

Matrice de confusion :

```
[[9 1]
 [3 7]]
```

[9 1] : Ce sont les prédictions pour les tweets négatifs (réellement négatifs) :

- **9** : Nombre de tweets négatifs correctement classés comme négatifs (True Negatives, TN).
- **1** : Nombre de tweets négatifs incorrectement classés comme positifs (False Positives, FP).

[3 7] : Ce sont les prédictions pour les tweets positifs (réellement positifs) :

- **7** : Nombre de tweets positifs correctement classés comme positifs (True Positives, TP).
- **3** : Nombre de tweets positifs incorrectement classés comme négatifs (False Negatives, FN).

Rapport de classification :

	precision	recall	f1-score	support
Négatif	0.75	0.90	0.82	10
Positif	0.88	0.70	0.78	10
accuracy			0.80	20
macro avg	0.81	0.80	0.80	20
weighted avg	0.81	0.80	0.80	20

1. Précision (Precision) :

La **précision** est la proportion de tweets positifs correctement prédits parmi tous les tweets prédits comme positifs. Elle mesure la qualité des prédictions positives.

- **Précision pour la classe Négatif : 0.75**
 - Cela signifie que sur tous les tweets prédits comme négatifs, 75% étaient réellement négatifs.
- **Précision pour la classe Positif : 0.88**
 - Cela signifie que sur tous les tweets prédits comme positifs, 88% étaient réellement positifs.

2. Rappel (Recall) :

Le **rappel** est la proportion de tweets positifs correctement identifiés parmi tous les tweets réellement positifs. Il mesure la capacité du modèle à détecter toutes les instances positives.

- **Rappel pour la classe Négatif : 0.90**
 - Cela signifie que sur tous les tweets réellement négatifs, 90% ont été correctement identifiés comme négatifs.
- **Rappel pour la classe Positif : 0.70**
 - Cela signifie que sur tous les tweets réellement positifs, 70% ont été correctement identifiés comme positifs.

3. F1-Score :

Le **F1-Score** est la moyenne harmonique de la précision et du rappel. Il est utilisé pour équilibrer les métriques lorsque les classes sont déséquilibrées.

- **F1-Score pour la classe Négatif : 0.82**
 - Cela montre un bon compromis entre la précision et le rappel pour la classe négative.
- **F1-Score pour la classe Positif : 0.78**
 - Cela montre un compromis raisonnable entre la précision et le rappel pour la classe positive.

Analyse des erreurs fréquentes et biais éventuels

- **Erreur fréquente pour les tweets négatifs** : Il y a une erreur relativement faible de 1 tweet négatif qui est classé comme positif. Cela peut être dû à des caractéristiques de langage dans le tweet qui ressemblent à un tweet positif (par exemple, sarcasme ou utilisation ambiguë de mots).
- **Erreur fréquente pour les tweets positifs** : Il y a 3 erreurs où des tweets positifs ont été classés comme négatifs. Cela peut être dû à des nuances subtiles dans le texte ou un vocabulaire trop proche des mots négatifs.

Biais potentiel :

- Le modèle semble mieux classer les **tweets négatifs** que les **tweets positifs** (90% de rappel contre 70% pour les positifs). Cela pourrait être dû à un déséquilibre dans le nombre de tweets négatifs et positifs dans l'ensemble de données d'entraînement.
- Le modèle semble avoir une **précision plus élevée pour la classe positive** (88%) que pour la classe négative (75%). Cela pourrait être dû à une meilleure capacité à détecter les caractéristiques des tweets positifs dans les données d'entraînement.

Recommandations pour améliorer le modèle

1. Augmenter la quantité de données :

Plus de données d'entraînement, en particulier pour les tweets négatifs, pourraient améliorer la précision pour cette classe.

2. Amélioration du nettoyage des données :

Travailler sur le nettoyage des données pour prendre en compte des cas spécifiques comme les sarcasmes ou les contextes où un tweet négatif peut sembler positif.