

Aplicación de modelos de lenguaje para la identificación de emociones presentes en twitter durante el periodo de elecciones presidenciales en Colombia 2022

Tesis presentada para optar por el título de

**Magister en Explotación de Datos y
Descubrimiento del Conocimiento**

por

Juan Jose Iguaran Fernandez



Universidad de Buenos Aires

**Facultad de Ciencias Exactas y Naturales
Departamento de Ciencias de la Computación**

[Insert Month and Year]

AGRADECIMIENTOS

Agradezco

Resumen

Dentro del texto existe información objetiva, como hechos verificables e información subjetiva, que corresponde a los procesos internos que los individuos experimentan y son plasmados en el texto, tal como las opiniones. Las emociones son parte de esta información subjetiva, y su clasificación en términos generales ha sido definida en seis emociones básicas: miedo, rabia, tristeza, alegría, sorpresa y disgusto. La detección de las emociones presentes en el texto es un sub campo del análisis de sentimiento en el texto que busca determinar la polaridad y el grado de las distintas dimensiones de la subjetividad presentes en el texto. El estudio del análisis de sentimientos en general y de emociones en particular se ha hecho usualmente a través tradicionales de NLP tales como el empleo de modelos de aprendizaje supervisado a partir de features construidos a partir del texto. Durante los últimos años, estas técnicas están siendo remplazadas por modelos de lenguajes usando redes neuronales, en particular arquitecturas como los transformers debido a su mejor desempeño y robustez. Un medio particularmente interesante para la detección de emociones son las redes sociales pues son capaces de captar a una gran cantidad de usuarios sobre una gran numero de tópicos en una cantidad limitada de palabras con estilo propio. Esto ha sido estudiado en el pasado en por ejemplo, como los análisis de texto provenientes de estas, coincide con lo que arrojan otros modelos de las realidades sociales tales como las encuestas de opinión. En español existen pocos casos de detección de emociones en redes sociales, y no se conoce de ninguno que use modelos de lenguaje basado en redes neuronales para este fin en un contexto político. EL presente trabajo tiene por objetivo el empleo de modelos de lenguaje, específicamente BERT que es un una red neuronal pre entrenada con la wikipedia basada en transformers para detectar emociones presentes en twitter durante las elecciones en Colombia.

Palabras Clave: [aquí van]

Índice general

Índice de figuras	5
Índice de cuadros	6
1. Introducción	1
1.1. Motivación	1
1.2. Marco Teórico	1
2. Metodología	6
3. Datos	7
Bibliografía	8

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

1.1. Motivación

Este trabajo es importante por que

1.2. Marco Teórico

En [Ekman, 1993] Ekman realiza un estudio de la respuesta fisiológica en general y de las expresiones faciales del ser humano en diferentes culturas ante distintas circunstancias. Esto lo lleva a concluir que existen grandes grupos en donde las distintas expresiones faciales pueden ser agrupadas ya que estas reflejan el estado emocional interno de los individuos. A estos grupos los denomino emociones básicas y son los siguientes: Alegría, enojo, sorpresa, asco, miedo y tristeza. Este modelo

de emociones básicas es comúnmente usado para los estudios relacionados con las emociones.

En el libro [Picard, 2000] Picard expone su punto de vista en donde parte de recientes investigaciones en psicología cognitiva que exponen las emociones como un componente fundamental de la inteligencia humana, por lo que la búsqueda de una inteligencia artificial capaz de interactuar eficazmente con los seres humanos, debe traer consigo la capacidad de reconocer, entender, tener y expresar emociones. A partir de ahí expone los avances hechos hasta la fecha en el terreno de lo que ella llama, computación afectiva.

En [Ortony et al., 1987] se habla de como a pesar de que las emociones son procesos mentales internos que no residen en el lenguaje, este es el medio no fenomenológico mas conveniente a través del cual podemos acceder a ellas. Elabora entonces una serie condiciones que deben estar presentes en los términos para poder referirse de una manera acertada a los estados emocionales, y las aplica sobre términos presentes en la literatura con respecto a las emociones, construyendo así un léxico emocional.

En [Hatzivassiloglou and McKeown, 1997] a partir de un corpus extenso de adjetivos que vienen en parejas con usando distintos conectores, y un etiquetado manual de algunos de ellos, se establece un algoritmo para determinar lo que los autores denominan la orientación semántica de los mismos, esto es determinar si determinado adjetivo tiene una connotación negativa o positiva de la característica que describe. Este método permite la creación automática de un corpus extenso de adjetivos cuyo sentimiento se encuentra identificado.

En [Strapparava et al., 2004] se realiza una anotación manual de estados emocionales basados en las categorías de [Ortony et al., 1987] sobre algunos términos

encontrados en WordNet [Miller, 1995], que es una base de datos de términos en ingles agrupados por grupos de sinónimos y con relación semántica entre grupos. A partir de ahí se establece la categoría emocional de nuevos términos gracias a los sinónimos y las relaciones semánticas, construyendo así una base de datos de estados emocionales llamada WordNet-Affect

En [Wiebe, 1994] se pone de manifiesto que en un tipo particular de texto, la ficción, la narración puede ocurrir desde un punto de vista objetivo, es decir, una descripción de hechos comprobables, y también desde un punto de vista subjetivo, es decir, poniendo de manifiesto los hechos atravesados por los estados mentales internos el narrador, y que la distinción entre un tipo de narración y otra no esta siempre clara, por lo que propone un algoritmo capaz de hacer esta distinción de manera automática.

