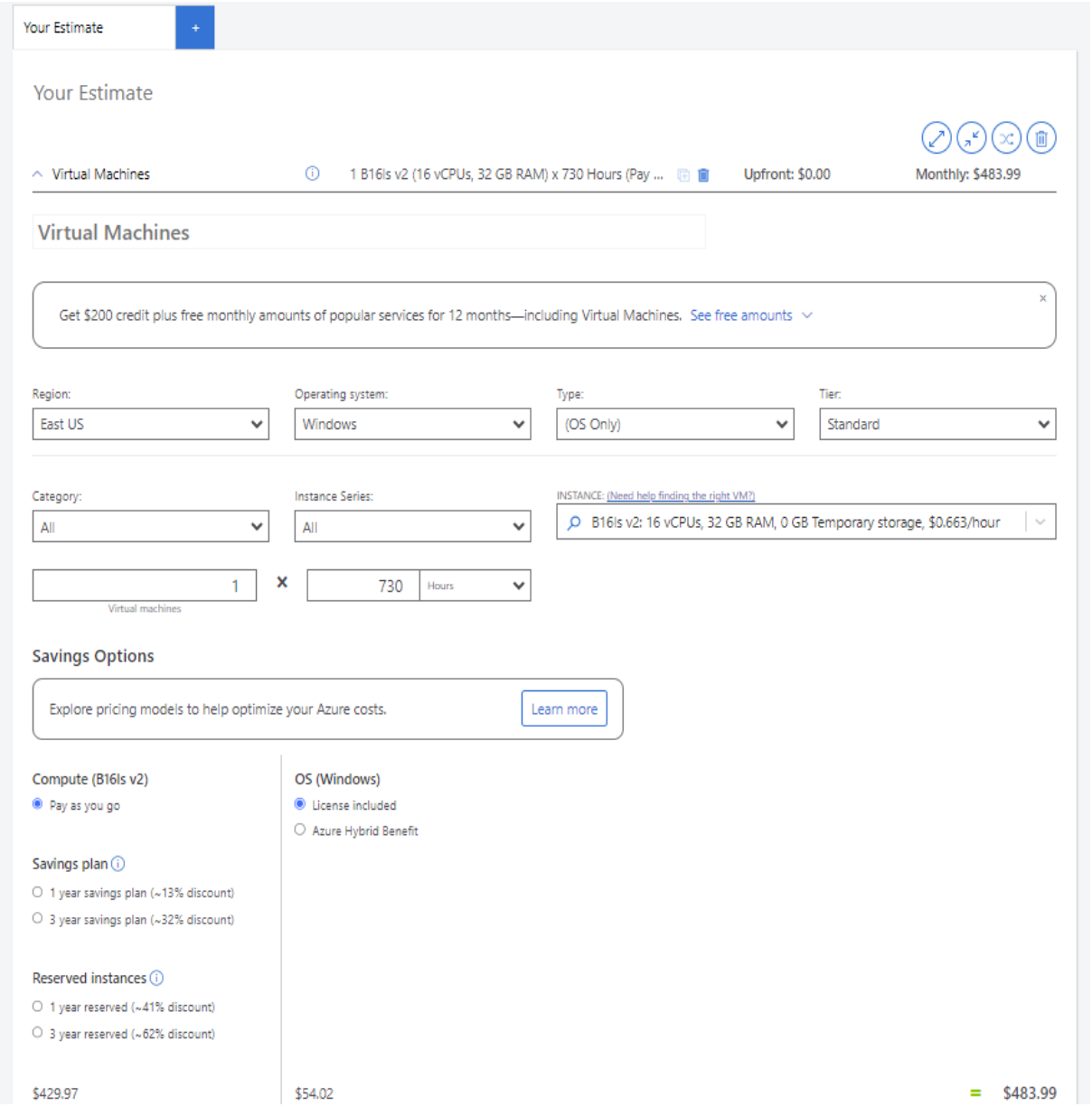


Figure 3.4.1

Instance	vCPU(s)	RAM	Linux VM Price	Machine Learning Service Surcharge	Pay As You Go Total Price	1 year savings plan	3 year savings plan
D2 v3	2	8 GiB	₹11.166/hour	N/A	N/A	₹9.68/hour	₹7.64/hour
D4 v3	4	16 GiB	₹22.414/hour	N/A	N/A	₹19.42/hour	₹15.32/hour
D8 v3	8	32 GiB	₹44.745/hour	N/A	N/A	₹38.76/hour	₹30.59/hour
D16 v3	16	64 GiB	₹89.489/hour	N/A	N/A	₹77.51/hour	₹61.17/hour
D32 v3	32	128 GiB	₹178.977/hour	N/A	N/A	₹155.02/hour	₹122.33/hour
D64 v3	64	256 GiB	₹357.953/hour	N/A	N/A	₹310.03/hour	₹244.67/hour

Figure 3.4.2

D2-64 v3

Instance	vCPU(s)	RAM	Linux VM Price	Machine Learning Service Surcharge	Pay As You Go Total Price	1 year savings plan	3 year savings plan
D2 v3	2	8 GiB	₹6,339.435/month	N/A	N/A	₹5,490.56/month	₹4,333.16/month
D4 v3	4	16 GiB	₹12,678.869/month	N/A	N/A	₹10,981.11/month	₹8,666.31/month
D8 v3	8	32 GiB	₹25,357.737/month	N/A	N/A	₹21,962.22/month	₹17,332.02/month
D16 v3	16	64 GiB	₹50,715.473/month	N/A	N/A	₹43,924.43/month	₹34,664.03/month
D32 v3	32	128 GiB	₹1,01,430.945/month	N/A	N/A	₹87,849.47/month	₹69,328.06/month
D64 v3	64	256 GiB	₹2,02,861.890/month	N/A	N/A	₹1,75,698.93/month	₹1,38,656.11/month

Figure 3.4.3

- This is how pricing is done on a month basis or according to hour basis.
- For each and every region we have different pricing systems and depends on the region that they apply for different countries.
- Figure 3.4.2 represents pricing based on cost per hour and figure 3.4.3 represents the costing based on month wise and price varies according to various region and different types of virtual machines.
- We have to first select the options according to our choice and it will calculate the pricing based on that.

3.7 Alternatives of Azure Machine Learning

- There are many other Machine Learning in the market other than Azure Machine Learning. Some of them are below:

1. **Amazon SageMaker (AWS SageMaker):**
