

UNIVERSIDADE SÃO JUDAS TADEU

Andressa Batista Carvalho de Sousa R.A: 819168556

Gabriel Baptistucci Garcia R.A. 819168700

Pedro Henrique Rocha Farias R.A: 81912492

Victor Tavares Talge R.A: 819145045

Projeto A3 – Inteligência Artificial: Modelo de Classificação de Diabetes

São Paulo, SP

2022

A sabedoria suprema é ter sonhos bastante grandes
para não se perderem de vista enquanto
os perseguimos.

(FAULKNER, William, 1929)

RESUMO

Durante a Unidade Curricular de Inteligência Artificial foram abordados temas indispensáveis para o entendimento e desenvolvimento do conhecimento voltado a IA, onde esse conhecimento foi voltado a desenvolver e empregar máquinas para que elas realizem atividades humanas de maneira autônoma, que desenvolvam seu próprio pensar de acordo com as instruções.

Desta forma, o tema escolhido para desenvolver uma IA que pudesse reconhecer, classificar ou até mesmo prever determinadas situações foi “Diabetes”. O desafio foi criar um algoritmo de Inteligência Artificial capaz de prever se determinada pessoa possui diabetes de acordo com diversas características físicas coletadas e dispostas em um robusto Dataset

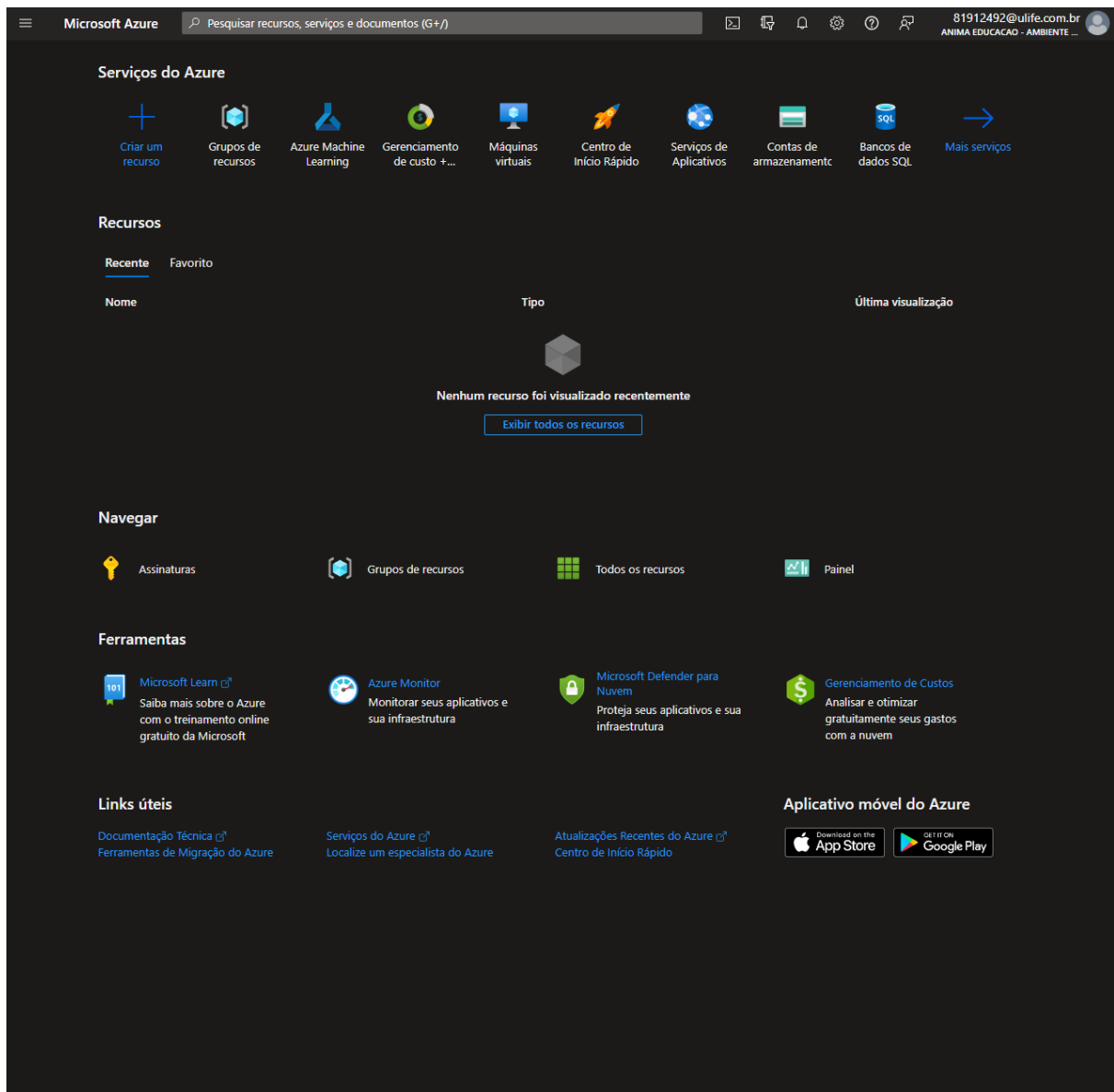
SUMÁRIO

1 CRIANDO RECURSO NO AZURE	05
2 AZURE AUTOMATED ML JOB	08
3 ACURACIA DOS MODELOS.....	14

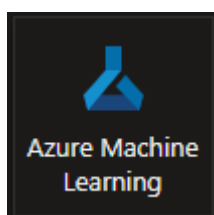
CRIANDO RECURSO NO AMBIENTE AZURE

Inicialmente deve-se acessar o portal.azure.com e realizar o login com as credenciais corretamente, após acessar o ambiente Azure será necessário criar um recurso no Azure Machine Learning.

