

Desafio 1 Microsoft

Integrantes:

Bruno Cássio Fonseca - RA 319142254

Débora Cristina Piani do Nascimento - RA: 31916025

Tiago Gilberto dos Santos - RA 320143021

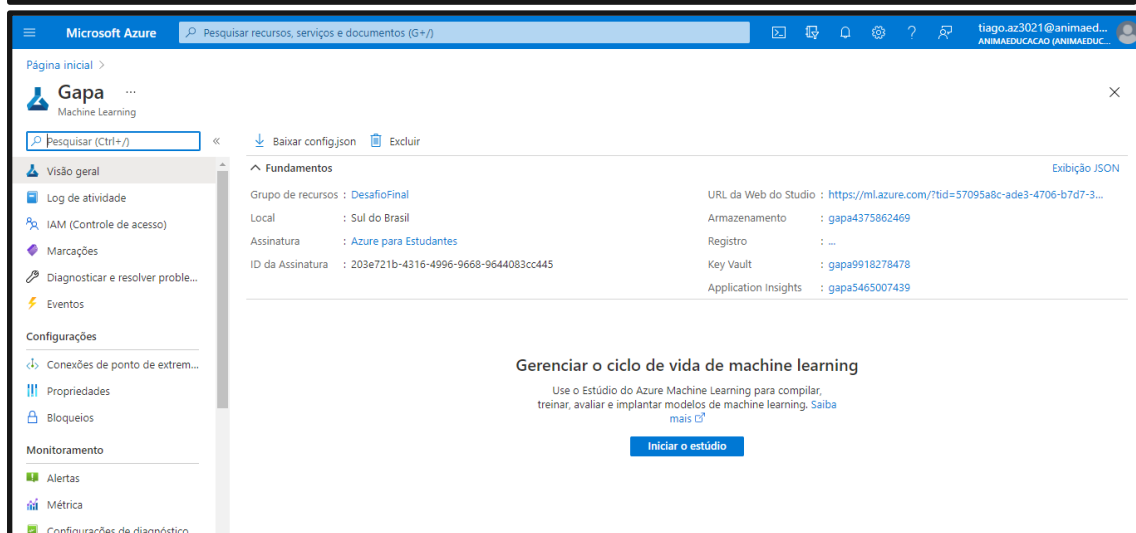
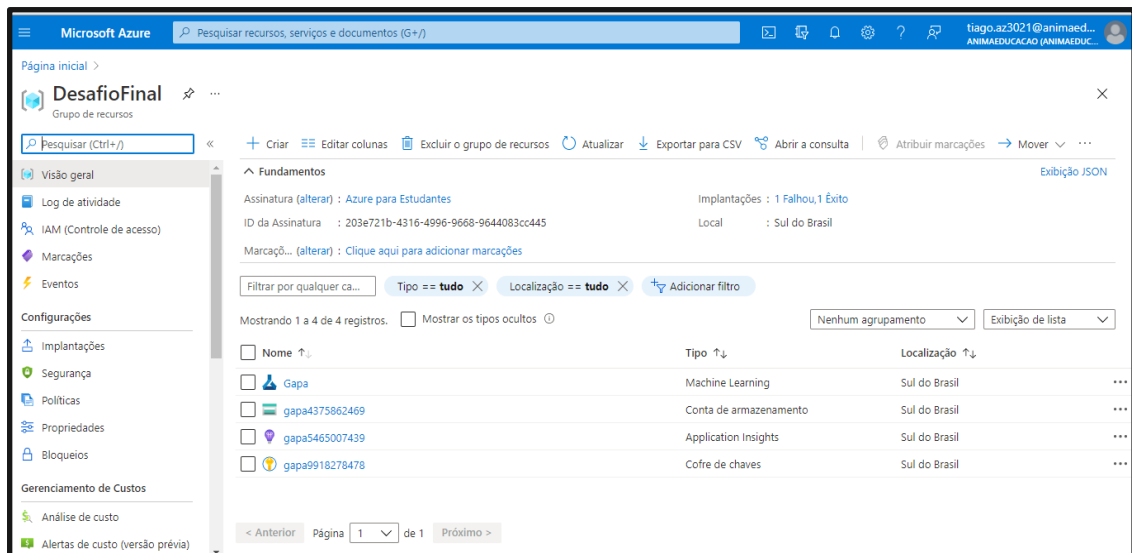
Para realização do trabalho fizemos as seguintes tarefas:

Criar um workspace do Azure Machine Learning

Entramos no portal do Azure usando nossas credenciais da Microsoft.

