

Comparison of Relation Extraction Models for Generating Knowledge Graphs in the Oil Drilling Domain

Gabriel H. G. Ticianeli¹, Arnaldo Candido Junior¹, Ivan Rizzo Guilherme¹,
Bruno Elias Penteado, Stephan Ribeiro Perrout³, Luis Henrique Morelli¹,
Pedro Henrique Paiola¹, Gabriel Lino Garcia¹

¹Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP)

²Universidade de São Paulo (USP)

³Petrobras

gh.garcia@unesp.br, arnaldo.candido@unesp.br, ivan.guilherme@unesp.br,
bpenteado@gmail.com, sperrout@petrobras.com.br, luis.morelli@unesp.br,
pedro.paiola@unesp.br, gabriel.lino@unesp.br

Abstract. *Knowledge Graphs are structures that provide explicit knowledge, symbolic reasoning, and interpretable results, and they can also be enhanced over time. For this reason, the automatic extraction of relations between entities from unstructured texts is one of the current research areas in Natural Language Processing. In this paper, we selected three state-of-the-art relation extraction models trained on general domains and compared their performance on a dataset from the oil well drilling area. The results demonstrated the low effectiveness of general models when applied to the technical language of this domain.*

Resumo. *Grafos de Conhecimento são estruturas que oferecem conhecimento explícito, raciocínio simbólico e resultados interpretáveis, além de poderem ser aprimoradas com o tempo. Por isso, a extração automática de relações entre entidades a partir de textos não estruturados é uma das atuais áreas de pesquisa do Processamento de Linguagem Natural. Neste artigo, selecionamos três modelos de extração de relações do estado da arte treinados em domínio geral e comparamos seus resultados obtidos em um dataset da área de perfuração de poços de petróleo. Os resultados demonstraram a baixa eficácia dos modelos gerais quando aplicados à linguagem técnica deste domínio.*

1. Introdução

Projetar algoritmos capazes de interagir com a linguagem humana é um desafio cobiçado pelos cientistas da computação, e modelos que o fazem estão em constante evolução. O Processamento de Linguagem Natural (PLN) é um campo de estudo que desde os anos 50 une Inteligência Artificial e Linguística para realizar tarefas como tradução de textos e verificação sintática [Nadkarni et al. 2011]. Avançando para os tempos atuais, dá-se destaque aos Grandes Modelos de Linguagem (GML), alimentados com quantidades massivas de parâmetros. Eles podem realizar tarefas ainda mais complexas devido a novas formas de treinamento como Aprendizagem em Contexto e Aprendizagem por Reforço a Partir de *Feedback* Humano.

As aplicações de GMLs variam desde responder perguntas até a geração de códigos baseada em uma demanda, o que consegue impactar áreas como a educação, o desenvolvimento de *software* e a pesquisa científica [Abdullah et al. 2022]. No entanto, por mais que robustos, eles ainda sofrem com um grande problema chamado de alucinação, nomeado a partir do fenômeno na psicologia. Um texto alucinado é uma resposta gerada pelo sistema que parece basear-se na realidade, porém é factualmente incorreta [Ji et al. 2023].

Uma proposta para solucionar o problema da alucinação e da inexplicabilidade das respostas de modelos de linguagem é a incorporação de Grafos de Conhecimento (GC), que podem ser incorporados tanto no período de treinamento quanto no período de inferência para fornecer conhecimento externo [Pan et al. 2024].

Por meio das conexões entre entidades (nós) e relações (arestas), GCs oferecem conhecimento explícito e acurado, raciocínio simbólico, resultados interpretáveis e flexibilidade, podendo ser melhorados com o tempo [Hogan et al. 2021] [Zhong et al. 2023].

A construção automática de GCs é uma área de pesquisa ativa que busca a extração de entidades e relações a partir de textos, unindo Processamento de Linguagem, Mineração de Dados e Aprendizado de Máquina [Melnyk et al. 2022]. A Extração de Relações (ER) de um texto é uma tarefa tradicionalmente dividida em duas etapas: reconhecer as entidades, fase conhecida como Reconhecimento de Linguagens Nomeadas (REN); e a Classificação de Relação (CR), que verifica se existe alguma ligação entre as entidades extraídas [Cabot and Navigli 2021].

