



Inteligencia Artificial

Aspectos básicos

Dr. José Ramón Iglesias

DSP-ASIC BUILDER GROUP

Director Semillero TRIAC

Ingenieria Electronica

Universidad Popular del Cesar

¿Qué es?



Artificial intelligence (AI)

Cualquier técnica que permite a las computadoras imitar la inteligencia humana
Uso de lógica, instrucciones if-then y aprendizaje automático



Machine learning (ML)

Un subconjunto de IA que utiliza máquinas para buscar
Patrones en los datos para crear modelos lógicos automáticamente



Deep learning (DL)

Un subconjunto de ML compuesto por redes neuronales profundamente multicapa
que realizan tareas como el reconocimiento de voz e imagen



Reinforcement learning (RL)

El aprendizaje por refuerzo entrena modelos mediante el uso de una recompensa
función de un resultado deseado



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Problemas que se resuelven con ML

Anticiparse a la cancelación (**abandono/churn**) de un contrato de servicios por parte de un cliente.

Ofrecer recomendaciones de productos a los clientes basadas en compras anteriores.

Determinar el **precio** de venta de una casa

Determinar si una transacción bancaria es un **possible fraude**



¿Qué es un modelo?

$$f(x) = a_0x_0 + a_1x_1 + \cdots + a_nx_n$$

Modelo en ML es una representación matemática

Características y pesos



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

¿Qué es un modelo?

$$f(\textcolor{brown}{x}) = \textcolor{violet}{a}_0x_0 + \textcolor{violet}{a}_1x_1 + \cdots + \textcolor{violet}{a}_n x_n$$

Label o Target: es el valor “ y ” real/actual/histórico que el modelo aprenderá a predecir

f($\textcolor{brown}{x}$) = precio real de una casa que depende de variables $\textcolor{brown}{x}$, “sospechosas”, que influyen en el precio



Más características y ponderaciones

$$f(x) = a_0x_0 + a_1x_1 + \cdots + a_nx_n$$

x_1 : ¿número de habitaciones?

x_2 : ¿en cuál zona de la ciudad se ubica?

x_3 : ¿año de construcción?

...

x_n : ¿tiene piscina?



Salida

$f(\textcolor{brown}{x}) = \hat{y}$, predicción del precio .

$$f(\textcolor{brown}{x}) = \textcolor{violet}{1000} * 100(m^2) + 5000 * 3(\text{habitaciones}) + \dots \\ \dots + 3000 * 1(\text{sí tiene piscina})$$

$$f(\textcolor{brown}{x}) = 118000 \text{ usd}$$



Salida

$f(\textcolor{brown}{x}) = \hat{y}$, predicción.

\hat{y} = precio de la casa (número decimal)

\hat{y} = Sí/No ¿la casa está muy barata?

\hat{y} = económica, mansión ¿qué tipo de casa es?

¿Qué tipo de algoritmo se podría utilizar?



Problemas

Enfoque de programación clásica



Requiere que los programadores establezcan explícitamente las reglas.

- Engorroso, propenso a sesgos y errores humanos.
- No siempre está claro para los humanos cómo resolver el problema.

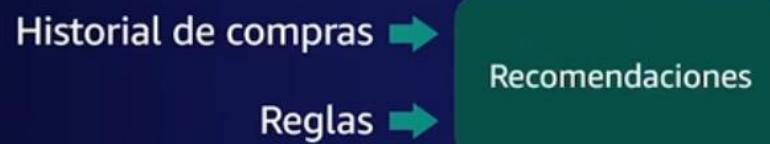
Los gustos de los clientes pueden ser sumamente impredecibles.

Cada regla agrega más tiempo al proceso.

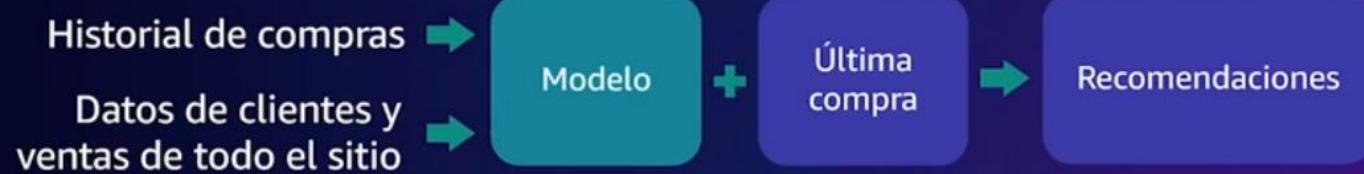


¿Qué ocurre con machine learning?

Enfoque de programación clásica



Enfoque de machine learning



Tres categorías de algoritmos de machine learning

Supervisado

No supervisado

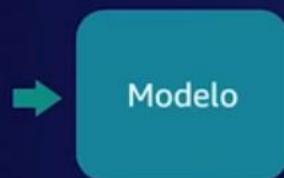
Por refuerzo



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Supervisado

Aprende identificando patrones en datos que **ya están etiquetados**.



Datos nuevos



Predictión

Fotos de automóviles,
etiquetados por personas



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Los diferentes tipos de problemas de ML supervisados

Supervisado

Clasificación binaria

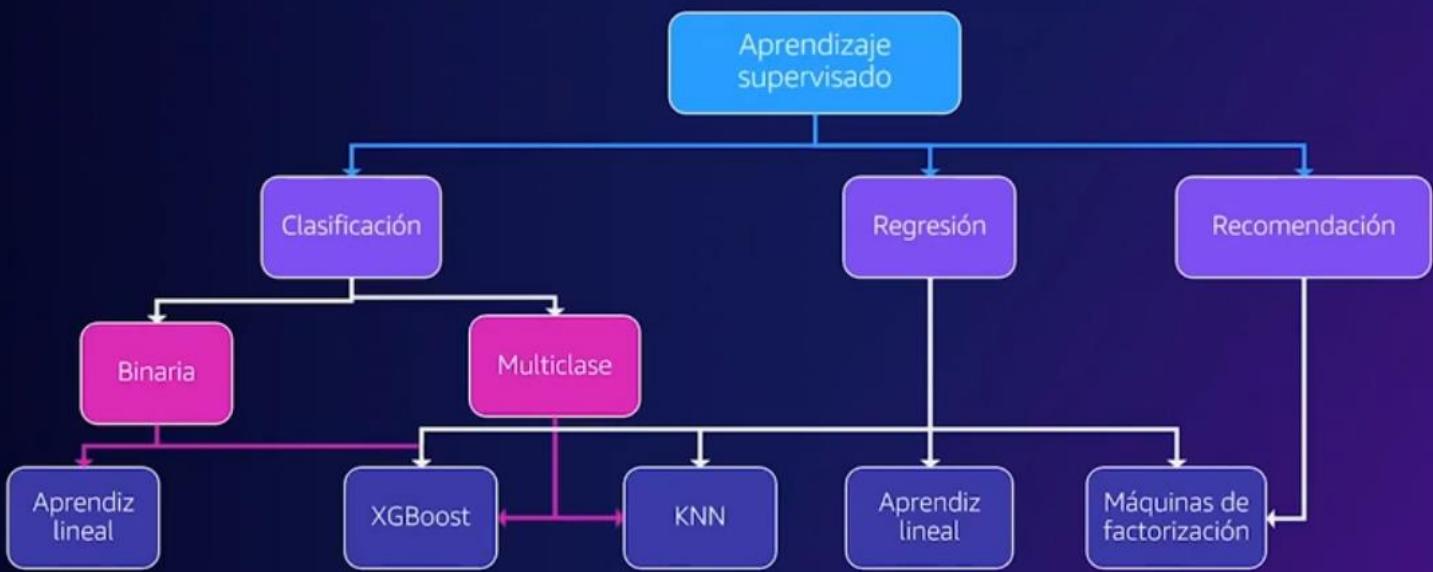
Clasificación multiclas

Problema de regresión



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Algoritmos integrados de Amazon SageMaker



Tres categorías de algoritmos de machine learning

Supervisado

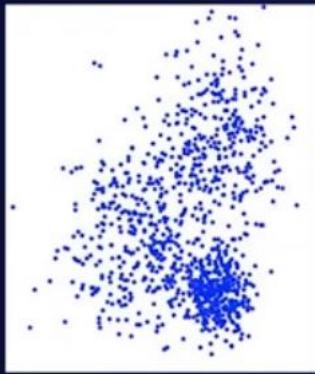
No supervisado

Por refuerzo

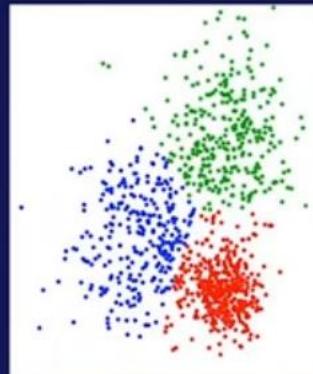


No supervisado

La máquina debe descubrir y **crear las etiquetas** por sí misma.



Datos originales



Datos agrupados



Caso práctico: detección de un error en el radar



¿Estas anomalías son simples **valores atípicos** o **indicadores** de errores de hardware?

Algoritmos integrados de Amazon SageMaker



Tres categorías de algoritmos de machine learning

Supervisado

No supervisado

Por refuerzo



Aprendizaje por refuerzo

Aprender a través de prueba y error

Mejor cuando se conoce el resultado deseado, pero se desconoce el camino exacto para lograrlo.



AWS DeepRacer



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Aprendizaje por refuerzo



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

¿Qué es el aprendizaje profundo? (deep learning)



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Una subcategoría de machine learning

Inteligencia artificial

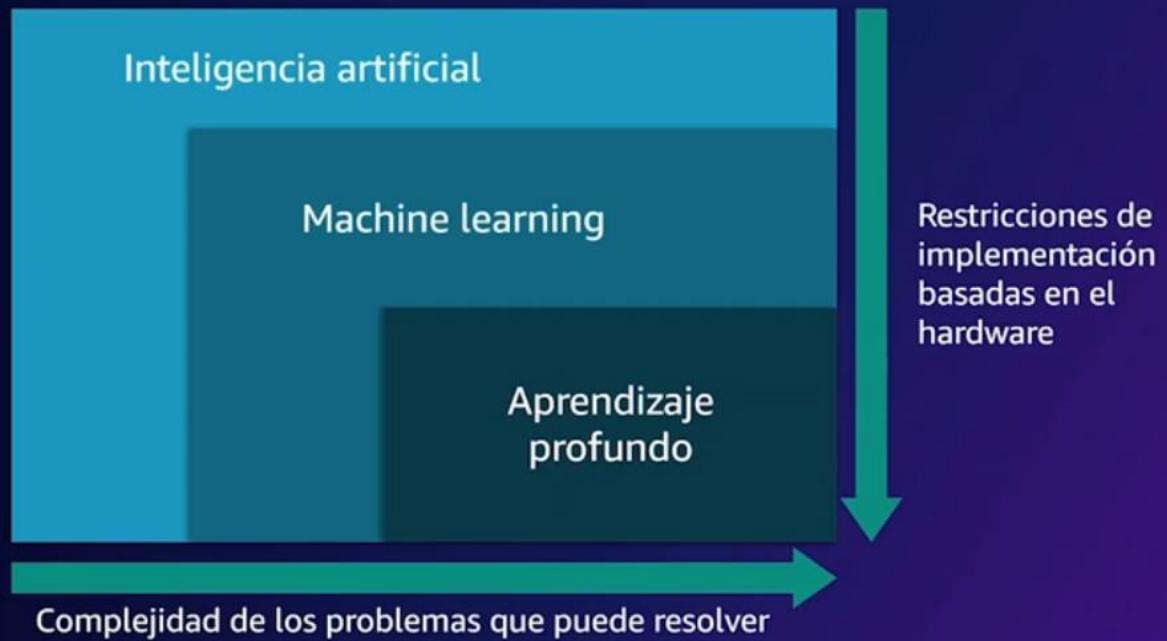
Machine learning

Aprendizaje
profundo



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Una subcategoría de machine learning



Ejemplos

Máquinas que pueden jugar al ajedrez basándose en reglas.

Máquinas que aprenden a jugar al ajedrez a partir del análisis de partidas anteriores jugadas por humanos.

Máquinas que pueden aprender a jugar al ajedrez jugando contra sí mismas.

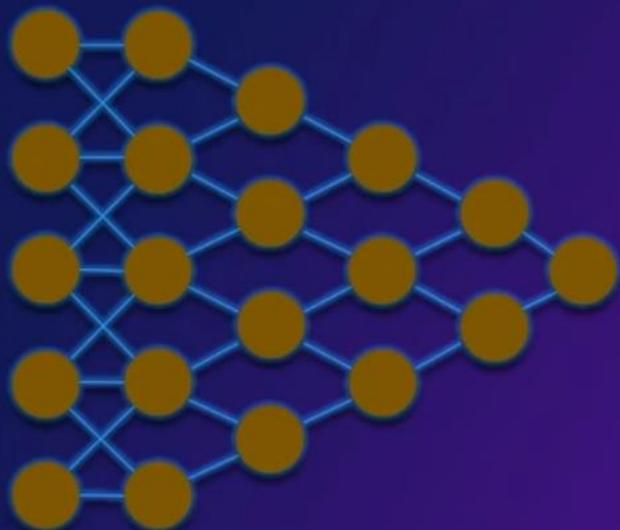
Restricciones de implementación basadas en el hardware

Complejidad de los problemas que puede resolver



Entonces, ¿cómo aprende el aprendizaje profundo?

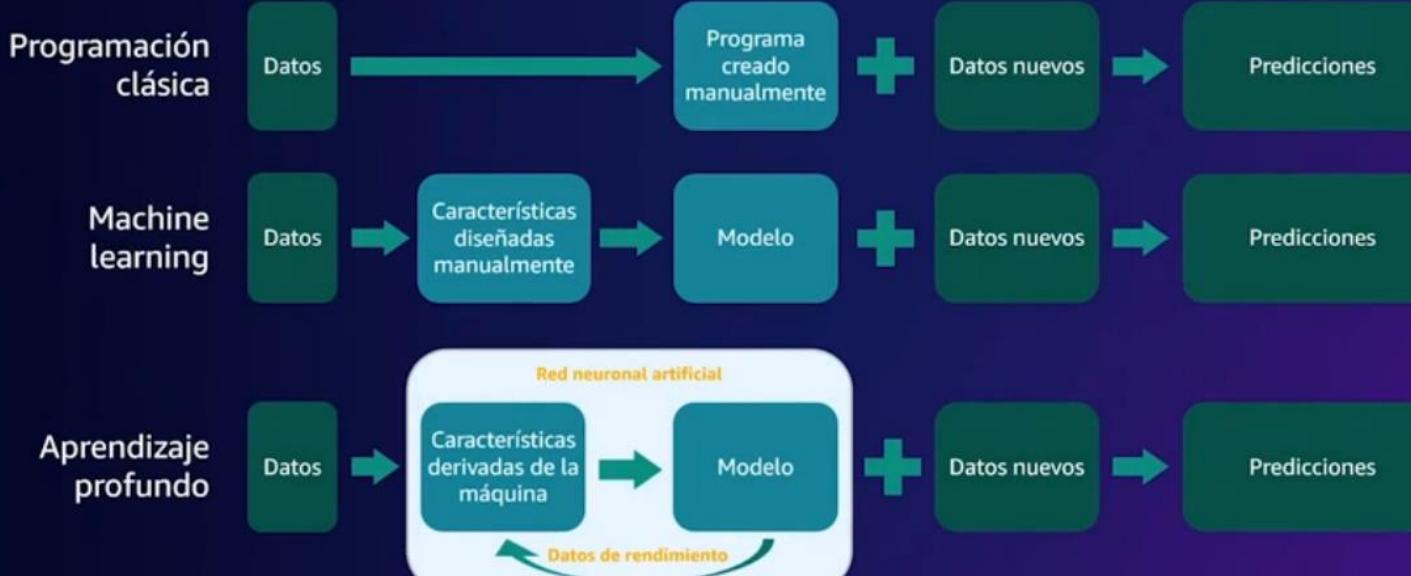
- Las tareas se aprenden mediante **redes neuronales artificiales**.
- Los modelos pueden recibir entrenamiento sobre datos sin procesar.
- La extracción de características se realiza automáticamente; el proceso es iterativo, con varias capas.



