



códigofacilito

---

# CódigoFacilito.

- Inauguración e introducción al bootcamp
- Introducción a los servicios de IA y Data en Azure
- Diseño de una solución de aprendizaje automático

Oscar Santos – ML Engineer



## Temario:

- >\_ Inauguración
- >\_ Servicios de Data
- >\_ Servicios de IA
- >\_ Diseño de solución





## Azure Machine Learning

<https://learn.microsoft.com/es-es/credentials/certifications/azure-data-scientist/?practice-assessment-type=certification>



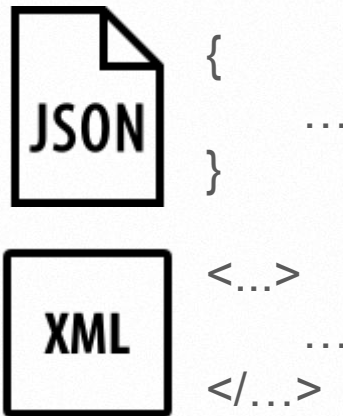


## >\_ Servicios de Data: Conceptos

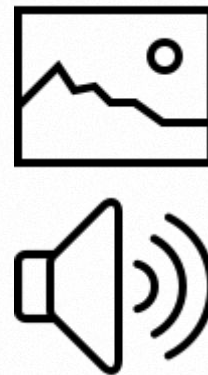
Customer				
ID	FirstName	LastName	Email	Address
1	Joe	Jones	joe@litware.com	1 Main St.
2	Samir	Nadoy	samir@northwind.com	123 Elm Pl.

Product		
ID	Name	Price
123	Hammer	2.99
162	Screwdriver	3.49
201	Wrench	4.25

Estructurado



Semi-estructurado



No estructurado







## >\_ Servicios de Data: Storage

Customer				
ID	FirstName	LastName	Email	Address
1	Joe	Jones	joe@litware.com	1 Main St.
2	Samir	Nadoy	samir@northwind.com	123 Elm Pl.

