PROJECT 5

Change model path in notebook 3

NOTES FOR PREZ

The practical distinction between stemming and lemmatization is that, where **stemming merely removes common suffixes from the end of word tokens, lemmatization ensures the output word is an existing normalized form of the word** (for example, lemma) that can be found in the dictionary.

Vu les résultats la technique LDA n’est pas très adapté à notre problématique.

****- Bag of words très bons résultats car mots techniques deviennent bien discriminants. Moins besoin du contexte ici comme avec USE ou Bert.****

* Présentation de la problématique et du prétraitement effectué, requête SQL et requête API de récupération des questions, exploration, cleaning et feature engineering (5 min).
* Présentation de l’approche non supervisée (4 min).
* Présentation de l’approche supervisée, de la visualisation avec MLFlow UI de la synthèse des résultats, et du modèle final sélectionné (5 min).
* Présentation des principes du MLOps et du pipeline de mise en production réalisé (3 min).
* Démonstration de l’API et conclusion (3 min).