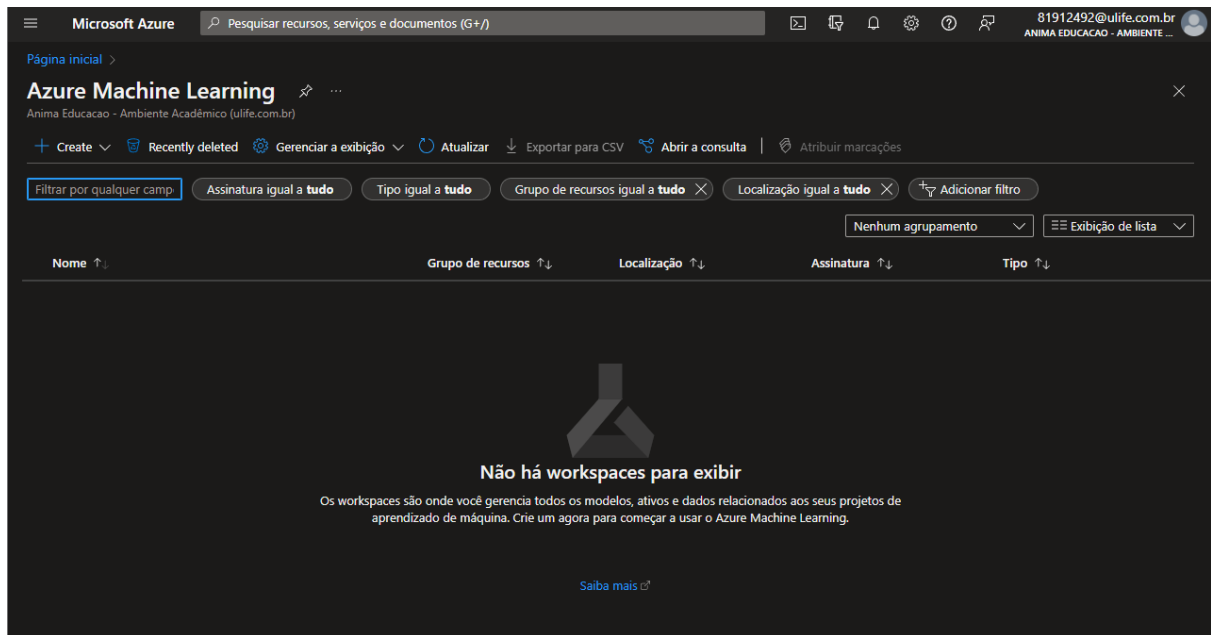
1. Portal Azure



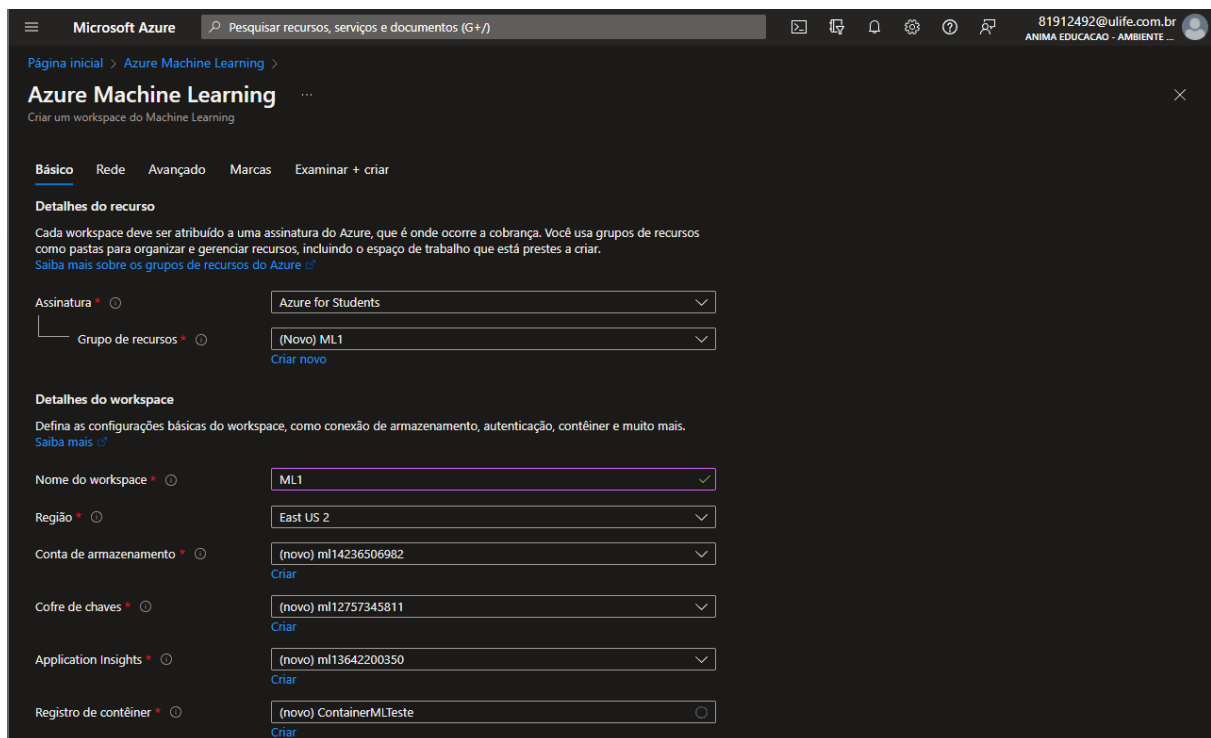
2. Azure Machine Learning



3. Criação do novo recurso dentro do Azure Machine Learning



4. Argumentos para a criação do recurso



Em resumo basta seguir o passo a passo já descrito na página, atribuindo inicialmente um nome ao seu Workspace, escolher uma região de deploy do seu recurso e guardar essas informações para utilizar posteriormente na etapa de criação do job automatizado de Machine Learning e selecionar a opção de criar um Job Automatizado de Machine Learning.

Para este caso utilizamos grande parte das configurações padrão pois o Azure ML já nos direciona para a região com o menor tempo de resposta disponível no momento.

AZURE AUTOMATED ML JOB

Após realizar a criação do recurso, deve-se acessar o ambiente Machine Learning Studio:

5. Machine Learning Studio

Microsoft Azure Machine Learning Studio

Search within your workspace (preview) This workspace

Anima Educacao - Ambiente Académico > ML1

ML1

Create new

Notebooks
Code with Python SDK and run sample experiments.
Start now

Automated ML
Automatically train and tune a model using a target metric.
Start now

Designer
Drag-and-drop interface from prepping data to deploying models.
Start now

Recent resources

Jobs Compute Models Data

Display name	☆	Experiment	Status	Logs	Submitted time	Submitte...	Job type
No jobs to display							

[View all jobs](#)

