

IDENTIFICAÇÃO DE PERFIS COM IDEAÇÃO SUICIDA NAS REDES SOCIAIS

XXXVIII Encontro de Iniciação Científica

George Abreu Masullo, Alex Lima Silva

O suicídio está entre as principais causas de morte no mundo, a frente de câncer de mama, guerra e homicídio e sendo a segunda principal causa de morte entre pessoas com 15 a 29 anos. No mundo, cerca de 800.000 pessoas morrem todos os anos. O presente trabalho tem como objetivo geral, obter um instrumento capaz de identificar perfis de brasileiros, que contenham ideação suicida (pensamento recorrente de cometer suicídio) no Twitter por meio de postagens de texto. Os dados utilizados neste estudo foram extraídos de um banco de dados da rede social Twitter, utilizando sua Application Programming Interface (API). O texto coletado é identificado como positivo, se existe algum indício de ideação suicida, caso contrário, negativo. Também é identificado a emoção: alegria, medo, tristeza, raiva, desgosto e surpresa. A identificação é realizada manualmente pelo autor. Os dados coletados são submetidos ao pré-processamento. Esta etapa é composta pelos seguintes estágios: transformação do texto para letras minúsculas, remoção da acentuação, remoção dos caracteres especiais, remoção das stopwords e remoção de sufixos e/ou prefixos de palavras. Por fim, a base de dados é submetida aos algoritmos naive bayes, k-means, SVM e árvore de decisão, separadamente. O desempenho de cada é comparado por meio da acurácia, precisão e recall. Como resultado parcial o naive bayes obtém melhor resultado com acurácia de 0.58. Para trabalhos futuros será testado o desempenho de redes neurais artificiais.

Palavras-chave: APRENDIZAGEM DE MÁQUINA. REDES SOCIAIS. SUICÍDIO. TWITTER.