Modelos atuais são capazes de fazer todo o processo de ER, texto como entrada e gerando triplas (Entidade, Relação, Entidade) como saída, processo chamado de extração *End-to-end* (ou *sequence to sequence*). Eles geram retornos com alta precisão em *datasets* do mundo real, alcançando *F-scores* de mais de 90, como exemplificado na Figura 1.

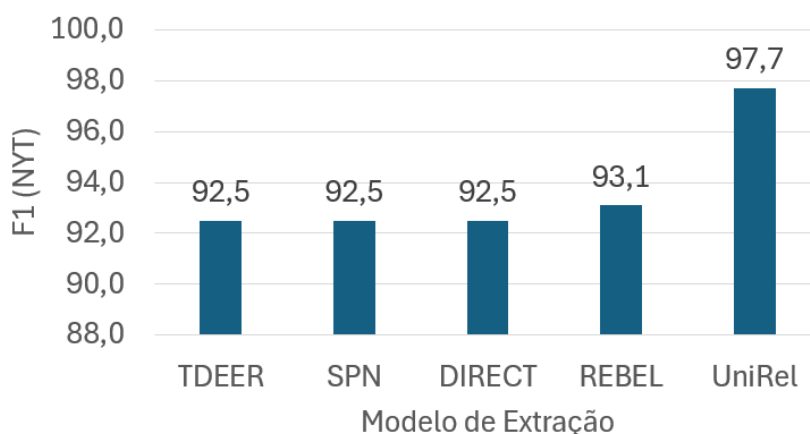


Figura 1. F-score obtido por modelos de ER no dataset NYT (2024)

O objetivo desta pesquisa foi verificar como esses algoritmos com alto desempenho em frases de domínio geral se comportam em textos sobre perfuração de poços de petróleo. Além disso, revelar a possível necessidade de abordagens personalizadas dependentes da área de atuação. Foram efetuadas tentativas de extração em um conjunto de dados privado utilizando os modelos Rebel [Cabot and Navigli 2021], UniRel

[Tang et al. 2022] e DIRECT [Zhao et al. 2021]. Os resultados obtidos foram avaliados a partir de uma classificação manual, onde foi considerada a quantidade de relações extraídas e a similaridade delas com um conjunto de triplas esperadas para cada sentença.

2. Trabalhos Relacionados

Uma revisão bibliográfica feita em [Ye et al. 2022] resume o progresso recente na Construção de Grafos de Conhecimento (CGC) usando ferramentas generativas, separando as abordagens em cinco paradigmas, sendo eles:

- **Sequência baseada em cópia:** Modelos que copiam as entidades correspondentes diretamente do texto de entrada. Usado em aplicações como [Zeng et al. 2018] e [Huang et al. 2021];
- **Sequência linearizada por estrutura:** Utilização de conhecimento estrutural e semântica de rótulo para lidar com um formato de saída unificado. Usado em aplicações como [Cabot and Navigli 2021] e [Lu et al. 2022];
- **Sequência baseada em rótulo:** Utilização de marcadores extras para indicar entidades ou relacionamentos específicos, utilizando colchetes ou outros identificadores para especificar a sequência de marcação da entidade de interesse. Usado em aplicações como [Athiwaratkun et al. 2020] e [De Cao et al. 2020];
- **Sequência baseada em índice:** Este paradigma gera diretamente os índices das palavras no texto de entrada de interesse e codifica rótulos de classe como índices de rótulos. Usado em aplicações como [Yan et al. 2021a] e [Yan et al. 2021b]; e
- **Sequência baseada em branco:** Utilização de modelos para definir a ordem e o relacionamento apropriados para os vãos em branco gerados no texto de entrada. Usado em aplicações como [Bosselut et al. 2019] e [Liu et al. 2022].

No que diz respeito à extração de conhecimento no domínio petrolífero, relatórios de perfuração são considerados uma fonte rica de informações com potencial significativo para análise de padrões e mineração de texto. [Ribeiro et al. 2020] e [Cinelli et al. 2021] trabalharam com a identificação e extração de eventos de boletins de perfuração utilizando uma abordagem contextual baseada em Redes Neurais Recorrentes e uma baseada em Redes Neurais Profundas, respectivamente. Esforços também estão sendo colocados na criação de bases de conhecimento e vocabulários específicos para o domínio de petróleo e gás [Freitas et al. 2015] [Gomes et al. 2021].

