Pràctica 4: Azure

JORGE OSARENKHOE PETRO

Introducció

En aquesta pràctica, treballarem amb Azure Machine Learning per crear un model de classificació utilitzant dades de pingüins. A més, explorarem el servei de veu d'Azure Al per convertir veu a text i text a veu. Aquesta pràctica està dividida en dos apartats principals:

- 1. Classificació de pingüins amb Azure Machine Learning
- 2. Servei de veu d'Azure Al

Objectius

- Utilitzar Azure Machine Learning per entrenar i desplegar un model de classificació.
- Documentar tot el procés d'entrenament, selecció de models, i desplegament del model.
- Utilitzar el servei de veu d'Azure Al per convertir veu a text i text a veu.

Requisits

- Compte d'Azure actiu.
- Coneixements bàsics de Machine Learning i Python.
- Accés a les dades dels pingüins (penguins_size.csv).

Estructura del document

Aquest document està estructurat de la següent manera:

1. Apartat 1: Classificació de pingüins amb Azure Machine Learning

- Creació del recurs
- Creació del workspace
- o Importació del dataset
- o Entrenament del model
- Visualització de resultats
- Neteja dels recursos

2. Apartat 2: Servei de veu d'Azure Al

- o Conversió de veu a text en temps real
- o Conversió de text a veu utilitzant la galeria de veus

Nota

Assegureu-vos de seguir tots els passos indicats i de documentar cada part del procés amb captures de pantalla i explicacions detallades. Això us ajudarà a entendre millor cada etapa i a tenir una referència clara per a futures pràctiques.

Apartat 1: Classificació de pingüins amb Azure Machine Learning

Tornarem a treballar amb un classificador sobre el conjunt de dades dels pingüins de l'arxipèlag Palmer. Recordau que podeu trobar les dades originals a Kaggle (fitxer penguins_size.csv) i una còpia al GitHub del curs.

Recordem que aquestes són les dades de cada individu:

```
species (variable objectiu, l'espècie de pingüí: Chinstrap, Adélie o Gentoo)
island (illa: Dream, Torgersen o Biscoe)
culmen_length_mm (longitud del bec en mm)
culmen_depth_mm (profunditat del bec en mm)
flipper_length_mm (longitud de l'aleta en mm)
body_mass_g (massa corporal en grams)
sex (sexe: Male o Female)
```

Per fer-ho, heu de treballar amb Azure Machine Learning amb l'Azure Machine Learning Studio i les funcionalitats de Machine Learning automatitzat, d'una manera semblant a com ho hem fet en el cas pràctic 1 dels apunts. Ara, però, es tracta d'un problema de classificació, no de regressió.

Per fer el treball de ML automatitzat heu d'emprar l'exactitud (accuracy). I heu de considerar els següents models que hem vist en el curs: LogisticRegression, SVM, KNN, DecisionTree i RandomForest. Posau els mateixos límits a l'entrenament dels models que hem emprat en el cas pràctic.

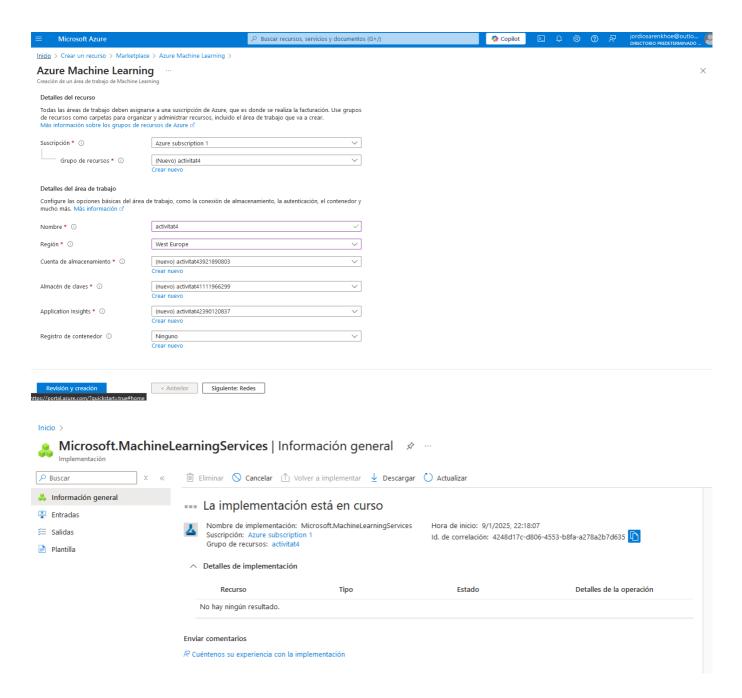
Documenta, incloent-hi explicacions i captures de pantalla, tot el procés, passa a passa. Explica quin és el model que s'ha seleccionat i mostra'n els detalls, incloent-hi els hiperparàmetres i totes les mètriques (amb gràfics). Mostra també la matriu de confusió i comenta-la.