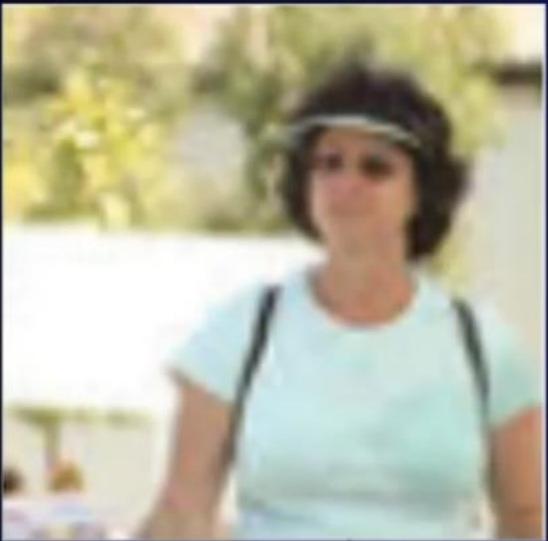
En qué se diferencia el aprendizaje profundo



En qué se diferencia el aprendizaje profundo

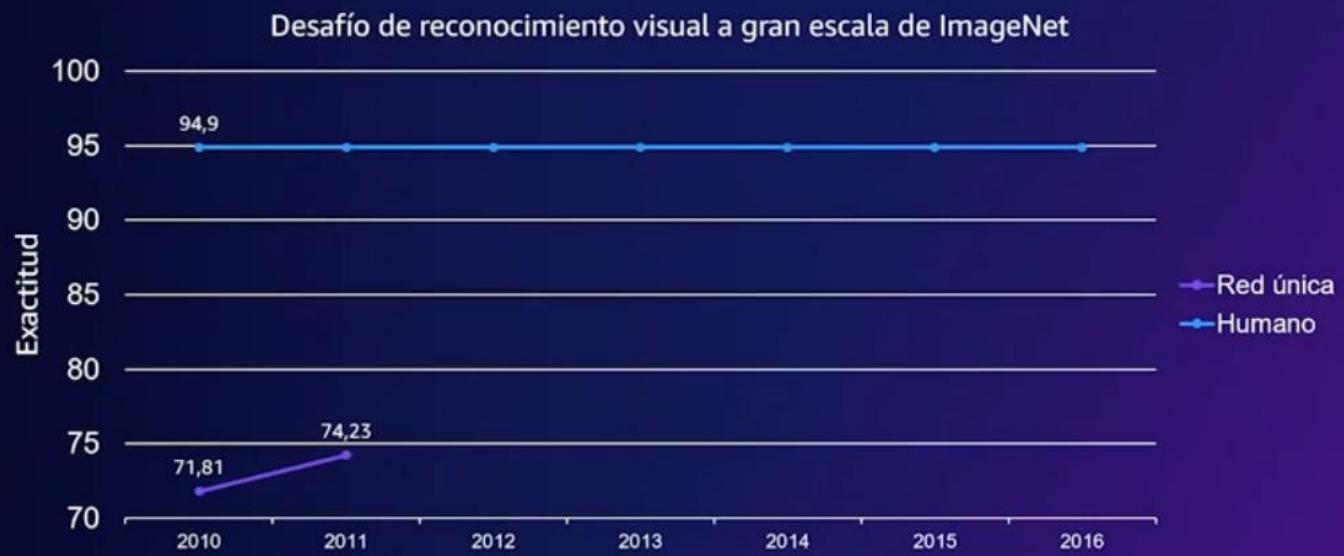


Ejemplo: clasificación de imágenes mediante ANN



Un ejemplo del efecto del aprendizaje profundo

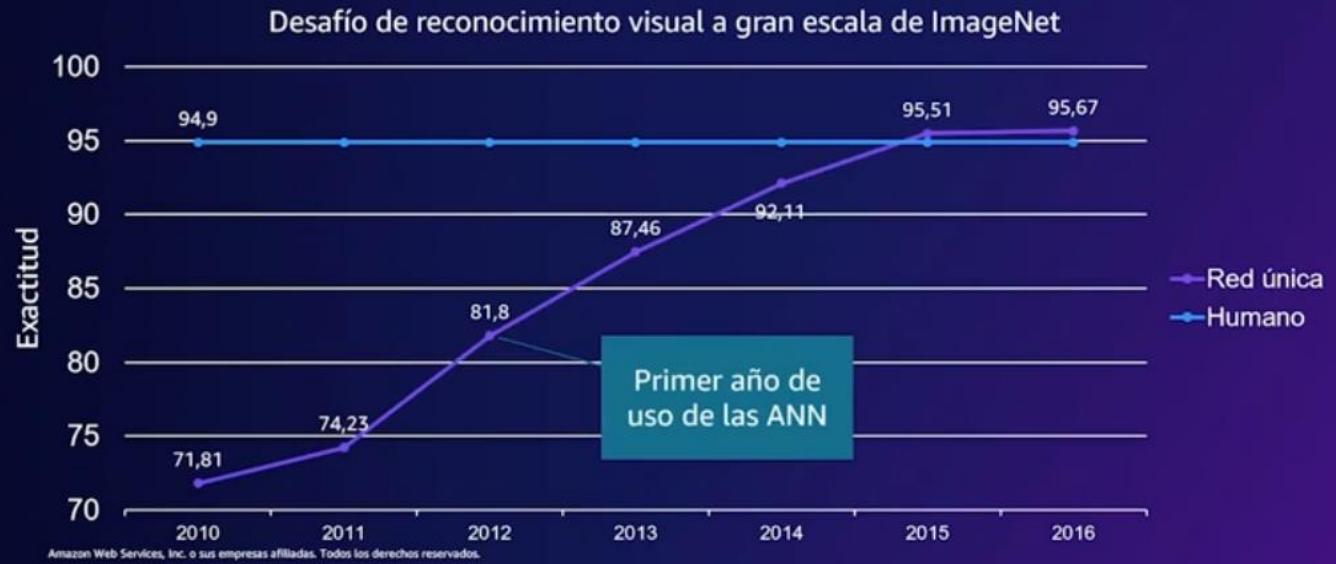
2010 y 2011: incluso los mejores programas presentados seguían estando muy por debajo de la exactitud humana.



Un ejemplo del efecto del aprendizaje profundo

2010 y 2011: incluso los mejores programas presentados seguían estando muy por debajo de la exactitud humana.

Desde 2012: la exactitud aumenta cada año hasta superar la exactitud humana en 2015.



Algoritmos integrados de Amazon SageMaker



Algoritmos integrados de Amazon SageMaker



¿Cómo se realiza un proyecto de ML?



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

La canalización de machine learning. (pipeline)



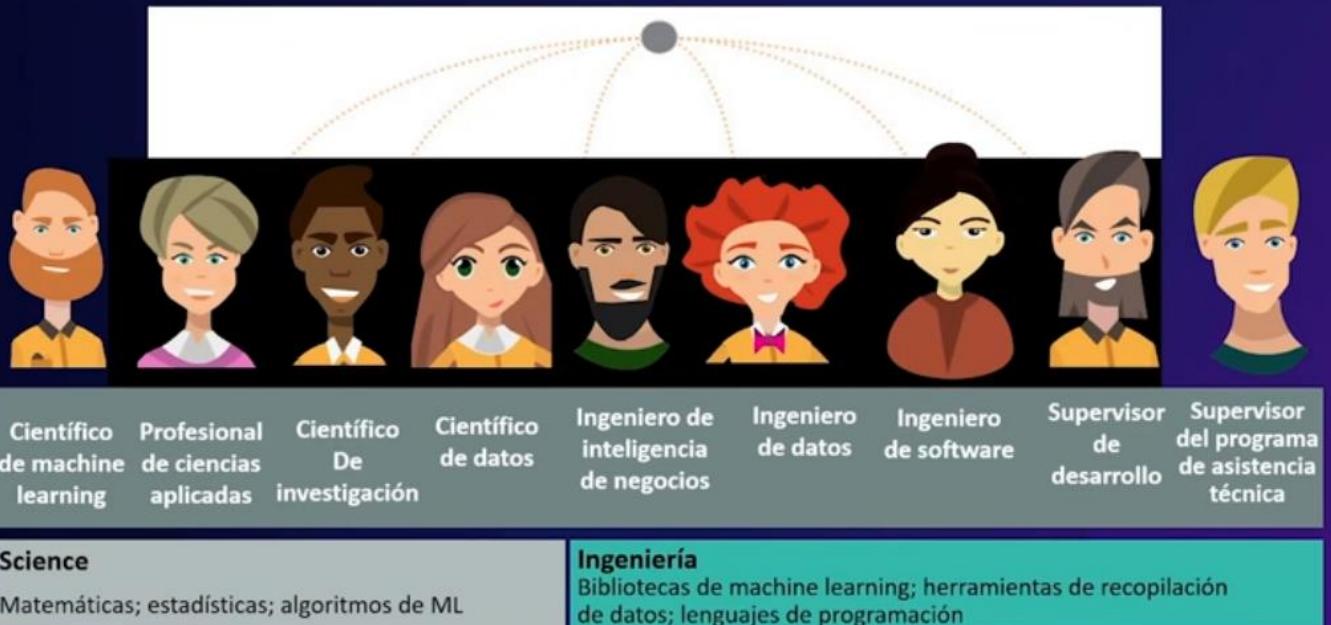
Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

La canalización de ML



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Son necesarias una variedad de habilidades



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Entendimiento del problema



Entendimiento del problema

Problema empresarial



Predicción inexacta de la demanda que causa pérdidas de dinero en la empresa

Objetivo



Mantenga bajos el inventario no vendido y las ventas perdidas

Métrica de éxito



Al final de cada mes, no tener más del 15 % de inventario sin vender sin agotarse



¿Qué modelo elegimos?

Utilice la información para determinar el tipo de problema de machine learning con el que está trabajando.



Otros problemas: predicciones de ventas de productos

Desea determinar si debe llevar un producto en stock en absoluto.

Ha decidido descartar cualquier producto que **tenga** menos de 100 ventas.



Predicción: No

Se trata de un problema de **clasificación binaria**.



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Ejemplo: predicciones de ventas de productos

Desea determinar el mejor mes para poner cada producto a la venta.



Predicción: junio



Predicción: noviembre



Predicción: enero

Este es un problema de **clasificación multiclasé**.



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Enmarque la solución más simple...

... pero trate de no perder información importante.



Demandas: 239



¿> 100 ventas? Sí

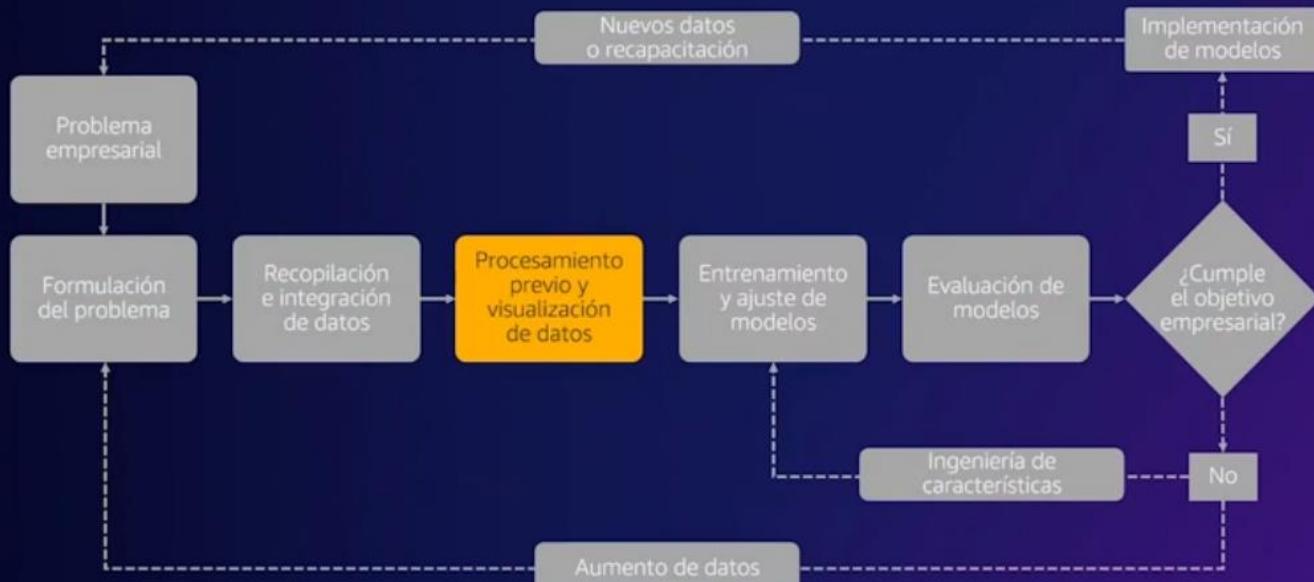
Ayuda a administrar
el suministro y el
inventory

Más simple, pero
 pierde información
 relevante de
 pronósticos y ventas

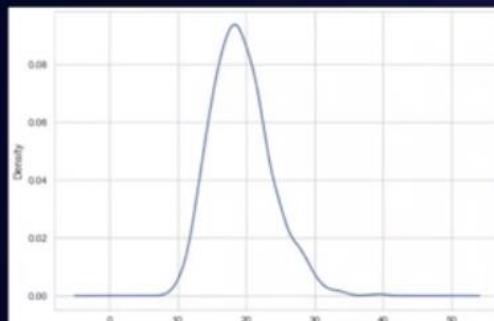


Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

La canalización de ML



Las técnicas de visualización comunes para las características numéricas son las siguientes



