

# Méthodologie du pilotage de la performance du modèle en production

## I) Métrique

Comme on ne dispose pas de métrique calculatoire pour évaluer le modèle, l'idée est de s'appuyer sur les retours des utilisateurs. On peut remonter à Application Insights les retours positifs et négatifs (via les méthodes TrackException et TrackEvent du telemetry client), pour calculer la fréquence de retours négatifs au cours d'une échéance donnée (semaine passée, par exemple).

Dans le cas d'un retour négatif, la conversation est archivée pour permettre une inspection future.

## II) Modalités de mise à jour du modèle

Après test du modèle, il s'avère que les prédictions incorrectes sur des messages génériques occupent moins de 5% du temps. Par conséquent, on peut placer un seuil de 10% de taux de retour négatif, au-delà duquel une alerte sera levée. Ce seuil permettra de constater efficacement d'éventuelles chutes de performances du modèle en production.

La phase de mise à jour consistera à étudier les conversations insatisfaisantes pour améliorer le modèle.