Una vegada tenguis el model, desplega'l en un servei web i envia-li almenys dues peticions de prova, que donin com a resultat flors diferents. Inclou tot això també en la teva documentació.

Per últim, quan ho hagis acabat tot, elimina tots els recursos. Documenta també aquest procés de neteja.

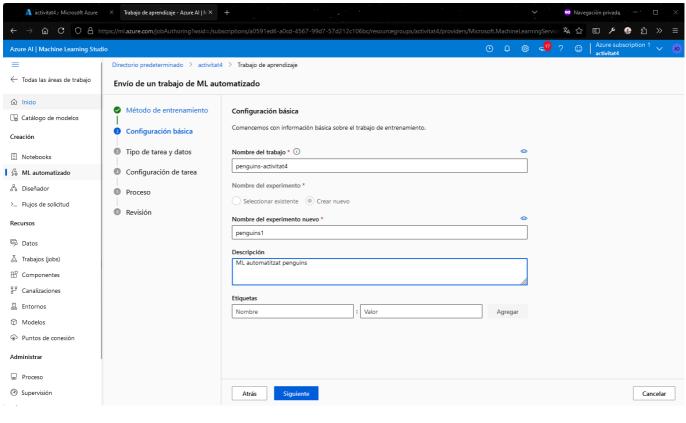
Creació del recurs

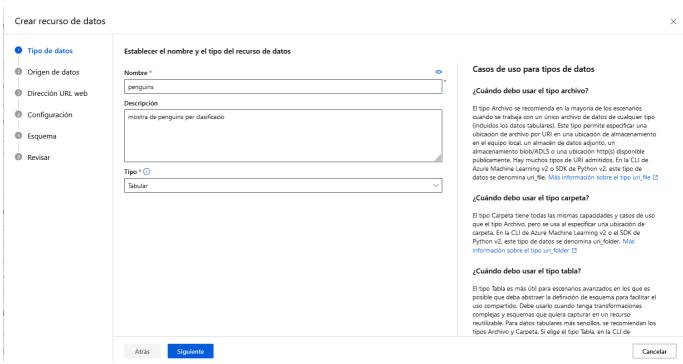
Després de obtenir la nostra conta de azure, ens dirigim a portal.azure.com i començam a crear un nou recurs. Per fer-ho, seleccionam la opció de crear un recurs i seleccionam Machine Learning o AutoLM.



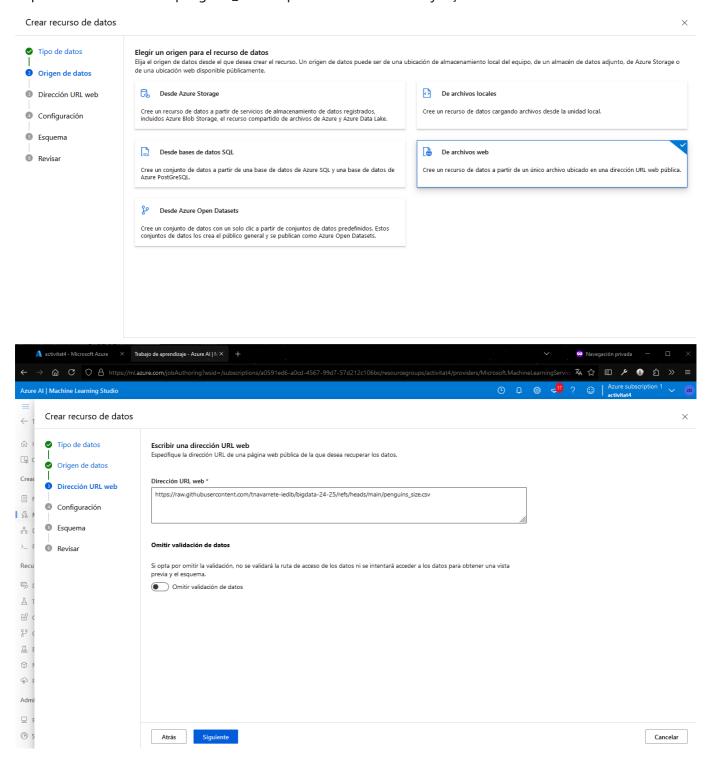
Creació del workspace

Després de que s'implementi el recurs, hem de configurar el workspace. Se ha seguit el tutorial donat als recursos del Curs.