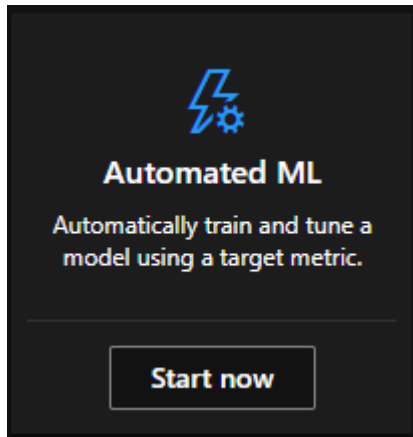
Documentation

Learning components Tutorials Additional resources

- Build AI solutions with Azure Machine Learning
- Orchestrate machine learning with pipelines
- Introduction to the Azure Machine Learning SDK
- Use automated machine learning in Azure ML
- Train a machine learning model with Azure ML
- Get started with artificial intelligence on Azure

[View all learning components](#)

6. New Automated ML Job



Para realizar a criação do ML Job é necessário realizar algumas etapas inserindo corretamente as informações para que o Job possa rodar corretamente.

Inicialmente deve-se realizar a carga do Dataset no Ambiente Azure, pois posteriormente este Dataset será indicado para que a IA possa trabalhar encima dele.

7. Carga do Dataset

Create a new Automated ML job

- Select data asset**
- Configure job
- Select task and settings
- Hyperparameter configuration (Computer Vision only)
- Validate and test

Select data asset
Select an input data asset from the list below, or create a new data asset. AutomatedML currently only supports tabular data for authoring jobs.

✓ Success: MLTest data asset created successfully. It may take a few seconds for lists to be u... ✕

+ Create Refresh | ☒ Show supported data assets only

Search All filters ✕ Clear all

Showing 1-1 of 1 data assets Page size: 25

Name	Dataset type	Created on ↓	Modified on
✓ MLTest	Tabular	Dec 5, 2022 8:...	Dec 5, 2022

Após realizar a carga do Dataset, será necessário configurar o Job com base nas colunas presentes no Dataset, atente-se ao inserir corretamente a coluna Target onde o algoritmo irá se basear inicialmente para realizar as predições e testes com os modelos.

Para o Dataset de Diabetes, a coluna target escolhida foi a “Diabetes_012”, pois é onde está localizada a resultante se a pessoa possui diabetes ou não, esta coluna leva em consideração

uma junção de diversos outros atributos de outras colunas presentes no Dataset, como por exemplo Peso, IMC, Hábitos Alimentares, Faixa Etária entre outros.

8. Configurações essenciais do ML Job

Create a new Automated ML job

- ✓ Select data asset
- Configure job**
- Select task and settings
- Hyperparameter configuration (Computer Vision only)
- Validate and test

Configure job
Select from existing experiments or create a new experiment, then select the target column and training compute.
[Learn more on how to configure the experiment.](#)

Data asset
MLTest ([View data asset](#))

Experiment name
☒ Create new
New experiment name *
MLTest

Target column *
Diabetes_012 (String)

Select compute type
Compute cluster

Select Azure ML compute cluster *
Compute1

[+ New](#) [Refresh computes](#)

Com o Dataset configurado e as devidas atribuições realizadas no ML Job, basta avançar para a próxima etapa para definirmos qual será o modelo que utilizaremos juntamente com as configurações de cada um dos modelos

Para o Dataset escolhido, o modelo escolhido foi o de Classificação, pois devemos prever a resultante com base na classificação combinada de diversas colunas presentes no Dataset.

9. Modelos

Create a new Automated ML job

Select data asset

Configure job

Select task and settings

Hyperparameter configuration
(Computer Vision only)

Validate and test

Select task and settings

Select the machine learning task type for the experiment. To fine tune the experiment, choose additional configuration or featurization settings.

Classification
To predict one of several categories in the target column.
yes/no, blue, red, green.

☒ Enable deep learning

Regression
To predict continuous numeric values.

Time series forecasting
To predict values based on time.

Natural language processing
Predict based on text-only data types using multi-class or multi-label classification.




Computer vision
Multi-class or multi-label image classification, object detection, and instance segmentation.

