

Processamento de Linguagem Natural para Detecção de Comportamento Depressivo

Gabriel Filipe da Silva Melo¹, Kayke Bonafé², Guilherme Wachs Lopes³

^{1,2,3} Ciência da Computação, Centro Universitário FEI

dev.gabrielomelo@gmail.com¹, kayke.bonafe98@gmail.com², gwachs@gmail.com³

Resumo: A depressão é um assunto que vem ganhando destaque nos últimos anos. De acordo com WHO [1], a depressão afeta mais de 294 milhões de pessoas ao redor do mundo. Trabalhos como o de [2] indicam que o diagnóstico precoce é um importante campo de pesquisa, uma vez que, em casos mais graves, a depressão pode levar ao suicídio. Portanto, este trabalho desenvolve, implementa e avalia um modelo computacional baseado em processamento de linguagem natural para classificação de tendências depressivas de usuários do *Twitter* por meio de suas postagens no decorrer do tempo. Como resultado, foi obtido um *F-Measure* de 83,58% utilizando-se do conteúdo textual acrescido da análise de sentimento dos documentos.

1. Introdução

Como a maioria das doenças, o diagnóstico precoce de um quadro de depressão é um fator fundamental para aumentar a taxa de sucesso do tratamento, minimizando as chances do quadro evoluir para um quadro de depressão profunda ou até para o suicídio.

Nesse sentido, as redes sociais podem ser um grande campo de estudo para detecção precoce de comportamentos depressivos. Um exemplo disso é o estudo de [3], que propõe inferir se há uma relação entre o uso de múltiplas redes sociais e o desenvolvimento de quadros como ansiedade e depressão. Os resultados deste estudo mostraram que de fato há uma correlação entre estes fatores.

Seguindo por esse ponto, é possível criar um modelo que consiga mostrar as informações relacionadas aos sentimentos contidos no texto, no caso, a variação de sentimentos, de modo que seja possível inferir acerca de quanta influência essa variação afeta a classificação.

Portanto, o objetivo deste trabalho é desenvolver, implementar e avaliar um modelo computacional baseado em processamento de linguagem natural para classificação de tendências depressivas de usuários do *Twitter* por meio de suas postagens no decorrer do tempo.

2. Metodologia

A metodologia deste trabalho é dividida em 4 partes, sendo elas: Modelagem de Dados, Pré-Processamento, Vetorização e Treinamento e Classificação. Uma representação da metodologia pode ser vista na Figura 1.

Na etapa da modelagem de dados ocorre a modelagem e formatação dos dados da base utilizada neste trabalho. A base em si é composta por mais de 3 milhões de tweets colhidos do *Twitter* entre 2009 e 2016 e dividida em 3 categorias diferentes: D1, que consiste em usuários que explicitamente declaram possuir depressão; D2, que consiste em usuários que não se classificam como depressivos e D3, que consiste em dados não classificados.

Na etapa de pré-processamento, o objetivo é realizar todos os pré-processamentos necessários para aumentar a qualidade dos dados obtidos na etapa anterior, onde foram aplicadas diversas técnicas para deixar os textos mais densos, com palavras mais uniformes e uma maior co-ocorrência, visando uma precisão maior na hora de processar esses dados.

Já na etapa de vetorização, o objetivo é transformar todos os textos pré-processados em vetores. Para isso, são utilizadas duas estratégias: o *Doc2Vec* e o *Bag of Words*. A primeira delas, o *Doc2Vec*, é utilizada para análise textual, indo diretamente para a rede neural recorrente *LSTM* e a segunda, o *Bag of Words*, é utilizada para a análise de sentimentos e polaridade sentimental, que é realizado pelo classificador *SVM*.

Por fim, na etapa de treinamento e classificação, o objetivo é formatar os dados para que possam ser treinados tanto no *LSTM* quanto no *SVM*, visando a detecção de possíveis padrões relacionados à problemática deste trabalho.

Para o *SVM*, irão os dados que serão treinados para a análise de sentimentos e polarização sentimental, já para o *LSTM*, irão os dados que serão treinados visando a análise textual dos mesmos, retornando 0 se não há traços de depressão detectados no documento analisado e 1 se há traços de depressão.

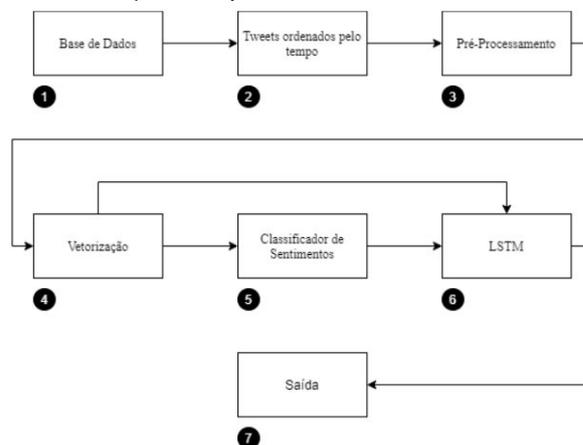


Figura 1 – Fluxograma da Metodologia

3. Experimentos

Os experimentos realizados no decorrer deste trabalho foram: o treinamento do *SVM*, a fim de realizar a classificação de polaridade sentimental; a vetorização, que tem como objetivo testar a semelhança entre documentos através da escolha de um número *n* de tweets colhidos aleatoriamente de uma lista gerada durante o experimento para calcular a semelhança delas com suas possíveis representações geradas no treinamento do *Doc2Vec*; um comparativo entre os

resultados obtidos pelo *Doc2Vec* e pelo *SVM* a fim de

Foram feitas duas execuções dos experimentos com

Tabela II – Resultados dos Experimentos de Tendência Depressiva.