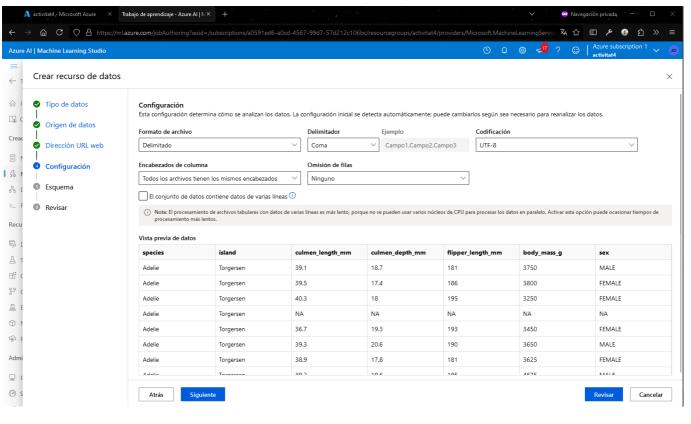


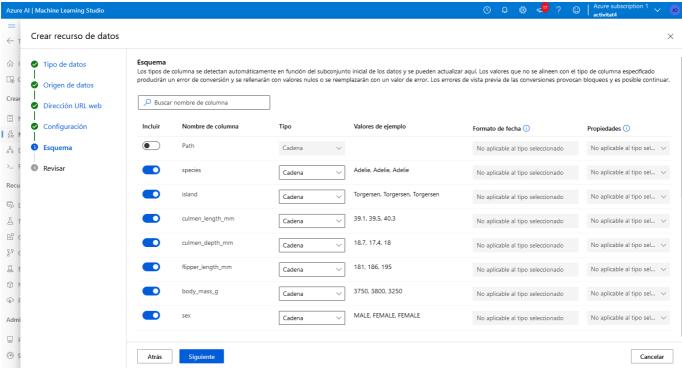


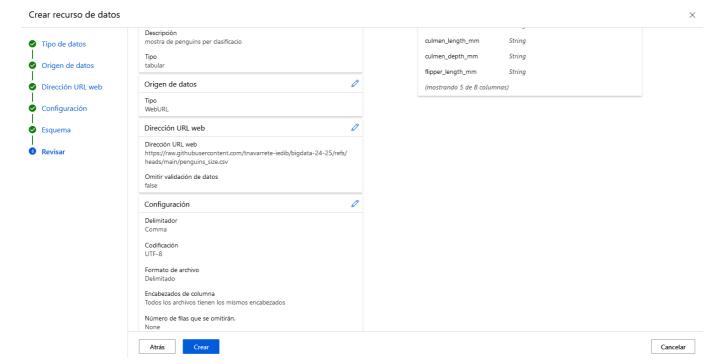
Importarem el dataset de penguins_size.csv per a la classificació mitjançant la url.



Aqui podem observar com se ha importat correctament les dates. De totes maneres, se ha de revisar que les dades siguin correctes.



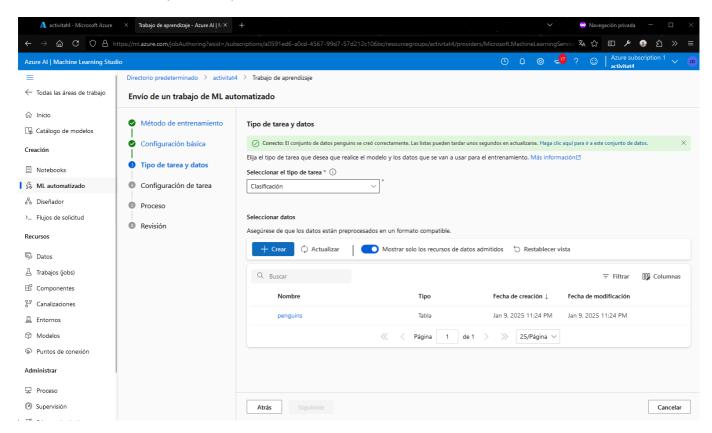




image_caption

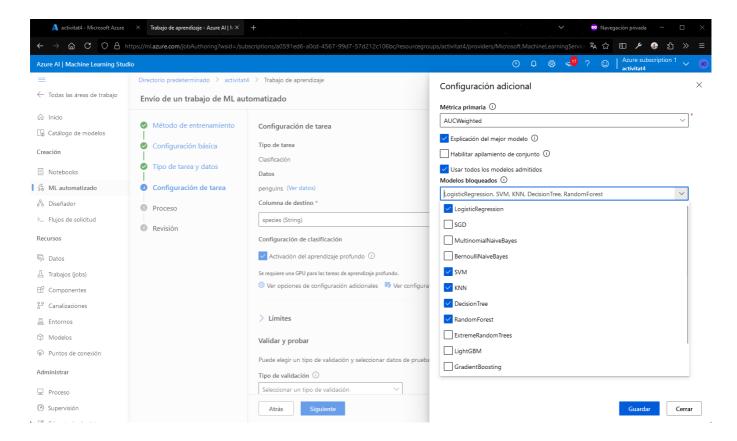
Entrenament del model

Haurem de marcar que volem fer un entrenament per a classificació, i seleccionar les columnes que volem que el model utilitzi per a fer la predicció. Haurem de marcar el tipus de entrenament com classificació i seleccionar la columna que volem predir.

