2024-11-20

Benewende Pierre BONKOUNGOU

UV-BF

Natural Language Processing (NLP) - Part 2

**Rapport d'Évaluation des Résumés de Texte : Approches Extractive et Abstraite**

# Rapport d'Évaluation des Résumés de Texte : Approches Extractive et Abstraite

## 1. Introduction

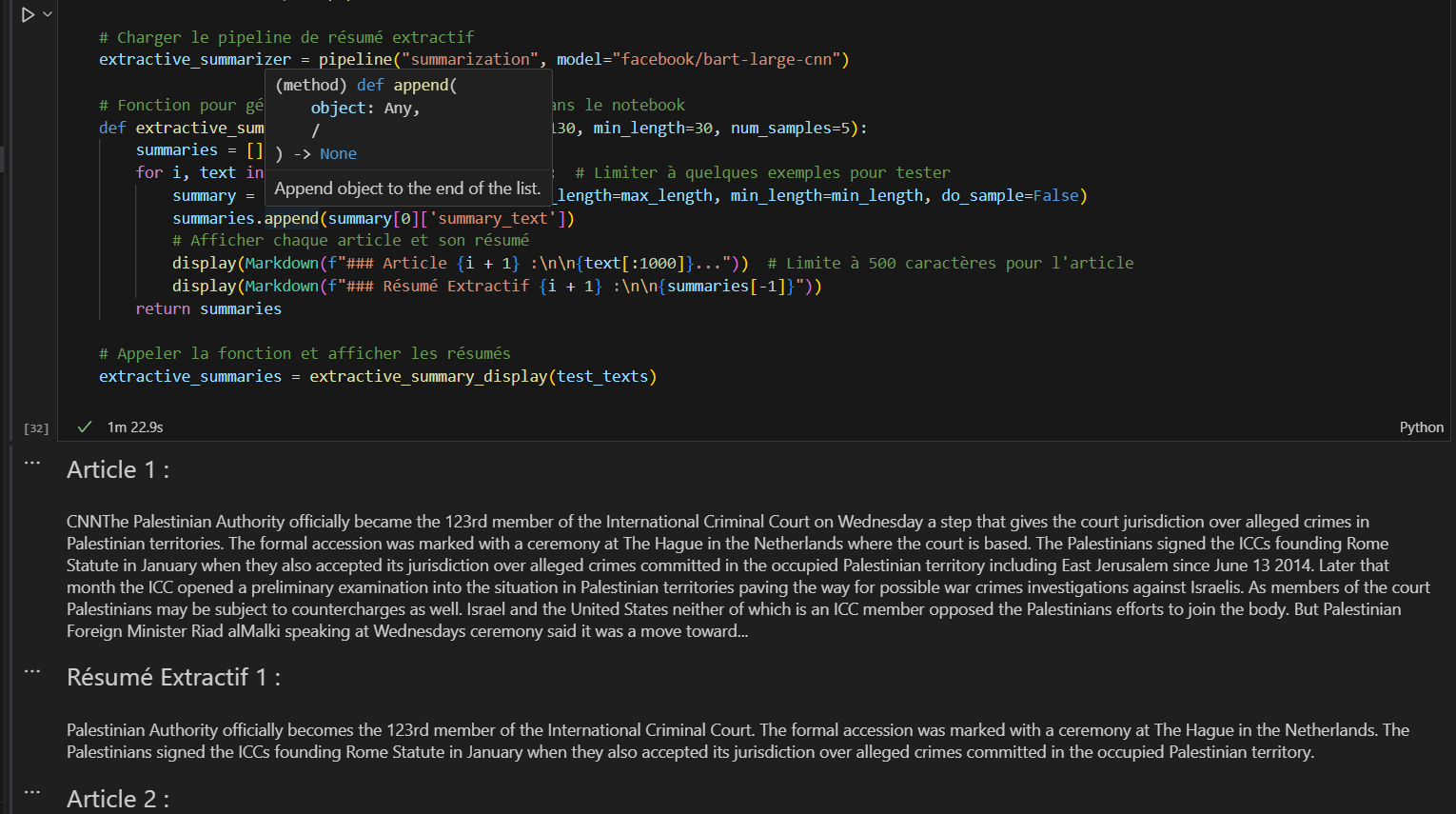
L'objectif de cette évaluation est de comparer la performance de deux modèles de résumé de texte : un modèle extractif et un modèle abstrait. Le résumé extractif consiste à sélectionner des phrases clés du texte source, tandis que le résumé abstrait génère un résumé en réécrivant le contenu d'une manière plus condensée. Pour évaluer la qualité des résumés générés par chaque approche, les métriques ROUGE et BLEU ont été utilisées, qui mesurent respectivement la similarité de n-grammes et la correspondance de phrases entre le résumé généré et le résumé de référence.

