

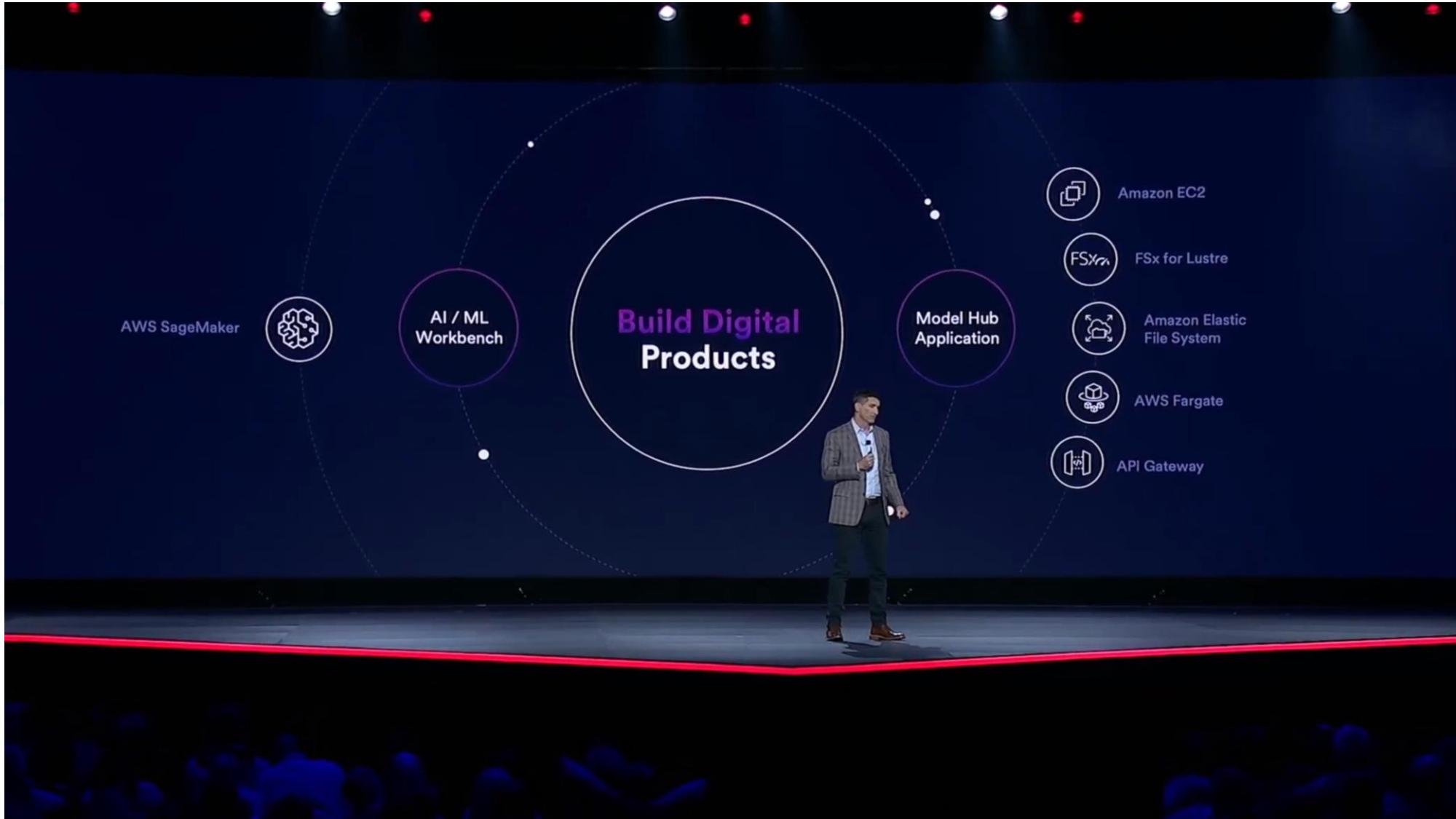


Human-Centered Data & AI

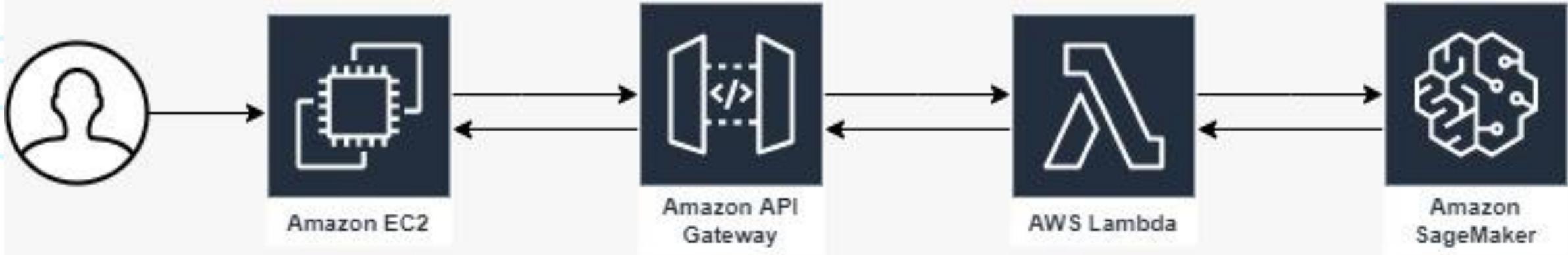
“

# Classificação de sentimento no Texto end-to-end

# Digital Products / NLP



# Digital Products / NLP



## Website

templates

App.py

Requirements.txt

## Sagemaker

SageMaker Project.ipynb

### Train

Model.py

train.py

Requirements.txt

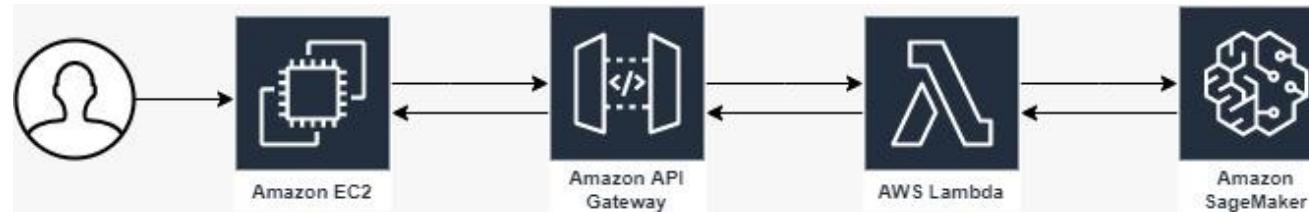
### Serve

Model.py

Predict.py

Utils.py

Requirements.txt



# Digital Products / NLP



Data &amp; AI

Amazon SageMaker

us-east-2.console.aws.amazon.com/sagemaker/home?region=us-east-2#/notebook-instances

AWS Services Search for services, features, blogs, docs, and more [Alt+S]

Ohio vfcarida

Amazon SageMaker X

Amazon SageMaker > Notebook instances

Notebook instances

Search notebook instances

Name Instance Creation time Status Actions

There are currently no resources.

Dashboard

Search

SageMaker Domain

Studio

RStudio

Canvas

Images

▶ Ground Truth

▼ Notebook

Notebook instances (selected)

Lifecycle configurations

Git repositories

▶ Processing

▶ Training

▶ Inference

