**La structure générale de rapport de Projet (Text mining )**

1. **Page de garde**
2. **Résumé et Abstract**
3. **Table des matières**
4. **Introduction**
   * Contexte du problème
   * Objectifs
   * Problématique et hypothèses
   * Méthodologie générale
5. **État de l’art / Revue de la littérature**
   * Travaux précédents sur le domaine
   * Méthodes et outils existants

(au moins 3 travaux)

1. **Méthodologie**
   * Détails techniques : outils, langages, bibliothèques
   * Architecture de la solution
   * Description des étapes de traitement --> il est souhaitable d'avoir un flowchart ou bien un graphique représentatif
2. **Partie centrale : Étude et expérimentation**
   * **Jeu de données utilisé**
   * **Prétraitement des données textuelles**
     + Nettoyage (suppression de ponctuation, stopwords, etc.)
     + Normalisation (lemmatisation, stemming)
     + Encodage (TF-IDF, Word2Vec, etc.)
   * **Exploration et visualisation des textes**
     + Nuages de mots, histogrammes, analyse de fréquence
   * **Application d’algorithmes de *text mining***
     + Classification (Naive Bayes, SVM)
     + Clustering (K-means, LDA pour topic modeling)
     + Analyse de sentiments, reconnaissance d’entités nommées (NER)
   * **Évaluation**
     + Métriques utilisées (précision, rappel, F-mesure)
     + Comparaison avec d’autres approches ou benchmarks
3. **Résultats et Interprétation**
   * Présentation des résultats (graphes, tableaux)
   * Analyse critique
   * Interprétations
   * Limites rencontrées
4. **Conclusion et perspectives**
   * Résumé des contributions
   * Voies d'amélioration / extensions possibles
5. **Bibliographie**
6. **Annexes**
   * Code source (éventuellement en GitHub)
   * Captures d’écran des outils développés

**NB : La similarité de rapport doit être inférieur à 20 %. L'indice d'intelligence artificielle doit être inférieur à 20 %.**