ICB@Défi TextMine'25 : Extraction de relations pour l'analyse des rapports de renseignement

Hussam Ghanem, Daren Hacbekri, Christophe Cruz

▶ **To cite this version:**

Hussam Ghanem, Daren Hacbekri, Christophe Cruz. ICB@Défi TextMine'25 : Extraction de relations pour l'analyse des rapports de renseignement. TextMine'25 (EGC 2025), Jan 2025, Strasbourg, France. hal-04928614

HAL Id: hal-04928614
https://hal.science/hal-04928614v1

Submitted on 4 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

ICB@Défi TextMine'25 : Extraction de relations pour l'analyse des rapports de renseignement

Hussam Ghanem*, Daren Hacbekri*
Christophe Cruz*

*ICB, UMR 6306, CNRS, Université de Bourgogne, 21000 Dijon, France
https://icb.u-bourgogne.fr/

1 Introduction

Le défi TextMine'25 (Prieur et al., 2025), proposé par Airbus Defence and Space, se concentre sur l'extraction automatique des relations entre entités dans des rapports de renseignement. Cette tâche, cruciale pour l'analyse dans les domaines de la défense et du renseignement, reste un défi scientifique nécessitant souvent une intervention humaine. Le défi vise à comparer des solutions technologiques publiques et privées à l'aide de données textuelles fictives annotées, issues du projet POPCORN (Giordano, 2024). Les participants doivent identifier et annoter des relations parmi 37 catégories, telles que LOCATED_IN ou PART_OF. Ces données, simulant des scénarios réels, permettent de tester les capacités des modèles à comprendre les interactions complexes entre acteurs, événements et attributs, avec des implications directes pour l'analyse de rapports de renseignement.

Le jeu de données TextMine'25 comprend 800 documents factices annotés pour l'entrainement et 400 documents pour l'évaluation, organisés en fichiers CSV : **train.csv** (textes avec entités et relations), **test.csv** (textes avec entités uniquement) et **sample_submission.csv** (exemple de soumission). Les entités sont représentées par des dictionnaires (id, mentions, type), tandis que les relations dans l'ensemble d'entraînement sont des triplets (id source, type de relation, id cible). Le corpus inclut 38 types de relations, avec plusieurs types d'identités et d'attributs. L'évaluation utilise le score Macro F1, calculé comme la moyenne des scores F1 pour chaque type de relation. L'ensemble des informations sont disponibles sur le site Kaggle [1]

2 Prompt Engineering

Nous avons abordé le défi TextMine'25 [2] en exploitant le **Prompt Engineering** pour extraire des relations à partir de textes en utilisant le grand modèle de langage (LLM) **Gemini-1.5-pro-002** [3]. En définissant des instructions explicites et en normalisant les formats de sor-

1. Défi TextMine 2025 : https://www.kaggle.com/competitions/defi-text-mine-2025/
2. Notre code sur Github : https://github.com/ChristopheCruz/2024_Kaggle_competition
3. Modèles Gemini : https://ai.google.dev/gemini-api/docs/models/gemini?hl=fr#gemini-1.5-flash

tie, nous avons structuré les relations sous forme de triplets : `[Subject_id, Relation Type, Object_id]`.

Notre approche combine le **zero-shot prompting (ZSP)** (Caufield et al., 2023) et le **few-shot prompting (FSP)** (Han et al., 2023). Le ZSP correspond à des prompts formulés sans fournir d'exemples d'entrée-sortie, tandis que le FSP implique l'utilisation de prompts enrichis avec des exemples illustratifs. Dans le prompt, le texte et les entités (ainsi que leurs attributs) sont fournis comme entrée, et le modèle est invité à extraire les relations correspondantes. Une hiérarchie stricte a été adoptée pour structurer les entités et attributs, renforçant la cohérence des relations extraites. Des règles ont également été imposées, telles que la priorisation des entités spécifiques (enfants) et un formatage rigoureux des résultats. Par exemple, la relation `IS_BORN_IN` était limitée aux types `Person` (sujet) et `Place` (objet).

Initialement, le modèle rencontrait des difficultés avec le format et la validité des relations. Par itérations successives, dans notre prompt, nous avons affiné les instructions et mis en place des directives pour éliminer les incohérences, éviter les explications superflues, et garantir une sortie conforme.

Défis et solutions. Au départ, Gemini ajoutait des explications superflues autour des triplets. L'interdiction explicite de tels textes garantissait des résultats plus propres. La restriction des sujets à la liste des entités autorisées a réduit les erreurs d'identification des sujets. Donner la priorité aux entités enfants par rapport aux types parents a amélioré la granularité des relations extraites. Des règles de formatage strictes ont permis de remédier aux incohérences en matière de représentation d'ID incorrecte et de structure de relation.

3 Résultats et conclusion

Les résultats de notre approche, présentés dans le tableau 1, montrent une amélioration significative des performances du modèle Gemini-1.5-pro-002 à mesure que des exemples supplémentaires sont incorporés dans les prompts. En zero-shot, le modèle atteint un Macro F1 de 0.307, ce qui reflète une capacité limitée à extraire correctement les relations sans exemples. L'ajout de 5 exemples structurés (5-shots) entraîne une augmentation notable du score, atteignant 0.457. Cette progression indique que le modèle bénéficie fortement de directives explicites et de contextes démonstratifs.

Cependant, l'ajout de 10 exemples (10-shots) n'entraîne qu'une amélioration marginale, avec un Macro F1 de 0.464. Cette faible différence entre les scénarios 5-shots et 10-shots suggère que l'utilisation de plus d'exemples ne garantit pas nécessairement une amélioration substantielle des performances. Cela pourrait s'expliquer par une saturation des capacités d'apprentissage du modèle dans ce cadre particulier, où un petit nombre d'exemples bien structurés suffit pour atteindre des performances optimales.

Ainsi, nos résultats indiquent qu'une approche concise, exploitant un nombre limité mais pertinent d'exemples, est à la fois efficace et efficiente pour cette tâche spécifique. Cette stratégie réduit également les coûts de traitement tout en maintenant des performances élevées.

	Zero-shot	5-shots	10-shots
Macro F1	0.307	0.457	0.464

TAB. 1 – *Résults with Gemini-1.5-pro-002.*

Références

Prieur, M., G. Gadek, H. Rawsthorne, A. Guille, P. Cuxac, et C. Lopez (2025). Défi textmine'25 -extraction de relations pour analyser des rapports de renseignement. actes de l'atelier textmine'25, p. à paraître, extraction et gestion des connaissances 2025 (egc'25).

Giordano, A., e. a. (2024). Fictional and synthetic intelligence reports for named entity recognition and relation extraction tasks. in proceedings of kes'24, to appear.

Caufield, J. H., H. Hegde, V. Emonet, N. L. Harris, M. P. Joachimiak, N. Matentzoglu, H. Kim, S. A. Moxon, J. T. Reese, M. A. Haendel, et al. (2023). Structured prompt interrogation and recursive extraction of semantics (spires) : A method for populating knowledge bases using zero-shot learning. *arXiv preprint arXiv :2304.02711*.

Han, J., N. Collier, W. Buntine, et E. Shareghi (2023). Pive : Prompting with iterative verification improving graph-based generative capability of llms. *arXiv preprint arXiv :2305.12392*.

Summary

The TextMine'25 challenge, by Airbus Defence and Space, focuses on extracting relationships in intelligence reports. We apply Large language models (Gemini) to identify relations between entities, reducing manual efforts and advancing relation extraction methods on annotated data.