

EXTRAÇÃO DE CARACTERÍSTICAS EM DOCUMENTOS PARA IDENTIFICAÇÃO DE TEXTOS OFENSIVOS

Cleiton de Lima Pinto¹

Guilherme Dal Bianco²

Resumo: As mídias sociais estão cada vez mais presente na vida das pessoas, incluindo ferramentas que permitem que o usuário colabore com a criação do conteúdo nela exposto. Muitos usuários se aproveitam dessa funcionalidade para disseminar conteúdo ilícito ou criminoso. Caso não seja removido, este conteúdo será visto por cada vez mais pessoas e poderá ser propagado pela internet, atingindo um número maior de vítimas, e incentivando a ocorrência de outros crimes. Com isso, através da classificação de textos juntamente com o aprendizado de máquina, é possível identificar discurso de ódio em documentos de forma automática. Para tal tarefa, métodos supervisionados de aprendizagem de máquina são aplicados para a criação de modelos que predizem se determinado documento é possui conteúdo ofensivo ou não. No aprendizado supervisionado, é fornecido ao sistema de aprendizado um conjunto de exemplos (textos), sendo que cada exemplo possui um rótulo associado. O rótulo determina se o texto é ofensivo ou não. Através de um nova entrada não rotulada, o classificador é capaz de prever a classe a qual o dado se assemelha. Para o funcionamento dos algoritmos de aprendizagem de máquina, é preciso que os dados possuam determinadas características para realizar a classificação dos textos. Essas características nada mais são que informações que descrevem determinado documento, como por exemplo, a quantidade de palavras no documento, o número de ocorrências de determinado termo nos documentos, etc. Além das características dos documentos, o pré-processamento dos textos, como a remoção de sufixos e acentos das palavras, é importante para que a classificação seja mais precisa e a dimensão do conjunto de dados minimizada. Este trabalho tem como objetivo explorar a criação de características utilizando técnicas de processamento de linguagem natural e aprendizado de máquina para detectar automaticamente discursos de ódio. Esse tipo de crime geralmente é voltado aos grupos mais vulneráveis da sociedade, e seus efeitos nocivos pode causar o aumento da exclusão social e da violência praticada contra esses grupos.

Palavras-chave: Mineração de dados. Aprendizado de Máquina. Classificação de Texto.

¹ Acadêmico de Ciência da Computação, UFFS, Chapecó, contato: cleiton.limapin@gmail.com

² Doutor, UFFS, Chapecó, , contato: dbguilherme@gmail.com



Anais do SEPE – Seminário de Ensino, Pesquisa e Extensão
Vol. VIII (2018) – ISSN 2317-7489



Categoria: Pesquisa

Área do Conhecimento: Ciências Exatas e da Terra

Formato: Comunicação Oral