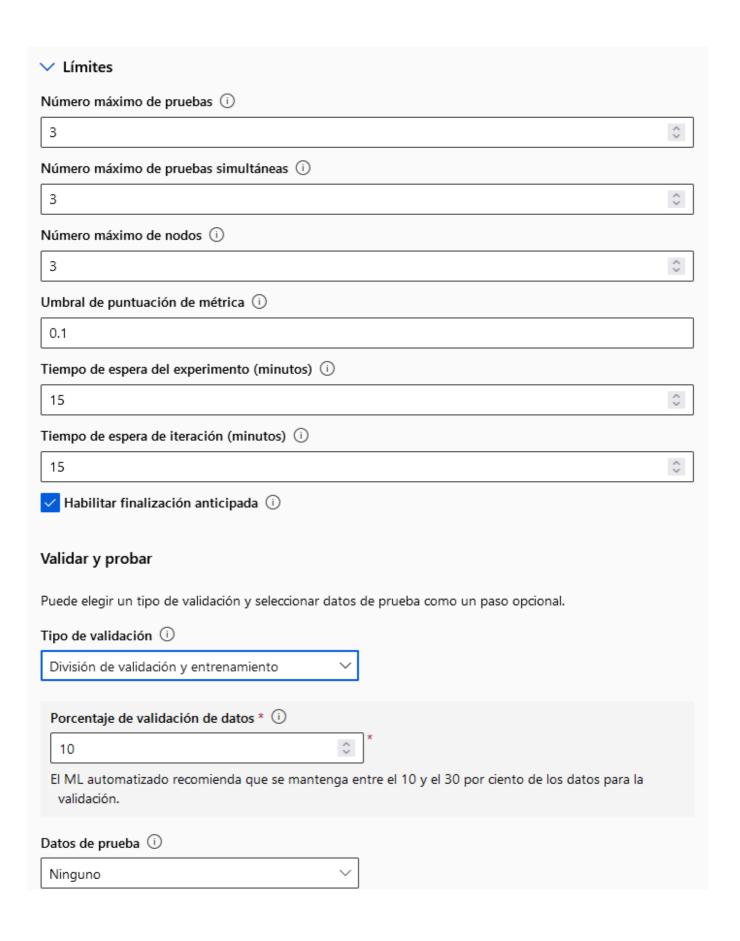


A la seguent passa, seleccionarem els models que volem que s'entrenin. En aquest cas, s'han seleccionat els següents models:

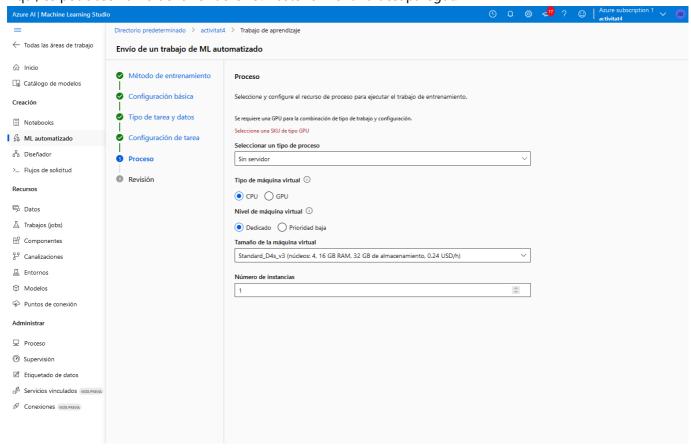
Esta activat el aprenentatge profund, posteriorment s'ha desactivat perque no es podia fer servir amb els models seleccionats sense GPU i se ha possat com Metrica principal la exactitud. Tambe se desactiva Expli



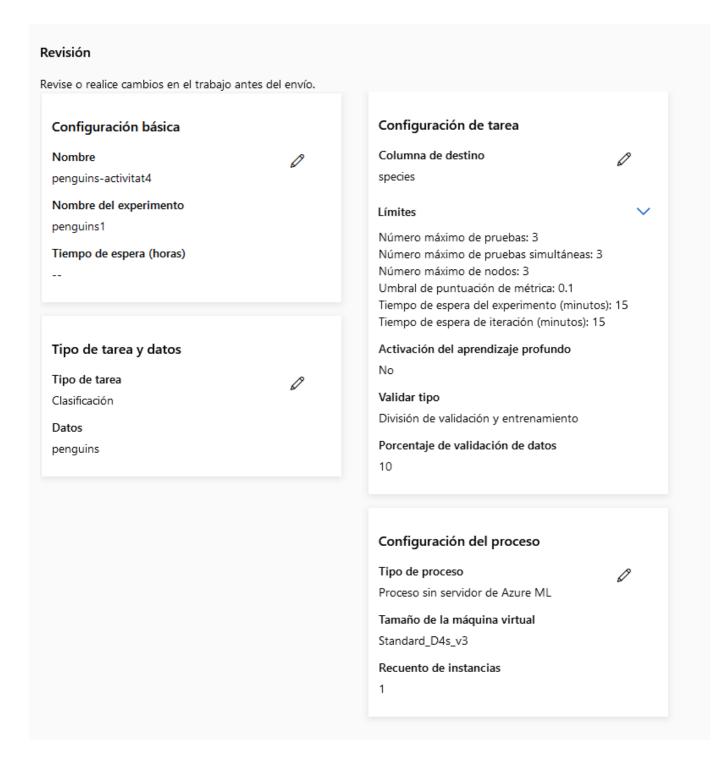
Es configuren els limits com estan especificats a l'enunciat.



