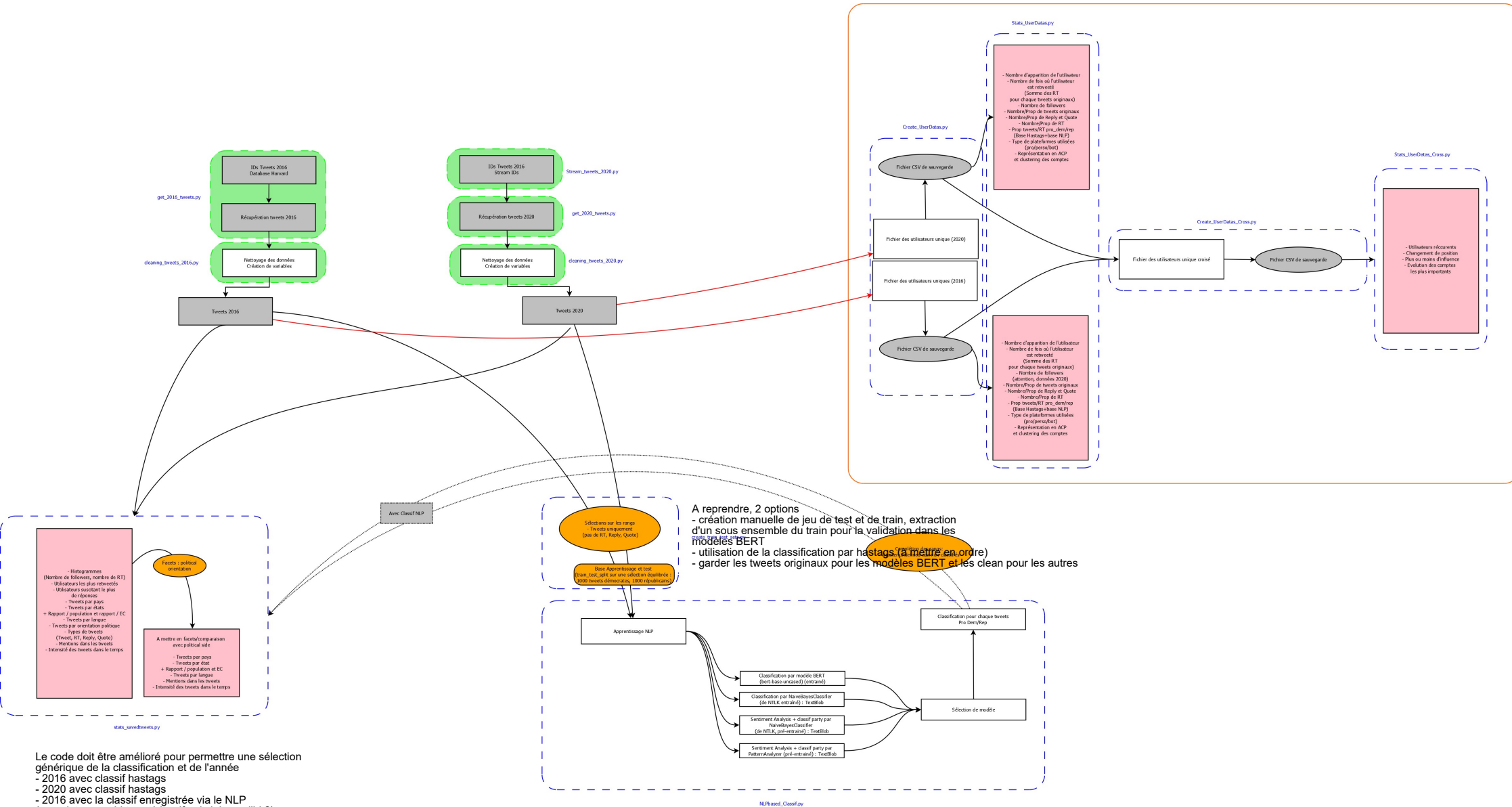


Etat du travail sur les individus ?



Le code doit être amélioré pour permettre une sélection générique de la classification et de l'année

- 2016 avec classif hastags
- 2020 avec classif hastags
- 2016 avec la classif enregistrée via le NLP (enregistrement id tweet/classif puis join sur l'id ?)
- 2020 avec la classif enregistrée via le NLP (de même)

Il faut ajouter les données de l'élection de 2020 (vote par état etc...) et les utiliser de manière automatique quand on est sur les tweets de 2020

Vérifier les sorties, améliorer les intitulés et les commentaires

Remettre en ordre avec :
1ere phase :

- Utilisation cohérente des jeux de données train/test pour l'évaluation des performances, stockage des score mis en forme (création d'un tableau, commentaires, graphes, traiter simultanément les jeux train/test de 2016 et 2020 pour les comparaisons...)
- Détermination des hyperparamètres optimaux pour les modèles BERT
- Utilisation d'un modèle cased pour BERT base

2eme phase

- Sélection d'un meilleur modèle et application pour la classif de tous les tweets 2016 et 2020 puis export