1. Seleccionamos + **Criar um recurso**, pesquisamos **Machine Learning** e criamos um recurso de Machine Learning com as seguintes configurações:

- **Assinatura:** Azure for students
- **Grupo de recursos:** Criamos o recurso DesafioFinal
- **Nome do workspace:** Criamos o workspace Gapa
- **Região:** Sul do Brasil
- **Conta de armazenamento:** gapa4375862469
- **Cofre de chaves:** gapa9918278478
- **Application Insights:** gapa5465007439
- **Registro de contêiner:** nenhum



2. Acessamos o recurso criado

3. Iniciamos o estúdio do Azure Machine Learning

URL do estudo: <https://ml.azure.com/?tid=57095a8c-ade3-4706-b7d7-30379e58918b&wsid=/subscriptions/203e721b-4316-4996-9668-9644083cc445/resourcegroups/DesafioFinal/workspaces/Gapa>

Criar recursos de computação Criar destinos de computação

1. No estúdio do Machine Learning, fomos em **computação**
2. Na guia **Instâncias de Computação**, criamos uma nova instância de computação com as configurações a seguir:
 - a) **Tipo de máquina virtual:** CPU
 - b) **Tamanho da máquina virtual:** Standard_DS11_v2
 - c) **Nome de computação:** MyVM1
 - d) **Habilitar o acesso SSH:** não selecionado

3. Na guia **Clusters de Cálculo**, criamos um novo cluster de cálculo com as configurações a seguir:
 - a) **Prioridade da máquina virtual:** dedicada
 - b) **Tipo de máquina virtual:** CPU
 - b) **Tamanho da máquina virtual:** Standard_DS11_v2
 - c) **Nome de computação:** MyVC1
 - d) **Número mínimo de nós:** 0
 - e) **Número máximo de nós:** 2
 - f) **Segundos de espera antes de reduzir verticalmente:** 120
 - g) **Habilitar o acesso SSH:** não selecionado

Criar um conjunto de dados

1. No estúdio do Azure Machine Learning, fomos na página **Conjuntos de Dados**.
2. Criamos um conjunto de dados com base nos arquivos baixados do Enem:

Informações Básicas:

Arquivo importado: MICRODADOS_ENEM_2019

Nome: microdados-enem2

Tipo de conjunto de dados: tabular

Descrição: microdados-enem

Configurações e visualização:

Formato de arquivo: delimitado

Delimitador: Ponto e vírgula

Codificação: ISO-8859-1

Cabeçalhos de coluna: usar cabeçalhos do primeiro arquivo

Ignorar linhas: Nenhum

Esquema:

incluir todas as colunas que não sejam Caminho

Confirmar os detalhes:

não criar o perfil do conjunto de dados após a criação


Informações básicas	Seleção de armazenamento de dados e de a...
Nome microdados-enem2	Armazenamento de dados workspaceblobstore
Versão do conjunto de dados 1	Arquivos selecionados (1) MICRODADOS_ENEM_2019.csv
Tipo do conjunto de dados Tabular	Caminho UI/06-19-2021_055510.UTC/MICRODADOS_ENEM_2019.csv
Descrição microdados-enem	

