





## ANÁLISIS DE SENTIMIENTOS Y EMOCIONES SOBRE GASTRONOMÍA PERUANA USANDO MINERÍA DE TEXTO CON PYTHON

### SENTIMENT AND EMOTION ANALYSIS ABOUT PERUVIAN GASTRONOMY USING TEXT MINING WITH PYTHON

Jesús Espinola Gonzales<sup>1</sup>  Ángel Cobo Ortega<sup>2</sup>  Rocío Rocha Blanco<sup>2</sup>   
Elisa Baraibar Diez<sup>2</sup> 

<sup>1</sup>Universidad Tecnológica de los Andes, Abancay, Perú

<sup>2</sup>Universidad de Cantabria, España

#### Correspondencia:

Dr. Jesús Espinola Gonzales  
[jespinolag@utea.edu.pe](mailto:jespinolag@utea.edu.pe)

#### Como citar este artículo:

Espinola, J., Cobo, A., Rocha, R., & Baraibar, E. (2023). Análisis de sentimientos y emociones sobre gastronomía peruana usando minería de texto con Python. *Hatun Yachay Wasi, 3(1)*, 126 - 136. <https://doi.org/10.57107/hyw.v3i1.63>

#### RESUMEN

El objetivo del estudio fue mostrar el proceso de aplicación de técnicas de análisis de sentimiento, conocido también como minería de opinión (MO), para la clasificación de sentimientos y emociones en textos de opinión de manera automática. Como caso de estudio se presenta el análisis de sentimientos y emociones en reseñas, que provienen de TripAdvisor y Google, sobre la gastronomía peruana. Para la consecución del objetivo se han seguido las fases de extracción de la información (opiniones), preprocesado y clasificación; esto se complementa con el uso de librerías en Python disponibles, con especial énfasis en las que permiten el análisis sobre textos en español. Como resultado se muestra la utilidad de una librería disponible como software libre con muy buenos resultados constatados para la automatización del proceso de clasificación de reseñas. Por otro lado, se muestra que las opiniones acerca de la gastronomía peruana se clasifican como positivas con 68 % vs. 11% de reseñas clasificadas como negativas.

**Palabras clave:** Análisis de sentimiento, emociones, minería de opinión, gastronomía peruana.

#### ABSTRACT

The aim of the study was to show the process of applying sentiment analysis techniques, also known as opinion mining (OM), for the classification of feelings and emotions in opinion texts automatically. As a case study, the analysis of feelings and emotions in reviews, which come from TripAdvisor and Google, about Peruvian gastronomy is presented. To achieve the objective, the phases of information extraction (opinions), preprocessing and classification have been followed; This is complemented using available Python libraries, with special emphasis on those that allow analysis of texts in Spanish. As a result, the usefulness of a library available as free software with very good results for automating the review classification process is shown. On the other hand, it shows that opinions about Peruvian gastronomy are classified as positive with 68 % vs. 11 % of reviews classified as negative.

**Keywords:** Sentiment analysis, emotions, opinion mining, Peruvian gastronomy.



## INTRODUCCIÓN

Es bien sabido que la cantidad de información generada en Internet por usuarios de diferentes servicios es extremadamente alta. Buena parte de esa opinión corresponde a reseñas, comentarios, críticas o publicaciones que los usuarios realizan sobre una enorme variedad de servicios o productos. Esta información (opiniones) proviene de plataformas de comercio electrónico como Amazon, Booking, Netflix, TripAdvisor, Google, etc. Por otro lado, en las redes sociales, como X (antes twitter), Instagram, u otros, también se escriben cantidades ingentes de comentarios u opiniones sobre muchos temas específicos (Henriquez et al., 2016).

