

Détection automatique de fausses critiques de produits

Dataset : kaggle Fake-Reviews Dataset

Colonnes importantes :

- category : catégorie de produit (optionnelle, à utiliser si nécessaire pour enrichir le modèle)
- rating : note du produit (nombre entier entre 1 et 5)
- label : classe de la critique (frauduleuse ou authentique) avec les valeurs possibles : fake ou real
- text : contenu textuel de la critique

Travaux à réaliser

1. Prétraitement des données

- Chargement du dataset

Nettoyage du texte : Appliquez les étapes suivantes sur la colonne text :

- Suppression de la ponctuation
- Conversion en minuscules
- Suppression des stopwords
- Lemmatisation (ou stemming)
- Encodage des labels (fake = 0, real = 1) dans la colonne label.

Analyse statistique initiale : Longueur des critiques , Distribution des classes label, Détection et suppression des doublons, si nécessaire.

2. Analyse exploratoire (EDA)

- Distribution des classes
- Longueur des critiques par classe (histogramme)
- Visualisation des mots les plus fréquents (WordCloud par classe)
- Présenter au moins une autre visualisation pertinente (ex : heatmap, bigrammes fréquents)

3. Vectorisation du texte

- Appliquer une vectorisation TF-IDF sur la colonne du Text.
- Comparer les dimensions, vocabulaire et effets sur la mémoire

4. Modélisation

Implémenter et comparer les modèles suivants :

- Régression logistique (baseline)
- Réseau de neurones simple (MLP)

5. Évaluation des performances

- Affichage des scores : accuracy, précision, rappel, F1-score, AUC
- Courbe ROC pour au moins deux modèles
- Matrice de confusion et analyse qualitative des erreurs

6. Interprétation des résultats

- Appliquer LIME ou SHAP pour expliquer le comportement du modèle le plus performant
- Présenter au moins deux critiques avec leur prédiction et les mots clés influents

Questions théoriques à traiter dans le notebook

1. Pourquoi le nettoyage du texte est-il crucial avant vectorisation ?
2. Que mesure le F1-score et pourquoi est-il souvent plus pertinent que la simple accuracy ?
3. Quelle est la différence entre TF-IDF et les embeddings comme Word2Vec ?
4. Comment le déséquilibre des classes peut-il impacter les performances d'un modèle ?
5. Quels sont les enjeux éthiques liés à la détection automatisée de contenus frauduleux ?

Livrables attendus

- Un notebook Jupyter .ipynb propre et structuré contenant :
 - Tout le code exécuté
 - Les visualisations pertinentes
 - Les réponses aux questions théoriques
 - Un compte rendu détaillé