Product		
ID	Name	Price
123	Hammer	2.99
162	Screwdriver	3.49
201	Wrench	4.25

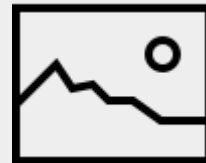
Bases de datos



{  
...  
}

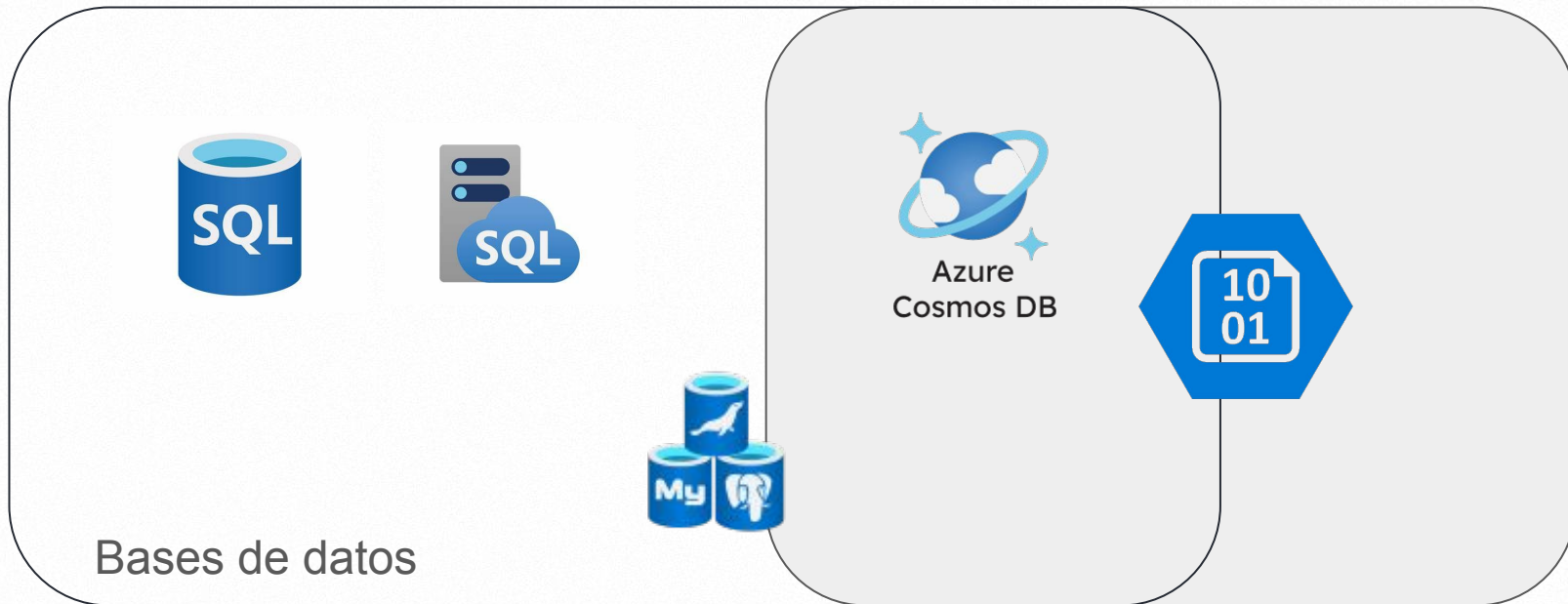


<...>  
...  
</...>



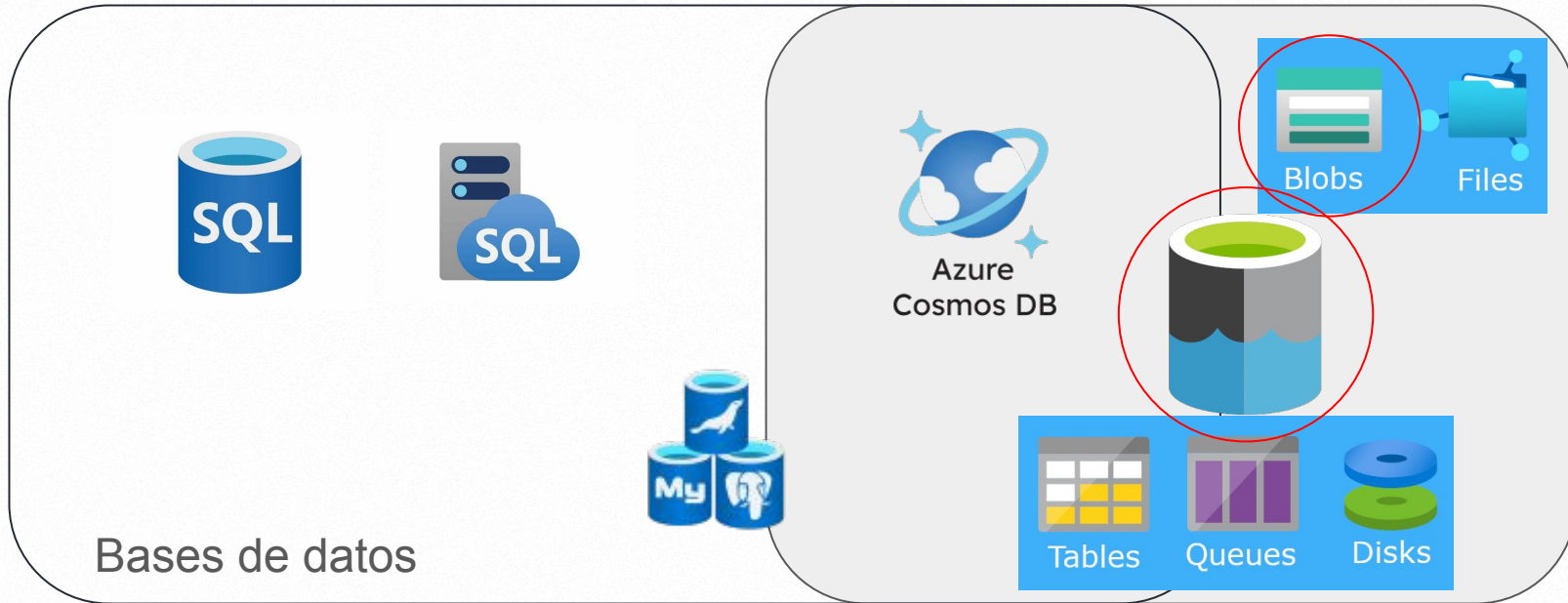


## >\_ Servicios de Data: Storage





## >\_ Servicios de Data: Storage







## >\_ Servicios de Data: Otros







## >\_ Servicios de IA: Conceptos





## >\_ Servicios de IA: Conceptos

Implementación  
completa o  
parcial con alta  
personalización y  
control



Azure Machine Learning



Son sencillos de  
implementar y no  
requieren conocimientos  
especializados de  
inteligencia artificial.





# Azure Machine Learning



- >\_ Exploración de los datos y preparación para su modelización.
- >\_ Entrenamiento y evaluación de modelos de Machine Learning.
- >\_ Registro y administración de modelos entrenados.
- >\_ Implementación de modelos entrenados para su uso por aplicaciones y servicios.
- >\_ Revisión y aplicación de los principios y prácticas de inteligencia artificial responsables.