View additional configuration settings

View featurization settings

Com o ML Job perfeitamente configurado, podemos acompanhar os status do deploy na seguinte tela:

10. Deploy do ML Job

witty_guitar_qpg2069p    Running


Overview


Data guardrails


Models


Outputs + logs


Child jobs



 Refresh

 Edit and submit (preview)

 Register model



 Cancel

 Delete

 Compare (preview) 

Properties

Status

 Running 

Model training

Created on

Dec 5, 2022 8:38 PM

Start time

Dec 5, 2022 8:38 PM

Compute target

Compute1

Name

AutoML_a7ca6f2a-c8e3-47e6-bc56-0d6e1a08cf80

Script name

--

Created by

Pedro Henrique Rocha Farias

Job type

Automated ML


Experiment

MLTest


Arguments

None

See all properties


 Raw JSON

See YAML job definition

 Job YAML


Inputs

Input name: training_data


Dataset: [MLTest:1](#) 

Outputs


Output name: full_training_dataset

Dataset: [ec6ee3c6-3bfe-4e8d-a44f-bcb8a0d51200](#) 

Output name: full_validation_dataset


Dataset: [c99a3e65-9b2a-420d-9f2b-838961945980](#) 

Best model summary

 No data

Run summary

Task type

Classification  [View configuration settings](#)

Featurization

Auto

Primary metric

AUC weighted

Experiment name

MLTest

10.1.

witty_guitar_qpg2069p ★ ✔ Completed

Overview

Data guardrails

Models

Outputs + logs

Child jobs

Refresh

Edit and submit (preview)

+ Register model

ⓧ Cancel

🗑 Delete

|

🔍 Compare (preview) ▾

Properties

Status

✔ Completed ▾

⚠ Warning: No scores improved over last 20 iterations, so experiment stopped early. This early stopping behavior can be disabled by setting enable_early_stopping = False in AutoMLConfig for notebook/python
[See more details](#)

Created on

Dec 5, 2022 8:38 PM

Start time

Dec 5, 2022 8:38 PM

Duration

56m 57.30s

Compute duration

56m 57.30s

Compute target

Compute1

Name

AutoML_a7ca6f2a-c8e3-47e6-bc56-0d6e1a08cf80

Script name

--

Created by

Pedro Henrique Rocha Farias

Job type

Automated ML

Experiment

MLTest

Arguments

None

See all properties

📄 Raw JSON

See YAML job definition

📄 Job YAML

Tags

Inputs

Input name: training_data
Dataset: [MLTest:1](#) 🔗

Outputs

Output name: best_model
Model: [azureml_AutoML_a7ca6f2a-c8e3-47e6-bc56-0d6e1a08cf80_40_output_mlflow_log_model_2119407539:1](#)

Output name: full_training_dataset
Dataset: [ec6ee3c6-3bfe-4e8d-a44f-bcb8a0d51200](#) 🔗

Output name: full_validation_dataset
Dataset: [c99a3e65-9b2a-420d-9f2b-838961945980](#) 🔗

Best model summary

Algorithm name
[VotingEnsemble](#)

Ensemble details
[≡ View ensemble details](#)

AUC weighted
0.82907 [≡ View all other metrics](#)

Sampling
100.00 % ⓘ

Registered models
No registration yet

Deploy status
No deployment yet

Run summary

Task type
Classification [≡ View configuration settings](#)