El análisis de estas opiniones, actualmente conocido como Minería de Opinión (MO), resulta de sumo interés con fines de marketing, o para la toma de decisiones en muchas áreas como la política (Blasco, 2020). Además, teniendo en cuenta la enorme cantidad de reseñas que pueden llegar a publicarse cada día, se hace necesario disponer de metodologías y herramientas que permitan automatizar la clasificación de los comentarios publicados. En la actual sociedad, gracias a las redes sociales, los clientes se han vuelto más poderosos, y para las empresas resulta esencial identificar a tiempo posibles valoraciones negativas o aprovechar los comentarios positivos que realicen sus clientes. Ese es precisamente el objetivo de las técnicas de MO. El análisis ha girado principalmente en torno a un análisis del sentimiento implícito en los comentarios, pero últimamente se ha añadido el análisis de emoción implícito también en el comentario. Tanto en el análisis de sentimientos como en el de emociones, el objetivo es clasificar la opinión.

Para la clasificación en el pronóstico de análisis de sentimiento, llamado polaridad, se consideran tres categorías: positivo, neutro y negativo (Pang & Lee, 2008); aunque se acompaña con una escala

de valor para cada categoría, esta escala suele estar comprendida entre -1 y 1, otros autores suelen expresar esta escala entre 0 y 1, a modo de probabilidades. Mientras que, para la clasificación de emociones se suelen considerar 6 tipos: alegría (joy), sorpresa (surprise), susto (fear), disgusto (disgust), enfado (anger), tristeza (sadness) (Ekman, 1999), y se añade otra etiqueta para casos que no encajan claramente en las categorías mencionadas, y que se identifica como otros (others).

Es importante notar que actualmente existen diferentes movimientos y acciones para que esta gastronomía sea reconocida como patrimonio inmaterial de la humanidad por parte de la UNESCO (Guardia, 2020); además del gran impacto que tiene la gastronomía peruana en el turismo del país (Moreno, 2019).

Teniendo en cuenta que estas opiniones se cuentan por millones, es impensable hacer el análisis personalmente; por esto en las últimas décadas se han desarrollado diferentes algoritmos, y consecuentemente diferentes librerías para lenguajes de programación como Python (Pérez et al., 2023), para pronosticar en automático el sentimiento y las emociones implícitas en las opiniones, también conocidas como reseñas.

Actualmente existentes diferentes librerías en Python, como software libre, para analizar los sentimientos y emociones en textos en general, aunque es importante recordar que estos textos “en bruto” deben ser previamente preprocesados para poderlos posteriormente procesar mediante aplicaciones informáticas y haciendo uso de los algoritmos respectivos. Sin embargo, también es necesario tener en cuenta que la mayoría de estas librerías están preparadas para el análisis de texto en inglés, y sólo últimamente se ha ido incrementando la disponibilidad de librerías para el análisis de texto en español (Pérez et al., 2023). El

caso de aplicación que se aborda en este trabajo recoge las reseñas en español, por lo que es muy importante la selección de la(s) librería(s) a usar.

Las reseñas con las que se han trabajado corresponden a comentarios de usuarios de restaurantes de comida peruana, concretamente de restaurantes en la provincia de Lima en Perú, incluye sus diferentes distritos. El dataset con 1258435 reseñas ha sido tomado de la página web (Kaggle, s.f.), este dataset ha sido creado con reseñas obtenidas de TripAdvisor y Googleplaces (actualmente Google).

El objetivo del estudio fue mostrar el proceso de aplicación de técnicas de minería de opinión para identificar, en el sentido de predicción, los sentimientos y emociones en textos de opinión; a su vez realizar un análisis de sentimiento y emociones en reseñas sobre la gastronomía peruana.