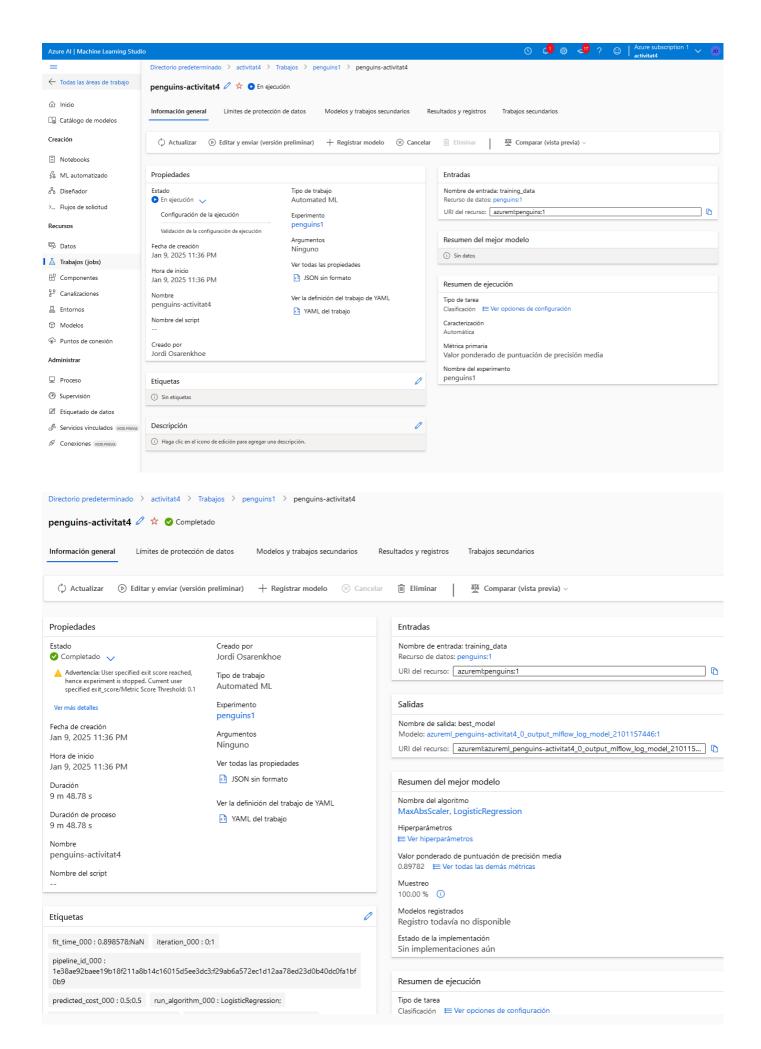
Aqui, es pot observar lo del error de GPU. Posteriorment ha desaparegut.



Resum



I aqui tenim el dashboard:



Hiperparàmetres

Hiperparámetros

Transformación de datos:

```
1  {
2    "class_name": "MaxAbsScaler",
3    "module": "sklearn.preprocessing",
4    "param_args": [],
5    "param_kwargs": {},
6    "prepared_kwargs": {},
7    "spec_class": "preproc"
8  }
```

Algoritmo de entrenamiento:

```
1
         "class_name": "LogisticRegression",
 2
         "module": "sklearn.linear_model",
 3
 4
         "param_args": [],
 5
         "param_kwargs": {
             "C": 4714.8663634573895,
 6
 7
             "class_weight": "balanced",
             "multi class": "multinomial",
 8
9
             "penalty": "12",
             "solver": "lbfgs"
10
11
12
         "prepared_kwargs": {},
         "spec_class": "sklearn"
13
     }
14
```