The screenshot shows the Amazon SageMaker console interface. A red box highlights the 'Amazon SageMaker' tab in the top navigation bar. Another red box highlights the 'Notebook instances' link under the 'Notebook' section of the left sidebar. A third red box highlights the 'Create notebook instance' button in the top right of the main content area. The sidebar also includes sections for 'Ground Truth', 'Lifecycle configurations', 'Git repositories', 'Processing', 'Training', and 'Inference'.

# Digital Products / NLP

Amazon SageMaker

us-east-2.console.aws.amazon.com/sagemaker/home?region=us-east-2#/notebook-instances/create

Services Search for services, features, blogs, docs, and more [Alt+S]

Amazon SageMaker > Notebook instances > Create notebook instance

## Create notebook instance

Amazon SageMaker provides pre-built fully managed notebook instances that run Jupyter notebooks. The notebook instances 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:

Elastic Inference [Learn more](#):

**Amazon SageMaker Notebook Instance is ending its standard support on Amazon Linux AMI (AL1). [Learn more](#)**

Amazon SageMaker

us-east-2.console.aws.amazon.com/sagemaker/home?region=us-east-2#/notebook-instances/create

Services Search for services, features, blogs, docs, and more [Alt+S]

Encryption key - optional  
Encrypt your notebook data. Choose an existing KMS key or enter a key's ARN.