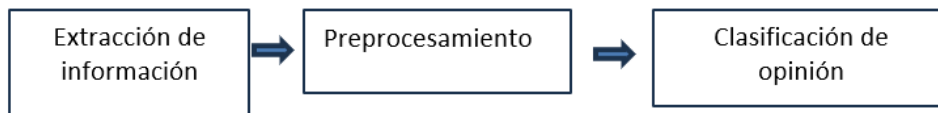
## MATERIALES Y MÉTODOS

Las reseñas originalmente están escritas en algún idioma, como español, inglés o cualquier otro idioma; es decir están escritas en un lenguaje natural, por lo que para poder clasificar los sentimientos y emociones en tales reseñas hace falta procesarlo, esto se conoce como procesamiento de lenguaje natural (Dale, 2000). Así, el análisis de sentimientos es la técnica de procesamiento de lenguaje natural (NLP) que se utiliza para identificar y evaluar las emociones, actitudes y opiniones expresadas en un texto o discurso.

El proceso para el análisis de sentimientos y emociones se puede representar por tres fases: Extracción de la información, preprocesamiento y finalmente la clasificación de opinión, como se muestra en la Figura 1.

### FIGURA 1

*A ChatGPT answer to a question regarding photosynthesis*



#### Extracción de la información

La primera tarea en el proceso del análisis de sentimiento es la recolección o extracción de la información (opiniones). Esta información puede tener diferentes orígenes, como reseñas en plataformas de venta de productos o servicios, las redes sociales, blogs, encuestas, etc.

En este trabajo los datos extraídos corresponden a reseñas (opiniones) acerca de la gastronomía peruana, escritas por usuarios de restaurantes ubicados en los distritos de la provincia de Lima,

Perú. El dataset que se usará en este trabajo, llamado Gastronomía Peruana, almacena 1258435 reseñas escritas en español y está disponible en la página web Kaggle (Kaggle, s.f.). Este dataset ha sido creado a partir de reseñas recogidas por las plataformas TripAdvisor y Googleplaces (actualmente Google) entre los años 2010 y 2021. Dado el volumen de reseñas y el amplio periodo en la recolección de la información el dataset constituye una buena representación de opinión.

### Preprocesamiento de datos

Dado las diferentes fuentes de donde se extrae la información, normalmente el texto no está estructurado, además tiene una morfología muy diversa, incluye espacios en blanco innecesarios, contienen mayúsculas y minúsculas no necesariamente correctas, tiene caracteres no deseados, etc. por lo que se hace necesario preprocesar el texto (Weiss & Indurkha, 1998).

Por lo que, el primer paso consiste en remover caracteres especiales, el segundo paso normalizar el texto en minúsculas o mayúsculas. Un tercer paso puede ser eliminar espacios en blanco que puede implicar eliminar toda la reseña porque no exista ningún comentario registrado. Un cuarto paso puede consistir en vectorizar “tokenizar” cada documento, es decir “tokenizar” cada reseña, identificando palabras, frases o expresiones.

Finalmente, un quinto paso puede consistir en remover las palabras innecesarias (stopwords) para el objetivo, que dependiendo del idioma básicamente son artículos, conjunciones, adverbios y preposiciones; este tipo de palabras realmente no aportan significado al texto, sino que actúan simplemente como elementos de conexión. Todo este proceso se puede considerar una vectorización simple. Existe una página web con un ejemplo bastante ilustrativo, con código Python incluido, de Cheng (2023) que muestra claramente el proceso; en el Ejemplo 1 se puede ver la transformación del texto de entrada y de salida que aparece en la página web mencionada (Cheng, 2023).

#### Ejemplo 1. Vectorización de texto

['The sky is blue and beautiful.' 'Love this blue and beautiful sky!'

'The quick brown fox jumps over the lazy dog.'

"A king's breakfast has sausages, ham, bacon, eggs, toast and beans"

'I love green eggs, ham, sausages and bacon!'

'The brown fox is quick and the blue dog is lazy!'

'The sky is very blue and the sky is very beautiful today'

'The dog is lazy but the brown fox is quick!']