Gráfica de densidad

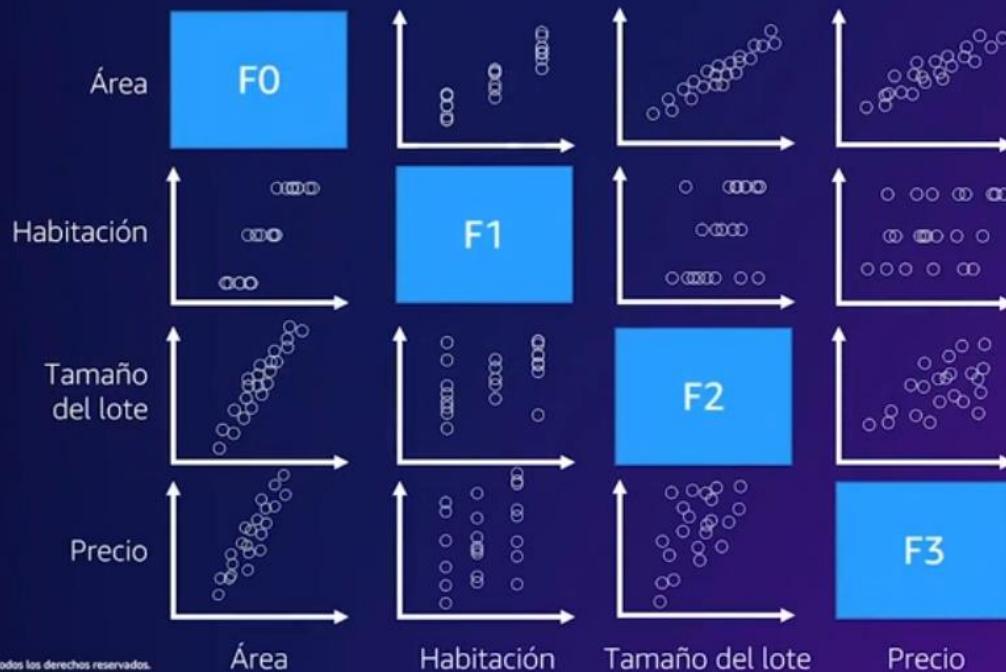


Histograma



Gráfica de caja

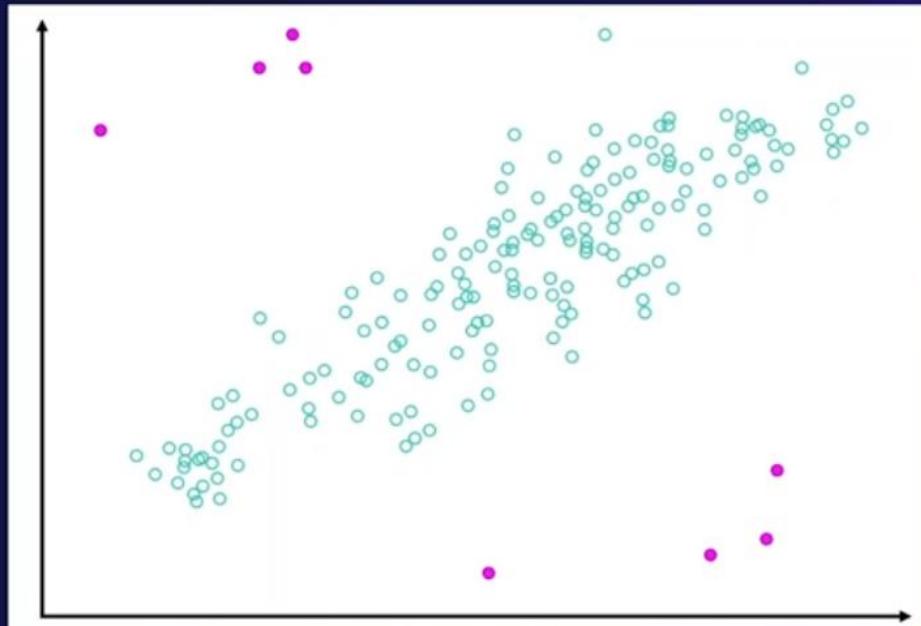
Los gráficos de dispersión visualizan las relaciones entre las variables numéricas



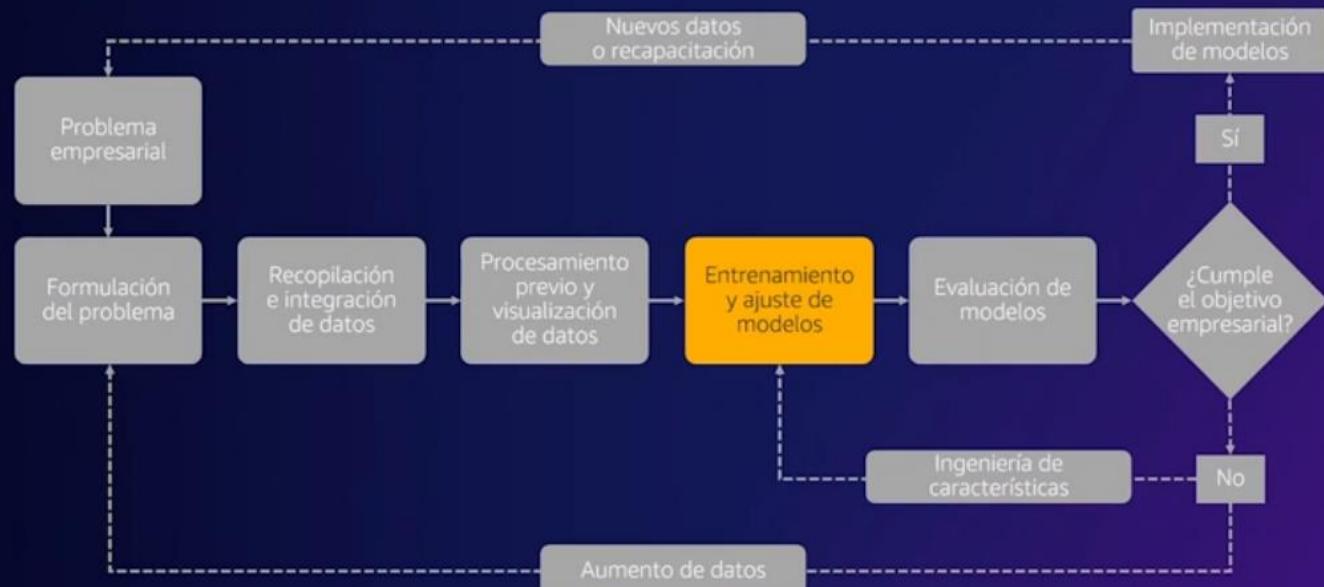
Los valores atípicos son puntos de datos que se encuentran a una distancia anormal de otros valores

Los valores atípicos pueden hacer lo siguiente:

- agregar riqueza a los datos
- hacer predicciones precisas más difíciles
- indicar que el punto de datos pertenece a otra columna



La canalización de ML



Entrenamiento del modelo

Validación:

Conjunto de entrenamiento
(80 %)

Conjunto de validación
(10 %) Conjunto de prueba
(10 %)

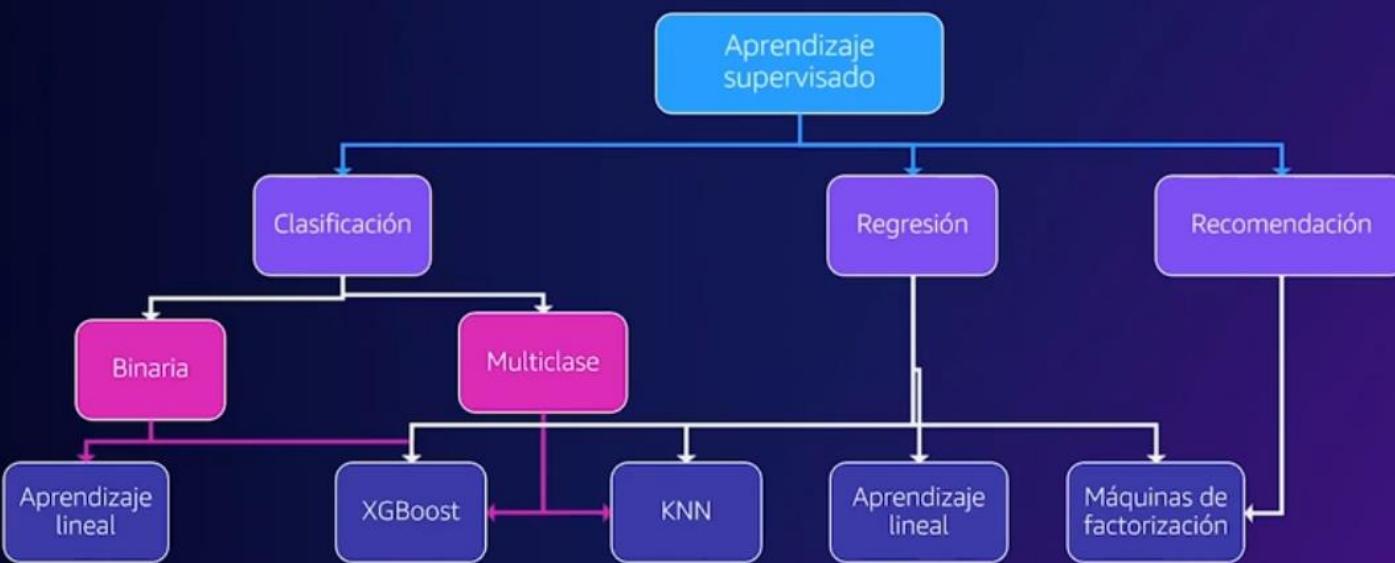
Entrenamiento

Pruebas



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Cómo utilizar un algoritmo integrado de SageMaker



Evaluación del modelo



Para los problemas de clasificación se suele utilizar una **matriz de confusión**, la cual compara los valores verdaderos con los resultados previstos

Predicciones	Valores reales
→ Gato	→ Gato
→ No es gato	→ No es gato
→ Gato	→ No es gato
→ No es Gato	→ Gato
→ No es gato	→ No es Gato
→ No es gato	→ Gato



Predicción

P

Total de
Positivos
verdaderos
(Σ TP)

N

Total de
Negativos
falsos
(Σ FN)

Valor real

N

Total de
Positivos falsos
(Σ FP)

Total de
Negativos
verdaderos
(Σ TN)

**La matriz de confusión
le proporciona
una comparación general
acerca de cómo las clases
previstas concuerdan
con las clases reales**

Métricas para evaluar el desempeño del modelo

		Clase prevista	
		P	N
Clase real	P	VP	FN
	N	FP	VN

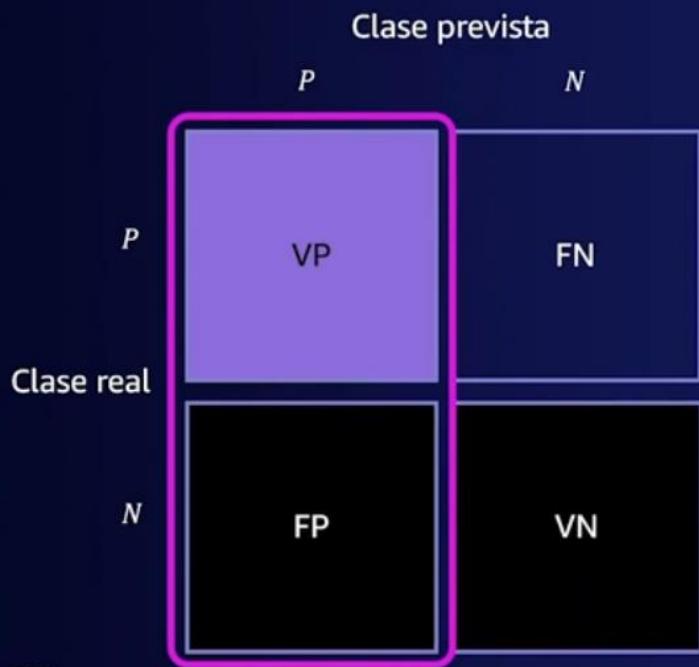
$$\text{Exactitud} = \frac{VP+VN}{TP+FP+FN+FP}$$

Accuracy

La precisión es menos efectiva cuando hay varios negativos verdaderos (p. ej.: predecir fraude con pocos datos de fraude o con ninguno)



Métricas para evaluar el desempeño del modelo



$$\text{Precisión} = \frac{VP}{VP+FP}$$

Precision

La precisión es una buena métrica cuando el costo de los falsos positivos es alto



Métricas para evaluar el desempeño del modelo

		Clase prevista	
		P	N
Clase real	P	VP	FN
	N	FP	VN

$$\text{Recuperación} = \frac{VP}{VP+FN}$$

Recall, Sensibilidad

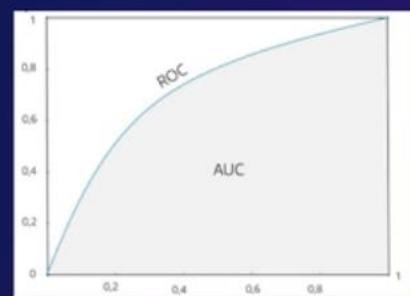
La recuperación es una buena métrica cuando el costo de los falsos negativos es alto



Métricas para evaluar el desempeño del modelo

$$\text{Puntuación F1} = \frac{2 \cdot \text{Precisión} \cdot \text{Recuperación}}{\text{Precisión} + \text{Recuperación}}$$

Área debajo la curva: curva del operador del receptor (AUC-ROC)