En [Yu and Hatzivassiloglou, 2003] se plantea la necesidad de sistemas algoritmos capaces de discernir opiniones de hechos con el objetivo de lograr sistemas capaces de responde preguntas. Esta necesidad se plante a nivel de documentos, por ejemplo diferenciar editoriales de artículos de noticias, así como a nivel de frases. se plantean distintos modelos de aprendizaje supervisado para estas tareas en particular.

En [Pang et al., 2002] se establece la importancia de desarrollar, ante textos que reflejen opiniones subjetivas, sistemas capaces de identificar si dicha opinión es negativa o positiva. Para este propósito, se emplea el dominio de las reviews online de películas, construyendo algoritmos de aprendizaje supervisado que utiliza como features principalmente unigramas.

En [Turney, 2002] se pretende determinar a traves de aprendizaje no supervisado, si determinada review sobre diversos temas online, presenta un sentimiento

negativo o positivo. Para ello, emplea el concepto de orientación semántica presente en [Hatzivassiloglou and McKeown, 1997] para determinar si una frase tiene orientación negativa o positiva para luego determinar si la review en su conjunto es positiva o negativa.

En [Wiebe et al., 2005] se elabora una anotación manual de los estados emocionales presentes en las oraciones de un gran volumen de noticias, en donde se tiene en cuenta el contexto. El objetivo de esta anotación es crear un corpus etiquetado lo suficientemente grande para generar avances en el campo de la detección de emociones

En [Alm et al., 2005] se utilizan los cuentos infantiles para desarrollar un modelo de aprendizaje supervisado capaz de detectar emociones en las frases del texto. Para ello se elabora una anotación manual de las frases que constituyen el set de datos y luego, se generan un set de features para estas que pasaran a entrenar un clasificador lineal.

En [Aman and Szpakowicz, 2007] se utilizan texto proveniente de blogs para realizar detección de emociones presentes en las oraciones de estos. Para ello recurren primero a una anotación manual de las mismas y luego a la construcción de features para entrenar distintos modelos supervisados.

Luego, en [Pang et al., 2008] se muestra como los foros de Internet son una fuente de información de la cual se puede extraer valiosa información, entre esos detectar emociones.

En [Pak and Paroubek, 2010], [Kouloumpis et al., 2011] y en [Go et al., 2009] se aprecia como twitter puede ser usado como fuente para identificar sentimientos positivos, negativos y neutros.

En [O'Connor et al., 2010] se muestra como los sentimientos encontrados en twitter corresponden con resultados de encuestas de opinión.

En [Davidov et al., 2010] se utilizan los hashtags y los emoticones para la clasificación

En , [Wang et al., 2012]y en [Roberts et al., 2012] se plantea la clasificación mediante distintos algoritmos de las emociones en los tweets.

En [Bollen et al., 2011] se observa la relación entre los eventos sociales, políticos y económicos y las emociones detectadas.

En [Tumasjan et al., 2010] se realiza un análisis de sentimientos durante una campaña política.

En [Cerón-Guzmán and León-Guzmán, 2016] explora el análisis de sentimiento en twitter durante las elecciones 2014 en Colombia.

En [Hochreiter and Schmidhuber, 1997] se desarrollan las redes LSTM que son un tipo de RNN.

En [Chung et al., 2014] se utilizan las GRU que son otro tipo de RNN y superan a las LSTM

En [Vaswani et al., 2017] se desarrollan la tecnica de trasnformers que combina LSTM y GRU, de las cuales bert es un ejemplo

En [Devlin et al., 2018] se desarrolla BERT

En [Acheampong et al., 2021] se hace un recuento de el uso de transformers para detectar emociones

EN [Canete et al., 2020] se propone una aplicación de BERT para español

En [Gonzalez et al., 2021], [Huang et al., 2019] se utiliza bert en twitter para detectar emociones

En [Plaza-del Arco et al., 2020], [Gil et al., 2013] se hace una clasificación de emociones en español.

En [Sidorov et al., 2012] propone un léxico de palabras en español asociadas a emociones ‘

Capítulo 2

Metodología

Capítulo 3

Datos

Bibliografía

- [Acheampong et al., 2021] Acheampong, F. A., Nunoo-Mensah, H., and Chen, W. (2021). Transformer models for text-based emotion detection: a review of bert-based approaches. *Artificial Intelligence Review*, 54(8):5789–5829.
- [Alm et al., 2005] Alm, C. O., Roth, D., and Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 579–586.
- [Aman and Szpakowicz, 2007] Aman, S. and Szpakowicz, S. (2007). Identifying expressions of emotion in text. In *International Conference on Text, Speech and Dialogue*, pages 196–205. Springer.
- [Bollen et al., 2011] Bollen, J., Mao, H., and Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 450–453.