► Network - optional

▼ Git repositories - optional

▼ Default repository  
Repository  
Jupyter will start in this repository. Repositories are added to your home directory.  
 [G](#)

Git repository URL  
Clone a repository to use for this notebook instance only.

Add additional repository

► Tags - optional

# Digital Products / NLP

Not secure | nilc.icmc.usp.br/nilc/index.php/repositorio-de-word-embeddings-do-nilc

Início Downloads Como Utilizar

## Download Word Embeddings Pré-treinadas

Para cada modelo, foram disponibilizados vetores de palavras gerados em várias dimensões. Alguns modelos como Word2vec, FastText e Wang2vec possuem as variações CBOW e Skip-Gram, que diferenciam-se pela forma como preveem as palavras. Em "Ver Detalhes" pode-se ter acesso à rotinas de pré-processamento, limpeza e avaliação. No córpus, foram feitas tratativas de tokenização, remoção de stopwords, stemmização e outras.

### Word2Vec

#### Modelo

- CBOW 50 dimensões
- CBOW 100 dimensões
- CBOW 300 dimensões**
- CBOW 600 dimensões
- CBOW 1000 dimensões
- SKIP-GRAM 50 dimensões
- SKIP-GRAM 100 dimensões
- SKIP-GRAM 300 dimensões
- SKIP-GRAM 600 dimensões
- SKIP-GRAM 1000 dimensões

[Ver Detalhes »](#)

#### Corpora STIL 2017

- [download](#)

### FastText

#### Modelo

- CBOW 50 dimensões
- CBOW 100 dimensões
- CBOW 300 dimensões
- CBOW 600 dimensões
- CBOW 1000 dimensões
- SKIP-GRAM 50 dimensões
- SKIP-GRAM 100 dimensões
- SKIP-GRAM 300 dimensões
- SKIP-GRAM 600 dimensões
- SKIP-GRAM 1000 dimensões

[Ver Detalhes »](#)

#### Corpora STIL 2017

- [download](#)

### Wang2Vec

#### Modelo

- CBOW 50 dimensões
- CBOW 100 dimensões
- CBOW 300 dimensões

#### Corpora STIL 2017

- [download](#)
- [download](#)
- [download](#)

### Glove

#### Modelo

- GLOVE 50 dimensões
- GLOVE 100 dimensões
- GLOVE 300 dimensões

#### Corpora STIL 2017

- [download](#)
- [download](#)
- [download](#)

# Digital Products / NLP

testedc.notebook.us-east-2.sagemaker.aws/notebooks/NLP\_TDC2021/NLP-producao/sagemaker/SageMaker%20Project.ipynb#

Jupyter SageMaker Project Last Checkpoint: 11 minutes ago (unsaved changes)

File Edit View Insert Cell Kernel Widgets Help

Run Markdown git nbdiff

Not Trusted conda\_pytorch\_p36

## Criando um Web App de Análise de Sentimento

### Usando PyTorch e SageMaker

*Esse tutorial foi feito com base no projeto final do Nanodegree em Deep Learning da Udacity*

Agora que vocês já conhecem um pouco de AWS e de Sagemaker, iremos construir um projeto completo (end to end) com o intuito de desenvolver um web app onde o usuário insira um texto de avaliação de filme e saiba se essa avaliação é positiva ou negativa.

### Linhas Gerais

Esboço geral para projetos SageMaker usando uma instância de notebook.

1. Baixe ou recupere os dados.
2. Processe / prepare os dados.
3. Faça upload dos dados processados para S3.
4. Treine um modelo escolhido.
5. Teste o modelo treinado (normalmente usando um trabalho de transformação em lote).
6. Implante o modelo treinado.
7. Use o modelo implantado.

Para este projeto, seguiremos todas as etapas gerais

In [1]: # Make sure that we use SageMaker 1.x  
!pip install sagemaker==1.72.0

Collecting sagemaker==1.72.0  
Download complete: 1.72.0.tar.gz (207 kB)

