

Méthodologie d'évaluation de la performance hors ligne, de pilotage en production et de mise à jour en production du modèle du chatbot



Introduction

Fly Me est une agence qui propose des voyages clé en main pour les particuliers ou les professionnels.

Fly Me a lancé un projet ambitieux de développement d'un chatbot pour aider les utilisateurs à choisir une offre de voyage.

Nous avons donc construit un MVP pour aider les employés de Fly Me à réserver facilement un billet d'avion pour leurs vacances. Le chatbot est disponible sur

<https://agreeable-tree-09bfe3f03.2.azurestaticapps.net/>

1 - Description des critères d'évaluation du modèle

Dans ce projet, afin d'identifier l'intention de l'utilisateur, ainsi que les informations importantes permettant de répondre à sa demande, nous avons utilisé le service de compréhension du langage LUIS de Microsoft Azure. Luis est un service d'IA conversationnelle qui applique une intelligence de machine learning personnalisée au texte en langage naturel des conversations d'un utilisateur afin d'en prédire le sens général et d'en extraire des informations (entités).

Pour évaluer les performances du modèle, on mesure le pourcentage d'entités correctement détectées pour l'intention BookFlight

Détails de la transaction de bout en bout

Résultats de la recherche

Filtré par : timestamp > 27/10/2022 18:...

timestamp < 28/10/2022 18:...

Résultats de la recherche

En savoir plus

Copier le lien

Commentaires

Quitter la prévision

Transaction de bout en bout

ID d'opération : 00000000000000000000000000000000

8 Tours 8 Traces 0 Événements

[Afficher la chronologie](#)

Filtrer sur un composant spécifique et appeler

Tout [Composant] [Appel]

ERROR 18:11:12.779

The customer is not satisfied with the Bot's proposition

ERROR 18:11:12.779

The customer is not satisfied with the Bot's proposition

ERROR 18:11:45.957

The customer is not satisfied with the Bot's proposition

ERROR 18:11:45.957

The customer is not satisfied with the Bot's proposition

ERROR 18:12:43.015

The customer is not satisfied with the Bot's proposition

ERROR 18:12:43.015

The customer is not satisfied with the Bot's proposition

ERROR 18:13:39.454

The customer is not satisfied with the Bot's proposition

ERROR 18:13:39.454

The customer is not satisfied with the Bot's proposition

Créer un élément de travail

TRACE

Error

Propriétés de Trace

Tout afficher

Event time	28/10/2022 18:13:39.454 (Heure locale)	...
Device type	PC	...
Message	The customer is not satisfied with the Bot's proposition	...
Severity level	Error	...

Custom Properties

process	MainProcess	...
module	booking_dialog	...
filename	/tmp/8da67ecbe93950/dialogs/booking_dialog.py	...
lineNumber	326	...
level	ERROR	...
destination	Berlin	...
origin	Toulouse	...
start_date	2022-10-28T00:00:00.0000000Z	...
end_date	2022-10-27T00:00:00.0000000Z	...
budget	6665	...

Éléments associés

Trouver les données de Métrix disponibles 5 minutes avant et après

Nous pouvons ensuite paramétrer des alertes (dans la section Alertes) :

The screenshot shows the 'Alertes' (Alerts) section of the 'p10_insights' Application Insights resource. The left sidebar contains navigation links for various monitoring features. The main area displays a summary of alert counts and a table of active alerts.

Nombre total d'alertes	Critique	Erreur	Avertissement	Information	Commentaires
1	0	0	0	1	0

Nom ↑	Gravité ↑	Condition d'alerte ↑	Réponse de l'utilisateur ↑	Temps d'incendie ↑
Chatbot Alert	3 - Information	Déclenché	Nouveau	28/10/2022 18:16

Ces alertes peuvent déclencher un envoi d'emails et/ou de SMS. Par exemple, ici l'alerte Chatbot Alert est déclenchée lorsqu'un utilisateur du chatbot clique sur "non" trois fois d'affilée lors du prompt de confirmation des informations pour le booking :

The screenshot shows a chatbot interface for 'Fly Me'. The chatbot asks for departure and return dates, then asks for a travel budget. It then displays a confirmation summary and asks for confirmation. The 'Nope' button is highlighted with a red arrow, indicating the action that triggers the alert.