¿Qué podemos hacer en Azure Machine Learning?



- >\_ Almacenamiento centralizado
- >\_ Recursos de proceso a petición
- >\_ Aprendizaje automático automatizado (AutoML)
- >\_ Herramientas visuales (low code/no code)
- >\_ Integración con marcos de aprendizaje automático comunes
- >\_ Compatibilidad integrada para visualizar y evaluar métricas

Características y capacidades

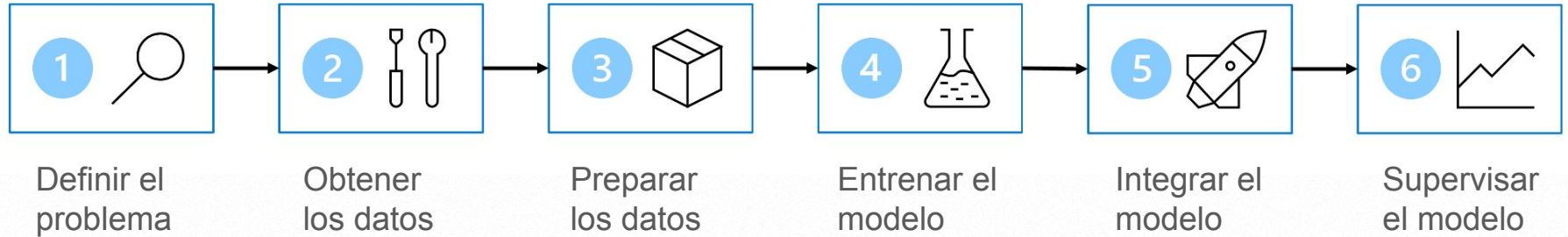


# Diseño de una solución de aprendizaje automático





## Identificación del origen de datos y el formato





## Identificación del problema

En el primer paso, quiere definir el problema que resolverá el modelo mediante la comprensión de lo siguiente:

- Cuál debe ser la salida del modelo.
- Qué tipo de tarea de Machine Learning usará.
- Qué criterios hacen que un modelo sea correcto.







## Identificación del origen de datos y el formato

1

```
{ "deviceId": 29482, "machine": "Machine1", "time"="2021-07-14T12:47:39Z", "temperature": 20.4 }  
{ "deviceId": 38273, "machine": "Machine1", "time"="2021-07-14T12:47:49Z", "temperature": 20.6 }  
{ "deviceId": 43819, "machine": "Machine2", "time"="2021-07-14T12:47:33Z", "temperature": 23 }  
{ "deviceId": 38273, "machine": "Machine1", "time"="2021-07-14T12:47:54Z", "temperature": 20.4 }  
{ "deviceId": 29482, "machine": "Machine1", "time"="2021-07-14T12:48:35Z", "temperature": 20.6 }  
{ "deviceId": 58291, "machine": "Machine2", "time"="2021-07-14T12:48:36Z", "temperature": 23 }
```



2

Device ID	Machine	Date	Time	Temperature
29482	Machine1	2021-07-14	12:47:39	20.4
38273	Machine1	2021-07-14	12:47:49	20.6
43819	Machine2	2021-07-14	12:47:33	23
38273	Machine1	2021-07-14	12:47:54	20.6
58291	Machine2	2021-07-14	12:48:36	23



3

Machine	Date	Time	Temperature
Machine1	2021-07-14	12:47	20.5
Machine2	2021-07-14	12:47	23
Machine1	2021-07-14	12:48	20.5
Machine2	2021-07-14	12:48	23