# Digital Products / NLP

testedc.notebook.us-east-2.sagemaker.aws/notebooks/NLP\_TDC2021/NLP\_BERT.ipynb#

jupyter NLP\_BERT Last Checkpoint: a few seconds ago (autosaved) Logout

In [1]:

```
#%%capture
!pip install torchvision cudatoolkit=10.0 -f
!pip install -U sentence-transformers
!pip install transformers
!pip install datasets

thon-dateutil>=2.7.3->pandas->datasets) (1.16.0)
Installing collected packages: tqdm, fsspec, xxhash, datasets
  Attempting uninstall: tqdm
    Found existing installation: tqdm 4.61.1
    Uninstalling tqdm-4.61.1:
      Successfully uninstalled tqdm-4.61.1
  Attempting uninstall: fsspec
    Found existing installation: fsspec 2021.4.0
    Uninstalling fsspec-2021.4.0:
      Successfully uninstalled fsspec-2021.4.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour
is the source of the following dependency conflicts.
fastai 1.0.61 requires nvidia-ml-py3, which is not installed.
spacy 3.0.6 requires pydantic<1.8.0,>=1.7.1, but you have pydantic 1.8.2 which is incompatible.
s3fs 2021.4.0 requires fsspec==2021.04.0, but you have fsspec 2021.11.1 which is incompatible.
Successfully installed datasets-1.16.1 fsspec-2021.11.1 tqdm-4.62.3 xxhash-2.0.2
WARNING: You are using pip version 21.2.4; however, version 21.3.1 is available.
You should consider upgrading via the '/home/ec2-user/anaconda3/envs/pytorch_p36/bin/python -m pip install --upgrade pip' command.
```

In [2]:

```
import torch
from transformers import BertTokenizerFast, BertModel

from transformers import logging
logging.set_verbosity_error()

model_weights = 'bert-base-multilingual-cased'
```

# Digital Products / NLP

← → C testetdc.notebook.us-east-2.sagemaker.aws/notebooks/NLP\_TDC2021/NLP\_intro.ipynb#

jupyter NLP\_intro Last Checkpoint: 2 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted conda\_tensorflow\_p36

Código da apresentação de NLP

In [1]: #%%capture  
!pip install -U pip setuptools wheel  
!pip install -U spacy  
!python -m spacy download en\_core\_web\_sm  
!python -m spacy download en  
!pip install gensim  
!pip install awswrangler

```
Requirement already satisfied: packaging in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from redshift-connector<2.1.0,>=2.0.887->awswrangler) (4.9.3)
Requirement already satisfied: requests<2.26.1,>=2.23.0 in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from redshift-connector<2.1.0,>=2.0.887->awswrangler) (21.2)
Requirement already satisfied: charset-normalizer~2.0.0 in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from redshift-connector<2.1.0,>=2.0.887->awswrangler) (2.0.7)
Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from requests<2.26.1,>=2.23.0->redshift-connector<2.1.0,>=2.0.887->awswrangler) (3.3)
Requirement already satisfied: asn1crypto>=1.4.0 in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from scrapy>=1.4.1->pg8000<1.22.0,>=1.16.0->awswrangler) (1.4.0)
Requirement already satisfied: pyparsing<3,>=2.0.2 in /home/ec2-user/anaconda3/envs/tensorflow_p36/lib/python3.6/site-packages (from packaging->redshift-connector<2.1.0,>=2.0.887->awswrangler) (2.4.7)
Installing collected packages: scrapy, python-utils, requests-aws4auth, redshift-connector, pymysql, progressbar2, pg8000, opensearch-py, jsonpath-ng, awswrangler
Successfully installed awswrangler-2.12.1 jsonpath-ng-1.5.3 opensearch-py-1.0.0 pg8000-1.21.3 progressbar2-3.55.0 pymysql-1.0.2 python-utils-2.5.6 redshift-connector-2.0.900 requests-aws4auth-1.1.1 scramp-1.4.1
```

# Digital Products / NLP

console.aws.amazon.com/iamv2/home#/roles

aws Services Search for services, features, blogs, docs, and more [Alt+S] Global vfa...

**Identity and Access Management (IAM)**

Introducing the new Roles list experience  
We've redesigned the Roles list experience to make it easier to use. [Let us know what you think.](#)

IAM > Roles

**Roles (8) Info**

An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.