What is your desired departure date ? (example: 10/28/2022 is today)

À l'instant

today

À l'instant

What is your desired return date ? (example: 10/28/2022 is today)

À l'instant

tomorrow

À l'instant

Now, let's talk about money... what's your travel budget?

(example: 3.14€)

À l'instant

3.14\$

À l'instant

Please confirm :

Departure : Toulouse

Destination : Tchernobyl

Starting on: 2022-10-28

Ending on: 2022-10-29

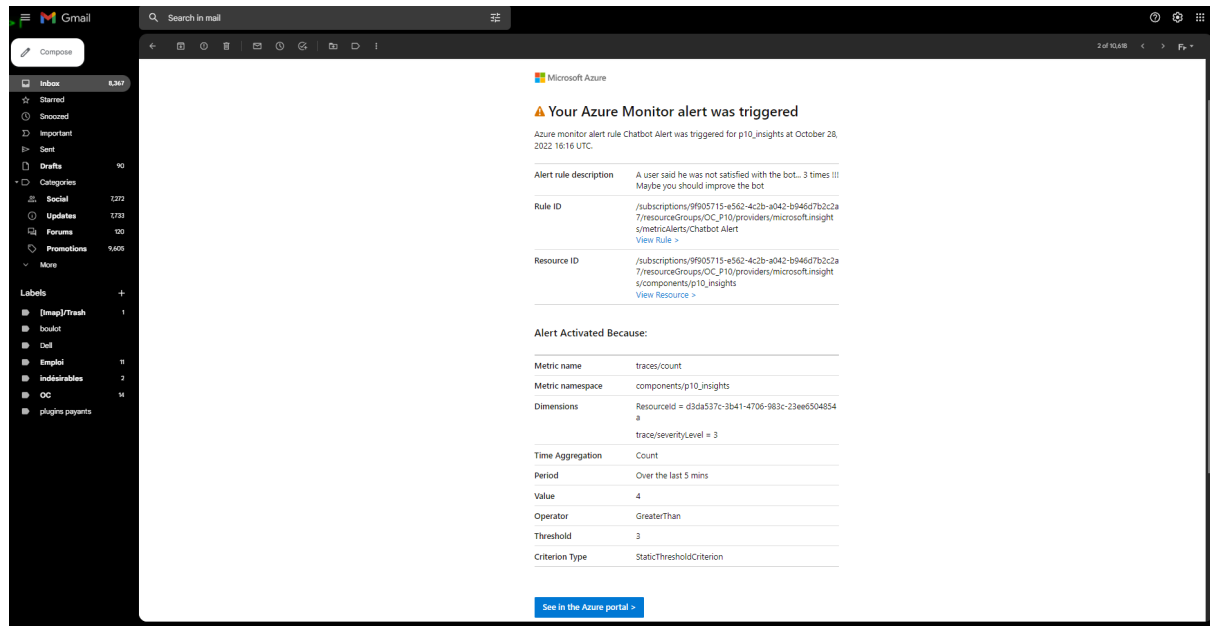
Budget: 3.14\$.