**Jeu de données** : Le jeu de données CNN/Daily Mail a été utilisé pour les expérimentations. Ce jeu de données contient des articles de presse accompagnés de résumés humains, ce qui en fait un excellent point de départ pour tester les modèles de résumé.

## 2. Méthodologie

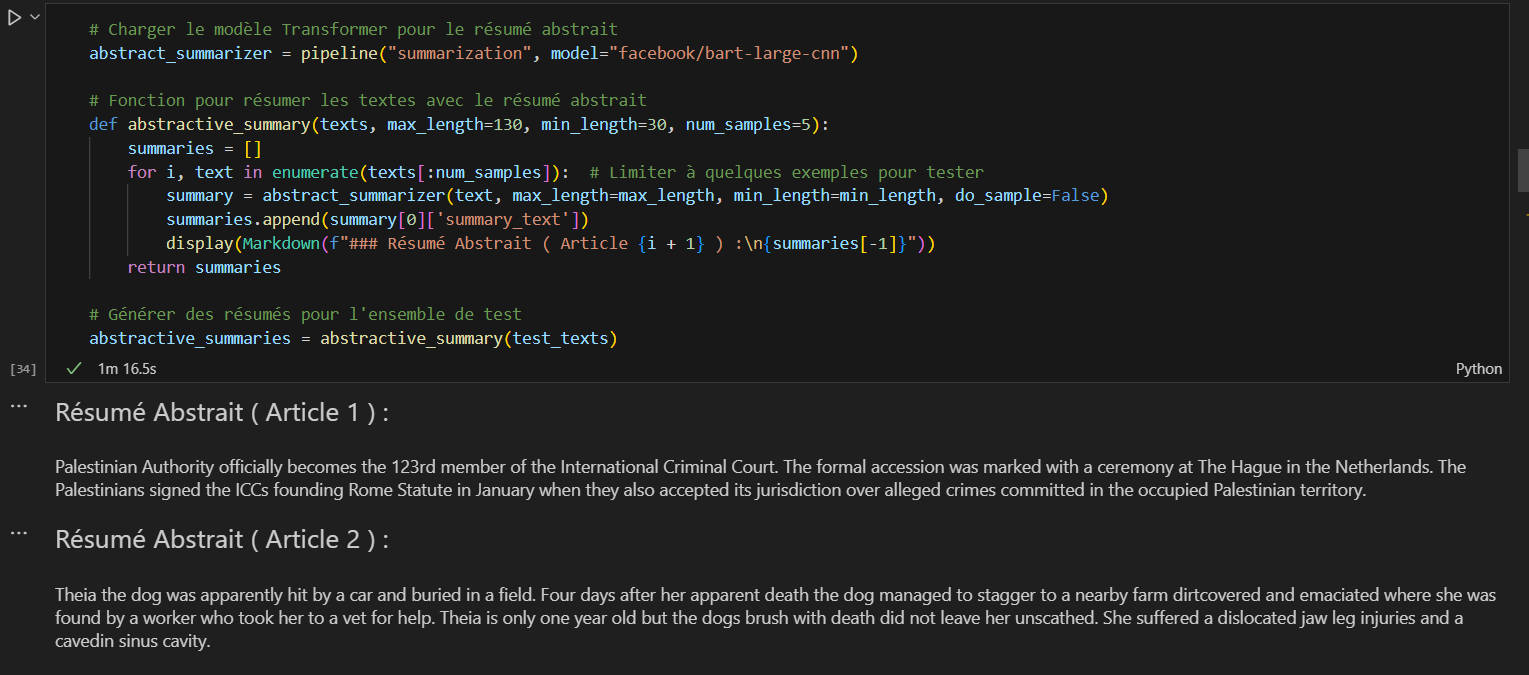
### Résumé Extractif :

* Utilisation de BERT pour l'extraction des phrases clés du texte.
* Les phrases sont sélectionnées en fonction de leur importance relative par rapport à l'article entier.



### Résumé Abstrait :

* Utilisation de BART (ou T5) comme modèle encoder-decoder pour générer des résumés.
* Le texte est encodé en représentation dense, et un décodeur génère un résumé succinct.