Search

Role name Trusted entities Last activity

Role name	Trusted entities	Last activity
acesso_s3	AWS Service: sagemaker	27 minutes ago
AmazonSageMaker-ExecutionRole-20200812T222132	AWS Service: sagemaker	111 days ago
AmazonSageMakerServiceCatalogProductsLaunchRole	AWS Service: servicecatalog	-
AmazonSageMakerServiceCatalogProductsUserRole	AWS Service: glue, and 9 more. ↗	1 hour ago
AWSServiceRoleForAmazonSageMakerNotebooks	AWS Service: sagemaker (Service-Linked Role)	111 days ago
AWSServiceRoleForAPIGateway	AWS Service: ops.apigateway (Service-Linked Role)	-
AWSServiceRoleForSupport	AWS Service: support (Service-Linked Role)	-
AWSServiceRoleForTrustedAdvisor	AWS Service: trustedadvisor (Service-Linked Role)	-

# Digital Products / NLP

← → C console.aws.amazon.com/iam/home#/roles\$new?step=type

aws Services Search for services, features, blogs, docs, and more [Alt+S]

1 2 3 4

## Create role

### Select type of trusted entity

**AWS service**  
EC2, Lambda and others

**Another AWS account**  
Belonging to you or 3rd party

**Web identity**  
Cognito or any OpenID provider

**SAML 2.0 federation**  
Your corporate directory

Allows AWS services to perform actions on your behalf. [Learn more](#)

### Choose a use case

#### Common use cases

**EC2**  
Allows EC2 instances to call AWS services on your behalf.

**Lambda**  
Allows Lambda functions to call AWS services on your behalf.

Or select a service to view its use cases

API Gateway	CloudWatch Events	EMR	IoT SiteWise	RAM
AWS Backup	CodeBuild	EMR Containers	IoT Things Graph	RDS
AWS Chatbot	CodeDeploy	ElastiCache	KMS	Redshift
AWS Marketplace	CodeGuru	Elastic Beanstalk	Kinesis	Rekognition
AWS Support	CodeStar Notifications	Elastic Container Registry	Lake Formation	RoboMaker

\* Required

Cancel Next: Permissions

# Digital Products / NLP

us-east-2.console.aws.amazon.com/lambda/home?region=us-east-2#/functions

AWS Services Search for services, features, blogs, docs, and more [Alt+S]

AWS Lambda Functions

Lambda > Functions

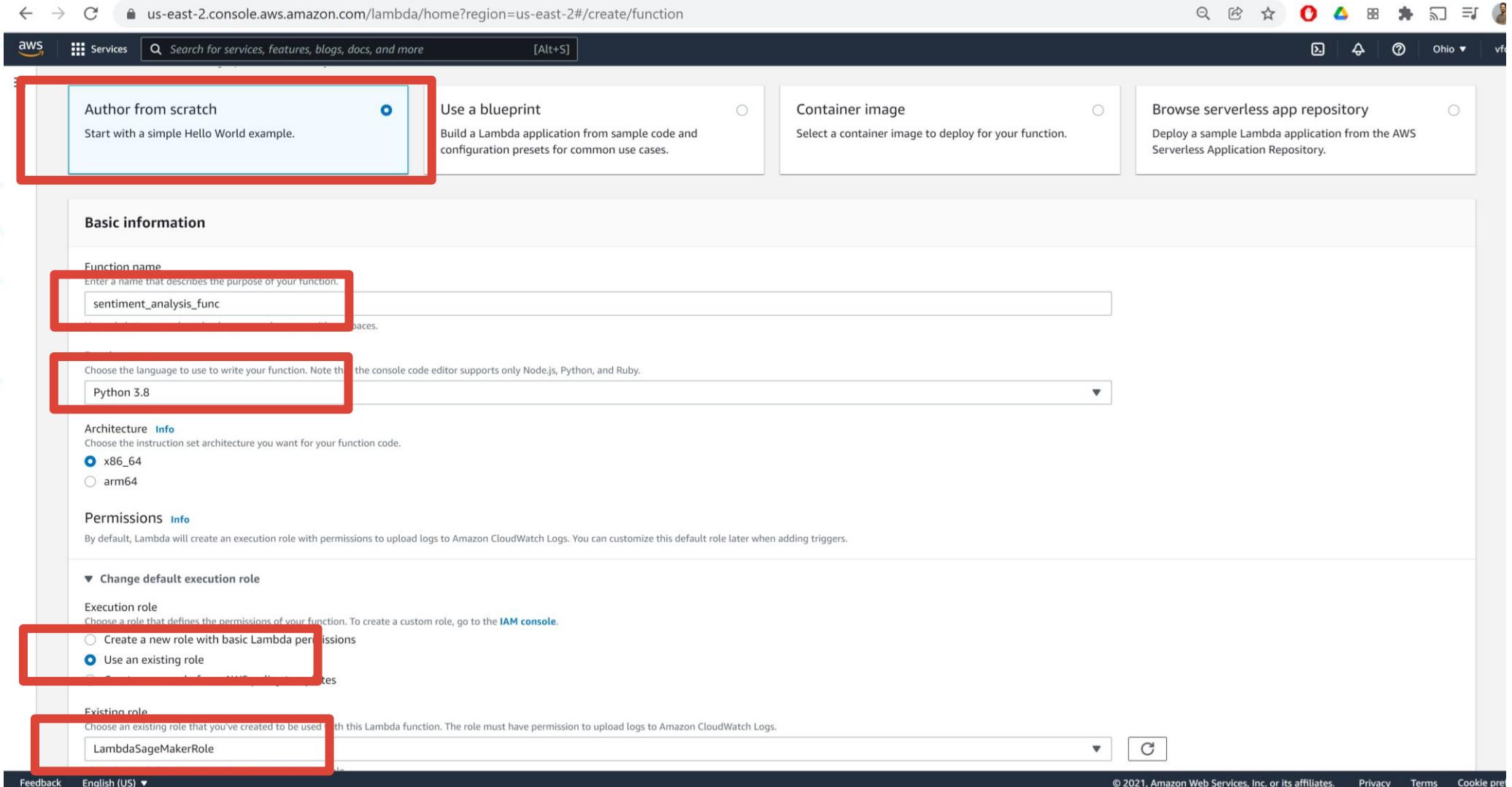
Functions (0) Last fetched in 0 seconds Actions Create function