À l'instant

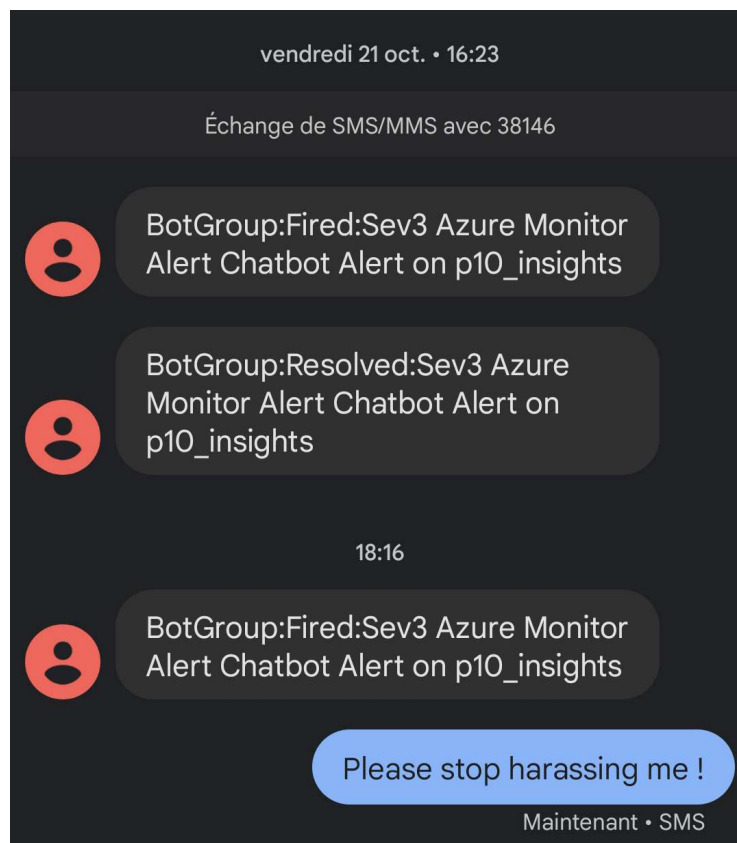
Yep Nope

Tapez votre message

un email est envoyé :

