

ARTÍCULO CIENTÍFICO
CIENCIAS DE LA EDUCACIÓN

Análisis comparativo para seleccionar una herramienta de reconocimiento de emociones aplicando el modelo AHP

Comparative analysis to select an emotion recognition tool applying the AHP model

Quiroz Martínez, Miguel Ángel ^I; Guachizaca Hernández, Jahaira María ^{II}; Ortega Veintimilla, Shakir Sebastián ^{III}.

^I. mquiroz@ups.edu.ec, Carrera de Ingeniería en Sistema, Universidad Politécnica Salesiana, Guayaquil, Ecuador

^{II}. jguachizaca@est.ups.edu.ec, Carrera de Ingeniería en Sistema, Universidad Politécnica Salesiana, Guayaquil, Ecuador

^{III}. ortegav@est.ups.edu.ec, Carrera de Ingeniería en Sistema, Universidad Politécnica Salesiana, Guayaquil, Ecuador

Recibido: 23/02/2019

Aprobado: 13/06/2019

Como citar en normas APA el artículo:

Quiroz Martínez, M. A.; Guachizaca Hernández, J. M.; & Ortega Veintimilla, S. S. (2019). Análisis comparativo para seleccionar una herramienta de reconocimiento de emociones aplicando el modelo AHP. *Uniandes Episteme*, 6(3), 453-463.

RESUMEN

La importancia del comportamiento no verbal y lo que se transmite mediante el lenguaje corporal, juega un papel muy importante en nuestro entorno. Como una de las áreas éste trabajo se enfoca en el reconocimiento de emociones para niños de 2 a 4 años, debido a que es importante revisar el estado de emoción que representa un niño dependiendo del ambiente en el cual se encuentra sumergido. El objetivo de este artículo es evaluar la eficiencia de tres herramientas de reconocimiento de emociones, que son: Face++, Microsoft Azure Emotion API y Google Vision API al inferir atributos de niños. Para poder realizar esta investigación, se realizó un análisis comparativo de las herramientas utilizadas junto con la ayuda de una especialista en psicología infantil. Se tomaron como muestra un grupo de 20 niños, de la Fundación “Los chavitos” del MIES en Ecuador. Los resultados experimentales mostraron que Face++ presentó mayor precisión con relación a Microsoft Azure Emotion API y Google Vision API. Se espera que el conjunto de datos que se presenta en los resultados pueda ayudar a

allanar el camino para futuras investigaciones sobre el uso de las herramientas de reconocimiento de emociones.

PALABRAS CLAVE: Expresión facial; reconocimiento facial; emociones; detección de rostro; reconocimiento de emociones.

ABSTRACT

The importance of nonverbal behavior and what is transmitted through body language plays a very important role in our environment. As one of the areas This work focuses on the recognition of emotions of children from 2 to 4 years of age. The objective of this article is to evaluate the efficiency of the emotions recognition tools, which are Face ++, Microsoft Azure Emotion API and Google Vision API to infer the attributes of children. In order to carry out this research, a comparative analysis of the tools used was carried out. They were taken as a sample in a group of 20 children, from the "Los chavitos" Foundation of MIES in Ecuador. The experimental results benefited from Face ++ being presented more accurately in relation to Microsoft Azure Emotion API and Google Vision API. The use of emotion recognition tools.

KEYWORDS: facial expression; face recognition; emotions; face detection; facial recognition of emotions.

INTRODUCCIÓN

El campo del reconocimiento facial empieza a crecer en el mundo con una atención especial a las emociones. Las cuales son la parte desencadenante de la acción en el proceso de toma de decisiones (Matsumoto, Hwang, López, & Pérez Nieto, 2013). El enfoque es hacer uso de la inteligencia artificial para la identificación de rostros con las herramientas de software de reconocimiento de emociones (Kwak & An, 2016), Considerando la cara como un factor importante para identificar los sentimientos (Al-Hajjar & Syed, 2015). Mediante algoritmos que incluyen diferentes herramientas para análisis de rostros de personas para posiblemente descubrir la emoción que emiten (Bakshi & Singhal, 2014).

Es necesario trabajar con los sentimientos y emociones desde la infancia como una actitud que se debe tener presente en todo momento (Cebrian Chulia, 2016), las expresiones faciales mejoran la autoconciencia y la confianza del ser humano. Los niños son muy vulnerables y los dominan las emociones y los impulsos, que les permite aumentar la empatía y la colaboración entre ellos (Bakshi & Singhal, 2014).