Filter by tags and attributes or search by keyword

Function name	Description	Package type	Runtime	Code size	Last modified
There is no data to display.					

Dashboard Applications Functions Additional resources Code signing configurations Layers Related AWS resources Step Functions state machines

# Digital Products / NLP



The screenshot shows the AWS Lambda console interface for creating a new function. The page is titled "Create function" and includes the following sections:

- Author from scratch**: Start with a simple Hello World example. This section is highlighted with a red box.
- Use a blueprint**: Build a Lambda application from sample code and configuration presets for common use cases.
- Container image**: Select a container image to deploy for your function.
- Browse serverless app repository**: Deploy a sample Lambda application from the AWS Serverless Application Repository.

**Basic information**

- Function name**: sentiment\_analysis\_func
- Language**: Python 3.8

**Architecture**: Info

- x86\_64 (selected)
- arm64

**Permissions**: Info

By default, Lambda will create an execution role with permissions to upload logs to Amazon CloudWatch Logs. You can customize this default role later when adding triggers.

**Change default execution role**

Execution role

- Create a new role with basic Lambda permissions (radio button selected)
- Use an existing role

**Existing role**

- LambdaSageMakerRole

Feedback English (US) © 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie pre

# Digital Products / NLP

us-east-2.console.aws.amazon.com/apigateway/main/apis?region=us-east-2



## API Gateway

- APIs
- Custom domain names
- VPC links

### Choose an API type

#### HTTP API

Build low-latency and cost-effective REST APIs with built-in features such as OIDC and OAuth2, and native CORS support.

Works with the following:  
Lambda, HTTP backends

[Import](#)[Build](#)

#### WebSocket API

Build a WebSocket API using persistent connections for real-time use cases such as chat applications or dashboards.

Works with the following:  
Lambda, HTTP, AWS Services

[Build](#)

#### REST API

Develop a REST API where you gain complete control over the request and response along with API management capabilities.

Works with the following:  
Lambda, HTTP, AWS Services

[Import](#)[Build](#)

#### REST API Private

Create a REST API that is only accessible from within a VPC.

Works with the following:  
Lambda, HTTP, AWS Services

[Import](#)[Build](#)

# Digital Products / NLP

us-east-2.console.aws.amazon.com/apigateway/home?region=us-east-2#/apis/create

AWS Services Search for services, features, blogs, docs, and more [Alt+S]

Amazon API Gateway APIs > Create Show all hints ?

### Choose the protocol

Select whether you would like to create a REST API or a WebSocket API.

REST  WebSocket

### Create new API

In Amazon API Gateway, a REST API refers to a collection of resources and methods that can be invoked through HTTPS endpoints.

New API  Import from Swagger or Open API 3  Example API

### Settings

Choose a friendly name and description for your API.