   SageMaker is Amazon Web Services' (AWS) machine learning platform. It offers a comprehensive set of tools for building, training, and deploying machine learning models. It includes managed Jupyter notebooks, model training and tuning, and model deployment capabilities.

2. **Google Cloud AI Platform:**
   Google Cloud's AI Platform provides a suite of tools and services for machine learning, including data preparation, model training, and deployment. It supports popular ML frameworks and integrates with *Google's BigQuery* and *Cloud Dataflow*.

3. **IBM Watson Studio:**
   Watson Studio is IBM's data science and machine learning platform. It offers collaboration features, model building, automated machine learning (AutoAI), and deployment options. It integrates with *IBM Watson services*.

4. **Databricks:**
   Databricks Unified Analytics Platform combines data engineering, data science, and machine learning. It's built on Apache Spark and provides collaborative notebooks, automated machine learning, and model deployment capabilities.

5. **DataRobot:**
   DataRobot is an automated machine learning platform that aims to simplify and accelerate the machine learning process. It automates tasks such as feature engineering, model selection, and hyperparameter tuning.

6. **H2O.ai:**
   H2O.ai offers open-source and commercial machine learning platforms. H2O.ai's Driverless AI automates machine learning and feature engineering tasks, while H2O-3 provides open-source machine learning libraries.

7. **KNIME:**
   KNIME is an open-source platform for data analytics, reporting, and integration. It offers a range of machine learning and data science extensions for building and deploying workflows.

8. **RapidMiner:**
   RapidMiner is an open-source data science platform that provides an integrated environment for data preparation, modeling, and deployment. It supports various machine learning algorithms.

9. **SAS Analytics:**
   SAS offers a suite of advanced analytics and machine learning tools, including SAS Enterprise Miner and SAS Viya, which provide end-to-end machine learning capabilities.

