AutoML Blueprint Quick Start Guide

Author: Dylan Tong, AWS

The AutoML Blueprint is a no-code, serverless workflow that performs data preparation, automated machine learning model delivery, and validation.



Amazon SageMaker Data Wrangler is used for data preparation. It empowers customers to create nocode data flows. Experimentation and model building are automated using Amazon SageMaker Autopilot, a fully-managed tabular AutoML service.

Predictions are automatically batched to support granular error analysis. Amazon SageMaker Clarify provides bias detection and model explainability (XAI). A notebook is packaged along with the solution to guide you through the model inspection process and equip you with tools to help get your consumers to trust your model predictions.

Users deploy this blueprint using CloudFormation to create the AWS StepFunctions workflow above. The user is required to provide the following:

• An Amazon SageMaker DataWrangler flow definition file created through the GUI.

- A dataset for model training and validation.
- Blueprint configurations. The default configurations can be downloaded here.

The blueprint delivers an AutoML optimized model along with pre-processed predictions, bias metrics and XAI results that you can use for detailed model inspection.

Quick Launch Steps

1. Launch the CloudFormation template using this S3 URL:

https://dtong-public-fileshare.s3-us-west-2.amazonaws.com/automlblueprint/code/deploy/cf/automl-blueprint.yml

Provide a stack name and use the provided *default settings*.

cloudronnation / stacks /	concrete on the second se
Step 1 Specify template	Create stack
Step 2 Specify stack details	Prerequisite - Prepare template
Step 3 Configure stack options	Prepare template Every stack is based on a template. A template is a JSON or YAMI. If le that contains configuration information about the AWS resources you want to include in the stack.
Step 4 Review	Template is ready Use a sample template Create template in Designer
	Specify template A template is a JSON or YAML file that describes your stack's resources and properties.
	Template source Selecting a template generates an Amazon S3 URL where it will be stored.
	Amazon S3 URL Upload a template file
	Amazon S3 URL
	https://dtong-public-fileshare.s3-us-west-2.amazonaws.com/automi-blueprint/code/deploy/cf/automi-blueprint.yml Amazon S3 template URL
	S3 URL: https://dtong-public-fileshare.s3-us-west-2.amazonaws.com/automl-blueprint/code/deploy/cf/autom View in I-blueprintyml
	Court Here

Check the box at the end to acknowledge that this template will create IAM resources.



 Launch Amazon SageMaker Studio if you haven't done so already. You need to ensure that the IAM role attached to your Studio instance has permissions to run Amazon S3, Amazon SageMaker and AWS StepFunctions.

aws Services 🔻	Q	Search for services, features, marketplace	products, and docs [Op	otion+S] 🗘 🎝 DylanTong @	8032-3586-9972 ▼ Singapore ▼ Support ▼		
Amazon SageMaker	×	User name \bigtriangledown	Last modified	Created \bigtriangledown			
Amazon Sagenaker	~	default-1615426817523	Mar 11, 2021 01:43 UTC	Mar 11, 2021 01:42 UTC	Open Studio [2]		
Amazon SageMaker Studio							
Dashboard		 Studio Summary 		How to delete Studio	Delete Studio Edit Settings		
Search							
Images		Status Ready	Studio ID d-or97xlava0fx	Execution role	Authentication method		
 Ground Truth 		() heady	a oror nardony		Management (IAM)		
Notebook							
Processing							
Training		Use the Studio ID for troubleshooti	Use the Studio ID for troubleshooting and tracking usage.				
Inference				and of compute resources such as even in			
Augmented AI		Projects					
AWS Marketplace		Launch constraint role: arn	aws:iam::803235869972:role	/service-role/AmazonSageMakerSer	viceCatalogProductsLaunchRole		
		Product use role: arn:aws:ia	am::803235869972:role/servie templates enabled for Studio	ce-role/AmazonSageMakerServiceCa o users	talogProductsUseRole		
		Execution role: arn:aws:iam	::803235869972:role/service	-role/AmazonSageMaker-ExecutionR	tole-20210302T182009		
		Custom images attach	ed to domain How to	attach SageMaker images 🗹 🛛 📿	Detach Attach image		
		You can attach up to 30 image request will fail. To attach add	es. Only the latest version for a litional images, you need to de	each image is displayed below. If you etach an image and all its versions ar	a try to attach more than 30 images, the nd then try again.		
		Name 🔺 D	isplayName ⊽ Li	atest attached version \bigtriangledown	Created V Actions		
Feedback English (US) 🔻			© 2008 - 2021, Amazon Web Serv	ices, Inc. or its affiliates. All rights reserved.	Privacy Policy Terms of Use Cookie preferences		

Navigate to the IAM console and attach the following managed policies.