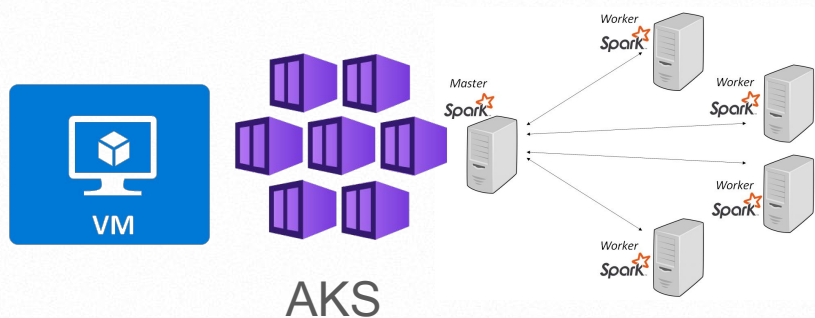




## > Elección de cómo servir datos a flujos de trabajo de aprendizaje automático

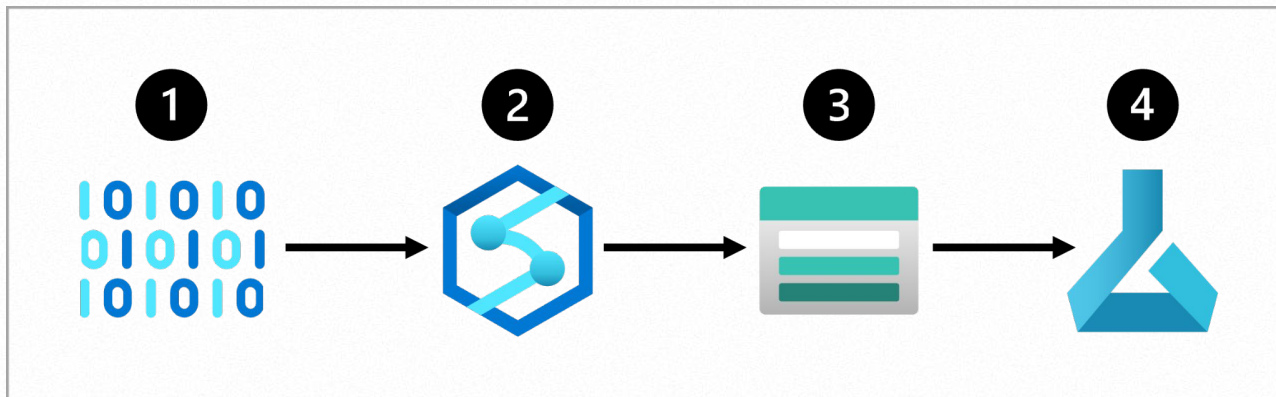


Cómputo  
Almacenamiento





## Diseño de una solución de ingesta de datos



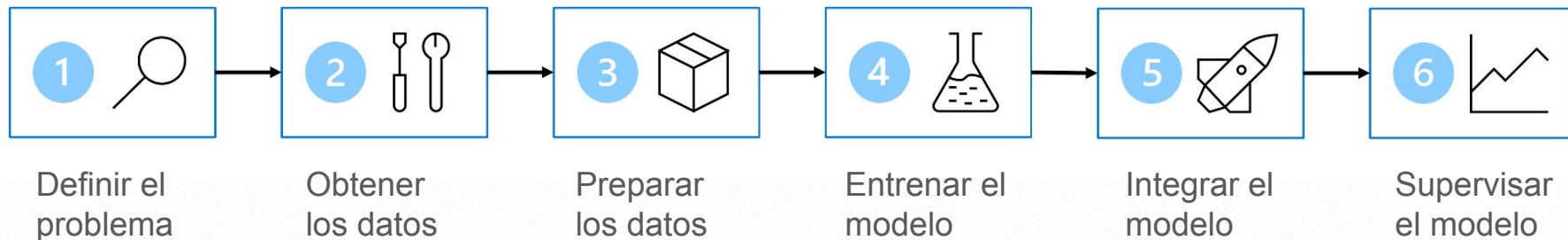
1. Extraer los datos sin procesar de su origen (como un sistema CRM o un dispositivo IoT).
2. Copiar y transformar los datos con Azure Synapse Analytics.
3. Almacenar los datos preparados en Azure Blob Storage.
4. Entrenar el modelo con Azure Machine Learning.







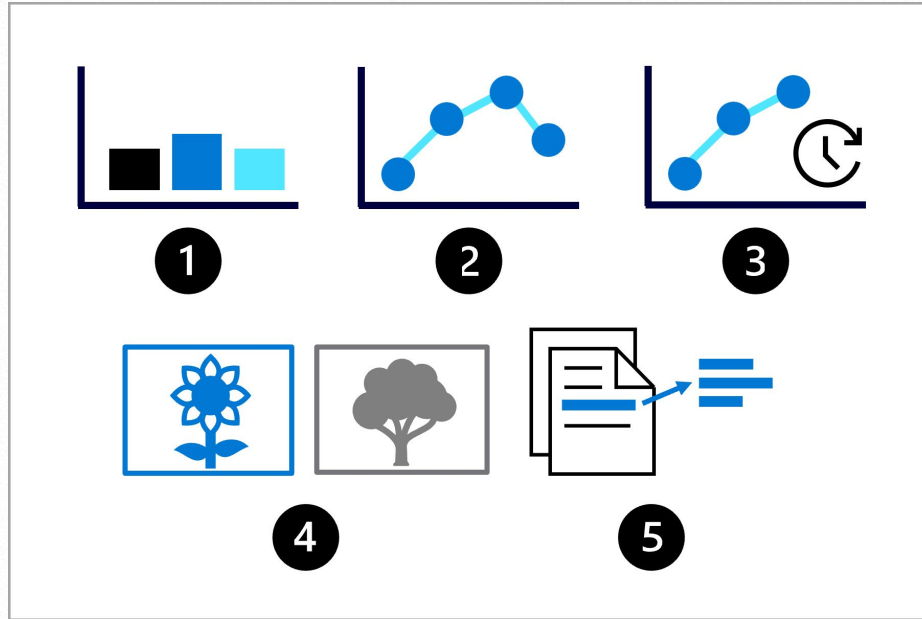
## Identificación de tareas de Machine Learning







