



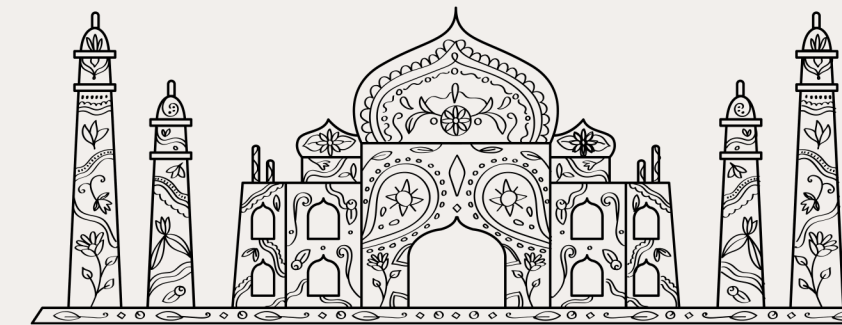
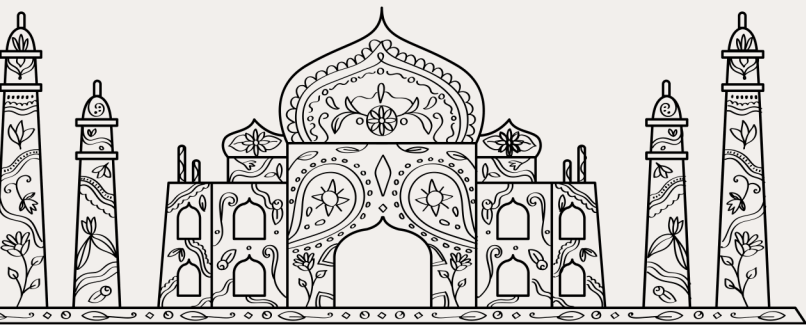
Mini-Projet

TALN

Ghada Bouchaddekh

Wafa Boughanmi

Ines Bouhelal



Etape 1

Collecte de Données

Manuelle

Extraction de données en ligne (data scrapping)

Etape 2

Annotation

Manuelle

Echelle de 1-5

- 1 étant une phrase incohérente
- &
- 5 étant une phrase cohérente

ص

ف

خ

Etape 3

Prétraitement

Suppression

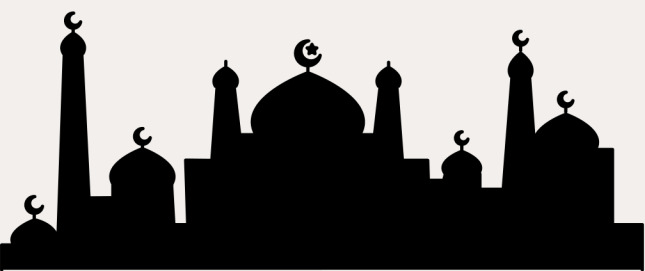
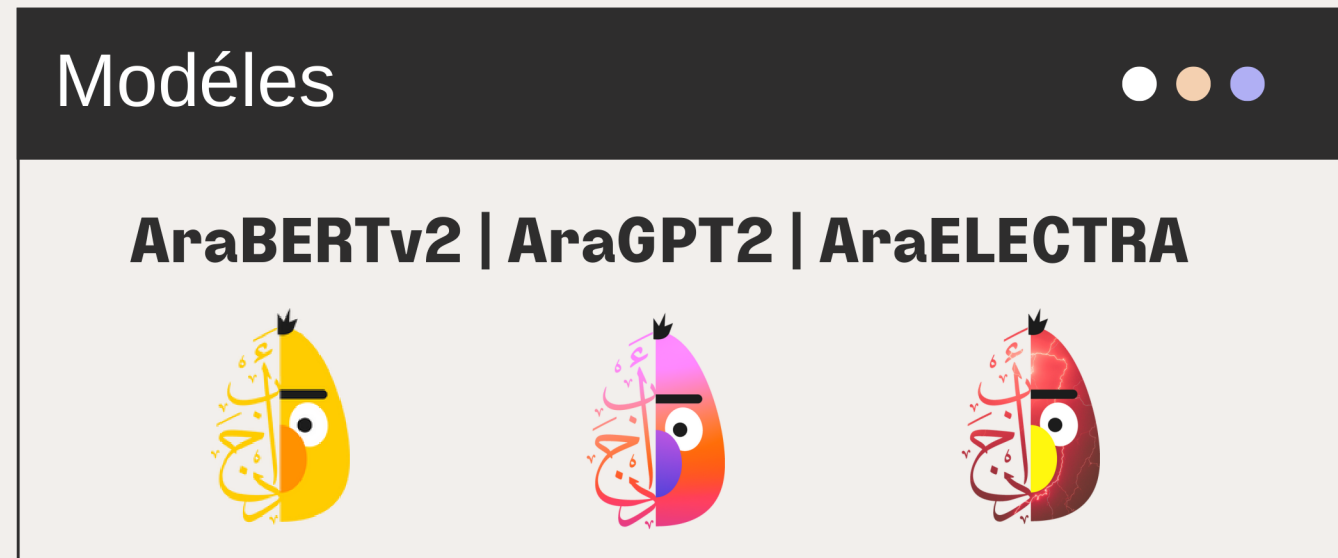
- Des caractères non-alphabétiques
- Les chiffres
- Les espaces superflus

ز

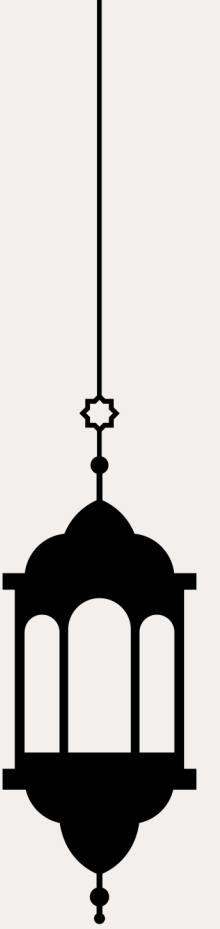
لا

أ

Etape 4



س
ش



Etape 5