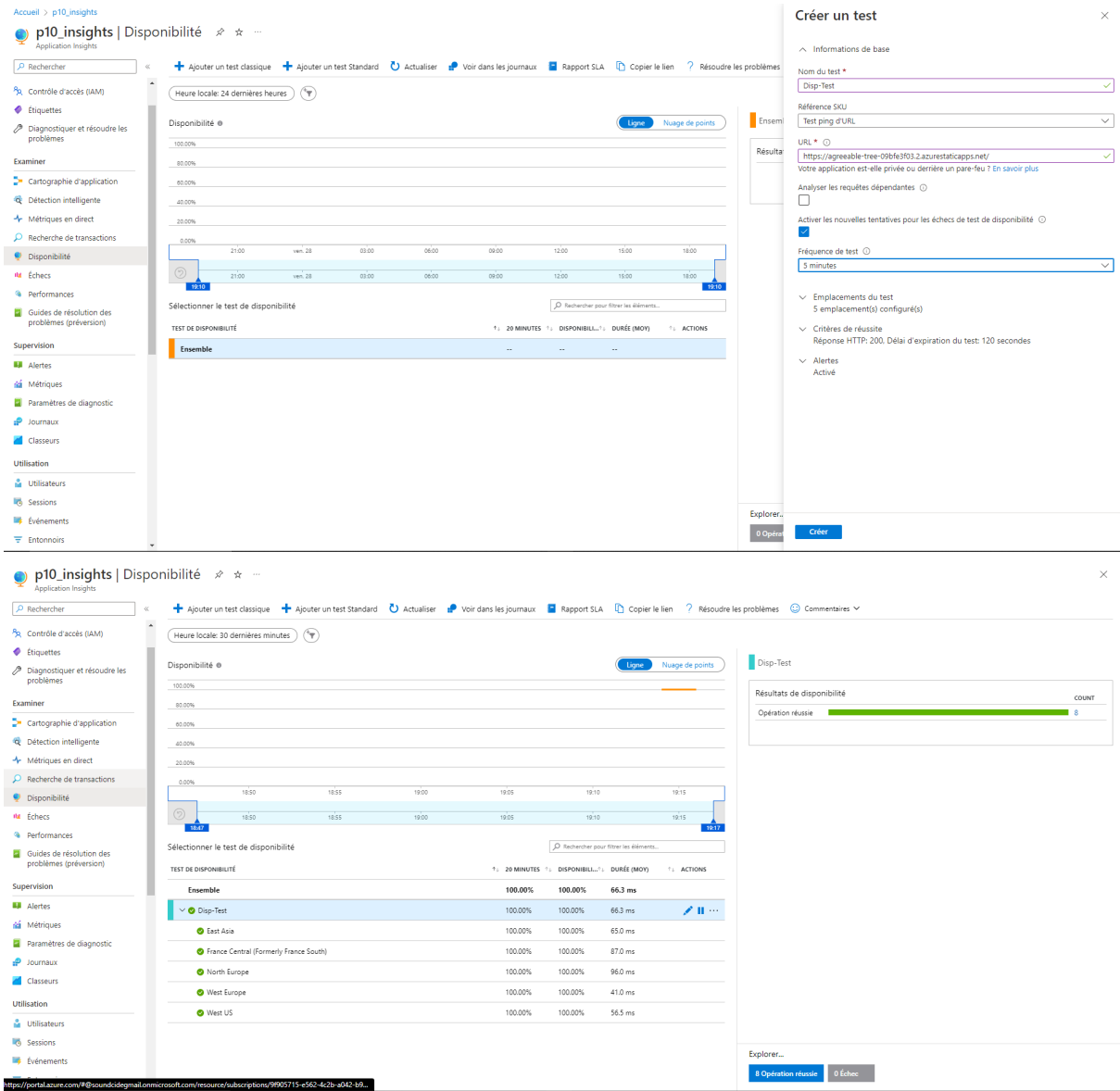


ainsi qu'un SMS :

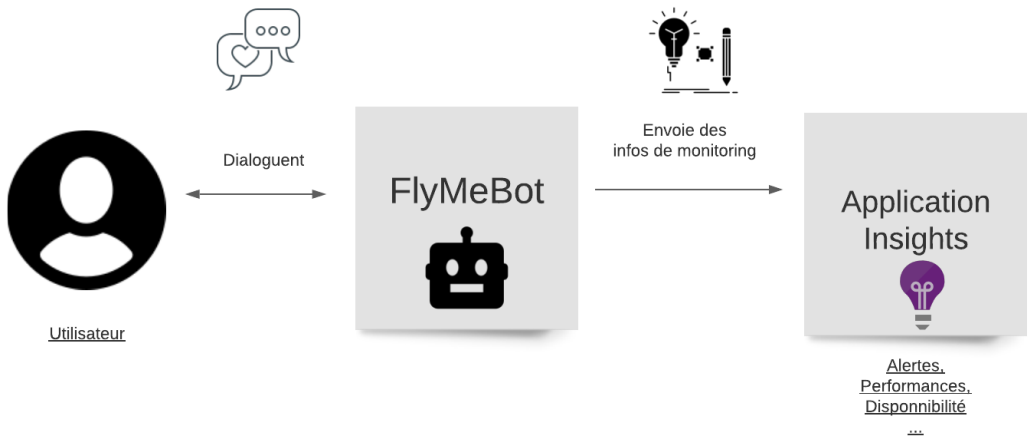


D'autres indicateurs peuvent être utilisés pour suivre les performances de l'application, comme le nombre de sessions, le nombre d'utilisateurs, la disponibilité, le temps de réponse moyen aux requêtes...

Un test de disponibilité a été mis en place (réponse de l'application à un ping):



Voici un schémas récapitulatif de l'évaluation du modèle en production :



3 - Modalités de mise à jour du modèle

Si le nombre d'erreurs enregistrées est trop important (Chatbot Alert), il peut être judicieux :

1. Sauvegarder les échanges entre l'utilisateur et le chatbot, par exemple dans une base de donnée (par exemple Azure SQL database)
2. Créer une nouvelle ressource LUIS
3. Entraîner le modèle de la manière suivante :
 - a. collecter dans la base de données les échanges pour lesquels les intentions n'ont pas été correctement identifiées par LUIS
 - b. labelliser ces échanges (intention et entités)
 - c. Entraîner LUIS avec ces échanges, comme indiqué dans le notebook
4. Mettre à jour LUIS_APP_ID et LUIS_API_KEY dans la ressource App Service, dans l'onglet configuration :

Nom	Valeur	Source	Paramètre de l'emplacement ...	Supprimer	Modifier
APPINSIGHTS_INSTRUMENTATION_KEY	Valeur masquée. Cliquez pour l'afficher	App Service			
LUIS_API_HOST_NAME	Valeur masquée. Cliquez pour l'afficher	App Service			
LUIS_API_KEY	Valeur masquée. Cliquez pour l'afficher	App Service			
LUIS_APP_ID	Valeur masquée. Cliquez pour l'afficher	App Service			
MicrosoftAppId	Valeur masquée. Cliquez pour l'afficher	App Service			
MicrosoftAppPassword	Valeur masquée. Cliquez pour l'afficher	App Service			
SCM_DO_BUILD_DURING_DEPLOYMENT	Valeur masquée. Cliquez pour l'afficher	App Service			

