Sentiment Analysis of COVID-19 Tweets Using Deep Learning and Lexicon-Based Approaches

Team no. 33

Name: Razin Sumyta Monsoor

ID: 20101529

ST: Ehsanur Rahman Rhythm

RA: Humaion Kabir Mehedi

Introduction:

This paper delves into an examination of the sentiments regarding the COVID-19 pandemic and the vaccination campaign, utilizing text messages shared on the Twitter platform using sentiment analysis.

- Social media usage & Twitter
- Sentiment analysis/ opinion mining
- Deep learning based models- i) Bidirectional long/short-term memory (Bi-LSTM)

  ii)Gated Recurrent Unit (GRU) networks

- Lexicon analysers- i)National Research Council of Canada Emotion Lexicon (NRCLex)

  ii)Valence Aware Dictionary for Sentiment Reasoning (VEDAR)

IMPORTANCE:

- Understanding people perspective regarding COVID-19 vaccines
- Taking initiatives through acknowledging misunderstandings.

A large dataset of around 1,000,000 tweets from India related to COVID-19 and vaccination was used as a case study along with datas from some other countries.

- Two used approaches- i) lexicon-based

ii) deep learning methods.

- classification accuracy obtained is higher than the results of existing works.
- Analyze people's emotion into positive, negative or neutral sentiments.

Literature Review

Twitter sentiments:

i) lexicon-based approach -   # Latent Dirichlet Allocation (LDA)(1st phase)

VADER (2nd phase)   #NRCLex

ii) machine learning techniques- #Support Vector Machine (SVM)

#Bayesian classifier   #Entropy classifier

iii) deep learning- negative, positive, neutral

iv) hybrid methods-   #LDA model   #Bi-LSTM (evaluating sentiment score)

v)Neuro Symbolic based on Artificial Intelligence (AI)

Proposed Methodology with Relevant Case Studies

- Data Collection and Preprocessing

Table 2. Overview of the COVID-19 dataset.

User Name	Screen Name	Location	Tweet At	Original Tweet	Sentiment
16	44968	Bangaluru	4/3/2020	#AirSewa	Extremely Positive
24	44976	Chandigarh	6/3/2020	Sellers are	Extremely Positive
1838	46790	Bidar, India	13/3/2020	Don't Panic, Take care	Extremely Positive
1935	46887	Bangaluru, India	13/3/2020	For More Details---	Neutral
2280	47232	Bangaluru, India	14/3/2020	#nifty50 the	Positive
2736	47688	Chandigarh, Indian	14/3/2020	We should stock up on food in case cities we live in shutdown supermarkets because of this damn #Coronavirus? ehhh	Negative
3425	48377	Bangaluru, India	16/3/2020	Babu Don't Think	Extremely Positive
3430	48382	Bangaluru, India	16/3/2020	Some online Shop	Positive

Sentiment Analysis Using Lexicon Based Approaches

- Depends on sentiment dictionary
- Semantic orientation (SO)
- Sentiment Analysis with VADER9 (labelled with valence score for sentiments; -4,+4,0)
- Emotion Effects of Vaccine Tweets with NRCLex (based on NLTK library, NRC Canada affect lexicon)

Table 4. Words and their Valence Score.

Word	Polarity	Valence/Intensity
Okay	positive	0.9
Good	positive	1.9
Great	positive	3.1
Horrible	negative	−2.5
Frowning emoticon :(	negative	−2.2
Sucks	negative	−1.5

Sentiment Analysis Using Deep Learning Approach

- Recurrent Neural Network (RNN)

- two problems with RNN: 1. Vanishing gradient   2. Exploding gradient.

- two variants of RNN: 1.LSTM [60]   2.GRU

- $f(x)=\tanh(x)=(2/\ 1+e^{\wedge}-2x)-1$

- $\tanh(x)=2 sigmoid(2x)-1$

- Model Evaluation (accuracy, specificity, sensitivity)

- Results and Discussion

Conclusion

- Sentiment analysis performed using - i)lexicon-based techniques using the tools: VADER and NRCLex

(ii) deep learning methods such as Bi-LSTM and GRU.

- tweets were classified into positive, negative, and neutral categories
- sentiment scores and different emotional effects of vaccination tweets were calculated
- classification accuracy achieved was 92.7%, and with GRU, the accuracy was 91.24% for COVID-19 tweets.
- Vaccination tweets- accuracy obtained was 92.48% with Bi-LSTM and with GRU modeL it's 93.03%

THANK YOU !!!