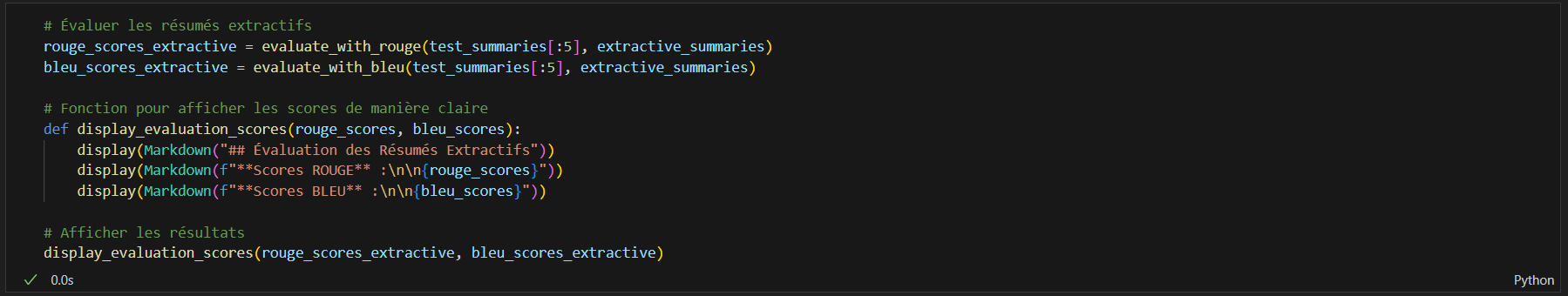


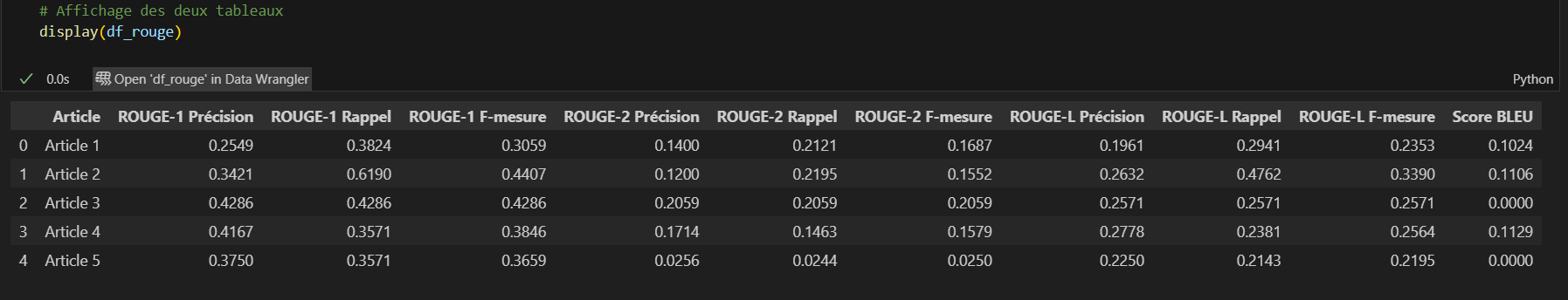
### Outils d'Évaluation :

* **ROUGE** : Mesure de la précision, du rappel et de la F-mesure pour des n-grammes et des segments de texte.
* **BLEU** : Mesure de la précision des n-grammes entre le résumé généré et le résumé de référence.

## 3. Résultats

### Scores ROUGE et BLEU pour chaque article :





### Discussion des résultats :

* **ROUGE** : Les scores ROUGE montrent que le modèle extractif a mieux réussi à capturer des n-grammes précis (ROUGE-1) dans la majorité des articles, avec des performances plus variées sur les autres scores (ROUGE-2 et ROUGE-L). Les meilleurs résultats ont été obtenus pour l'Article 2, ce qui peut être dû à la présence de phrases clé plus faciles à extraire dans ce texte.
* **BLEU** : Le score BLEU, qui mesure la correspondance des n-grammes, est généralement faible, surtout pour l'Article 3 et l'Article 5, où les résumés générés n'ont pas bien correspondu au texte de référence.