[Hoffmann et al. 2018] propõem outra metodologia para a classificação de boletins baseada no Processamento de Linguagem Natural Profundo. Os autores também identificam os principais desafios da análise do corpus textual, que incluem a alta frequência de símbolos técnicos, erros de digitação, abreviatura de termos e a presença de frases incompletas nas descrições.

O REBEL, o UniRel e o DIRECT foram escolhidos por serem modelos do estado da arte com maior desempenho em extração de relações usando como *benchmark* o *dataset* do *New York Times* (NYT). A Figura 2 traz uma visão geral das diferentes abordagens de cada um.

REBEL é um modelo que aplica uma abordagem autorregressiva que enquadra a ER como uma tarefa *sequence-to-sequence*. Segundo a pesquisa, essa abordagem oferece vantagens em relação às anteriores graças à adoção de uma decomposição de tripla em

REBEL (Cabot e Navigli 2021)	Enquadra a tarefa de forma auto-regressiva
UniRel (Tang et al. 2022)	Unifica as entidades e relações em linguagem natural
DIRECT (Zhao et al. 2021)	Utiliza listas de adjacência para coordenar o processo

Figura 2. Resumo das características únicas dos modelos selecionados

uma sequência de texto. A sua simplicidade torna-o altamente flexível para se adaptar a novos domínios ou documentos mais longos.

O UniRel se preocupa com a representação heterogênea dos relacionamentos entre entidades que trabalhos anteriores apresentavam como problema, não conseguindo captar as relações semânticas mais ricas. Como solução proposta, ele unifica as representações de entidades e relações, codificando-as conjuntamente em uma sequência concatenada de linguagem natural.

Por fim, o DIRECT é um modelo orientado a listas de adjacência. Matrizes de adjacência são um meio de representar quais vértices de um grafo são adjacentes a quais outros vértices, e uma lista de adjacências é a representação de todas as arestas ou arcos de um grafo na forma de lista [Singh and Sharma 2012]. Se o grafo não for direcionado, cada entrada é um conjunto (ou multiconjunto) de dois nós contendo as duas extremidades da aresta correspondente; se for direcionado, cada entrada é uma tupla de dois nós, um denotando o nó de origem e o outro denotando o nó de destino do arco correspondente.

3. Metodologia

Conforme mencionado na introdução, o objetivo dessa pesquisa foi fazer um comparativo entre modelos generativos de ER para a construção de GCs. Mais especificamente, observar a qualidade das relações extraídas de textos do domínio da perfuração de poços de petróleo, por modelos que apenas receberam dados de domínio geral. Para isso, foi feita uma avaliação manual entre os resultados obtidos de cada modelo, utilizando como base triplas de um conjunto já existente. A Figura 3 apresenta um visão geral dos procedimentos realizados.

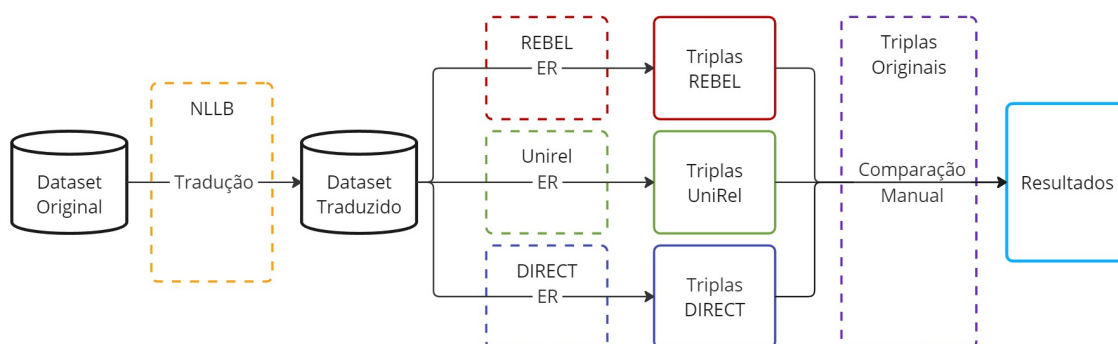


Figura 3. Fluxo dos experimentos realizados

3.1. Dataset

O *dataset* de teste foi montado a partir de uma coleção de relatórios de atividade de perfuração. Cada uma das 200 entradas é composta por descrições em texto livre e uma

lista de triplas esperadas para cada uma. Essas triplas, que serão usadas como padrão de ouro, foram obtidas a partir de uma ferramenta de extração que utiliza uma ontologia do domínio construída por especialistas. Considerando que os três modelos escolhidos para comparação foram treinados na língua inglesa e os relatórios estão escritos em português, foi necessário fazer a tradução das frases e triplas como parte do pré-processamento dos dados. Portanto, algumas nuances semânticas foram perdidas.

