

Detectando Avaliações Falsas na Internet Usando Classificação de Texto

Daniel Atkinson

Universidade Federal do Rio de Janeiro, Rio de Janeiro RJ, Brasil
danatkho@gmail.com

Abstract. Nos últimos anos o problema de notícias falsas tem se tornado cada vez mais aparente. A abundância desse tipo de notícia na internet apresenta um grande risco para a maioria dos usuários, especialmente para aqueles que só lêem manchetes. Um outro problema muito relacionado ao de notícias falsas é o de avaliações falsas. Ambos os tipos de publicações têm como objetivo apresentar informações ou opiniões falaciosas e são difíceis de ser detectados por humanos. No caso das avaliações falsas tanto consumidores quanto prestadores de serviços são afetados. Ambas essas questões, inicialmente muito discutidas no ambiente popular e midiático, estão ganhando cada vez mais atenção no ambiente acadêmico pois esse tipo de conteúdo continua a crescer, tornando-o um problema maior, porém ao mesmo tempo, aumentando as possibilidades de estudos. Neste artigo é feita a implementação de um modelo n-grama para detectar automaticamente avaliações falsas. O modelo implementado é baseado no modelo criado por Ahmed et al.(1) e pode ser usado também para detecção automática de notícias falsas. No estudo original duas técnicas de extração de atributos e seis técnicas de classificação por aprendizado de máquina são comparados. Também são testados diferentes valores para os n-gramas e diferentes faixas de atributos. Neste artigo somente será implementado e analisado o caso que resultou na maior taxa de acerto no artigo original.

Keywords: avaliações falsas online · notícias falsas · classificação de texto · segurança de redes sociais online.

1 Introdução