**Veamos un ejemplo
del centro de llamadas de Amazon**

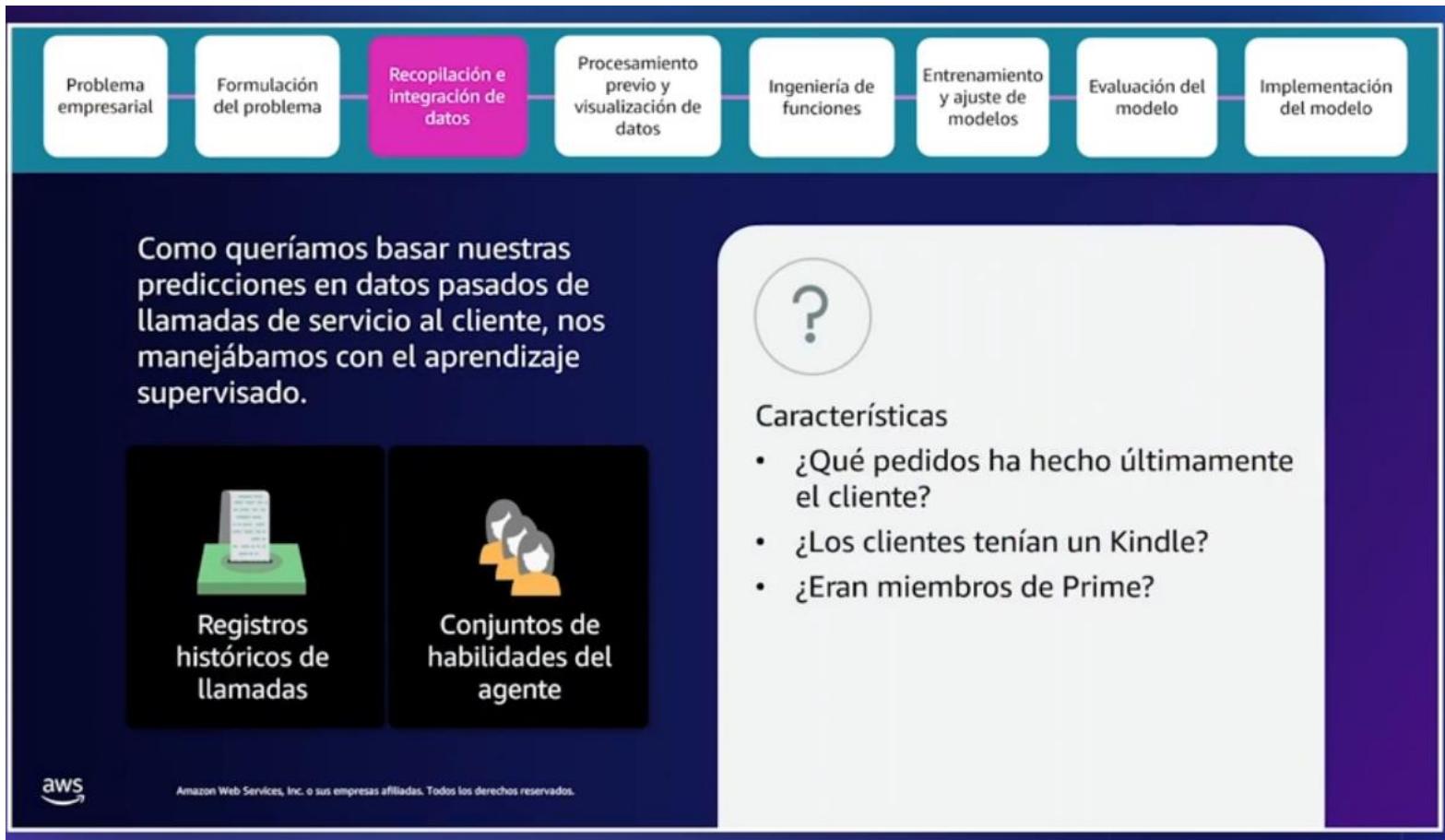


Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

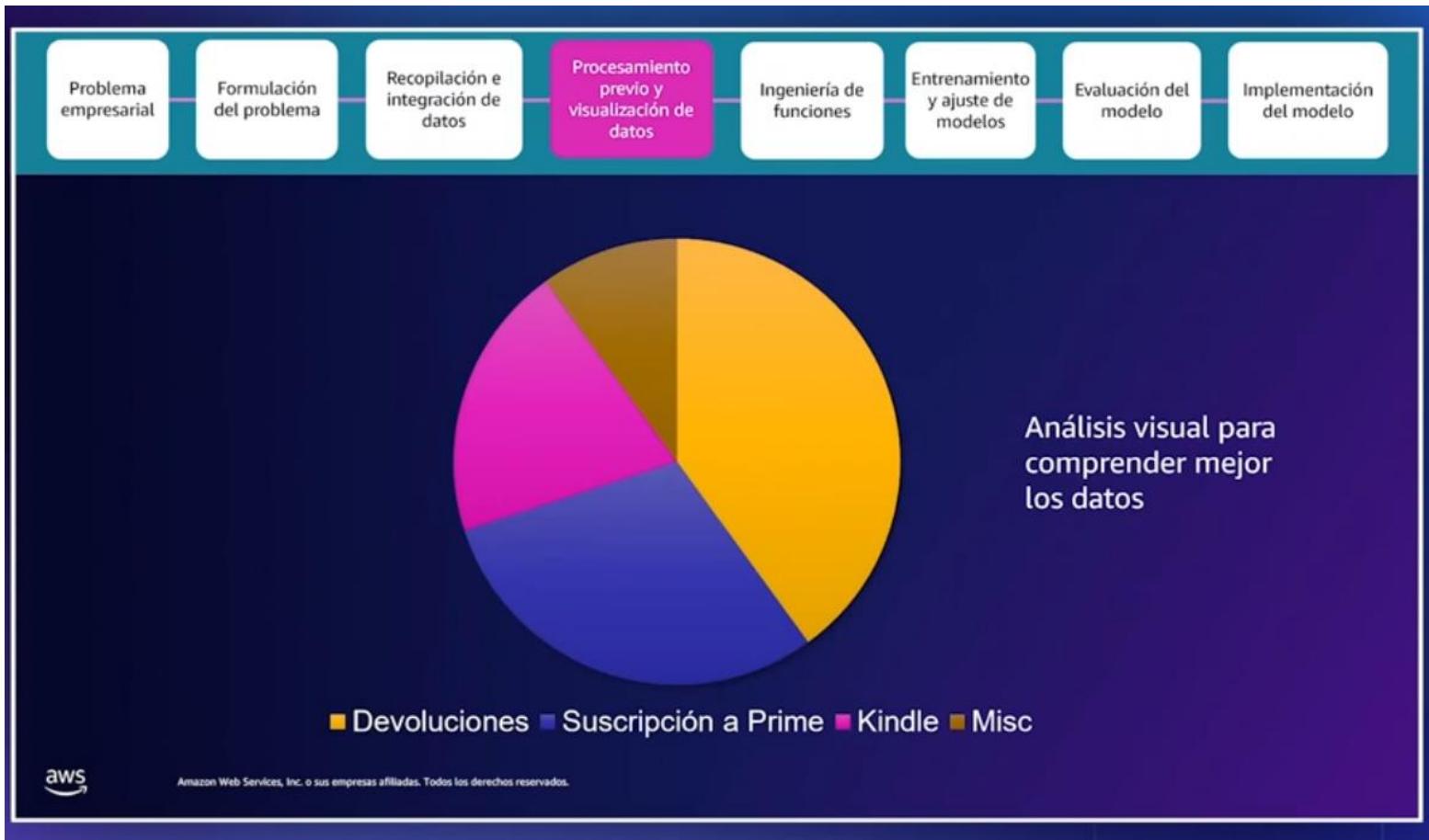
El problema del centro de llamadas de Amazon