Featurization
Auto

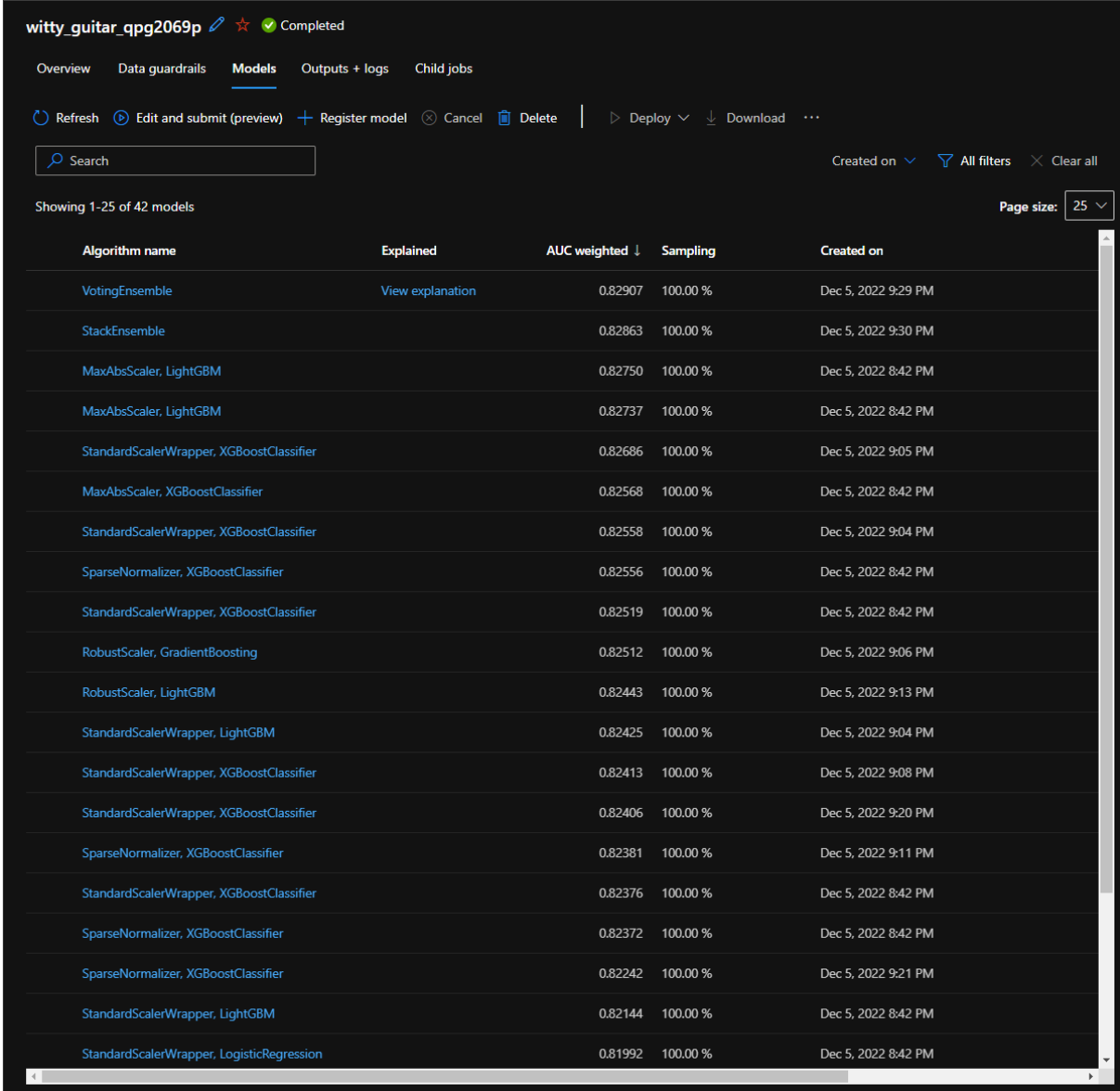
Primary metric
AUC weighted

Experiment name
MLTest

ACURÁCIA DOS MODELOS

Para visualizar os resultados da execução do Job basta acessar a aba “Models” para visualizar todos os modelos que foram testados com base no Dataset. O Azure por padrão já realiza uma classificação em ordem decrescente contemplando os modelos que obtiveram maior acurácia com base no algoritmo escolhido, que no caso foi Classificação.

11. Acurácia



witty_guitar_qpg2069p 🔗 ★ ✅ Completed

Overview Data guardrails **Models** Outputs + logs Child jobs

🔄 Refresh ⏸ Edit and submit (preview) + Register model ⌂ Cancel 🗑 Delete | ▶ Deploy ⬇ Download ...

