

Universidade de Pernambuco Programa de Pós-Graduação em Engenharia da Computação (PPGEC)
Proposta de Dissertação de Mestrado
Área: Computação Inteligente
Título: Integração Neuro-Simbólica: Estratégias para combinação de <i>background knowledge</i> com algoritmos inteligentes para otimização de modelos XAI
Orientador – Cleyton Mário de Oliveira Rodrigues (cleyton.rodrigues@upe.br) Co-orientadora – Patrícia Takako Endo (patricia.endo@upe.br)
<p>Descrição – A Inteligência Computacional compreende a construção de algoritmos de aprendizagem para identificar e extrair padrões, regularidades ou tendências a partir de <u>diferentes tipos de grandes volumes de dados</u>, desde os mais organizados até os mais desestruturados. Em geral, estas tarefas podem envolver outras capacidades, como raciocínio, aprendizagem, tomada de decisão, otimização e autonomia. Se por um lado, as técnicas passaram a ser mais acuradas e complexas, por outro lado, os mecanismos internos de decisão tornaram-se menos transparentes, levando a um <i>trade-off</i> entre desempenho e interpretabilidade.</p> <p>Em virtude da falta de transparência dos modelos inteligentes (<u>vistos como uma caixa-preta</u>), a área conhecida como <i>eXplainable Artificial Intelligence</i> (XAI) busca criar modelos de aprendizagem de máquina capazes de tanto obter alto <u>desempenho</u> <u>acurácia</u> nas predições, quanto explicar as decisões em certos níveis de detalhes. Este alinhamento traz diversos benefícios como aumento na confiança das soluções, atendimento às questões regulatórias, e às necessidades dos domínios que requerem tais explicações (saúde, negócio, jurídico, entre outros). Uma possibilidade para prover as explicações é o uso do conhecimento de domínio, os quais podem ser devidamente formalizados em modelos ontológicos livres de ambiguidades. Ontologias podem ser convenientemente representadas por recursos lógicos-matemáticos, que garantem a possibilidade de tarefas de raciocínio, úteis inclusive para justificar inferências realizadas.</p> <p>Portanto, este projeto de pesquisa tem como objetivo propor mecanismos de alinhamento entre modelos de aprendizagem com elementos ontológicos (conceitos, relações, axiomas e regras) sobre o domínio no qual o modelo inteligente foi construído para que seja possível gerar justificativas simbólicas das decisões. Também faz parte da pesquisa propor abordagens para avaliação das explicações geradas, e sua conformidade com a ontologia de domínio.</p>
Referências Bibliográficas <ol style="list-style-type: none">1. ANGELOV, Plamen P. et al. Explainable artificial intelligence: an analytical review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, v. 11, n. 5, p. e1424, 2021.2. CHARI, Shruthi et al. Explanation ontology: A model of explanations for user-centered ai. In: International Semantic Web Conference. Springer, Cham, 2020. p. 228-243.3. HOLZINGER, Andreas. From machine learning to explainable AI. In: 2018 world symposium on digital intelligence for systems and machines (DISA). IEEE, 2018. p. 55-66.

[PE1] Comentário: Tirei o “grandes volumes” para não ficar preso em de learning...

Formatado: Fonte: Itálico