

Azure ML Studio

Azure雲端服務 個人體驗練習



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

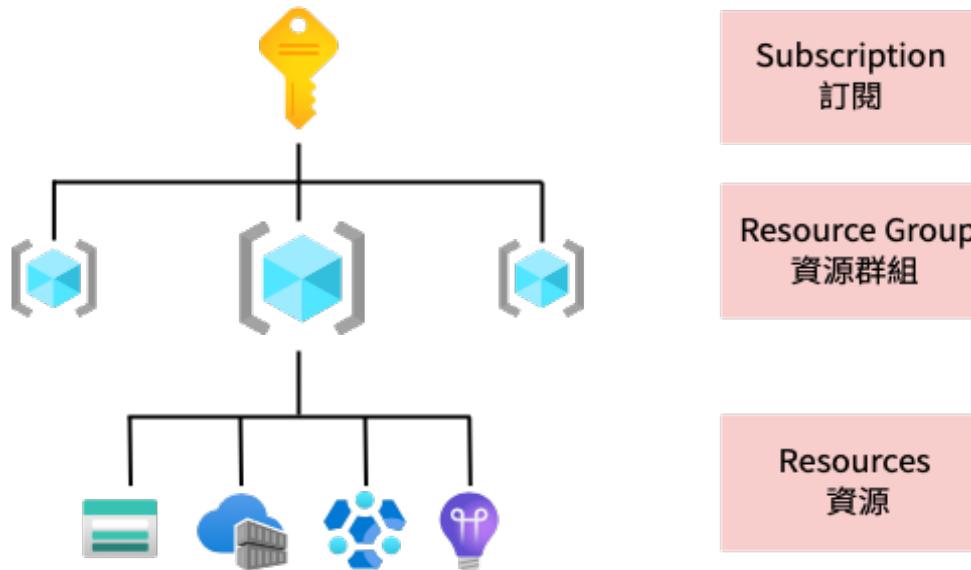
Azure雲端個人建置體驗

- 本課程的目的要體驗如何使用 Azure 雲端服務的各項工具
- 本次課程主要介紹Azure Machine Learning中的：
 - ◆設計工具 Designer
 設計工具是一項拖放工具，讓您以拖拉的方式建立機器學習模型，建立機器學習可以不需要任何一行程式碼，也可以透過編寫程式碼調整您的模型。

Azure環境介紹 -1



- Azure環境可分為多層次的架構，可以分層進行授權管理
- 友達有一個企業戶的訂閱Subscription，其下依照單位、專案等設定不同資源群組，在每一個資源群組內按照需求建立不同的資源(AutoML, Azure Databricks等)
- 建立資源一定要有相對應的訂閱、資源群組



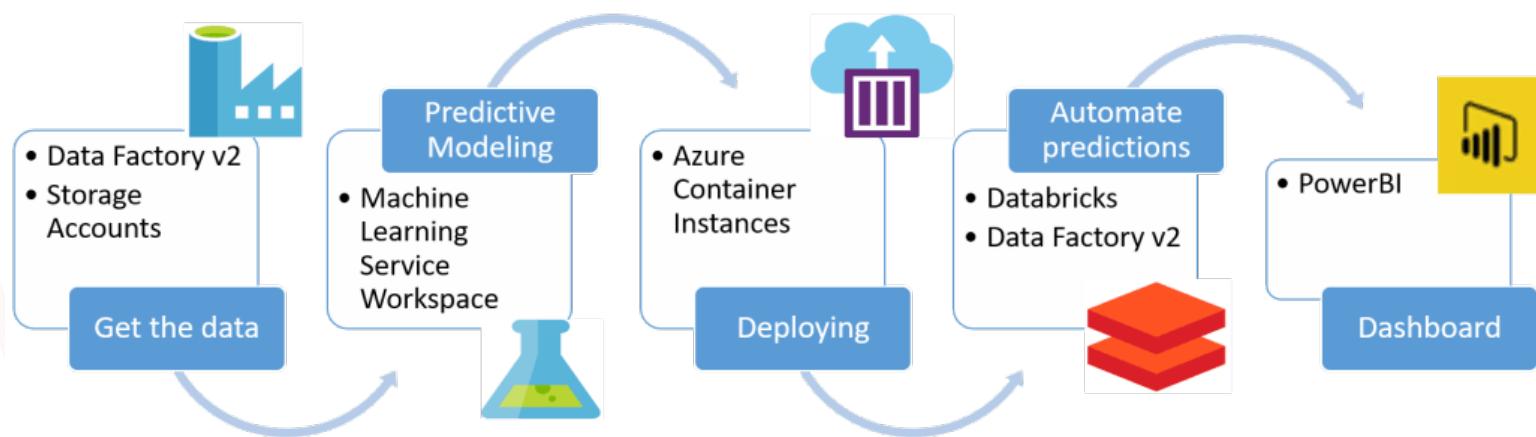
Azure環境介紹 -2

- Azure雲端上的服務資源是各自建立的，例如建立虛擬機器同時需要建立或連結到現有的儲存空間、虛擬網路等

Name	Type	Location
 MediaWebServer	Virtual machine	Japan East
 MicrosoftMonitoringAgent (Me...)	Microsoft.Compute...	Japan East
 MediaWebServer_OsDisk_1_b70...	Disk	Japan East
 mediawebserver285	Network interface	Japan East
 MediaWebServer-ip	Public IP address	Japan East
 MediaWebServer-nsg	Network security gr...	Japan East
 mslearn-vmupdate-vnet	Virtual network	Japan East

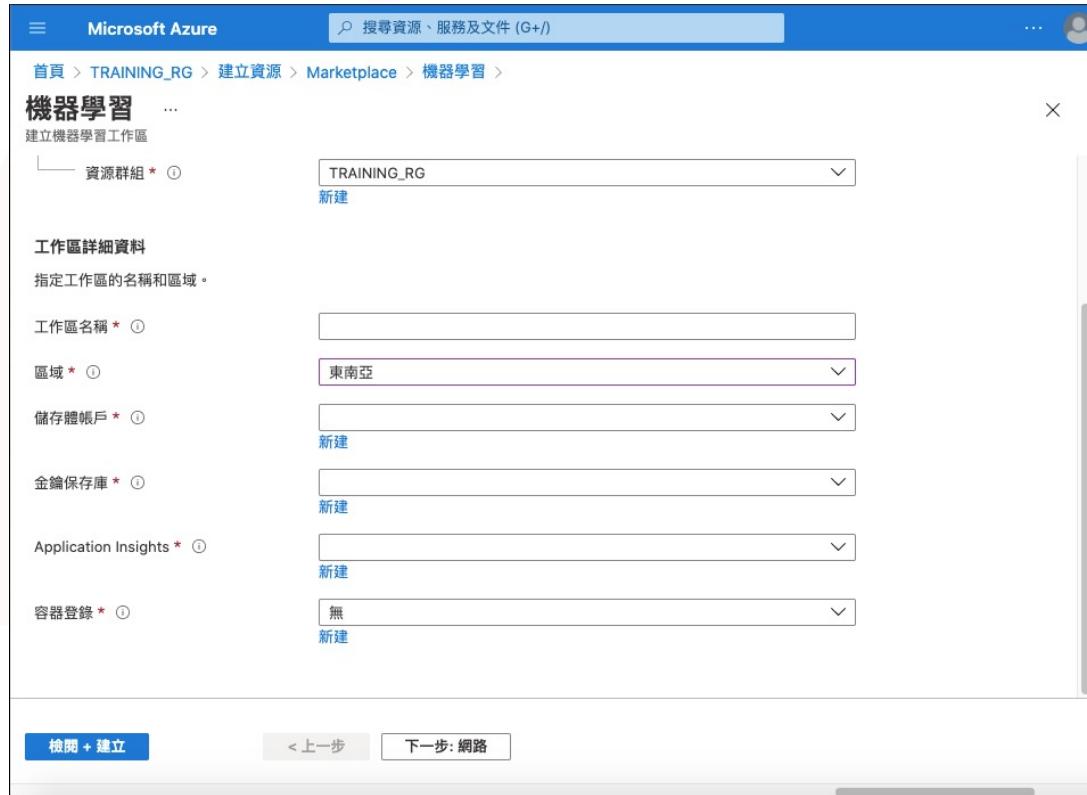
Azure雲端建置說明 -1

- Azure 雲端提供一系列的機器學習服務資源，從Raw Data的存放、資料處理、機器學習及視覺化呈現，都可以在Azure上找到相對應的資源。



Azure雲端建置說明 -2

- 建立Machine Learning Studio，會需要有相依的服務，如，存放紀錄的Storage Account、金鑰保存庫、給予模型建議的Application Insights、註冊機器學習產生Docker Image的容器登錄資源等



Azure雲端建置說明 -3

- 透過本課程練習使用 Azure Machine Learning Studio，可以理解並熟悉 Azure 雲端各項服務的操作與使用，包括
 - ◆ 建立儲存體
 - ◆ 使用 Azure 儲存體總管傳輸資料
 - ◆ 註冊應用程式
 - ◆ 設定資源存取控制
 - ◆ 建立與連接機器學習工作區
 - ◆ 使用 Azure Machine Learning Studio 的設計工具

AUO訓練 實作環境



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

AUO訓練實作環境介紹 -1

- 登入Azure Portal <https://portal.azure.com/>，點選Resource Group資源群組 "TRAINING_RG"

The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header bar with the Microsoft Azure logo, a search bar, and various navigation icons. On the right side of the header, the user's email (lihsiang.tu@caloudi.com) and company name (AU OPTRONICS CORP...) are displayed.

The main content area is titled "Azure 服務" (Azure Services). On the left, there's a sidebar with sections for "建立資源" (Create Resource), "最近的資源" (Recent Resources), and "瀏覽" (Browse). The "Recent Resources" section lists several items, including "TRAINING_RG".

The central part of the screen is the "資源群組" (Resource Group) blade for "TRAINING_RG". It shows a summary card with a "建立" (Create) button and a "檢視" (View) button. Below this, there's a "最近的資源" (Recent Resources) list with "TRAINING_RG" at the top, followed by other resources like "trainingml", "trainingml5524565781", "auotrainingadls", and "trainingdbs".

On the right side of the blade, there are links to "SQL 資料庫" (SQL Database), "適用於 PostgreSQL" (For PostgreSQL), "Azure Cosmos DB", "Kubernetes 服務" (Kubernetes Service), and a "更多服務" (More Services) button. Below these links, there's a table showing the "上次檢視時間" (Last checked time) for various resources, such as "32 分鐘之前" (32 minutes ago) for "TRAINING_RG".

At the bottom of the blade, there are links for "概觀" (Overview), "開始使用" (Get started), and "文件" (Files). The footer of the blade includes "所有資源" (All resources) and "儀表板" (Dashboard) buttons.

AUO訓練實作環境介紹 -2

- 進到資源群組”TRAINING_RG”的”概觀”畫面，您可以看到在此資源群組下的所有資源，如儲存體帳戶、機器學習等等

The screenshot shows the Microsoft Azure Resource Group Overview page for the 'TRAINING_RG' resource group. The left sidebar contains navigation links for Home, Overview, Activity Log, IAM, Tags, Events, Settings (Deployment, Security, Principles, Properties, Locks), Cost Management (Cost Analysis, Budgets, Cost Allocations), and Monitoring (Insights, Alerts, Metrics). The main content area displays the 'Overview' tab for the 'TRAINING_RG' resource group. It shows deployment details: '訂用帳戶 (變更) : auobigdata' (Subscription changed to: auobigdata), '部署 : 4 成功' (Deployment: 4 successful), and '位置 : 東南亞' (Location: Southeast Asia). A 'Tags' section indicates '按一下這裡即可新增標籤' (Click here to add a tag). Below this, a table lists 8 resources:

名稱	類型	位置	操作
ames-ml-deploy-4iNLQWdEVEO_tnaLTv1i_A	容器執行個體	東南亞	...
autrainingadls	儲存體帳戶	東南亞	...
trainingdbs	Azure Databricks 服務	東南亞	...
trainingml	機器學習	東南亞	...
trainingml3253432003	金鑰保存庫	東南亞	...
trainingml5524565781	儲存體帳戶	東南亞	...
trainingml6944426921	Application Insights	東南亞	...
trainingmlcr	容器登錄	東南亞	...

At the bottom, there are navigation buttons for '上一頁' (Previous page), '下一頁' (Next page), and a page number indicator '1 / 1'.

AUO訓練實作環境介紹 -3



^ 程式集

JSON 檢視

訂用帳戶 (變更) : auobigdata

部署 : 4 成功

訂用帳戶識別碼 : a7bdf2e3-b855-4dda-ac93-047ff722cbbd

位置 : 東南亞

標籤 (變更) : 按一下這裡即可新增標籤

篩選任何欄位...

類型 == 全部 ×

位置 == 全部 ×

+ 加入篩選

正在顯示 1 至 8 筆記錄 (共 8 筆)。 顯示隱藏的類型 ⓘ

未分組

清單檢視

<input type="checkbox"/> 名稱 ↑↓	類型 ↑↓	位置 ↑↓	更多
<input type="checkbox"/> trainingml6944426921	Application Insights	東南亞	...
<input type="checkbox"/> trainingdbs	Azure Databricks 服務	東南亞	...
<input type="checkbox"/> auotrainingadls	儲存體帳戶	東南亞	...
<input type="checkbox"/> trainingml5524565781	儲存體帳戶	東南亞	...
<input type="checkbox"/> ames-ml-deploy-4iNLQWdEVEO_tnaLTv1i_A	容器執行個體	東南亞	...
<input type="checkbox"/> trainingmlcr	容器登錄	東南亞	...
<input type="checkbox"/> trainingml	機器學習	東南亞	...
<input type="checkbox"/> trainingml3253432003	金鑰保存庫	東南亞	...

AUO訓練實作環境介紹

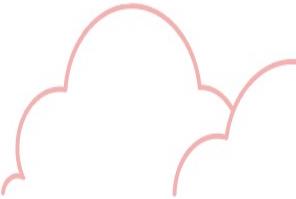
-4

資源名稱	類型	內容
auoautotrain小組編號	Application Insights	Azure對資源部署的智慧建議
dllakedemo	儲存體帳戶 Storage Account	ADLS Gen2，主要放置上課使用的講義和資料集
auoautotrain小組編號	儲存體帳戶 Storage Account	主要存放ML相關Log紀錄
建立模型後產生	容器執行個體 Container Instances	部署ML設計工具產生的模型
建立模型後產生	容器登錄 Container Registry	Container Registry，因Cluster產生的資源
MLStudio員工編號	機器學習 Machine Learning	機器學習的工作站 workspace
auoautotrain小組編號	金鑰保存庫 Key Vault	儲存資源群組的金鑰

Azure 儲存體介紹

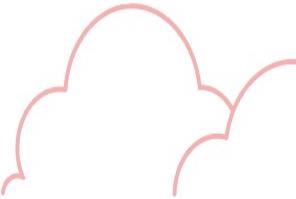


- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具



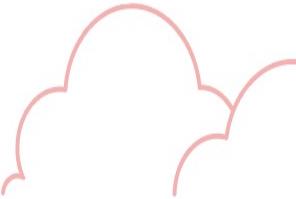
Azure Blob 儲存體 -1

- Azure Blob 儲存體是 Microsoft 針對雲端推出的物件儲存體解決方案，Blob 儲存體已針對儲存大量非結構化資料進行最佳化
- Blob 儲存體設計用來：
- 直接提供映像或文件給瀏覽器
- 儲存檔案供分散式存取
- 串流影片和音訊
- 寫入記錄檔
- 儲存備份和還原、災害復原和封存資料
- 儲存資料供內部部署或 Azure 裝載服務進行分析