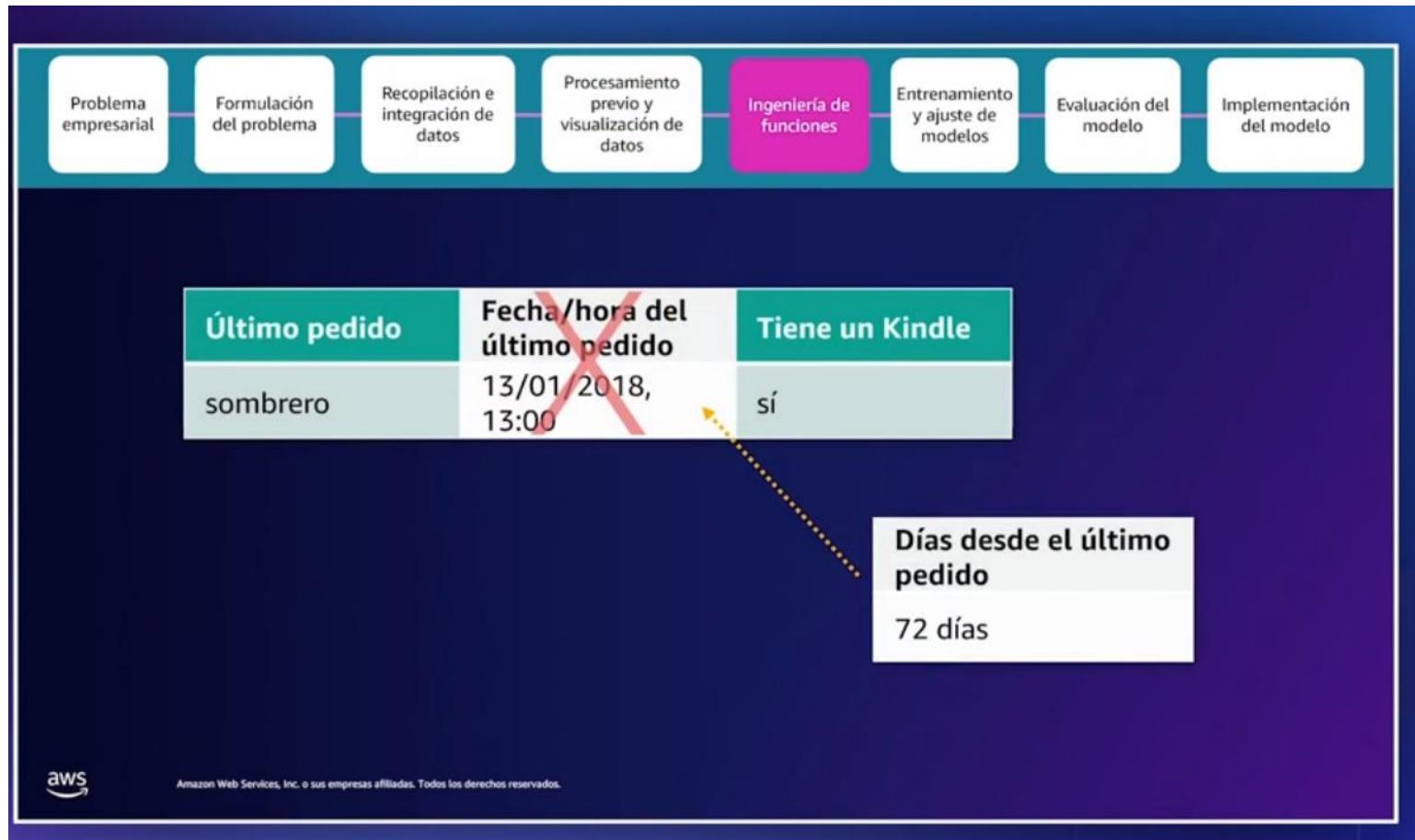




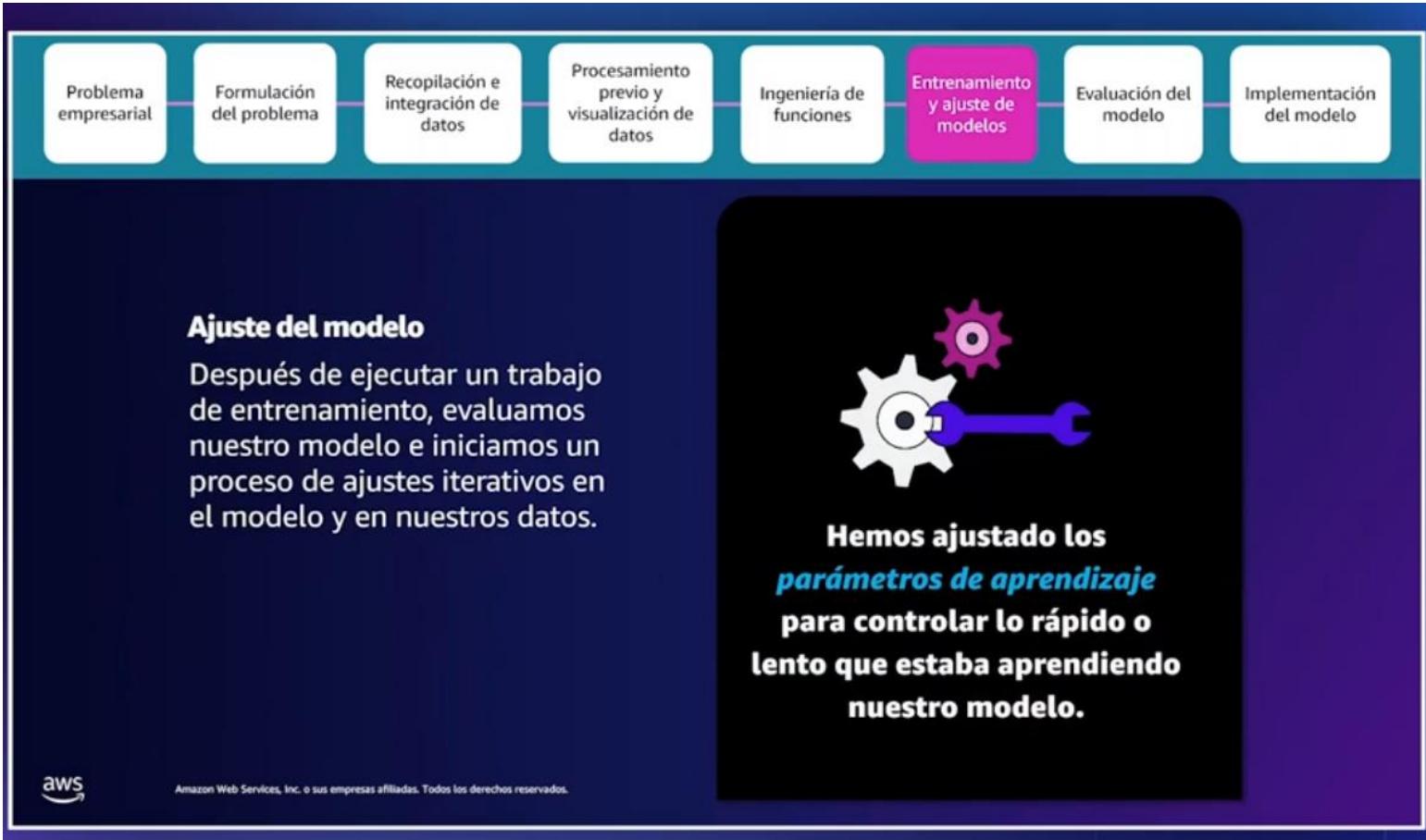


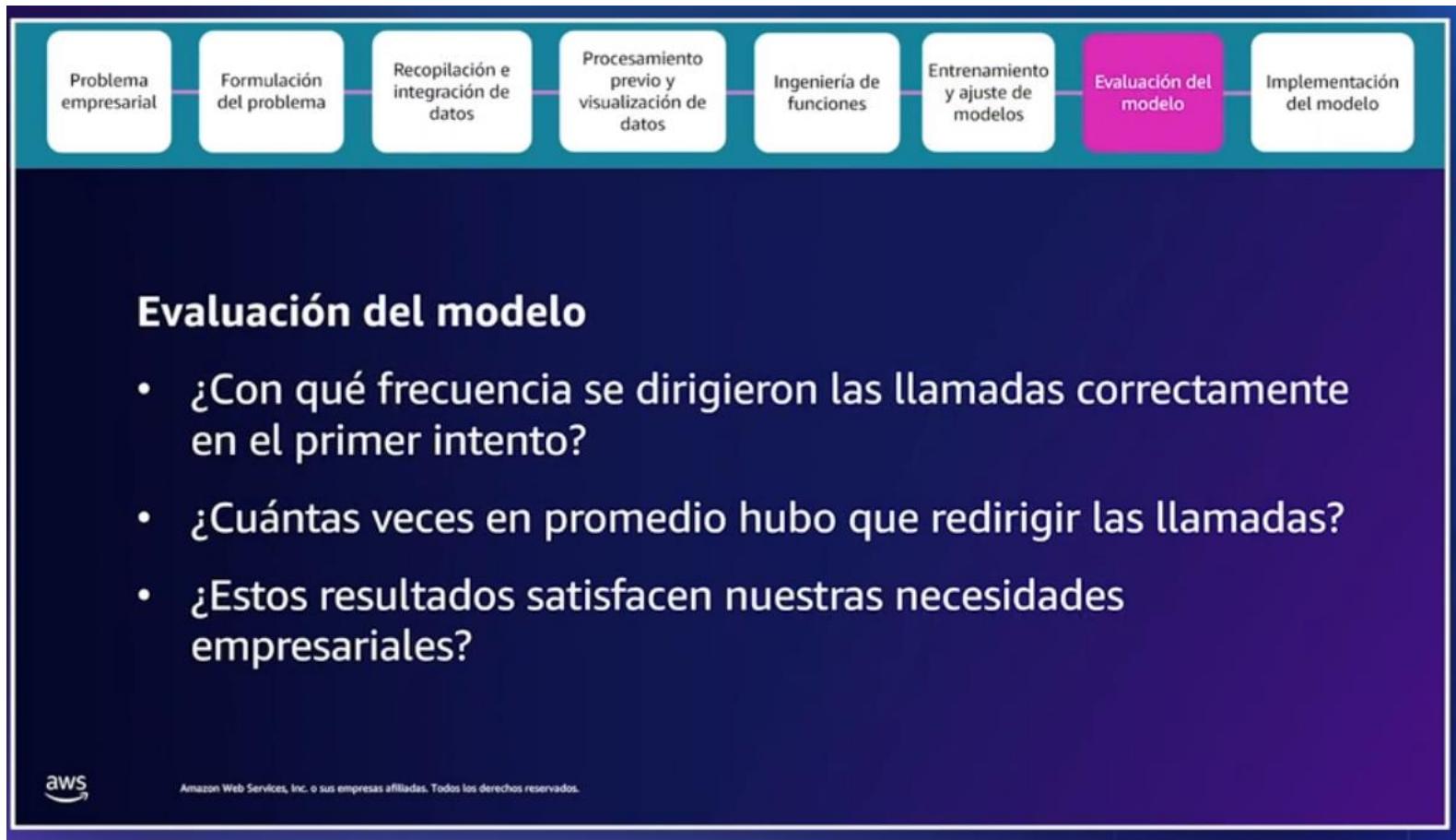


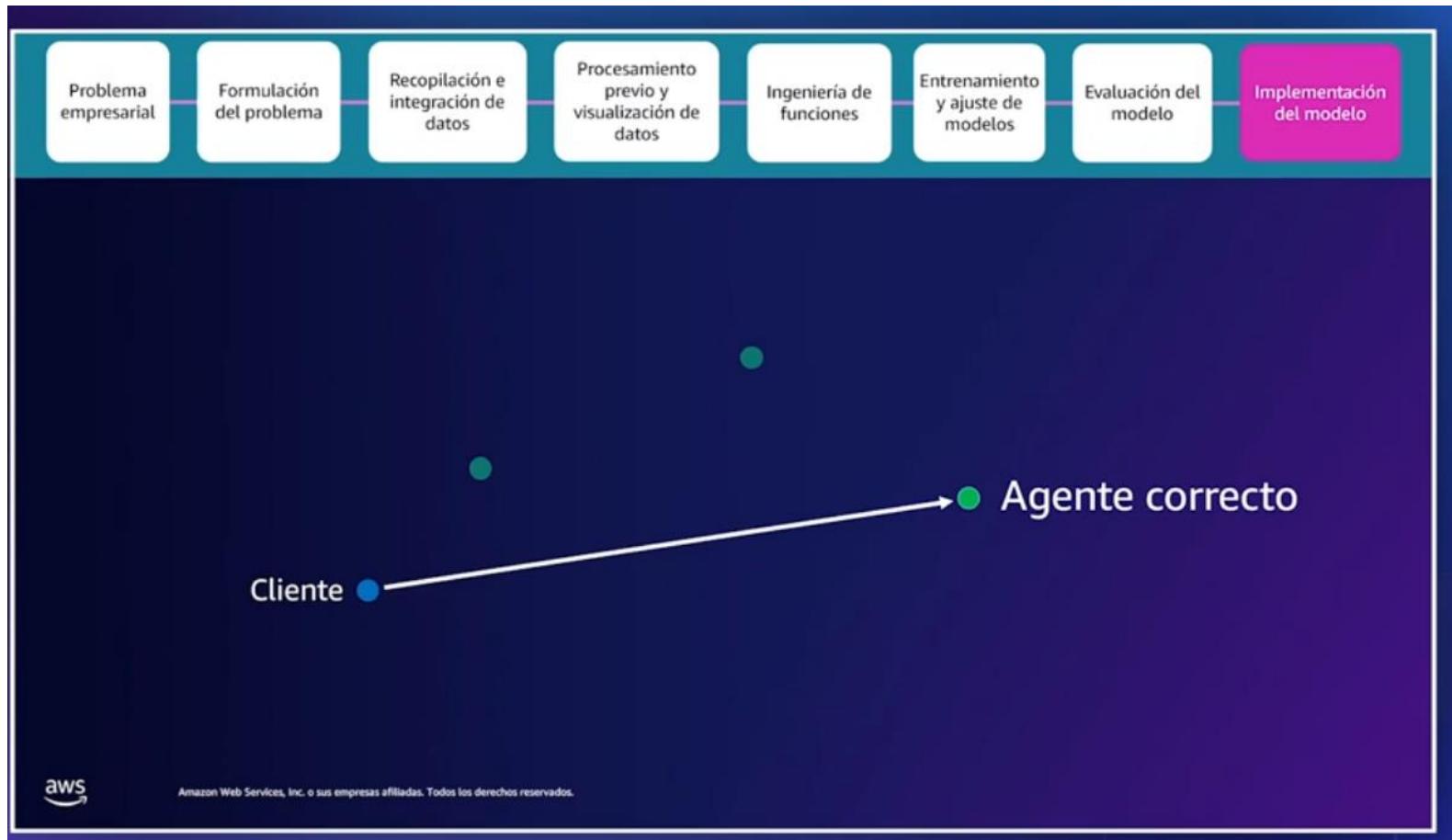




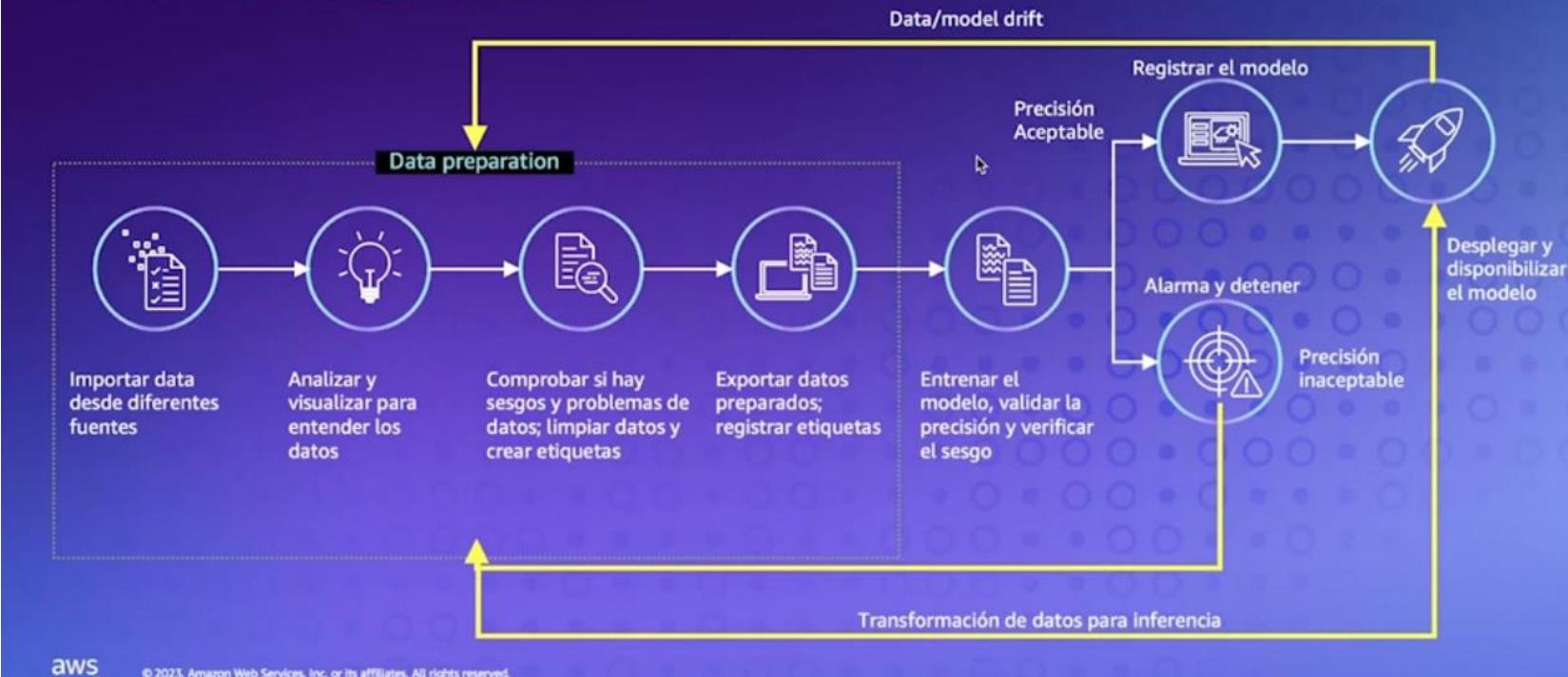








Preparación de los Datos : Parte critica del workflow de ML



Como ML impulsa hoy la creación de valor



Por lo general, lleva de semanas a meses resolver los casos de uso priorizados



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

¿Cómo podemos escalar la creación de valor?

1

Expandir el equipo de desarrollo de ML

Crecer los equipos técnicos en proporción a las necesidades, sin embargo, el talento de ML tiene una gran demanda

+74%

Crecimiento compuesto anual en los últimos 4 años

2x

el crecimiento de la demanda de cualquier otro puesto de trabajo emergente

2

Mejorar la productividad del equipo de ML

Aprovechando las herramientas low-code / no-code que hacen que los equipos de ciencia de datos sean más productivos

Habilitar a los [equipos de ciencia de datos](#) para experimenten más rápido con capacidades low-code / no-code para Machine Learning

3

Democratizar la innovación con ML

Habilitar a mas grupos de personas, incluyendo analistas de negocio a que construyan modelos de ML

Empoderar a los [analistas de negocio](#) para tomar decisiones más inteligentes con no-code Machine Learning

Desafíos que enfrentan los analistas en la construcción de ML



Los analistas carecen de experiencia profunda en ML y la curva de aprendizaje es larga

- Es necesario tener comprensión de los conceptos de ML a través de la preparación de datos, el desarrollo de modelos y la optimización
- Se requiere experiencia en la selección de la combinación correcta de ingeniería de características, tipo de modelo y técnica de optimización
- Por lo general, se necesita aprender a escribir o descifrar código



El negocio necesita explicación y validación de expertos

- Los analistas prefieren trabajar con los científicos de datos para aprender y generar confianza en el proceso, pero el tiempo de los científicos de datos es limitado y, por lo general, se dedican a algunos proyectos clave de ML
- Los analistas deben poder explicar las predicciones de los modelos de ML a los ejecutivos de negocios



Las herramientas de ML sin código disponibles tienden a carecer de transparencia

- Muchas de las opciones de ML sin código carecen de transparencia a nivel de código, lo que dificulta la inspección y la producción de modelos
- El UX para analistas y científicos de datos tiende a ser el mismo, lo que requiere que los analistas conozcan los conceptos y la jerga de ML
- Con frecuencia, las herramientas de ML sin código vienen con tarifas de licencias, por lo que la experimentación requiere de una inversión inicial



Problema de abandono (churn) de un cliente.

Ejemplo de algoritmo supervisado de clasificación binaria



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Bibliotecas usadas

NumPy

Scikit-Learn

Matplotlib

Seaborn

```
import numpy as np    from sklearn      import matplotlib.pyplot
                      import datasets   as plt
                                         import seaborn
                                         as sns
```



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

**Jupyter notebook en
una instancia de
Amazon SageMaker
(incluido en capa gratuita por 2 meses)**



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

The screenshot shows the Amazon SageMaker console interface. On the left, there is a navigation sidebar with sections like 'Getting started' (Studio, Studio Lab, Canvas, RStudio), 'Domains', 'SageMaker dashboard' (Images, Lifecycle configurations, Search), and 'JumpStart' (Foundation models, Computer vision models, Natural language processing models). The main content area is titled 'Amazon SageMaker > Notebook instances'. It displays a table of 'Notebook instances' with columns: Name, Instance, Creation time, Status, and Actions. There are three entries:

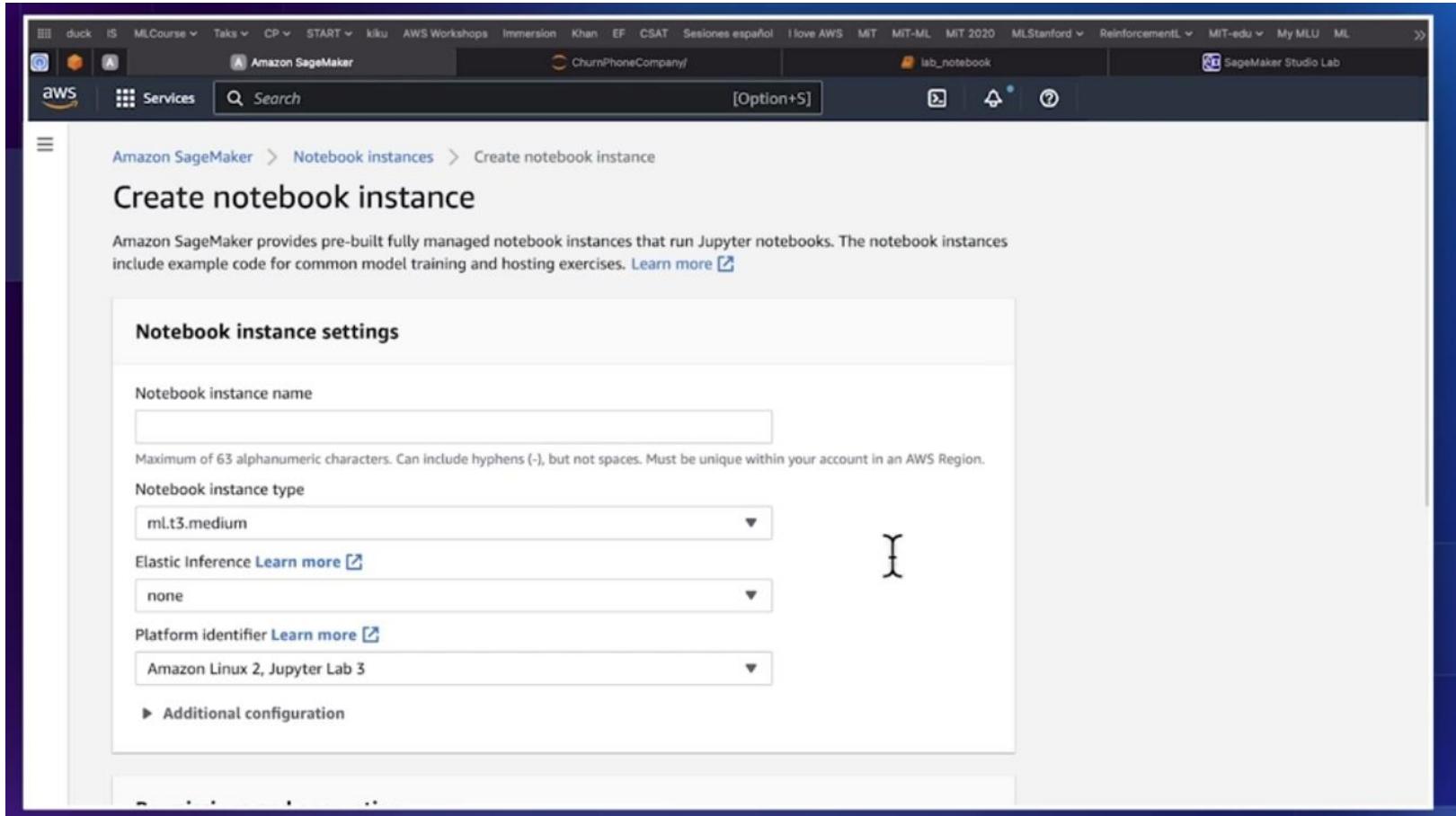
Name	Instance	Creation time	Status	Actions
notebook-nueva	ml.t3.medium	Jan 12, 2023 16:00 UTC	Stopped	Start
deeplearning-virginia-cesarml	ml.m4.xlarge	Jun 22, 2020 02:10 UTC	InService	Open Jupyter Open JupyterLab
sagemaker-virginia-cesarml	ml.t3.medium	May 29, 2020 08:33 UTC	InService	Open Jupyter Open JupyterLab

A cursor is visible near the 'notebook-nueva' entry. The top navigation bar includes tabs for 'duck', 'IS', 'MLCourse', 'Tasks', 'CP', 'START', 'kiku', 'AWS Workshops', 'Immersion', 'Khan', 'EF', 'CSAT', 'Sesiones español', 'I love AWS', 'MIT', 'MIT-ML', 'MIT 2020', 'MLStanford', 'ReinforcementL', 'MIT-edu', 'My MLU', 'ML', and 'ChurnPhoneCompany/ lab_notebook'. The search bar contains 'Search [Option+S]'.

The screenshot shows the Amazon SageMaker console interface. On the left, there is a navigation sidebar with various sections like JumpStart, Foundation models (NEW), Computer vision models, Natural language processing models, Governance, Ground Truth, Notebook (which is expanded to show Notebook instance, Git repositories, Processing, Training, Inference, Edge Manager, Augmented AI, and AWS Marketplace), and AWS Marketplace. The main content area is titled "Notebook instances" and shows a table of existing instances. The table has columns for Name, Instance, Creation time, Status, and Actions. There are three instances listed:

Name	Instance	Creation time	Status	Actions
notebook-nueva	ml.t3.medium	Jan 12, 2023 16:00 UTC	Stopped	Start
deeplearning-virginia-cesarml	ml.m4.xlarge	Jun 22, 2020 02:10 UTC	InService	Open Jupyter Open JupyterLab
sagemaker-virginia-cesarml	ml.t3.medium	May 29, 2020 08:33 UTC	InService	Open Jupyter Open JupyterLab

At the top of the main content area, there is a search bar labeled "Search notebook instances" and a "Create notebook instance" button.



The screenshot shows the AWS SageMaker console interface. At the top, there is a navigation bar with various links like 'duck', 'IS', 'MLCourse', 'Tasks', 'CP', 'START', 'kiku', 'AWS Workshops', 'Immersion', 'Khan', 'EF', 'CSAT', 'Sesiones español', 'I love AWS', 'MIT', 'MIT-ML', 'MIT 2020', 'MLStanford', 'ReinforcementL', 'MIT-edu', 'My MLU', and 'ML'. Below the navigation bar is a header with tabs for 'Amazon SageMaker', 'ChurnPhoneCompany', 'lab_notebook', and 'SageMaker Studio Lab'. The main content area has a dark background with white text. It displays the 'Notebook instance settings' page. The 'Notebook instance name' field contains the value 'mi-equipo-con-jupyter'. The 'Notebook instance type' dropdown menu is open, showing several options: 'ml.t3.medium' (selected), 'ml.t2.medium', 'ml.t2.large', 'ml.t2.xlarge', 'ml.t2.2xlarge', 'ml.t3.medium', 'ml.t3.large', 'ml.t3.xlarge' (with a hand cursor icon over it), 'ml.t3.2xlarge', 'ml.m5.xlarge', and 'ml.m5.2xlarge'. A note at the bottom right of the dropdown says 'Or let us create a role with the required permissions...'. There is also a note above the dropdown stating 'Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.'

include example code for common model training and hosting exercises. [Learn more](#)

Notebook instance settings

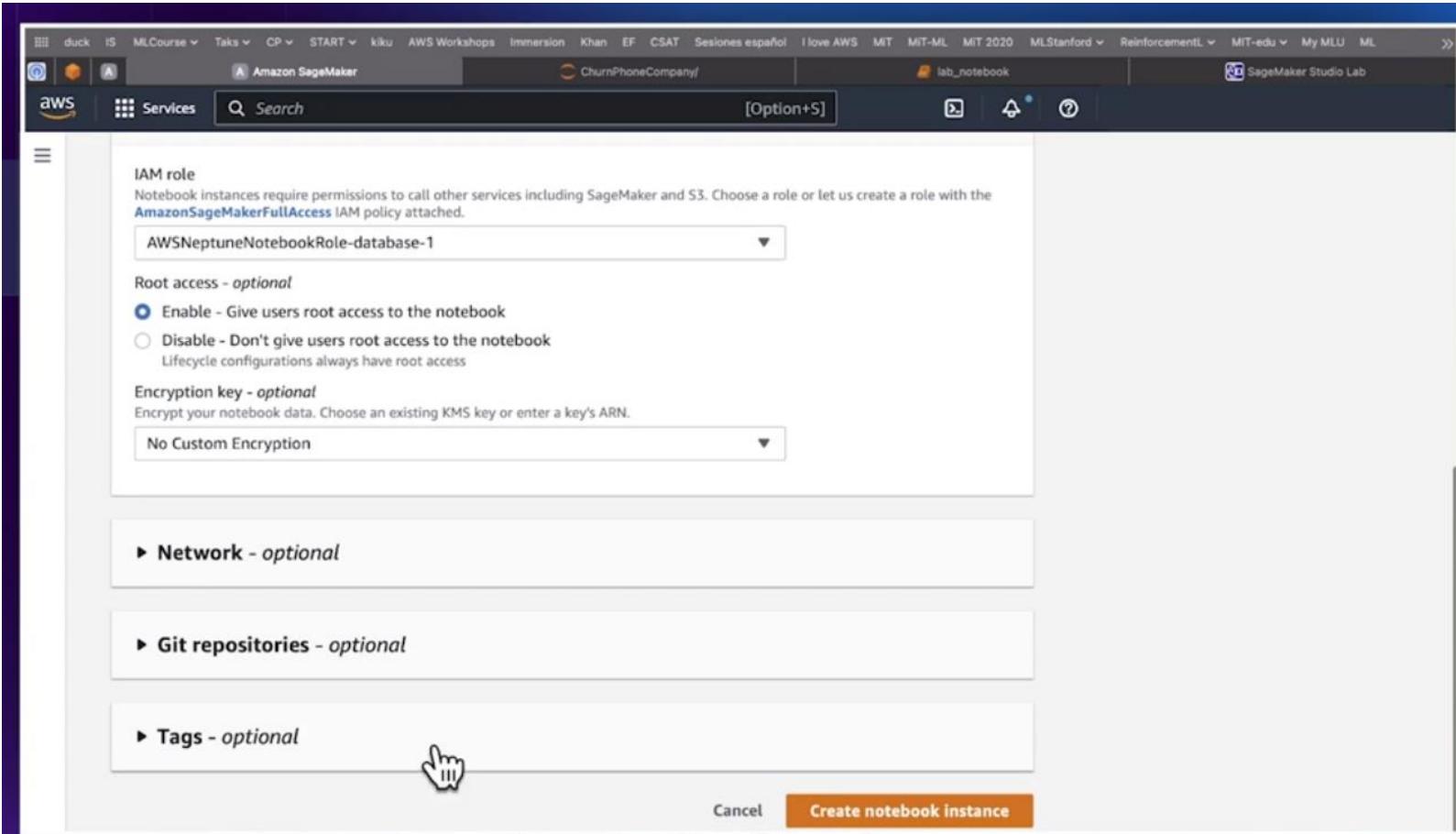
Notebook instance name

Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.

Notebook instance type

- ml.t3.medium
- ml.t2.medium
- ml.t2.large
- ml.t2.xlarge
- ml.t2.2xlarge
- ml.t3.medium
- ml.t3.large
- ml.t3.xlarge
- ml.t3.2xlarge
- ml.m5.xlarge
- ml.m5.2xlarge

Or let us create a role with the required permissions...



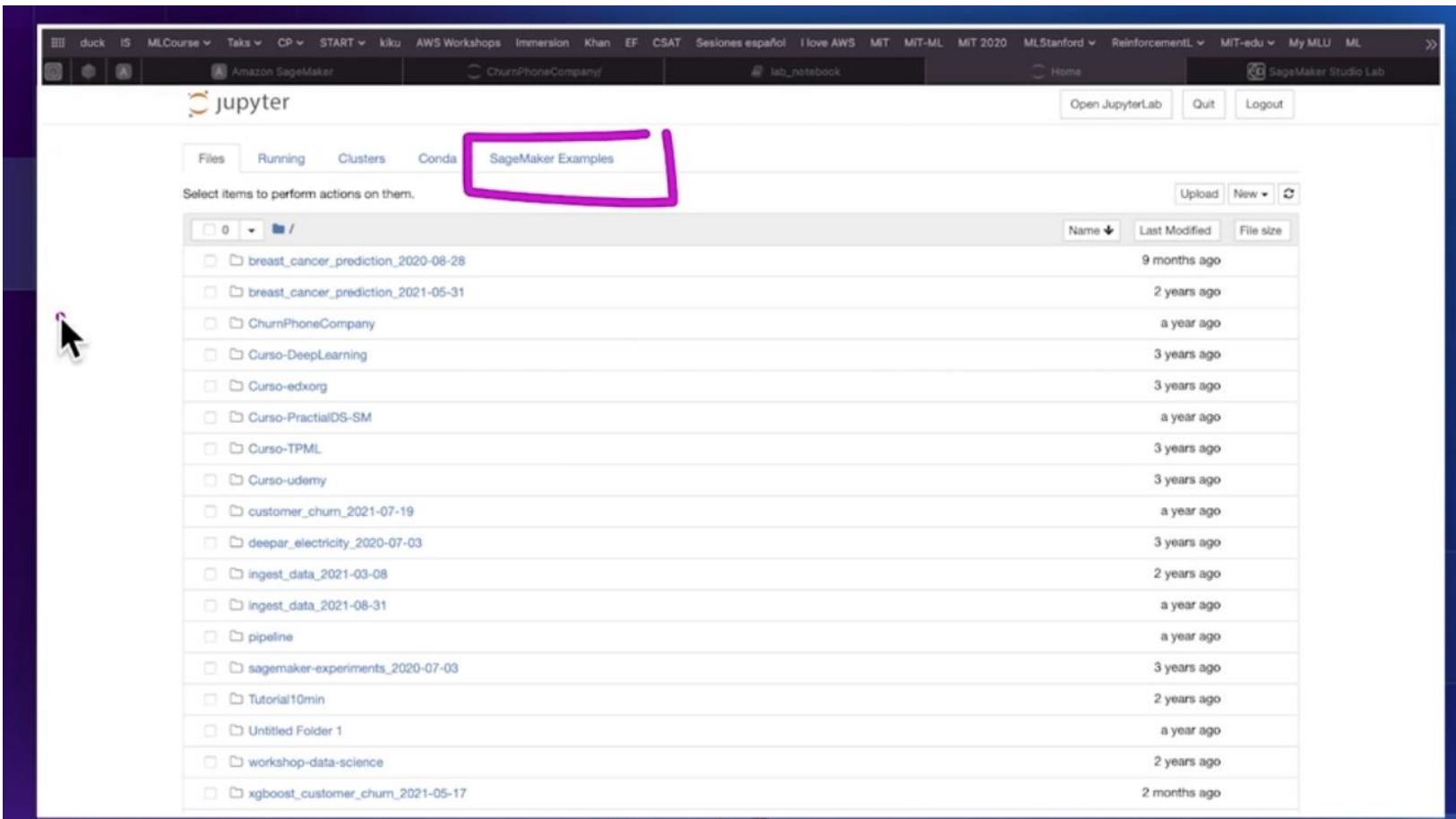
The screenshot shows the Amazon SageMaker console interface. A green success message at the top states: "Success! Your notebook instance is being created. Open the notebook instance when status is InService and open a template notebook to get started." Below this, the "Notebook instances" section displays a table of existing instances:

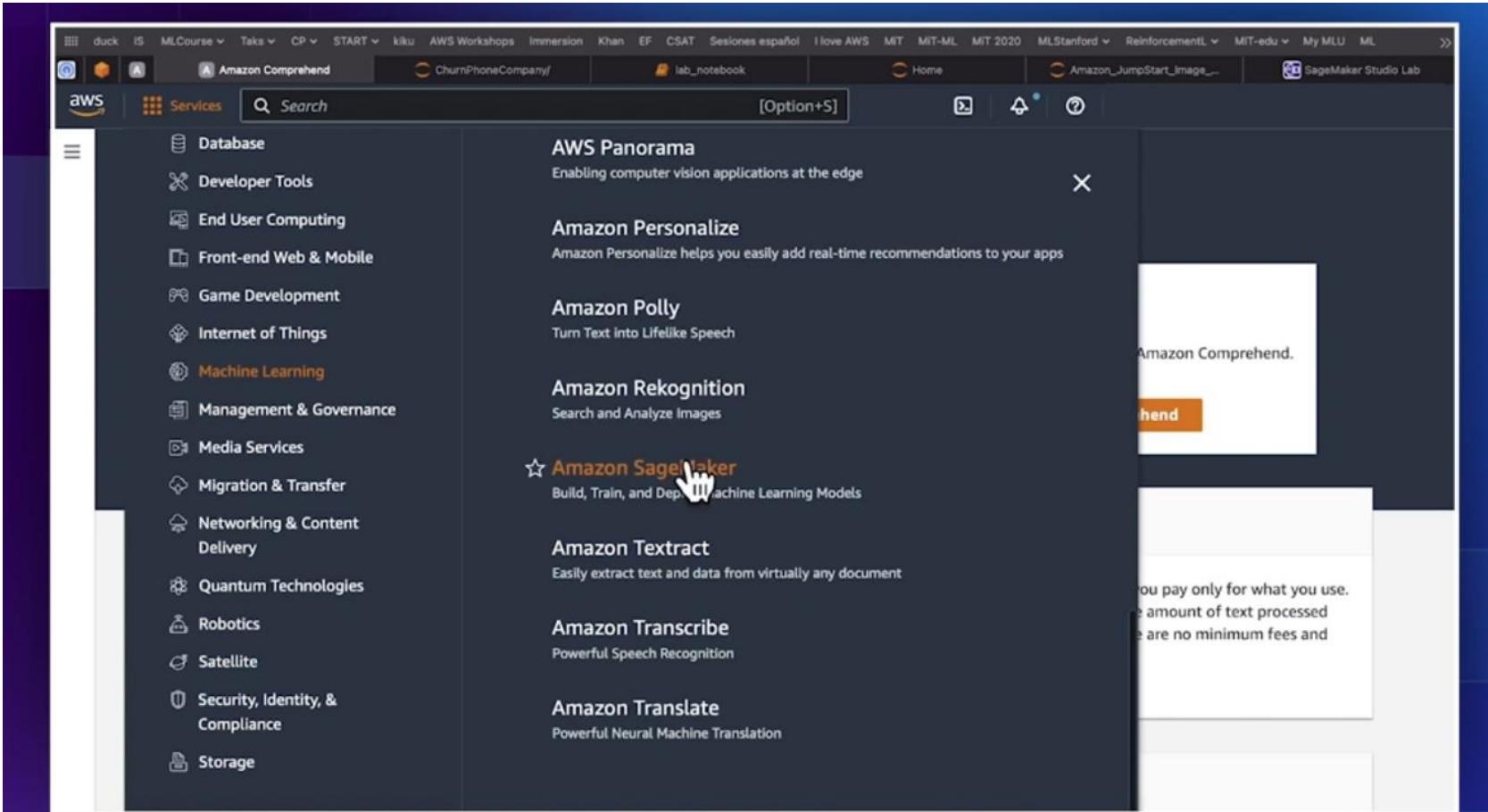
Name	Instance	Creation time	Status	Actions
mi-equipo-con-jupyter	ml.t3.medium	Jan 12, 2023 19:41 UTC	Pending	-
notebook-nueva	ml.t3.medium	Jan 12, 2023 16:00 UTC	Stopped	Start
deeplearning-virginia-cesarml	ml.m4.xlarge	Jun 22, 2020 02:10 UTC	InService	Open Jupyter Open JupyterLab
sagemaker-virginia-cesarml	ml.t3.medium	May 29, 2020 08:33 UTC	InService	Open Jupyter Open JupyterLab

The left sidebar includes links for "Getting started" (Studio, Studio Lab, Canvas, RStudio), "Domains" (highlighted with a pink box), "SageMaker dashboard", "Images", "Lifecycle configurations", "Search", and "JumpStart" (Foundation models, Computer vision models, Natural language processing models). A "Create notebook instance" button is located in the top right of the main content area.

The screenshot shows the Amazon SageMaker console interface. On the left, there's a sidebar with various navigation options like 'Getting started', 'Studio', 'Canvas', 'RStudio', 'Domains' (which is highlighted with a purple box), 'SageMaker dashboard', 'Images', 'Lifecycle configurations', 'Search', 'JumpStart' (with 'Foundation models' and 'Computer vision models' under it), and 'Natural language processing models'. The main area shows a green success message: 'Success! Your notebook instance is being created. Open the notebook instance when status is InService and open a template notebook to get started.' Below this, the 'Notebook instances' section is displayed. It has a search bar, an 'Actions' dropdown, and a 'Create notebook instance' button. A table lists four notebook instances:

Name	Instance	Creation time	Status	Actions
mi-equipo-con-jupyter	ml.t3.medium	Jan 12, 2023 19:41 UTC	Pending	-
notebook-nueva	ml.t3.medium	Jan 12, 2023 16:00 UTC	Stopped	Start
deeplearning-virginia-cesarml	ml.m4.xlarge	Jun 22, 2020 02:10 UTC	InService	Open Jupyter Open JupyterLab
sagemaker-virginia-cesarml	ml.t3.medium	May 29, 2020 08:33 UTC	InService	Open Jupyter Open JupyterLab





Instancia gratuita de dos meses

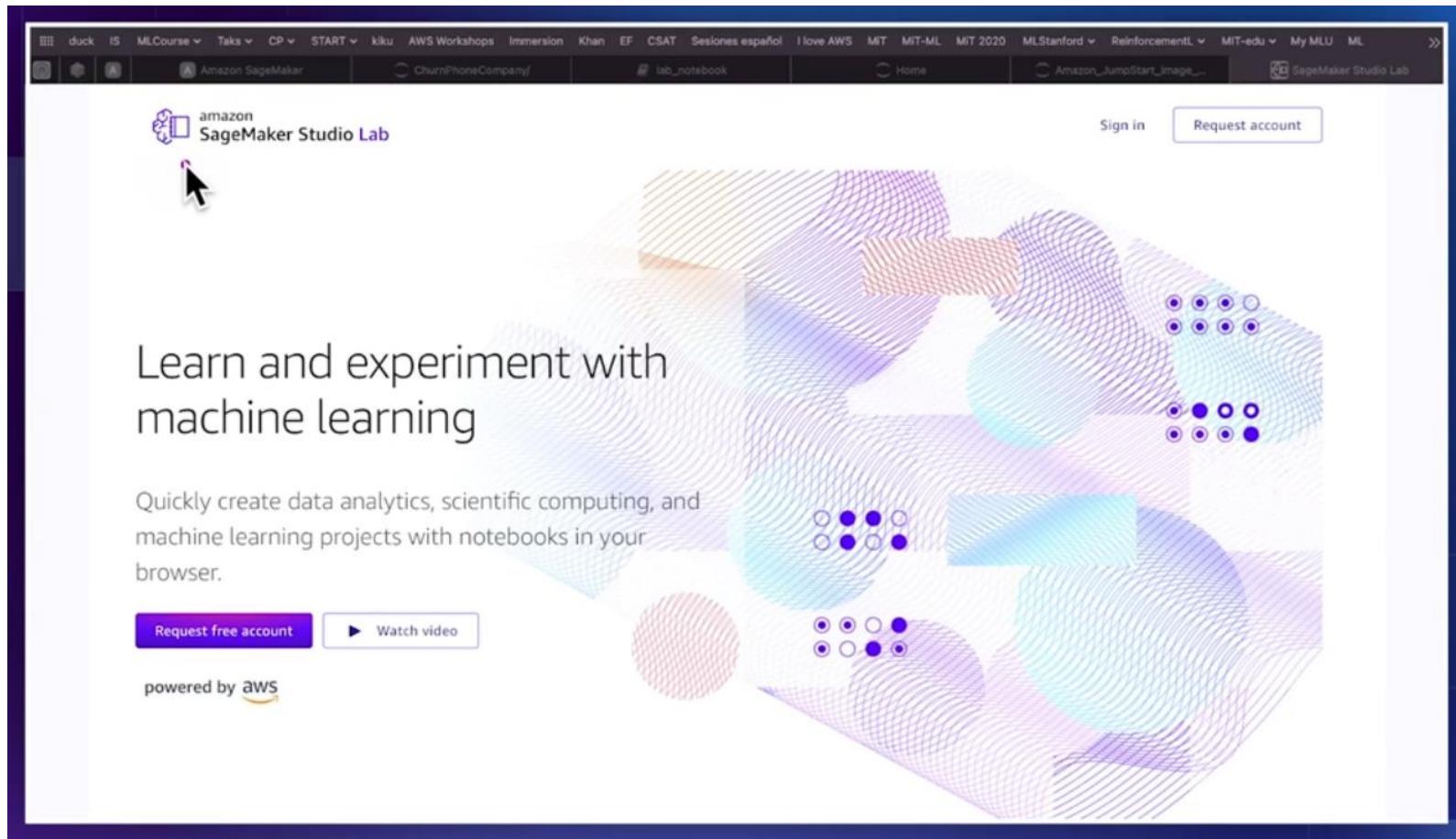
The screenshot shows the Amazon SageMaker console interface. On the left, there is a navigation sidebar with the following sections:

- SageMaker dashboard
- Images
- Lifecycle configurations
- Search
- ▼ JumpStart
 - Foundation models NEW
 - Computer vision models
 - Natural language processing models
- Governance
- Ground Truth
- ▼ Notebook
 - Notebook instances** (highlighted in blue)
 - Git repositories
- Processing
- Training
- Inference

The main content area is titled "Amazon SageMaker > Notebook instances". It displays a table of "Notebook instances" with the following columns: Name, Instance, Creation time, Status, and Actions. The table contains four rows:

Name	Instance	Creation time	Status	Actions
mi-equipo-con-jupyter	ml.t3.medium	Jan 12, 2023 19:41 UTC	InService	Open Jupyter Open JupyterLab
notebook-nueva	ml.t3.medium	Jan 12, 2023 16:00 UTC	Stopped	Start
deeplearning-virginia-cesarml	ml.m4.xlarge	Jun 22, 2020 02:10 UTC	InService	Open Jupyter Open JupyterLab
sagemaker-virginia-cesarml	ml.t3.medium	May 29, 2020 08:33 UTC	InService	Open Jupyter Open JupyterLab

A magenta box highlights the "ml.t3.medium" instance in the fourth row. A mouse cursor is positioned over the "Actions" column for this row.



The screenshot shows the landing page of the Amazon SageMaker Studio Lab. At the top, there is a navigation bar with various tabs like 'duck', 'IS', 'MLCourse', 'Taks', 'CP', 'START', 'kiku', 'AWS Workshops', 'Immersion', 'Khan', 'EF', 'CSAT', 'Sesiones español', 'I love AWS', 'MIT', 'MIT-ML', 'MIT 2020', 'MLStanford', 'ReinforcementL', 'MIT-edu', 'My MLU', and 'ML'. Below the navigation bar, the main heading is 'amazon SageMaker Studio Lab' with a logo. To the right are 'Sign in' and 'Request account' buttons. A large, stylized graphic of overlapping colored ellipses and data points (circles) serves as the background. Overlaid on this graphic is the text 'Learn and experiment with machine learning' and 'Quickly create data analytics, scientific computing, and machine learning projects with notebooks in your browser.' At the bottom left are 'Request free account' and 'Watch video' buttons. The bottom right corner features the AWS logo with the text 'powered by AWS'.

Sign in [Request account](#)

