

Engenharia Física

O uso de redes complexas semânticas aplicadas à linguagem natural

Maria Clara Oliveira Annoni Farah - 8º módulo de Engenharia Física, UFLA, bolsista PIBIC/FAPEMIG

Angélica Sousa da Mata - Orientadora DFI, UFLA - Orientador(a)

Aline Viol - Coorientadora, Donders Institute for Brain, Cognition and Behaviour, Radboud University

Resumo

As redes complexas constituem um campo interdisciplinar de investigação, caracterizado pela representação de sistemas como conjuntos de pontos (nós) interconectados por ligações (arestas), cuja dinâmica resulta em comportamentos coletivos e emergentes. Entre suas aplicações contemporâneas, destaca-se a análise semântica no processamento de linguagem natural, que, embora à primeira vista possa parecer pouco intuitiva no âmbito de ferramentas matemáticas, revela-se promissora para a modelagem de textos. Nesse contexto, um texto pode ser traduzido para uma representação computacional por meio de modelos de aprendizado de máquina, como o BERT, que convertem sentenças em hipervetores distribuídos em um espaço semântico de alta dimensionalidade. A partir desses vetores, calcula-se uma matriz de similaridade, comumente baseada no cosseno, a fim de quantificar relações semânticas entre sentenças. Para viabilizar esse processo, emprega-se a linguagem de programação Python, juntamente com bibliotecas voltadas à manipulação de dados e ao uso de modelos de aprendizado profundo (deep learning). O procedimento consiste em extrair informações linguísticas dos textos, convertê-las em representações vetoriais e, em seguida, estruturá-las sob a forma de redes complexas. A partir dessa modelagem, tornam-se aplicáveis medidas clássicas do estudo de redes, como medidas de centralidade e entropia, possibilitando a análise quantitativa de aspectos semânticos e estruturais do discurso.

Palavras-Chave: redes complexas, processamento de linguagem natural, aprendizado de máquina.

Instituição de Fomento: FAPEMIG

Link do pitch: <https://youtu.be/C5h7Zet06zc>