Azure Blob 儲存體 -2

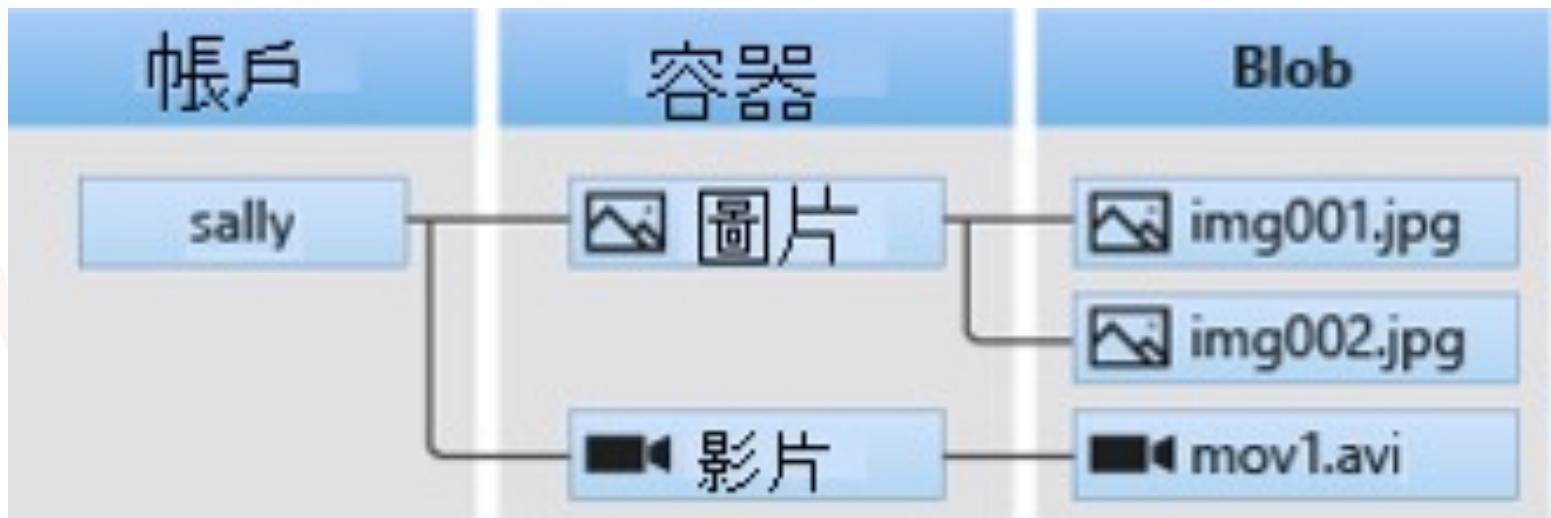
- 使用者或用戶端應用程式可以從世界各地透過 HTTP/HTTPS 存取 Blob 儲存體中的物件
- Blob 儲存體中的物件可透過 Azure 儲存體 REST API、Azure PowerShell、Azure CLI 或 Azure 儲存體用戶端程式庫存取
- 用戶端程式庫適用於不同的語言，包括：
 - ◆.NET
 - ◆Java
 - ◆Node.js
 - ◆Python
 - ◆Go
 - ◆PHP
 - ◆Ruby

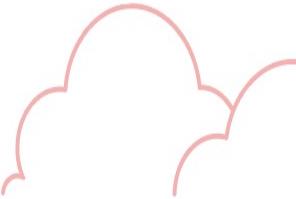


Blob 儲存體資源 -1

- Blob 儲存體提供三種類型資源：

- ◆ 儲存體帳戶
- ◆ 儲存體帳戶中的容器
- ◆ 容器中的 Blob





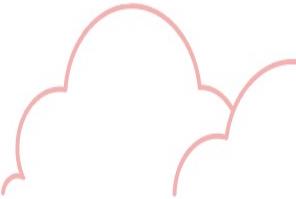
Blob 儲存體資源 -2

- 儲存體帳戶

- ◆ 儲存體帳戶會在 Azure 中為您的資料提供唯一命名空間。每個儲存在 Azure 儲存體中的物件都有一個位址，其中包含您的唯一帳戶名稱。帳戶名稱與 Azure 儲存體 Blob 端點的組合會形成儲存體帳戶中物件的基底位址
 - ◆ 例如: <http://mystorageaccount.blob.core.windows.net>

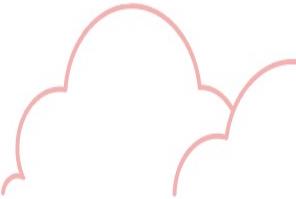
- 容器

- ◆ 容器會組織一組 Blob，類似於檔案系統中的目錄。儲存體帳戶可以包含無限數量的容器，而一個容器則可儲存無限數量的 Blob



Blob 儲存體資源 -3

- Azure 儲存體支援三種 blob 類型：
 - ◆ 區塊 Blob，可儲存文字和二進位資料。區塊 Blob 是由可個別管理的資料區塊所組成。區塊 blob 最多可以儲存約 190.7 TiB
 - ◆ 附加 Blob，和區塊 Blob 相似，由區塊所組成，但已針對附加作業最佳化。附加 Blob 很適合某些案例，例如記錄虛擬機器中的資料
 - ◆ 分頁 Blob，可儲存隨機存取檔案 (大小上限為 8 TiB)。分頁 Blob 存放區可儲存虛擬硬碟 (VHD) 檔案，以作為 Azure 虛擬機器的磁碟



Azure Data Lake Store -1

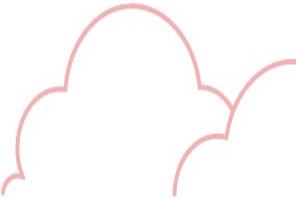
- Azure Data Lake Store提供了一個單一的存儲庫，無論大小型組織都可以在其中上傳無限大小的數據
- 用於Hadoop分佈式檔案系統工具和應用程式的高性能處理和分析，包括對低延遲工作負載的支源
- 允許使用本機格式的結構化和非結構化數據
- 允許巨大的資料流量以提高分析性能
- 提供了高可用性，耐用性和可靠性
- ADLS的類型
 - ◆ADLS Gen1
 - ◆ADLS Gen2

ADLS Gen1 -1

- 全名為 Azure Data Lake Storage Gen1
- Azure Data Lake Storage Gen1 是容納巨量資料分析工作負載的企業級超大規模存放庫。 Azure Data Lake 可讓您在單一位置擷取任何大小、類型和擷取速度的資料，以便進行運作和探究分析
- 使用與 WebHDFS 相容的 REST API，可以從 Hadoop (HDInsight 叢集所提供之) 存取 Data Lake Storage Gen1
- 為預存資料分析而設計，並針對資料分析案例而調整效能
- Data Lake Storage Gen1 包含所有企業級功能：安全性、管理性、延展性、可靠性和可用性
- 2024 年 2 月 29 日 Azure Data Lake Storage Gen1 將會淘汰

ADLS Gen1 -2

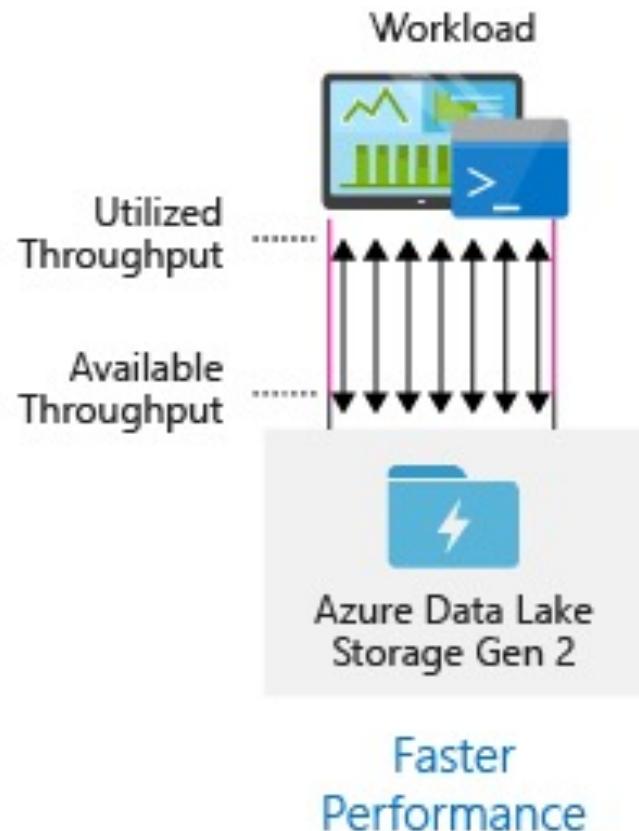
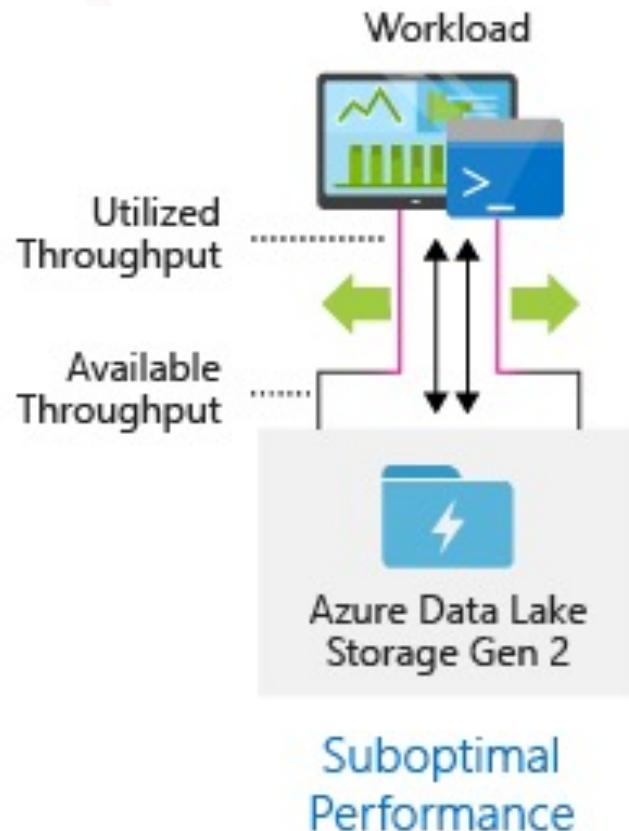




ADLS Gen2 -1

- Azure Data Lake Storage Gen2 是一組巨量資料分析的專屬功能，內建於 Azure Blob 儲存體
- Data Lake Storage Gen2 包含 Azure Data Lake Storage Gen1 的功能和 Azure Blob 儲存體
- 例如，Data Lake Storage Gen2 會提供檔案系統語法、檔案層級安全性和規模調整。因為這些功能是建置於 Blob 儲存體之上，所以您也可以取得低成本的分層式儲存體，以及高可用性/災害復原功能
- 效能經過最佳化，相較於 Blob 儲存體上的一般命名空間，階層命名空間大幅提高目錄管理作業的效能，從而提高整體作業效能。
- 管理更容易，可以透過目錄和子目錄整理和操作檔案
- 安全性是強制的，可以在目錄或個人檔案上定義 POSIX 權限

ADLS Gen2 -2



AUO實作環境說明

- 建立ADLS Gen2 作為Big Data儲存空間
- 建立與Machine Learning相依的Storage Account，儲存機器學習中的預設資料
- 本次課程以 **dllakedemo** 為放置資料集的儲存體

<input type="checkbox"/> 名稱 ↑↓	類型 ↑↓	位置 ↑↓
<input type="checkbox"/> trainingml6944426921	Application Insights	東南亞
<input type="checkbox"/> trainingdbs	Azure Databricks 服務	東南亞
<input type="checkbox"/> auotrainingadls	儲存體帳戶	東南亞
<input type="checkbox"/> trainingml5524565781	儲存體帳戶	東南亞
<input type="checkbox"/> ames-ml-deploy-4iNLQWdEVEO_tnaLTv1i_A	容器執行個體	東南亞
<input type="checkbox"/> trainingmlcr	容器登錄	東南亞
<input type="checkbox"/> trainingml	機器學習	東南亞
<input type="checkbox"/> trainingml3253432003	金鑰保存庫	東南亞

建立 Data Lake Storage Gen2 存放區



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

建立資源群組 -1

- 建立新的資源群組來保存 Data Lake 儲存體
- 使用資源群組來管理相關服務和應用程式
- 資源群組也會讓清除資源的過程更為容易
- 登入 Azure 入口網站
- 選取左上方的 [+建立資源] 。
- 在搜尋方塊中，鍵入“Resource”，並從結果選取“Resource Group”
- 選取 [建立] 來新增新的資源群組
- 在 [基本] 索引標籤上，選取您想要使用的訂用帳戶，本次課程使用的是“**azuretrain01**”
- 將資源群組命名，您在此處使用的是“**autotrain小組編號**”
- 選擇資源群組的區域 (位置)在東南亞 Southeast Asia

建立資源群組 -2

- 選取 [檢閱 + 建立] > [建立]

首頁 > 新增 > 資源群組 >

建立資源群組 ...

基本 標籤 檢閱 + 建立

資源群組 - 握有 Azure 解決方案相關資源的容器。該資源群組可包含解決方案的所有資源，或是只包含您想以群組方式管理的資源。您可以最適合您組織的方式，決定如何將資源配置到資源群組。[深入了解](#)

專案詳細資料

訂用帳戶 * ①

timyang

資源群組 * ①

mslearn-test

資源詳細資料

區域 * ①

(Asia Pacific) 東南亞

檢閱 + 建立

< 上一步

下一步 : 標籤 >

建立 Data Lake 帳戶 -1

- 在 Azure 入口網站功能表的選單，選取 [儲存體帳戶] > [建立]

The screenshot shows the Azure portal interface. On the left, there is a sidebar with various service links. The 'Storage Accounts' link is highlighted with a red box. The main content area is titled 'Azure 服務' (Azure Services) and shows a grid of service icons. One icon, '儲存體帳戶' (Storage Account), is highlighted with a red box. Below the grid, there is a detailed view of a Storage Account named '儲存體帳戶'. This view includes sections for '建立' (Create), '檢視' (View), '描述' (Description), '類型' (Type), and '上次檢視時間' (Last checked time). The '建立' button is also highlighted with a red box. At the bottom of the main content area, there are sections for 'Microsoft 提供的免費訓練' (Free Microsoft training), 'Azure 監視器' (Azure Monitor), '資訊安全中心' (Information Security Center), and '成本管理' (Cost Management).

建立 Data Lake 帳戶 -2

- 選取您的“訂用帳戶azuretrain01”和“資源群組Training_RG”。
- 輸入儲存體帳戶的名稱。此名稱在整個 Azure 中必須是唯一的。您可能必須嘗試數種變化，才能找到唯一的名稱。本次課程使用“**dlakedemo**”為例，您在命名時可以使用首碼並搭配一些數字。綠色核取記號表示您輸入的名稱有效。
- 選擇資源群組的區域(位置)為東南亞
- 確認效能為標準(一般用途 V2)
- 備援更改為本地備援儲存體(LRS)
- 其餘欄位則保留預設值
- 選取上方的 [進階] 標籤

建立 Data Lake 帳戶 -3

建立儲存體帳戶 ...

基本 進階 網路 資料保護 標籤 檢閱 + 建立

選取要在其中建立新儲存體帳戶的訂閱。請選擇新的或現有的資源群組，以便一併整理和管理您的儲存體帳戶與其他資源。

訂用帳戶 *

azuretrain01

資源群組 *

autoDemo

新建

執行個體詳細資料

如果您需要建立舊版儲存體帳戶類型，請按一下[這裡](#)。

儲存體帳戶名稱 ① *

dlakecaloudi

區域 ① *

(Asia Pacific) 東南亞

效能 ① *

標準: 建議用於大多數案例 (一般用途 v2 帳戶)

進階: 建議用於需要低延遲的案例。

備援 ① *

本地備援儲存體 (LRS)

檢閱 + 建立

< 上一步

下一步 : 進階 >

建立 Data Lake 帳戶

-4

- 在 [Data Lake Storage Gen2] 區段的 [階層命名空間] 旁邊，選取 [啟用]。
- 選取 [檢閱 + 建立] 建立儲存體帳戶
- 在驗證建立詳細資料後，選取 [建立] 來啟動部署
- 當您收到訊息，指出您的部署已完成時，請選取 [前往資源] 來確認部署

建立儲存體帳戶

需要安全傳輸 已停用 已啟用

允許共用金鑰存取 已停用 已啟用

TLS 最低版本

基礎結構加密 已停用 已啟用

● 目前必須註冊，才能以訂用帳戶為單位啟用基礎結構加密。[註冊基礎結構加密](#)

Blob 儲存體

允許 Blob 公開存取 已停用 已啟用

Blob 存取層 (預設) 非經常性 經常性

NFS v3 已停用 已啟用

● 目前必須註冊，才能以訂用帳戶為單位使用 NFS v3 功能。 [註冊 NFS v3](#)

Data Lake Storage Gen2

階層式命名空間 已停用 已啟用

Azure 檔案儲存體

大型檔案共用 已停用 已啟用

< 上一步 下一步 : 標籤 >

建立 Data Lake 帳戶 -5



首頁 > Microsoft.StorageAccount-20210402095428 >

dlakedata8809 儲存體帳戶

搜尋 (Ctrl+ /) 在總管中開啟 移動 重新整理 刪除 意見反應

Microsoft recommends upgrading to the new alerts platform to ensure no interruptions in your alerts. Classic alerts will be retired starting in 2021. Upgrade to the new alerts platform. [深入了解](#).

