Documents

Puertas, E., Martinez-Santos, J.C., Andres Pertuz-Duran, P.

Presidential preferences in Colombia through Sentiment Analysis

Universidad Tecnológica de Bolívar, Faculty of Engineering, Cartagena, Colombia

Abstract
This work carries out the sentiment analysis of the social network Twitter regarding the presidential debate on May 23, where a hashtag was left open so viewers could give their points of view on these three candidates: Gustavo Petro, Federico Gutierrez, and Rodolfo Hernández. Once we extracted these Tweets contained in the hashtag, they were manually classified. They then went through all the pre-processing and elimination of special characters, links, URLs, images, or videos. Next, the TextVectorization layer from the TensorFlow library was used to convert these tweets to vectors and finally to go through the two models. The results show the best results for the BERT model with an accuracy of 76% and an F1 score of 85%. © 2022 IEEE.

Index Keywords
Classifieds, Colombia, Hashtags, Language model, Language processing, Natural language processing, Natural languages, Presidential debate, Sentiment analysis, Sentimental analyse; Sentiment analysis

References

- Alva-Segura, D.A.
  (2021) *Análisis Del Sentimiento Político en Twitter Durante Las Elecciones Congresales 2020 en El Perú*,
  Master's thesis

- Cuervo, M.C., Guerrero, M.A.V.
  Predicción electoral usando un modelo híbrido basado en análisis sentimental: Elecciones presidenciales de Colombia
  (2019) *Revista Politécnica*, 15 (30), pp. 94-104.

- Andriot, F, Park, Gudivada
  Sentiment analysis of democratic presidential primaries debate tweets using machine learning models
  (2020) *Springer*, 11 (55), pp. 339-349.

- Boiy, E., Hens, P., Deschacht, K., Moens, M.-F.
  Automatic sentiment analysis in on-line text
  (2007) *Elpub*, pp. 349-360.

- Naiknaware, B.R., Kawathekar, S.S.
  (2018) *Prediction of 2019 Indian Election Using Sentiment Analysis.*, pp. 660-665.

- Abdullah, M., Hadzikadic, M.
  (2017) *Sentiment Analysis of Twitter Data: Emotions Revealed Regarding Donald Trump during the 2015-16 Primary Debates.*, pp. 760-764.

- Puertas, E., Moreno-Sandoval, L.G., Redondo, J., Alvarado-Valencia, J.A., Pomares-Quimbaya, A.
  Detection of sociolinguistic features in digital social networks for the detection of communities
  (2021) *Cognitive Computation*, 13 (2), pp. 518-537.

- Moreno-Sandoval, L.G., Del Castillo, E.A.P., Quimbaya, A.P., Alvarado-Valencia, J.A.
  Assembly of polarity, emotion and user statistics for detection of fake profiles
  (2020) *CLEF (Working Notes)*,

- Puertas, E., Martinez-Santos, J.C.
  (2021) *Phonetic Detection for Hate Speech Spreaders on Twitter*,

- Smagulova, K., James, A.P.
  A survey on lstm memristive neural network architectures and applications
  (2019) *The European Physical Journal Special Topics*, 228 (10), pp. 2313-2324.

- Chicco, D., Jurman, G.
  The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation
  (2020) *Bmc Genomics*, 21 (1), pp. 1-13.

- Zou, Q., Xie, S., Lin, Z., Wu, M., Ju, Y.
  Finding the best classification threshold in imbalanced classification
  (2016) *Big Data Research*, 5, pp. 2-8.