Para realizar a tradução das entradas do *dataset*, tanto para as descrições de operação quanto para as triplas esperadas de cada sentença, foi escolhido usar modelo NLLB-200, da Meta AI. O NLLB, que significa No Language Left Behind (Nenhuma língua deixada para trás), é um projeto de IA que cria modelos de tradução para mais de 200 idiomas, incluindo aqueles com poucos recursos [Meta 2024].

A versão do modelo utilizada, chamada NLLB-200's distilled 600M, é um modelo focado em pesquisa na tradução de Máquina, disponível na biblioteca de modelos do HuggingFace¹, possuindo perto de 600 milhões de parâmetros.

3.2. Métricas de Avaliação

Avaliar as triplas resultantes dos processos de ER se mostrou ser um desafio. Métricas comuns para *benchmark* como precisão e revocação não foram aplicadas, pois elas falhariam em reconhecer resultados com o mesmo significado semântico, porém escritos de maneiras diferentes. Logo, optou-se por uma avaliação manual qualitativamente das relações obtidas. As triplas extraídas foram comparadas ao conjunto de ouro, de forma exemplificada na Tabela 1, colocando-as em uma das seguintes categorias:

- **Categoria A:** A tripla representa uma relação que aparece nos resultados esperados ou é equivalente;
- **Categoria B:** A tripla representa uma relação correta que não está inclusa nos resultados esperados, ou seja, um conhecimento novo; e
- **Categoria C:** A tripla representa uma relação incorreta.

Tabela 1. Exemplo da avaliação das triplas obtidas

Gold Triplets	Generated Triplets	Triplets by Category		
		A	B	C
<drilling, involves, hole>, <hole, diameter, 12.25">	<hole, parameter, 12.25>, <circulated, parameter, clean>	1	0	1
<pullout, involves, hole>, <layingdown, involves, drillcollars>	<drillcollars, diameter, 8">	0	1	0

Por fim, também foi calculado o número de triplas extraídas e a porcentagem das entradas do *dataset* no qual o modelo conseguiu obter alguma relação.

3.3. Funcionamento dos modelos

A metodologia do REBEL aborda a extração e classificação das relações como uma tarefa de geração usando um modelo autorregressivo baseado no BART-large

¹<https://huggingface.co/facebook/nllb-200-distilled-600M>

[Lewis et al. 2019]. O processo consiste na tradução da sentença de entrada, suas entidades e relações implícitas para um conjunto de triplas que se refere explicitamente a elas. Estas triplas então são decodificadas e linearizadas pelo modelo para a geração da saída, como mostra a Figura 4.

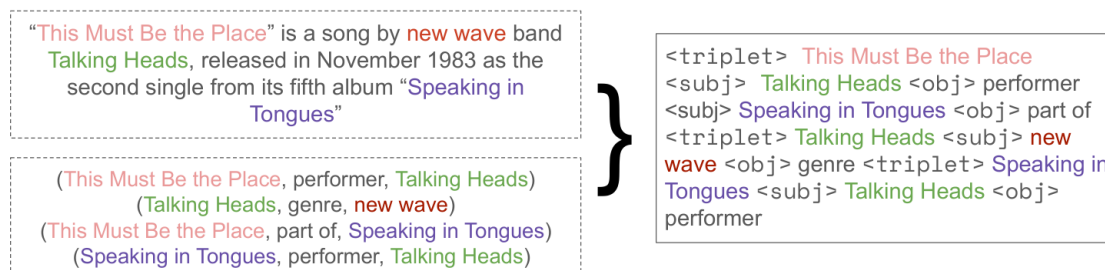


Figura 4. Exemplo da linearização de triplas do REBEL [Cabot and Navigli 2021]

No caso do UniRel, a relação e as entidades são codificadas em *embeddings* de sequências significativas para construir as representações unificadas. Elas são então convertidas em textos de linguagem natural para formar uma sequência consecutiva em conjunto da frase de entrada, com base na definição semântica. Depois, um modelo pré-treinado baseado em *transformer*, no caso o BERT [Devlin 2018], codifica a sequência e captura suas correlações. As dependências entre entidade-entidade e entre entidade-relação são então modeladas simultaneamente em um mapa de interação, exemplificado na Figura 5