=====  
=====

['sky blue beautiful' 'love blue beautiful sky'

'quick brown fox jumps lazy dog'

'kings breakfast sausages ham bacon eggs toast beans'

'love green eggs ham sausages bacon' 'brown fox quick blue dog lazy'

'sky blue sky beautiful today' 'dog lazy brown fox quick']

Hay un nivel más avanzado de vectorización que consiste en convertir texto en forma numérica o vectorial, de forma que pueda ser utilizada por algoritmos de aprendizaje automático. Para esto existente diferentes opciones como la Bolsa de Palabras (Bag of Words) o TF+IDF (Term Frequency-Inverse Document Frequency). En Cheng, (2023), también se muestran ejemplos para estas técnicas. Después de haber procesado el dataset GastronomíaPeruana, el número de reseñas válidas quedó en 755066; las otras fueron eliminadas debido a la ausencia de comentario de opinión, aunque sí contenían otra información como puntaje otorgado al restaurante, fechas, etc.

#### Clasificación de opinión

La tarea de clasificación consiste en, haciendo uso de algoritmos específicos, identificar la positividad, negatividad o neutralidad de las opiniones (Li et al., 2019; Liu, 2010).

Para el análisis de sentimiento existen diferentes algoritmos (Henriquez et al., 2016), a su vez diferentes librerías en Python (Insua, 2019), que permiten clasificar la polaridad (positivo, neutro o negativo).

A la hora de abordar este problema de determinar la polaridad de las opiniones existen dos tipos

de aproximaciones metodológicas (Martínez et al, 2011): a) basada en aprendizaje automático, que utiliza un conjunto de datos preclasificados para entrenar el algoritmo de clasificación y, b) basada en orientación semántica, que no necesita un entrenamiento previo, sino que analiza la orientación de las palabras del comentario u opinión. El primer enfoque puede resultar más eficaz, pero resulta más costoso y complejo debido a la necesidad de contar con un buen conjunto de datos de entrenamiento.

El segundo enfoque puede resultar más sencillo si se dispone de los recursos lingüísticos adecuados, por ejemplo, diccionarios, o léxicos catalogados con su sentimiento. Por ejemplo, el hecho de que en un comentario aparezcan palabras como “fenomenal”, “genial”, “disfrutar” parece señalar un sentimiento positivo por parte del usuario que publicó el comentario.

Para el lenguaje de programación Python existen librerías de libre acceso, tales como las muy conocidas NLTK y TextBlob, que pueden ser utilizadas para tareas de OM. Estas librerías suelen estar orientadas a textos en inglés, aunque incorporan funciones que realizan las traducciones automáticas de texto a través del traductor de Google. Pero también existen algunas librerías Python específicamente desarrolladas para el análisis de textos en español, entre ellas se podrían citar a Sentiment-analysis-spanish y Pysentimiento, esta última con una buena explicación y sustentación en (Pérez et al., 2023), es la que se usa en este trabajo para el análisis de sentimientos y emociones.

**Caso práctico de aplicación: Gastronomía peruana.** Considerando que el dataset GastronomíaPeruana preprocesado contenía 755066 reseñas válidas se determinó una muestra de trabajo para agilizar el proceso de cálculo. Fijando un nivel de confianza de 99% y un margen de error de 2 % la muestra de trabajo quedó determinada en 4138 reseñas. Por lo

que, de forma aleatoria, se seleccionó en el dataset GastronomíaPeruana las reseñas de acuerdo con el número indicado. Estas reseñas corresponden con 3784 recogidas a través de Google places (ahora Google) y 354 a través de TripAdvisor.

Con este dataset preprocesado se pasó a la fase de clasificación de opinión. Haciendo uso de la librería Pysentimiento se logró clasificar en cada reseña la polaridad de sentimiento (positivo, neutro o negativo); así como, la emoción ('joy', 'surprise', 'fear', 'disgust', 'anger', 'sadness', 'others').

