

# 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. Datos	9
3. Metodología	10
Bibliografía	11

# 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 inglés 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 del narrador, y que la distinción entre un tipo de narración y otra no está 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 responder preguntas. Esta necesidad se plantea 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 través 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 hace manifiesto la importancia que ha venido ganando el campo del análisis de sentimiento debido al auge del Internet, tanto para usuarios individuales como para la industria de la publicidad, el mercado financiero y la academia, por lo que se hace un recuento de las distintas técnicas y aplicaciones que son consideradas relevantes por los autores hasta la fecha.

En [Pak and Paroubek, 2010] se identifica twitter como una plataforma útil para

extraer texto de diferentes usuarios sobre distintos temas por lo que se propone realizar un análisis de sentimientos sobre la misma. Para ello se procede a la identificación de tweets que contengan emoticones felices y tristes, así como tweets provenientes de cuentas de medios de noticias para tweets neutrales. luego se procede a la construcción de features usando n-gramas a partir de las palabras presentes en el tweets con estos se entrenaron varios clasificadores.

En [O'Connor et al., 2010] se plantea la pregunta si existe una correlación entre el sentimiento encontrado en twitter y las encuestas de opinión. Para ello, se toma una muestra del de mil millones de tweets entre 2008 y 2009 y se toman aquellos tweets que contengan palabras claves asociadas a los temas que se están investigando. EL sentimiento se determina a partir de la proporción de palabras con asociación negativa o positiva presentes en el tema que se esta analizando en un día en particular. Los resultados dan una correlación alta entre el sentimiento encontrado a través del texto y las encuestas.

En [Davidov et al., 2010] se parte del supuesto de que los hashtags y los emoticones contienen información relevante el cuanto al sentimiento del tweet, por lo que se hace una selección de 50 hashtags que tengan una asociación fuerte al sentimiento y se entrena un modelo supervisado a partir de los tweets que contengan estos hashtags, construyendo un vector de features para cada uno. El modelo es posteriormente empleado para clasificar otros tweets y jueces humanos verifican su eficacia.

En [Wang et al., 2012] se entrena un modelo de supervisado distante de emociones presentes en twitter. Para ellos, se utilizan hashtags que contengan terminos claves provenientes de las 5 emociones basicas de propuestas por [Ekman, 1993] para realizar el llamado de la api. A partir de ahi se entrena el modelo a partir de los features

construidos para el texto de los tweets.

En [Roberts et al., 2012] se identifica la necesidad de contar con un corpus que sirva de base para la tarea de identificación de emociones en twitter. Para ello, se seleccionan 14 temas que para los autores tienen un fuerte contenido emocional y las palabras clave asociados a estos para ser usados como hashtags en las extracción. A partir de ahí, manualmente se etiquetaron los tweets con su respectiva emoción. Esto sirvió de base para entrenar un modelo de aprendizaje supervisado.

En [Bollen et al., 2011] se plantea que a través del uso de twitter, los usuarios reflejan sus estados emocionales, ya sea de una manera explícita al indicar su emoción o de una manera implícita al hablar sobre un tema de interés general. En ese contexto, se genera procede a realizar una medida del estado emocional de una muestra de tweets entre agosto y diciembre del 2008, en donde el mismo se mide a través de la similaridad presente entre las palabras de los tweets y ciertos términos claves asociados a estados emocionales. Esto permite encontrar que determinados eventos relevantes generan un impacto emocional significativo y durante un periodo de tiempo en los usuarios.

En [Tumasjan et al., 2010] se plante el uso de twitter como plataforma de medición de la sensación política durante las elecciones parlamentarias en Alemania en el 2009. Una de sus preguntas de investigación estuvo relacionada con los sentimientos que se reflejan en los tweets que mencionan a los políticos que hacen parte de las campañas y para esto, se empleo sobre el texto proveniente de twitter, un software capaz de identificar palabras claves asociadas a estados emocionales y cognitivos en el texto. El resultado fue un perfil emocional para cada político que ve de acuerdo con su discurso político.