O DIRECT aplica uma estratégia adaptativa de aprendizagem multitarefa. Seu *framework* é composto por um codificador BERT compartilhado e três camadas de saída: extração de sujeito, extração de objeto e classificação da relação, todas baseadas em listas de adjacência, exemplificadas na Figura 6. Cada sujeito extraído pela primeira camada é concatenado com a sentença e enviado para a segunda, formando pares de sujeito-objeto, os quais são enviados à terceira usando o mesmo procedimento para obter a relação entre eles.

4. Experimentos e Resultados

Os experimentos realizados consistem na formatação e tradução do *dataset* original para elaborar o conjunto de teste e na execução da extração usando cada um dos três modelos escolhidos. O processo de tradução e inferência nos modelos UniRel, Direct e REBEL estão descritos a seguir, respectivamente.

4.1. Ambiente de Teste

O *hardware* utilizado para a realização dos experimentos foi um notebook Acer do modelo Nitro 5 AN515-46-R5WF com especificações descritas na Tabela 2. O ambiente de programação e os códigos foram desenvolvidos nos seguintes *softwares*:

- Python 3.12 como a linguagem de programação principal;
- PyTorch 2.2.2, uma estrutura de código aberto para construção e treinamento de modelos de aprendizado profundo baseados em redes neurais;

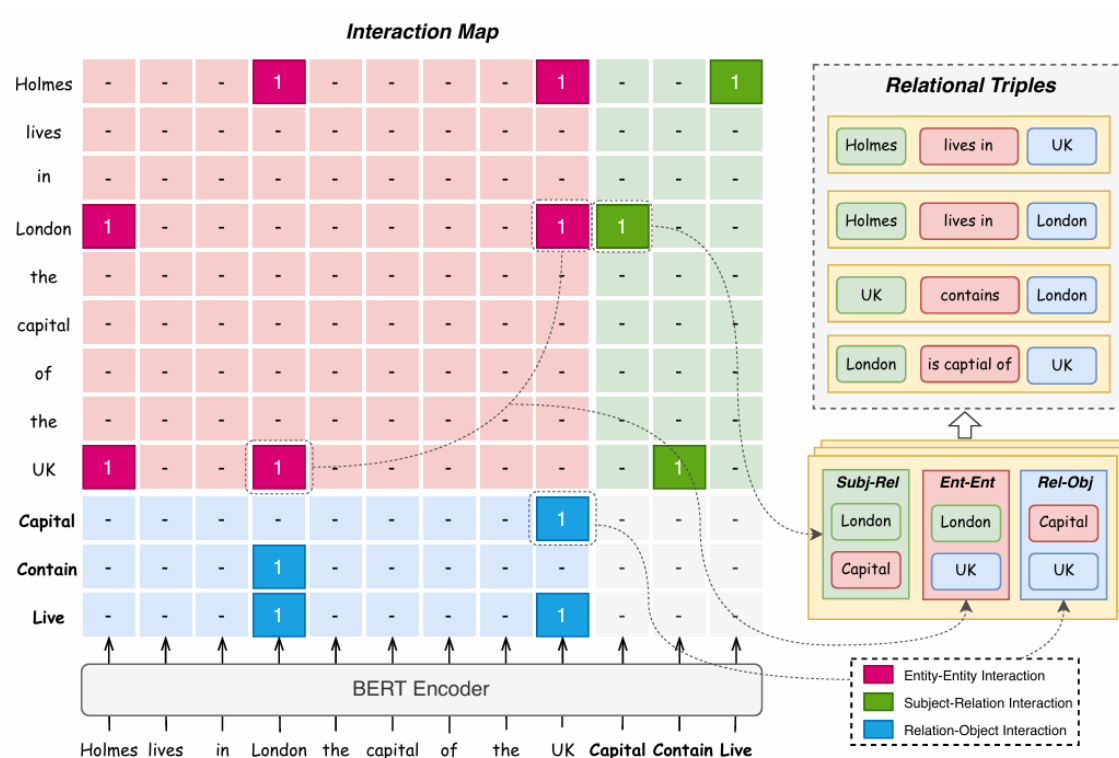


Figura 5. Exemplo de um mapa de interação gerado pelo UniRel [Tang et al. 2022]

- CUDA 12.1, uma plataforma de computação paralela e modelo de programação desenvolvido pela NVIDIA para computação geral em unidades de processamento gráfico (GPUs); e
- HuggingFace, um repositório especializado em tarefas de PLN de onde é possível obter modelos pré-treinados.

