



IV Jornadas de Calidad de Software y Agilidad

JCSA | 2021

12 y 13 de noviembre de 2021



LIBRO DE ACTAS



Libro de Actas de las Cuartas Jornadas de Calidad de Software y Agilidad / Gladys Noemí Dapozo, Emanuel Irrazábal, María de los Ángeles Ferraro, Horacio D. Kuna, Eduardo Zamudio, Alice Rambo, César Acuña, Verónica Bollati y Noelia Pinto ; compilación de Gladys Noemí Dapozo ; Emanuel Agustín Irrazábal. - 1a ed compendiada. - Corrientes : Universidad Nacional del Nordeste. Facultad de Ciencias Exactas, 2021. Libro digital, PDF

Editorial de la Universidad Nacional del Nordeste (EUDENE)

ISBN 978-987-3619-72-4

1. Software. 2. Jornadas. 3. Argentina. I. Dapozo, Gladys Noemí, comp. II. Irrazábal, Emanuel Agustín, comp.

CDD 004.0711

Fecha de catalogación: 03/01/2022

Autoridades

Universidad Nacional del Nordeste

Rectora: Prof. María Delfina Veiravé

Vicerrector: Dr. Mario H. Urbani

Facultad de Ciencias Exactas y Naturales y Agrimensura

Decana: Mgter. María Viviana Godoy Guglielmone

Vicedecano: Dr. Enrique Laffont

Universidad Nacional de Misiones

Rectora: Mgter. Alicia Violeta Bohren

Vicerrector: Ing. Fernando Luis Kramer

Facultad de Ciencias Exactas, Químicas y Naturales

Decano: Dr. Luis Brumobsky

Vicedecano: Dr. Marcelo Marinelli

Universidad Tecnológica Nacional

Rector: Ing. Héctor Aiassa

Vicerrector: Ing. Haroldo Avetta

Facultad Regional Resistencia

Decano: Jorge A. De Pedro

Vicedecano: Dr. Ing. Walter Gustavo Morales

IV Jornadas de Calidad de Software y Agilidad

Este evento es organizado en forma conjunta por las universidades de la región que tienen unidades académicas que ofrecen carreras de Informática, UNNE, UTN y UNaM, a través de un acuerdo específico institucional, con el objetivo de difundir avances significativos en el campo de conocimiento de la Ingeniería de Software, Calidad y Agilidad; y propiciar el encuentro entre las universidades, las empresas y los organismos del Estado para contribuir al desarrollo de la industria del software en la región de influencia de las universidades participantes. Las Jornadas de Calidad de Software y Agilidad (JCSA) se inician en el año 2017, bajo el nombre original de Jornadas de Calidad de Software, en esta edición, las jornadas se consolidaron como un foro regional de referencia, ampliando su espectro para enfatizar también temas relacionados al uso de la agilidad. Además, se propuso la publicación de los trabajos académicos presentados, previa evaluación de pares, que se incorporan en este libro de actas. Durante el evento se realizaron tres talleres vinculados con los temas de la Jornada, se expusieron los artículos y posters académicos aceptados y se presentaron experiencias de la industria del software, con la participación de empresas y organismos del Estado. La conferencia inaugural denominada “Estrategias pruebas de aceptación para Entrega Continua”, estuvo a cargo de Diego Fontdevilla (UNTREF). Las actividades propuestas estuvieron destinadas a ingenieros y licenciados en sistemas, estudiantes y docentes de estas especialidades, profesionales y empresarios del sector del software y servicios informáticos, así como también público interesado en la temática. En total participaron más de 200 personas en las actividades programadas, lo que evidencia el interés que estos temas suscitan en los destinatarios.

Comité Organizador

Mgter. Gladys Noemí Dapozo
(FaCENA - UNNE)

Dr. Emanuel Irrazábal
(FaCENA - UNNE)

Lic. María de los Ángeles Ferraro
(FaCENA - UNNE)

Dr. Horacio D. Kuna
(FQCEyN - UNaM)

Dr. Eduardo Zamudio
(FQCEyN - UNaM)

Ing. Alice Rambo
(FQCEyN - UNaM)

Dr. César Acuña
(FRRe - UTN)

Dra. Verónica Bollati
(FRRe - UTN)

Dra. Noelia Pinto
(FRRe - UTN)

Comité de Programa

Mgter. Cristina Greiner (GICS - FaCENA - UNNE)	Dr. Diego Godoy (UGD)
Mgter. Gladys Noemí Dapozo (GICS - FaCENA - UNNE)	Ing. Edgardo Belloni (UGD)
Dr. Emanuel Irrazábal (GICS - FaCENA - UNNE)	Mag. Liliana Cuenca Pletsch (CINApTIC - UTN - FRRe)
Dr. Rubén Bernal (GICS - FaCENA - UNNE)	Dr. César J. Acuña (CINApTIC - UTN - FRRe)
Dr. David la Red Martínez (FaCENA - UNNE)	Dra. Verónica Bollati (CINApTIC - UTN - FRRe / CONICET)
Dra. Sonia Mariño (FaCENA - UNNE)	Dra. Noelia Pinto (CINApTIC - UTN - FRRe)
Mgter. M. Viviana Godoy Guglielmono (FaCENA - UNNE)	Esp. Gabriela Tomaselli (CINApTIC - UTN - FRRe)
Mgter. Mónica Tugnarelli (FCAD - UNER)	Ing. Nicolas Tortosa (CINApTIC - UTN - FRRe)
Dra. Gabriela Arévalo (DCyT - UNQ)	Ing. Valeria Sandobal Verón (GIESIN- UTN - FRRe)
Dra. María Fernanda Golobisky (UTN - FRStaFe)	Ing. Germán Gaona (CINApTIC - UTN - FRRe)
Master Ariel Pasini (LIDI - UNLP)	Dr. Horacio Leone (INGAR - UTN - FRSF)
Mgter. Pablo Thomas (LIDI - UNLP)	Dr. Silvio Gonnet (INGAR - UTN - FRSF)
Dr. Fernando Emmanuel Frati (DCByT - UNdeC)	Dr. Gustavo Rossi (LIFIA - UNLP)
Dra. Marcela Genero Bocco (Grupo Alarcos - UCLM)	Dra. Alejandra Garrido (LIFIA - UNLP)
Dr. Jorge Andrés Díaz Pace (ISISTAN - CONICET)	Dr. Andrés Rodríguez (LIFIA - UNLP)
Dr. Nazareno Aguirre (FCEfQyN - UNRC / CONICET)	Dr. Marcelo Estayno (UNSAM)
Dra. Nancy Ganz (IIDII-FCEQyN - UNaM)	Dr. Luis Olsina (GIDIS - UNLPam)
Esp. Ing. Alice Rambo (IIDII-FCEQyN - UNaM)	Dra. Luciana Ballejos (CIDISI - UTN - FRSF)
Ing. Selva Nieves Ivaniszyn (FCEQyN - UNaM)	Dra. Luciana Roldan (INGAR - UTN - FRSF)
Lic. Sergio Caballero (FCEQyN - UNaM)	Dra. Milagros Gutierrez (UTN - FRSF)
Lic. Martín Rey (IIDII-FCEQyN - UNaM)	Dra. Mariel Alejandra Ale (CIDISI - UTN - FRSF)
Dr. Eduardo Zamudio (IIDII - FCEQyN - UNaM)	Dra. Elsa Estevez (UNSur / CONICET)
Dr. Horacio Kuna (IIDII - FCEQyN - UNaM)	Dra. Alicia Mon (UNLaM)

Evaluación de la calidad de visualizaciones de datos basada en el Proceso Analítico Jerárquico y Redes Neuronales de Propagación hacia Atrás

Andrea Lezcano Airaldi¹, Joaquín Acevedo¹, María Laura Godoy¹

¹ Universidad Nacional del Nordeste, Corrientes, Argentina
{alezcano, jacevedo}@exa.unne.edu.ar, marialgodoy15@gmail.com