	Técnicas	Tamanho da entrada	Tamanho da camada escondida	Precisão	Revocação	Acurácia	F-Measure	Limiar
Execução 1	Sentimento	2	2	0,5230	0,5242	0,5230	0,5236	0,4933
			4	0,5237	0,5197	0,5235	0,5217	0,4933
	Textual	300	300	0,8087	0,8202	0,8134	0,8144	0,5401
			600	0,8248	0,8415	0,8315	0,8331	0,5241
	Concatenado	301	301	0,8225	0,8193	0,8214	0,8209	0,9266
			602	0,8241	0,8404	0,8306	0,8322	0,4246
Execução 2	Sentimento	2	2	0,5227	0,5002	0,5231	0,5112	0,4967
			4	0,5213	0,5298	0,5231	0,5255	0,4935
	Textual	300	300	0,8162	0,8275	0,8212	0,8219	0,6234
			600	0,8265	0,8421	0,8330	0,8342	0,7241
	Concatenado	301	301	0,8189	0,8213	0,8199	0,8201	0,8044
			602	0,8287	0,8431	0,8346	0,8358	0,6092

inferir se a detecção de depressão está mais relacionada com a variação constante de emoções ou com a mensagem transmitida pelo texto e, por fim, avaliar o impacto da utilização de *stemming* no resultado obtido pelo modelo.

4. Resultados e Discussão

Durante este trabalho, 3 resultados diferentes foram alcançados, sendo eles referentes à análise de sentimentos do *SVM*, à vetorização do *Doc2Vec* e à análise de tendência depressiva do *LSTM*.

4.1. Análise de Sentimentos

Foram realizadas diversas execuções do treinamento e da classificação do *SVM*, visando encontrar o melhor resultado e também foram utilizadas duas versões da base: uma com *stemming* e a outra sem.

Levando em consideração a métrica *F-Measure*, o melhor resultado obtido foi utilizando a base com *stemming* e 300 dimensões no qual obteve-se 70,11% de precisão, 78,35% de revocação, 72,48% de acurácia e 74,00% de *F-Measure*, conforme pode ser observado na Tabela I.

Base de Dados	Dimensões	Precisão	Revocação	Acurácia	F-Measure
Com Stemming	25	0,6096	0,6883	0,6224	0,6465
	50	0,6309	0,7382	0,6546	0,6804
	75	0,6560	0,7572	0,6801	0,7030
	100	0,6695	0,7565	0,6908	0,7104
	200	0,6875	0,7811	0,7132	0,7314
	300	0,7011	0,7835	0,7248	0,7400
Sem Stemming	25	0,5718	0,7239	0,5904	0,6389
	50	0,6259	0,7235	0,6451	0,6711
	75	0,6306	0,7480	0,6563	0,6843
	100	0,6581	0,7664	0,6841	0,7082
	200	0,6801	0,7687	0,7033	0,7217
	300	0,6960	0,7777	0,7183	0,7346

Tabela I – Resultados da Classificação do *SVM*.

4.2. Vetorização do Texto

Foram realizados 7 experimentos, cada um com uma porcentagem crescente da base textual, indo de 1000 até o conteúdo da base inteira. Após os experimentos, as inferências realizadas no modelo gerado pelo *Doc2Vec* atingiram uma média de 95,81% de similaridade entre as palavras recém vetorizadas e as anteriormente vetorizadas e presentes no modelo.

4.3. Classificação de Tendência Depressiva

Para realizar a classificação de tendência depressiva, o *LSTM* foi executado em três estratégias diferentes, sendo elas: classificação sentimental dos textos feita pelo *SVM*; o conteúdo textual vetorizado pelo *Doc2Vec* (textual); o conteúdo textual vetorizado pelo *Doc2Vec* e concatenado com a classificação sentimental do *SVM* (concatenado).

cada variação da base de dados aleatorizada tanto na fase de treino quanto de avaliação.

Conforme pode-se observar na Tabela II, o resultado foi de 82,87% de precisão, 84,31% de revocação, 83,46% de acurácia e 83,58% de *F-Measure*.

5. Conclusões

Esse trabalho teve como objetivo desenvolver um modelo computacional baseado em processamento de linguagem natural para classificação de tendências depressivas em textos do *Twitter*. Para isso, foram utilizadas 3 abordagens diferentes: classificação pela análise de sentimentos contidos nos textos; classificação pela análise do conteúdo textual; classificação utilizando ambas abordagens citadas de forma conjunta.

Os resultados mostraram que foi possível realizar a detecção de tendência depressiva utilizando-se do conteúdo textual, e textual adjunto da classificação de sentimentos.

Além disso, um importante achado foi que modelos de representação de documentos foram mais relevantes para a classificação do que informações sobre análise de sentimento, já que quando utilizado somente a informação de sentimento, o *F-Measure* obtido foi próximo de 50%. Isso indica que o sentimento contido no texto não é uma informação discriminante para a classificação de tendências depressivas, e sim todo o conteúdo textual. Com isso, é notável que os dados concatenados e exclusivamente textuais obtiveram resultados semelhantes, tornando assim irrelevante a utilização da classificação sentimental junto com conteúdo textual na análise.

6. Referências

- [1] WHO, Depression, 2020. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/depression>
- [2] Y. Wang, A neural network approach to early risk detection of depression and anorexia on social media text, CLEF, 2018
- [3] B. Primack, Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among u.s. young adults, Computers in Human Behavior, 2017

Agradecimentos

Ao Centro Universitário FEI pelo suporte e a todo o corpo docente do curso de Ciência da Computação.