Por medio de las expresiones faciales se logra representar el estado mental y emocional (Hatem, Beiji, & Majeed, 2015). Los sistemas automáticos de reconocimiento facial de emociones tienen el objetivo de detectar la emoción. Las expresiones faciales se clasifican en siete estados básicos: alegría, tristeza, ira, sorpresa, asco, miedo y desprecio.

La detección de emociones se basa en expresiones faciales, las herramientas o APIs utilizan la detección facial de los ojos, boca, nariz, etc., para lograr determinar el estado de ánimo (Jung, An, Kwak, Salminen & Jansen, 2017).

El reconocimiento de emociones se puede aplicar en temas relacionados con la psiquiatría, como es el comportamiento humano, así también en la robótica, al manejar la interacción entre humanos y máquinas (Ma Xiaoxi, Lin Weisi, Huang Dongyan, Dong Minghui, 2017). Uno de los puntos importantes del reconocimiento emocional es el uso en los softwares educativos para poder trabajar de una mejor manera en la atención del niño y su aprendizaje. Actualmente se está usando en los estudios de mercadeo donde se puede captar las emociones de los consumidores en tiempo real (Padhi, Palo, Mishra, & Mohanty, 2016), mientras realizan la prueba de un producto para concluir con métricas claves al observar el nivel de satisfacción de los usuarios estando en contacto con los productos. Las herramientas para el reconocimiento de emociones utilizan el análisis de las expresiones faciales (Padhi et al., 2016), que permiten detectar las emociones a partir de imagen o video, estas expresiones aparecen y desaparecen del rostro en fracción de segundos sin que el lente de la cámara pueda capturar la emoción real de la persona (Bhattacharya, 2016).

La presente investigación se basa en los trabajos que sirven de apoyo y referencia, como:

“...Revisando nuestros resultados específicos, vemos una tendencia de alta precisión para el género, con tres de las herramientas que realizan con una precisión de arriba 90 por ciento para todos los conjuntos de datos. En cuanto a la raza, solo una herramienta. ofrece esta capacidad, Face ++, y la precisión es bastante alta, por encima del 90 por ciento” (Jung, An, Kwak, Salminen, & Jansen, 2018, p.625).

Esto significa que Face++ se destaca por su tasa de precisión en evaluación de detección de raza es la más alta. Por otra parte, indica que buscó discutir la aplicación de análisis de sentimientos en Microsoft y Amazon, donde sus resultados presentan la polaridad positiva más alta en Microsoft Azure que en AWS y en cuanto al análisis de emociones sus indicadores presentan semejanza. En el caso (Qaisi & Aljarah, 2016) discuten algunos beneficios que se puede obtener con el uso de Google Cloud Vision para escenarios de tecnología asistencial con la funcionalidad del reconocimiento facial. Sin embargo, ante la gran variedad de herramientas o aplicaciones desarrolladas (Brodny, et al., 2016), nos lleva a un mundo desconocido sobre la precisión y fiabilidad de los resultados que generan las herramientas. En el presente trabajo, Utiliza Face++ para el reconocimiento de emociones, Microsoft Azure y Google Cloud Vision (Mulfari, Celesti, Fazio, Villari, & Puliafito, 2016), siendo el principal objetivo realizar una comparación con otras APIs similares y proponer cuál de las herramientas seleccionadas para esta investigación es más eficiente para reconocer las emociones en los niños, mediante el uso del proceso de análisis jerárquico AHP (Delgado & Romero, 2015). Para la obtención de resultados.

En la actualidad es un desafío optar por una alternativa cuando se emplea algún proceso de toma de decisión, debido que las diferentes alternativas por lo general tienen diferentes características o criterios con aspectos tanto cuantitativas como cualitativas, por lo que resulta difícil la comparación. Sin embargo, el modelo AHP permite tomar una decisión tomando en cuenta estos aspectos mediante herramientas matemáticas, reduciendo así la subjetividad del proceso de decisión (Herrera & Osorio, 2006), y el peso que da el experto o decisor a un criterio sobre otro, mediante escala psicométrica de Saaty (Saaty, 1980).