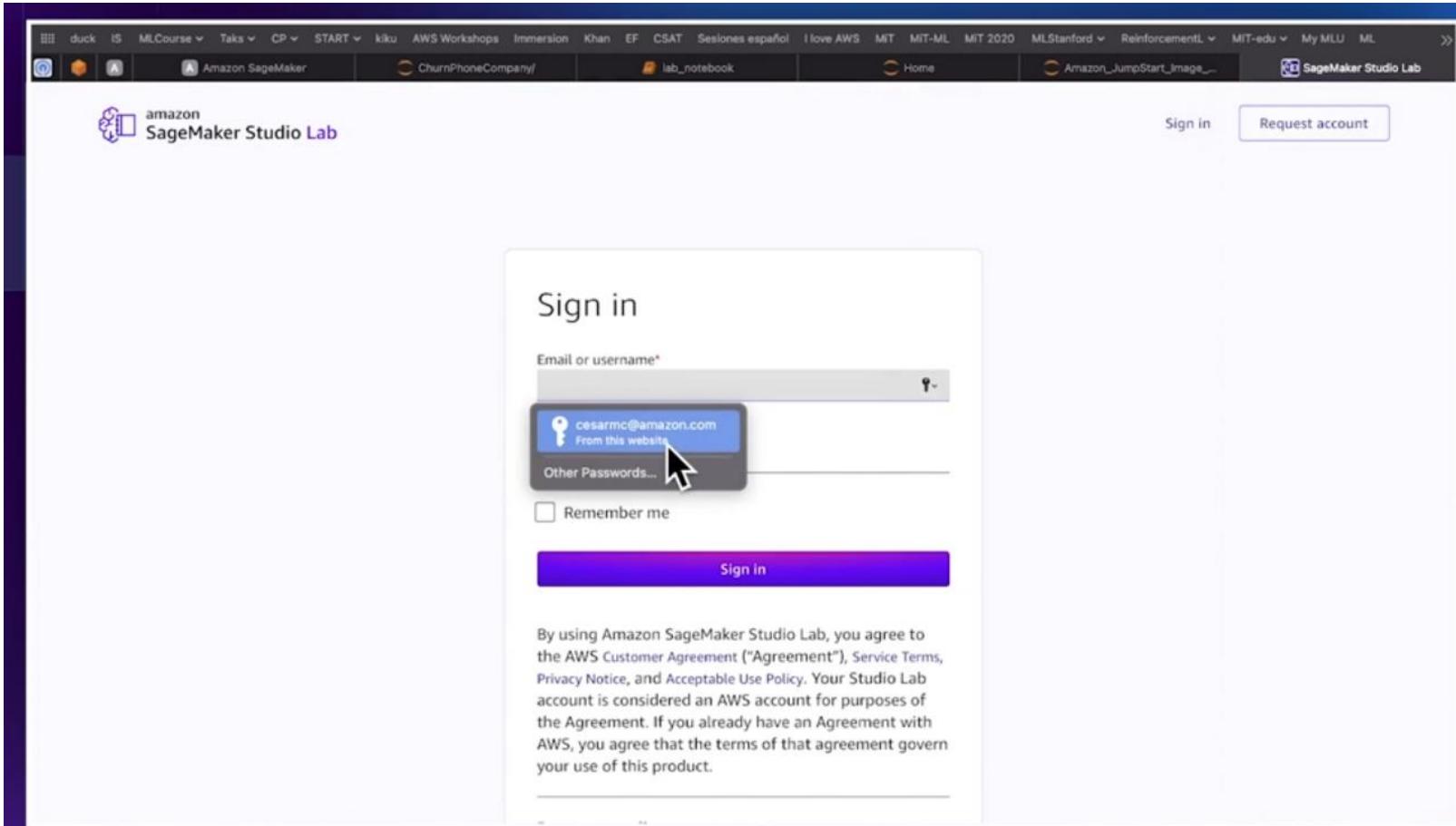
amazon SageMaker Studio Lab

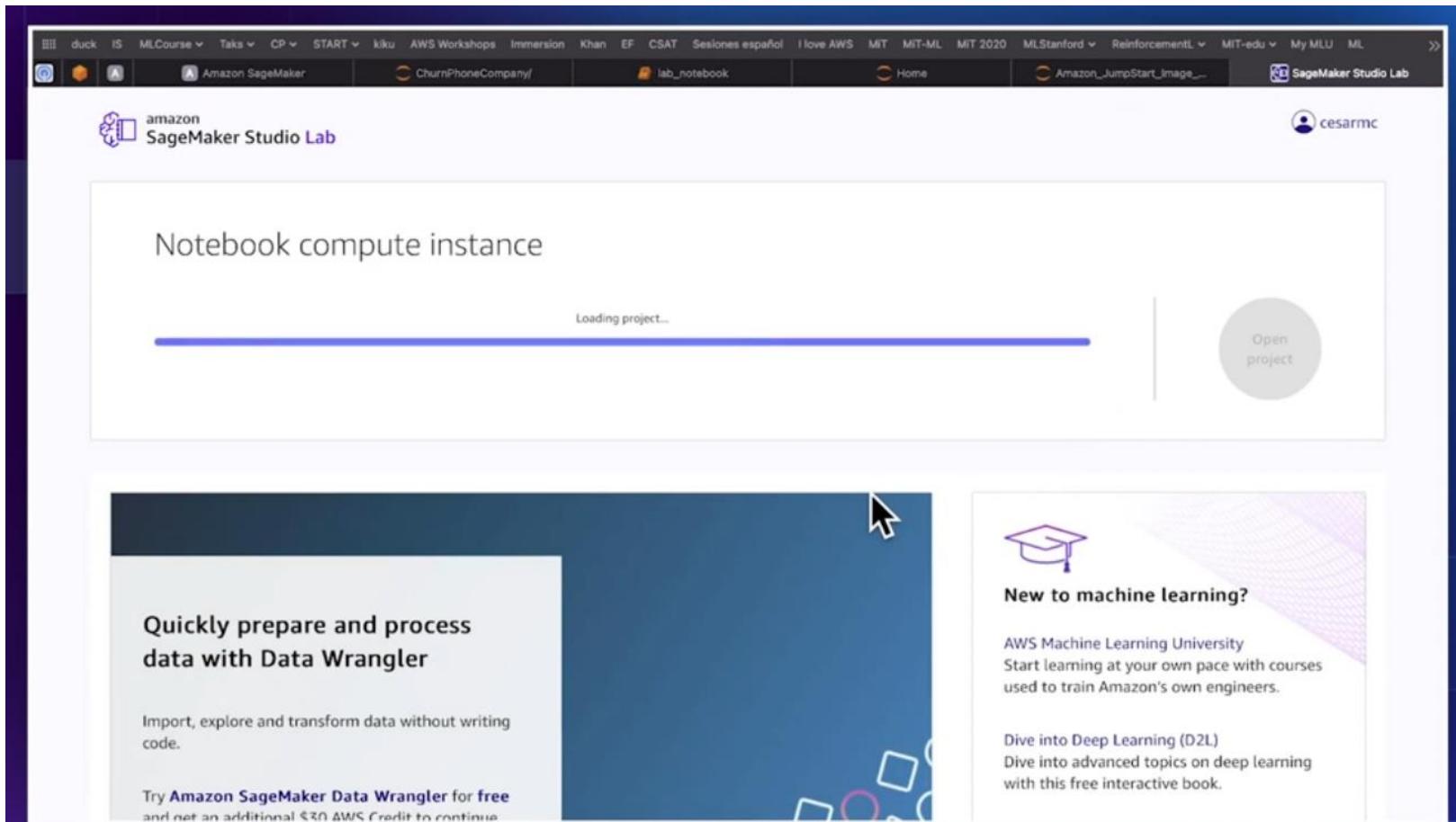
Learn and experiment with machine learning

Quickly create data analytics, scientific computing, and machine learning projects with notebooks in your browser.

[Request free account](#) [Watch video](#)

powered by







Amazon SageMaker

ChurnPhoneCompany/

lab_notebook

Home

Amazon_JumpStart_Image_...

SageMaker Studio Lab



cesarmc

My project



SageMaker Studio Lab now requires users to provide their phone number for compute. See FAQ for details.



Runtime status

Stopped

Runtime remaining

—

Compute type

 CPU GPU

Quickly prepare and process
data with Data Wrangler

Import, explore and transform data without writing
code



New to machine learning?

AWS Machine Learning University
Start learning at your own pace with courses
used to train Amazon's own engineers.

Dive into Deep Learning (D2L)

que esto decíamos

The screenshot shows the Amazon SageMaker Studio Lab interface. At the top, there is a navigation bar with various links like 'duck', 'IS', 'MLCourse', 'Taks', 'CP', 'START', 'kiku', 'AWS Workshops', 'Immersion', 'Khan', 'EF', 'CSAT', 'Sesiones español', 'I love AWS', 'MIT', 'MIT-ML', 'MIT 2020', 'MLStanford', 'ReinforcementL', 'MIT-edu', 'My MLU', and 'ML'. Below the navigation bar, the main content area has a header 'amazon SageMaker Studio Lab' and a user profile icon for 'cesarmc'. The main section is titled 'My project' and displays runtime information: 'Runtime status: Running', 'Runtime remaining: 11 h 59 m', 'Compute type: CPU' (selected), and a 'Stop runtime' button. To the right of this is a large purple circular button with a white hand icon and the text 'Open project'. Below this, there are two promotional cards. The left card is for 'Data Wrangler' with the text 'Quickly prepare and process data with Data Wrangler', 'Import, explore and transform data without writing code.', and a call-to-action 'Try Amazon SageMaker Data Wrangler for free and get an additional \$20 AWS Credit to continue.' The right card is for 'AWS Machine Learning University' with the text 'New to machine learning?', 'Start learning at your own pace with courses used to train Amazon's own engineers.', and a call-to-action 'Dive into Deep Learning (D2L) Dive into advanced topics on deep learning with this free interactive book.'

duck IS MLCourse Taks CP START kiku AWS Workshops Immersion Khan EF CSAT Sesiones español I love AWS MIT MIT-ML MIT 2020 MLStanford ReinforcementL MIT-edu My MLU ML

Amazon SageMaker ChurnPhoneCompany/ lab_notebook Home Amazon_JumpStart_Image_ SageMaker Studio Lab

amazon SageMaker Studio Lab cesarmc

My project

Runtime status: Running Runtime remaining: 11 h 59 m Compute type: CPU

CPU GPU

Stop runtime

Open project


Quickly prepare and process data with Data Wrangler
Import, explore and transform data without writing code.
Try Amazon SageMaker Data Wrangler for free and get an additional \$20 AWS Credit to continue.


New to machine learning?
AWS Machine Learning University
Start learning at your own pace with courses used to train Amazon's own engineers.
Dive into Deep Learning (D2L)
Dive into advanced topics on deep learning with this free interactive book.

The screenshot shows the Amazon SageMaker Studio Lab interface. The top navigation bar includes links like duck, IS, MLCourse, Tasks, CP, START, kiku, AWS Workshops, Immersion, Khan, EF, CSAT, Sesiones español, I love AWS, MIT, MIT-ML, MIT 2020, MLStanford, ReinforcementL, MIT-edu, My MLU, and ML. The main area has tabs for ChurnPhoneCompany, lab_notebook, Home, Amazon_JumpStart, SageMaker Studio Lab, and linear-regre... - Jupyter. A file browser on the left shows a directory structure under /sagemaker-studiolab-notebooks/d2l-pytorch-sagemaker-studio-lab/. The current notebook, linear-regression.ipynb, is open in a code editor. The content starts with a section on Linear Regression:

3.1 Linear Regression

Regression refers to a set of methods for modeling the relationship between one or more independent variables and a dependent variable. In the natural sciences and social sciences, the purpose of regression is most often to *characterize* the relationship between the inputs and outputs. Machine learning, on the other hand, is most often concerned with *prediction*.

Regression problems pop up whenever we want to predict a numerical value. Common examples include predicting prices (of homes, stocks, etc.), predicting length of stay (for patients in the hospital), demand forecasting (for retail sales), among countless others. Not every prediction problem is a classic regression problem. In subsequent sections, we will introduce classification problems, where the goal is to predict membership among a set of categories.

3.1.1 Basic Elements of Linear Regression

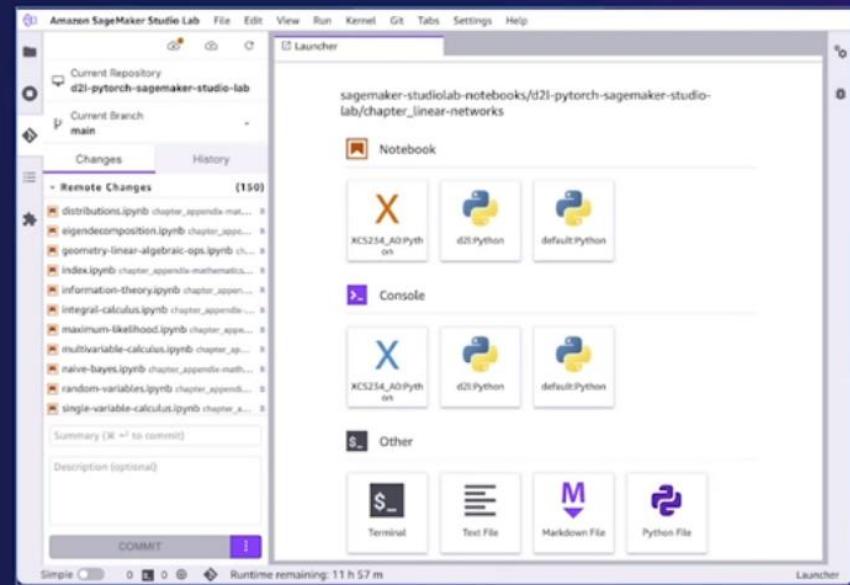
Linear regression may be both the simplest and most popular among the standard tools to regression. Dating back to the dawn of the 19th century, linear regression flows from a few simple assumptions. First, we assume that the relationship between the independent variables x and the dependent variable y is linear, i.e., that y can be expressed as a weighted sum of the elements in x , given some noise on the observations. Second, we assume that any noise is well-behaved (following a Gaussian distribution).

To motivate the approach, let us start with a running example. Suppose that we wish to estimate the prices of houses (in dollars) based on their area (in square feet) and age (in years). To actually develop a model for predicting house prices, we would need to get our hands on a dataset consisting of sales for which we know the sale price, area, and age for each home. In the terminology of machine learning, the dataset is called a *training dataset* or *training set*, and each row (here the data corresponding to one sale) is called an *example* (or *data point*, *data instance*, *sample*). The thing we are trying to predict (price) is called a *label* (or *target*). The independent variables (age and area) upon which the predictions are based are called *features* (or *covariates*).

Typically, we will use n to denote the number of examples in our dataset. We index the data examples by i , denoting each input as $\mathbf{x}^{(i)} = [x_1^{(i)}, x_2^{(i)}]^T$ and the corresponding label as $y^{(i)}$.

Amazon SageMaker Studio Lab

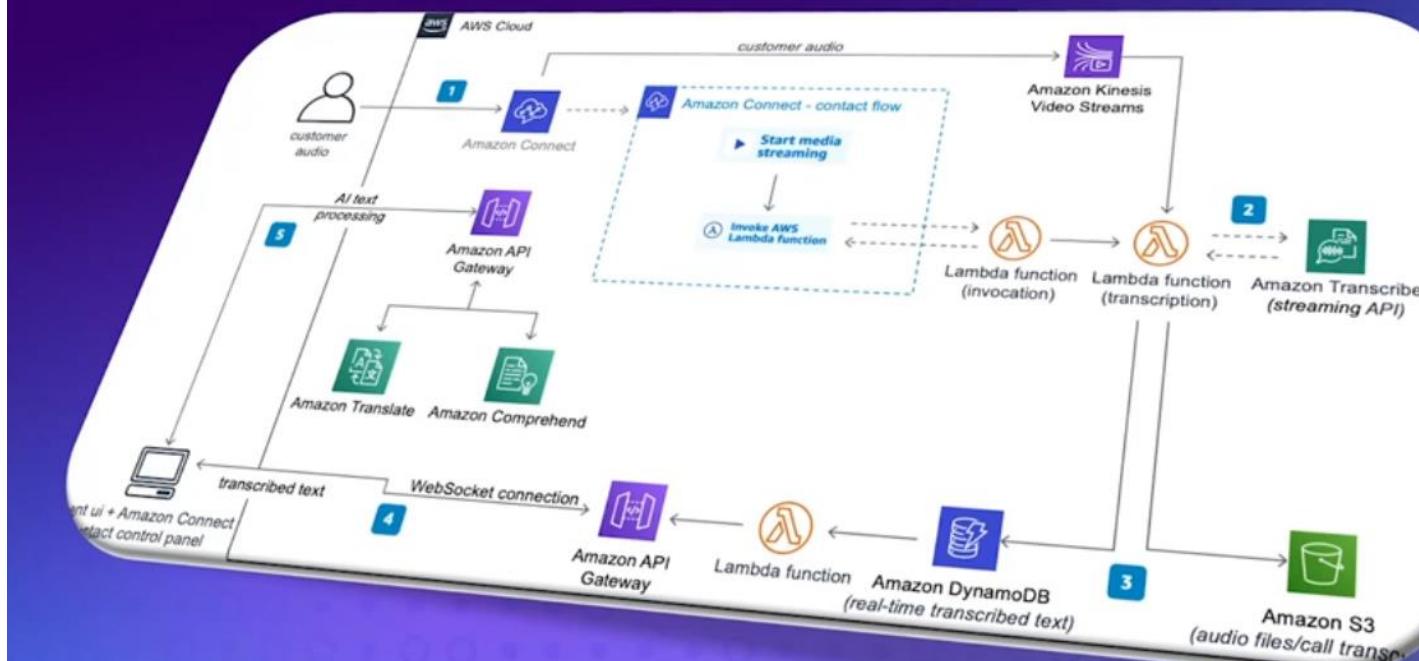
Crea y ejecuta código en
Amazon SageMaker Studio Lab
sin costo



Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

Caso de uso: Servicio al cliente

Procesamiento de llamadas



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.