Projeto de Formatura – 2021 – Press Release

PCS - Departamento de Engenharia de Computação e Sistemas Digitais

Engenharia de Computação

Tema: ZeroBERTo - Zero-Shot BERT based on Topic Modeling applied to the Portuguese Parliament

Estudantes da Escola Politécnica da USP desenvolvem algoritmo de Inteligência Artificial para interpretação de textos em português no contexto de Democracia Digital

São Paulo, 08 de dezembro de 2021

Um dos princípios da Democracia Digital é informar e mobilizar cidadãos para que participem do debate político ativamente. Nesse contexto, Inteligência Artificial (IA) pode ser uma grande aliada, auxiliando na interpretação da quantidade cada vez maior de dados públicos disponíveis. Dentre as tecnologias de IA, uma área que vem crescendo muito é a relacionada ao Processamento de Linguagem Natural (PLN), que lida com algoritmos capazes de entender aspectos da comunicação humana. Algoritmos de PLN baseados em aprendizado costumam requerer uma grande quantidade de dados rotulados para treinamento, o que dificulta trabalhar em domínios restritos ou línguas com baixa disponibilidade de dados.

Foi nesse contexto que os alunos Alexandre Teodoro De Siqueira Guedes Alcoforado e Rodrigo Elizardo Gerber, orientados pela Professora Dra. Anna Helena Reali Costa e pelo Prof. Dr. Fabio Levy Siqueira, desenvolveram o **ZeroBERTo**: um modelo de IA capaz de classificar dados textuais, preparado para lidar com baixa disponibilidade de dados. **ZeroBERTo** faz uso de duas etapas para promover a classificação: primeiro, dado um conjunto de dados textuais, aprende a representá-los de forma compacta, em tópicos; depois, classifica os dados utilizando essa representação de tópicos ao invés de utilizar todo o texto.

ZeroBERTo foi então aplicado nas atas das reuniões plenárias do Parlamento Português. O texto relatado nas atas - a fala de cada deputado - é classificado entre temas de interesse público, como Educação, Infraestrutura, Economia e Direitos Humanos, viabilizando a obtenção de estatísticas dos temas debatidos pelos parlamentares, assim como seus autores, datas e contexto. Os resultados, ilustrados na figura a seguir, reforçam a oportunidade e a importância de se utilizar técnicas modernas de IA para facilitar o acesso à informação pública pela população.

Integrantes: - Alexandre Teodoro De Siqueira Guedes Alcoforado
- Rodrigo Elizardo Gerber