10. **Scikit-learn:**
    Scikit-learn is an open-source machine learning library for Python. While it's not a complete platform like Azure ML, it's widely used for building and training machine learning models.

11. Kubeflow:

Kubeflow is an open-source machine learning platform designed for Kubernetes. It enables us to build, deploy, and manage machine learning workflows in containerized environments.

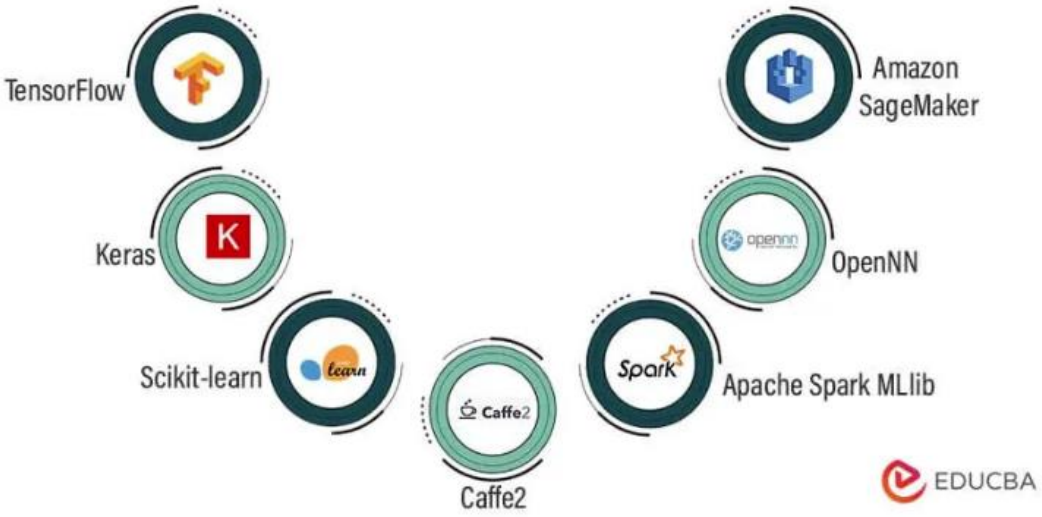


Figure 3.6.1

3.8 Steps to reduce the costing of Azure Machine Learning

- **Right-Size Compute Resources:**
  Choose the appropriate virtual machine (VM) sizes and GPU configurations for your machine learning workloads. Use smaller instances for development and testing and larger instances only when necessary for training.

- **Pause/Stop Resources:**
  Pause or stop compute resources when they are not in use, especially during evenings, weekends, or when we don't have active experiments or training jobs.

- **Budget and Cost Controls:**
  Establish clear budgets and spending controls within your organization to ensure cost-consciousness.

- **Review Pricing Regularly:**
  Stay informed about changes to Azure Machine Learning pricing and services, as Azure's offerings evolve over time.

- **Azure Cost Management and Billing:**

  Use Azure Cost Management and Billing tools to track and analyze your Azure Machine Learning expenses. Set up alerts to stay informed about cost overruns.

- **Establish Budgets and Cost Controls:**
  Implement budget limits and cost controls within your organization to ensure spending remains within acceptable levels.

➢ **Monitor cost:**

As we use Azure resources with Azure Machine Learning, we incur costs. Azure resource usage unit costs vary by time intervals (seconds, minutes, hours, and days) or by unit usage (bytes, megabytes, and so on.) As soon as Azure Machine Learning use starts, costs are incurred and we can see the costs in cost analysis.