El objetivo de este artículo es evaluar la eficiencia de tres herramientas de reconocimiento de emociones, que son: Face++, Microsoft Azure Emotion API y Google Vision API al inferir atributos de niños

MÉTODOS

Se recibió el apoyo por parte de la fundación “Los chavitos” del Ministerio de Inclusión Económica y Social, donde se seleccionó un grupo experimental de 20 niños entre 2 a 4 años, que participaron de manera voluntaria con el consentimiento de sus padres. En total se obtuvieron 223 imágenes durante un periodo de 10 días en el mes de mayo, 2018. Las imágenes fueron capturadas en un entorno no controlado y de manera más natural mediante una cámara profesional Nikon Coolpix P520 con una resolución de 4320p x 3240p, en el instante que los niños realizaban sus actividades diarias. De las 223 imágenes, se descartaron 163, tomando los siguientes criterios: similitud de emoción, sombras, mala iluminación, nitidez y brillo. Quedando con un total de 60 imágenes, de las cuales, con ayuda de 2 expertos en el área de psicología en la niñez, éstas fueron clasificadas de la siguiente manera:

- 20 imágenes demuestran rostros de felicidad,
- 10 imágenes rostro de tristeza,
- 10 imágenes presentaron un rostro con miedo,
- 10 de ellas fueron neutrales y
- 10 con sentimiento de sorpresa.

Las herramientas utilizadas para el reconocimiento facial de emociones utilizadas fueron:

1. Face++ que reconoce emociones como: felicidad, sorpresa, tristeza, asco, ira, miedo y neutral,
2. Google Vision Api que reconoce felicidad, tristeza, ira, sorpresa y neutral, y
3. Microsoft Azure Emotion API que reconoce felicidad, tristeza, sorpresa, ira, asco, miedo, neutral y desprecio.

Las herramientas en conjunto nos ayudaron a identificar el total de emociones que nos interesaban captar, como: Felicidad, Tristeza, Miedo, Sorpresa y Neutral, que se encuentran

presente en niños con un rango de edad entre 2 a 4 años. Se realiza el estudio comparativo de la mejor herramienta mediante la aplicación del modelo AHP, siguiendo la siguiente metodología:

1. Modelo Jerárquico,
2. Definición de criterios,
3. Análisis de herramientas y
4. Matriz comparativa.

Este es el desarrollo de la estructura jerárquica AHP, para establecer la definición de los criterios comunes entre cada una de las herramientas y determinar cuál de ellas es la óptima. Por consiguiente, se analiza y evalúa cada una de las herramientas que mayor relevancia presenten y se crea la matriz de preferencias y cálculo de pesos. En nuestro análisis se contemplan tres alternativas y cuatro criterios para la comparación (Figura 1).

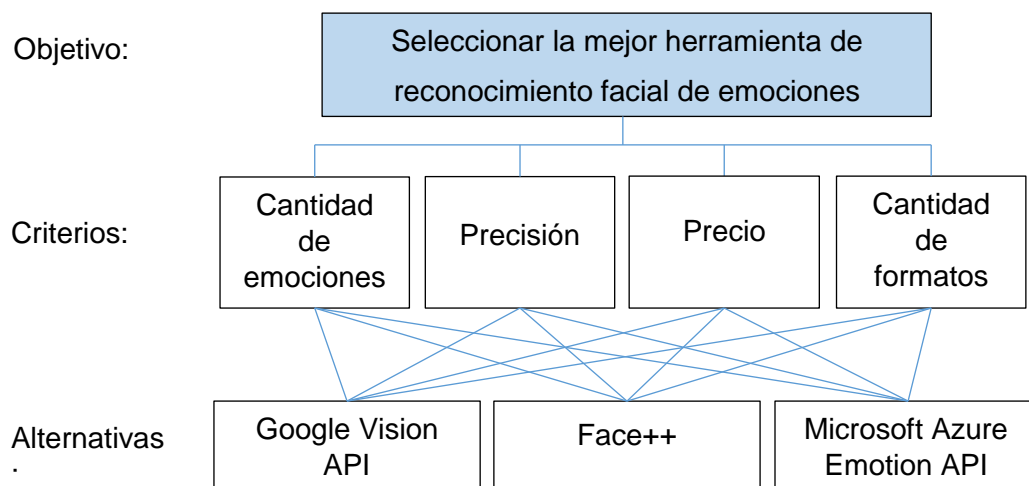


Figura 1. Nivel Jerárquico de las tres capas del modelo AHP.