En [Cerón-Guzmán and León-Guzmán, 2016] se pretende realizar un análisis de sentimientos sobre tweets relacionados con las elecciones presidenciales en Colombia en el 2014 y comparar los resultados de este con las encuestas de opinión. Para ellos, inicialmente obtiene tweets que tengan palabras claves y hashtags relacionados con las elecciones, luego hace un filtrado de spam y finalmente entrena un modelo supervisado con los features que genera para los tweets restantes. Los resultados no son consistentes por lo que se plantea en futuros trabajos una caracterización demográfica.

Las RNN (redes neuronales recurrentes) son un tipo de arquitectura de red neuronal cuyo uso es especial para datos secuenciales, tales como las tareas de NLP. En estas, para un ejemplo nuevo, es posible utilizar el resultado del procesamiento de un dato anterior, para el procesamiento del siguiente dato (como en una secuencia de caracteres por ejemplo), es decir la misma topología de pesos recibirá para cada ejemplo nuevo, el resultado del ejemplo anterior además del ejemplo nuevo de entrada. Sin embargo, debido a su arquitectura, la optimización de los pesos que constituyen la topología se vuelve compleja pues para el cálculo del gradiente del error, se deberá pasar tantas veces por los pesos de la red como pasos en el tiempo haya (como número de caracteres en una palabra) haciendo que en secuencias particularmente largas, este cálculo crezca o disminuya en demasía. Para enfrentar este problema, [Hochreiter and Schmidhuber, 1997] y [Chung et al., 2014] plantean nuevas arquitecturas de RNN conocidas como LSTM (Long-Short term memory) y GRU (Gated Recurrent Unit) respectivamente, en donde a través de nuevas unidades que permiten la activación/cancelación de las señales que constituyen la red, se puede realizar la optimización de manera directa sin pasar por los pesos de la red.

En general, las arquitecturas de RNN, al tener una estructura secuencial, se impi-

de la paralelización de su cómputo, pues se necesitan las salidas de ejemplos anteriores para llevar a cabo el siguiente paso en el tiempo. Además, debido a este mismo funcionamiento secuencial, si existe un gran número de pasos en el tiempo, es poco probable que la red tenga en cuenta información presente al inicio. Para solucionar estos inconvenientes, [Vaswani et al., 2017] desarrollan los Transformers, que son un tipo de arquitectura en la que todos los datos de entrada, son ingresados a la red de manera simultánea, como todas las palabras en una frase por ejemplo, y a través de una arquitectura que definen como atención, se realiza una transformación de los datos de manera que la representación de cada dato, en este caso cada palabra, tenga en cuenta que tan importante es la misma para todas las demás palabras de la oración. Esta transformación, aplicada tanto en los datos de entrada como en los de salida, es luego usada para entrenar los pesos de la red. De esta manera la red es capaz de procesar datos en simultáneo así como tener en cuenta toda la información disponible.

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

## Datos

En [Pak and Paroubek, 2010] los autores utilizan distintos emoticones en la query de la API para traer tweets que tengan sentimientos positivos y negativos.

En [Davidov et al., 2010] se utiliza un dataset de 450 millones de tweets se hace calcula la frecuencia de los hashtags hasta encontrar los mas comunes y de estos se seleccionan aquellos que tengan una asociación mas fuerte con los sentimientos para luego traer los tweets que los contengan.

En [Wang et al., 2012] se utilizan distintos términos relacionados con emociones que serán usados como hashtags para hacer el llamado a la API. luego se hace una preseleccion de los tweets en donde se verifique el hashtag este al final, que no contenga URL, que tenga mas de 5 palabras y que no tenga mas de 3 hashtags. Además, se agregaron variaciones lingüísticas a los términos: esperanza, esperanzador. inicialmente eran 5 millones de tweets que quedaron reducidos a 2 y medio

# Capítulo 3

## Metodología

En [O'Connor et al., 2010] se utiliza una media móvil para atenuar el sentimiento percibido diariamente debido a lo ruidoso que puede ser

# 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.