API name\* apitdc

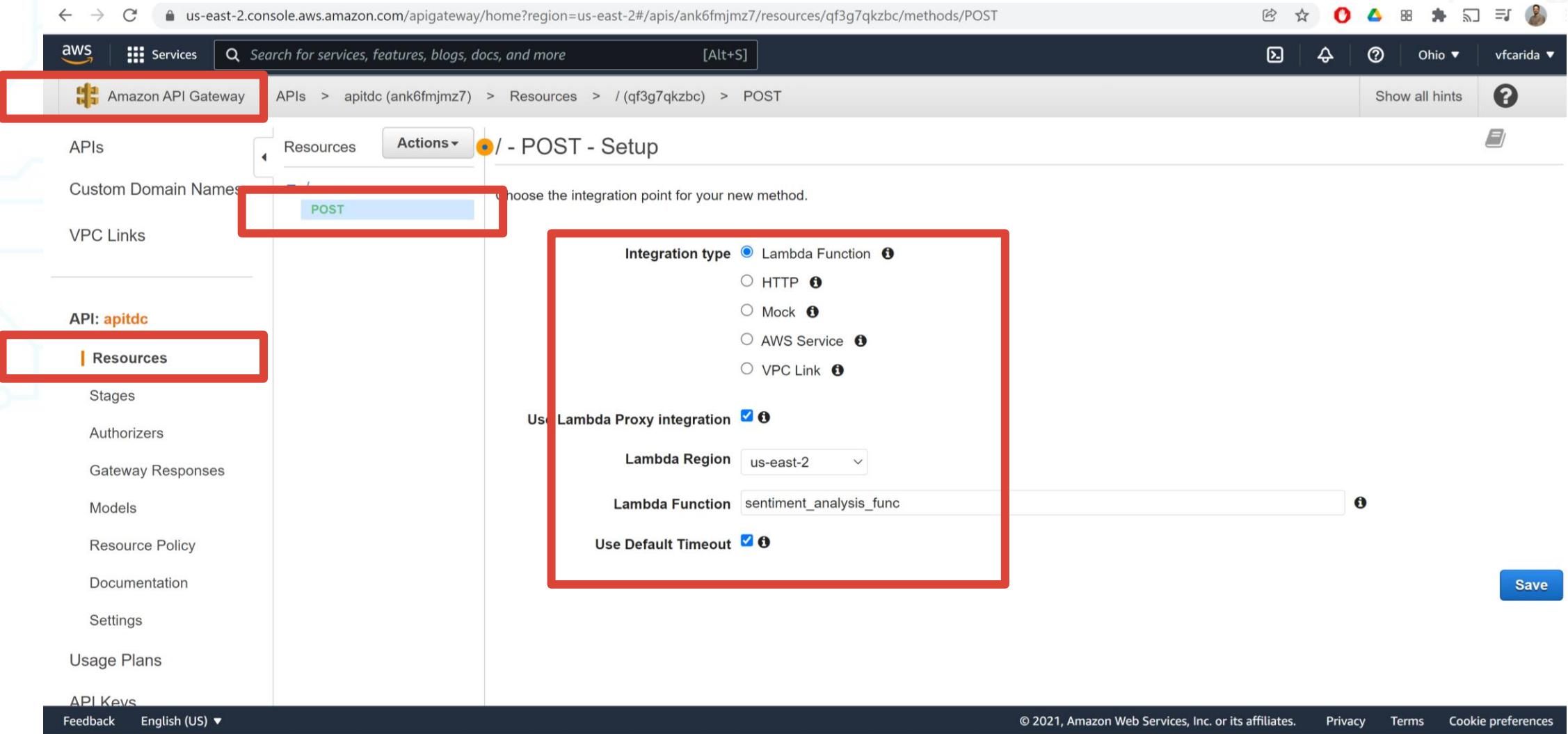
Description

Endpoint Type Regional

\* Required

Create API

# Digital Products / NLP



The screenshot shows the AWS API Gateway interface for setting up a POST method. The URL in the browser is `us-east-2.console.aws.amazon.com/apigateway/home?region=us-east-2#/apis/ank6fmjmz7/resources/qf3g7qkzbc/methods/POST`.

**Left Sidebar:** Shows the navigation menu with 'APIs' selected, followed by 'Resources', 'Actions', 'POST', 'Custom Domain Names', and 'VPC Links'. Below this, under 'API: apitdc', are 'Resources' (selected), 'Stages', 'Authorizers', 'Gateway Responses', 'Models', 'Resource Policy', 'Documentation', 'Settings', and 'Usage Plans'.