Para la selección de los criterios y su peso respecto a uno del otro, se realizó un cuestionario a 5 expertos en desarrollo de software de la Universidad Politécnica Salesiana, en el cual se estableció que la precisión es el factor más importante de decisión con respecto a los otros criterios, quedando en segundo el número de emociones que pueden detectar, y como últimos con igual preferencia en ambas, el precio y la cantidad de formatos. Los datos de los criterios a excepción de la precisión, fueron obtenidas de la página oficial de cada aplicación, de los cuales se pueden observar en la Tabla 1.

Tabla 1: Datos de los criterios de decisión.

Herramientas	Cantidad de emociones reconocidas	Cantidad de formatos de imágenes	Precio
Face ++	7	2	\$1,00
Google Vision API	4	8	\$13,50
Microsoft Azure Emotion API	8	4	\$2,50

El criterio de precisión se obtuvo mediante el análisis de las imágenes de los niños por medio de los datos obtenidos de las herramientas, que se presentan en la Tabla 2.

RESULTADOS

Se analizaron las 60 imágenes por parte de 2 expertos en psicología infantil, y mediante su criterio profesional evaluaron y clasificaron las imágenes obteniendo los valores de porcentaje mostrados en la Tabla 2.

Tabla 2: Tabla de evaluación de expertos

	Felicidad	Neutral	Tristeza	Miedo	Sorpresa
Experto 1	93,38 %	84,2 %	89,88 %	80 %	100 %
Experto 2	83,75 %	82 %	85,92 %	83,33 %	85 %

Las 60 imágenes que fueron analizadas y categorizadas por emociones por parte de los expertos en psicología infantil fueron ajustadas a un formato en común de las 3 herramientas de reconocimiento de emociones, de manera que se obtengan resultados fiables, para posteriormente medir el valor de precisión individual de las emociones reconocidas por cada imagen y como resultado final obtener el valor de precisión promedio de cada emoción por herramienta, mostrados en la Tabla 3.

Tabla 3: Tabla de precisión de las herramientas

Herramientas	Felicidad	Neutral	Tristeza	Miedo	Sorpresa
Face ++	82,96%	74%	81,65%	77%	99,45%
Google Vision API	60,83%	0%	77,50%	10%	63,98%
Microsoft Azure Emotion API	79,73%	84%	90,71%	24%	90,91%

A. Aplicación de AHP en selección de la mejor herramienta para el reconocimiento de emociones.

Para la comparación de las tres herramientas de software seleccionadas se utilizó la escala psicométrica de Saaty que propone el grado de importancia de cada criterio respecto a los otros criterios como se muestra a continuación:

$$R = \begin{bmatrix} 1 & 1/7 & 5 & 5 \\ 7 & 1 & 9 & 9 \\ 1/5 & 1/9 & 1 & 1 \\ 1/5 & 1/9 & 1 & 1 \end{bmatrix}$$

Sea R una matriz de comparaciones pareadas de 4 alternativas, si r_{ij} es el peso del criterio i con respecto a la columna j , donde las columnas son representadas por el criterio de las siguientes maneras:

- [1] Cantidad de emociones reconocidas,
- [2] precisión,
- [3] Precio, y
- [4] Cantidad de formatos de imágenes.

Para comprobar la consistencia de la matriz se calculó mediante la fórmula de Razón de Consistencia (RC) de Saaty, siendo una matriz de dimensión 4, el Índice de Consistencia Aleatorio (IA) es 0.882 y el Índice de Consistencia (IC) igual a 0.082, por lo tanto, la Razón de Consistencia es igual a 0.093. Comprobando así, que la Matriz R es consistente mediante la Tabla 4, la cual nos indica que para una matriz de dimensión 4, la Razón de Consistencia debe ser menor a 9 %.

Tabla 4: Limites de consistencia

Tamaño de la matriz (n)	Ratio de Consistencia
3	5 %
4	9 %
5 o mayor	10 %

Fuente: (Saaty, 1980)

Se consideró la matriz de prioridad de criterio en término de la meta global definida, donde W es la matriz de prioridad w_i es la prioridad del criterio i con respecto a la meta global.

$$W = \begin{bmatrix} 0,212 \\ 0,673 \\ 0,058 \\ 0,058 \end{bmatrix}$$

En la tabla 5 se define la matriz de valoración de las alternativas para cada criterio. Las cuales se presentan de la siguiente manera:

- [A1] Face ++,
- [A2] Google Vision API, y
- [A3] Microsoft Azure Emotion API.