Nom	Valeur	Source	Type	Paramétr...	Supprimer	Modifier
(aucune chaîne de connexion à afficher)						



Veillez à bien conserver l'ancienne ressource LUIS:

Pour faire un rollback en cas de problème (si le nouveau modèle LUIS s'avère moins performant), simplement remettre les anciennes clés LUIS_APP_ID et LUIS_API_KEY),

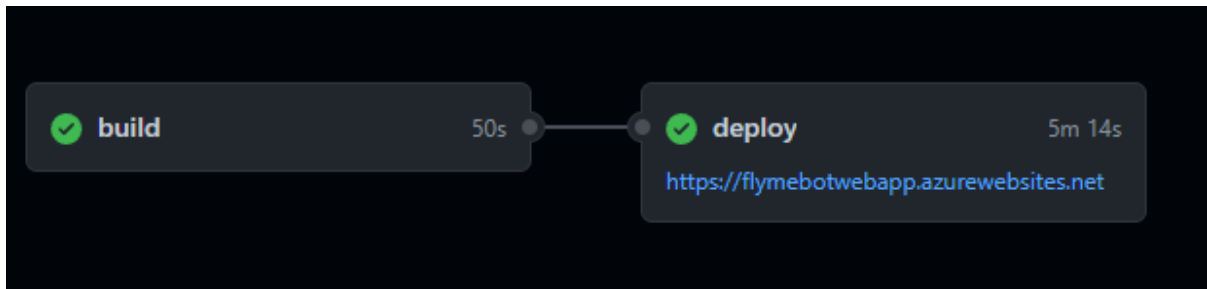
Concernant la fréquence de mise à jour, elle peut être continue dès qu'une alerte a été déclenchée (l'alerte Chatbot Alert ou d'autres alertes que vous aurez configurées selon les performances souhaitées). Ainsi, le modèle pourra être amélioré de façon continue.

4 - Modalités de mise à jour de l'application

La mise à jour de l'application se fait via Github. L'adresse du repository est

https://github.com/AnodeGrindYo/FlyMeBot_App.git

Lorsque vous pousserez un commit, les tests unitaires seront lancés. Ainsi, le nouveau code ne sera déployé en production que si les tests unitaires sont tous validés



This screenshot shows the GitHub Actions interface for the repository 'AnodeGrindYo / FlyMeBot_App'. The 'All workflows' tab is selected, displaying a list of workflow runs. The table includes columns for the workflow name, status, branch, and actor. The most recent run is titled 'date hints modified to conform to the strange and confusing way ameri...' and was pushed by 'AnodeGrindYo'.

Workflow Name	Status	Branch	Actor
date hints modified to conform to the strange and confusing way ameri...	Success	main	AnodeGrindYo
The bot was boring, so I've made it more funny and unpredictable. BTW...	Success	main	AnodeGrindYo
Build and deploy Python app to Azure Web App - FlyMeBotWebApp #30	Success	main	AnodeGrindYo
Confirmation prompt customized due to an inversion of boolean values ...	Success	main	AnodeGrindYo
Some dialog improvements	Success	main	AnodeGrindYo
debug logs added	Success	main	AnodeGrindYo

This screenshot shows the details of a specific workflow run for the 'date hints modified to conform to the strange and confusing way ameri...' workflow. The 'build' job is highlighted, showing its steps and duration. The 'test application' step is expanded, displaying the command output.

```
1 Run pytest
11 ===== test session starts =====
12 platform linux -- Python 3.8.14, pytest-7.1.3, pluggy-1.0.0
13 rootdir: /home/runner/work/FlyMeBot_App/FlyMeBot_App
14 collected 4 items
15
16 tests/test_bot.py .... [100%]
17
18 ===== 4 passed in 2.45s =====
19
20 Upload artifact for deployment jobs
21 Post Run actions/checkout@v2
22 Complete job
```

5 - Modalités de mise à jour de la page web

La mise à jour de la page web se fait également via Github. L'adresse du repository est : https://github.com/AnodeGrindYo/FlyMe_WebPage.git