- [Canete et al., 2020] Canete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., and Pérez, J. (2020). Spanish pre-trained bert model and evaluation data. *Pml4dc at iclr*, 2020:1–10.
- [Cerón-Guzmán and León-Guzmán, 2016] Cerón-Guzmán, J. A. and León-Guzmán, E. (2016). A sentiment analysis system of spanish tweets and its application in colombia 2014 presidential election. In *2016 IEEE international conferences on big data and cloud computing (BDCloud), social computing and networking (socialcom), sustainable computing and communications (sustaincom)(BDCloud-socialcom-sustaincom)*, pages 250–257. IEEE.
- [Chung et al., 2014] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- [Davidov et al., 2010] Davidov, D., Tsur, O., and Rappoport, A. (2010). Enhanced sentiment learning using twitter hashtags and smileys. In *Coling 2010: Posters*, pages 241–249.
- [Devlin et al., 2018] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [Ekman, 1993] Ekman, P. (1993). Facial expression and emotion. *American psychologist*, 48(4):384.
- [Gil et al., 2013] Gil, G. B., Jesús, A. B. d., and López, J. M. M. (2013). Combining machine learning techniques and natural language processing to infer emotions

- using spanish twitter corpus. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 149–157. Springer.
- [Go et al., 2009] Go, A., Bhayani, R., and Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N project report, Stanford*, 1(12):2009.
- [Gonzalez et al., 2021] Gonzalez, J. A., Hurtado, L.-F., and Pla, F. (2021). Twilbert: Pre-trained deep bidirectional transformers for spanish twitter. *Neurocomputing*, 426:58–69.
- [Hatzivassiloglou and McKeown, 1997] Hatzivassiloglou, V. and McKeown, K. (1997). Predicting the semantic orientation of adjectives. In *35th annual meeting of the association for computational linguistics and 8th conference of the european chapter of the association for computational linguistics*, pages 174–181.
- [Hochreiter and Schmidhuber, 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- [Huang et al., 2019] Huang, C., Trabelsi, A., and Zaïane, O. R. (2019). Ana at semeval-2019 task 3: Contextual emotion detection in conversations through hierarchical lstms and bert. *arXiv preprint arXiv:1904.00132*.
- [Kouloumpis et al., 2011] Kouloumpis, E., Wilson, T., and Moore, J. (2011). Twitter sentiment analysis: The good the bad and the omg! In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 538–541.
- [Miller, 1995] Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.
- [Mohammad, 2012] Mohammad, S. (2012). # emotional tweets. In ** SEM 2012: The First Joint Conference on Lexical and Computational Semantics–Volume 1:*

- Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255.
- [O’Connor et al., 2010] O’Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *Fourth international AAAI conference on weblogs and social media*.
- [Ortony et al., 1987] Ortony, A., Clore, G. L., and Foss, M. A. (1987). hatzivassiloglou1997predicting. *Cognitive science*, 11(3):341–364.
- [Pak and Paroubek, 2010] Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*.
- [Pang et al., 2008] Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, 2(1–2):1–135.
- [Pang et al., 2002] Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.
- [Picard, 2000] Picard, R. W. (2000). *Affective computing*. MIT press.
- [Plaza-del Arco et al., 2020] Plaza-del Arco, F. M., Martín-Valdivia, M. T., Ureña-López, L. A., and Mitkov, R. (2020). Improved emotion recognition in spanish social media through incorporation of lexical knowledge. *Future Generation Computer Systems*, 110:1000–1008.
- [Roberts et al., 2012] Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., and Harabagiu, S. (2012). Empatweet: Annotating and detecting emotions on twitter.

In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 3806–3813.

- [Sidorov et al., 2012] Sidorov, G., Miranda-Jiménez, S., Viveros-Jiménez, F., Gelbukh, A., Castro-Sánchez, N., Velásquez, F., Díaz-Rangel, I., Suárez-Guerra, S., Trevino, A., and Gordon, J. (2012). Empirical study of machine learning based approach for opinion mining in tweets. In *Mexican international conference on Artificial intelligence*, pages 1–14. Springer.
- [Strapparava et al., 2004] Strapparava, C., Valitutti, A., et al. (2004). Wordnet affect: an affective extension of wordnet. In *Lrec*, volume 4, page 40. Lisbon, Portugal.
- [Tumasjan et al., 2010] Tumasjan, A., Sprenger, T., Sandner, P., and Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 4, pages 178–185.
- [Turney, 2002] Turney, P. D. (2002). Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*.
- [Vaswani et al., 2017] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [Wang et al., 2012] Wang, W., Chen, L., Thirunarayan, K., and Sheth, A. P. (2012). Harnessing twitter”big data”for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pages 587–592. IEEE.

- [Wiebe et al., 2005] Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2):165–210.
- [Wiebe, 1994] Wiebe, J. M. (1994). Tracking point of view in narrative. *arXiv preprint cmp-lg/9407019*.
- [Yu and Hatzivassiloglou, 2003] Yu, H. and Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 129–136.