概觀

資源群組 (變更) : mslearn-test
狀態 : 主要: 可用, 次要: 可用
位置 : 東南亞, 東亞
訂用帳戶 (變更) : timyang
訂用帳戶識別碼 : f4c08b1d-5835-4123-804a-6d01bb0446ab
標籤 (變更) : 按一下這裡即可新增標籤

儲存體基本資訊
雙能/仔取/層 : 標準/經常性
複寫 : 讀取權限異地備援儲存體 (RA-GRS)
帳戶種類 : StorageV2 (一般用途 v2)

程式集

容器 可大幅調整的 Data Lake Storage [深入了解](#)

檔案共用 無伺服器 SMB 和 NFS 檔案共用 [深入了解](#)

資料表 表格式資料儲存體 [深入了解](#)

佇列 根據流量有效率地縮放應用程式 [深入了解](#)

工具和 SDK

儲存體總管 (預覽) PowerShell Azure CLI .NET Java Python Node.js

監視

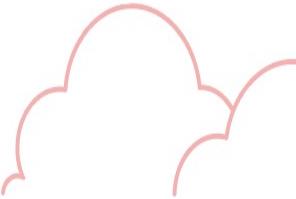
顯示下列時間範圍的資料: 1 小時 2 小時 4 小時 1 天 7 天

This screenshot shows the Azure Storage Account management portal for the account 'dlakedata8809'. The left sidebar contains navigation links for monitoring, activity logs, tags, troubleshooting, IAM, data migration, events, and storage overview. The main content area displays basic account information, including resource group, location, and usage details. It also features four service cards: Container (Data Lake Storage), File sharing, Table storage, and Queue storage. Below these are links for tools and SDKs (PowerShell, Azure CLI, .NET, Java, Python, Node.js) and a monitoring section with time range filters.

上傳檔案



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

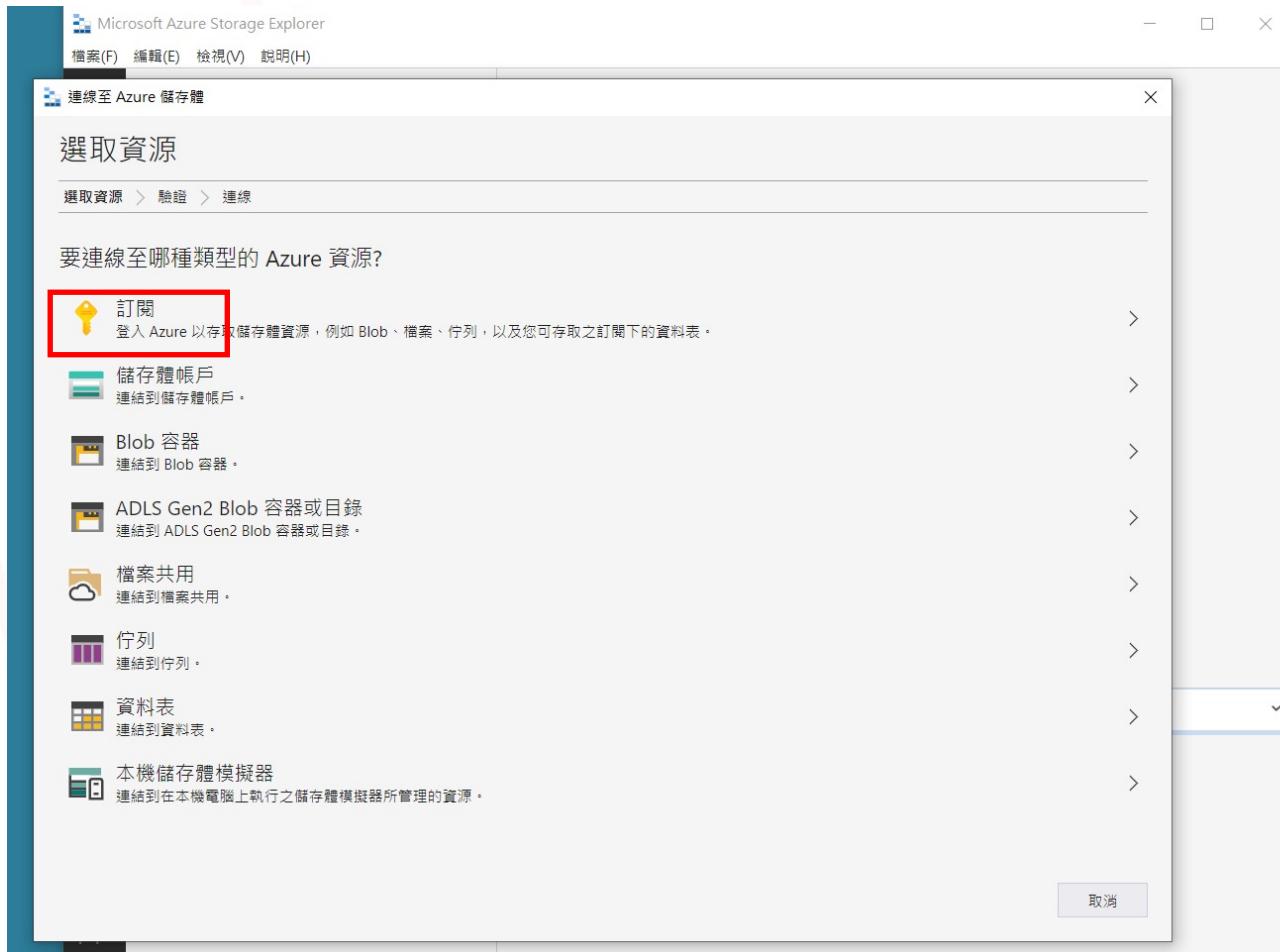


Azure 儲存體總管

- 若要對 Azure Data Lake 存放區執行臨機操作資料傳輸，請使用 Azure 儲存體總管。
- 儲存體總管是適用於 Windows、macOS 和 Linux 的免費應用程式
- 這項 Azure 應用程式旨在管理非結構化的資料，例如資料表、Blob、佇列和檔案
- 它也支援 Azure Cosmos DB 和 Azure Data Lake Storage Gen2 中的資料
- 下載及安裝儲存體總管
 - ◆ <https://azure.microsoft.com/zh-tw/features/storage-explorer/>

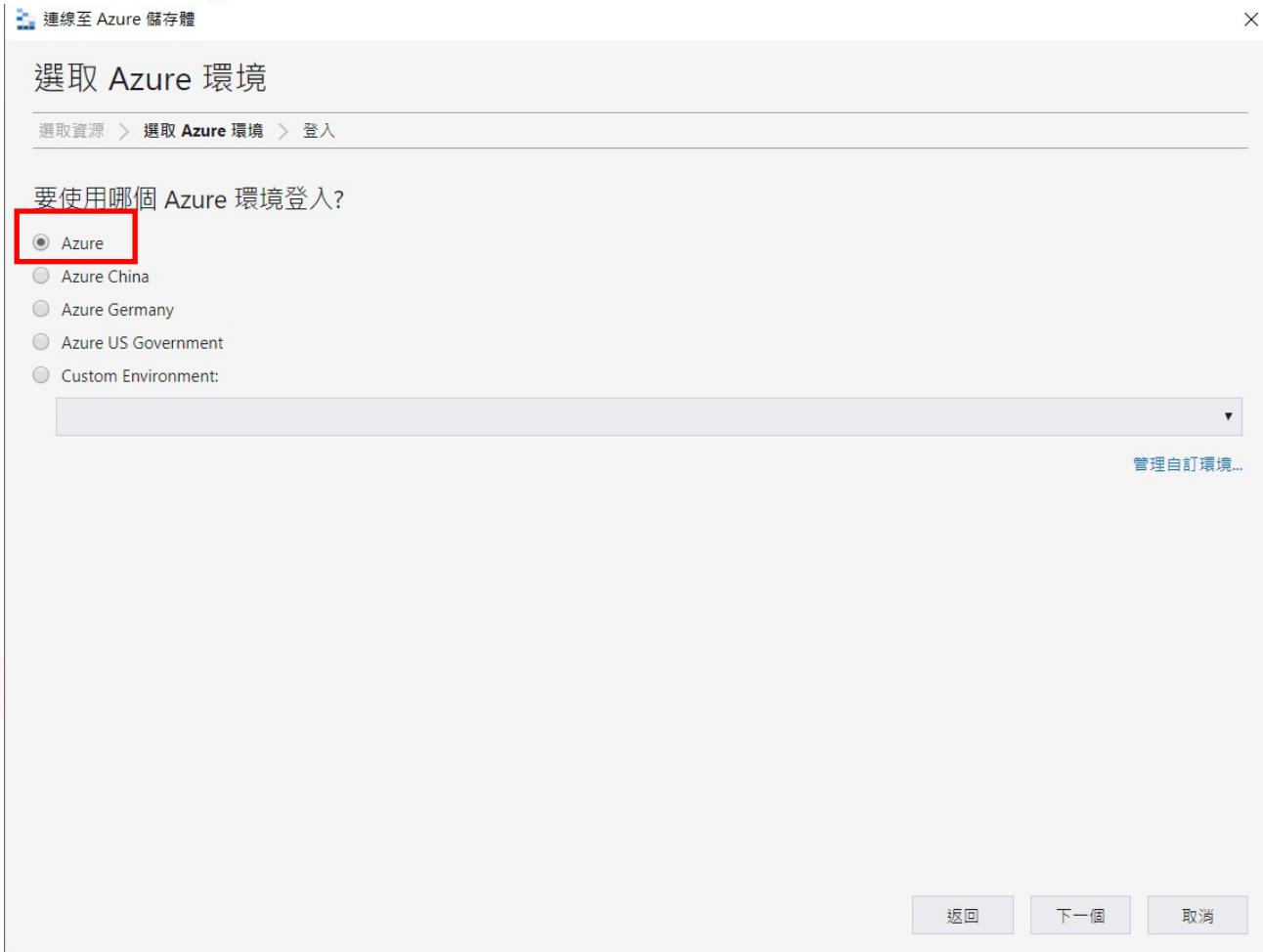
設定帳號 -1

- 若要開始使用 [儲存體總管]，請先新增您的 Azure 帳戶
- 選擇 [訂閱]



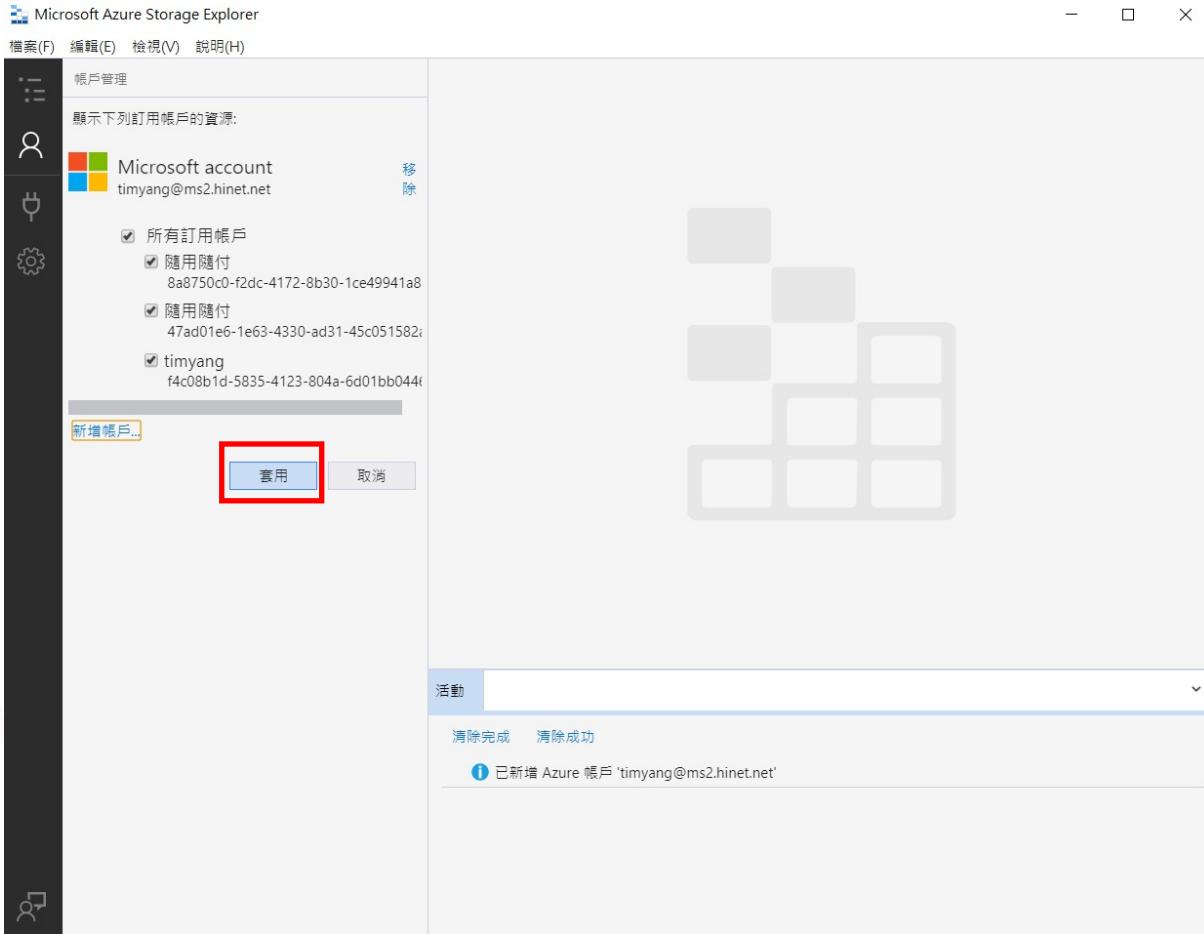
設定帳號 -2

● 選擇 [Azure] 登入



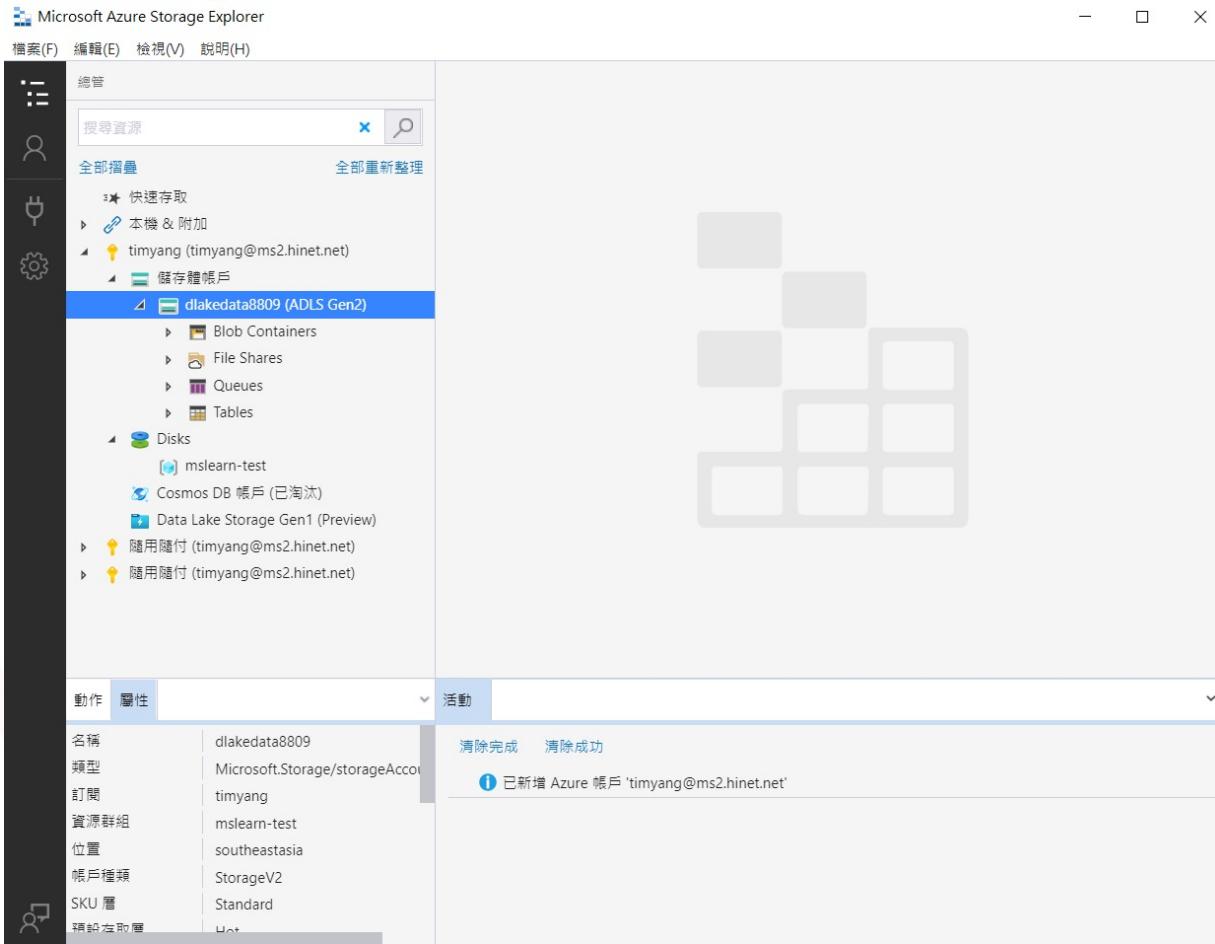
設定帳號 -3

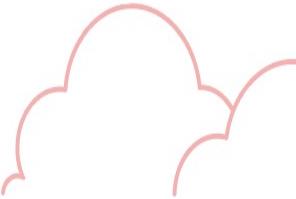
- 在您登入後，請選取您要使用的訂用帳戶。請務必選取您已在其中建立 Azure 儲存體帳戶的訂用帳戶。



設定帳號 -4

- 應用程式會顯示您可以從訂用帳戶中使用的儲存體區域清單。您的 Azure 儲存體帳戶應該會出現在清單上



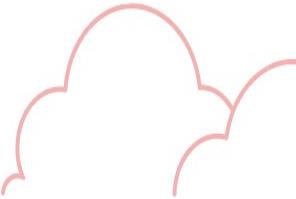


建立檔案系統

- 因為 Blob 一律會上傳至資料夾，因此您可以組織 Blob 群組，其方式就跟組織您電腦上檔案的方式雷同
- 在 Data Lake Storage Gen2 中，您會從建立「檔案系統」開始。這會定義將保存您 Data Lake 的 Blob 儲存體容器。您會在此專用區域內建立資料夾和檔案
- 在 [儲存體總管] 中，展開您的訂用帳戶，然後展開您的儲存體帳戶
- 展開您先前建立的儲存體帳戶，然後選取 [Blob 容器]
- 以滑鼠右鍵按一下 [Blob 容器]，然後選取 [建立 Blob 容器]
- 在 [Blob 容器] 下方，輸入 demodata
- 建立容器後，請選取 demodata

建立資料夾

- 資料夾可協助為您的資料提供階層結構。您可以在帳戶中建立多個層級。您只需要確保在您建立子資料夾前，主要資料夾已存在
- 在位於 [儲存體總管] 頂端的功能表上，選取 [新增資料夾]
- 將資料夾命名為 sample。然後選取 [確定] 來建立目錄。若您收到訊息，詢問您是否要重新整理您的檢視，請選取 [是]
- 按兩下新的資料夾。此動作會進入資料夾，此時資料夾應該是空白的
- 建立另一個名為 data 的資料夾。



上傳檔案

- 您可以將檔案和資料夾從本機電腦上傳至位於您 [儲存體總管] 右側檔案共用中的目錄
- 在 [儲存體總管] 中，按兩下 data 資料夾
- 在視窗頂端的功能表上，選取 [上傳]。從 [上傳] 功能表中，您可以上傳資料夾或檔案
- 選取 [上傳檔案]
- 選取您要上傳檔案
- 在 [上傳檔案] 對話方塊中，確認目的地目錄是 sample/data。然後選取 [上傳]

註冊應用程式



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

應用程式註冊 -1

- Microsoft 身分識別平臺只會對已註冊的應用程式執行 (IAM) 的身分識別和存取管理
- 無論是用戶端應用程式 (例如 web 或行動應用程式) , 或支援用戶端應用程式的 Web API , 註冊會在您的應用程式與身分識別提供者 (Microsoft 身分識別平台) 之間建立信任關係
- 登入 Azure 入口網站
- 搜尋並選取 [Azure Active Directory]
- 在“管理”下選取 [應用程式註冊] > [新增註冊]
- 輸入應用程式的顯示“名稱” , 例如“MyDemoAPP”
- 指定可以使用應用程式的人員 , 預設為“僅此組織目錄中的帳戶”

應用程式註冊 -2

● 選取 [註冊] 以完成伺服器初始註冊

首頁 > 預設目錄 >
註冊應用程式 ...

* 名稱

此應用程式的使用者互動顯示名稱 (之後可以變更)。

MyDemoApp



支援的帳戶類型

誰可以使用此應用程式或存取此 API?

- 僅此組織目錄中的帳戶 (僅 預設目錄 - 單一租用戶)
- 任何組織目錄中的帳戶 (任何 Azure AD 目錄 - 多租用戶)
- 任何組織目錄中的帳戶 (任何 Azure AD 目錄 - 多租用戶) 和個人 Microsoft 帳戶 (例如 Skype、Xbox)
- 僅限個人 Microsoft 帳戶

[協助我選擇...](#)

重新導向 URI (選用)

我們會在成功驗證使用者之後，將驗證回應傳回給此 URI。現在已不一定需要在此時提供此 URL，可以在之後變更，但在大多數的情況下會需要值。

Web



e.g. https://example.com/auth

[繼續即表示您同意《Microsoft 平台原則》](#)

[註冊](#)

應用程式註冊 -3

- 註冊完成時，Azure 入口網站會顯示應用程式註冊的 [總覽] 窗格
- 您會看到“應用程式 (用戶端) 識別碼”，此值也稱為“用戶端 識別碼”，可在 Microsoft 身分識別平臺中唯一識別您的應用程式

The screenshot shows the Microsoft Azure portal interface for managing application registrations. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information (meganb@contoso.com, CONTOSO AD (DEV)). Below the navigation is a breadcrumb trail: Home > Contoso AD (dev) | App registrations > Contoso App 1.

The main content area displays the details for 'Contoso App 1'. On the left, a sidebar lists navigation options: Overview (selected), Quickstart, Integration assistant (preview), Manage (with sub-options: Branding, Authentication, Certificates & secrets, Token configuration, API permissions, Expose an API), and a search bar.

The right pane shows the following details for the application:

- Display name:** Contoso App 1
- Application (client) ID:** 11111111-1111-1111-1111-111111111111 (highlighted with a red box)
- Supported account types:** My organization only
- Directory (tenant) ID:** 00000000-0000-0000-0000-000000000000
- Object ID:** 22222222-2222-2222-2222-222222222222
- Supported account types:** My organization only
- Redirect URLs:** Add a Redirect URI (with the URL https://contoso.com/oauth2/authorize)
- Application ID URI:** Add an Application ID URI
- Managed application in local directory:** Contoso App 1

A note at the bottom states: "Starting June 30th, 2020 we will no longer add any new features to Azure Active Directory Authentication Library (ADAL) and Azure AD Graph. We will continue to provide technical support and security updates but we will no longer provide feature updates. Applications will need to be upgraded to Microsoft Authentication Library (MSAL) and Microsoft Graph. Learn more".

新增用戶端密碼 -1

- 用戶端密碼也稱為“應用程式密碼”
- 它是一個字串值，您的應用程式可以使用此值來取代憑證本身的身分識別
- 在 Azure 入口網站的“應用程式註冊(App Registrations) 中，選取您的應用程式
- 選取 [憑證及祕密] > [新增用戶端密碼]
- 新增用戶端密碼的描述
- 選取持續時間
- 選取 [新增]。
- 記錄秘密的值，以便在您的用戶端應用程式程式碼中使用。離開此頁面後，就不會再次顯示此秘密值

新增用戶端密碼 -2



所有服務 > 應用程式註冊 > MyDemoApp

MyDemoApp | 憑證及祕密 ⚙ ...

The screenshot shows the Azure portal interface for managing application secrets. On the left, the navigation menu is visible with options like Overview, Quick Start, Integration assistant, Management (Brands, Verification, Certificates & Secrets), Permissions (OAuth 2.0 Client ID, API Permissions, Public API), App roles, Owners, Roles and System Administrators, and Information (Information, Support tickets). The 'Certificates & secrets' option is selected and highlighted.

In the main content area, a modal window titled '新增用戶端密碼' (Add Client Secret) is open. It contains fields for '描述' (Description) and '到期' (Expiration). Under '到期', the '1 年內' (Within 1 year) radio button is selected. At the bottom of the modal are '新增' (Add) and '取消' (Cancel) buttons, with '新增' being the active button.

Below the modal, the '用戶端密碼' (Client Secret) section is displayed. A note states: '應用程式在要求權杖時，用以證明其身分識別的祕密字串，也稱為應用程式密碼。' (A secret string used by the application to prove its identity when requesting tokens, also known as the application password.)

A table lists the existing client secret:

描述	到期	值	ID
Password uploaded on Sat Apr 03 2021	3/4/2022	~_n*****	902ba91b-c42e-4757-acfd-0b308ccfe28f

The '值' (Value) column for the secret is highlighted with a red box.

上課示範用密碼

- Application/Client ID:

cf446eaf-3099-4179-94e3-20b51bc4a78f

- Object ID:

cece5a9f-8e44-4398-82af-4b39d6548af5

- Directory/Tenant ID:

e7c24d00-7479-4343-8247-0204699693e8

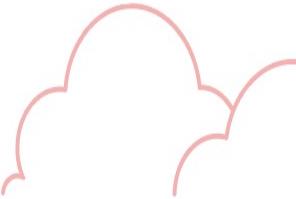
- Secrets Value:

SBoFq1TfYG2~Z1E_4N3i_ialkLs0bgN~ZU

設定資源 存取控制



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具



存取 Blob

- Azure Active Directory (Azure AD) 透過 AZURE RBAC(角色型存取控制) , 來授與受保護資源的存取權限
- Azure 儲存體會定義一組 Azure 內建角色，其中包含用來存取 blob 或佇列資料的一般許可權集
- 適用於 Blob 的 Azure 角色
 - ◆ 儲存體 Blob 資料擁有者：
用來設定擁有權和管理 Azure Data Lake Storage Gen2 的 POSIX 存取控制
 - ◆ 儲存體 Blob 資料參與者：
用來授與 Blob 儲存體資源的讀取/寫入/刪除權限
 - ◆ 儲存體 Blob 資料讀者：
用來授與 Blob 儲存體資源的唯讀權限

授與存取權 -1

- 在 [資源群組] 清單中，選擇資源群組”**autotrain小組編號**”
- 按一下 [存取控制 (IAM)]
 - ◆ Identity and access management
- 按一下 [角色指派] 索引標籤，以查看目前的角色指派清單
- 按一下 [新增] > [新增角色指派]，以開啟 [新增角色指派] 窗格
- 在 [角色] 下拉式清單中選取 [儲存體 Blob 資料參與者]
- 在 [選取] 文字框中，輸入選取您之前註冊的應用程式(例如 MyDemoApp)
- 按一下 [儲存] 以指派角色

授與存取權 -2

新增角色指派

X

角色 ①

儲存體 Blob 資料參與者 ①

存取權指派對象為 ①

使用者、群組或服務主體

選取 ①

My

找不到使用者、群組或服務主體。

選取的成員:



MyDemoApp

移除

儲存

捨棄

授與存取權 -3

首頁 > mslearn-test

mslearn-test | 存取控制 (IAM)

資源群組

搜尋 (Ctrl+ /) 新增 下載角色指派 編輯資料行 重新整理 移除 想提供意見反應嗎?

檢查存取權 角色指派 角色 角色 (預覽) 否定性指派 傳統管理員

此訂用帳戶的角色指派數 6 2000

依姓名或電子郵件搜尋 類型：全部 角色：全部 範圍：所有範圍 分組依據：角色

2 個項目 (1 服務主體, 1 受控識別)

名稱	類型	角色	範圍
MyDemoApp	應用程式	儲存體 Blob 資料參與者	此資源

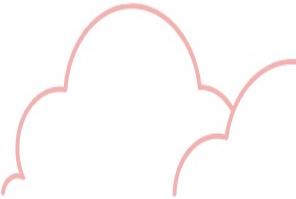
參與者

名稱	類型	角色	範圍
mls-demo	應用程式	參與者	此資源

Azure Machine Learning Studio



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具



什麼是 Azure 機器學習

- Azure 機器學習是一項雲端服務
- 用來做資料建模，測試和部署預測性分析的解決方案
- Machine Learning Studio 是一種拖放工具，可用來構建 ML 模型，並將其作為 Web 服務發布，也可以輕鬆地由 MS Excel 等客製化的應用程式使用
- Azure MLS 是一個互動式工作區，可以輕鬆地使用它來開發 ML 模組
- 可以將資料集，分析模組拖放到互動式桌面中，將它們連接起來以形成一個實驗
- 您可以隨時編輯該實驗，保存並再次運行
- 也可以將訓練實驗轉換為預測性實驗，然後發佈為 Web 服務，以便其他人可以訪問您的模型

機器學習實作

- 一般而言，機器學習實作通常會有一些進行步驟
- 準備資料集
- 分析資料集
- 使用資料集來訓練模型
- 評估模型
- 檢驗模型的精確度

建立ML工作區



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

建立工作區 -1

- 使用您 Azure 訂閱的認證來登入 Azure 入口網站
- 在 Azure 入口網站的左上側，選取 [+建立資源]
- 使用搜尋列尋找“Machine”，選取 [Machine Learning] 或[機器學習]
- 在 [Machine Learning] 窗格中選取 [建立] 來開始操作
- 提供下列資訊來設定新的工作區
 - ◆ 工作區名稱，輸入可識別您工作區的唯一名稱，例如**“automi員工編號”**
 - ◆ 訂用帳戶，選取課程使用的訂用帳戶**azuretrain01**
 - ◆ 資源群組，選取課程使用的資源群組**“autotrain小組編號”**
 - ◆ 儲存體帳戶，選取課程使用的**“auoautotrain小組編號”**
 - ◆ 金鑰保存庫，選取課程使用的**“auoautotrain小組編號”**
 - ◆ Application Insights，選取課程使用的**“auoautotrain小組編號”**
 - ◆ 區域請依照小組名單選擇**東南亞/東亞**

建立工作區 -2

Microsoft Azure 搜尋資源、服務及文件 (G+)

首頁 > autotrain10 > 建立資源 > Marketplace > 機器學習 >

機器學習 ...

建立機器學習工作區

選取用以管理部署資源及成本的訂用帳戶。使用像資料夾這樣的資源群組來安排及管理您的所有資源。

訂用帳戶 * ⓘ

azuretrain01

資源群組 * ⓘ

autotrain10

新建

工作區詳細資料

指定工作區的名稱和區域。

工作區名稱 * ⓘ

Demo

區域 * ⓘ

東南亞

儲存體帳戶 * ⓘ

auoautotrain10

新建

金鑰保存庫 * ⓘ

auoautotrain10

Application Insights * ⓘ

auoautotrain10

新建

檢閱 + 建立 < 上一步 下一步: 網路

建立工作區 -3

- 當您完成工作區的設定時，請選取 [檢閱 + 建立]
- 檢查設定，並進行任何額外的變更或修正。當您對設定感到滿意之後，請選取 [建立]
- 程序完成後，會出現部署成功訊息，要檢視新的工作區，選取 [前往資源]

連接工作區

● 利用“工作室 Web URL”連線到工作室

es > X

[下載 config.json](#) [刪除](#)

程式集

資源群組 : mslearn-test

位置 : 東南亞

訂用帳戶 : timyang

訂用帳戶識別碼 : f4c08b1d-5835-4123-804a-6d01bb0446ab

工作室 Web URL : <https://ml.azure.com/?tid=f872b66e-ba23-4c79-9521-ec54328fcbdd&wsid=/subscriptions/f4c08b1d-5835-4123-804a-6d01bb0446ab/resourcegroups/mslearn-test/workspaces/ml-demo>

JSON 檢視 [複製]

金鑰保存庫 : mldemo8819786522

Application Insights : mldemo4263821986

管理您的機器學習生命週期

使用 Azure Machine Learning Studio (傳統版) 來建置、訓練、評估及部署機器學習模型。深入了解 [\[深入了解\]](#)

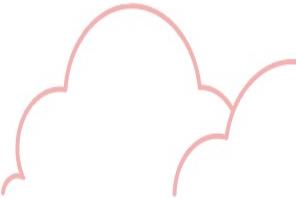
[啟動工作室](#)

-9521-ec54328fcbdd&wsid=/subscript...

連接資料



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具



Azure ML 資料

- Azure Machine Learning 的資料有兩類
- 資料存放區
 - ◆ 將您的連線資訊(例如您的訂用帳戶識別碼和 token 授權)儲存在與工作區相關聯的 Key Vault，以安全地連線到您在 Azure 上的儲存體服務
- 資料集
 - ◆ 藉由建立資料集，您可以建立資料來源位置的參考，以及其中繼資料的複本
- 資料集的使用
 - ◆ 在模型定型期間存取資料
 - ◆ 共用資料並與其他使用者共同作業。
 - ◆ 利用開放原始碼程式庫(例如 pandas) 來探索資料

建立資料存放區 -1

- 登入 Azure Machine Learning Studio
- 在左窗格中，選取 [管理] 底下的 [資料存放區]
- 選取 [+ 新增資料存放區]，在同一個ML Studio內，資料存放區的命名不能重複
- 完成表單，以建立並註冊新的資料存放區。此表單會根據您選取的 Azure 儲存體類型和驗證類型，以智慧方式自行更新

建立資料存放區 -2



- 資料存放區名稱:
sample_data_員工編號
- 資料存放區類型:
Azure Data Lake Storage Gen2
- 訂用帳戶ID: (使用預設)
azuretrain01
- 存放區名稱:
dllakedemo (Training_RG)
- Azure Data Lake Gen2 檔案系統名稱:
designertool

新增資料存放區

資料存放區名稱 *

sample_data_demo

資料存放區類型 *

Azure Data Lake Storage Gen2

選取帳戶的方法

來自 Azure 訂用帳戶

手動輸入

訂用帳戶 ID *

azuretrain01 (a278626c-1d76-422d-a5b2-4c2cf191a0ac)

存放區名稱 *

dllakedemo (Training_RG)

Azure Data Lake Gen2 檔案系統名稱 *

designertool

建立資料存放區 -3

- 用戶端識別碼：之前註冊的應用程式識別碼

The screenshot shows the Microsoft Azure portal interface. At the top, there's a navigation bar with 'Microsoft Azure' and a search bar. Below it, the breadcrumb navigation shows 'Home > Contoso AD (dev) | App registrations >'. The main content area displays a single application registration named 'Contoso App 1'. On the left, there's a sidebar with 'Overview' selected, along with 'Quickstart' and 'Integration assistant (preview)'. The 'Overview' section contains the following details:

Display name	Supported account types
Contoso App 1	My organization only
Application (client) ID	Redirect URIs
11111111-1111-1111-1111-111111111111	Add a Redirect URI

- 用戶端密碼：之前註冊的應用程式設定的密碼

The screenshot shows a table titled '新增用戶端密碼' (Add User Secret). The table has columns for '描述' (Description), '到期' (Expires), '值' (Value), and 'ID'. There is one row of data:

描述	到期	值	ID
Password uploaded on Sat Apr 03 2021	3/4/2022	~_n*****	902ba91b-c42e-4757-acfd-0b308ccfe28f

- 選取“建立”

上課示範用密碼

- Application/Client ID:

cf446eaf-3099-4179-94e3-20b51bc4a78f

- Object ID:

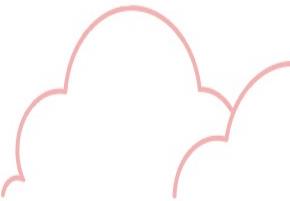
cece5a9f-8e44-4398-82af-4b39d6548af5

- Directory/Tenant ID:

e7c24d00-7479-4343-8247-0204699693e8

- Secrets Value:

SBoFq1TfYG2~Z1E_4N3i_ialkLs0bgN~ZU



建立資料集 -1

- 在左窗格中，選取 [管理] 底下的 [資料存放區]
- 選取要建立資料集的資料存放區
- 選取 [建立資料集]
- 基本資訊
 - ◆名稱: ames_train_demo
 - ◆資料集類型: 表格式
 - ◆描述: Demo Dataset Ames Training Data
- 選取 [下一個]
- 資料存放區選取
 - ◆路徑: AmesHousingNew.csv
- 選取 [下一個]
- 出現“設定與預覽”頁面

建立資料集 -2



- “設定與預覽”頁面的設定是自動偵測而來。請驗證選取項目是否正確
 - ◆ 正確的設定“分隔符號”、“資料行標頭”
- 選取 [下一個]

從資料存放區建立資料集 X

設定與預覽
這些設定是自動偵測而來。請驗證選取項目是否正確，否則請予以更新。

檔案格式
分隔符號 逗號 範例 Field1,Field2,Field3

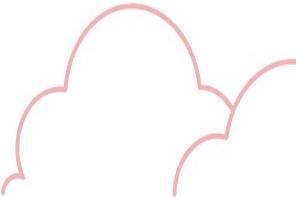
編碼
UTF-8

資料行標頭
使用第一個檔案的標頭

略過資料列
無

#	識..	Order	MS Zoning	Lot Frontage	Lot Area	Utilities	Ov
1	1		RL	141	31770	AllPub	6
2	2		RH	80	11622	AllPub	5
3	3		RL	81	14267	AllPub	6
4	4		RL	93	11160	AllPub	7
5	5		RL	74	13830	AllPub	5
6	6		RL	78	9978	AllPub	6

上一步 下一個 取消



建立資料集 -3

- 進入“結構描述”頁面
 - ◆ 資料行類型會根據資料的前 200 個資料列自動偵測，可以進行必要的調整
 - ◆ 與指定資料行類型不一致的值將無法轉換，而且會填入 Null 或以錯誤值取代
- 選取 [下一個]
- 確認詳細資料後，選取“建立”來產生資料集
- 可以利用 Azure Machine Learning [首頁] > [資料集] 來檢視剛建立的資料集
 - ◆ 選擇要查閱的資料集
 - ◆ 選取“瀏覽”，可以檢視資料內容

建立資料集 -4

Microsoft Azure Machine Learning

首頁 > Datasets > ames_train_demo

ames_train_demo 版本 1 (最新) ▾

詳細資料 取用 瀏覽 模型

重新整理 產生設定檔 取消註冊 新增版本

預覽 設定檔

資料行數目: 21 列數: 50 (的 2930)

#	識..	Order	MS Zoning	Lot Frontage	Lot Area	Utilities	Overall Qual	Overall Cond	Roof Matl	Total Bsmt SF	Gr Liv J
1	1		RL	141	31770	AllPub	6	5	CompShg	1080	1656
2	2		RH	80	11622	AllPub	5	6	CompShg	882	896
3	3		RL	81	14267	AllPub	6	6	CompShg	1329	1329
4	4		RL	93	11160	AllPub	7	5	CompShg	2110	2110
5	5		RL	74	13830	AllPub	5	5	CompShg	928	1629
6	6		RL	78	9978	AllPub	6	6	CompShg	926	1604
7	7		RL	41	4920	AllPub	8	5	CompShg	1338	1338
8	8		RL	43	5005	AllPub	8	5	CompShg	1280	1280
9	9		RL	39	5389	AllPub	8	5	CompShg	1595	1616
10	10		RL	60	7500	AllPub	7	5	CompShg	994	1804
11	11		RL	75	10000	AllPub	6	5	CompShg	763	1655
12	12		RL	Null	7980	AllPub	6	7	CompShg	1168	1187
13	13		RL	63	8402	AllPub	6	5	CompShg	789	1465
14	14		RL	85	10176	AllPub	7	5	CompShg	1300	1341
15	15		RL	Null	6820	AllPub	8	5	CompShg	1488	1502
16	16		RL	47	53504	AllPub	8	5	CompShg	1650	3279

使用設計工具 建立機器學習模型



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

建立管線

- 登入 <https://ml.azure.com>，並選取您要使用的工作區
- 選取 [設計工具]
- 選取 [易於使用的預建模組]
- 在畫布頂端選取預設管線名稱 Pipeline-Created-on，可以考慮重新命名為“Ames Housing Price”
- 在管線名稱旁邊，選取畫布頂端的“齒輪”圖示，可開啟 [設定] 窗格
- 在畫布右側的 [設定] 窗格中，選取 [選取計算目標]
 - ◆ 選取既有的或是建立新的計算目標，選取“儲存”
 - ◆ 建立計算資源大約需要五分鐘

設定計算目標

設定計算目標

×

- 選取現有項目 建立新的

計算目標組態

建議您使用下方預先定義的組態，快速設定訓練計算目標。

組態	vCPU	RAM	儲存體	節點計數
Predefined	2	7GB	8GB	2

不符合您的需求嗎? 請前往 [\[計算\] -> \[訓練叢集\]](#)，以建立自訂計算目標。

新增計算目標名稱 ②



MLCompute

儲存

取消

準備資料 -1

- 汇入資料
 - ◆ 選取資料集“ames_train_demo”，並將其拖曳到畫布上
- 將資料視覺化
 - ◆ 以滑鼠右鍵按一下“ames_train_demo”，然後選取 [視覺化 > 資料集輸出]
 - ◆ 選取資料視窗中的不同資料行，以檢視各個資料行的相關資訊
 - ◆ 每個資料列分別代表一筆交易，而與每筆交易相關聯的變數會顯示為資料行。此資料集中有 2930 個資料列和 21 個資料行
- 準備資料
 - ◆ 資料集在分析之前通常需進行一些前置處理，才能讓模型正確地分析資料

準備資料 -2

● 資料視覺化

ames_train_demo 結果視覺效果

X

資料列 ⑦ 資料行 ⑦
2,930 21

1	TotRms AbvGrd	Garage Cars	Garage Area	Wood Deck SF	Open Porch SF	Mo Sold	Yr Sold	SalePrice
7	2	528	210	62	5	2010	215000	
5	1	730	140	0	6	2010	105000	
6	1	312	393	36	6	2010	172000	
8	2	522	0	0	4	2010	244000	
6	2	482	212	34	3	2010	189900	
7	2	470	360	36	6	2010	195500	
6	2	582	0	0	4	2010	213500	
5	2	506	0	82	1	2010	191500	
5	2	608	237	152	3	2010	236500	
7	2	442	140	60	6	2010	189000	

SalePrice

統計資料

平均數	180796.0601
中位數	160355.2174
最小值	12789
最大值	755000
標準差	79886.6924
唯一值	0
遺漏值	0
功能類型	Numeric Feature

視覺效果



關閉

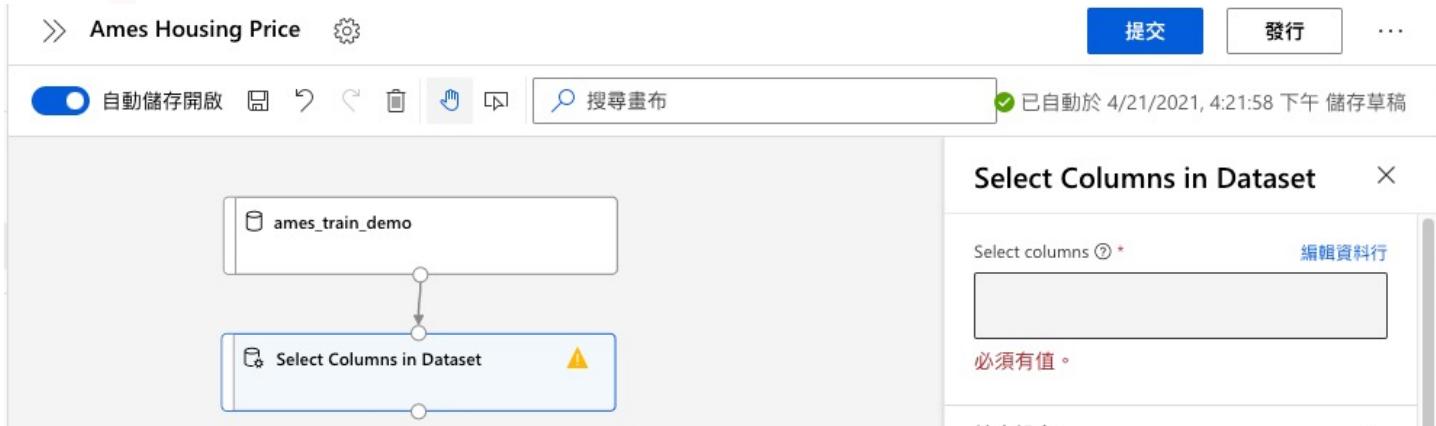
準備資料 -3



● 移除資料行

- ◆當您定型模型時，您必須對遺漏的資料採取某些動作
- ◆在畫布左側的模組選擇區中，展開 [Data Transformation] 區段，然後尋找 [**Select Columns in Dataset**] 模組
- ◆將“Select Columns in Dataset”模組拖曳到畫布上。將模組放在資料集模組下
- ◆將“ames_train_demo”資料集連線至“Select Columns in Dataset”模組
- ◆從資料集的輸出連接埠 (即畫布上位於資料集底部的小圓圈) 拖曳到“選取資料集中的資料行”的輸入連接埠 (即模組頂端的小圓圈)
- ◆選取“Select Columns in Dataset”模組

準備資料 -4



- ◆ 在畫布右側的 [模組詳細資料] 窗格中，選取 [編輯資料行]
- ◆ 展開 [包含] 旁邊的 [資料行名稱] 下拉式清單，然後選取 [所有資料行]
- ◆ 選取 + 以新增規則
- ◆ 從下拉式功能表中，選取 [排除] 和 [資料行名稱]。
- ◆ 在文字方塊中輸入“Order, MS Zoning, Utilities, RoofMatl”
- ◆ 在右下方選取 [儲存] 按鈕，以關閉資料行選取器

準備資料 -5

Select columns

選取欄

具有規則

依名稱

×

允許選取範圍中有重複項並保留資料行順序

包含

所有資料行

+

排除

資料行名稱

+

-

Order X MS Zoning X

Utilities X Roof Matl X

儲存

取消

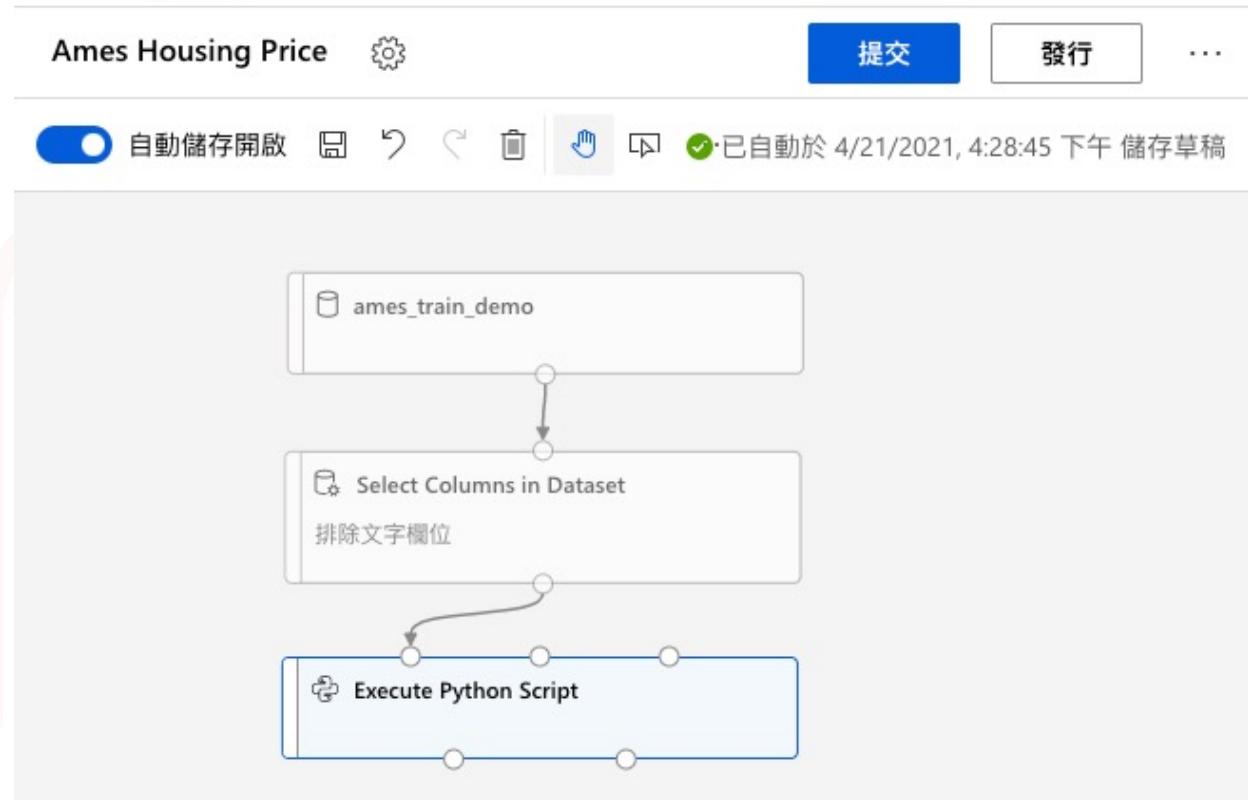
- ◆ 選取“Select Columns in Dataset”模組
- ◆ 在畫布右側的 [模組詳細資料] 窗格中，選取 [註解] 文字方塊，輸入“排除文字欄位”
- ◆ 圖形上會出現註解，以協助您組織管線

使用Python指令碼 -1

- 在設計工具選擇區中，尋找 [Execute Python Script] 模組，您可以在 [Python Language] 區段中找到該模組
- 將模組拖放到管線畫布上
- 將“Select Columns in Dataset”模組的輸出連接埠連線至“Execute Python Script”模組左上角的輸入連接埠(dataset1)
- 設計工具會將輸入公開為進入點指令碼的參數
- 記下您使用的輸入連接埠。 設計工具會將左側的輸入連接埠指派給變數 dataset1，並將中間的輸入連接埠指派給 dataset2

使用Python指令碼 -2

- 設計工具會提供初始的進入點指令碼，讓您編輯和輸入自己的 Python 程式碼



使用Python指令碼 -3

- 選取“Execute Python Script”模組
- 在顯示於畫布右側的窗格中，選取 [Python Script] 文字方塊下方的 [編輯程式碼] 按鈕
- 在“Python Script”窗格中，找到程式碼# Execution logic goes here，輸入以下程式碼，並注意縮排
 - ◆# 新增一個欄位
```python  
dataframe1['TotBath'] = dataframe1['Full Bath'] +  
dataframe1['Half Bath']
  - ◆# 移除兩個欄位  
```python  
dataframe1 = dataframe1.drop(columns=['Full Bath', 'Half Bath'])

使用Python指令碼 -4

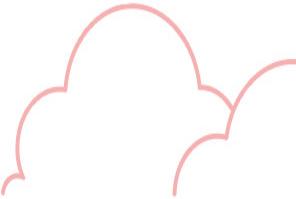
● 選取 [儲存]，將程式碼儲存起來

Python script

```
9  # If the input port is not connected, the corresponding
10 # dataframe argument will be None.
11 # Param<dataframe1>: a pandas.DataFrame
12 # Param<dataframe2>: a pandas.DataFrame
13 def azureml_main(dataframe1 = None, dataframe2 = None):
14
15     # Execution logic goes here
16     dataframe1['TotBath'] = dataframe1['Full Bath'] + dataframe1['Half Bath']
17     dataframe1 = dataframe1.drop(columns=['Full Bath', 'Half Bath'])
18
19     # If a zip file is connected to the third input port,
20     # it is unzipped under "./Script Bundle". This directory is added
21     # to sys.path. Therefore, if your zip file contains a Python file
22     # mymodule.py you can import it using:
23     # import mymodule
24
25     # Return value must be of a sequence of pandas.DataFrame
26     # E.g.
27     #   - Single return value: return dataframe1,
28     #   - Two return values: return dataframe1, dataframe2
29     return dataframe1,
```

儲存

取消

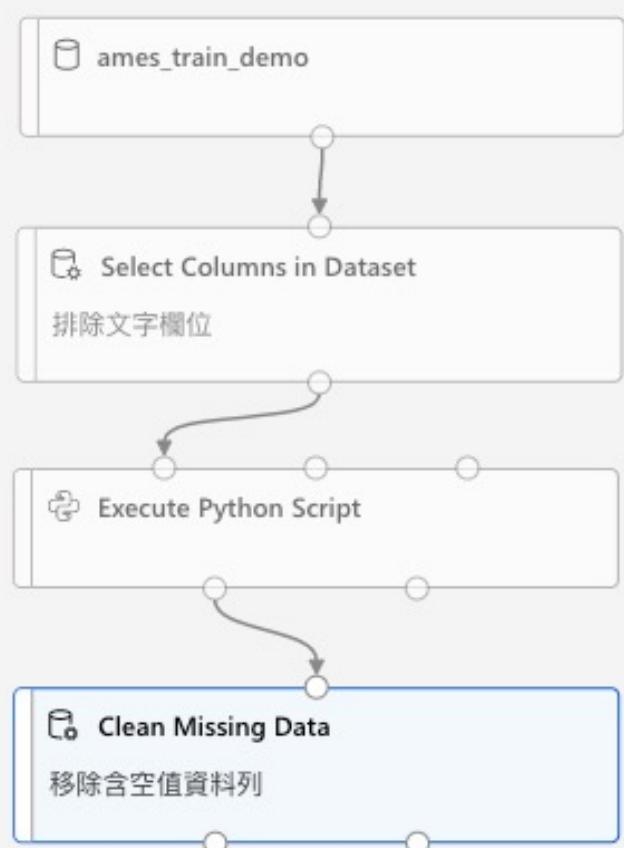


處理空值 -1

- 在畫布左側的模組選擇區中，展開 [Data Transformation] 區段，然後尋找 [Clean Missing Data] 模組
- 將“Clean Missing Data”模組拖曳至管線畫布上，將其連線至“Execute Python Script”模組的左下角輸出連接埠(Result dataset)
- 選取 [Clean Missing Data] 模組
- 在畫布右側的 [模組詳細資料] 窗格中，選取 [編輯資料行]
- 在顯示的 [Columns to be cleaned] 視窗中，展開 [包含] 旁的下拉式功能表。選取 [所有資料行]
- 選取 [儲存]

處理空值 -2

- 在畫布右邊的 [模組詳細資料] 窗格中，選取 [Cleaning mode] 底下的 [Remove entire row]
- 在畫布右側的 [模組詳細資料] 窗格中，選取 [註解] 方塊，然後輸入“移除含空值資料列”
- 您的管線此時應會顯示如右：

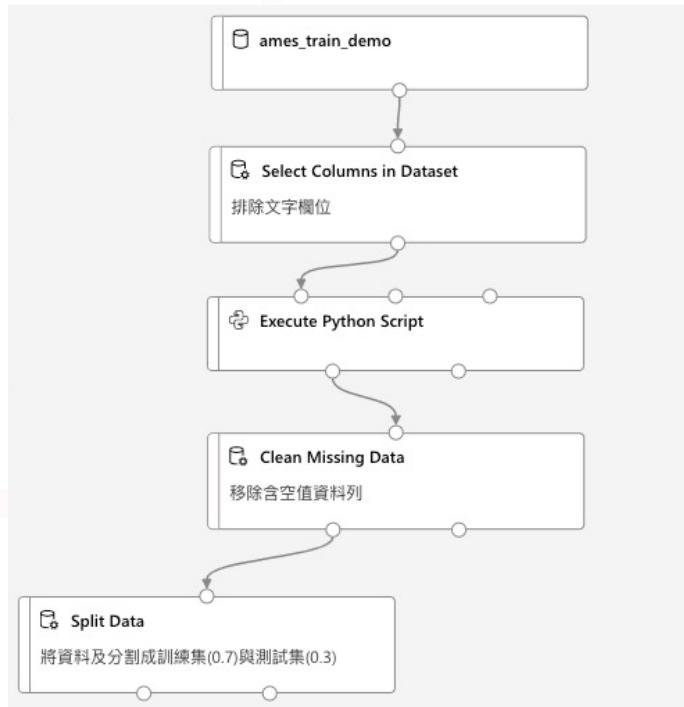


分割資料 -1

- 分割資料是機器學習服務中常見的工作。您會將資料分割成兩個不同的資料集。一個資料集會定型模型，另一個則會測試模型的執行效果
- 在模組選擇區中展開 [Data Transformation] 區段，然後尋找 [Split Data] 模組
- 將 [Split Data] 模組拖曳到管線畫布上
- 將“Clean Missing Data”模組左側的連接埠連線至“Split Data”模組
- 選取“Split Data”模組
- 在畫布右側的 [模組詳細資料] 窗格中，將 [Fraction of rows in the first output dataset] 設定為 0.7

分割資料 -2

- 此選項會分割 70% 的資料來定型模型，而 30% 供測試之用。70% 的資料集將透過左側輸出連接埠來存取，其餘資料可透過右側輸出連接埠取得
- 在畫布右側的 [模組詳細資料] 窗格中，選取 [註解] 方塊，然後輸入“將資料集分割成訓練集 (0.7) 與測試集 (0.3)”



訓練模型 -1

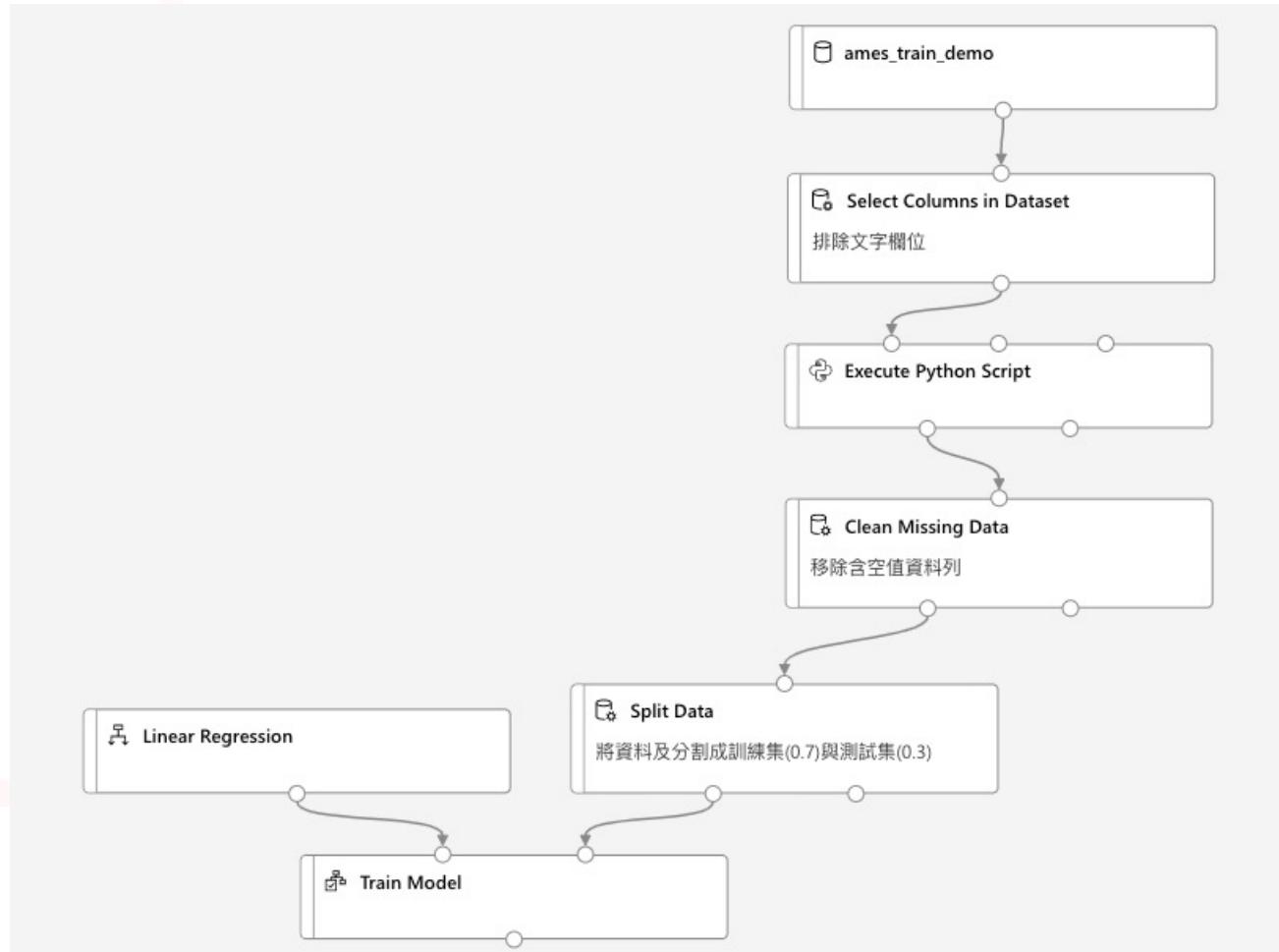
- 藉由提供一個包含價格的資料集來訓練模型
- 此演算法會建立一個模型，用以說明訓練資料所呈現的特徵值與價格之間的關聯性
- 在模組選擇區中，展開 [Machine Learning Algorithms]
 - ◆此選項會顯示數個可用來初始化學習演算法的模組類別
- 選取 [Regression] > [Linear Regression]，然後將其拖曳到管線畫布上
- 在模組選擇區中，展開 [Model training] 區段，然後將 [Train Model] 模組拖曳至畫布
- 將 [Linear Regression] 模組的輸出連接到 [Train Model] 模組的左側輸入

訓練模型 -2

- 將“Split Data”模組的訓練資料輸出 (左側連接埠) 連接到“Train Model”模組的右側輸入
- 選取“Train Model”模組
- 在畫布右側的 [模組詳細資料] 窗格中，選取 [編輯資料行] 選取器
- 在 [Label column] 對話方塊中，展開下拉式功能表，然後選取 [資料行名稱]
- 在文字方塊中輸入“SalePrice”，以指定您的模型要預測的值
- 選擇 [儲存]

訓練模型 -3

- 您的管線應會顯示如下：



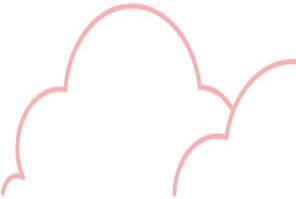
模組評分與評估 -1

- 使用 70% 的資料來定型模型後，您即可將該模型用來為其他 30% 的資料評分，以了解模型的運作是否理想
- 在搜尋方塊中輸入“Score Model”以尋找 [Score Model] 模組。將此模組拖曳到管線畫布上
- 將“Train Model”模組的輸出連線至“Score Model”的左側輸入連接埠。將“Split Data”模組的測試資料輸出 (右側連接埠) 連線至“Score Model”的右側輸入連接埠
- 使用 [Evaluate Model] 模組，評估您的模型在測試資料集下的評分
- 在搜尋方塊中輸入“Evaluate”，以尋找“Evaluate Model”模組，將此模組拖曳到管線畫布上

模組評分與評估 -2

- 將“Score Model”模組的輸出連線至“Evaluate Model”的左側輸入
- 最終的管線應會顯示如右





提交管線 -1

- 現在管線已全部設定完成，您可以提交管線執行來定型您的機器學習模型
- 您可以在任何時間點提交有效的管線執行，以便在開發期間用來檢閱管線的變更
- 在畫布頂端，選取 [提交]
- 在 [設定管線回合] 對話方塊中，選取 [建立新的]
- 針對 [新增實驗名稱]，輸入“Test-AmesHousing”
- 選取 [提交]
- 您可以在畫布右上方檢視執行狀態和詳細資料
- 如果這是第一次執行，您的管線可能需要 20 分鐘的時間才能完成執行

提交管線 -2

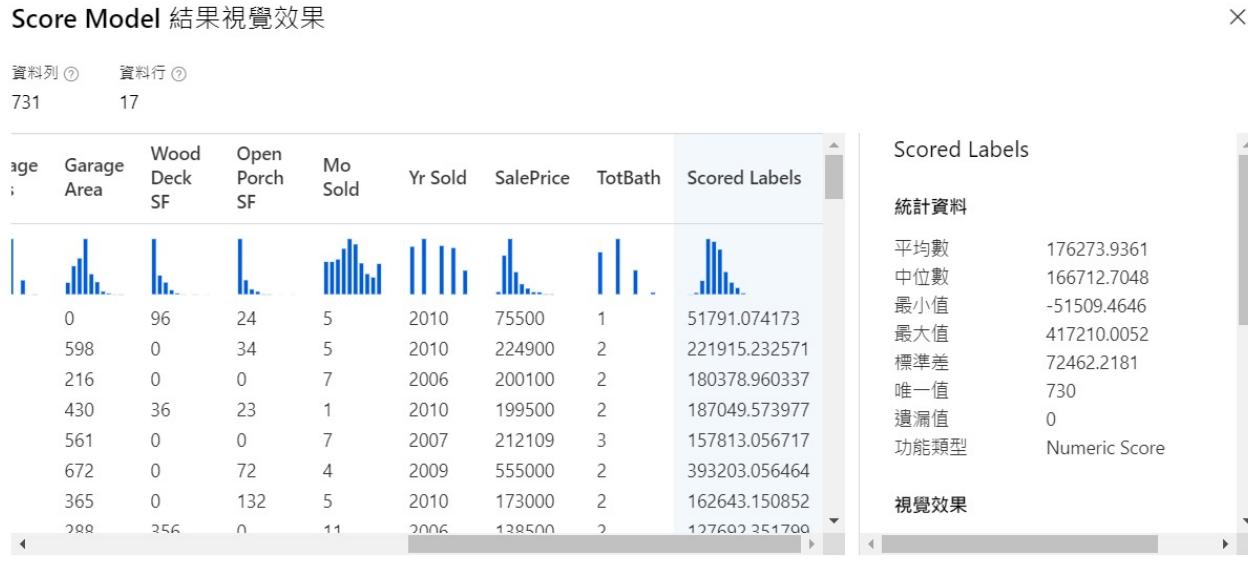
- 預設計算設定的最小節點大小為 0，這表示設計工具必須在閒置之後配置資源
- 重複的管線執行花費較少的時間，因為已經配置計算資源
- 選取“檢視執行概觀”查看執行過程

The screenshot shows a 'Execution Overview' window from the Cloudi platform. At the top, there's a status bar with a blue circle icon indicating '正在執行' (Running) and the text '檢視執行概觀' (View Execution Overview). The main area is titled '執行概觀' (Execution Overview) and contains tabs for '屬性' (Properties), '管線參數' (Pipeline Parameters), '輸出 + 記錄檔' (Outputs + Log Files), '計量' (Metrics), '子回合' (Subcycles), '影像' (Images), and more. The '屬性' tab is selected. Below it, the window displays various details about the execution:

| 狀態 | 正在執行 |
|---------|--------------------------------------|
| 提交時間 | 2021年4月21日 17:32 |
| 結束時間 | - |
| 持續時間 | 2 分 37 秒 |
| 描述 | Demo Ames Housing Price |
| 執行 ID | 49dda1ec-4426-408c-80ce-16cc18b84882 |
| 實驗 | Test-AmesHousing |
| 查看所有屬性: | 原始 JSON |

檢視評分

- 執行完成後，您可以檢視管線執行的結果
- 首先，請查看迴歸模型產生的預測
- 以滑鼠右鍵按一下 [Score Model] 模組，然後選取 [視覺化] > [Scored dataset] 以查看其輸出
- 您可以在這裡看到測試資料中的預測價格和實際價格



評估模型 -1

- 以滑鼠右鍵按一下 [Evaluate Model] 模組，然後選取 [視覺化] > [Evaluation results] 以查看其輸出
- 您的模型會顯示下列統計資料
 - ◆ 平均絕對誤差 (MAE)：絕對誤差的平均值。誤差是指預測值與實際值之間的差異
 - ◆ 均方根誤差 (RMSE)：對測試資料集所做之預測的平方誤差的評分根平均值
 - ◆ 相對絕對誤差：相對於實際值與所有實際值之平均值之間的絕對差異的絕對誤差平均值
 - ◆ 相對平方誤差：相對於實際值與所有實際值之平均值之間的平方差異的平方誤差平均值
 - ◆ 決定係數：也稱為 R 平方值，這是一個統計計量，可指出模型對於資料的適用程度

評估模型 -2

- 針對每個誤差統計資料，越小越好。值越小，表示預測越接近實際值
- 就決定係數而言，其值愈接近一 (1.0)，預測就愈精準

Evaluate Model 結果視覺效果

資料列 ② 資料行 ②

1 5

| Squared_Error | Relative_Squared_Error | Relative_Absolute_Error | Coefficient_of_Determination |
|---------------|------------------------|-------------------------|------------------------------|
| 4 | 0.160389 | 0.391116 | 0.839611 |

Coefficient_of_Determination

統計資料

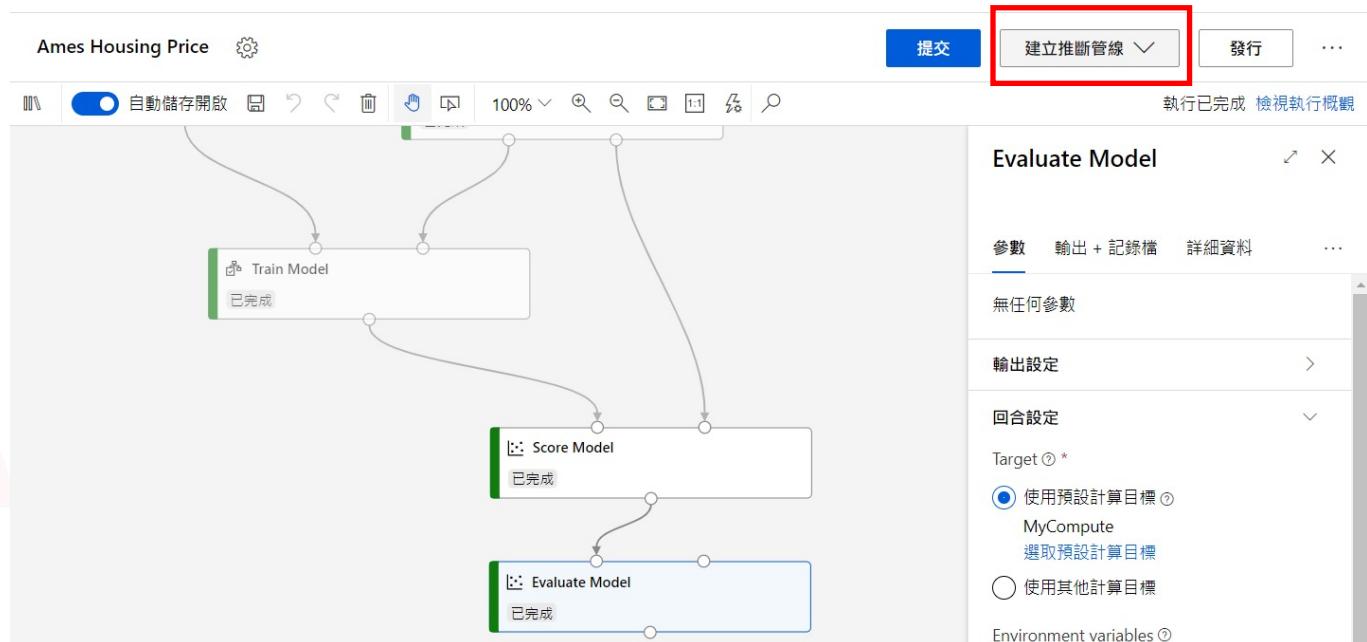
| | |
|------|-----------------|
| 平均數 | 0.8396 |
| 中位數 | 0.8396 |
| 最小值 | 0.8396 |
| 最大值 | 0.8396 |
| 標準差 | - |
| 唯一值 | 1 |
| 遺漏值 | 0 |
| 功能類型 | Numeric Feature |

視覺效果

關閉

建立即時推斷管線 -1

- 若要部署您的管線，您必須先將訓練管線轉換成即時推斷管線
- 此程序會移除訓練模組，並新增 Web 服務輸入和輸出以處理要求
- 在管線畫布上方，選取 [建立推斷管線] > [即時推斷管線]



建立即時推斷管線 -2

- 選取 [建立推斷管線] 後，畫布會發生一些變化
 - ◆ 定型的模型會儲存為模組選擇區中的“資料集”模組，您可以在 [我的資料集] 下方找到該項目
 - ◆ 訓練模組 (例如“訓練模型”和“分割資料”) 會被移除
 - ◆ 儲存的定型模型會加回管線中
 - ◆ 會新增“Web 服務輸入”和“Web 服務輸出”模組，這些模組會顯示使用者資料將在何處進入管線，以及資料會在何處傳回
- 根據預設，“Web 服務輸入”會預期與用來建立預測管線的訓練資料相同的資料結構描述。在此情況下，價格會包含在結構描述中。不過，在預測期間並不會以價格作為因素

建立即時推斷管線 -3

首頁 > 設計工具 > 正在撰寫

訓練管線 即時推斷管線

Ames Housing Price-real time inference

提交 部署 ...

依名稱、標籤和描述搜尋

自動儲存開啟 搜尋畫布

總計 98 個資產

Datasets (3)

- TD-Ames_Housing_Price-Clean_Missing_Data...
This is a dataset promoted by inference graph generation automatically on 04/18/2021 12:09:12...
4/18/2021
- MD-Ames_Housing_Price-Train_Model-Train...
This is a dataset promoted by inference graph generation automatically on 04/18/2021 12:09:11...
4/18/2021
- ames_train
Ames Training Data
4/18/2021

Sample datasets (16)

Data Input and Output (3)

Data Transformation (19)

Feature Selection (2)

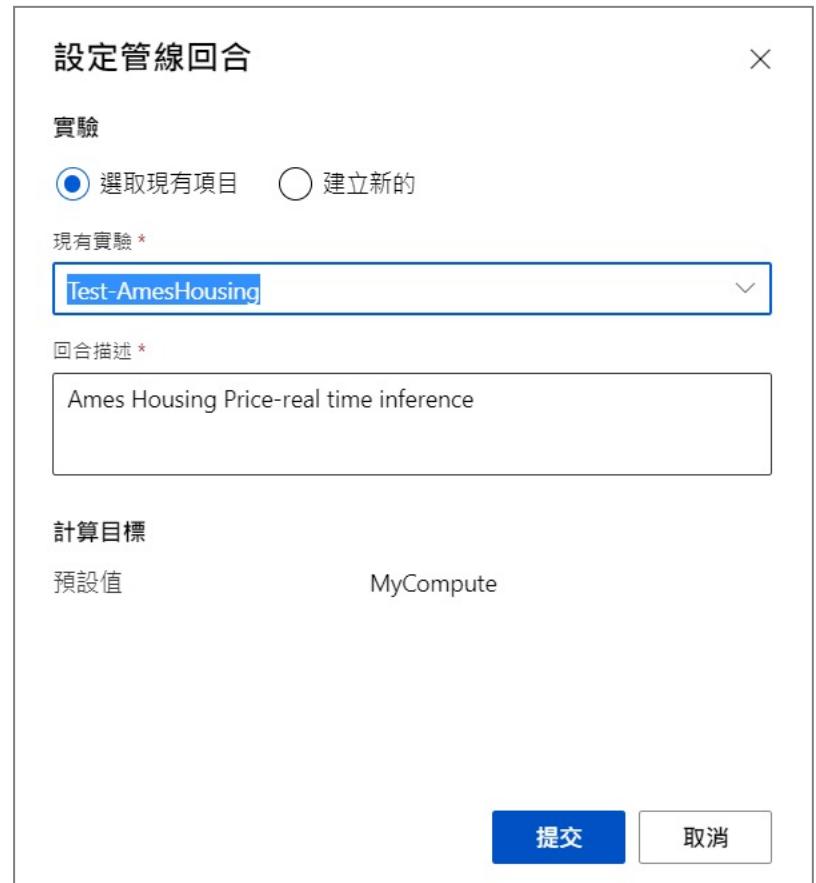
Statistical Functions (1)

Machine Learning Algorithms (10)

```
graph TD
    ames_train[ames_train] --> SC[Select Columns in Dataset]
    WebServiceInput[Web Service Input] --> SC
    SC --> TD[TD-Ames_Housing_Price-Clean_Missing_Data...]
    SC --> MD[MD-Ames_Housing_Price-Train_Model-Train...]
    TD --> EP[Execute Python Script]
    MD --> AT[Apply Transformation]
    EP --> SM[Score Model]
    AT --> SM
    SM --> WSO[Web Service Output]
    AT --> EM[Evaluate Model]
    EM --> WSO
```

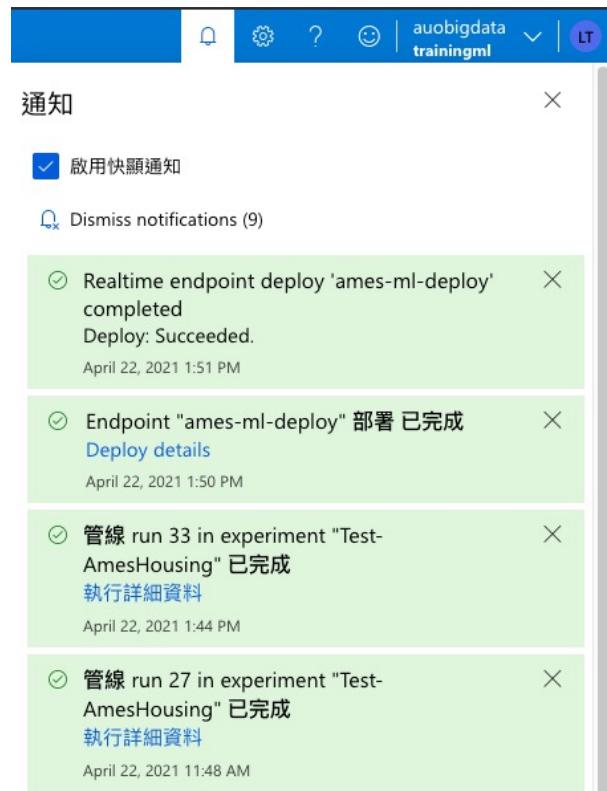
建立即時推斷管線 -4

- 選取 [提交]，並使用您在第一部分中使用的相同計算目標和實驗(例如 Test-AmesHousing)
- 選取 [提交]
- 如果這是第一次執行，您的管線可能需要 20 分鐘的時間才能完成執行
- 預設計算設定的最小節點大小為 0，這表示設計工具必須在閒置之後配置資源



部署模型

- 選取 Azure Machine Learning 右上方的 [部署]
- 進入“設定即時端點”的畫面
 - ◆ 選取“部署新的即時端點”
 - ◆ 名稱: **ames員工編號**
 - ◆ 描述: Ames Housing Evaluate
 - ◆ 計算類型: Azure 容器執行個體
- 選取 [部署]
- 完成部署之後，畫布上方會出現成功通知，這可能需要幾分鐘的時間



測試即時端點 -1

- 部署完成後，您即可移至 [端點] 頁面，以檢視您的即時端點
- 在 [端點] 頁面上，選取您部署的端點

首頁 > 端點

端點

即時端點 管線端點

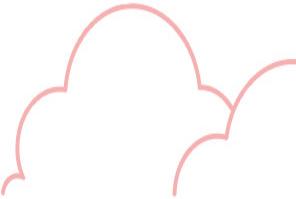
刪除 編輯資料行 重新整理 重設為預設檢視

+ 新增筛选

顯示第 1-1 個 端點

頁面大小: 25

| 名稱 | 描述 | 建立日期 | 建立者 | 更新日期 | 計算類型 | 計算目標 | 標籤 |
|----------------|-----------------------|-----------------------|----------|-----------------------|------|------|-------------------------------|
| ames-ml-deploy | Ames Housing Evaluate | Apr 12, 2021 10:48 PM | Tim Yang | Apr 12, 2021 10:48 PM | ACI | | CreatedByAMLStudio : true ... |



測試即時端點 -2

- 在 [詳細資料] 索引標籤中，您可以看到詳細資訊，例如 REST URI、Swagger 定義、狀態及標記
- 在 [取用] 索引標籤中，您可以找到範例耗用量程式碼、安全性金鑰，以及設定驗證方法
- 在 [部署記錄] 索引標籤中，您可以找到即時端點的詳細部署記錄
- 若要測試您的端點，請移至 [測試] 索引標籤
- 可以從這裡輸入測試資料，然後選取 [測試] 以驗證端點的輸出

測試即時端點 -3

首頁 > 端點 > ames-ml-deploy

ames-ml-deploy

詳細資料 測試 取用 部署記錄

輸入要測試即時端點的資料 測試

WebServiceInput0

| | |
|--------------|--------|
| Order | 1 |
| MS Zoning | RL |
| Lot Frontage | 141 |
| Lot Area | 31770 |
| Utilities | AllPub |
| Overall Qual | 6 |

測試結果 已剖析 原始

WebServiceOutput0

| key | value |
|---------------|-------|
| Lot Frontage | 141 |
| Lot Area | 31770 |
| Overall Qual | 6 |
| Overall Cond | 5 |
| Total Bsmt SF | 1080 |
| Gr Liv Area | 1656 |
| Bedroom AbvGr | 3 |
| TotRms AbvGrd | 7 |
| Garage Cars | 2 |
| Garage Area | 528 |
| Wood Deck SF | 210 |
| Open Porch SF | 62 |

清除資源



- Azure 雲端介紹
- Azure各項資源介紹
- ML Studio介紹
- Lab – 設計工具

資源上限 -1



- 閒置的資源仍會被收費：
 - ◆ Azure是使用者付費 Pay-as-You-Go
 - ◆ 使用多少資源就會收多少費用，閒置的計算資源也會被收費，因為是佔用了雲端的計算資源

- 資源的使用上限：
 - ◆ 每個群組/資源都會有個別資源的使用上限
 - ◆ ML Workspace也有自己的資源使用上限
 - ◆ 虛擬機器VM的核心會有區域總數量限制及每個系列的區域總數量限制，兩個限制會同時生效

資源上限 -2

- 資源群組和資源所在的區域可以不同，在建立機器學習ML Studio時就須決定資源所在區域，ML Studio中所使用的計算叢集會和ML Studio會架設在同一個區域
- 查詢ML Studio的資源上限：
[資源群組] TRAINING_RG > [機器學習] trainingml
> [支援與疑難排解] 使用量 + 配額
- Azure會自動帶入資源群組所在地理位置，在使用量 + 配額的欄位可以看到該區域各系列虛擬機器核心的使用量

資源上限 -3

Microsoft Azure 搜尋資源、服務及文件 (G+/)

lihsiang.tu@caloudi.com AU Optronics CORP...

首頁 > TRAINING_RG > trainingml

trainingml | 使用量 + 配額

機器學習

搜尋 (Cmd +/)

重新整理

設定

- 私人端點連線
- 屬性
- 鎖定

監視

- 警訊
- 計量
- 診斷設定
- 記錄

自動化

- 工作 (預覽)
- 匯出範本

支援與疑難排解

- 使用量 + 配額
- 新增支援要求

專用核心使用量 低優先順序核心使用量

專用配額: 已使用 20 個核心 | 剩餘 280 個核心
叢集配額: 已使用 13 個叢集和 CI | 剩餘 187 個

要求配額增加 設定配額

| VM 系列 | 使用量 | |
|-------------------|-----|----------------------|
| 標準 D 系列 vCPU | 0% | 已使用 0 個核心 (共 300 個) |
| 標準 DDSv4 系列 vCPU | 0% | 已使用 0 個核心 (共 0 個) |
| > 標準 DSv2 系列 vCPU | 6% | 已使用 20 個核心 (共 300 個) |
| 標準 DSv3 系列 vCPU | 0% | 已使用 0 個核心 (共 300 個) |
| > 標準 Dv2 系列 vCPU | 0% | 已使用 0 個核心 (共 300 個) |
| 標準 Dv3 系列 vCPU | 0% | 已使用 0 個核心 (共 300 個) |
| 標準 EAv4 系列 vCPU | 0% | 已使用 0 個核心 (共 0 個) |
| 標準 Ev3 系列 vCPU | 0% | 已使用 0 個核心 (共 0 個) |
| 標準 FSv2 系列 vCPU | 0% | 已使用 0 個核心 (共 300 個) |
| 標準 H 系列 vCPU | 0% | 已使用 0 個核心 (共 0 個) |
| 標準 HCS 系列 vCPU | 0% | 已使用 0 個核心 (共 0 個) |

清除資源 -1

- Azure的雲端計算資源都有AutoScaling的功能，如果將節點數目下限設定為0時，閒置的資源就會被解除配置而不佔用資源。
- 在建立設計工具管線時需要設定計算資源，設定的節點nodes數目下限為0，當管線部署完後節點會自動解除配置以避免產生閒置資源，但產生的從集會保留，可能造成叢集的數量超過限制



清除資源 -2

- 從ML Studio檢查所有計算資源：
[管理] 計算 > 計算叢集

| 名稱 | 狀態 | 大小 | 建立日期 ↓ | 使用中的執行...
節點 | 閒置節點 |
|-----------|------------|------------------|------------------|-----------------|------|
| MLCompute | 成功 (0 個節點) | STANDARD_D2_V2 | 2021年4月21日 16:08 | 0 | 0 |
| testDS12 | 成功 (0 個節點) | STANDARD_DS12_V2 | 2021年4月21日 15:28 | 0 | 0 |

- 確認計算的狀態都已經顯示為成功，找到閒置的計算叢集 MLCompute，點入確認資源是閒置的狀態

MLCompute

詳細資料 節點 執行

⟳ 重新整理 ✎ 刪除

叢集節點狀態

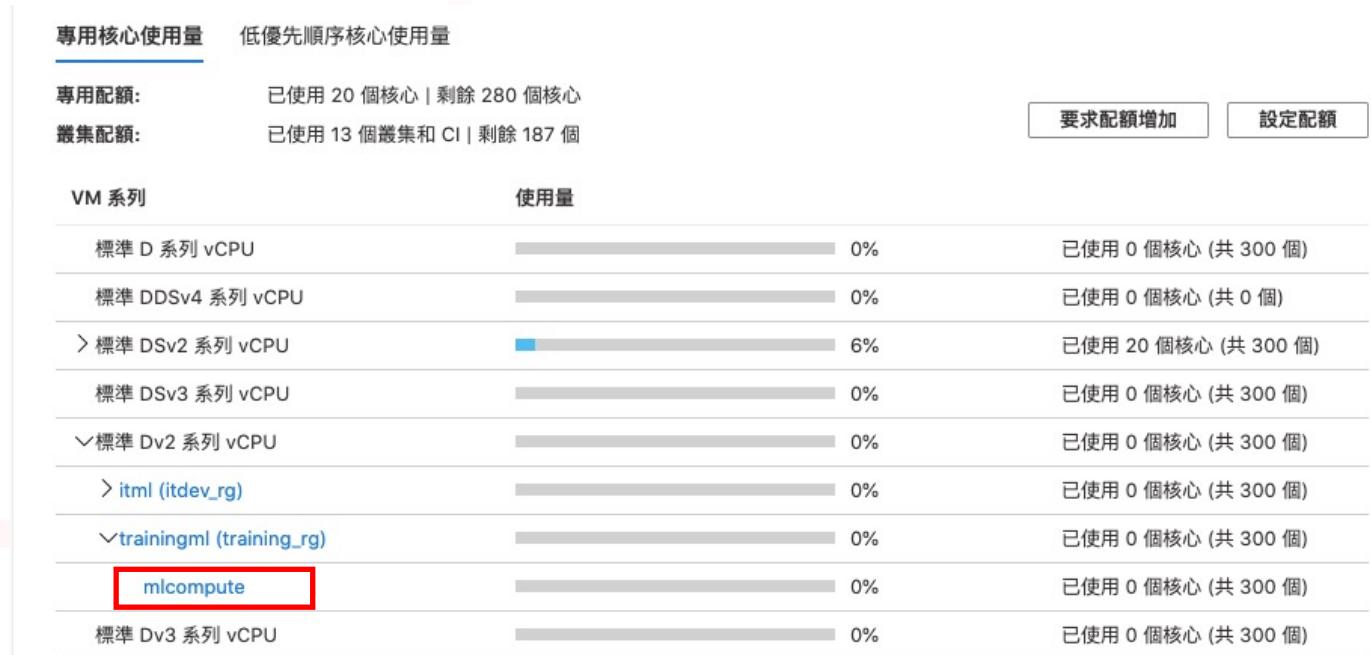
- 閒置
- 正在離開
- 正在準備
- 正在執行

0

清除資源 -3

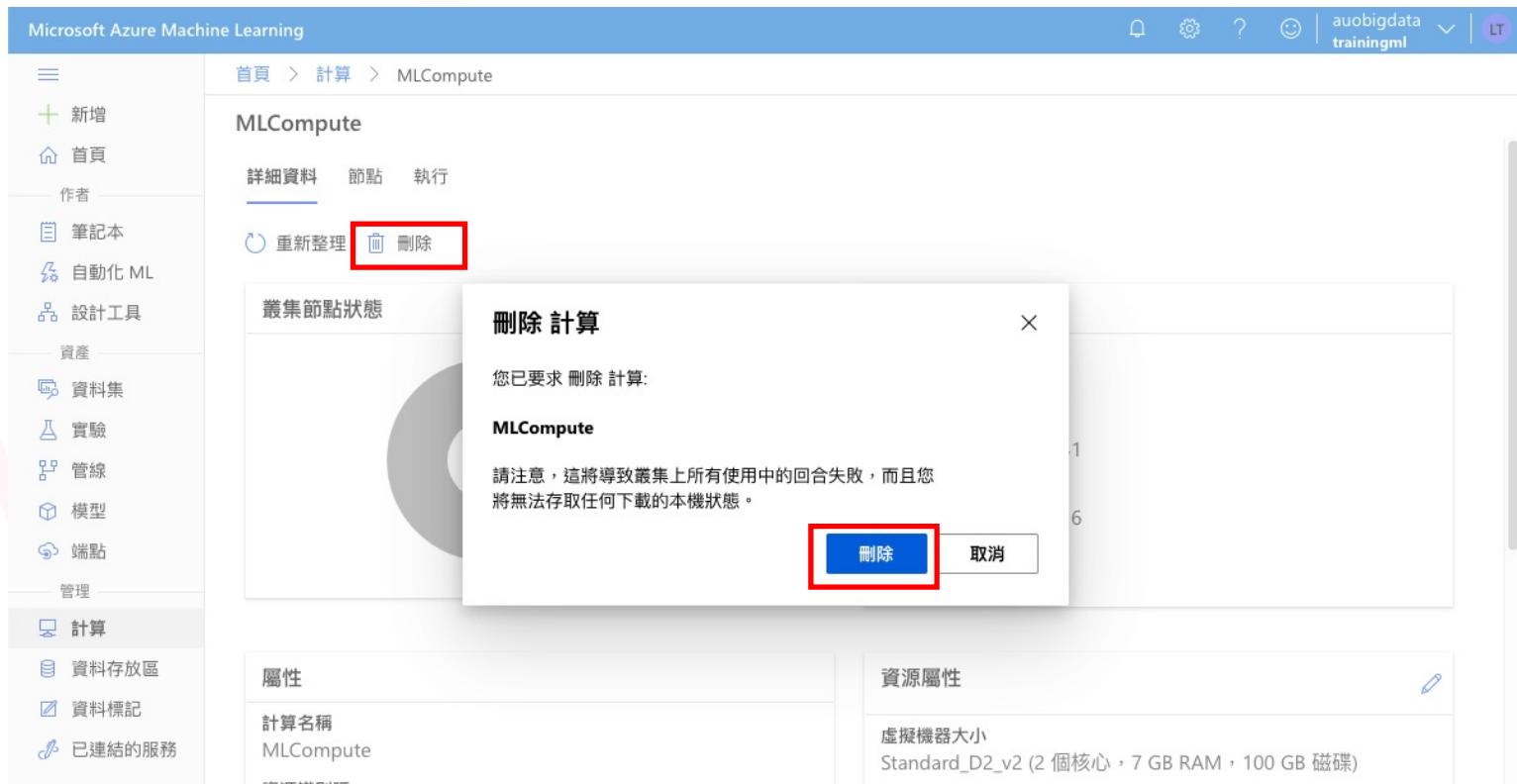
- 找到閒置資源後，回到 Azure Machine Learning 的頁面確認佔用的狀態

[資源群組] TRAINING_RG > [機器學習] trainingml
> [支援與疑難排解] 使用量 + 配額



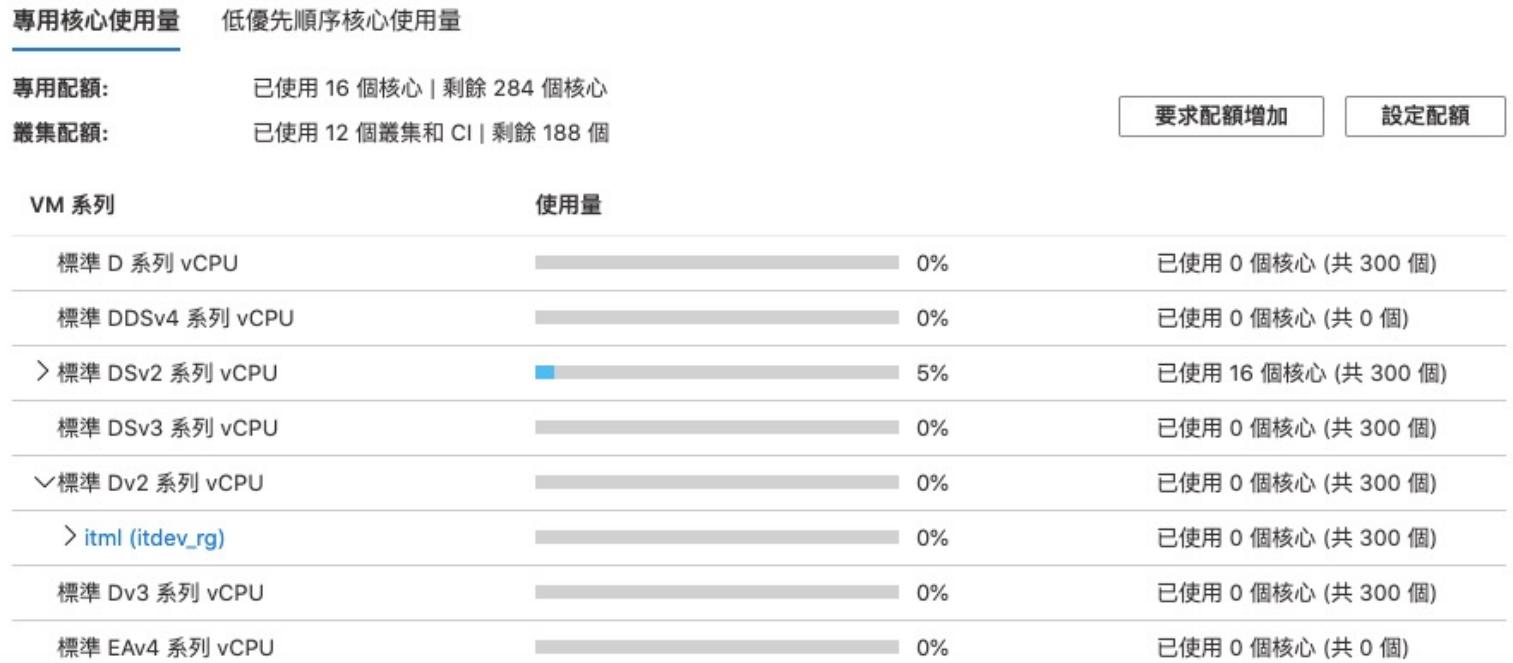
清除資源 -4

- 回到ML Studio的計算叢集MLCompute頁面，點選上方的刪除，確認[刪除]，可以在該畫面重新整理確認執行狀態，直到資源被刪除出現錯誤消息(無法載入計算)



清除資源 -5

- 再在[支援與疑難排解] 使用量 + 配額頁面重新整理，確認該組叢集已經不再佔用資源



清除資源 -6

- 提醒您，Azure雲端服務是依照使用的資源計費，閒置未使用的資源亦可能因相依的服務被收費(例如儲存體帳戶)，除了產生不必要的費用之外，佔用各項資源的配額也可能導致他人無法建立所需資源。
- 建議您，在確認無需使用該項服務時，解除配置或刪除閒置資源