Permissions	Trust relationships	Tags	Access Advisor	Revoke sessions		
 Permissions policies (3 policies applied) 						
Attach polic	ies					
Policy	Policy name -					
Policy						
Ama	azonS3EullAccess					
 Ama Awa 	azonS3FullAccess SStepFunctionsFullAccess	6				

3. Login to Amazon SageMaker Studio. Clone the following Github <u>repository</u>: <u>https://github.com/aws-samples/automl-blueprint.git</u>



4. Under automl-blueprint/notebook, open the Jupyter notebook user_guide.ipynb. Use the Python3 (Data Science) kernel.

\otimes	Amazon SageMaker Studio File	Edit View Run Kerni	Start Preferred Kernel	12	Feedback
		🕅 user_guide.ipynb	PySpark (SparkMagic)		
			Python 3 (Base Python)		
			Python 3 (Data Science)		
			Python 3 (MXNet 1.6 Python 3.6 CPU Optimized)		
		Auton	Python 3 (MXNet 1.6 Python 3.6 GPU Optimized)	it also	
		Αυτορ	Python 3 (MXNet 1.8 Python 3.7 CPU Optimized)	llae	
0			Python 3 (MXNet 1.8 Python 3.7 GPU Optimized)		
		This noteboo	Bython 3 (PyTorch 14 Python 3.6 CPU Optimized)	(blueprint) to deliver an optimized ML model. The	
		workflow use	Python 3 (PyTorch 1.6 Python 3.6 CPU Optimized)		
			Python 3 (PyTorch 16 Python 36 GPU Optimized)		
		Amazon	Python 3 (SageMaker JumpStart Data Science 1.0)	nterface for creating data prep flows.	
		Amazon	Python 3 (SageMaker JumpStart MXNet 1.0)	abular data through a fully managed experience.	
		Amazon	Python 3 (SageMaker JumpStart PyTorch 1.0)	e Trusted Al.	
			Python 3 (SageMaker JumpStart TensorFlow 1.0)		
		Leaning	Python 3 (TensorFlow 1.15 Python 3.6 CPU Optimized)		
			Python 3 (TensorFlow 1.15 Python 3.6 GPU Optimized)		
		 Learn 	Python 3 (TensorFlow 1.15 Python 3.7 CPU Optimized)		
		Learn	Python 3 (TensorFlow 1.15 Python 3.7 GPU Optimized)		
		 Deploy 	Python 3 (TensorFlow 2.1 Python 3.6 CPU Optimized)	your analytics systems.	
			Python 3 (TensorFlow 2.1 Python 3.6 GPU Optimized)		
		The Data	Python 3 (TensorFlow 2.3 Python 3.7 CPU Optimized)		
			Python 3 (TensorFlow 2.3 Python 3.7 GPU Optimized)		
		This noted	Use No Kernel	dataset contains the results of a marketing	
		campaign to	V No Kernel	Mation about a prospect. The last column is the	
		target variab	Use Kernel from Preferred Session	the term deposit.	
			Use Kernel from Other Session		
		Citation: Mor	uci-bank-marketing-dataset flow	ach to Predict the Success of Bank Telemarketing.	
		Decision Sup	untitled flow		
			undicanow		
1	🛐 0 🌐 Git: idle 🛛 No Kernel Unk	known		Mode: Command 🛞 Ln 1, Col 1 user_	_guide.ipynb

5. Confirm that your CloudFormation template has successfully launched.

Stacks (1)	C	Delete Update Sta	ck actions 🔻 Create stack 🔻
Q Filter by stack name		Active	View nested
Stack name	Status	Created time	Description
) autopilot-bp	⊘ CREATE_COMPLETE	2021-04-08 15:36:43 UTC-070	CloudFormation template to deploy the Autopilot Blueprint

You're good to go! Follow the steps in the notebook.