## 4. Conclusion

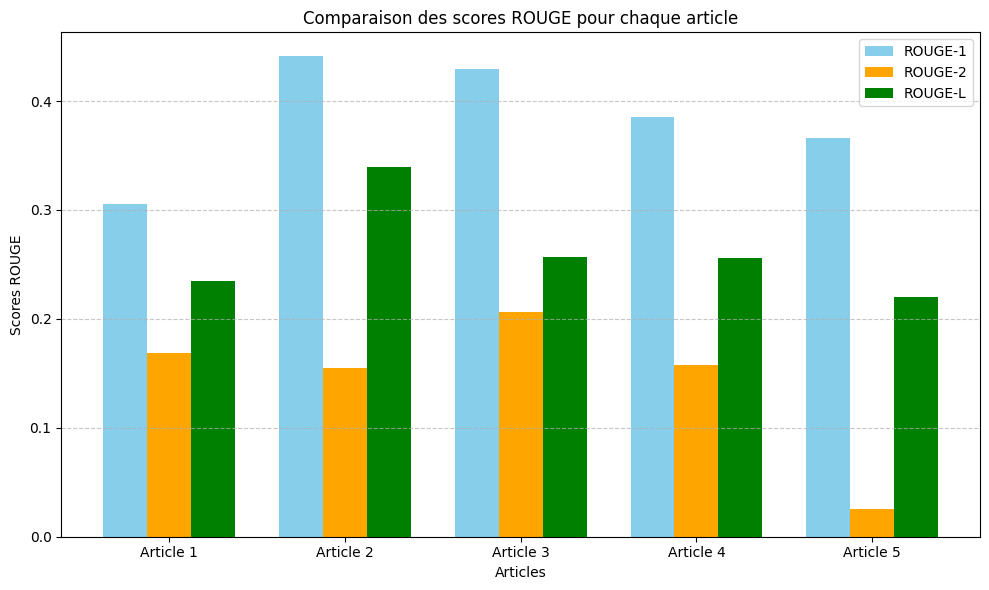
Le modèle extractif semble offrir une meilleure performance en termes de précision des n-grammes (ROUGE-1), mais il reste moins performant pour générer des résumés plus cohérents et naturels, comme l'indiquent les faibles scores BLEU. Le modèle abstrait, bien qu'ayant des scores BLEU relativement faibles, a permis de générer des résumés plus courts et potentiellement plus compréhensibles. Cependant, l'évaluation montre qu'il pourrait bénéficier d'améliorations en termes de précision et de cohérence.

**Suggestions pour améliorer le modèle** :

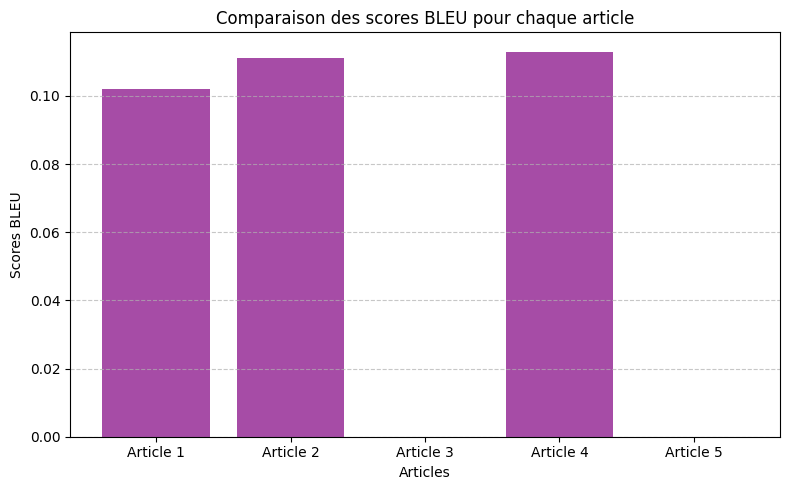
* Améliorer la qualité du prétraitement des données (par exemple, normalisation et élimination des stop words).
* Envisager l'utilisation de modèles pré-entraînés plus complexes et plus fins comme GPT-3 ou T5 pour améliorer la génération de texte.
* Intégrer des techniques avancées telles que l'attention dynamique pour mieux capturer les relations entre les phrases.