Tabela 2. Ficha técnica do notebook utilizado

Processador	Ryzen 7 6800H
Placa de Vídeo	RTX 3070 Ti (150W)
RAM	16GB DDR5 4800MHz
SO	Windows 11

4.2. Processo de preparação dos modelos para extração

Esta seção irá explicar os passos adotados para preparar o ambiente com os modelos para realizar a extração.

4.2.1. UniRel

Para a extração com o uniRel, foi utilizado o *checkpoint* disponibilizado em seu repositório², treinado em um recorte de 56196 sentenças do *dataset* "New York Times Annotated Corpus". Este *dataset* possui 24 classes, sendo composto por artigos escritos e

²<https://github.com/wtangdev/UniRel>

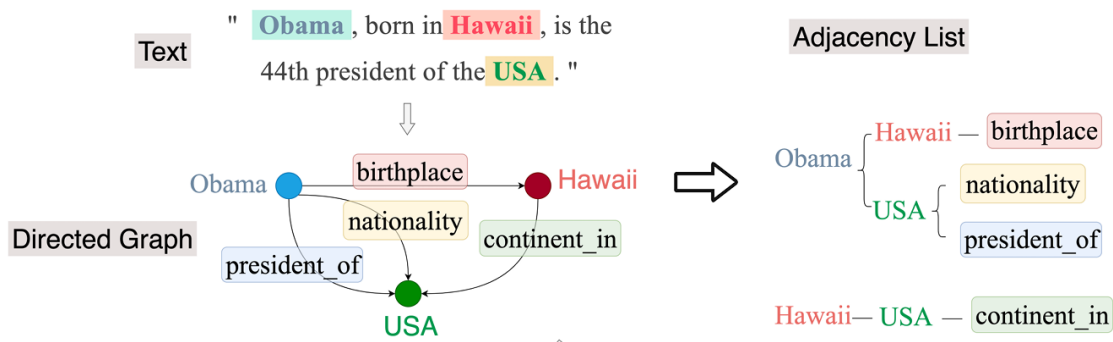


Figura 6. Relacionamentos na lista de adjacência [Zhao et al. 2021] (adaptado)

publicados pelo New York Times (NYT) entre 1º de janeiro de 1987 e 19 de junho de 2007.

Dado o *checkpoint*, foi possível fazer a extração das triplas buscando a possibilidade do modelo encontrar alguma relação no *dataset* de teste que represente uma das classes do NYT.

4.2.2. Direct

Para o processo de extração com o Direct, foi necessário fazer o *fine tuning* do modelo a partir do bert-base-cased, um modelo baseado em transformers pré-treinado em um grande corpo de dados em inglês de forma auto-supervisionada.

Os códigos e *datasets* disponibilizados no repositório do modelo³ para o pré-treinamento foram o NYT, mencionado anteriormente, e o WebNLG, um corpus composto por conjuntos de triplas e os fatos correspondentes na forma de texto em linguagem natural. Por ser mais robusto, com 170 classes diferentes, o *fine tuning* foi feito com o WebNLG.

4.2.3. Rebel

Por fim, o processo de extração com o REBEL foi o mais simples dos três devido ao modelo estar disponibilizado diretamente no repositório da plataforma HuggingFace para o uso.

Mais especificamente, a versão usada foi o mRebel-large⁴, um *checkpoint* treinado em cima de um extenso *dataset* multilíngue de ER chamado SRED^{FM}, baseado na interconexão de dados da Wikipédia e Wikidata, filtrados por um sistema crítico treinado em anotações humanas [Cabot et al. 2023].

4.3. Resultados

Como resultado dos experimentos de extração, realizados, foi elaborada a Tabela 3 contendo os valores obtidos para as métricas escolhidas e detalhadas anteriormente. Para

³<https://github.com/fyubang/direct-ie>

⁴<https://huggingface.co/Babelscape/mrebel-large>

cada um dos três modelos, estão exibidos o número de triplas extraídas no total, o número e a porcentagem relativa de triplas avaliadas em cada categoria, e a porcentagem de linhas onde o modelo obteve algum retorno.