Y los criterios

- [C1] Cantidad de emociones,
- [C2] Precisión,
- [C3] Precio y
- [C4] Cantidad de formatos de imágenes.

Teniendo en cuenta que los valores de C1, C3 y C4 fueron obtenidas de los respectivos sitios web de las herramientas, y C2 fue obtenido como valores promedio de precisión de cada herramienta contemplando las 5 emociones reconocidas.

Tabla 5. Matriz de valoración de alternativas para cada criterio

Alternativas	C1	C2	C3	C4
A1	7	83,01%	\$ 1.00	2
A2	4	42,46%	\$ 13,50	8
A3	8	73,87%	\$ 2,50	4

La normalización de la matriz de valoración de alternativas M se multiplicó por la matriz prioridad de criterio W quedando de la siguiente manera:

$$M_{norm} = \begin{bmatrix} 0,875 & 1 & 0,926 & 0,25 \\ 0,5 & 0,512 & 0 & 1 \\ 1 & 0,889 & 10,815 & 0,51 \end{bmatrix} \times \begin{bmatrix} 0,212 \\ 0,673 \\ 0,058 \\ 0,058 \end{bmatrix}$$

A continuación, se observa la matriz resultado R , donde las alternativas se exponen de la siguiente manera:

[Fila 1] Face ++,

[Fila 2] Google Vision API, y

[Fila 3] Microsoft Azure Emotion API.

$$R = \begin{bmatrix} 0,926 \\ 0,508 \\ 0,887 \end{bmatrix}$$

En base a la matriz resultante obtenida a partir de los pesos ponderados entre las alternativas y criterios mencionados, se puede observar que la mejor opción para el reconocimiento facial de emociones en niños de 2 a 4 años es la herramienta Face ++ con un 92,6% de elección, seguida de Microsoft Azure Emotion API con 88,7% y por último con un bajo índice la herramienta de Google Vision API con un 50.8%. Cabe indicar que la diferencia entre las herramientas Face++ Y Microsoft Azure Emotion API, es mínima.

DISCUSIÓN

Desde el análisis realizado en los resultados, podemos confirmar que se muestra una tendencia positiva para la herramienta Face++.

Jung, et al. (2018) indicaron en su trabajo: En cuanto a la raza, solo una herramienta. ofrece esta capacidad, Face ++, y la precisión es bastante alta, por encima del 90 por ciento. Sin embargo, esto fue evaluado usando un conjunto de datos de imágenes de alta calidad. Se necesita investigación futura para determinar si tal precisión se mantiene para imágenes

ruidosas. Todas las herramientas analizadas para la edad, incluso con la tarea relajada de determinar un intervalo de edad en lugar de la edad exacta (Jung, et al. 2018).

De esta manera podemos deducir que las herramientas de reconocimiento de emociones pueden ser evaluadas para determinar diversas características como la raza, y en como Face++ se destaca una vez más por su precisión y en como su evaluación fue con la recolección de imágenes aleatorias de celebridades, considerando que ninguno de ellos era un niño. Hemos podido apreciar que hay algunas limitaciones en la elaboración de nuestro trabajo. Debido a que, solo se examinaron un conjunto de imágenes preseleccionadas de un mes, Los datos con un periodo de tiempo más largo, darían un estudio más robusto.

No se encontraron estudios equivalentes a la comparación sobre las tres herramientas analizadas, ni similitud de los datos obtenidos en los resultados.

Esperamos alentar con nuestro trabajo a más investigadores en las diferentes disciplinas con el interés de llegar a evaluar más herramientas y conocer la gran variedad que hay en la actualidad.

CONCLUSIONES

Los resultados obtenidos al utilizar y comparar Face++, Google Vision API y Microsoft Emotion API, mediante el muestreo de fotografías obtenidos de los distintos niños y clasificadas por un experto en psicología infantil, nos permite afirmar que, Face++ es la más precisa al momento de identificar las emociones en niños de 2 a 4 años, teniendo un 83.01% de asertividad en las 5 emociones descritas en el artículo.