Mètriques

```
Precisión
0.80000
```

NA---- --- ALIC

o.93992

Micro de AUC 0.94653

Valor ponderado de AUC 0.93445

Macro de puntuación de precisión media 0.87973

Micro de puntuación de precisión media 0.90107

Valor ponderado de puntuación de precisión media 0.89782

Precisión equilibrada 0.76825

Macro de puntuación de F1 0.76753

Micro de puntuación de F1 0.80000

Valor ponderado de puntuación de F1 0.79310

Pérdida de registro 0.69754

Correlación de Matthews 0.68777

Coincidencia de macro normalizada 0.65238

Macro de puntuación de precisión 0.77302

Micro de puntuación de precisión 0.80000

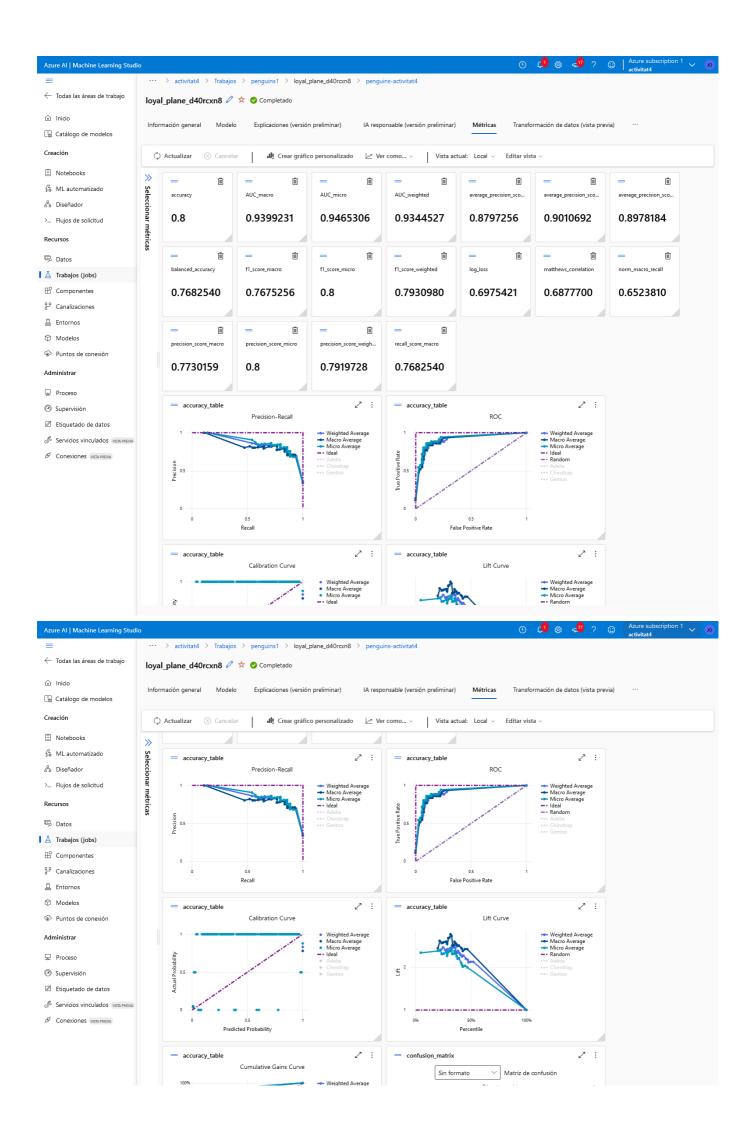
Valor ponderado de puntuación de precisión 0.79197

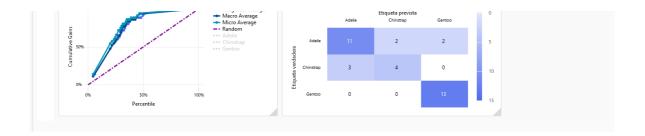
Coincidencia de macro de puntuación 0.76825

Coincidencia de micro de puntuación 0.80000

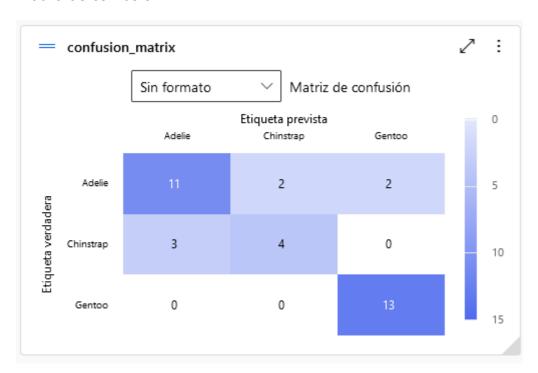
Coincidencia de valor ponderado de puntuación 0.80000

