

Universidade de Pernambuco Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Computação Inteligente

Título: Validação de Coerência Contextual em Sistemas de *Retrieval Augmented Generation*

Orientador – Fernando Buarque de Lima Neto (fbln@ecomp.poli.br)

Co-orientador – Luís Filipe de Araújo Pessoa (lfap@ecomp.poli)

Contexto

Os recentes avanços em *Large Language Models* (LLMs) [1] têm viabilizado o processamento contextualizado e semântico de dados textuais, visuais e auditivos. Eles são centrais para sistemas atuais de *Retrieval Augmented Generation* (RAGs) [2] e Agentes Inteligentes (ou Agentes LLMs), viabilizando a indexação semântica dos dados (i.e., *embeddings*) [3], a extração de informações, a geração de respostas contextuais às requisições dos usuários e até mesmo a tomada de decisões contextuais autônomas. Embora as respostas sejam com frequência gramaticalmente corretas e semanticamente plausíveis, a integração entre conhecimento geral dos LLMs e informações específicas de bases de dados privativas, além do problema inerente de alucinações, introduz desafios críticos de coerência contextual, como inconsistências factuais, contradições ou inadequações terminológicas. A coerência contextual é particularmente crítica em domínios especializados como o jurídico que exige extrema precisão factual.

Problema

Sistemas RAG e de Agentes Inteligentes atuais carecem de mecanismos eficazes para validar e corrigir automaticamente incoerências contextuais que emergem da integração entre conhecimento geral dos LLMs e informações contextuais recuperadas. A ausência de validação contextual automatizada pode resultar em informações incorretas que comprometem a confiabilidade do sistema, especialmente em aplicações críticas.

Hipótese

Uma abordagem integrativa para validação e correção de coerências contextuais baseadas em modelos de *deep learning* (e.g., *transformers*, *embeddings* e LLMs) [4] e métricas e métodos tradicionais de Processamento de Linguagem Natural (NLP) pode melhorar significativamente a qualidade, consistência e precisão factual e terminológica das respostas.

Pergunta Principal

Como desenvolver uma abordagem integrativa que combine eficazmente modelos de *deep learning* e métodos tradicionais de NLP para validação e correção automática de (in)coerência contextual em sistemas RAG e Agentes Inteligentes?

Perguntas Secundárias

- Quais métricas e métodos tradicionais de NLP podem ser usadas para avaliar e validar a coerência textual?
- Quais abordagens baseadas em modelos de *deep learning* podem ser usadas para avaliar e validar a coerência contextual?
- Como integrar as métricas e métodos tradicionais de NLP com modelos de *deep learning* para validar e corrigir a (in)coerência contextual?

Objetivos

Desenvolver, implementar e avaliar uma abordagem integrativa que combine modelos de *deep learning* e métodos tradicionais de NLP para validação e correção automática de coerência contextual em sistemas RAG e Agentes Inteligentes, com validação experimental no domínio jurídico brasileiro.

Produtos esperados

- Revisão quase-sistemática das métricas e métodos tradicionais de NLP, bem como de abordagens baseadas em modelos de *deep learning* para a avaliação e validação de coerência contextual.
- Modelo para integração das métricas e métodos tradicionais de NLP com as abordagens baseadas em modelos de *deep learning* para validação e correção de incoerências contextuais.
- Teste e avaliação de aplicação da contribuição em estudos de caso no domínio jurídico.

Referências

- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., ... & Mian, A. (2023). A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology*.
- Wang, X., Wang, Z., Gao, X., Zhang, F., Wu, Y., Xu, Z., ... & Huang, X. (2024). Searching for best practices in retrieval-augmented generation. *arXiv preprint arXiv:2407.01219*.
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.
- Gardazi, N. M., Daud, A., Malik, M. K., Bukhari, A., Alsahfi, T., & Alshemaimri, B. (2025). BERT applications in natural language processing: a review. *Artificial Intelligence Review*, 58(6), 1-49.