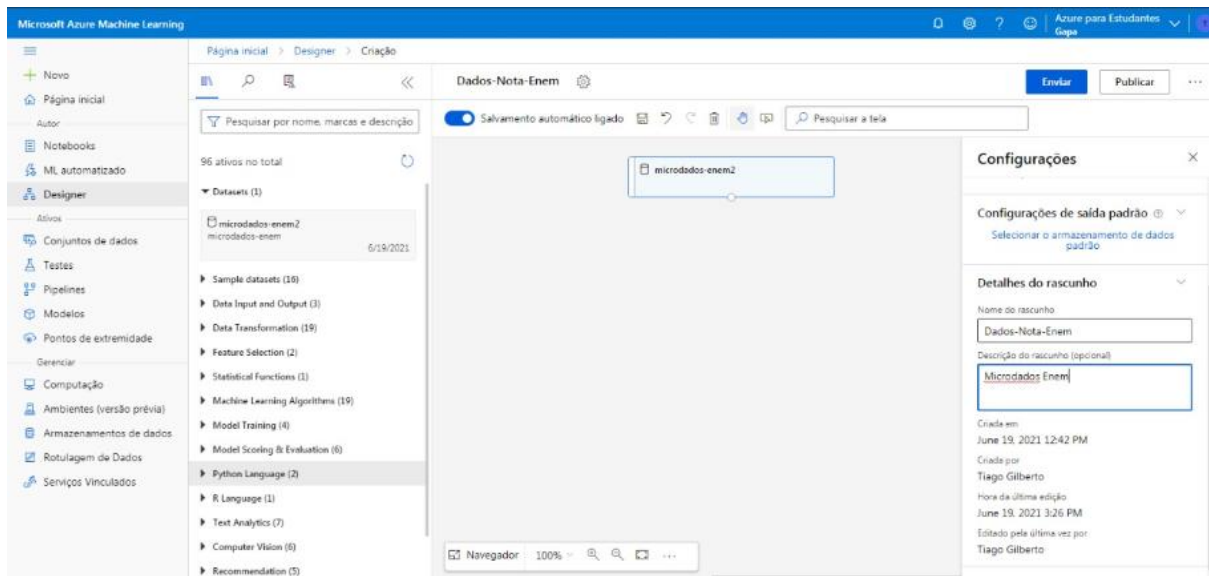
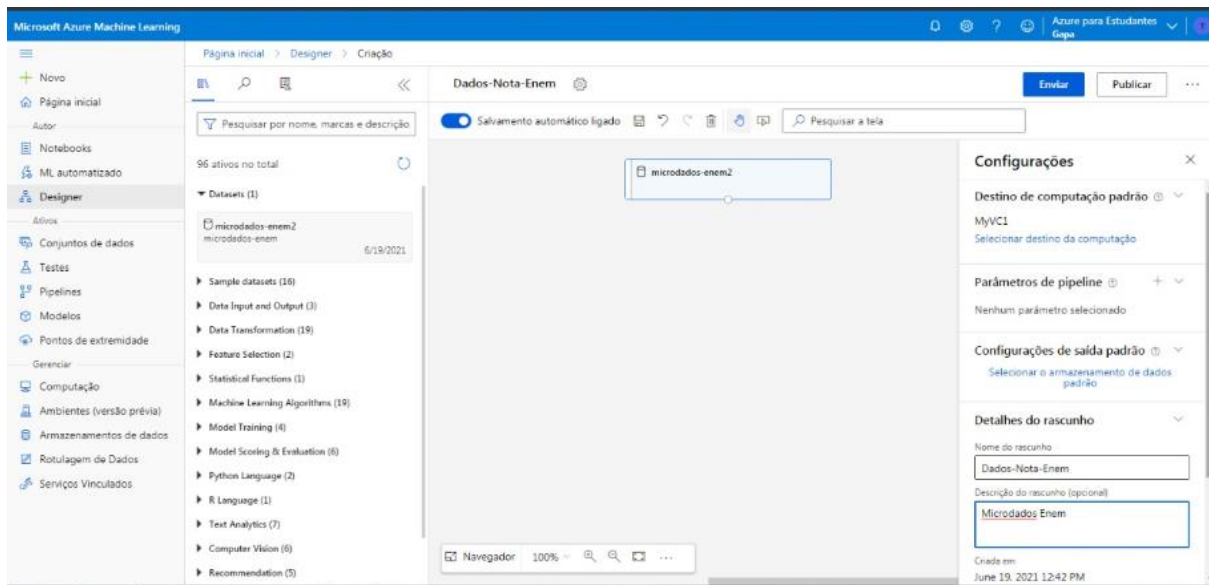
Configurações de arquivo
Formato do arquivo Delimitado
Delimitador Ponto e vírgula
Codificação ISO-8859-1
Cabeçalhos de coluna Somente o primeiro arquivo tem cabeçalhos
Ignorar linhas Nenhuma

- Depois que o conjunto de dados tiver sido criado, abra-o e exiba a página Explorar para ver uma amostra dos dados. Esses dados representam detalhes de pacientes que foram testados para diabetes.