Precisión ponderada 0.81716

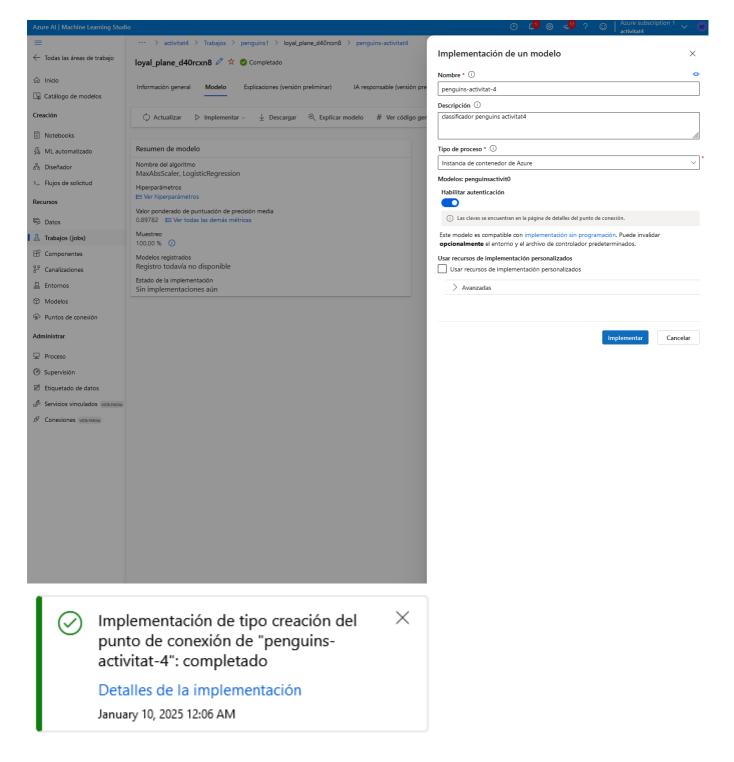




Matriu de confusió



Neteja



Se ha mandat la request des de la meva maquina local, modificant un poc el script perque donava error. El script es el seguent:

```
import urllib.request
import json
import os
import ssl

def allowSelfSignedHttps(allowed):
    # bypass the server certificate verification on client side
    if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl,
'_create_unverified_context', None):
    ssl._create_default_https_context = ssl._create_unverified_context
```

```
allowSelfSignedHttps(True) # this line is needed if you use self-signed
certificate in your scoring service.
# Request data goes here
# The example below assumes JSON formatting which may be updated
# depending on the format your endpoint expects.
# More information can be found here:
# https://docs.microsoft.com/azure/machine-learning/how-to-deploy-advanced-entry-
script
data = {
    "Inputs": {
        "data": [{
            'island': 'Biscoe',
            'culmen_length_mm': 50.0,
            'culmen_depth_mm': 15.0,
            'flipper_length_mm': 220,
            'body_mass_g': 5000,
            'sex': 'Male'
        },
            'island': 'Dream',
            'culmen_length_mm': 35.0,
            'culmen_depth_mm': 17.0,
            'flipper_length_mm': 185,
            'body_mass_g': 3400,
            'sex': 'Female'
        }]
    },
}
body = str.encode(json.dumps(data))
url = 'http://ala2cba1-a1ce-4c2c-b3c7-
75b18f73de0b.westeurope.azurecontainer.io/score'
# Replace this with the primary/secondary key, AMLToken, or Microsoft Entra ID
token for the endpoint
api key = '###'
if not api key:
    raise Exception("A key should be provided to invoke the endpoint")
headers = {'Content-Type':'application/json', 'Authorization':('Bearer '+
api key)}
req = urllib.request.Request(url, body, headers)
try:
    response = urllib.request.urlopen(req)
    result = response.read()
    print(result)
except urllib.error.HTTPError as error:
    print("The request failed with status code: " + str(error.code))
```

```
# Print the headers - they include the requert ID and the timestamp, which are
useful for debugging the failure
print(error.info())
print(error.read().decode("utf8", 'ignore'))
```

