

I - Workshop de Dissertações do Programa de Pós-graduação em Sistemas de Informação (PPgSI)

Análise de viés em notícias na língua portuguesa

Aluno: Gabriel Domingos de Arruda NUSP 6776827, **Orientador:** Dr. Norton Trevisan Roman

Contexto

O viés é um problema inerente a produção de notícias, nenhum canal de comunicação está livre desse problema. As novas formas de consumir conteúdo, utilizando pesquisas na internet e *feeds* de redes sociais, estão agravando o problema da “Câmera de Eco” ao recomendar conteúdo de preferência dos usuários.

Objetivo

O objetivo desse trabalho é aplicar técnicas de *Sentiment Analysis*, baseadas em aprendizado de máquina, em notícias da língua portuguesa para identificar se são positivas ou negativas. Dessa forma, averiguar se o desempenho de técnicas aplicadas em outros idiomas se mantém na língua portuguesa. Além da aplicação das técnicas, pretende-se utilizar meta-dados das notícias para melhorar o processo de classificação junto com as técnicas de *Sentiment Analysis*.

Método

Utilizando um corpus criado para esse projeto, o desempenho de algoritmos de classificação baseados em aprendizado de máquina serão comparados entre si. Meta-dados, como a quantidade de palavras de uma notícia, serão incluídos nos algoritmos junto com as informações no texto para verificar a influência no processo de classificação.

Resultados

Espera-se que técnicas mais sofisticadas de classificação obtenham melhor desempenho na língua portuguesa, como ocorre em outros idiomas. Além disso, a inclusão dos meta-dados também deveria aumentar o desempenho do processo de classificação.

Conclusões

Novas formas de consumir conteúdo e o grande volume de informação produzida estão agravando o problema do consumo de notícias enviesadas. O uso de técnicas de *Sentiment Analysis* em notícias é uma alternativa para apresentar melhores sistemas de recomendação, que foquem em anular o viés e não em trazer mais conteúdo que agrada ao usuário.