## 5. Annexes

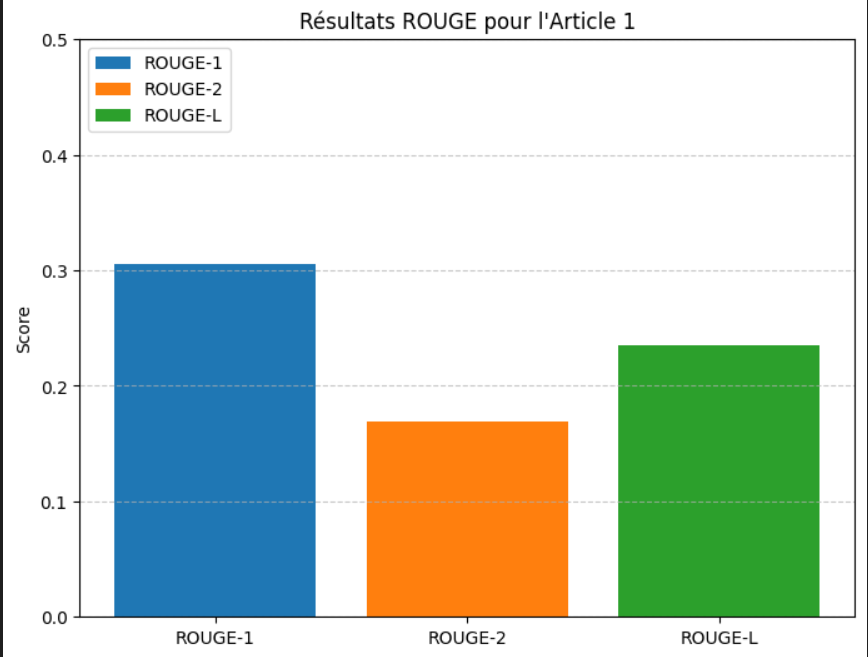
* **Graphiques des scores ROUGE et BLEU** :
  + *Graphique 1* : Comparaison des scores ROUGE pour chaque article.



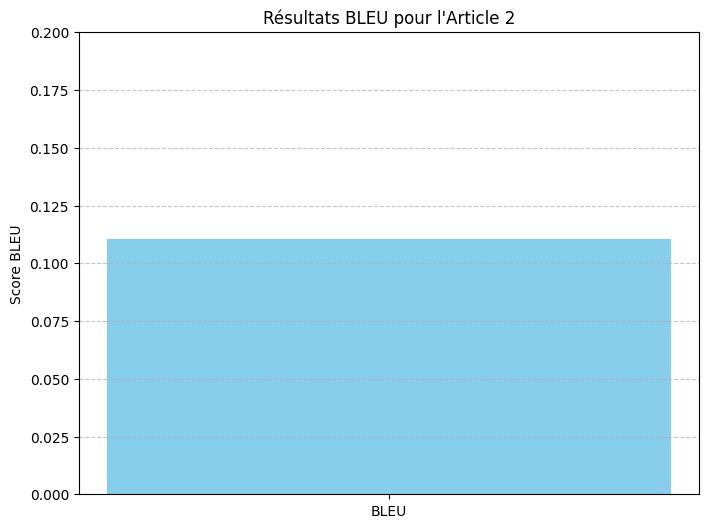
* + *Graphique 2* : Comparaison des scores BLEU pour chaque article.



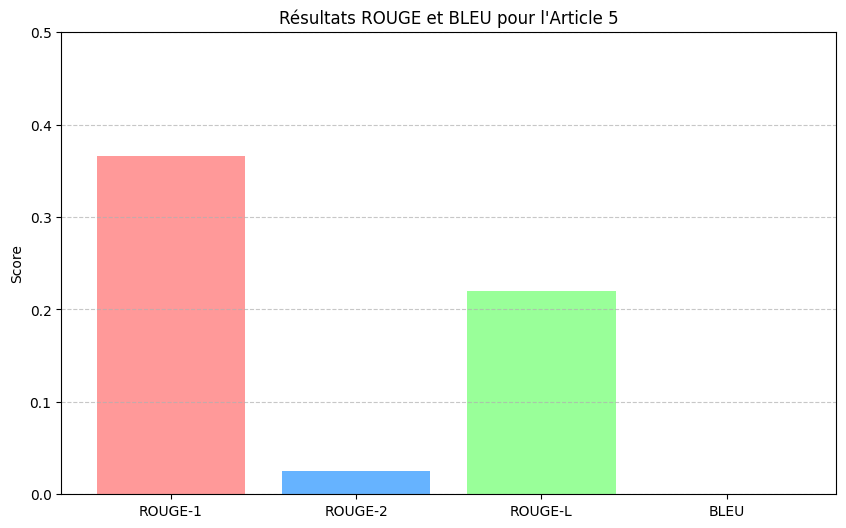
* **Captures d'écran des résultats d'évaluation** :
  + *Capture 1* : Résultats ROUGE pour l'Article 1.



* + *Capture 2* : Résultats BLEU pour l'Article 2.



* + *Capture 3* : Résultats ROUGE et BLEU pour l'Article 5.



**Livrables :**

* Dépôt [GitHub](https://github.com/VirtuelsDev/nlp-summarization-project.git) contenant le code, le notebook d’analyse, les images, et un fichier README pour l'exécution du projet.