Microsoft Azure Ma	Ichine Learning		🎯 🕞 🗘 ?	Someone@contoso.com
	Model Monitor > Model List > Model Detail			
Monitor	RandomeForestClassifier (TEST_3_try_2)			
I BE Endpoints	Detail Info of some modification			
La Alerts	RandomForestClassifier (TEST_3_try_2) \checkmark 🔿 Refresh	Edit Blocks		Generate report
Reports Description Reports	Overview Feature			
	(Dn) d de ted.			\times
	Prediction Distribution A Durrent Witching 08/21/2021	Deploy Info		Drift Info
	0.625	Deployed at	GPU utilization	Drift Count
		10/09/2020	85%	3
		Running since last modification	GPU utilization	Prediction Drift (1)
	O Class Class Class Class Class	21d 14:04	78%	Feature Drift (2)
rovide essential information to aid decision-	makings	*		
	Feature Distribution (Numeric)	raining 08/21/2021 Feature Di	stribution (Categorical)	Current Training 08/21/2021
	0.5	<u>≥</u> 0.75		
	20.25 0.25	ige 0.5		
		č 0.25		Feature
	C Feature num 🔥 Feature num Feature	Feature_	Feature_cate Feature_	cate reature_cate reature.



Once a machine learning model is released into the wild, it works with new data. The model performance will inevitably decline over time as data deviates from what the model has seen in training.

Machine learning engineers lack a leading signal to help determine if/when to retrain their models and get the most up-to-date performance statistics in the dev/test/ production environment.

Solution

Users can set up model monitoring jobs at deployment time via config files and commands, which invokes services that calculate model quality metrics and drifts.

Key Features



Users can view the endpoint list initially, then select a specific deployment to look at its detail. Users could also edit columns and blocks to customize the dashboard.



In each deployment, sections of Overview, Data, Infrastructure, Performance, Alert and Report would be provided, along with several Azure built-in features like Responsible AI. Users can see drifts, metrics on those pages.

Add report schedule

• • • Model Monitor × +

Then, users can view all metrics in one customizable dashboard. Also, we provide tools to compare the performance of different model versions.

The detailed features include:

- Model detail: model performance statistics
- Model comparison: compare multiple models' performance on one page
- Alert configuration: customize when & how to be notified of model failure
- Report generation: download and share statistics

Design Process

We followed a user-centered and customer-centered design process. We have two types of key stakeholders, the ML engineers as direct users and the manager in tech companies as the customers.

We tested the design prototype by usability tests and heuristic evaluations with Azure Machine Learning Studio engineers. We also collected advice and needs from

										Rule Name
										Daily Report Check
										Rule Range
										Last
📕 🌕 📄 📄 Model	Monitor × -	+								24 û hours
Microsoft Azure Mach	nine Learning					© 0	₽ Q ?	😳 some	one@contoso.com сомтоso	Resolution
=	Model Monito	or > Reports								Daily
Monitor	Report									Recipients
Monitor	Schedule	Report								Mona Kane X 🚇 Robin Counts X
Reports	+ Add	🔟 Delete	📅 Edit Column					Search		Kobin Counts A
 Settings 	Status	Report ↑		Range	Resolution	Recipients	Frequency	Time	Created On	Frequency
	On	Daily Report Check		Last 24 hours	Daily	4	Every workday	9:00 AM	12/02/2021	Every workday
	On	Daily Report Check		Last 24 hours	Daily	4	Every workday	9:00 AM	12/02/2021	9:00 AM
	On	Daily Report Check		Last 24 hours	Daily	4	Every workday	9:00 AM	12/02/2021	Elements
	• Off	f Daily Report Check		Last 24 hours	Daily	4	Every workday	9:00 AM	12/02/2021	SKU
										Prediected P-value
										Save

Users could generate reports from selected deployments, and even schedule automatic report generation.

rosoft Azure	Machine Learning	Learning							omeone@contoso.com сохтоso					
	Model Monitor > End	points > Deployment Comparison												
nitor	Deployment Cor	Deployment Comparison												
Monitor	Select again							💍 Refresh						
Comparison	Basic Information													
Reports Settings	🗋 Name î		SKU		Training Data Location			n Traffic						
	Blue - Endpoin CNN-model1 X-	Blue - Endpoint_12131 CNN-model1 X-ray image classification						eastus2	80%					
	Green - Endpo CNN-model2 X-	int_12131 ray image classification		Standard v2, Basic_	_DS4_ A2	<u>URL</u>		eastus2	20%					
	Endpoint_1213 CNN-model2 X-	2 ray image classification		Standard v2, Basic_	_DS4_ A2	<u>URL</u>		eastus2	0					
	Performance 01/11/2021 Image: State of the s	Performance 01/11/2021 Image: Today Yesterday 7D 30D 3M Image: Solution of the second state of the second stateo												
		Blue - Endpoint_12131	Green - Endpoint_12131	Endpoint_12132										
	Prediction P-value	0.46*	0.78	0.66										
	Accuracy	0.92	0.91	0.91										
	Precision	0.92	0.91	0.91										
	Latency	1.32ms	1.00ms	1.92ms										
	Prediction P-value	AA-AAA	A	Accuracy										

Users are allowed to select different deployments from different endpoints to compare. This feature could aid their decision-making.

Software Architecture





GLOBAL INNOVATION EXCHANGE Microsoft Azure

Jerry WangMSTIDec. 2021Qintian LiGIX Dual DegreeJuly 2022Yanzhang LiGIX Dual DegreeJuly 2022Ziqi WangGIX Dual DegreeJuly 2022