## Identificación de tareas de Machine Learning



1. Clasificación
2. Regresión
3. Previsión de series temporales
4. Computer Vision
5. Procesamiento del lenguaje natural (NLP)





## Elección de un servicio para entrenar un modelo de Machine Learning





## > Elección de un servicio para entrenar un modelo de Machine Learning

1. El tipo de modelo que necesita entrenar.
2. Si necesita un control total sobre el entrenamiento del modelo.
3. El tiempo que desea invertir en el entrenamiento del modelo.
4. Los servicios que ya están dentro de su organización.
5. El lenguaje de programación con el que se siente cómodo.



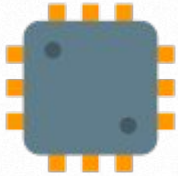
Azure Machine Learning



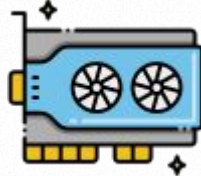




## Elección entre opciones de proceso



VS



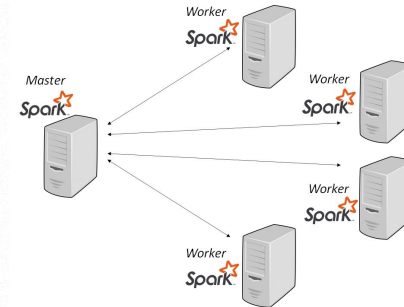
VS



General

Optimizado  
para memoria

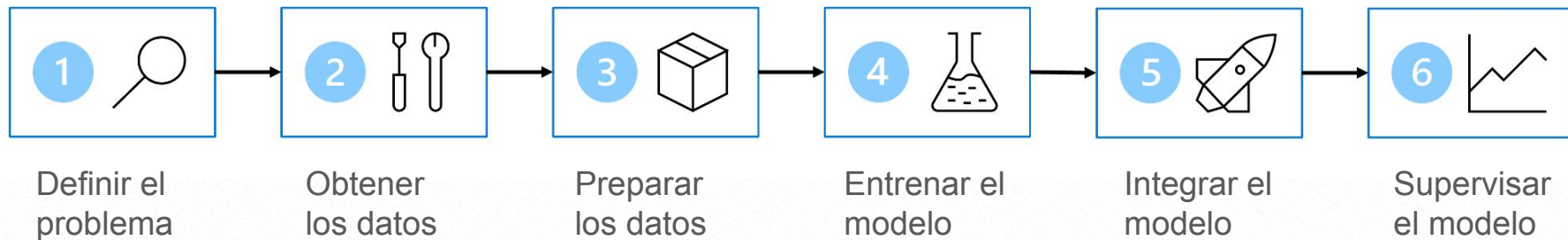
VS







## Comprender cómo se consumirá el modelo





## Comprender cómo se consumirá el modelo

Al implementar un modelo en un punto de conexión, tiene dos opciones:

1. Obtener predicciones en tiempo real
2. Obtener predicciones por lotes





## Obtención de predicciones en tiempo real

Imagine que tiene un sitio web que contiene un catálogo de productos:

1. Un cliente selecciona un producto en su sitio web, como una camisa.
2. En función de la selección del cliente, el modelo recomienda otros elementos del catálogo de productos inmediatamente. El sitio web muestra las recomendaciones del modelo.







## Obtención de predicciones en tiempo real







## Obtención de predicciones por lotes

Ejemplo:

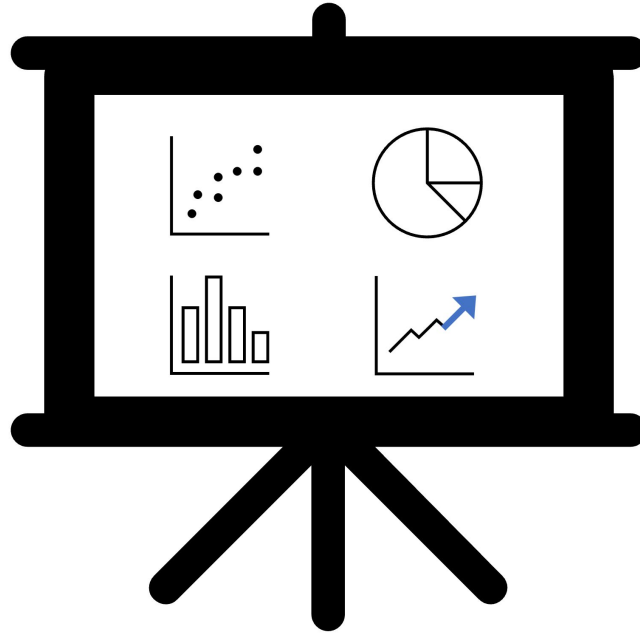


Imagine que está visualizando todos los datos históricos de ventas en un informe. Querrá incluir las ventas previstas en el mismo informe.





## Obtención de predicciones por lotes





## >\_ Decidir la implementación por lotes o en tiempo real

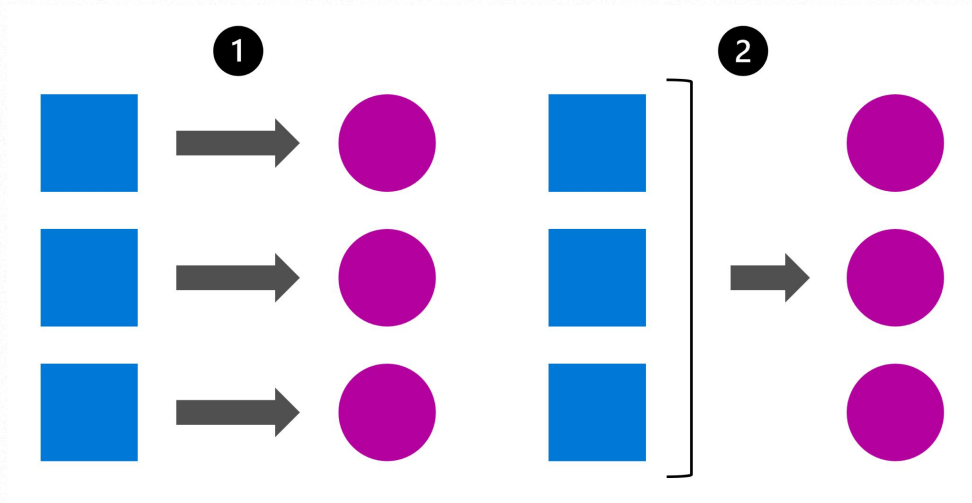
- ¿Con qué frecuencia se deben generar las predicciones?
- ¿Con qué puntualidad se necesitan los resultados?
- ¿Se deben generar predicciones de forma individual o por lotes?
- ¿Cuánta capacidad de proceso se necesita para ejecutar el modelo?







## >\_ Decidir la implementación por lotes o en tiempo real





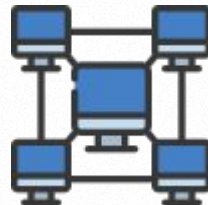


## >\_ Decidir la implementación por lotes o en tiempo real

- Decisión del número de predicciones
- Consideración del costo del proceso
- Decidir la implementación por lotes o en tiempo real



vs





## Caso



¿Mi paciente tiene diabetes?

Pregnancies	<input type="text" value="0"/>
PlasmaGlucose	<input type="text" value="104"/>
DiastolicBloodPressure	<input type="text" value="51"/>
TricepsThickness	<input type="text" value="7"/>
SerumInsulin	<input type="text" value="24"/>
BMI	<input type="text" value="23"/>
DiabetesPedigree	<input type="text" value="1"/>
Age	<input type="text" value="43"/>

Analyze

Yes, with 89% certainty



## Caso

La primera característica que tenemos prevista es que la aplicación indique al médico si debe seguir examinando al paciente o tratarlo de diabetes.

Ya hemos recopilado datos que se correlacionan con la diabetes, como el número de embarazos, la edad y el índice de masa corporal (IMC). También tenemos un equipo de científicos de datos que trabajan en el entrenamiento de un modelo capaz de señalar las probabilidades de que un paciente tenga diabetes.

Necesitamos su ayuda para decidir cómo implementar el modelo para integrarlo en nuestra aplicación móvil.




Esperamos su consejo sobre cómo diseñar la solución de implementación del modelo.