## RESULTADOS Y DISCUSIÓN

### Análisis de sentimiento

Aplicando la librería Pysentimiento en Python, ver código en Figura 2, se obtuvo el resultado de clasificación para las 4138 reseñas de la muestra que se presenta en la Tabla 1, se observa que 2822 reseñas se clasifican como positivas, 874 reseñas se clasifican como neutras, y 442 reseñas se clasifican como negativas.

**TABLA 1**

*Polaridad de sentimientos*

Polaridad	$F_i$	$f_i$
Positivo	2822	0,68
Neutro	874	0,21
Negativo	442	0,11
Total	4138	1

Los resultados obtenidos con el análisis en las reseñas de la muestra se pueden extrapolar al conjunto total de reseñas válidas (Fig. 3). Es decir, se puede predecir que el 68 % de las reseñas son clasificadas como positivas, 21 % neutras, y sólo el 11 % de las reseñas son clasificadas con polaridad negativa. Esto es un buen indicador a favor de la gastronomía peruana.

**FIGURA 2**

*Código Python para análisis de sentimiento*

```
# Lectura de archivo
peruvianfoodreviews = pd.read_csv("Dataset/GastronomiaPeruana.csv")
pfr=peruvianfoodreviews

# Análisis de SENTIMIENTOS con pysentimiento
from pysentimiento import create_analyzer
sentiment_analyzer = create_analyzer(task="sentiment",lang="es")
pfr['Sentiment']=pfr['review'].apply(lambda x: sentiment_analyzer.predict(x))

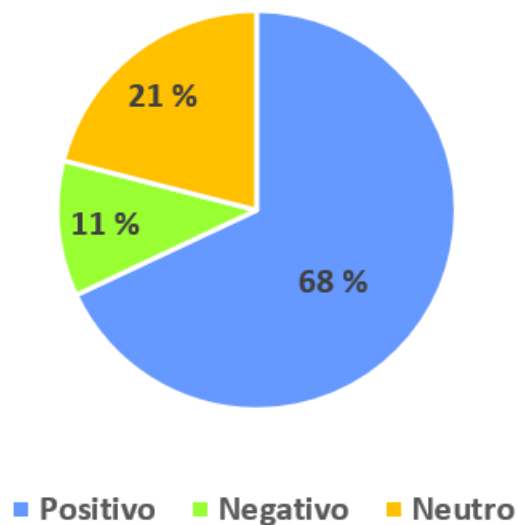
# Añade la columna Classif_Sentimet
pfr['Classif_Sentimet']=pfr['review'].apply(lambda x: sentiment_analyzer.predict(x).output)

# Añadir columnas ('POS', 'NEU', 'NEG', ...) con sus valores (tomados del objeto Sentiment)
pfr["Sentiment"]=pfr["Sentiment"].astype(str)
sentiments = ['POS', 'NEU', 'NEG']
for sentiment in sentiments:
    pfr[sentiment] = pfr['Sentiment'].str.extract(fr'{sentiment}: (\d+\.\d+)')

pfr[sentiments]=pfr[sentiments].astype(float)
```

**FIGURA 3**

*Prevalencia de polaridad de sentimientos*



### Análisis de emociones

Para el análisis de emociones también se aplicó la librería Pysentimiento, ver código respectivo en Figura 4.

Los resultados de clasificación de la muestra formada por las 4138 reseñas se presentan en la Tabla 2. En esta tabla se observa que la mayoría, 2223 reseñas,

se clasifican como de alegría; también se puede ver otro número representativo, 1718 reseñas no se pueden clasificar dentro de las 6 establecidas y por ello se les agrupa con el identificador "otros"; por lo demás, salvo 148 reseñas que son clasificadas como de enfado, las otras prácticamente no tienen una representación significativa.

**FIGURA 4***Código Python para análisis de emociones*

```

# Análisis de EMOCIONES con Pysentimiento
from pysentimiento import create_analyzer

# Se toma el Dataset con la información de sentimientos
pfr1=pfr

emotion_analyzer = create_analyzer(task="emotion",lang="es")
pfr1['Emotion']=pfr1['review'].apply(lambda x: emotion_analyzer.predict(x))
pfr1['Classif_Emotion']=pfr1['review'].apply(lambda x: emotion_analyzer.predict(x).output)

# Añadir columnas ('others', 'joy', 'surprise', 'fear', 'disgust', 'anger', 'sadness') con sus valores.
pfr1["Emotion"]=pfr1["Emotion"].astype(str)
emotions = ['others', 'joy', 'surprise', 'fear', 'disgust', 'anger', 'sadness']
for emotion in emotions:
    pfr1[emotion] = pfr1['Emotion'].str.extract(fr'{emotion}: (\d+\.\d+)')

pfr1[emotions]=pfr1[emotions].astype(float)

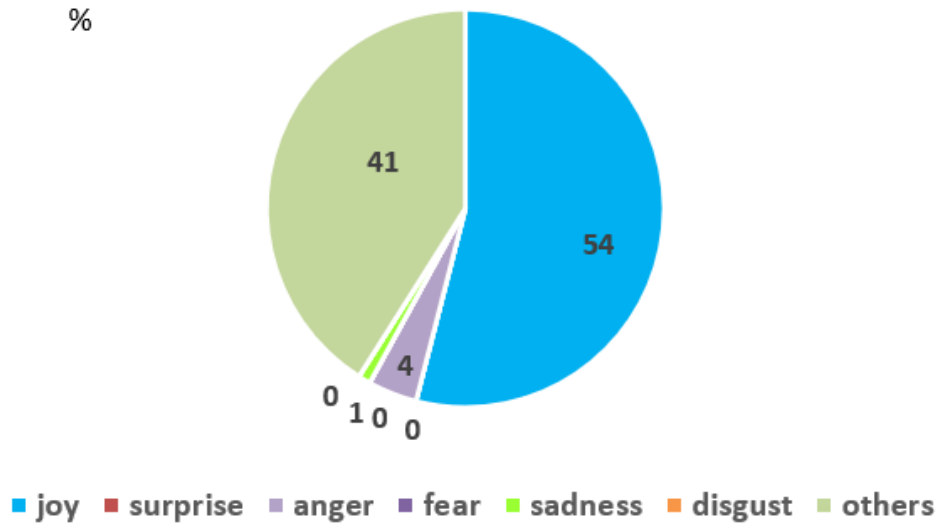
```

**TABLA 2***Clasificación de emociones*

Emociones	Emotions	$F_i$	$f_i$ (%)
Alegría	joy	2223	54
Sorpresa	surprise	4	0
Enfado	anger	148	4
Miedo	fear	3	0
Tristeza	sadness	42	1
Disgusto	disgust	0	0
Otros	others	1718	41
Total		4138	100

Como se puede observar en la Figura 5, los resultados de la muestra extrapolados al conjunto total de reseñas válidas muestran que 54 % de las reseñas se pueden clasificar como de alegría, mientras que 42 %, no encajan dentro de las emociones consideradas; Un 4 % se clasifican como

de enfado, las otras emociones no aparecen con un porcentaje significativo. El resultado de este análisis de emociones también constituye un buen indicador a favor la gastronomía peruana.

**FIGURA 5***Prevalencia de emociones***Análisis de calificación asignada**

Es habitual que los usuarios, juntamente con las reseñas, califiquen con un valor entero entre 1 y 5 a los restaurantes visitados. Por lo que, otra información importante que se ha obtenido por simple análisis es la frecuencia de puntajes asignados. En la Tabla 3 se observa que, de las

4138 reseñas consideradas en la muestra, en 2185 reseñas se calificó a la gastronomía peruana con el máximo puntaje: 5 puntos; en 1152 reseñas se calificó con 4 puntos; en 481 reseñas se calificó con 3 puntos, mientras que con valores menores a 3 puntos se calificó en un total de 320 reseñas.

**TABLA 3***Clasificación de calificaciones*

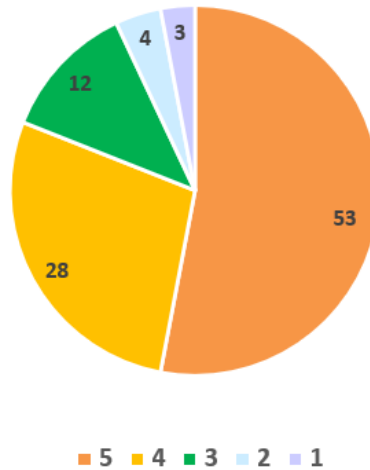
Puntajes	$F_i$	$f_i$
5	2185	0,53
4	1152	0,28
3	481	0,12
2	182	0,04
1	138	0,03
<b>Total</b>	<b>4138</b>	<b>1</b>

Extrapolando estos resultados a la totalidad de las reseñas válidas, observando la Figura 6, se puede afirmar que más del 50 % de las reseñas se ha calificado a la gastronomía peruana con una calificación de 5 puntos; por otro lado, en 28 % de las reseñas se ha dado una calificación de

4 puntos, en 12 % de las reseñas se ha calificado con 3 puntos; mientras que sólo en un 7 % se ha dado una calificación menor que 3 puntos. Esto es muy coherente con los resultados del análisis de sentimientos y emociones obtenidos.



**FIGURA 6**  
*Prevalencia de emociones*



#### Relación entre sentimiento, emociones y calificación asignada

También es importante la relación existente en los resultados del análisis de sentimiento, análisis de emociones y la calificación otorgada. Sólo a modo de ilustración se muestran algunas reseñas elegidas de manera aleatoria (Tabla 4).

En esta tabla se puede apreciar una estrecha relación entre comentarios totalmente positivos asociados con alegría como emoción, y con las calificaciones otorgadas de 5 o 4 puntos. Por otro lado, se puede

observar un comentario que expresan una mezcla de opinión entre lo positivo y negativo, esto está clasificado con una polaridad neutra, asociado con "otros" como emoción, y asociado con una calificación de 3 puntos. Finalmente, también se puede observar una reseña claramente negativa, asociada con enfado como emoción y 1 punto en la calificación otorgada por el usuario. Este simple análisis permite ver la coherencia que existe entre la calificación dada por el usuario y la predicción de sentimiento y emoción dada por el programa.

**TABLA 4**  
*Relación reseña (opinión)-polaridad-emoción-calificación*

Reseña (opinión)	Polaridad (sentimiento)	Emociones	Calificación dada
Muy buenas carnes y acompañamientos en un lugar cálido y de excelente servicio, tienen valet parking y la zona es excelente	Positivo	Alegría	5
Muy rico	Positivo	Otros	5
El lugar es muy limpio la atención buena pero algunas cosas con respecto a la comida no me gustaron	Neutro	Otros	3
La torta de chocolate muy agradable tiene el toque perfecto de dulce y acompañado de un buen café	Positivo	Alegría	4
Pésima la atención falta trato e higiene	Negativo	Enfado	1

## CONCLUSIONES

La utilidad y potencialidad de este tipo de metodologías y herramientas facilita la toma de decisiones en relación con estrategias comerciales, conocimiento de clientes, optimización de servicios y mejora de la reputación corporativa.

La implementación con Python permite automatizar el proceso de identificación de reacciones de clientes, constituyendo un buen instrumento para diseñar estrategias de marketing o políticas encaminadas a mejorar la reputación on-line de marcas, productos, servicios, o en este caso, establecimientos de restaurantes.

Las reseñas sobre la gastronomía peruana, mediante plataformas como Tripadvisor y Google reflejan un resultado positivo en 68 % de opiniones.

En las reseñas la emoción predominante fue la alegría (joy), en un 54 % de las opiniones.

