

# 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

El presente trabajo .....

**Palabras Clave:** [aquí van ]

# Índice general

Índice de figuras	4
Índice de cuadros	5
<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Marco Teórico . . . . .	1
<b>2. Metodología</b>	<b>4</b>
<b>3. Datos</b>	<b>5</b>
<b>Bibliografía</b>	<b>6</b>

# 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 habla sobre las seis emociones básicas que sirven de base para el estudio.

En [Ortony et al., 1987] se hace una referencia al relacion que existe entre los estados emocionales y el lexico utilizado.

En [Strapparava et al., 2004] y [Esuli and Sebastiani, 2006] se expande este concepto para elaborar un léxico robusto asociado a emociones.

En [Wilson et al., 2009] se muestra como el contexto de una frase puede cambiar el sentimiento de una palabra en particular

En el libro [Picard, 2000] Picard da un vistazo general sobre el uso de computadoras para detectar emociones.

En [Pang et al., 2002], [Pang and Lee, 2004], [Dave et al., 2003], [Wilson et al., 2005], [Turney, 2002], [Nasukawa and Yi, 2003] y se analiza la detección de sentimiento en el texto

En [Wiebe et al., 2005][Strapparava and Mihalcea, 2008], [Strapparava and Mihalcea, 2007], [Alm et al., 2005],[Aman and Szpakowicz, 2007], [Liu et al., 2003] se puede apreciar como el texto puede ser utilizado para detectar emociones.

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 esas detectar emociones.

En [Read, 2005] se utilizan emoticones en los blogs de internet para detectar sentimiento.

En [Pak and Paroubek, 2010], [Kouloumpis et al., 2011] y en [Go et al., 2009], [Barbosa and Feng, 2010] 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 [Hasan et al., 2014], [Wang et al., 2012] y en [Roberts et al., 2012] se plantea la clasificación mediante distintos algoritmos de las emociones en los tweets.

En [Mohammad, 2012] se hace uso de los hashtags para identificar emociones y entrenar los modelos.

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.

# Capítulo 2

## Metodología

# Capítulo 3

## Datos

# Bibliografía

- [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.
- [Barbosa and Feng, 2010] Barbosa, L. and Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. In *Coling 2010: Posters*, pages 36–44.
- [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.
- [Dave et al., 2003] Dave, K., Lawrence, S., and Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews.

- In *Proceedings of the 12th international conference on World Wide Web*, pages 519–528.
- [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.
- [Ekman, 1993] Ekman, P. (1993). Facial expression and emotion. *American psychologist*, 48(4):384.
- [Esuli and Sebastiani, 2006] Esuli, A. and Sebastiani, F. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC’06)*.
- [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.
- [Hasan et al., 2014] Hasan, M., Rundensteiner, E., and Agu, E. (2014). Emotex: Detecting emotions in twitter messages.
- [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.
- [Liu et al., 2003] Liu, H., Lieberman, H., and Selker, T. (2003). A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 125–132.
- [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.

- [Nasukawa and Yi, 2003] Nasukawa, T. and Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture*, pages 70–77.
- [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). The referential structure of the affective lexicon. *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 and Lee, 2004] Pang, B. and Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *arXiv preprint cs/0409058*.
- [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.
- [Read, 2005] Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL student research workshop*, pages 43–48.
- [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.
- [Strapparava and Mihalcea, 2007] Strapparava, C. and Mihalcea, R. (2007). Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 70–74.
- [Strapparava and Mihalcea, 2008] Strapparava, C. and Mihalcea, R. (2008). Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560.
- [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 Welp, 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*.

- [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.
- [Wilson et al., 2005] Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 347–354.
- [Wilson et al., 2009] Wilson, T., Wiebe, J., and Hoffmann, P. (2009). Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational linguistics*, 35(3):399–433.