**Top Bar:** Shows the AWS logo, Services search bar, and user profile 'vfcarida'.

**Central Content:** The main area shows the 'Actions' dropdown set to 'POST - Setup'. It displays the 'Integration type' section, which is currently set to 'Lambda Function' (radio button selected). Other options include 'HTTP', 'Mock', 'AWS Service', and 'VPC Link'. Below this, there is a checkbox for 'Use Lambda Proxy integration' (checked) and dropdowns for 'Lambda Region' (set to 'us-east-2') and 'Lambda Function' (set to 'sentiment\_analysis\_func'). There is also a checkbox for 'Use Default Timeout' (checked).

**Bottom Right:** A blue 'Save' button is located at the bottom right of the configuration area.

# Digital Products / NLP



us-east-2.console.aws.amazon.com/ec2/v2/home?region=us-east-2#Instances:

New EC2 Experience  Tell us what you think X

Instances Info C Connect Instance state Actions  Launch instances ▾

Filter instances

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
You do not have any instances in this region							

Instances New

- Instance Types
- Launch Templates
- Spot Requests
- Savings Plans
- Reserved Instances New
- Dedicated Hosts
- Capacity Reservations

Images

- AMIs

Elastic Block Store

Select an instance above

Feedback English (US) ▾ © 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

# Digital Products / NLP

Sentiment Analysis Web App

Not secure | 3.138.112.109

Incognito (2)

## Is your review positive, or negative?

Enter your review below and click submit to find out...

**Review:**

Submit

# Embedding / NLP

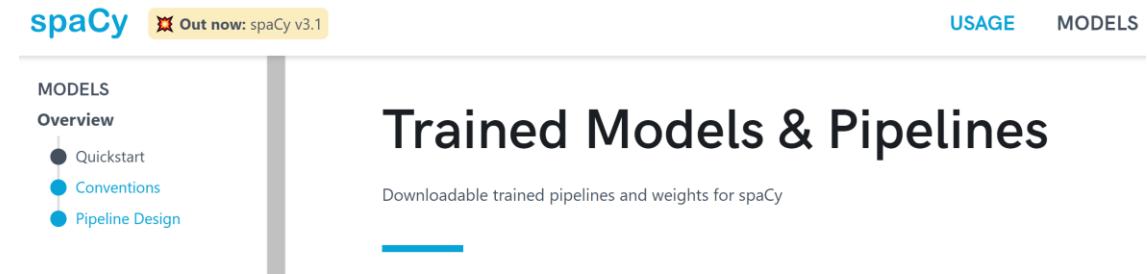
<https://www.nltk.org/>

## NLTK 3.6.2 documentation

[NEXT](#) | [MODULES](#) | [INDEX](#)

## Natural Language Toolkit

<https://spacy.io/models>



The screenshot shows the spaCy v3.1 "Trained Models & Pipelines" page. At the top, it says "Out now: spaCy v3.1". Below that is a sidebar with "MODELS" and "Overview" sections, including links for "Quickstart", "Conventions", and "Pipeline Design". The main content area is titled "Trained Models & Pipelines" and describes "Downloadable trained pipelines and weights for spaCy".

[https://www.tensorflow.org/text/tutorials/classify\\_text\\_with\\_bert](https://www.tensorflow.org/text/tutorials/classify_text_with_bert)

TensorFlow > Resources > Text > Tutorials

Was this helpful  

## Text classification with an RNN

 Run in Google Colab

 View source on GitHub

 Download notebook

This text classification tutorial trains a [recurrent neural network](#) on the [IMDB large movie review dataset](#) for sentiment analysis.

[https://www.tensorflow.org/text/tutorials/classify\\_text\\_with\\_bert](https://www.tensorflow.org/text/tutorials/classify_text_with_bert)

TensorFlow > Resources > Text > Tutorials

Rate and review  

## Classify text with BERT

 Run in Google Colab

 View on GitHub

 Download notebook

 See TF Hub model

This tutorial contains complete code to fine-tune BERT to perform sentiment analysis on a dataset of plain-text IMDB movie reviews. In addition to training a model, you will learn how to preprocess text into an appropriate format.

# Thanks !



Vinicius Fernandes Caridá

[vfcarida@gmail.com](mailto:vfcarida@gmail.com)



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida