



UNIVERSIDAD DE MÁLAGA



GRADO EN INGENIERÍA INFORMÁTICA

DISEÑO E IMPLEMENTACIÓN DE ÍNDICES DE
VALORACIÓN DE PRODUCTOS/SERVICIOS CON
INFORMACIÓN NO SOLICITADA DE LOS SOCIAL MEDIA

DESIGN AND IMPLEMENTATION OF VALUATION
INDICES OF PRODUCTS/SERVICES WITH UNSOLICITED
INFORMATION FROM SOCIAL MEDIA

Realizado por
FRANCISCO MANUEL CUEVAS MÁRQUEZ

Tutorizado por
JOSE IGNACIO PELÁEZ SÁNCHEZ

Departamento
LENGUAJES Y CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE MÁLAGA

MÁLAGA, OCTUBRE 2020



UNIVERSIDAD
DE MÁLAGA





UNIVERSIDAD
DE MÁLAGA





UNIVERSIDAD
DE MÁLAGA





UNIVERSIDAD
DE MÁLAGA



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA
GRADO EN INGENIERÍA INFORMÁTICA

**DISEÑO E IMPLEMENTACIÓN DE ÍNDICES DE VALORACIÓN
DE PRODUCTOS/SERVICIOS CON INFORMACIÓN NO
SOLICITADA DE LOS SOCIAL MEDIA**

**DESIGN AND IMPLEMENTATION OF VALUATION INDICES OF
PRODUCTS/SERVICES WITH UNSOLICITED INFORMATION
FROM SOCIAL MEDIA**

Realizado por
D. Francisco Manuel Cuevas Márquez
Tutorizado por
Dr. Jose Ignacio Peláez Sánchez
Departamento
Lenguajes y Ciencias de la Computación

UNIVERSIDAD DE MÁLAGA
MÁLAGA, OCTUBRE 2020

Fecha defensa:



UNIVERSIDAD
DE MÁLAGA





Resumen

La revolución tecnológica de estos últimos años ha dado lugar a lo que se conoce como “sociedad de la información”, la cual no solo ha originado cambios en aspectos puramente económicos, sino que también lo ha hecho en aspectos sociales y organizacionales. La Sociedad de la Información se ha convertido en un fenómeno de transformación que ha alcanzado a la sociedad en su conjunto, caracterizada por la capacidad de sus miembros (ciudadanos, empresas y administraciones) para obtener, compartir y procesar cualquier información por medios tecnológicos, desde cualquier lugar y en la forma que se prefiera.

Este cambio de modelo social, en el que Internet se ha convertido en un espacio de interacción humana en el que los usuarios/consumidores realizan valoraciones de productos y servicios basadas en sus experiencias y sentimientos, se está convirtiendo en la principal fuente de prescripción utilizada para la adquisición de productos y servicios.

En este contexto, las metodologías de análisis multicriterio (MCDA), así como los índices de valoración (ORIs), que utilizan valoraciones subjetivas o juicios, están experimentando un gran interés desde diferentes ámbitos, como el académico o el empresarial.

El objetivo de este trabajo es implementar las principales metodologías MCDA: método AHP y método EMPV; y los principales índices ORIs: QV-ORIs, CMGI-Index; y crear una librería de código abierto en Python, que puedan ser utilizadas por la comunidad universitaria, para su aplicación académica y de investigación.

Palabras Clave

Decisión Multicriterio · Consistencia · Proceso Analítico Jerárquico · Operadores de Agregación · Índices de Reputación Online.



UNIVERSIDAD
DE MÁLAGA





Abstract

The technological revolution of recent years has given rise to what is known as the “Information Society”, which has not only led to changes in purely economic aspects, but also in social and organizational aspects. The Information Society has become a phenomenon of transformation that has reached society as a whole, characterized by the capacity of its members (citizens, companies, and administrations) to obtain, share and process any information by technological means, from any place in any form.

This change of social model, in which Internet has become a space for human interaction in which users/consumers make evaluations of products and services based on their experiences and feelings, is becoming the main source of prescription used for the acquisition of products and services.

In this context, multi-criteria analysis methodologies (MCDA), as well as valuation indices (ORIs), which use subjective assessments or judgments, are experiencing great interest from different fields, such as academia or business.

The objective of this work is to implement the main MCDA methodologies: AHP method and EMPV method; and the main ORIs indexes: QV-ORIs, CMGI-Index; and to create an open source library in Python, which can be used by the university community, for its academic and research applications.

Keywords

Multicriteria Decision · Consistency · Analytic Hierarchy Process · Aggregation Operators · Online Reputation Indices



UNIVERSIDAD
DE MÁLAGA





A mis padres, que me han apoyado desde el primer hasta el último minuto.

A María, por su apoyo y paciencia en los momentos más complicados.

A Isabel, que allá donde esté, sabrá que terminé este camino.

A Jose Ignacio, por confiar en mí y haber estado siempre ahí cuando le he necesitado.

A Francisco y Fabián, por aguantar todos esos días de dudas.



UNIVERSIDAD
DE MÁLAGA





Índice

| | |
|---|----|
| 1.Introducción | 15 |
| Contexto..... | 15 |
| Motivación | 17 |
| Objetivo del proyecto..... | 17 |
| Estructura del documento | 18 |
| 2.Estudio del estado del arte | 19 |
| Toma de decisión multicriterio..... | 19 |
| Teoría de la Elección Social..... | 20 |
| Métodos de agregación OWA | 21 |
| Método de Condorcet..... | 23 |
| Método de Kemeny-Young..... | 23 |
| Método de Borda..... | 24 |
| 3.Tecnologías Utilizadas | 25 |
| Python | 25 |
| Numpy..... | 26 |
| Nltk..... | 26 |
| Sklearn | 26 |
| Archi | 27 |
| Bitbucket | 28 |
| Git..... | 29 |
| Pycharm Python | 30 |
| 4.Metodología | 33 |
| 5.Diseño | 37 |
| Métodos de toma de decisión multicriterio | 37 |
| Proceso Analítico Jerárquico (AHP)..... | 38 |
| Método Harker para la Construcción de Matrices Consistentes | 42 |
| Eigenvector Method for Pairwise Voting (EMPV)..... | 43 |
| Índices de Reputación Online | 44 |
| Índice QV-ORI..... | 44 |
| Reprtrak-Pulse | 46 |
| Índice de Sinergia CMGI | 48 |



| | |
|--|-----------|
| Índice de Reputación por Intervalos..... | 50 |
| Índices de consistencia de datos..... | 52 |
| Índice de consistencia Saaty..... | 52 |
| Ratio de consistencia de Saaty..... | 52 |
| Índice de consistencia CI*..... | 53 |
| Índice de consistencia CI+..... | 55 |
| Ratio de consistencia CR+..... | 56 |
| Índice de Consistencia Intervalar (ICI+)..... | 57 |
| Índice de Fiabilidad de Intervalos (IRI)..... | 57 |
| Operadores de agregación..... | 58 |
| Operadores Básicos: Media Aritmética, Geométrica y Armónica..... | 58 |
| Majority Additive Ordered Weighting Averaging Operator (MA-OWA)..... | 59 |
| Selective Majority Additive Ordered Weighting Averaging Operator (SMA-OWA) | 59 |
| Interval Selective Majority Additive Ordered Weighting Averaging Operator (ISMA-OWA)..... | 60 |
| Continuous Ordered Weighted Geometric (C-OWG)..... | 60 |
| Additive Continuous Ordered Weighted Geometric (AC-OWG)..... | 61 |
| Análisis semántico de texto..... | 62 |
| Naive-Bayes..... | 62 |
| Red Neuronal..... | 63 |
| Captura de datos a través de páginas web..... | 63 |
| Twitter y YouTube..... | 64 |
| 6.Implementación | 65 |
| Clase Ahp..... | 66 |
| Clase AhpConsistency..... | 66 |
| Clase Cmgj..... | 66 |
| Clase Empv..... | 67 |
| Clase Interval..... | 67 |
| Clase Harkerprediction..... | 67 |
| Clase Operators..... | 68 |
| Clase Qvori..... | 68 |
| Clase Reprakpulse..... | 68 |
| Analizador Semántico..... | 68 |



| | |
|--|-----------|
| Clase LazyTrainer | 69 |
| Clase SentimentClassifier | 69 |
| Clase ClassifierFactory | 69 |
| Clase Semanticmethods | 70 |
| Captura de datos..... | 70 |
| 7.Pruebas | 71 |
| 8.Conclusión | 73 |
| Bibliografía | 75 |
| Anexo I: Listado de Figuras | 79 |



UNIVERSIDAD
DE MÁLAGA





1.Introducción

Contexto

La revolución tecnológica de estos últimos años ha dado lugar a lo que se conoce como “sociedad de la información”, la cual no solo ha originado cambios en aspectos puramente económicos, sino también en aspectos sociales y organizacionales (J. I. Peláez, Martínez, & Vargas, 2019). La Sociedad de la Información se ha convertido en un fenómeno de transformación que ha alcanzado a la sociedad en su conjunto, caracterizada por la capacidad de sus miembros (ciudadanos, empresas y administraciones) para obtener, compartir y procesar cualquier información por medios tecnológicos, desde cualquier lugar y en la forma que se prefiera (José I. Peláez, Martínez, & Vargas, 2019).

Este cambio de modelo social, en el que Internet se ha convertido en un espacio de interacción humana en el que los usuarios/consumidores realizan valoraciones de productos y servicios basadas en sus experiencias y sentimientos (J. I. Peláez, Bernal, & Karanik, 2016), (Carrasco, Sánchez-Fernández, Muñoz-Leiva, Francisca Blasco, & Herrera-Viedma, 2017), se está convirtiendo en la principal fuente de prescripción utilizada para la adquisición de productos y servicios (J. I. Peláez et al., 2019), (Chung & Koo, 2015), (Zhang, Tian, Fan, Liu, & Fan, 2019).

Las empresas/organismos/administraciones/etc., conscientes de este cambio social, diseñan sus sitios web con la intención de recopilar esta información para mejorar su relación con los distintos públicos, y, por consiguiente, su negocio (Hu & Liu, 2004). En este escenario, han surgido dos líneas de actualización/investigación, la primera relacionada con la valoración de los productos o servicios a través de índices de valoración, los denominados ORIs; y una segunda línea, relacionada con los Sistemas de Apoyo a la Decisión (DSS, siglas en Ingles), (J. I. Peláez et al., 2016), (J. I. Peláez et al., 2019).



Los ORIs, son índices que representan la valoración de los usuarios acerca de un producto/servicio a través de sus valoraciones (Carrasco et al., 2017). Utilizan escalas de valores discretos para que el usuario emita su opinión. Sitios como Amazon, eBay, TripAdvisor, Trivago, Booking, etc., califican libros, productos electrónicos, ropa, hoteles, etc., utilizando estas escalas que utilizan estrellas, gustos, etc., para que el usuario exprese su grado de satisfacción, junto con un valor global de satisfacción del producto/servicio que se obtiene como media de las valoraciones (Hu & Liu, 2004).

Un Sistema de Apoyo a las Decisiones, es un sistema informático que sustenta el proceso de toma de decisiones, lo cual implica la utilización de datos y modelos para la generación, la estimación, la evaluación y/o la comparación sistemática de alternativas, ayudando a los responsables de la toma de decisión a reunir inteligencia, generar opciones y tomar decisiones. En este contexto uno de los principales DSS es el modelo AHP.

El Modelo AHP (Thomas L. Saaty, 2008) es una técnica usada para la toma de decisiones, se basa en realizar una distribución de las decisiones a tomar en función de una prioridad o jerarquía que ayuda a visualizar cuál o cuáles son las decisiones que mayor impacto crean para el objetivo buscado, sea un problema y ajustándose a las necesidades del momento.

Este tipo de herramientas se engloba en un marco para encuadrar un problema de decisión, y poder realizar una visualización de como impactan cada una de las decisiones, y de esta manera poder evaluarlas por orden de impacto y poder buscar alguna alternativa cuando no se esté seguro de cómo van a impactar las decisiones elegidas sobre el problema o asunto a tratar.



Motivación

Las motivaciones que han llevado a realizar este trabajo final de grado han sido las siguientes:

1. Iniciación a la Investigación. Las dos líneas de investigación que se han presentado en la contextualización del proyecto, la primera relacionada con los índices de valoración, los denominados ORIs; y la segunda, relacionada con los Sistemas de Ayuda a la Decisión (DDS, siglas en Ingles), como se ha indicado son de gran actualidad y presentan un gran crecimiento a nivel de investigación, debido a la necesidad de nuevos y mejores sistemas de valoración y de toma de decisión.
2. Temática de Actualidad y de Interés Empresarial. Se ha indicado anteriormente, cada vez más las empresas/administraciones/etc., implementan sus sitios web con sistemas para la captura de las opiniones de los usuarios, para de esta forma, mejorar su negocio. Este hecho, hace necesario de especialistas en este campo, para incorporarse en el mercado laboral.
3. Aplicación de los Conocimientos Adquiridos en los Estudios de Grado. Este proyecto implica la aplicación de una amplia gama de conocimientos adquiridos durante los estudios de grado, y algunos más avanzados, como, por ejemplo: metodología de Ingeniería del Software, programación, análisis semántico, control de versiones, toma de decisión, entre otros.

Objetivo del proyecto

Dentro del objetivo de este proyecto se encuentra la implementación de metodologías MCDA entre las cuales se encuentran los métodos AHP y EMPV. Además, se implementarán índices de valoración ORIs, como, por ejemplo, Qv-ORI, CMGI-Index.

Otro objetivo surge de la necesidad de aplicar correctamente estos métodos e índices de valoración, para lo cual se incluye un apartado de extracción de datos (Scraping) y otro para el análisis semántico de los comentarios extraídos.



Junto con lo anterior, se pretende crear una librería de código abierto que incluya los métodos e índices anteriormente indicados, junto con un analizador semántico y un apartado de extracción de datos. Para que esta librería pueda ser utilizada por aquellas personas que requieran de su uso para necesidades académicas o de investigación.

Estructura del documento

El presente trabajo ha sido organizado como sigue: en la primera sección, se realiza una introducción a la valoración de productos y los modelos de decisión que se aplican en dicho contexto; en la segunda sección, se realiza un estudio del arte, centrándonos en aquellos aspectos que son de interés para el trabajo realizado; en las secciones tercera, cuarta, quinta, sexta y séptima, se presenta el desarrollo del trabajo realizado, desde las metodologías utilizadas, el diseño del sistema, implementación de la solución, y las pruebas realizadas; en la octava sección, se presentan las conclusiones; en la novena, la bibliografía utilizada; y finaliza el trabajo, con los anexos, donde se presenta un manual de usuario del sistema desarrollado, y una relación de figuras.



2. Estudio del estado del arte

Toma de decisión multicriterio

La vida de las personas se fundamenta en torno a las decisiones que deben tomar frente a las situaciones cotidianas que se les plantean, tanto en el ámbito profesional como en el personal.

Los problemas de toma de decisión se han abordado durante mucho tiempo desde la perspectiva monocriterio (un único criterio de decisión), pero este planteamiento poco a poco ha ido perdiendo protagonismo en beneficio de la perspectiva multicriterio (diversos criterios, a menudo en conflicto).

Los procesos de Toma de Decisión Multicriterio se caracterizan porque tienen en cuenta al menos dos criterios de decisión, que suelen ser en la mayoría de los casos contradictorios (el beneficio de uno supone la penalización de otro) y al menos dos alternativas de decisión. Además, permiten reflejar las preferencias del decisor o grupo de decisores y estas se tienen en cuenta durante el proceso de elección.

Los Métodos de Toma de Decisiones Multicriterio han desarrollado una terminología propia y común, que incluye conceptos tales como:

- Alternativas: Son las posibles soluciones al problema de decisión, entre las cuales el decisor puede elegir.
- Atributos: Son las características, rasgos, cualidades, o parámetros que describen cada una de las alternativas. El número de atributos que describe las alternativas será elegido por el decisor o grupo de elección.
- Criterios: Son los parámetros que permiten reflejar las preferencias del decisor respecto a un atributo. Los criterios pueden presentarse de dos formas, criterios cuantitativos (evaluaciones numéricas) y cualitativos (no existe una unidad de medida, la medida es subjetiva).

- **Objetivos:** Delimitan los deseos que se quieren satisfacer, indicando las direcciones de mejora según las preferencias del conjunto decisor. Pueden considerarse como ideales inalcanzables.
- **Metas:** Reflejan los ideales alcanzables. La alternativa que recogerá los atributos establecidos y pueda satisfacer los criterios, acercándose al máximo a los objetivos establecidos.

Los procesos de toma de decisiones comprenden las 5 primeras fases de cualquier proceso de resolución de problemas, el cual está compuesto por 7 etapas:

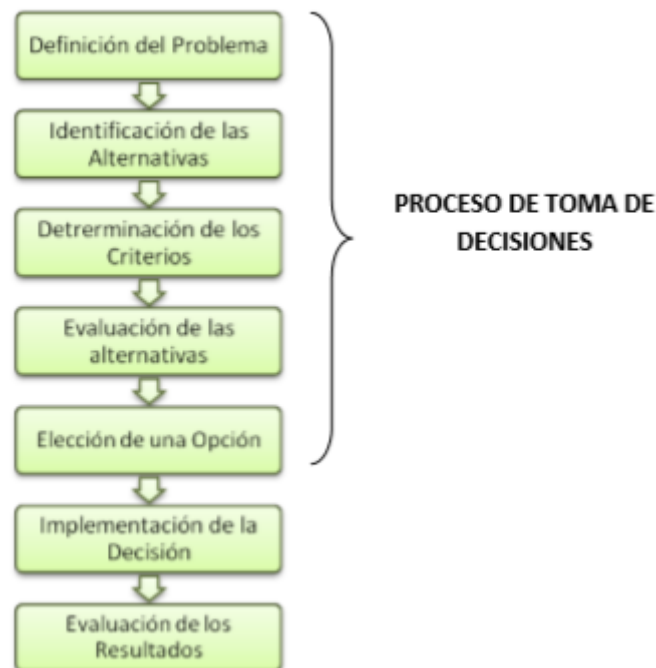


Figura 1. Proceso de toma de decisiones.

Teoría de la Elección Social

La teoría de la elección social es el apartado teórico sobre las opiniones, preferencias e intereses para alcanzar una decisión o un bienestar social. Dicha teoría tiene sus comienzos con la formulación de Condorcet y su paradoja de la votación (Doña, 2008). La elección social y los valores individuales, junto con el teorema de imposibilidad de Arrow son la base de la teoría de la elección social moderna (Kelly, 1988).

Pero el teorema de Arrow no es el único que pertenece a esta teoría, otros teoremas como el Teorema de Gibbard-Satterthwaite, Teorema del jurado de Condorcet, Teorema del votante mediano y el Teorema de May son los más conocidos sobre la teoría de elección social.

El teorema de “Elección social y los valores individuales”, es presentado por Kenneth Arrow en 1951, estableciendo las propiedades que debe satisfacer (Fishburn, 1979):

- Dominio no restringido: El resultado del voto debería poder ordenar entre sí todas las preferencias.
- No imposición o Pareto débil: Si A es preferido a B, al menos una persona votó que prefiere A sobre B.
- Ausencia de dictadura: La regla de elección social no sigue el orden de preferencia de un único individuo.
- Monotonía: Si un individuo cambia su orden de preferencia al promover una opción, el orden de preferencia de la sociedad responde promoviendo esa misma opción.
- Independencia de alternativas irrelevantes: Los cambios en las alternativas irrelevantes no deberían tener impacto en el ordenamiento que haga la sociedad en el conjunto relevante.

Métodos de agregación OWA

Dentro de la toma de decisión se han desarrollado diferentes esquemas que intentan dar respuesta al problema de la toma de decisión social, como son, entre otros: los grafos de preferencias, el proceso analítico jerárquico (AHP) (Thomas L. Saaty, 2008), métodos de fusión, toma de decisión lingüística, etcétera, teniendo como nexo común todos ellos el uso de operadores de agregación ya sea en una o en varias de sus fases de desarrollo.

La familia de operadores de agregación más extendida en la actualidad son los operadores OWA (Ordered Weighted Averaging) (Yager, 1988). Tanto por su importancia, así como por su influencia en todos los procesos de toma de decisión, se

desarrollan continuamente multitud de trabajos de investigación en el área de las ciencias de la computación con el objetivo de aportar nuevos elementos que sirvan como mecanismos útiles a los sistemas de apoyo a la decisión y que introduzcan mejoras desde un punto de vista social a los resultados que se obtienen.

Un operador OWA es una función de dimensión n a la que se le asocia un vector W de longitud n , tal que, $W=[w_1, w_2, w_3, \dots, w_n]^T$ donde (Doña, 2008):

1. $w_i \in [0,1]$
2. $\sum_{i=1}^n w_i = 1$

$f(a_1, a_2, \dots, a_n) = \sum_{j=1}^n w_j^* b_j$, donde b_j es el j -ésimo mayor valor de a_n .

La obligación de ordenar pesos permite que un elemento a_i no se asocie al peso w_i , sin embargo un peso si estará asociado a una posición una vez estas estén ordenadas. En función de la ponderación se dan tres casos para la clasificación de operadores OWA (Yager, 1988), (Doña, 2008):

1. F^* , entonces $W = W^* = [1, 0, \dots, 0]^T$
2. F_* , entonces $W = W_* = [0, 0, \dots, 1]^T$
3. F_{ave} , entonces $W = W_{ave} = [\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}]^T$

Por lo que se puede deducir que:

1. $F^*(a_1, a_2, \dots, a_n) = \text{Max}(a_i)$
2. $F_*(a_1, a_2, \dots, a_n) = \text{Min}(a_i)$
3. $F_{ave}(a_1, a_2, \dots, a_n) = \frac{1}{n} * \sum_{i=1}^n a_i$



Método de Condorcet

Creado por Marie Jean Antoine Nicolas de Caritat, Marqués de Condorcet es un sistema de votación el cual elige un ganador entre una serie de candidatos y además cumple con la Paradoja de Condorcet.

Este método fue creado por Marie Jean Antoine Nicolas de Caritat, Marqués de Condorcet. Las claves de este método son:

- Ordenar candidatos por orden de preferencia, pudiendo haber empate.
- Comparar cada candidato en la papeleta con cada uno del resto. El de mayor preferencia obtendrá un punto.
- El candidato que gana a cada uno de los otros candidatos más veces de las que ha perdido es el ganador.

Se puede dar el caso de que gane un candidato que no haya sido el preferido de ningún votante, ya que se elige un candidato el cual no tiene por qué ser la primera preferencia de los votantes, pero tampoco la última.

Método de Kemeny-Young

Creado por John Kemeny en 1959, es un método que mediante preferencias y comparación entre pares encuentra la opción más votada en una elección (Kemeny, Endriss, & Haan, n.d.).

Este método pertenece al método Condorcet, ya que en caso de que haya un ganador va a ser el más popular (que no quiere decir que sea el más votado). Los votantes realizan un ranking en función de la preferencia que ellos quieren. A raíz de esto, el método genera una matriz con los pares de preferencia, es decir, aquel porcentaje o número de votos que prefiere X con respecto a Y.

Una vez generada dicha matriz, se separa en comparación de pares. Comparando todas las posibilidades, para ello se usará el siguiente ejemplo:



Una matriz con 4 opciones a elegir (A, B, C o D) tiene 12 comparaciones. (A respecto a B, C, D + B respecto a C, D + C respecto a D y viceversa).

Usando estas comparaciones se obtiene otra matriz donde, mediante comparación de pares, se obtendrá que porcentaje de votantes prefiere cada alternativa con respecto

a las demás. Si de 4 alternativas, una de ellas es preferida con respecto a las otras 3, está claro que es la ganadora. Si otra es preferida con respecto a 2, será la segunda elegida, y así sucesivamente.

Al ser un método englobado en el método Condorcet cumplirá:

- No imposición: Libre elección de preferencias
- Criterio de Condorcet: Si una elección gana todas las comparaciones, es la ganadora.
- Criterio de la mayoría: Si la mayoría de los votantes prefiere X, esa es la ganadora.
- Ausencia de Dictador: Un único votante no puede controlar la votación.

Método de Borda

Propuesto por Jean Charles Borda en 1770 es uno de los métodos principales de recuento. Es un método válido para la elección de candidatos entre varios distintos.

Este método consiste en el orden de los distintos candidatos de mayor a menor puntuación dentro de la papeleta de voto. Una vez todos los votantes hayan finalizado la votación se tomarán todas las papeletas y se sumarán las puntuaciones de todas ellas, dando por ganador a aquel candidato con mayor puntuación global, segundo a aquel con la segunda mayor puntuación global y así sucesivamente.



3. Tecnologías Utilizadas

En esta sección se presentan las tecnologías utilizadas para el desarrollo del presente trabajo final de grado. En primer lugar, se presenta el lenguaje de programación Python, el cual ha sido elegido, entre otros motivos, por la proyección profesional, software libre, etc.

Además, se han usado las siguientes librerías incluidas en el lenguaje de programación Python, como pueden ser: Numpy, Nltk, Sklearn o Requests entre otras.

Para la gestión y el control de versiones se ha utilizado GIT y Bitbucket. Como herramienta de modelado se ha usado Archi, y, por último, se ha usado como entorno de desarrollo Pycharm.

Python

Python es un lenguaje interpretado, dinámico y simple. Fue creado por Guido Van Rossum a finales de los 80, publicándose su código por primera vez en 1991. Es de propósito general, es decir, no está orientado a una finalidad concreta.

Es un lenguaje multiparadigma, no solo se orienta a objetos, sino que también es posible usarlo en programación imperativa o programación funcional. Es interpretado, no tendemos que compilar nosotros el código, si no que un intérprete se encargará de eso por nosotros. Además, incluye una extensa cantidad de librerías y funciones que se usarán para la creación de los métodos de decisión y de los índices ORIs.

Debido a que es un lenguaje de código abierto, que dispone de las librerías y funciones matemáticas necesarias y que es utilizado para tareas de análisis numérico en la comunidad científica, así como para aplicaciones de propósito general se decide usar este lenguaje para la implementación de la librería.



Numpy

Numpy es una librería de Python, la cual contiene material fundamental para la computación científica. La mayor parte de esta librería es usada en el trabajo con matrices y vectores. Es una librería de código abierto y encontramos todo lo necesario para trabajar con ella en <https://numpy.org/>.

Se elige este paquete ya que contiene las funciones básicas para trabajar con matrices y álgebra lineal que es lo que más se ha necesitado dentro de la librería implementada. Numpy pertenece a Scipy, un conjunto de software abierto el cual comprende: Numpy, Librería SciPy, Matplotlib, IPython, SymPy y Pandas. Dentro de la web scipy.org está toda la información acerca de este software.

Nltk

Nltk es una biblioteca para lenguaje Python que contiene los métodos necesarios para poder trabajar con lenguaje humano. Proporcionará interfaces con las cuales se podrá trabajar con más de 50 corpus lingüísticos, además, procesará el texto para clasificación, tokenización, stemming, etiquetado, análisis sintáctico y razonamiento semántico.

Gracias a esta biblioteca se podrá trabajar en los analizadores sintácticos que permiten reconocer un texto para obtener su polaridad e implementar índices basados en esta como pueda ser el “Índice de Sinergia CMGI”. Se puede encontrar toda la información relacionada a esta biblioteca en su web: <https://www.nltk.org/>.

Sklearn

La biblioteca scikit-learn de Python es una potente herramienta para el análisis predictivo de datos. Construida sobre Numpy, Scipy y Matplotlib (Librerías de cálculo matemático). Dentro de sus múltiples usos se encuentra la clasificación, regresión, clustering, reducción dimensional, selección de modelos...



El principal uso se le da a una librería llamada “metrics”, el objetivo es cuantificar la calidad de las predicciones que se realizan. La información relacionada a esta biblioteca se encuentra en <https://scikit-learn.org/stable/index.html>.

Archi

Archi es una herramienta de modelado de código abierto. Basada en parte del estándar IEEE 1471 ofrece un lenguaje para describir distintos procesos dentro de una arquitectura empresarial. Tiene una gran orientación al modelado empresarial. Mientras que otros lenguajes de modelado, tales como UML y BPMN, son muy pesados, Archimate presume de ser un lenguaje ligero.

Es una herramienta muy sencilla ya que permite dividir por niveles el modelado, (negocio, aplicación y tecnología), como a su vez dividir en aspectos (estructura pasiva, comportamiento y estructura activa).

Dentro de cada capa se encuentra lo siguiente:

- Capa empresarial: Se encuentran los procesos de negocio, servicios, funciones y eventos de unidades de negocio.
- Capa de aplicación: Se encuentran las aplicaciones software.
- Capa de tecnología: Se encuentran hardware e infraestructura de comunicaciones.

En <https://www.archimatetool.com/> se puede encontrar la información relacionada a este software.

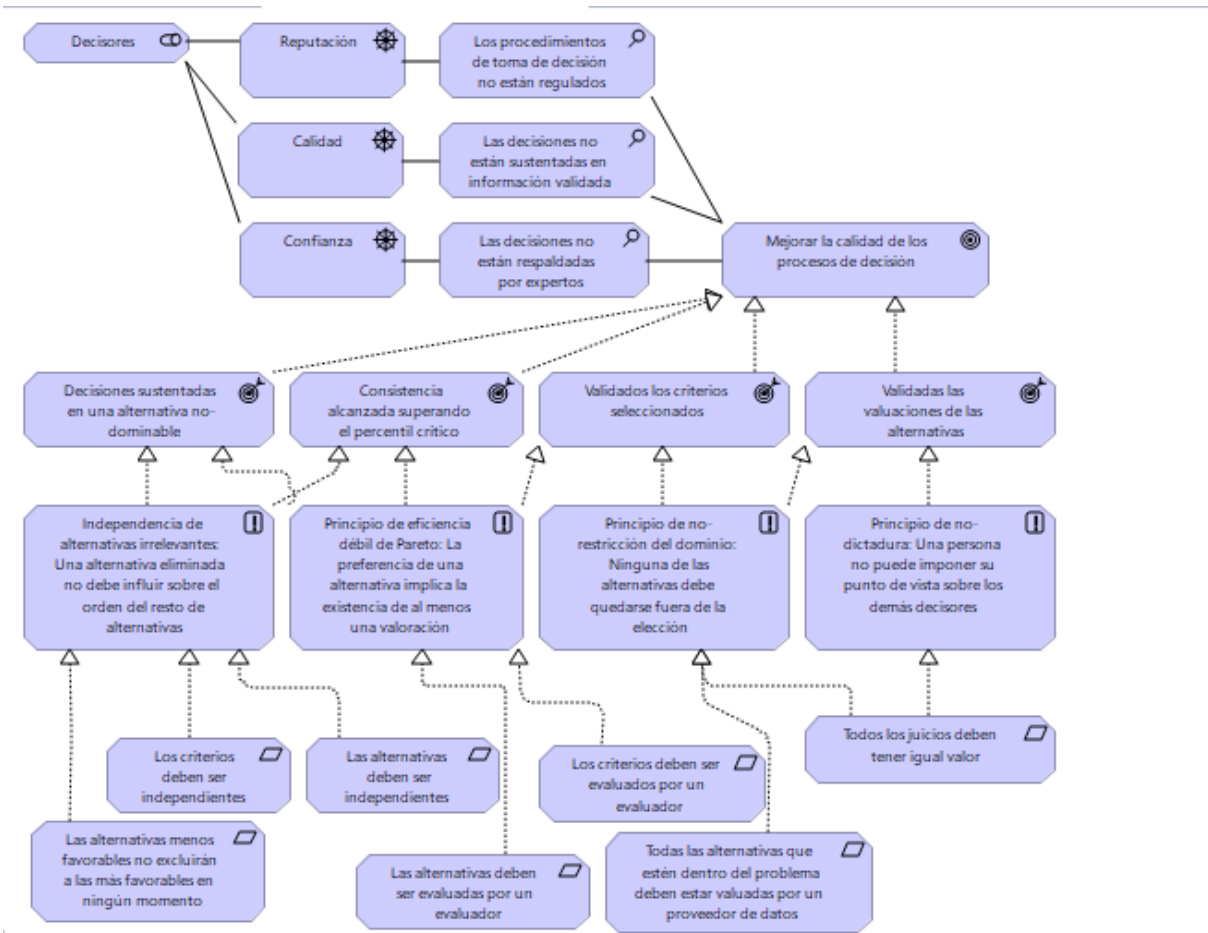


Figura 2. Modelado del método AHP con Archimate.

Bitbucket

Bitbucket es un repositorio, escrito en Python, que permite gestionar las acciones realizadas por el control de versiones Git. Es una gran herramienta ya que permite compartir dentro de una misma carpeta el trabajo de varios individuos por separado.

La información sobre este software puede encontrarse en <https://www.atlassian.com/es/software/bitbucket>.

| Author | Commit | Message | Date |
|------------------|---------|---|---------------|
| Francisco Cuevas | 66d8b6f | Added ICI+, QV-ORI and their tests | 2 minutes ago |
| Francisco Cuevas | 4d4b57f | Added ISMA-OWA operator | 2019-07-16 |
| Francisco Cuevas | a94718e | Added intervalar files and minor changes in tests | 2019-07-11 |
| Francisco Cuevas | 85db54d | Added first version of CR+ | 2019-07-05 |
| Francisco Cuevas | 87c8a1a | Added papers tests and comments | 2019-07-05 |
| Francisco Cuevas | 5ad993a | Added EMPV method, comments and up version | 2019-07-02 |
| Francisco Cuevas | 618b315 | First Version of EMPV without clean code | 2019-07-01 |
| Fabián Vaccaro | d44e519 | modificaciones en el modelo de motivación | 2019-06-19 |
| fabian | 72cb5e1 | avances en la arquitectura | 2019-06-18 |
| Francisco Cuevas | 0200270 | Added Archimate Models | 2019-06-18 |
| Francisco Cuevas | ae4a907 | Added AHP model and ci_star with percentile | 2019-06-14 |
| Francisco Cuevas | 287d439 | Fixed problem in ci_star with matrix 3x3 | 2019-06-07 |

Figura 3. Captura de cambios en el proyecto mediante Bitbucket.

| Name | Size | Last commit | Message |
|--------------------------------|----------|---------------|---|
| .ipynb_checkpoints | | 2019-06-14 | Added AHP model and ci_star with percentile |
| mcdm | | 3 minutes ago | Added ICI+, QV-ORI and their tests |
| notebooks | | 2019-06-14 | Added AHP model and ci_star with percentile |
| pruebas | | 3 minutes ago | Added ICI+, QV-ORI and their tests |
| tests | | 3 minutes ago | Added ICI+, QV-ORI and their tests |
| .gitignore | 338 B | 2019-05-09 | Added .gitignore |
| Arquitectura_AHP.archimate | 52.18 KB | 2019-07-01 | First Version of EMPV without clean code |
| Arquitectura_AHP.archimate.bak | 52.13 KB | 2019-07-01 | First Version of EMPV without clean code |
| NotebookAHP.ipynb | 1.15 KB | 2019-06-14 | Added AHP model and ci_star with percentile |
| README.md | 105 B | 2019-05-08 | Added readme file |
| requirements.txt | 5 B | 2019-05-14 | Made mcda package pipable |
| setup.py | 320 B | 2019-07-02 | Added EMPV method, comments and up version |

Figura 4. Carpetas del proyecto en Bitbucket.

Git

Git es un control de versiones de código abierto basado en BitKeeper y Monotone, escrito en C y mediante el cual se van a poder subir los archivos necesarios al

repositorio que tenemos en Bitbucket. Lo que hace este software es llevar un registro de cambios en los archivos que se estén modificando.

Es un software rápido ya que casi todas las operaciones se realizan localmente. Además, genera respaldos múltiples, lo que quiere decir que, si se trabaja entre varios usuarios, cada uno tendrá una copia de seguridad.

Se puede obtener más información sobre Git en <https://git-scm.com/>.

```
MINGW64:/c/Users/Francisco/Desktop/TFG/Repositorios/mcdmlib
no changes added to commit (use "git add" and/or "git commit -a")

Francisco@DESKTOP-CKTENJ2 MINGW64 ~/Desktop/TFG/Repositorios/mcdmlib (master)
$ git add -A

Francisco@DESKTOP-CKTENJ2 MINGW64 ~/Desktop/TFG/Repositorios/mcdmlib (master)
$ git commit -m "Added ICI+, QV-ORI and their tests"
[master 66d8b6f] Added ICI+, QV-ORI and their tests
12 files changed, 454 insertions(+), 130 deletions(-)
create mode 100644 mcdm/owg.py
create mode 100644 mcdm/qvori.py
create mode 100644 pruebas/pruebasintervalarori.py
create mode 100644 tests/test_intervalar.py
create mode 100644 tests/test_valuationindexes.py

Francisco@DESKTOP-CKTENJ2 MINGW64 ~/Desktop/TFG/Repositorios/mcdmlib (master)
$ git push
Enumerating objects: 28, done.
Counting objects: 100% (28/28), done.
Delta compression using up to 4 threads
Compressing objects: 100% (17/17), done.
Writing objects: 100% (17/17), 5.11 KiB | 872.00 KiB/s, done.
Total 17 (delta 8), reused 0 (delta 0)
To https://bitbucket.org/cmgi/mcdmlib
```

Figura 5. Control de versiones con el software Git.

Pycharm Python

PyCharm es un entorno de desarrollo de la compañía JetBrains el cual va a ser utilizado para la implementación de todo el código de la librería. Dispone de su propio terminal para poder realizar las pruebas e instalar paquetes de forma muy sencilla. Además, posee un asistente para la codificación el cual indicará errores durante la marcha, incluso refactorización de código automática.

Una de sus características es que ofrece soporte para desarrollo web y paquetes científicos (Véase Numpy).

Toda la información referente a Pycharm puede encontrarse en <https://www.jetbrains.com/es-es/pycharm/>.

```
1 import mcdm.ahp as ahp
2 import numpy as np
3
4 matrix = np.ones((3, 3))
5
6 matrix[0, 1] = 2
7 matrix[0, 2] = 6
8 matrix[1, 2] = 3
9 matrix[1, 0] = 1/2
10 matrix[2, 0] = 1/6
11 matrix[2, 1] = 1/3
12
13 a1 = [6, 7, 5]
14 a2 = [2, 6, 2]
15 a3 = [8, 6, 8]
16
17 a = ahp.Ahp(matrix)
18 b = ahp.AhpConsistency(a)
19
20 print("AHP alternativa 1: ", a.ahp(a1))
21 print("AHP alternativa 2: ", a.ahp(a2))
22 print("AHP alternativa 3: ", a.ahp(a3), "\n")
23 print("Autovector normalizado:", a.get_criteria_weights())
```

Figura 6. Captura del software Pycharm.



UNIVERSIDAD
DE MÁLAGA



4. Metodología

El apartado de la metodología es uno de los más importantes en un proyecto software, ya que es la guía para dicho proyecto. La metodología comprende desde el inicio del proyecto hasta la entrega del mismo, sin una buena metodología la probabilidad de fracaso es bastante alta.

La metodología usada en este proyecto es SCRUM. Para conocer esta metodología debemos saber que Ikujiro Nonaka y Takeuchi en 1986 publican “The new product development game”, a raíz de esta publicación surge una forma de llevar a cabo proyectos de forma flexible y ágil. (Trigás Gallego, 2012)

Partiendo de esa base, Nonaka y Takeuchi ven que empresas tecnológicas no disponen de un equipo especializado en todas las fases de creación del producto, si no que el equipo es multidisciplinar de inicio a fin. De ahí que ambos compararan el progreso del producto con una melé de jugadores de rugby y de ahí aparece la denominación Scrum. (Trigás Gallego, 2012)

Las características más importantes de esta metodología son las siguientes:

- **Transparencia:** Los miembros del equipo saben lo que pasa en todo momento.
- **Inspección:** Se comprueba frecuentemente el progreso para detectar posibles problemas.
- **Adaptación:** Si hay que cambiar algo, el equipo realiza el cambio para conseguir el objetivo del sprint correspondiente.

El método Scrum se puede aplicar a proyectos con:

- **Incertidumbre:** No se sigue un plan estrictamente detallado.
- **Auto-organización:** La organización es cosa de todos.
- **Control moderado:** Se crea un escenario que permita evitar los descontrolados.
- **Transmisión de conocimientos:** Todos aprenden de todos.

Esta metodología se basa en sprints. Para tratar sobre un sprint se realizarán reuniones diarias, breves y concisas para aclarar todas las dudas sobre la realización de dicho sprint, estos se desarrollarán en 5 fases:

1. Concepto: Definición genérica de producto y asignación de equipo.
2. Especulación: Se establecen límites de costes y se construirá el producto siguiendo las ideas principales.
3. Exploración: Se incrementan las funcionalidades de la fase de especulación.
4. Revisión: El equipo revisa y contrasta todo lo construido hasta el momento.
5. Cierre: Se entrega una versión del producto (no quiere decir que esté cerrado), se irán haciendo cambios hasta alcanzar el producto deseado.

Dentro de las reuniones mencionadas anteriormente tenemos distintos puntos (en cada una de las reuniones realizadas):

- Planificación: Se define un documento con los requisitos por prioridades. Se define el sprint 0 que decide cuales son los objetivos, se obtendrá un “Sprint Backlog” que será la lista de tareas y el objetivo principal de la planificación.
- Seguimiento: ¿Qué hemos avanzado?, ¿Qué se hará hasta la siguiente reunión?, ¿Qué problemas han surgido?, ¿Qué debemos solucionar?
- Revisión: Se revisará lo que se ha avanzado y se presentará el resultado con respecto a la reunión anterior, entregando una versión nueva del producto.

Dentro de un desarrollo Scrum se encuentran varios roles de entre los cuales destacan:

- Product Owner: Se encarga de que el equipo trabaja correctamente y ayuda al usuario a escribir las historias de usuario y llevarlas al “Product Backlog”
- ScrumMaster: Su función va a ser la de ayudar al equipo a alcanzar el objetivo en el sprint. No es un líder, pero si actuará de forma que evite distracciones del equipo.
- Equipo de desarrollo: Equipo entre 3 y 9 personas cuyas habilidades transversales permitan llevar correctamente el proyecto a cabo.



Por otro lado, se encuentran los Stakeholders (clientes, proveedores...) que no tienen una influencia directa en la creación del proyecto.

Todo esto conlleva unas ventajas las cuales son entre otras:

- Flexibilidad, reacción ante cambios.
- Uso del producto antes de su finalización.
- Aumento de calidad y productividad respecto a métodos tradicionales.
- Reducción de riesgos en el desarrollo.

En este proyecto se decide utilizar la metodología Scrum, principalmente por las distintas entregas que hay que realizar, las cuales se harán cada semana o cada quince días. En este proyecto habrá algunas fases en las que se deberá abandonar la implementación principal para llevar a cabo otra implementación secundaria, ya que así completaremos el funcionamiento de un método de toma de decisión o un índice ORI.

Las prioridades dentro del proyecto pueden variar entre una semana y otra, de ahí que se realicen reuniones todas las semanas para generar un nuevo sprint en el cual se observen todos los avances del proyecto y se establezcan nuevos objetivos. Además, al entregar periódicamente distintas partes de este proyecto, aseguramos que, en caso de existir algún error, este pueda subsanarse en la reunión del siguiente sprint.



UNIVERSIDAD
DE MÁLAGA





5. Diseño

En este capítulo se detalla el diseño que se le ha dado a la librería. Para cada método de toma de decisión/índice ORI se han investigado los artículos relacionados a cada uno de ellos.

La división del proyecto se va a realizar mediante una librería de toma de decisión multicriterio, la cual se dividirá en tres partes principales:

- Métodos de toma de decisión multicriterio
- Operadores de agregación
- Índices de reputación online (ORIs)

Las pruebas y test estarán separados en los directorios “Pruebas” y “Test”, dentro de cada uno habrá un archivo con las pruebas y test de cada índice por separado, las pruebas y test de cada apartado tendrán una pequeña explicación acerca del artículo del cual proceden.

Los extractores de datos para “Twitter” y “YouTube” irán incluidos en la carpeta “Extractors”.

Por último, se incluirá una carpeta “Semantics” para examinar textos y deducir su polaridad.

Métodos de toma de decisión multicriterio

La toma de decisión multicriterio es la disciplina que se encarga de elegir la mejor alternativa entre una serie de candidatos basándose unos de criterios. Cuando se realiza una toma de decisión se busca llegar a una solución óptima, pero eso no es posible salvo que solo haya un único criterio.



Un problema de este tipo es una situación en la cual una persona debe seleccionar una alternativa, pero no sabe cuál es la mejor. Estos procesos suelen ir englobados en tres fases:

1. Identificación del problema.
2. Elaboración y utilización de modelos.
3. Elaboración de plan de acción.

Los métodos de toma de decisión multicriterio se usan para ayudar a tomar decisiones de acuerdo con unas preferencias, encontrando la mejor opción entre las alternativas. No se debe confundir con la elección final, ya que el método no elige por la persona, sino que le aporta información numérica acerca de cada alternativa para facilitar la decisión.

Proceso Analítico Jerárquico (AHP)

El proceso analítico jerárquico fue desarrollado por Thomas L. Saaty a finales de los años 60. Thomas L. Saaty propone un método para tomar decisiones de una manera organizada y generando prioridades. Para ello necesitamos descomponer la decisión en los siguientes pasos (Thomas L. Saaty, 2008):

- Definir el problema y determinar el tipo de conocimiento que se busca.
- Estructurar la jerarquía de decisiones, empezando por el objetivo de la decisión, después criterios y por último alternativas.
- Construir las matrices de comparación. Los elementos de la parte superior se compararán con el nivel inmediatamente inferior, es decir, A_{01} se compara con A_{12} dentro de la matriz.
- Con las prioridades obtenidas se ponderan las del nivel inferior. Para el nivel inferior se suman los valores ponderados y se obtiene un global y se continua hasta llegar al nivel más bajo.

Una de las propiedades más importantes es que la información que se use sea fiable, es decir, sea consistente. En muchas ocasiones, cuando tenemos que tomar una

decisión, no usamos un criterio consistente, ya que, si tenemos que comparar entre varias alternativas, rellenar la matriz nos supone hacer un gran número de preguntas sobre preferencias y el cliente puede acabar por no ofrecer una información consistente. El ejemplo contrario lo encontramos al comparar dos alternativas ya que en este caso los juicios serán perfectos.

Una matriz recíproca es consistente si se cumple $a_{ij}a_{jk} = a_{ik}$, para todo $i, j, k = 1$ (Thomas L. Saaty, 2008).

Para que una matriz 3x3 sea perfecta debemos cumplir lo siguiente (J. I. Peláez & Lamata, 2003):

- $\det(M_{3 \times 3}) = \frac{a_{ik}}{a_{ij} \cdot a_{jk}} + \frac{a_{ij} + a_{jk}}{a_{ik}} - 2$
- Si los juicios son perfectos, entonces $a_{ik} = a_{ij} \cdot a_{jk}$ y $\det(M_{3 \times 3}) = 0$

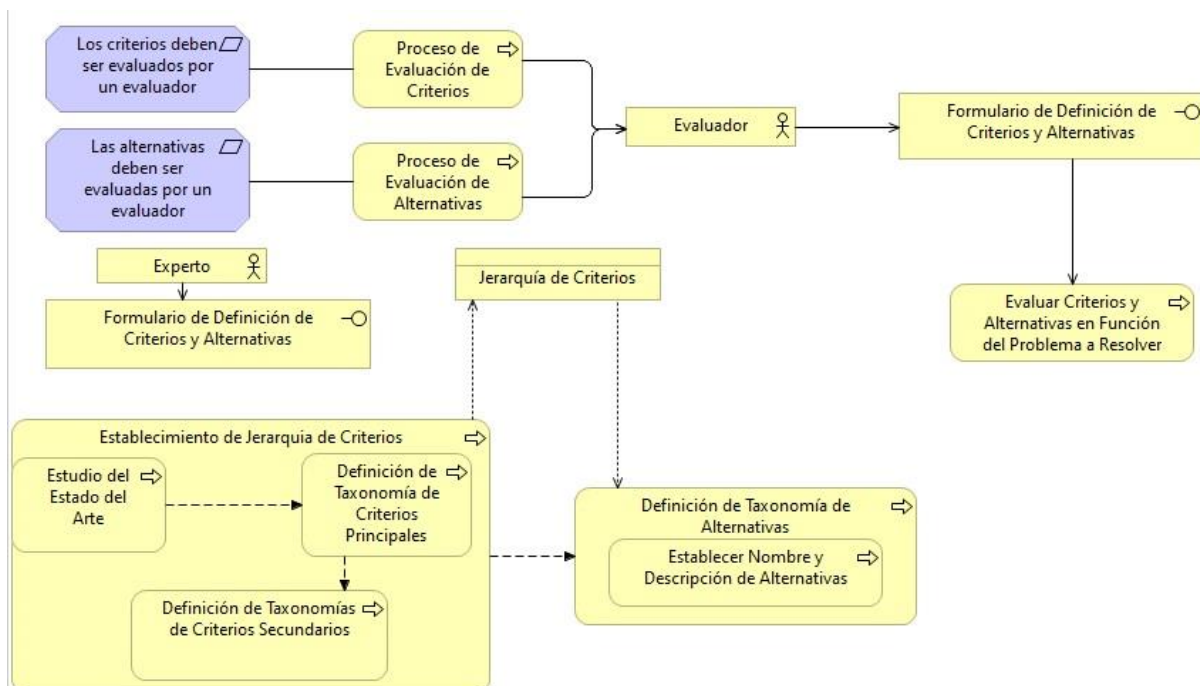


Figura 7. Funcionamiento del Método AHP.

Para hacer estas comparaciones se necesita una escala mediante la cual se determina la importancia un elemento sobre otro. Esta escala (Thomas L. Saaty, 2008):

| <i>Intensity of Importance</i> | <i>Definition</i> | <i>Explanation</i> |
|--------------------------------|---|---|
| 1 | Equal Importance | Two activities contribute equally to the objective |
| 2 | Weak or slight | |
| 3 | Moderate importance | Experience and judgement slightly favour one activity over another |
| 4 | Moderate plus | |
| 5 | Strong importance | Experience and judgement strongly favour one activity over another |
| 6 | Strong plus | |
| 7 | Very strong or demonstrated importance | An activity is favoured very strongly over another; its dominance demonstrated in practice |
| 8 | Very, very strong | |
| 9 | Extreme importance | The evidence favouring one activity over another is of the highest possible order of affirmation |
| Reciprocals of above | If activity i has one of the above non-zero numbers assigned to it when compared with activity j , then j has the reciprocal value when compared with i | A reasonable assumption |
| 1.1–1.9 | If the activities are very close | May be difficult to assign the best value but when compared with other contrasting activities the size of the small numbers would not be too noticeable, yet they can still indicate the relative importance of the activities. |

Figura 8. Escala de Saaty para el Método AHP (Thomas L. Saaty, 2008)

Los números impares desde 1 hasta 9 van en función de la importancia de un criterio sobre otro, siendo 1 igual importancia de ambos y siendo 9 la importancia absoluta de un criterio sobre otros. El significado de los números pares desde el 2 hasta el 8 es para dar una importancia media entre los números impares.

Otra de las características de este método es que la matriz es recíproca, es decir, el valor de la posición A_{12} va a ser inverso al valor de la posición A_{21} (Por ejemplo: 4 y $\frac{1}{4}$), y con todas las posiciones de la diagonal con valor 1 (no podemos dar importancia a una actividad sobre si misma).

Which drink is consumed more in the USA?

An example of examination using judgements

| <i>Drink consumption in US</i> | <i>Coffee</i> | <i>Wine</i> | <i>Tea</i> | <i>Beer</i> | <i>Sodas</i> | <i>Milk</i> | <i>Water</i> |
|--------------------------------|---------------|-------------|------------|-------------|--------------|-------------|--------------|
| Coffee | 1 | 9 | 5 | 2 | 1 | 1 | 1/2 |
| Wine | 1/9 | 1 | 1/3 | 1/9 | 1/9 | 1/9 | 1/9 |
| Tea | 1/5 | 2 | 1 | 1/3 | 1/4 | 1/3 | 1/9 |
| Beer | 1/2 | 9 | 3 | 1 | 1/2 | 1 | 1/3 |
| Soda | 1 | 9 | 4 | 2 | 1 | 2 | 1/2 |
| Milk | 1 | 9 | 3 | 1 | 1/2 | 1 | 1/3 |
| Water | 2 | 9 | 9 | 3 | 2 | 3 | 1 |

Note: The derived scale based on the judgements in the matrix is:

0.177 0.019 0.042 0.116 0.190 0.129 0.327

With a consistency ratio of 0.022.

the actual consumption (from statistical sources) is:

0.180 0.010 0.040 0.120 0.180 0.140 0.330

Figura 9. Creación de una matriz de comparación entre pares (Thomas L. Saaty, 2008).

Se aprecia en la imagen (Thomas L. Saaty, 2008) las propiedades explicadas anteriormente para la matriz de comparación de pares mediante escala de Saaty. Una vez generada la matriz ya se puede trabajar esta para el cálculo del método AHP.

Algunos ejemplos en los que el método AHP ha sido utilizado para la toma de decisiones multicriterio son (Thomas L. Saaty, 2008):

- El estado de Carolina del Norte lo usa para evaluar y asignar proveedores.
- El Departamento de Defensa de EEUU lo usa frecuentemente para asignar recursos a sus actividades.
- En 2001 se usa para elegir el mejor lugar para la reubicación de la ciudad turca de Adapazari después del terremoto que sacudió la ciudad.
- En 1998 lo usa British Airways para elegir el proveedor de sistemas de entretenimiento de todos sus aviones.
- En 1987 es usado por una empresa para elegir la mejor plataforma para perforar en busca de petróleo en el Atlántico Norte.

Método Harker para la Construcción de Matrices Consistentes

El Método AHP ha demostrado ser una ayuda a la decisión muy efectiva, pero las personas no suelen ser coherentes en sus juicios y el Método AHP no obliga a ser consistente a un individuo o un grupo. Además, AHP tiene el inconveniente de que mientras más aumenta el número de criterios más aumenta el número de comparaciones (Patrick T. Harker, 1987).

Cuando los juicios son perfectamente consistentes solo se harán $n-1$ comparaciones, sin embargo, cuando existen inconsistencias deben completarse $n(n-1)/2$ juicios (Patrick T. Harker, 1987).

Harker propone un método mediante el cual se pueden generar una matriz consistente si tener que recurrir a completar $n(n-1)/2$ juicios. Para ello, se va a implementar un método de cálculo de consistencia que, cuando se encuentra una posición de la matriz sin un juicio emitido, dará valor consistente a esa posición.

Posteriormente, y gracias a la transitividad, se calculan los valores que puede tener la posición que se desea calcular, por ejemplo, si se quiere hallar la posición (1,5) de una matriz M, se deben hacer las siguientes comparaciones:

- $M(1,4) * M(4,5) = M(1,5)$
- $M(1,3) * M(3,5) = M(1,5)$
- $M(1,2) * M(2,5) = M(1,5)$

Como se puede comprobar, no es necesario la comparación entre:

- $M(1,1) * M(1,5) = M(1,5)$, ya que estaríamos comparando con el valor a buscar.
- $M(1,5) * M(5,1) = M(1,1)$, ya que no estaríamos calculando la posición correcta.

Una vez encontrados todos los valores, se realiza una media geométrica de ellos y el valor que devuelve esa media geométrica es el que pertenece a la posición buscada (Patrick T. Harker, 1987), (P. T. Harker, 1987b), (P. T. Harker, 1987a).

Eigenvector Method for Pairwise Voting (EMPV)

El método EMPV es un método para la toma de decisiones multicriterio creado por Luis G. Vargas el cual se basa en la idea de votación por parejas para crear un ranking de las alternativas o candidatos preferidos por un grupo de votantes (Vargas, 2016)(Cabrera, Amaya, Vaccaro, & Peláez, 2019).

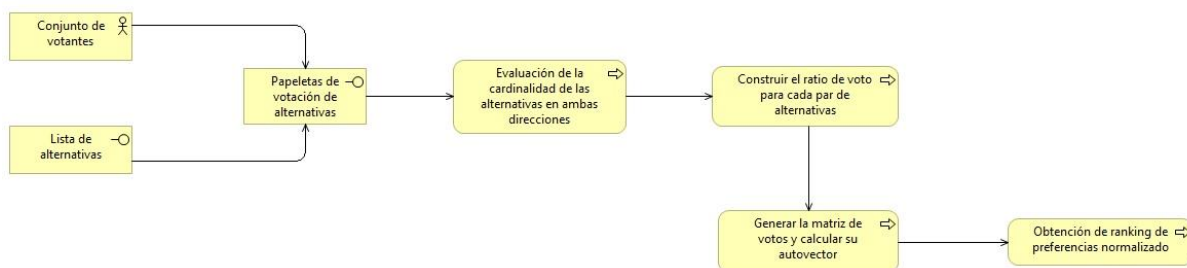


Figura 10. Modelado del Método EMPV.

Con este método se pretende que un número de votantes pueda establecer un orden de preferencias de los candidatos o alternativas. Para ello se necesitará una lista de alternativas y otra de votos emitidos.

El método EMPV se basa en un conteo de los votos emitidos por los votantes, un votante no tiene la obligación de votar por todas las alternativas de la lista. Supongamos que debemos decidir la ciudad a la que iríamos de viaje fin de grado y las alternativas son: Amberes, Roma y Dakar. Los votantes no tienen la obligación de introducir las tres alternativas en su papeleta de votación, podrían establecer:

- Dakar >> Roma
- Amberes
- Roma >> Amberes >> Dakar
- Roma >> Amberes
- Dakar >> Roma >> Amberes

Continuando con el conteo de los votos, este se desglosará usando permutaciones de tal forma que podamos ver los votos emitidos de cada permutación (para 3 alternativas tendríamos 6 permutaciones distintas) y su ranking.

Aun así, todavía se debe desglosar más cada permutación hasta obtener parejas de alternativas (en el caso del ejemplo son combinaciones de tamaño 2 dentro de la lista de alternativas), por lo que obtendríamos la siguiente tabla:

| Par de Alternativas | Cálculo de Alternativas | Preferido |
|----------------------------|-------------------------------|-----------|
| P1. Amberes > Roma | $r_{abc} + r_{acb} + r_{cab}$ | 1 |
| P2. Amberes > Dakar | $r_{abc} + r_{acb} + r_{bac}$ | 3 |
| P3. Roma > Amberes | $r_{bac} + r_{bca} + r_{cba}$ | 4 |
| P4. Roma > Dakar | $r_{abc} + r_{bac} + r_{bca}$ | 2 |
| P5. Dakar > Amberes | $r_{bca} + r_{cab} + r_{cba}$ | 2 |
| P6. Dakar > Roma | $r_{acb} + r_{cab} + r_{cba}$ | 2 |

Tabla 1. Cálculo de alternativas para el ejemplo EMPV anterior.

Una vez recogidos y clasificados los votos se crea una matriz de comparación por pares a la que se le hallará el autovector asociado al máximo autovalor y una vez se normalice el autovector se obtendrá un ranking de soluciones.

Este método se ha basado en otros métodos tales como Borda, Condorcet o Kemeny-Young.

Índices de Reputación Online

Índice QV-ORI

Este índice de reputación online (J. I. Peláez et al., 2019) devuelve una valoración con respecto a un servicio. La ventaja es el uso de valores discretos para obtener una valoración de un producto/servicio en una escala entre [-1, 1], siendo -1 una opinión negativa y 1 una opinión positiva.



La extracción de datos desde los social media puede ser a través de distintas alternativas que pueden incluir algoritmos que usen APIs, las cuales pueden ser

proporcionadas por la entidad correspondiente (Twitter, TripAdvisor...) o procesos de scraping. Estos análisis de la comunicación usan algoritmos de análisis semántico para extraer los sentimientos sobre un producto/servicio (J. I. Peláez et al., 2019) (Hu & Liu, 2004).

Los OWA tradicionales representan la opinión de la mayoría (J. Peláez & Doña, 2003) por lo tanto no tendremos una opinión completamente fundamentada. Por otro lado, usando las respuestas de los consumidores podemos encontrar verdades “distorsionadas”, es decir, respuestas no asemejadas a la realidad. Por ello, utilizar un único valor puede ser erróneo, y se propone usar un rango de valores. Para ello se propone la implementación de ISMA-OWA (J. I. Peláez et al., 2019), (J. I. Peláez et al., 2016), (J. Peláez & Doña, 2003), (J. Peláez & Doña, 2006), (Marcelo Karanik, Peláez, & Bernal, 2016), (M. Karanik, Pelaez, & Bernal, 2005).

Este índice a través de unas valuaciones y un parámetro delta usa un operador de agregación ISMA-OWA (Marcelo Karanik et al., 2016) para modelar la opinión de la mayoría. A diferencia del operador SMA-OWA devuelve un intervalo que representa la variación de las valuaciones de los consumidores y en función del parámetro delta contemplará la opinión de la mayoría ($\delta = 1$), minoría ($\delta = 0$) o media ($\delta = 0.5$).

Una vez obtenido el intervalo de opiniones, se realiza un índice de fiabilidad de intervalos (IRI), es decir, se busca que el intervalo obtenido sea fiable (J. I. Peláez et al., 2019).

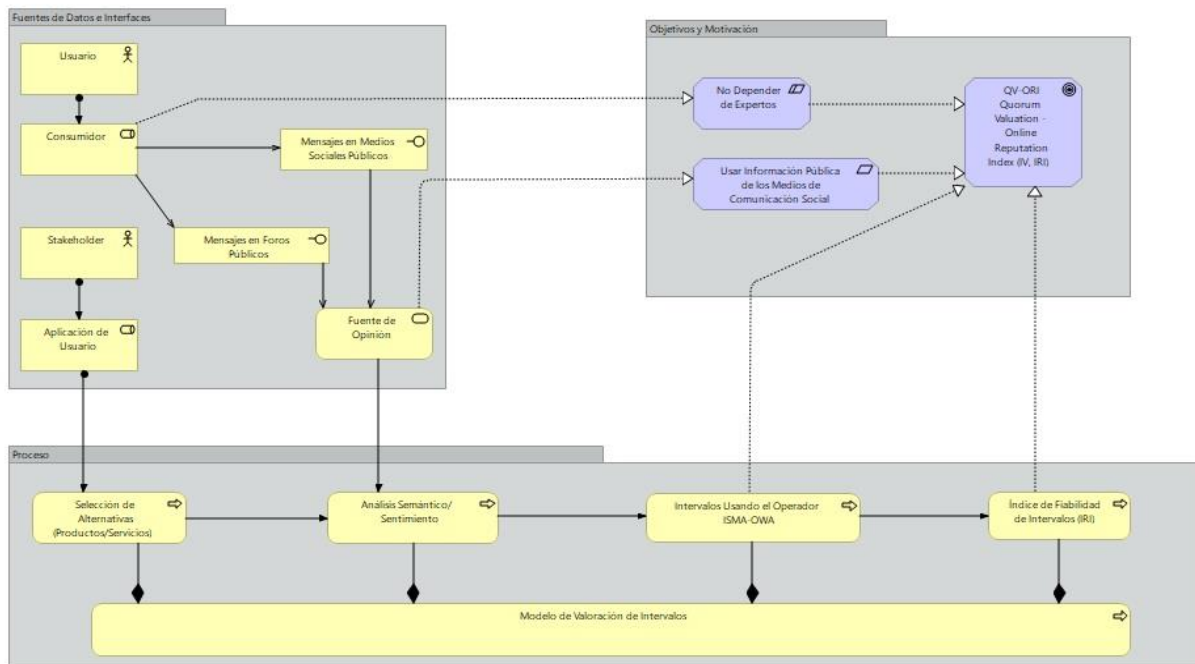


Figura 11. Modelado del Índice de Reputación Online QV-ORI (J. I. Peláez et al., 2019)

Reprtrak-Pulse

Reprtrak-Pulse es un indicador de reputación corporativa, este modelo permite medir la reputación corporativa de un grupo de Stakeholders. La reputación se construye en base a experiencias personales, comunicación de la empresa con el exterior y perspectivas de terceros (Ponzi, Fombrun, & Gardberg, 2011).

Para ello este modelo se basa en cuatro elementos (Ponzi et al., 2011):

1. Reputación general.
2. Buena impresión.
3. Confianza.
4. Admiración y respeto.

Para dar valor a estas dimensiones se usará una escala de 7 puntos, en la cual 1 será “Totalmente desacuerdo” y 7 será “Totalmente de acuerdo”. Por lo que aquellas personas entrevistadas deben estar familiarizadas con la empresa.

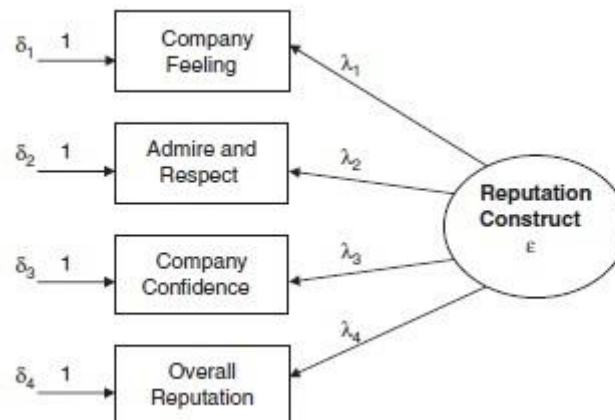


Figura 12. Modelado del método Reprtrak-Pulse (Ponzi et al., 2011)

Se propone una modificación del modelo original, el cual también se basará en los cuatro elementos expuestos anteriormente, sin embargo, el funcionamiento partirá de los siguientes elementos:

- 4 Dimensiones (Elementos) (Ponzi et al., 2011).
- Experiencias de usuarios con conocimientos sobre la empresa.
- Lista épsilon de experiencia general.

Para todas las respuestas sobre un elemento se realizará una correlación de Pearson con respecto al sumatorio de la puntuación total de cada experiencia.

$$\text{Corr. } a' \leftrightarrow t': \rho > p$$

Donde a' serán los valores de la dimensión a , mientras que el valor t' será $\sum_{i=0}^n a_i$ (ambos normalizados previamente). En el caso $\rho > 0.25$ y $p < 0,05$, la correlación será correcta. Se realizará esta comprobación para las 4 dimensiones, para posteriormente realizar estas correlaciones con respecto a épsilon.

Si los valores de correlación de Pearson son correctos, se realiza el cálculo del método EMPV para calcular el ranking de las alternativas.

$$S = \text{EMPV}(a) * a + \text{EMPV}(b) * b + \text{EMPV}(c) * c + \text{EMPV}(d) * d$$

La ecuación anterior obtiene el sentimiento en función del ranking de alternativas y las experiencias. Se realizará una nueva correlación entre los sentimientos obtenidos y épsilon.

$$S_T = \frac{\sum_{i=0}^n S}{n}$$

La fórmula anterior establece el sentimiento total para las alternativas y las dimensiones establecidas.

Índice de Sinergia CMGI

Este índice se presenta como una métrica cualitativa de reputación online, considera las emociones acumuladas y las comparaciones directas expresadas en los social media para inferir en las preferencias de los usuarios dentro de un grupo de alternativas. Es un índice muy útil para generar un ranking de alternativas cuando la cantidad de opiniones independientes sobre alternativas individuales supera las comparaciones por pares que emite la comunidad.

Se necesitan métodos eficaces que permitan medir tanto nuestra reputación, como la reputación de nuestros servicios y productos, además de medir la preferencia de los consumidores por nuestros productos sobre los productos de la competencia.

Para este índice hay que entender algunos conceptos:

- **Preferencia Positiva Fuerte:** Una preferencia polarizada positivamente de a_i sobre a_j es expresada explícitamente. Ejemplo: “La piscina del hotel i es mejor que la piscina del hotel j”.
- **Preferencia Positiva Débil:** Si se da una de las dos siguientes condiciones.
 - Se hace una evaluación positiva sobre a_i pero no es explícita sobre a_j .
 - No hay evaluación explícita sobre a_i , pero hay una evaluación negativa explícita sobre a_j .

Ejemplo: “El servicio del hotel j es muy bueno”

- **Preferencia Neutral:** La preferencia es neutral si se da uno de los siguientes casos:
 - Hay una evaluación positiva explícita sobre ambas alternativas.
 - Hay una evaluación negativa explícita sobre ambas alternativas.

- No hay evaluaciones explícitas sobre ambas alternativas, pero las hay sobre otras.

Ejemplo: El servicio del hotel j y del hotel i es excelente.

- Preferencia Negativa: Es la inversa de cualquiera de las preferencias positivas anteriores (fuerte o débil).

Este índice trata con las opiniones online de los usuarios (información no solicitada), a través de la cual lo que hacen los usuarios es comparar una serie de características de cada una de las alternativas.

El índice recibe un conjunto de opiniones sobre diferentes características acerca de un producto/servicio, las cuales estarán representadas por polaridades, siendo 1 una polaridad positiva, 0 una polaridad neutra y -1 una polaridad negativa.

Ejemplo: [0,1,0,-1] → La alternativa 2 tiene una preferencia positiva fuerte sobre la alternativa 4 (Se comparan 4 alternativas en este ejemplo).

Para cada característica se crea una Matriz de Preferencia de Opiniones (OPM), esta matriz representa la polaridad entre las distintas alternativas. Un ejemplo de matriz de este tipo es el siguiente:

| | | | |
|-------|-------|-------|----|
| np | np | np/wp | wp |
| np | np | np/wp | wp |
| wp | wp | np | sp |
| np/wp | np/wp | np/sp | np |

Tabla 2. Matriz de preferencia de opiniones.

Donde “np” es una opinión neutral, “wp” es una opinión positiva débil y “sp” es una opinión positiva fuerte.

De lo que se obtendrán tantas matrices como características se estén evaluando, pero para reducir esta cantidad de información se crea una Matriz de Preferencias Agregadas (APM). Esta matriz se obtiene realizando una media geométrica de cada una de las posiciones de todas las matrices OPM obtenidas.

Una vez se tiene la matriz APM, se busca el autovector asociado al máximo autovalor de dicha matriz mediante el método de Perron – Frobenius, y una vez que se obtiene este, se normaliza y se consigue el ranking de alternativas del índice CMGI.

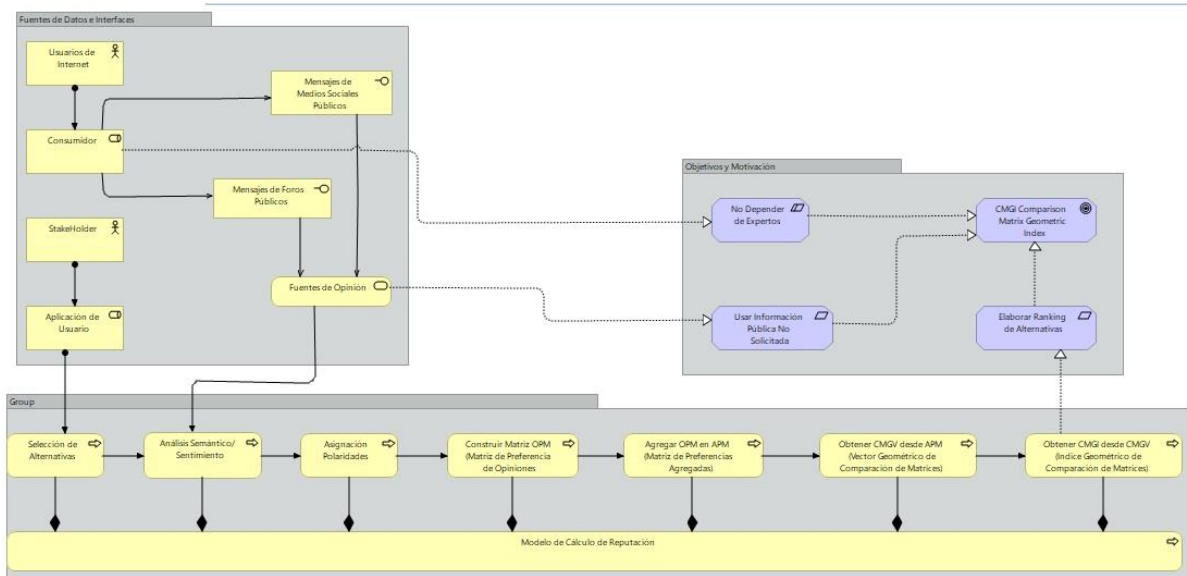


Figura 13. Modelado del índice de Sinergia CMGI.

Índice de Reputación por Intervalos

Este índice de reputación mediante intervalos para generar un ranking de alternativas (José I. Peláez et al., 2019). Uno de los problemas de los índices reputación online (ORIs) está en que varios de estos índices no consideran que los datos pueden ser emitidos desde el desconocimiento, lo cual puede dar lugar a decisiones mediante datos inconsistentes.

Únicamente el método AHP incluye un índice de inconsistencia para considerar la ignorancia en el proceso de toma de decisión, pero está diseñado para matrices recíprocas, si a esto le sumamos que muchos de esos métodos únicamente cogen un valor (a veces simplemente es la media) en vez de un intervalo, la fiabilidad desciende.

Para ello este índice va a usar un operador ISMA-OWA, que recibiendo un conjunto de valuaciones las transformará en un intervalo [ínfimo, supremo]. Posteriormente se

calculará un índice de consistencia intervalar (ICI+). Una vez obtenido comienza el proceso de deflación de matrices (José I. Peláez et al., 2019):

- Entrada: Preferencias Públicas
- Salida: Matriz deflacionada de Intervalos
 - Paso 1: Construir la matriz mediante un operador ISMA-OWA
 - Paso 2: Verificar la consistencia de la matriz mediante el índice ICI+
 - Paso 3: Si la matriz es consistente (\geq valor del percentil establecido), pasar al paso 5. En caso contrario, deflacionar intervalos sumando al ínfimo y restando al supremo un parámetro delta respectivamente, en caso de no funcionar ir al paso 4.
 - Paso 4: Cambiar el percentil de aceptación una vez transcurridas X operaciones (En la implementación son 1000).
 - Paso 5: Devolver la matriz deflacionada.

Una vez devuelta la matriz deflacionada, se le aplicará un operador AC-OWG mediante el cual se obtendrá el ranking final de alternativas con sus respectivos valores.

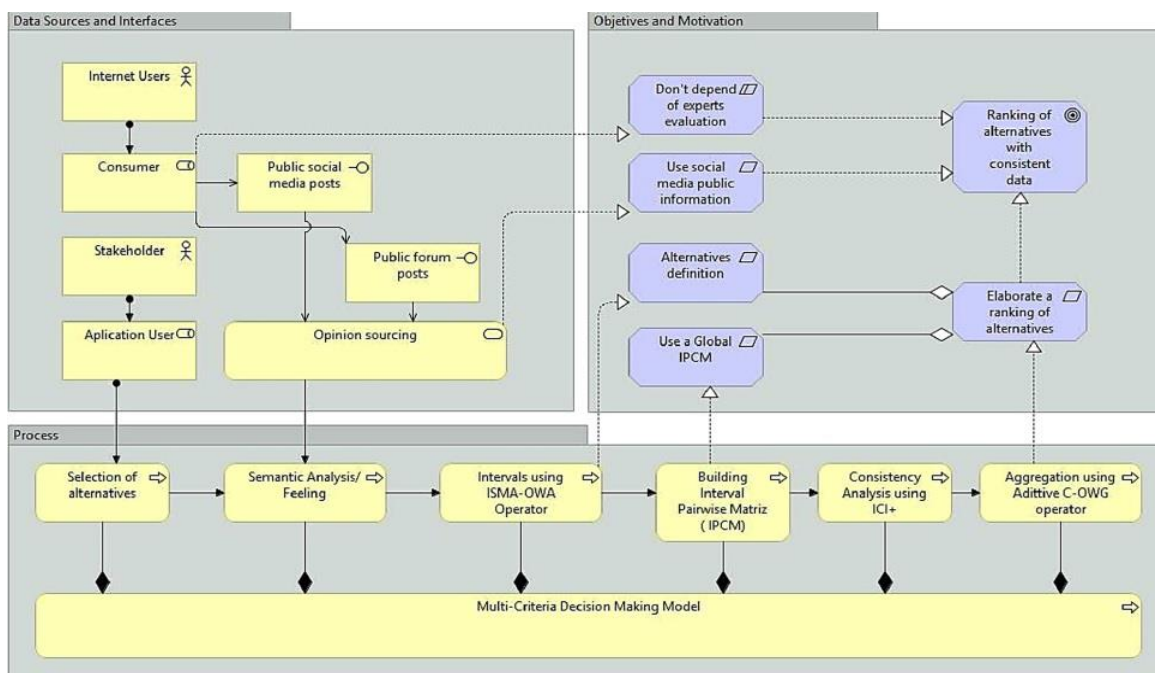


Figura 14. Modelado del Índice de Reputación Online mediante intervalos (José I. Peláez et al., 2019).



Índices de consistencia de datos

En este apartado se presentan los índices de consistencia de datos y ratios de consistencia de datos, los cuales permiten calcular la consistencia de las matrices generadas mediante los datos recibidos. Así se puede descubrir si estamos trabajando datos consistentes o por el contrario no hay una gran fiabilidad en los datos que tenemos.

Índice de consistencia Saaty

Este índice creado por Thomas L. Saaty en 1980 se basa en una solución mediante el autovector del máximo autovalor de la matriz y el orden de esta.

La fórmula de este índice de consistencia es la siguiente:

$$\text{Saaty CI} = (\lambda_{\max} - n) / (n-1)$$

Siendo n la dimensión de la matriz y λ_{\max} el autovector asociado al máximo autovalor de la matriz, en el caso de que la matriz sea consistente el índice devolverá el valor 0, en caso contrario devolverá un valor por encima de este (mientras más cercano a 0 más consistente serán los valores).

Ratio de consistencia de Saaty

La ratio de consistencia de Saaty aparece para normalizar el índice de consistencia de Saaty, para ello se va a usar la siguiente fórmula que relaciona el índice de consistencia de Saaty con la ratio del mismo:

$$\text{Saaty CR} = \text{CI} / \text{RI}(n)$$

Siendo CI el índice de consistencia de Saaty y RI(n) es un valor obtenido de la media de 10000 matrices de comparación aleatorias, siendo estos valores los siguientes (Pelaez, Martinez, & Vargas, 2018):

| n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| RI(n) | 0.52 | 0.88 | 1.11 | 1.25 | 1.34 | 1.40 | 1.48 | 1.50 | 1.51 | 1.53 | 1.55 | 1.57 | 1.58 | 1.59 |

Tabla 3. Valores de RI para matrices de dimensión entre 3 y 16 (Pelaez et al., 2018)

Para aquellas matrices con dimensión menor que 3 el valor es 0. Generalmente, para matrices de dimensión mayor que 9 no suele usarse, ya que decidir entre 9 alternativas conlleva un proceso muy largo y extenso con una cantidad de operaciones muy elevadas.

La lectura que se obtiene es que si el valor que devuelve CR es menor a 0.10 significa que el decisor está aportando datos consistentes para la toma de decisión (Pelaez et al., 2018)(J. I. Peláez & Lamata, 2003).

Índice de consistencia CI*

El índice de consistencia CI* (J. I. Peláez & Lamata, 2003) introduce el concepto de transitividad. Este concepto engloba la relación entre tres alternativas siendo la transitividad el menor elemento de la consistencia. La transitividad asegura que se cumple el Axioma de Arrow de alternativas irrelevantes, se asegura así que eliminando una alternativa no se produzca inversión de rango y asegura el dominio de fila en la matriz recíproca.

Para entenderlo mejor se tiene la siguiente imagen:

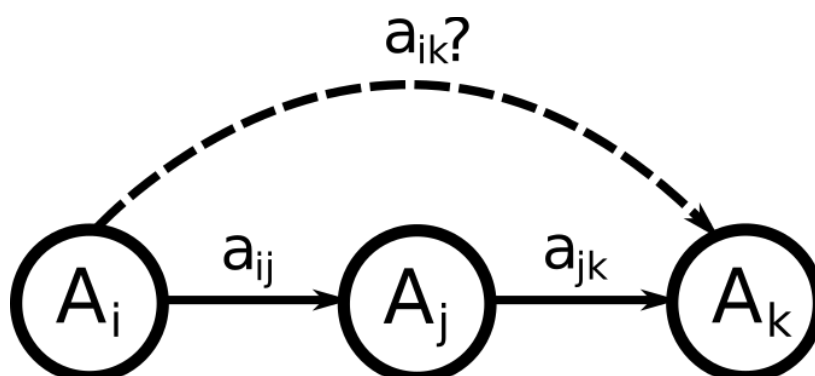


Figura 15. Concepto de transitividad (Γ)

Dos transitividades van a ser distintas si difieren en al menos un elemento, por lo que para AHP el menor elemento de consistencia será una matriz recíproca 3x3.

Una vez tenemos una comparación de tres alternativas, es muy sencillo entender cuando existe una inconsistencia entre ellas. Existe una inconsistencia si ocurre lo siguiente:

- Existe un ciclo entre las alternativas ($a_i > a_j > a_k > a_i$)
- a_{ik} es distinto al producto de $a_{ij} * a_{jk}$

| M | A | B | C |
|---|--------------------|--------------------|----------|
| A | 1 | M_{ab} | M_{ac} |
| B | $\frac{1}{M_{ab}}$ | 1 | M_{bc} |
| C | $\frac{1}{M_{ac}}$ | $\frac{1}{M_{bc}}$ | 1 |

Tabla 4. Matriz recíproca de comparación entre pares.

Como bien se indicaba en apartados anteriores, la matriz recíproca de comparación de pares es perfecta si y solo si el determinante es 0. Aplicando la fórmula:

$$\det(M_{3 \times 3}) = \frac{a_{ik}}{a_{ij} * a_{jk}} + \frac{a_{ij} + a_{jk}}{a_{ik}} - 2$$

Otro dato que debemos tener en cuenta es que, para cada matriz, en función de su dimensión, habrá un número de transitividades. Para poder medir la consistencia de una matriz NxN debemos hallar el número de transitividades, para lo cual se expone la siguiente fórmula:

$$NT(M_{n \times n}) = \begin{cases} 0, & n < 3 \\ n! \\ \frac{n!}{(n-3)!3!}, & \text{otherwise} \end{cases}$$

Figura 16. Cálculo del número de transitividades de una matriz NXN.

Una vez se tienen estos conceptos, se define un índice de consistencia llamado CI^* . Para calcular este índice se propone la siguiente fórmula:

$$CI^*(M_{n \times n}) = \begin{cases} 0, & n < 3 \\ \det(M_{n \times n}), & n = 3 \\ \frac{1}{NT(M_{n \times n})} \sum_{i=1}^{NT(M_{n \times n})} CI^*(\Gamma_i), & n > 3 \end{cases}$$

Figura 17. Cálculo del índice de consistencia CI^* (J. I. Peláez & Lamata, 2003).

Siendo:

- $\det(M_{n \times n})$ el determinante de la matriz.
- $NT(M_{n \times n})$ el número de transitividades diferentes.

Índice de consistencia CI^+

El siguiente índice tiene como objetivo una acotación de los resultados entre 0 y 1. Otra novedad con respecto al índice anterior es que será aplicable a diferentes escalas. Se repite el concepto de transitividad como menor elemento de consistencia.

Las escalas pueden ser aditiva, Saaty y difusa, para cada una habrá un método distinto de cálculo de la variable Phi (Φ) que será la siguiente:

- Escala Saaty: 531442/729
- Escala Aditiva y Difusa: 100.001

Para el cálculo de este índice en una matriz 3x3 se presenta la fórmula (Pelaez et al., 2018):

$$CI^+(A_{3 \times 3}) = \frac{\Phi - \left(\left| \frac{a_{13}}{a_{12} \cdot a_{13}} \right| + \left| \frac{a_{12} \cdot a_{23}}{a_{13}} \right| \right)}{\Phi - 2}$$

Para una matriz de dimensión $N \times N$ tendremos la siguiente fórmula:

$$CI^+(A_{n \times n}) = \begin{cases} 0, & n < 3 \\ CI^+(A_{n \times n}), & n = 3 \\ \frac{1}{\Omega(A_{n \times n})} \sum_{i=1}^{\Omega(A_{n \times n})} CI^+(\gamma_i), & n > 3 \end{cases}$$

Figura 18. Cálculo de CI+ para una matriz (Pelaez et al., 2018).

Siendo γ_i la i -ésima transitividad de la matriz $A_{n \times n}$ y siendo $\Omega(A_{n \times n}) = (n-3)!3!/n!$ el número de las distintas transitividades de la matriz $A_{n \times n}$.

La siguiente tabla muestra los valores para los que CI+ es consistente si es igual o mayor:

| n | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|----------|----------|----------|----------|----------|----------|----------|
| δ | 0.9985 | 0.9983 | 0.9982 | 0.9982 | 0.9978 | 0.9977 | 0.9976 |

Tabla 5. Valores para los que CI+ es consistente (Pelaez et al., 2018).

| n | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|----------|----------|----------|----------|----------|----------|----------|
| (%) | 22.06 | 3.36 | 0.2526 | 0.0074 | 0* | 0* | 0* |

Tabla 6. Porcentaje de matrices consistentes por dimensión, para CR = 0.1. 0* está por debajo de 0.007% (Pelaez et al., 2018).

Ratio de consistencia CR+

Este valor establece cuando una matriz es consistente, los valores son menores o iguales a los que aparecen en la siguiente tabla:

| n | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|----------|----------|----------|----------|----------|----------|----------|
| CR+ | 1.3834 | 1.5815 | 1.3677 | 1.2713 | 1.2130 | 1.2088 | 1.1657 |

Tabla 7. Valores para que la ratio CR+ sea consistente.

Aquella matriz cuya ratio sea mayor al número de la figura 5.16 no se aceptará como matriz consistente.

Índice de Consistencia Intervalar (ICI+)

Este índice está basado en el índice de consistencia CI+. Este índice va a dividir las matrices en matriz superior M^U y matriz inferior M^L .

La matriz superior tendrá diagonal 0.5 ya que se usa escala aditiva y la transitividad es de la forma $m_{ik} = m_{ij} - 0.5 + m_{jk}$. Para la parte superior a la diagonal se usará los valores obtenidos para la matriz, mientras que para la parte inferior se usará 1 menos el valor de la parte superior.

La matriz inferior será 0.5 en la diagonal, pero, sin embargo, será la parte inferior la que mantenga los números obtenidos para la matriz y la parte superior la que tenga la resta de 1 menos el valor de la parte inferior.

Para el cálculo de este índice se usará la siguiente fórmula:

$$ICI^+(M) = \begin{cases} 0, & n < 3 \\ \min\{CI^+(M^L), CI^+(M^U)\}, & n = 3 \\ \min\left\{\frac{1}{\Omega(M_{n \times n}^L)} \sum_{i=1}^{\Omega(M_{n \times n}^L)} CI^+(\gamma_i^L), \frac{1}{\Omega(M_{n \times n}^U)} \sum_{i=1}^{\Omega(M_{n \times n}^U)} CI^+(\gamma_i^U)\right\}, & n > 3 \end{cases}$$

Figura 19. Cálculo del índice ICI+ para matrices de dimensión NxN (José I. Peláez et al., 2019).

Por lo tanto, si $0 \leq CI^+(M) \leq 1$, la matriz será consistente si $\delta^* \leq ICI^+(M) \leq 1$, siendo δ^* el valor crítico de aceptación obtenido en la figura 5.17.

Índice de Fiabilidad de Intervalos (IRI)

Este índice evalúa la fiabilidad de las valuaciones. Para determinar las opiniones/sentimientos más fiables que otros es necesario este índice que mide la similitud entre límite superior y límite inferior de un intervalo.

Se realizará mediante la siguiente fórmula (J. I. Peláez et al., 2019):

$$IRI = \frac{|X^U| - |X^L|}{|X^U| + |X^L|}$$

Donde $-1 \leq X^L \leq 0$ y $0 \leq X^U \leq 1$, no se puede dar el caso en el que X^L y X^U sean iguales a 0.

Hay varios casos en los que el IRI tiene un valor fijo (J. I. Peláez et al., 2019):

- Si $X^L = 0$ y $X^U \neq 0 \rightarrow IRI = 1$
- Si $X^L \neq 0$ y $X^U = 0 \rightarrow IRI = -1$
- Si $X^L = X^U \rightarrow IRI = 0$
- Si $X^L \neq 0$ y $X^U \neq 0$ y además $X^L < X^U \rightarrow 0 < IRI < 1$
- Si $X^L \neq 0$ y $X^U \neq 0$ y además $X^L > X^U \rightarrow -1 < IRI < 0$

El valor del IRI cercano a 1 o -1 será más fiable ya que habrá una menor polarización de las valoraciones.

Operadores de agregación

Los operadores de agregación son fundamentales tanto en métodos de toma de decisión como en índices de reputación online. Un operador es una función a la que se le asocia un vector y devuelve un valor que será interpretado para un problema concreto. Dentro de estos operadores hay dos familias principales que son OWA y OWG, en las que se usa la media aritmética y la media geométrica respectivamente.

Los operadores básicos son: Media Aritmética, Media Geométrica y Media Armónica.

Operadores Básicos: Media Aritmética, Geométrica y Armónica.

Estos operadores, comúnmente llamados medias, son la media aritmética clásica y su ponderación con pesos, la media geométrica y la media armónica (Bernal, 2015).

Majority Additive Ordered Weighting Averaging Operator (MA-OWA)

El operador MA-OWA es una modificación de la media aritmética por lo que el resultado final es una media aritmética ponderada (J. Peláez & Doña, 2003). Está basado en el proceso de mayoría.

Este operador está calificado como neat OWA, sus pesos están calculados en función de la cardinalidad de los elementos y son independientes del orden.

Su funcionamiento es el siguiente (J. Peláez & Doña, 2003)(J. I. Peláez et al., 2016):

1. Se recibe una lista de valuaciones y se seleccionan las valuaciones con diferentes valores dentro de la lista. [0.3, 0.4, 0.4, 0.6, 0.6]
2. Cada valor único se divide entre el número de elementos de la lista de valores únicos y se suma. $(0.3/3 + 0.4/3 + 0.6/3)$ y se eliminan de la lista anterior.
3. Se repite el paso 1.
4. Se obtiene por un lado otra lista de valores únicos [0.4, 0.6] y se realiza la suma del paso 2, aparte se añade un 1 ya que el número 0,3 ya no aparece en la lista de valores $(0.4/3 + 0.6/3 + 1)$.
5. Se calcula un denominador cuya fórmula es $1 + \text{tamaño de la lista de únicos}$.
6. Se divide los valores obtenidos en los dos puntos anteriores y se repite el proceso hasta que la lista de valuaciones quede vacía.

Selective Majority Additive Ordered Weighting Averaging Operator (SMA-OWA)

Basado en el operador anterior (MA-OWA), este operador tiene en cuenta la opinión de una minoría en función de un parámetro delta (δ) (Marcelo Karanik et al., 2016).

Este operador aplicará el parámetro delta a aquellas valuaciones que sean únicas, mientras que aplicará $(1-\delta)$ a aquellas valuaciones que dejen de pertenecer a la lista de valuaciones únicas.

Cuando $\delta = 1$, solo se considerará la opinión de la mayoría, mientras que si el valor de $\delta = 0$ solo se considerará la opinión de la minoría, si el valor de δ queda en 0.5 se estará calculando una media aritmética.

Interval Selective Majority Additive Ordered Weighting Averaging Operator (ISMA-OWA)

Este operador, perteneciente a los operadores de mayoría, se desarrolla como operador de intervalos (José I. Peláez et al., 2019) (J. I. Peláez et al., 2019). Se construye desde un conjunto de valuaciones y sus cardinalidades.

$$F_{ISMA}(r) = [F_{SMA_{\sim}}(r_{\sim}), F_{SMA^{\sim}}(r^{\sim})]$$

En la fórmula anterior, se aplica un SMA-OWA inferior y otro superior para generar el intervalo, aparte este operador también depende de un parámetro δ .

Su funcionamiento se basa en recibir un conjunto de valuaciones y sus cardinalidades, para las cuales calculará el valor de la mediana, ordenará en dos listas los valores mayores o iguales a la mediana y en la otra lista los menores. Se aplica la fórmula anterior a cada uno de los valores, después de esto se genera el intervalo.

Continuous Ordered Weighted Geometric (C-OWG)

Basándose en el operador C-OWA (Continuous Ordered Weighted Averaging) (Yager, 1988), (Yager, 2004), Yager desarrolla un operador geométrico muy útil para trabajar con una colección finita de argumentos. (Yager & Xu, 2006). Para ambos operadores aparece una función (BUM).

La función de unidad básica monótona irá definida por la siguiente fórmula:

$$\lambda = \int_0^1 Q(y) dy$$

La familia OWA se basa en la siguiente fórmula, que, dado un vector de pesos con valores entre 0 y 1, además de $\sum_{j=1}^n w_j=1$, se obtiene: $f(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{j=1}^n w_j \beta_j$ siendo β_j el j-ésimo mayor valor de α_i (Yager & Xu, 2006).

A raíz de esta fórmula surge la siguiente, que, basándose en los operadores OWA y en la media geométrica da lugar a la familia de operadores OWG, dado un vector de pesos con valores entre 0 y 1 y $\sum_{j=1}^n w_j=1$, se obtiene: $g(\alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{j=1}^n \beta_j^{w_j}$ siendo β_j el j-ésimo mayor valor de α_i (Yager & Xu, 2006).

Como otros operadores recibe una matriz de preferencias, pero habrá un intervalo en cada posición de la matriz, ya que va a trabajar con un ínfimo y un supremo. Para esto, se desarrolla la siguiente fórmula, mediante la cual podremos calcular el grado de preferencia de cada alternativa sobre las demás mediante la media geométrica:

$$g_Q(\tilde{a}_{ij}) = g_Q([a_{ij}^-, a_{ij}^+]) = a_{ij}^- * \left(\frac{a_{ij}^+}{a_{ij}^-}\right)^\lambda, \text{ para } i > j$$

$$g_Q(\tilde{a}_{ji}) = \frac{1}{a_{ij}^-} * \left(\frac{a_{ij}^-}{a_{ij}^+}\right)^\lambda, \text{ para } i \leq j$$

Una vez aplicadas estas fórmulas, se aplica la media geométrica por filas:

$$g_Q(\tilde{a}_i) = \left(\prod_{j=1}^n g_Q(\tilde{a}_{ij})\right)^{1/n}, i = 1, 2, \dots, n.$$

Additive Continuous Ordered Weighted Geometric (AC-OWG)

Este operador es una modificación del operador C-OWG (Yager & Xu, 2006). Se presenta como un operador para calcular el ranking de alternativas de una matriz recíproca de comparación entre pares (José I. Peláez et al., 2019).

Su funcionamiento se basa en aplicar una fórmula como la siguiente a cada punto de la matriz:

$$\varphi_Q(m_{ij}) = \varphi_Q([m_{ij}^L, m_{ij}^U]) = m_{ij}^L \left(\frac{m_{ij}^U}{m_{ij}^L} \right)^\lambda$$

También se basará en una BUM (Unidad Monotónica Básica) que se verá reflejada en λ , ya que este símbolo será $\lambda = \int_0^1 Q(y) dy$.

Una vez calculada la fórmula anterior para cada punto de la matriz se realiza la media geométrica de las filas de la matriz:

$$AC-OWG(m_i) = \left(\prod_{j=1}^n \varphi_Q(m_{ij}) \right)^{1/n}$$

Para $i = 1, 2, 3, \dots, n$

Una vez realizada la media geométrica por filas se obtendrá un ranking de alternativas para la matriz que hayamos introducido.

Análisis semántico de texto

Dentro de la librería se incluye un analizador semántico de texto el cual, bien a través de un algoritmo Naive-Bayes o bien a través de una red neuronal, recibe un texto el cual lo polarizará entre -1 y 1 (siendo el texto negativo o positivo respectivamente).

Naive-Bayes

El algoritmo de Naive-Bayes es uno de los algoritmos usados para el entrenamiento de un analizador semántico de texto. Se trata de un clasificador Bayesiano ingenuo ya que está fundamentado en el teorema de Bayes.

Este algoritmo basa su funcionamiento en usar la regla de Bayes para expresar:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$

Una vez realizada esta regla el siguiente paso es realizar la suposición de Naive la cual supone que todas las características son independientes:

$$P(A_i|B) = \frac{P(A_i) * P(B_1|A_i) * \dots * P(B_n|A_i)}{P(B)}$$

Por último, calcula el numerador para cada etiqueta y las normaliza:

$$P(A_i|B) = \frac{P(A_i) * P(B_1|A_i) * \dots * P(B_n|A_i)}{\text{Sum}[l](P(l) * P(B_1|l) * \dots * P(B_n|l))}$$

Red Neuronal

Este sistema consta de dos partes, por un lado, existe un entrenador (Train) que recibe un conjunto de comunicaciones debidamente etiquetadas (corpus de entrenamiento) y mediante un proceso de entrenamiento genera un clasificador y lo guarda como un “Pickle” de Python para poder ser utilizado a posteriori. Este proceso es computacionalmente costoso (tanto en tiempo como en memoria) por lo que se incluye en el proyecto un clasificador ya entrenado y solo se recomienda un nuevo entrenamiento en caso de que se desee usar otro corpus de entrenamiento.

El clasificador permite predecir la polaridad de una comunicación mediante dos métodos que se encuentran en “semanticmethods.py”, según se desee clasificar un único texto o un conjunto de los mismos.

Captura de datos a través de páginas web

Las páginas web de acceso público son una de las fuentes de información más utilizadas en la toma de decisiones, tanto a nivel empresarial, así como por los consumidores. Destacan las páginas web que contienen opiniones de personas que han podido probar dicha actividad, también llamados “consumer reviews”.



En varios casos es posible utilizar las APIs provistas por dichas webs para poder extraer esta información; sin embargo, en la mayoría de las ocasiones se hace necesario emplear técnicas de captura activa de información (web scraping).

A partir de los datos recolectados de estas páginas web se puede calcular la polaridad de los comentarios, y posteriormente trabajar dichas polaridades con los índices reputacionales indicados anteriormente.

Twitter y YouTube

Twitter es una red social gratuita que permite enviar mensajes de texto con un máximo de 280 caracteres. Esta red social provee a los usuarios académicos e investigadores con acceso especial a la API, por medio de la cual se puede recolectar comunicaciones (tweets) que contengan la palabra clave deseada. Es una herramienta clave en la recolección de opiniones de los usuarios de internet a nivel mundial.

Por otra parte, YouTube es un sitio web dedicado a compartir videos, y que provee de herramientas especiales y APIs para la captura de comentarios de texto asociados a videos públicos. Desde el 2016, YouTube se ha convertido en una fuente significativa de comentarios, superando a Twitter en ciertas temáticas especialmente relacionadas con influenciadores y personajes públicos.

6. Implementación

En este apartado se van a generar unas listas de métodos divididos por clases y dentro de cada método se observará su funcionamiento, sus parámetros de entrada y su salida para comprender el correcto funcionamiento de la librería presentada en este proyecto.

Se va a presentar la siguiente división:

- Clase Ahp: Cálculo del Proceso Análítico Jerárquico (Thomas L. Saaty, 2008).
 - Clase AhpConsistency: Cálculo de índices y ratios relacionados con AHP.
- Clase Cmgj: Incluye métodos para calcular el Índice de Sinergia Cmgj.
- Clase Empv: Incluye métodos para calcular el método de toma de decisión multicriterio EMPV. (Vargas, 2016)(Cabrera et al., 2019)
- Clase Interval: Incluye métodos e índices para trabajar con intervalos.
- Clase Operators: Incluye los métodos que implementan los operadores usados en este trabajo.
- Clase Qvori: Incluye métodos para calcular el índice de reputación Qvori. (J. I. Peláez et al., 2019)
- Clase Reprakpulse: Incluye métodos para calcular una versión de Reprak-Pulse a través de Empv.
- Clase Harker: Incluye los métodos para calcular el método de Harker para predicción de matrices consistentes (Patrick T. Harker, 1987).

El parámetro “Self” se corresponde con una referencia a la propia clase, por lo que no aparecerá en la columna “Parámetros”. Aquellos métodos privados que se hayan

implementado no aparecerán tampoco. En la implementación se describe cada método individualmente.

Clase Ahp

| Método | Parámetros | Devuelve |
|----------------------|---------------------|---------------------------|
| get_criteria_weights | | Autovector normalizado |
| Ahp | Alternativa (Lista) | Valor de cada alternativa |

Clase AhpConsistency

| Método | Parámetros | Devuelve |
|----------------------|-----------------------------------|---|
| factorial | Valor numérico | Factorial |
| get_autovalue | | Máximo autovalor de la matriz. |
| ci | | Índice de consistencia de Saaty |
| cr | | Ratio de consistencia de Saaty |
| Recursive_decomposer | Matriz | Lista de matrices 3x3 |
| ci_star | | Índice de consistencia CI* |
| ci_plus | Escala (por defecto = 'Saaty') | Índice de consistencia CI+ y aceptación |
| cr_plus | | Aceptación del ratio (booleano) |

Clase Cmgj

| Método | Parámetros | Devuelve |
|---------------------|----------------------------|------------------------|
| calculateOPM | Lista polaridades (-1,0,1) | Matrices OPM |
| calculateAPM | Lista polaridades (-1,0,1) | Matriz APM |
| calculatenormalized | Lista polaridades (-1,0,1) | Autovector normalizado |
| calculatecmgj | Lista polaridades (-1,0,1) | Ranking alternativas |

En este caso, los parámetros siempre son una lista de listas. Para cada lista sus valores serán -1 (Polaridad negativa), 0 (Polaridad neutra), 1 (Polaridad positiva). Esto se debe a que cada método llamará a su anterior para realizar los cálculos necesarios.

Clase Empv

| Método | Parámetros | Devuelve |
|---------------|-------------------|-------------------------|
| empv | | Ranking de alternativas |

Clase Interval

| Método | Parámetros | Devuelve |
|--------------------------|---------------------------------|--|
| Recursive_decomposer | Matriz | Lista de matrices 3x3 |
| matrix_upper | Matriz de intervalos | Matriz suprema sin intervalos |
| matrix_lower | Matriz de intervalos | Matriz ínfima sin intervalos |
| ici_plus | | Índice de consistencia de intervalos |
| Interval_reability_index | Intervalo | Índice de fiabilidad de intervalos |
| Infla_defla_matrix | Matriz de intervalos | Matriz deflacionada |
| matrices_IPCM | Matriz de intervalos | Matriz IPCM y percentil usado |
| Interval_mcmd | Matriz de intervalos, lambda | Ranking de alternativas para matrices de intervalos. |

Clase Harkerprediction

| Método | Parámetros | Devuelve |
|---------------|-------------------------------|--|
| prediction | Matriz Posición a calcular | Matriz con posición indicada calculada |

Clase Operators

| Método | Parámetros | Devuelve |
|---------------------------|---------------------------------|--|
| calculate_arithmetic_mean | Lista de valores | Media aritmética |
| calculate_geometric_mean | Lista de valores | Media geométrica |
| calculate_harmonic_mean | Lista de valores | Media armónica |
| calculate_sma_owa | Lista de valores, delta | Valor operador sma-owa |
| calculate_ma_owa | Lista de valores (delta=1) | Valor operador ma-owa |
| calculate_isma_owa | Lista de valores, delta | Valor operador isma-owa (intervalo) |
| c_owg | Matriz de intervalos, lambda | Ranking de alternativas |
| ac_owg | Matriz de intervalos, lambda | Ranking de alternativas |

Clase Qvori

| Método | Parámetros | Devuelve |
|--------|---------------------------------|--|
| qv_ori | Valuación (intervalo), delta | Valor del índice qv-ori sobre una alternativa |

Clase Reprakpulse

| Método | Parámetros | Devuelve |
|-----------------------|--|---|
| calculate_sentiments | Dimensiones, Experiencias (Lista) | Valor de los sentimientos |
| calculate_reprakpulse | Dimensiones, Experiencias (Lista), Épsilon (Lista) | Valor del Reprak-Pulse (Media de sentimientos) y correlación entre sentimientos y épsilon. |

Analizador Semántico

En este apartado se detallará la implementación de las partes principales del analizador semántico explicado anteriormente. Lo primero hay que saber es que el

entrenamiento de una red neuronal es muy costoso y para ello se incluye “semantic_classifier.p”, que será el clasificador previamente entrenado.

Clase LazyTrainer

| Método | Parámetros | Devuelve |
|-------------|-------------------|-------------------------------------|
| stem_tokens | Tokens, stemmer | Raíces de los tokens |
| tokenize | Texto | Tokens con x raíces |
| Train | Textos, objetivos | Clasificador semántico entrenado |

Clase SentimentClassifier

| Método | Parámetros | Devuelve |
|------------------------|---|-----------------------|
| __init__ (Constructor) | Locale (idioma), clasificador de polaridad, métricas de rendimiento | ----- |
| Classify | Texto | Polaridad de un texto |

Clase ClassifierFactory

| Método | Parámetros | Devuelve |
|------------------------|--|-------------------------------------|
| __init__ (Constructor) | Locale (idioma), frases del corpus, polaridad del corpus | ----- |
| Train | | Clasificador semántico entrenado |

Clase Semanticmethods

| Método | Parámetros | Devuelve |
|--------------------|-------------------|--|
| Classify_text | Texto | Polaridad del texto (-1,0,1) |
| Classify_text_list | Lista de textos | Polaridades de una lista de textos (-1,0,1) |

Captura de datos

Para la implementación de la captura de datos mediante scraping, se han implementado métodos similares tanto en Twitter, como en YouTube. En el caso de Twitter será necesario una clave aportada dicha empresa, mientras que en el caso de YouTube además se necesitará una cuenta de Google y acceso a la API de YouTube. Únicamente se realiza la extracción del comentario para su posterior procesamiento a través del analizador semántico.

7.Pruebas

Dentro de la implementación de los métodos, índices y operadores se han realizado pruebas a cada uno de ellos, siempre comparando resultados con los de los artículos utilizados para realizarlos. Las siguientes capturas demuestran que las soluciones de los test de Unittest creados a partir de ejemplos de artículos científicos corresponden con sus respectivos ejemplos en la librería y estos se ejecutan correctamente y obtienen los resultados esperados.

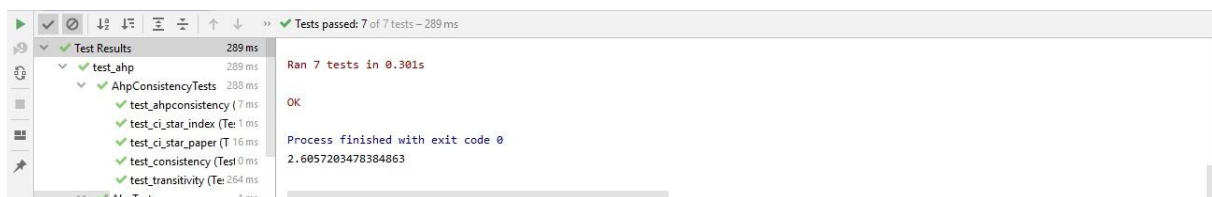


Figura 20: Test clase Ahp y AhpConsistency

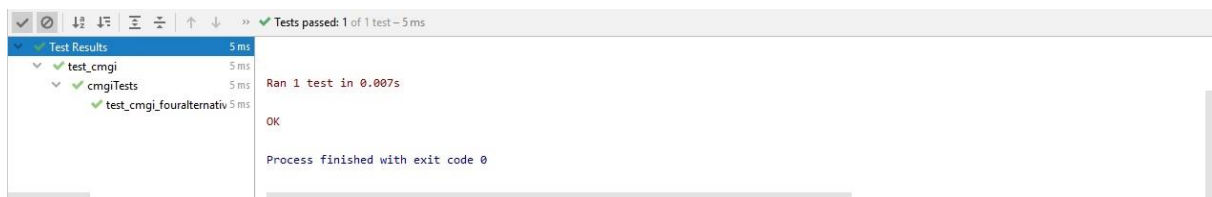


Figura 21. Test clase Cmgj.

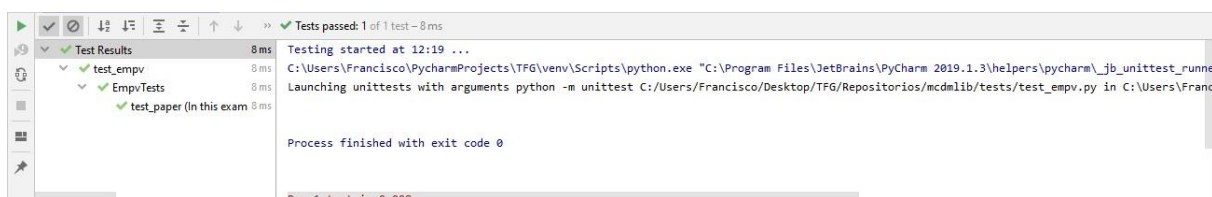


Figura 22. Test clase Empv.

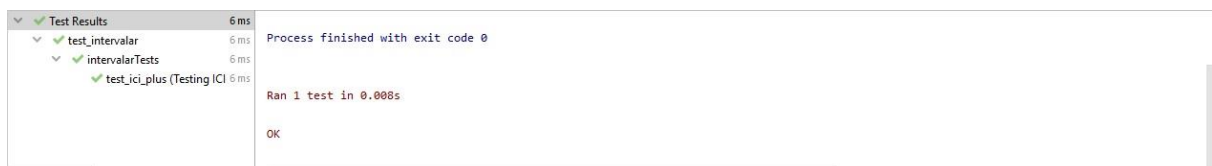


Figura 23. Test clase interval.



Figura 24. Test class operators.

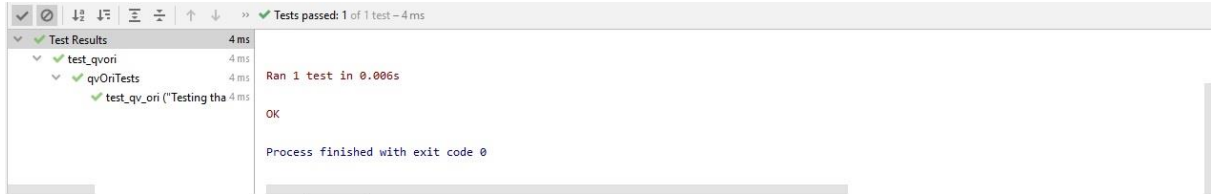


Figura 25. Test class Qvori.



8. Conclusión

La realización de este trabajo final de grado tiene implícitos varios objetivos: iniciación a la investigación, aplicación y puesta en práctica de los conocimientos adquiridos durante mis estudios, y finalmente, la utilización de tecnologías de actualidad en el mercado laboral para mi facilitar mi incorporación en el mismo.

En relación a los objetivos de iniciación a la investigación, las expectativas se han cumplido, ya que se ha desarrollado una librería en lenguaje Python que implementa los principales métodos de toma de decisión multicriterio, como son el método AHP y el método EMPV. Se ha desarrollado diferentes índices de valoración, como son el índice Qv-ORI, CMGI-Index, que permiten valorar los productos y servicios de internet, a través de información expresada en lenguaje natural. También, se han tenido que desarrollar técnicas de Scraping, para poder extraer información de sitios web, para posteriormente valorarla. Por último, dentro del marco teórico se ha aprendido distintos paradigmas y teorías expuestas en el apartado estudio del arte.

Los objetivos relacionados con mi formación durante mis años de estudio, también se han superado, ya que he tenido que aplicar técnicas de programación y algoritmia, pruebas con unittest aplicadas en la asignatura Introducción a la Ingeniería del Software que han sido muy útiles en este TFG. Usar metodología similar a la vista en Programación Orientada a Objetos, debido al parecido de java con Python. Algunos apartados de este proyecto también han tenido que ver con cosas aprendidas en Estructuras Algebraicas para la Computación (Autovectores, Autovalores, Perron-Frobenius...).

Finalmente, los objetivos iniciales relacionados con la incorporación al mercado laboral, también se han cumplido, ya que he desarrollado trabajos sobre una temática de gran interés, como es el procesamiento de la información generada por los públicos en internet, y su aplicación en modelos de toma de decisión. Las tecnologías utilizadas son muy solicitadas por las empresas y el mercado laboral en su conjunto. Además, las tareas desarrolladas me han permitido involucrarme con un equipo de desarrollo,



lo que me ha permitido poner en práctica y experimentar (aunque el resto de compañeros desarrollaba su propio trabajo) el trabajo en grupo, aspecto muy valorado por las empresas.

Bibliografía

- Bernal, R. A. (2015). *Modelado de Toma de Decisión con Coalición de Criterios e Información Lingüística*.
- Cabrera, F. E., Amaya, M., Vaccaro, G. F., & Peláez, J. I. (2019). Pairwise voting to rank touristic destinations based on preference valuation. *Sustainability (Switzerland)*, 11(21), 1–13. <https://doi.org/10.3390/su11215873>
- Carrasco, R. A., Sánchez-Fernández, J., Muñoz-Leiva, F., Francisca Blasco, M., & Herrera-Viedma, E. (2017). Evaluation of the hotels e-services quality under the user's experience. *Soft Computing*, 21(4), 995–1011. <https://doi.org/10.1007/s00500-015-1832-0>
- Chung, N., & Koo, C. (2015). The use of social media in travel information search. *Telematics and Informatics*, 32(2), 215–229. <https://doi.org/10.1016/j.tele.2014.08.005>
- Doña, J. M. (2008). *Modelado de los procesos de toma de decisión en entornos sociales mediante operadores de agregación OWA*. 41(1), 129–138.
- Fishburn, P. C. (1979). *Arrow's Impossibility Theorem: Concise Proof and Infinite Voters*. 20(2), 403–414.
- Harker, P. T. (1987a). Alternative modes of questioning in the analytic hierarchy process. *Mathematical Modelling*, 9(3–5), 353–360. [https://doi.org/10.1016/0270-0255\(87\)90492-1](https://doi.org/10.1016/0270-0255(87)90492-1)
- Harker, P. T. (1987b). Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling*, 9(11), 837–848. [https://doi.org/10.1016/0270-0255\(87\)90503-3](https://doi.org/10.1016/0270-0255(87)90503-3)
- Harker, Patrick T. (1987). Shortening the comparison process in the AHP. *Mathematical Modelling*, 8(C), 139–141. [https://doi.org/10.1016/0270-0255\(87\)90557-4](https://doi.org/10.1016/0270-0255(87)90557-4)
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. <https://doi.org/10.1145/1014052.1014073>
- Karanik, M., Pelaez, J. I., & Bernal, R. (2005). *Selective majority additive ordered*

weighting averaging operator. 12, 107–120.

- Karanik, Marcelo, Peláez, J. I., & Bernal, R. (2016). Selective majority additive ordered weighting averaging operator. *European Journal of Operational Research*, 250(3), 816–826. <https://doi.org/10.1016/j.ejor.2015.10.011>
- Kelly, J. (1988). *Arrow's Impossibility Theorem.* 58(4), 7–8.
- Kemeny, J. A., Endriss, U., & Haan, R. De. (n.d.). *Complexity of the Winner Determination Problem in.* 117–125.
- Peláez, J., & Doña, J. (2006). A majority model in group decision making using QMA–OWA operators: Research Articles. *International Journal of Intelligent Systems*, 21(2), 193–208. <https://doi.org/10.1002/int.v21:2>
- Peláez, J., & Doña, J. M. (2003). Majority additive-ordered weighting averaging: A new neat ordered weighting averaging operator based on the majority process. *International Journal of Intelligent Systems*, 18(4), 469–481. <https://doi.org/10.1002/int.10096>
- Peláez, J. I., Bernal, R., & Karanik, M. (2016). Majority OWA operator for opinion rating in social media. *Soft Computing*, 20(3), 1047–1055. <https://doi.org/10.1007/s00500-014-1564-6>
- Peláez, J. I., & Lamata, M. T. (2003). A new measure of consistency for positive reciprocal matrices. *Computers and Mathematics with Applications*, 46(12), 1839–1845. [https://doi.org/10.1016/S0898-1221\(03\)90240-9](https://doi.org/10.1016/S0898-1221(03)90240-9)
- Pelaez, J. I., Martinez, E. A., & Vargas, L. G. (2018). Consistency in Positive Reciprocal Matrices: An Improvement in Measurement Methods. *IEEE Access*, 6, 25600–25609. <https://doi.org/10.1109/ACCESS.2018.2829024>
- Peláez, J. I., Martínez, E. A., & Vargas, L. G. (2019). Products and services valuation through unsolicited information from social media. *Soft Computing*, 3. <https://doi.org/10.1007/s00500-019-04005-3>
- Peláez, José I., Martínez, E. A., & Vargas, L. G. (2019). Decision making in social media with consistent data. *Knowledge-Based Systems*, 172, 33–41. <https://doi.org/10.1016/j.knosys.2019.02.009>
- Ponzi, L. J., Fombrun, C. J., & Gardberg, N. A. (2011). *RepTrak™ Pulse : Conceptualizing and Validating a Short-Form Measure of Corporate Reputation.*



14(1), 15–35. <https://doi.org/10.1057/crr.2011.5>

Thomas L. Saaty. (2008). Decision making with the analytic hierarchy process. *Int. J. Services Sciences*, 1(1), 83. <https://doi.org/10.1504/IJSSCI.2008.017590>

Trigás Gallego, M. (2012). Metodología Scrum. *Gestion de Proyectos Informáticos*, 56. Retrieved from https://s3.amazonaws.com/academia.edu.documents/39164786/mtrigasTFC0612memoria_1.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1554050320&Signature=lv5tLOeYpbqQj0Q14ZHSM3k12H8%3D&response-content-disposition=inline%3B+filename%3DMtrigas_TFC0612memoria_1

Vargas, L. G. (2016). Voting with Intensity of Preferences. *International Journal of Information Technology & Decision Making*, 15(04), 839–859. <https://doi.org/10.1142/s0219622016400058>

Yager, R. R. (1988). On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1), 183–190. <https://doi.org/10.1109/21.87068>

Yager, R. R. (2004). OWA aggregation over a continuous interval argument with applications to decision making. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(5), 1952–1963. <https://doi.org/10.1109/TSMCB.2004.831154>

Yager, R. R., & Xu, Z. (2006). The continuous ordered weighted geometric operator and its application to decision making. *Fuzzy Sets and Systems*, 157(10), 1393–1402. <https://doi.org/10.1016/j.fss.2005.12.001>

Zhang, C., Tian, Y. X., Fan, Z. P., Liu, Y., & Fan, L. W. (2019). Product sales forecasting using macroeconomic indicators and online reviews: a method combining prospect theory and sentiment analysis. *Soft Computing*, 1(2009). <https://doi.org/10.1007/s00500-018-03742-1>



UNIVERSIDAD
DE MÁLAGA





Anexo I: Listado de Figuras

| | |
|---|----|
| Figura 1. Proceso de toma de decisiones. | 20 |
| Figura 2. Modelado del método AHP con Archimate. | 28 |
| Figura 3. Captura de cambios en el proyecto mediante Bitbucket. | 29 |
| Figura 4. Carpetas del proyecto en Bitbucket. | 29 |
| Figura 5. Control de versiones con el software Git. | 30 |
| Figura 6. Captura del software Pycharm. | 31 |
| Figura 7. Funcionamiento del Método AHP. | 39 |
| Figura 8. Escala de Saaty para el Método AHP (Thomas L. Saaty, 2008) | 40 |
| Figura 9. Creación de una matriz de comparación entre pares (Thomas L. Saaty, 2008). | 41 |
| Figura 10. Modelado del Método EMPV. | 43 |
| Figura 11. Modelado del Índice de Reputación Online QV-ORI (J. I. Peláez et al., 2019) | 46 |
| Figura 12. Modelado del método Reprtrak-Pulse (Ponzi et al., 2011) | 47 |
| Figura 13. Modelado del índice de Sinergia CMGI. | 50 |
| Figura 14. Modelado del Índice de Reputación Online mediante intervalos (José I. Peláez et al., 2019). | 51 |
| Figura 15. Concepto de transitividad (Γ). | 53 |
| Figura 17. Cálculo del número de transitividades de una matriz NXN. | 54 |
| Figura 18. Cálculo del índice de consistencia CI* (J. I. Peláez & Lamata, 2003). | 55 |
| Figura 19. Cálculo de CI+ para una matriz (Pelaez et al., 2018). | 56 |
| Figura 20. Cálculo del índice ICI+ para matrices de dimensión NxN (José I. Peláez et al., 2019). | 57 |
| Figura 21: Test clase Ahp y AhpConsistecy | 71 |
| Figura 22. Test clase Cmgj. | 71 |
| Figura 23. Test clase Empv. | 71 |
| Figura 24. Test clase interval. | 71 |
| Figura 25. Test clase operators. | 72 |
| Figura 26. Test clase Qvori. | 72 |



UNIVERSIDAD
DE MÁLAGA





UNIVERSIDAD
DE MÁLAGA





UNIVERSIDAD
DE MÁLAGA | **uma.es**

E.T.S. DE INGENIERÍA
INFORMÁTICA

E.T.S de Ingeniería Informática
Bulevar Louis Pasteur, 35
Campus de Teatinos
29071 Málaga