Tabela 3. Resultados das tentativas de extração

Modelo	A	B	C	Triplas	% de linhas
UniRel	0 0,0%	2 11,1%	16 88,9%	18	8,0%
Direct	0 0,0%	3 30,0%	7 70,0%	10	5,0%
Rebel	1 0%	48 22%	174 78%	223	99%

Como é possível observar, os resultados encontrados na extração com os três modelos foi insatisfatório, especialmente nos casos do UniRel e do Direct, ambos extraindo triplas em menos de 10% das entradas do dataset. Tal resultado foi de se esperar devido ao pré-treinamento em conjuntos pequenos de dados de domínio geral. O REBEL, graças ao treinamento mais extensivo, obteve um grande retorno de triplas, porém também falhou em reconhecer as relações corretamente.

5. Conclusão

Em conclusão, esta pesquisa reuniu modelos de IA com altos *benchmarks* na extração automática de relações de textos visando testá-los em uma base de dados muito mais técnica do que a de seu treinamento. O processo passou por algumas condições não ideais, sendo a principal delas a necessidade de fazer a tradução do dataset para a língua inglesa, o que naturalmente adicionou erros na coleção de dados. Outra deficiência foi na fase de anotação manual dos resultados. Por mais que os padrões de ouro tenham sido gerados por uma ferramenta de confiança, a comparação entre os resultados esperados e obtidos foi feita por um pesquisador sem a especialização ideal na área. Possivelmente, uma análise feita por um conjunto de avaliadores com proficiência no domínio resultaria em classificações diferentes para algumas triplas.

Apesar dos problemas mencionados, é seguro dizer que o objetivo do trabalho foi alcançado. Ficou aparente que, apesar do avanço da qualidade dos métodos generativos para construção automática de GCs, os resultados demonstraram a grande queda de eficácia quando eles são aplicados em um contexto diferente e especializado. Mesmo os modelos de ER mais atuais ainda precisam ser treinados em um banco de dados característico do domínio em que se deseja obter bons resultados. Dessa forma, mostra-se fundamental a criação de boas bases de dados para aproveitar todo o potencial dos métodos de extração avaliados.

Como sequência a esta pesquisa, esforços serão direcionados à elaboração de um *dataset* de treino extensivo contendo dados de perfuração de poços e suas respectivas relações. Futuros testes de modelos de extração serão feitos a partir do *fine tuning* nos dados específicos do domínio, o que resultará em uma comparação mais justa dos méritos de cada metodologia empregada.

6. Agradecimentos

Agradecemos o apoio do Unespetro e da Petrobras no financiamento do projeto Proto-Radiar.

Referências

- Abdullah, M., Madain, A., and Jararweh, Y. (2022). Chatgpt: Fundamentals, applications and social impacts. In *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pages 1–8. Ieee.
- Athiwaratkun, B., Nogueira dos Santos, C., Krone, J., and Xiang, B. (2020). Augmented natural language for generative sequence labeling. In Webber, B., Cohn, T., He, Y., and Liu, Y., editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 375–385, Online. Association for Computational Linguistics.
- Bosselut, A., Rashkin, H., Sap, M., Malaviya, C., Celikyilmaz, A., and Choi, Y. (2019). COMET: Commonsense transformers for automatic knowledge graph construction. In Korhonen, A., Traum, D., and Màrquez, L., editors, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4762–4779, Florence, Italy. Association for Computational Linguistics.
- Cabot, P.-L. H. and Navigli, R. (2021). Rebel: Relation extraction by end-to-end language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 2370–2381.
- Cabot, P.-L. H., Tedeschi, S., Ngomo, A.-C. N., and Navigli, R. (2023). Red^{FM}: a filtered and multilingual relation extraction dataset. *arXiv preprint arXiv:2306.09802*.
- Cinelli, L. P., de Oliveira, J. F., de Pinho, V. M., Passos, W. L., Padilla, R., Braz, P. F., Galves, B., Dalvi, D. P., Lewenfus, G., Ferreira, J. O., et al. (2021). Automatic event identification and extraction from daily drilling reports using an expert system and artificial intelligence. *Journal of Petroleum Science and Engineering*, 205:108939.
- De Cao, N., Izacard, G., Riedel, S., and Petroni, F. (2020). Autoregressive entity retrieval. *arXiv preprint arXiv:2010.00904*.
- Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Freitas, C., Souza, E., Castro, M. C., Cavalcanti, T., da Silva, P. F., and Cordeiro, F. C. (2015). Recursos linguísticos para o pln específico de domínio: o petrolês. *Linguamática*, 15(2):51–68.
- Gomes, D. d. S. M., Cordeiro, F. C., Consoli, B. S., Santos, N. L., Moreira, V. P., Vieira, R., Moraes, S., and Evsukoff, A. G. (2021). Portuguese word embeddings for the oil and gas industry: Development and evaluation. *Computers in Industry*, 124:103347.