Amb el resultat:

```
b'{"Results": ["Gentoo", "Chinstrap"]}'
```

```
    □ powershell - practica_4 + ∨ □ 
    □

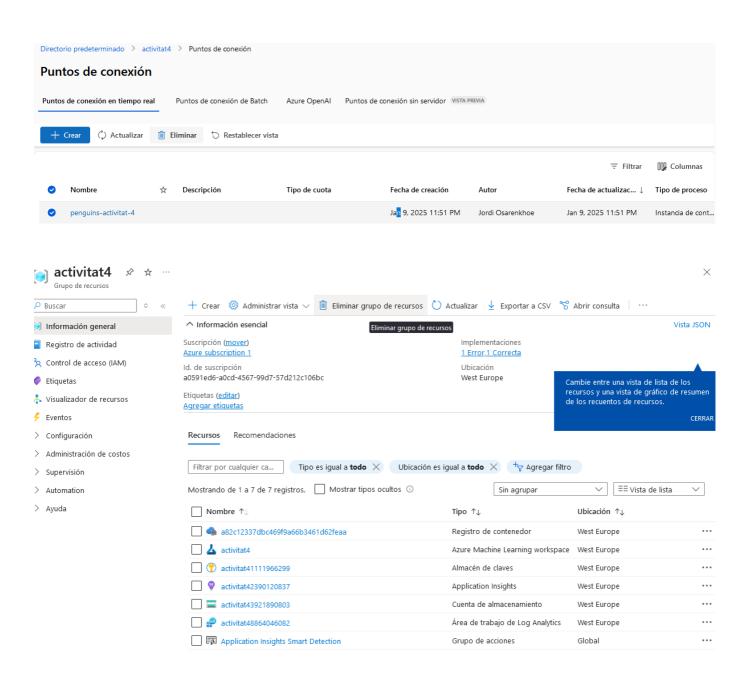
                                                                                                                                                                                                                                                                                                                                                                                                      PS C:\Users\jordi\Documents\Github\iedib\Programació (

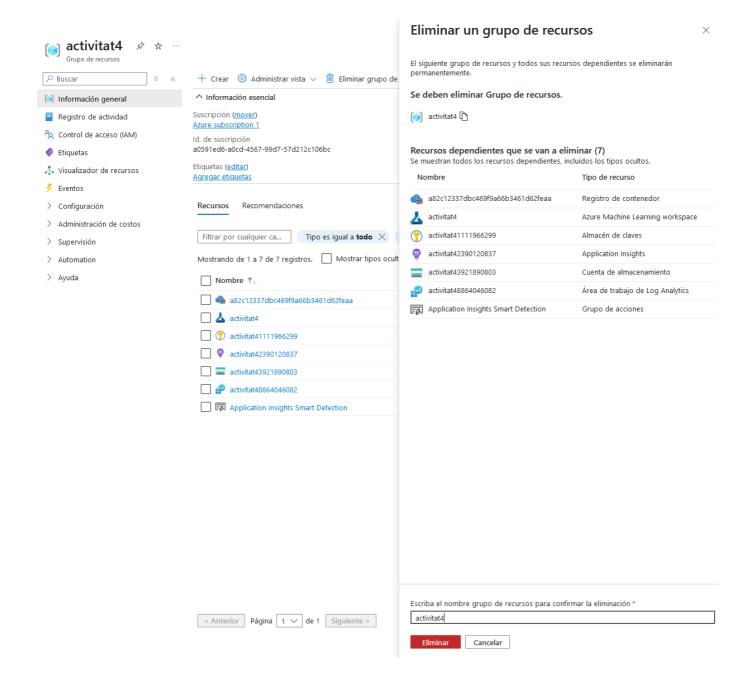
IA\practica_4> python .\script.py
b'{"Results": ["Gentoo", "Chinstrap"]}'
PS C:\Users\jordi\Documents\Github\iedib\Programació (

\[ IA\practica_4> \]
import os
import ssl
          bypass the server certificate verification on client side

if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl, '_create_unverified_context', None):

ssl._create_default_https_context = ssl._create_unverified_context
                                  ta": [{
'island': 'Biscoe',
'culmen_length_mm': 50.0,
'culmen_depth_mm': 15.0,
'flipper_length_mm': 220,
'body_mass_g': 5000,
'sex': 'Male'
                                  'island': 'Dream',
'culmen_length_mm': 35.0,
'culmen_depth_mm': 17.0,
'flipper_length_mm': 185,
'body_mass_g': 3400,
'sex': 'Female'
url = 'http://ala2cbal-alce-4c2c-b3c7-75b18f73de0b.westeurope.azurecontainer.io/score'
# Replace this with the primary/secondary key, AMLToken, or Microsoft Entra ID token for the endpoint
api_key = 'ZMbadJ6uHzZSAhkLOrv16owL1BauosJ7'
if not api_key:
    raise Exception("A key should be provided to invoke the endpoint")
   result = response.read()
print(result)
except unllib.error.HTTPError as error:
print('The request failed with status code: " + str(error.code))
```





Apartat 2: Servei de veu d'Azure Al

Anam a treballar amb el servei de veu d'Azure AI i el Speech Studio.

Heu de crear el recurs necessari per treballar amb aquest studio i anem a provar dues funcionalitats.

- 2.1. En el bloc de "Conversión de voz en texto", anam a provar la funcionalitat "Voz a texto en tiempo real". Grava un vídeo provant la funcionalitat en temps real, emprant l'entrada de micròfon del micro (tria l'idioma català o alguna de les variants d'espanyol per fer-ho). Xerra almenys 20 segons.
- 2.2. En el bloc "Texto a voz", anam a provar la funcionalitat de "Galería de voz", que permet passar text a veu, seleccionant entre diferents veus configurades. Grava un vídeo provant totes les veus disponibles en català llegint un text en català que hagis escrit tu (d'almenys 50 paraules).

Video 1