Haciendo un análisis de decisión más profunda mediante el método de análisis jerárquico multicriterio (AHP), se determinó también que la herramienta Face++ es la mejor opción para elegir con un 92.6% de acuerdo a la comparación entre 4 criterios diferentes, dando lugar a futuras investigaciones de estar seguros en usar Face++ para diferentes ámbitos que se relacionen a la medicina, pediatría, psicología, reconocimiento facial de emociones en niños, entre otros.

REFERENCIAS

- Al-Hajjar, D., & Syed, A. Z. (2015). *Applying sentiment and emotion analysis on brand tweets for digital marketing*. 2015. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT 2015. Amman, Jordan. Disponible en: <https://doi.org/10.1109/AEECT.2015.7360592>.
- Bakshi, U., & Singhal, R. (2014). A survey on face detection methods and feature extraction techniques of face recognition. *International Journal of Emerging Trends & Technology*

in Computer Science (IJETTCS), 3(3), 233–237.

- Bhattacharya, A. (2016). *Emotion recognition from facial image analysis using composite similarity measure aided bidimensional empirical mode decomposition*. IEEE First International Conference on Control, Measurement and Instrumentation, (CMI), (pág. 336–340). Kolkata, India.
- Brodny, G., Kolakowska, A., Landowska, A., Szwoch, M., Szwoch, W., & Wróbel, M. R. (2016). *Comparison of selected off-the-shelf solutions for emotion recognition based on facial expressions*. 9th International Conference on Human System Interactions (HSI). (pág. 397–404) Portsmouth, UK.
- Cebrian Chulia, L. (2016). *Reconocimiento de Emociones mediante técnicas de aprendizaje profundo*. (Tesis de grado). Escola Tècnica Superior d'Enginyeria Informàtica. Universitat Politècnica de València. Valencia, España.
- Delgado, A., & Romero, I. (2015). Selección de un método para la evaluación del impacto social usando AHP. *Revista ECIPerú*, 12(1), 84-91.
- Hatem, H., Beiji, Z., & Majeed, R. (2015). A survey of feature base methods for human face detection. *International Journal of Control and Automation*, 8(5), 61–78.
- Herrera, M., & Osorio, J. (2006). Modelo para la gestión de proveedores utilizando AHP difuso. *Estudios Gerenciales*, (99),69-88.
- Jung, S.G., An, J., Kwak, H., Salminen, J. & Jansen, B.J. (2017). *Inferring social media users' demographics from profile pictures: A Face++ analysis on Twitter users*. Proceedings of The 17th International Conference on Electronic Business (pp. 140-145). ICEB, Dubai, UAE.
- Jung, S., An, J., Kwak, H., Salminen, J., & Jansen, B. J. (2018). *Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race*. Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM), (pág 624–627). Palo Alto, Estados Unidos.
- Kwak, H., & An, J. (2016). *Revealing the hidden patterns of news photos: analysis of millions of news photos using GDELT and deep learning-based Vision APIs*. Disponible en: <https://arxiv.org/abs/1603.04531v2>.
- Ma Xiaoxi, Lin Weisi, Huang Dongyan, Dong Minghui, H. L. (2017). *Facial Emotion Recognition*. 2nd International Conference on Signal and Image Processing (ICSIP) Singapur, Singapur.
- Matsumoto, D., Hwang, H. S., López, R. M., & Pérez Nieto, M. Á. (2013). Lectura de la expresión facial de las emociones: Investigación básica en la mejora del reconocimiento de emociones. *Ansiedad y Estrés*, 19(2–3), 121–129.
- Mulfari, D., Celesti, A., Fazio, M., Villari, M., & Puliafito, A. (2016). *Using google cloud vision in assistive technology scenarios*. Proceedings - IEEE Symposium on Computers and

- Communications. (pág. 214–219). Messina, Italia
- Padhi, I., Palo, H. K., Mishra, S. K., & Mohanty, M. N. (2016). Statistical feature based child emotion analysis. *International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016*, (pág 1687–1689). Chennai, India.
- Qaisi, L. M., & Aljarah, I. (2016). *A twitter sentiment analysis for cloud providers: A case study of Azure vs. AWS*. Proceedings - CSIT 2016: 2016 7th International Conference on Computer Science and Information Technology. Amman, Jordania.
- Saaty, T. L. (1980). *The analytic hierarchy process: planning, priority setting, resource allocation*. McGraw-Hill.