🔍 Search Created on ▾ All filters ✕ Clear all

Showing 1-25 of 42 models Page size: 25 ▾

Algorithm name	Explained	AUC weighted ↓	Sampling	Created on
VotingEnsemble	View explanation	0.82907	100.00 %	Dec 5, 2022 9:29 PM
StackEnsemble		0.82863	100.00 %	Dec 5, 2022 9:30 PM
MaxAbsScaler, LightGBM		0.82750	100.00 %	Dec 5, 2022 8:42 PM
MaxAbsScaler, LightGBM		0.82737	100.00 %	Dec 5, 2022 8:42 PM
StandardScalerWrapper, XGBoostClassifier		0.82686	100.00 %	Dec 5, 2022 9:05 PM
MaxAbsScaler, XGBoostClassifier		0.82568	100.00 %	Dec 5, 2022 8:42 PM
StandardScalerWrapper, XGBoostClassifier		0.82558	100.00 %	Dec 5, 2022 9:04 PM
SparseNormalizer, XGBoostClassifier		0.82556	100.00 %	Dec 5, 2022 8:42 PM
StandardScalerWrapper, XGBoostClassifier		0.82519	100.00 %	Dec 5, 2022 8:42 PM
RobustScaler, GradientBoosting		0.82512	100.00 %	Dec 5, 2022 9:06 PM
RobustScaler, LightGBM		0.82443	100.00 %	Dec 5, 2022 9:13 PM
StandardScalerWrapper, LightGBM		0.82425	100.00 %	Dec 5, 2022 9:04 PM
StandardScalerWrapper, XGBoostClassifier		0.82413	100.00 %	Dec 5, 2022 9:08 PM
StandardScalerWrapper, XGBoostClassifier		0.82406	100.00 %	Dec 5, 2022 9:20 PM
SparseNormalizer, XGBoostClassifier		0.82381	100.00 %	Dec 5, 2022 9:11 PM
StandardScalerWrapper, XGBoostClassifier		0.82376	100.00 %	Dec 5, 2022 8:42 PM
SparseNormalizer, XGBoostClassifier		0.82372	100.00 %	Dec 5, 2022 8:42 PM
SparseNormalizer, XGBoostClassifier		0.82242	100.00 %	Dec 5, 2022 9:21 PM
StandardScalerWrapper, LightGBM		0.82144	100.00 %	Dec 5, 2022 8:42 PM
StandardScalerWrapper, LogisticRegression		0.81992	100.00 %	Dec 5, 2022 8:42 PM

Nota-se que o modelo com maior acurácia foi o “VotingEnsemble”, desta forma seguiremos com este modelo para realizar o deploy em API para que o modelo fique On através de um EndPoint.

12. Acurácia

Algorithm name	Explained	AUC weighted ↓
VotingEnsemble	View explanation	0.82907
StackEnsemble		0.82863

13. Modelo Detalhado

shy_balloon_hkb87fb8 [🔗](#) [★](#) [✅](#) Completed

Overview **Model** Explanations (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs Code Monitor

[🔄 Refresh](#)
[▶ Deploy](#)
[↓ Download](#)
[🔍 Explain model](#)
[# View generated code](#)
[✓ Test model \(preview\)](#)
[+ Register model](#)
[⌫ Cancel](#)
[🗑 Delete](#)

Model summary

Algorithm name
VotingEnsemble

Ensemble details
[View ensemble details](#)

AUC weighted
0.82907 [View all other metrics](#)

Sampling
100.00 % [?](#)

Registered models
No registration yet

Deploy status
No deployment yet

14. Deploy to web service

Overview **Model** Explanations (preview) Metrics

[🔄 Refresh](#)
[▶ Deploy](#)
[↓ Download](#)
[🔍 Explain model](#)