Criar um pipeline

Para começar a usar o designer do Azure Machine Learning, primeiro você precisa criar um pipeline e adicionar o conjunto de dados com o qual deseja trabalhar.

- No estúdio do Azure Machine Learning, para o seu workspace, veja a página Designer e selecione + para criar um pipeline.
- No painel Configurações, altere o nome do pipeline padrão (*Pipeline-Created-on- date***) (se o painel Configurações não estiver visível, clique no ícone  ao lado do nome do pipeline na parte superior).



3. Você precisa especificar um destino de computação no qual executar o pipeline. No painel Configurações, clique em Selecionar destino de computação e selecione o cluster de cálculo criado anteriormente.

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem [Enviar] [Criar um pipeline de inferência] [Publicar]

Salvamento automático ligado [Pesquisar a tela] Execução concluída [Exibir a visão geral de execução]

microdados-enem2

Select Columns in Dataset [Concluído]

Clean Missing Data [Concluído]

Normalize Data [Concluído]

Split Data [Concluído]

Train Model [Concluído]

Linear Regression [Concluído]

Navegador 60%

Select Columns in Dataset

Atualizar

Parâmetros Saídas + logs Detalhes

Select columns *

Nomes de colunas:
NU_NOTA_CN,NU_NOTA_MT

Configurações de saída

Configurações de execução

Comentário

Informações do módulo

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem [Enviar] [Criar um pipeline de inferência] [Publicar]

Salvamento automático ligado [Pesquisar a tela] Execução concluída [Exibir a visão geral de execução]

microdados-enem2

Select Columns in Dataset [Concluído]

Clean Missing Data [Concluído]

Normalize Data [Concluído]

Split Data [Concluído]

Train Model [Concluído]

Linear Regression [Concluído]

Normalize Data

Atualizar Baixar tudo

Parâmetros Saídas + logs Detalhes

Saídas de dados

Transformation function

Transformed dataset

Outras saídas

azureml-logs

- 55_azureml-execution-vmmp-304
- 65_job_prep-vmmp-304527544991
- 70_driver_log.txt
- 75_job_post-vmmp-304527544991
- process_info.json

Navegador 100%

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem [Enviar] [Publicar]

Salvamento automático ligado [Pesquisar a tela]

microdados-enem2

Select Columns in Dataset [Concluído]

Clean Missing Data [Concluído]

Normalize Data [Concluído]

Split Data [Concluído]

Train Model [Concluído]

Linear Regression [Concluído]

Split Data

Splitting mode *

Split Rows

Fraction of rows in the first output dataset *

0.7

Randomized split *

True

Random seed *

123

Stratified split *

False

Configurações de saída

Configurações de execução

Comentário

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Novo

Página inicial

Autor

Notebooks

ML automatizado

Designer

Ativos

Conjuntos de dados

Testes

Pipelines

Modelos

Pontos de extremidade

Gerenciar

Computação

Ambientes (versão prévia)

Armazenamentos de dados

Rotulagem de Dados

Serviços Vinculados

Dados-Nota-Enem

Enviar

Criar um pipeline de inferência

Publicar

Salvamento automático ligado

Pesquisar a tela

Execução concluída Exibir a visão geral de execução

microdados-enem2

Select Columns in Dataset

Clean Missing Data

Normalize Data

Linear Regression

Split Data

Train Model

Score Model

Reduzir

Navegador 60%

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Novo

Página inicial

Autor

Notebooks

ML automatizado

Designer

Ativos

Conjuntos de dados

Testes

Pipelines

Modelos

Pontos de extremidade

Gerenciar

Computação

Ambientes (versão prévia)

Armazenamentos de dados

Rotulagem de Dados

Serviços Vinculados

Dados-Nota-Enem

Enviar

Criar um pipeline de inferência

Publicar

Score Model visualização do resultado

Linhas 1.113.130

Colunas 3

NU_NOTA_CN	NU_NOTA_MT	Scored Labels
0.613776	0.609843	0.59594
0.464979	0.428818	0.50223
0.754675	0.74551	0.66617
0.500755	0.483714	0.530648
0.679289	0.608118	0.595047
0.473574	0.420599	0.497975
0.562782	0.548452	0.56416