- Hoffmann, J., Mao, Y., Wesley, A., and Taylor, A. (2018). Sequence mining and pattern analysis in drilling reports with deep natural language processing. In *SPE Annual Technical Conference and Exhibition?*, page D031S033R004. SPE.
- Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., et al. (2021). Knowledge graphs. *ACM Computing Surveys (Csur)*, 54(4):1–37.
- Huang, K.-H., Tang, S., and Peng, N. (2021). Document-level entity-based extraction as template generation. In Moens, M.-F., Huang, X., Specia, L., and Yih, S. W.-t., editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5257–5269, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., and Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38.
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.
- Liu, X., Huang, H., Shi, G., and Wang, B. (2022). Dynamic prefix-tuning for generative template-based event extraction. In Muresan, S., Nakov, P., and Villavicencio, A., editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5216–5228, Dublin, Ireland. Association for Computational Linguistics.
- Lu, Y., Liu, Q., Dai, D., Xiao, X., Lin, H., Han, X., Sun, L., and Wu, H. (2022). Unified structure generation for universal information extraction. In Muresan, S., Nakov, P., and Villavicencio, A., editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5755–5772, Dublin, Ireland. Association for Computational Linguistics.
- Melnyk, I., Dognin, P., and Das, P. (2022). Knowledge graph generation from text. *arXiv preprint arXiv:2211.10511*.
- Meta (2024). Meta ai research topic - no language left behind.
- Nadkarni, P. M., Ohno-Machado, L., and Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5):544–551.
- Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., and Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*.
- Ribeiro, L. C., Afonso, L. C., Colombo, D., Guilherme, I. R., and Papa, J. P. (2020). Evolving neural conditional random fields for drilling report classification. *Journal of Petroleum Science and Engineering*, 187:106846.
- Singh, H. and Sharma, R. (2012). Role of adjacency matrix & adjacency list in graph theory. *International Journal of Computers & Technology*, 3(1):179–183.

- Tang, W., Xu, B., Zhao, Y., Mao, Z., Liu, Y., Liao, Y., and Xie, H. (2022). Unirel: Unified representation and interaction for joint relational triple extraction. *arXiv preprint arXiv:2211.09039*.
- Yan, H., Dai, J., Ji, T., Qiu, X., and Zhang, Z. (2021a). A unified generative framework for aspect-based sentiment analysis. In Zong, C., Xia, F., Li, W., and Navigli, R., editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 2416–2429, Online. Association for Computational Linguistics.
- Yan, H., Gui, T., Dai, J., Guo, Q., Zhang, Z., and Qiu, X. (2021b). A unified generative framework for various NER subtasks. In Zong, C., Xia, F., Li, W., and Navigli, R., editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5808–5822, Online. Association for Computational Linguistics.
- Ye, H., Zhang, N., Chen, H., and Chen, H. (2022). Generative knowledge graph construction: A review. *arXiv preprint arXiv:2210.12714*.
- Zeng, X., Zeng, D., He, S., Liu, K., and Zhao, J. (2018). Extracting relational facts by an end-to-end neural model with copy mechanism. In Gurevych, I. and Miyao, Y., editors, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 506–514, Melbourne, Australia. Association for Computational Linguistics.
- Zhao, F., Jiang, Z., Kang, Y., Sun, C., and Liu, X. (2021). Adjacency list oriented relational fact extraction via adaptive multi-task learning. *arXiv preprint arXiv:2106.01559*.
- Zhong, L., Wu, J., Li, Q., Peng, H., and Wu, X. (2023). A comprehensive survey on automatic knowledge graph construction. *ACM Computing Surveys*, 56(4):1–62.