## Caso

Requisito	Descripción
	<b>Tenga en cuenta la frecuencia.</b> El plan es que un médico introduzca la información de un paciente en la aplicación, como su edad y el IMC. Tras ello, un médico puede seleccionar el botón <code>Analyze</code> , que hará que el modelo pronostique las probabilidades de que un paciente tenga diabetes o no.
	<b>Tenga en cuenta el proceso.</b> Una consulta médica suele tardar menos de 10 minutos. Si queremos que los médicos usen esta aplicación, necesitamos que obtengan las respuestas lo antes posible. El modelo implementado siempre debe estar disponible, ya que no sabemos cuándo lo puede usar un médico.
	<b>Tenga en cuenta el tamaño.</b> Un médico usará la aplicación exclusivamente para obtener un pronóstico sobre la situación de una persona. No es necesario generar pronósticos de varios pacientes a la vez.







## Caso

¿Qué tipo de pronósticos se necesitan en la aplicación móvil?

1. Predicciones en tiempo real
2. Predicciones por lotes
3. Pronósticos locales





## Caso

¿Qué tipo de pronósticos se necesitan en la aplicación móvil?

1. **Predicciones en tiempo real**
2. Predicciones por lotes
3. Pronósticos locales





## Caso

¿Qué tipo de proceso debe usar el modelo implementado?

1. Máquinas virtuales
2. Contenedores
3. Dispositivo local







## Caso

¿Qué tipo de proceso debe usar el modelo implementado?

1. Máquinas virtuales
2. **Contenedores**
3. Dispositivo local







## > Exploración de una arquitectura de MLOps

- En mi computadora si funciona
- Si pero no le vamos a dar tu computadora al cliente





## >\_ Exploración de una arquitectura de MLOps

Para preparar el modelo y hacerlo funcionar, quiere hacer lo siguiente:

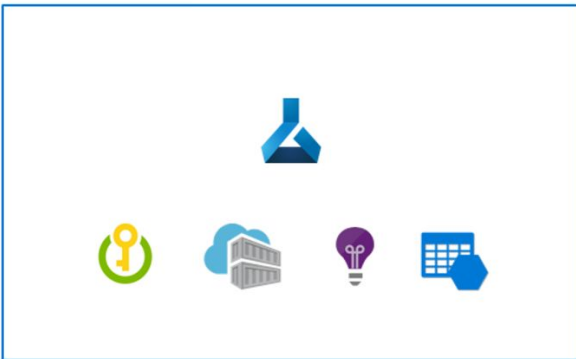
- Convertir el entrenamiento del modelo en una canalización (pipeline) sólida y reproducible.
- Probar el código y el modelo en un entorno de desarrollo.
- Implementar el modelo en un entorno de producción.
- Automatizar el proceso de un extremo a otro.