Scored Labels

Estatísticas

Média	0.555
Mediana	0.5435
Min.	0.2802
Máx.	0.7979
Desvio padrão	0.0573
Valores exclusivos	5855
Valores ausentes	0

Fechar

70_driver_log.txt

75_job_post-tvmps_8d9d385846d

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem

Enviar Criar um pipeline de inferência Publicar

Score Model visualização do resultado

Linhas 1.113.130 Colunas 3

NU_NOTA_CN	NU_NOTA_MT	Scored Labels
0.613776	0.609843	0.59594
0.464979	0.428818	0.50223
0.754675	0.74551	0.66617
0.500755	0.483714	0.530648
0.679289	0.608118	0.595047
0.473574	0.420599	0.497975
0.563783	0.548453	0.56416

Fechar

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem

Enviar Criar um pipeline de inferência Publicar

Salvamento automático ligado

Execução concluída Exibir a visão geral de execução

Score Model

Atualizar Baixar tudo

Parâmetros Saídas + logs Detalhes

Saídas de dados Ocultar as saídas de dados

Scored dataset

Outras saídas

- azureml-logs
 - 55_azureml-execution-tvmps_8d9
 - 65_job_prep-tvmps_8d9d385846c
 - 70_driver_log.txt
 - 75_job_post-tvmps_8d9d385846d

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem

Enviar Criar um pipeline de inferência Publicar

Exibir as Instruções de Saída

Siga as etapas abaixo para exibir a saída no Portal do Ibiza

- Clique em 'Abrir o Armazenamento de Dados' para acessar o armazenamento de dados
[Abrir o Armazenamento de Dados](#)
- Copie o caminho abaixo e cole-o na caixa de pesquisa para procurar a pasta
azureml/54cb852d-f3c8-4ae6-9ac0-2b1126a0b5fe/Scored_dataset

Fechar

azureml/54cb852d-f3c8-4ae6-9ac0-2b1126a0b5fe/Scored_dataset

Microsoft Azure

Pesquisar recursos, serviços e documentos (G+)

Página inicial >

azureml-blobstore-ef85f108-b2bb-456a-ae90-b3bd14fa089e

Contêiner

Pesquisar (Ctrl+/)

Carregar Alterar o nível de acesso Atualizar Excluir Alterar a camada Adquirir concessão Interromper concessão

Visão geral

Diagnosticar e resolver problemas

IAM (Controle de Acesso)

Configurações

Assinatura de acesso compartilh...

Política de acesso

Propriedades

Metadados

Método de autenticação: Tecla de acesso (Alternar para a Conta de Usuário do Azure AD)

Local: azureml-blobstore-ef85f108-b2bb-456a-ae90-b3bd14fa089e

Pesquisar blobs por prefixo (diferenciar maiúsculas de minúsculas)

Mostrar blobs excluídos

Adicionar o filtro

	Nome	Modificado	Camada de acesso	Tipo de blob	Tamanho	Estado de concessão
<input checked="" type="checkbox"/>	azureml					...
<input checked="" type="checkbox"/>	UI					...

Microsoft Azure Machine Learning

Página inicial > Designer > Criação

Dados-Nota-Enem

Salvamento automático ligado

Pesquisar a tela

Novo

Página inicial

Autor

Notebooks

ML automatizado

Designer

Ativos

Conjuntos de dados

Testes

Pipelines

Modelos

Pontos de extremidade

Gerenciar

Computação

Ambientes (versão prévia)

Armazenamentos de dados

Rotulagem de Dados

Serviços Vinculados

Selecionar Colunas de Dataset

Clean Missing Data

Normalize Data

Linear Regression

Split Data

Train Model

Score Model

Notificações

Habilitar as notificações pop-up

Ignorar tudo

Execução Pipeline 14 no teste "TesteDesafio1" Concluído

Detalhes da execução

June 19, 2021 6:30 PM

Execução Pipeline 3 no teste "TesteDesafio1" Concluído

Detalhes da execução

June 19, 2021 4:56 PM

Computar "MyVM1" provisionando com êxito

Detalhes da computação

June 19, 2021 12:40 PM