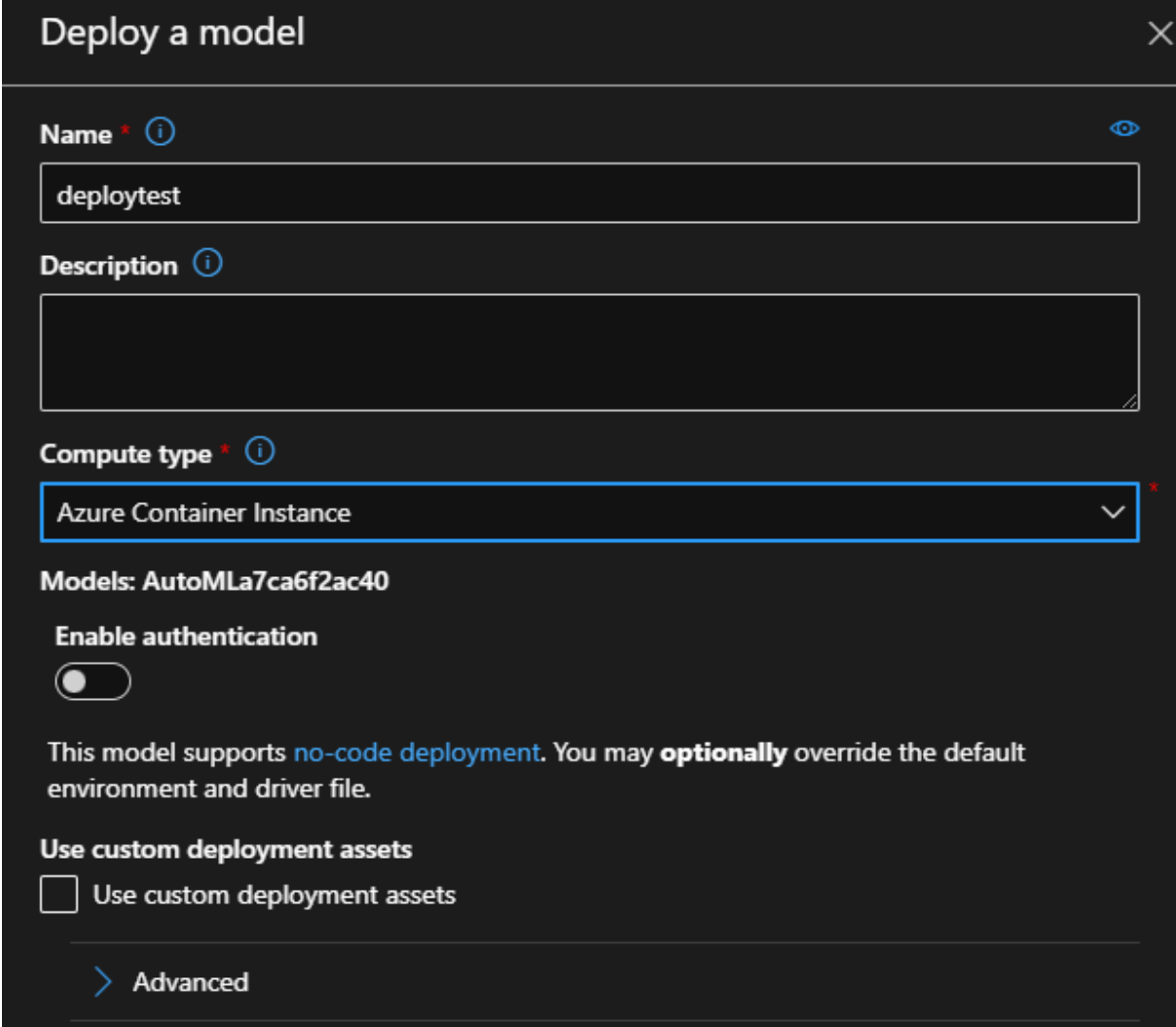
Model summary

Algorithm name
VotingEnsemble

Deploy to real-time endpoint
Deploy the model using the new real-time endpoint wizard

Deploy to web service
Deploy the model to a web service

15. Configurações do Deploy



The screenshot shows a dark-themed dialog box titled "Deploy a model" with a close button (X) in the top right corner. The dialog contains several fields and options:

- Name**: A text input field containing "deploytest". It has a red asterisk and an information icon (i) to its left, and a visibility icon (eye) to its right.
- Description**: A large, empty text area. It has an information icon (i) to its left.
- Compute type**: A dropdown menu showing "Azure Container Instance". It has a red asterisk and an information icon (i) to its left, and a red asterisk to its right.
- Models**: Labeled "AutoMLa7ca6f2ac40".
- Enable authentication**: A toggle switch that is currently turned off.
- Information text**: "This model supports [no-code deployment](#). You may **optionally** override the default environment and driver file."
- Use custom deployment assets**: A checkbox that is currently unchecked, followed by the text "Use custom deployment assets".
- Advanced**: A link with a blue chevron icon pointing right.

Após a realização de todas as etapas de configuração do Endpoint, basta aguardar que o Endpoint em breve estará disponível, é comum que leve alguns minutos até o fim da configuração e deploy total.

15. Endpoint Final

deploytest ☆

Details

Test

Consume

Deployment logs

Attributes

Service ID

deploytest

Description

--

Deployment state

Healthy ⓘ

Operation state

Running

Compute type

Container instance

Created by

Pedro Henrique Rocha Farias

Model ID

AutoMLa7ca6f2ac40:1

Created on

Dec 5, 2022 9:57 PM

Last updated on

Dec 5, 2022 9:57 PM

Image ID

--

REST endpoint

http://21431189-646a-48ab-a702-79d1c563b662.eastus2.azurecontainer.io...

 ⓘ

Key-based authentication enabled

false

Swagger URI

http://21431189-646a-48ab-a702-79d1c563b662.eastus2.azurecontain

er.io/swagger.json

 ⓘ

CPU

1.8

Memory

4 GB

Application Insights enabled

false

Tags

ⓘ No data

Properties

runId

AutoML_a7ca6f2a-c8e3-47e6-bc56-0d6e1a08cf80_40