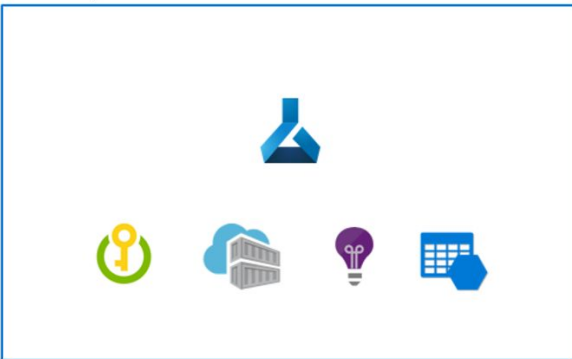


## Configuración de entornos para desarrollo y producción

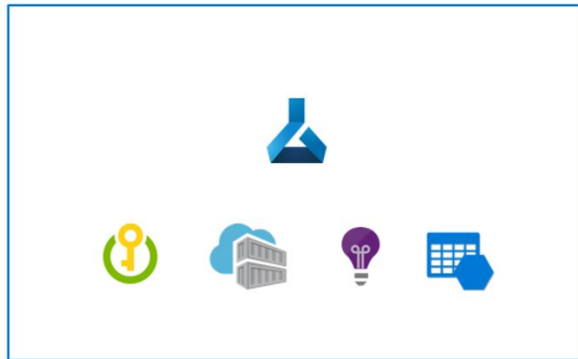
Desarrollo



Preproducción



Producción

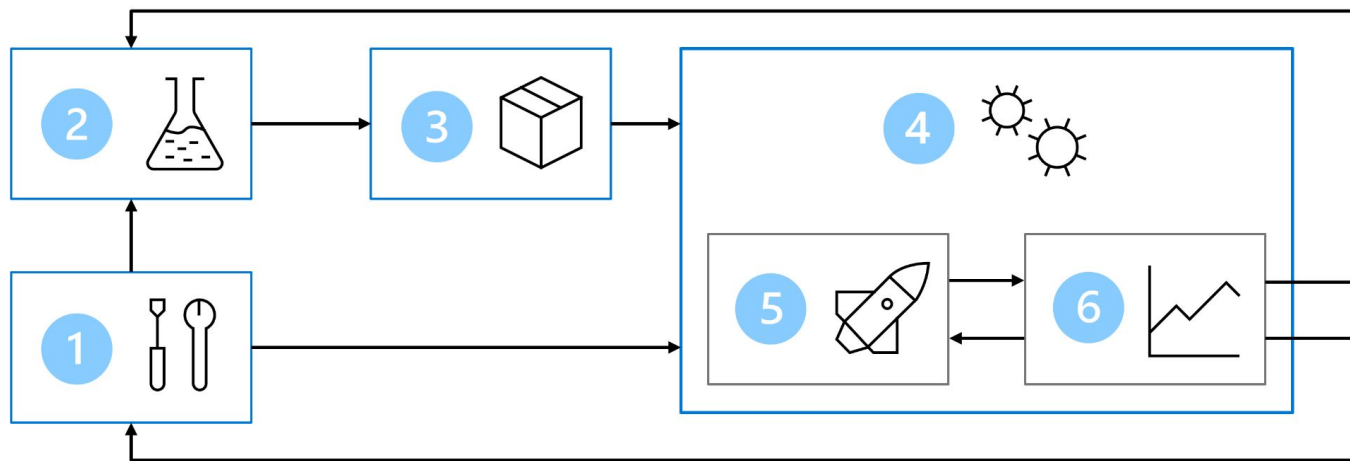


RBAC





## > Diseño de una arquitectura de MLOps



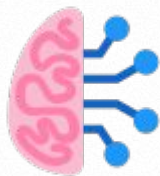
1. Configuración
2. Desarrollo de modelos (bucle interno)
3. Integración continua
4. Implementación de modelos (bucle externo)
5. Implementación continua
6. Supervisión



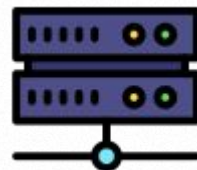
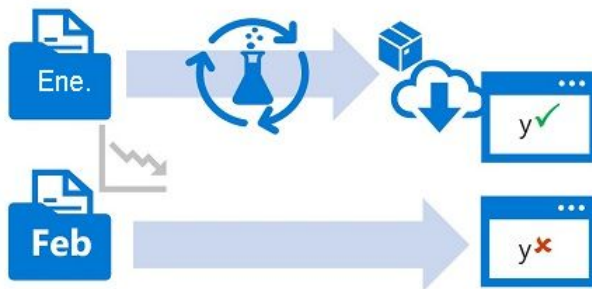




## >\_ Diseño de la supervisión



mlflow™





## >\_ Diseño del reentrenamiento





## > Preparación del código







## >\_ Automatización del código



GitHub Actions



Azure DevOps







## Caso

Esperamos su consejo sobre cómo diseñar la solución de operaciones de aprendizaje automático (MLOps).



¿Mi paciente tiene diabetes?

Pregnancies	0
PlasmaGlucose	104
DiastolicBloodPressure	51
TricepsThickness	7
SerumInsulin	24
BMI	23
DiabetesPedigree	1
Age	43

Analyze

Yes, with 89% certainty



## Caso

- Tenga en cuenta los entornos. Actualmente, estamos trabajando en un equipo pequeño y usted es el único científico de datos que participa. Queremos ver si este proyecto tiene éxito antes de escalarlo verticalmente y conseguir que un equipo grande participe.
- Tenga en cuenta el modelo. Como el modelo se usa para ayudar a los médicos, la precisión es importante para nosotros. El modelo solo debe estar en uso cuando sepamos que funciona según lo previsto.
- Tenga en cuenta los datos. Estamos comenzando poco a poco y usaremos principalmente el modelo implementado para probar la aplicación. Los datos en los que el modelo implementado genera predicciones no deben usarse para volver a entrenar el modelo, ya que pueden estar sesgados.





## Caso

¿Cuántas áreas de trabajo de Azure Machine Learning debe crear el equipo?

1. Uno
2. Dos
3. Tres







## Caso

¿Cuántas áreas de trabajo de Azure Machine Learning debe crear el equipo?

1. **Uno**
2. Dos
3. Tres







## Caso

¿Cuándo deberíamos volver a entrenar el modelo?

1. Cada semana.
2. Cuando las métricas del modelo están por debajo del punto de referencia.
3. Cuando haya un desfase de datos.





## Caso

¿Cuándo deberíamos volver a entrenar el modelo?

1. Cada semana.
2. **Cuando las métricas del modelo están por debajo del punto de referencia.**
3. Cuando haya un desfase de datos.