El puntaje asignado por los usuarios de restaurantes de comida peruana es de 5 o 4 puntos en un 81 % de opiniones.

## REFERENCIAS BIBLIOGRÁFICAS

Blasco, M. (2020). Análisis de sentimiento de la agenda de los partidos políticos españoles en Twitter durante la Moción de Censura de 2018. Un enfoque de datos composicionales. *Revista Mediterránea de Comunicación*, 11(2), 185-198. <https://dialnet.unirioja.es/servlet/articulo?codigo=7491324>

Cheng, A. (2023). *Text Vectorization Using Traditional Methods*. [https://alvinntnu.github.io/NTNU\\_ENC2045\\_LECTURES/nlp/text-vec-traditional.html](https://alvinntnu.github.io/NTNU_ENC2045_LECTURES/nlp/text-vec-traditional.html)

Dale, R. (2000). Symbolic Approaches to Natural Language Processing. En D. Robert, H. Moisi, & H. Somers, *Handbook of Natural Language*

*Processing*. Marcel Dekker, Inc.

Ekman, P. (1999). *Basic Emotions*. En T. Dalgleish, & M. J. Power, *Handbook of Cognition and Emotion* John Wiley & Sons Ltd. <https://doi.org/10.1002/0470013494.ch3>

Guardia, S. (2020). *Gastronomía Peruana Patrimonio Cultural de la Humanidad*. Universidad de San Martín de Porres. [https://catedraunesco.usmp.edu.pe/portfolio\\_item/gastronomia-peruana/](https://catedraunesco.usmp.edu.pe/portfolio_item/gastronomia-peruana/)

Henriquez, C., Guzmán, J., & Salcedo, D. (2016). Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles. *Procesamiento del Lenguaje Natural*, 56, 25-32. <https://www.redalyc.org/pdf/5157/515754423002.pdf>

Insua Yañez, A. (2019). *Sistema Deep Learning para el análisis de sentimientos en opiniones de productos para la ordenación de resultados de un buscador semántico*. [Tesis de Licenciatura, Universidade da Coruña]. <http://hdl.handle.net/2183/25152>

Kaggle. (s.f.). *Kaggle*. <https://www.kaggle.com/>

LI, L., LI, W., & CONG, D. (2019). Naive Bayesian Automatic Classification of Railway Service Complaint Text Based on Eigenvalue Extraction. *Technical Gazette*, 26(3), 778-785. <https://doi.org/10.17559/TV-20190420161815>

Liu, B. (2010). *Sentiment Analysis and Subjectivity*. En N. Indurkha, & F. J. Damerou, *Handbook of Natural Language Processing*. Chapman and Hall/CRC. <https://www.cs.uic.edu/~liub/FBS/NLP-handbook-sentiment-analysis.pdf>

Martínez, E., Martín, M., Perea, J., & Ureña, L. (2011). Técnicas de clasificación de opiniones aplicadas a

un corpus en español. *Revista de Procesamiento del Lenguaje Natural*, 47, 163-170. <https://www.redalyc.org/pdf/5157/515751747017.pdf>

Moreno, A. (2019). *El impacto de la gastronomía peruana en el turismo y el mercado del país* [Tesis de Licenciatura, Universidad Santo Tomás]. <http://hdl.handle.net/11634/17755>

Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval*, 2(1), 1-135. <https://doi.org/10.1561/1500000011>

Perez, J., Rajngewerc, M., Giudici, J., Furman, D., Luque, F., Alemany, L., & Martínez, M. (2023). Pysentimiento: A Python Toolkit for Opinion Mining and Social NLP tasks. *arXiv preprint arXiv:2106.09462*. DOI:10.21203/rs.3.rs-3570648/v1

Weiss, S. M., & Indurkha, N. (1998). *Predictive Data Mining. A Practical Guide*. Morgan Kaufmann Publishers, Inc.