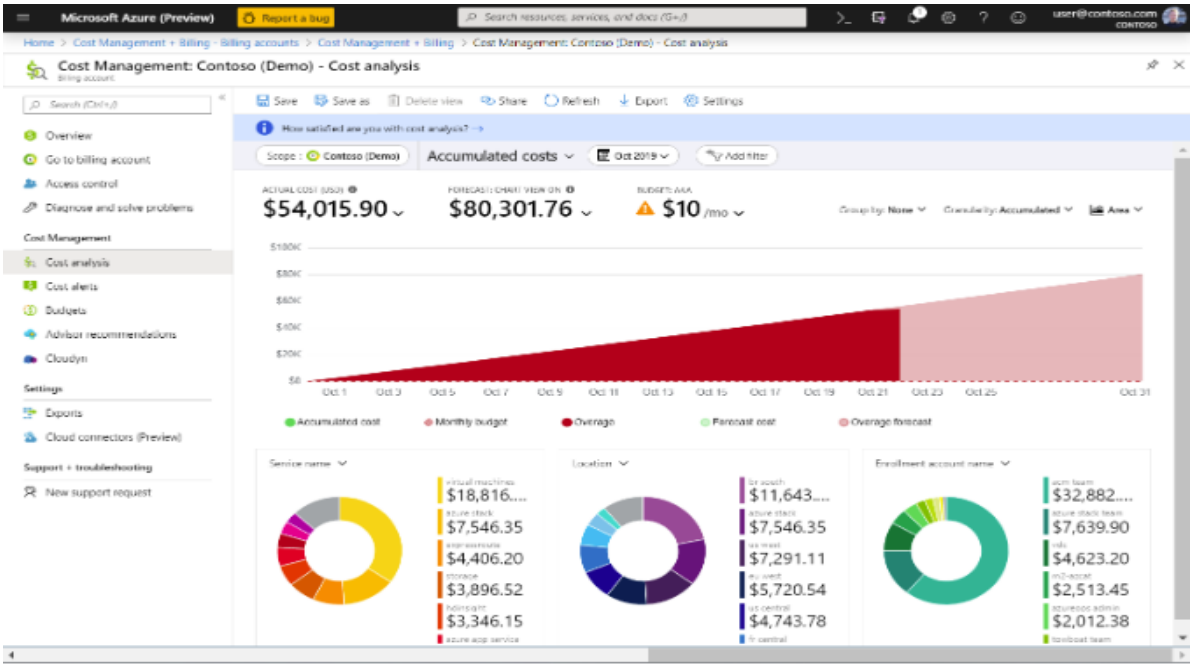


Figure 3.8.1

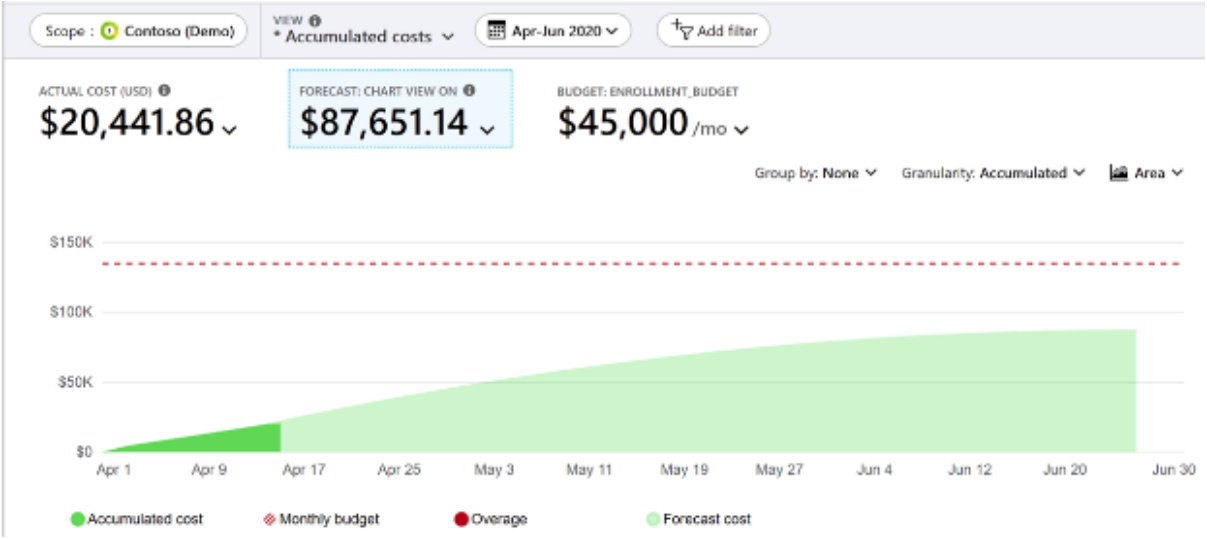


Figure 3.8.2