Resumen. La visualización de datos se ha vuelto esencial para comprender grandes conjuntos de datos y comunicar hallazgos. La narrativa basada en datos se utiliza cada vez más para dar soporte a la toma de decisiones en diferentes ámbitos, por lo cual resulta fundamental evaluar su calidad. Este trabajo propone la implementación del proceso analítico jerárquico en combinación con redes neuronales de propagación hacia atrás (AHP-BP), teniendo en cuenta el conocimiento y experiencia de los evaluadores. Se analizaron estudios existentes para conocer los factores de mayor influencia en la calidad de visualizaciones narrativas y se elaboró el sistema de índices de evaluación. Se seleccionaron diferentes usuarios como expertos evaluadores y se calculó el peso de cada uno. Se ajustaron los valores iniciales de las evaluaciones con los pesos de los expertos y luego se ponderaron para mitigar la subjetividad de los cálculos. Finalmente, con los resultados de la evaluación de 20 visualizaciones se entrenó un modelo de red neuronal BP. Los resultados demuestran que el método refleja la competencia de los expertos y reduce la subjetividad de la evaluación. La capacidad de aprendizaje de la red neuronal proporciona resultados precisos y confiables.

Palabras clave: proceso analítico jerárquico, redes neuronales, visualización de datos, evaluación de calidad.

1 Introducción

La visualización de datos se ha vuelto esencial para comprender grandes conjuntos de datos y comunicar hallazgos. La narrativa basada en datos o «data storytelling» se utiliza cada vez más para la toma de decisiones[1], por lo cual es importante considerar las perspectivas de los usuarios finales en la construcción de dichas narrativas.

En los últimos años, la evaluación se convirtió en un eje central en el campo de las visualizaciones de datos. Hay diversos métodos cualitativos y cuantitativos para evaluar diferentes aspectos de las visualizaciones narrativas[2]. Algunos de estos incluyen experimentos controlados y pruebas de usabilidad[3].

El proceso analítico jerárquico [4] es un método para toma de decisiones que permite estructurar un problema como una jerarquía de criterios con distintos “pesos” y seleccionar los más adecuados a través de comparaciones por pares realizadas por expertos,

lo que permite traducir opiniones subjetivas en métricas precisas. El método tradicional de AHP evalúa únicamente a través de las valoraciones de los expertos. Sin embargo, cuando estos tienen diferentes conocimientos, competencias y experiencias, la valoración de cada uno frente a un mismo criterio puede no ser igual y la equidad del cálculo se pierde. Por ello, es importante considerar estas diferencias e incorporarlas a los cálculos de ponderación.

La red neuronal de propagación hacia atrás (*back propagation*, BP) es un modelo típico de red neuronal de retroalimentación multicapa compuesto por una capa de entrada, una capa oculta y una capa de salida. Fue propuesto por Rumelhart y McClelland en 1986 y es uno de los algoritmos de red neuronal más utilizados[5][6].

El objetivo de este trabajo es implementar un modelo de evaluación que combine el método analítico jerárquico con redes neuronales de propagación hacia atrás, asegurando un análisis efectivo de los factores que influyen en la calidad de las visualizaciones y reduciendo la subjetividad de los expertos.

Se analizó la literatura existente para conocer los factores de mayor influencia en la calidad de visualizaciones narrativas y se elaboró el sistema de índices de evaluación. Se seleccionaron diferentes usuarios como expertos evaluadores y se calculó el peso de cada uno. Se ajustaron los valores iniciales de las evaluaciones con los pesos de los expertos y luego se ponderaron para mitigar la subjetividad de los cálculos. Finalmente, con los resultados de la evaluación de 20 visualizaciones se entrenó un modelo de red neuronal de propagación hacia atrás.

El trabajo se encuentra organizado de la siguiente manera. En la Sección 2 se describen los antecedentes y los trabajos relacionados. En la Sección 3 se enumeran las decisiones tomadas para la metodología de investigación propuesta, la recolección y el análisis de los datos, mientras que en la Sección 4 se detallan los resultados y en la Sección 5 se presenta las conclusiones.

2 Antecedentes y Trabajos Relacionados

2.1 Evaluación de calidad de visualizaciones de datos

Las visualizaciones de datos efectivas se elaboran en base al mensaje que los autores pretenden comunicar y considerando el mejor medio para mostrar las variables [7][8]. Actualmente existen varios enfoques para evaluar la efectividad de las visualizaciones de datos. Los métodos tradicionales incluyen estudios de percepción gráfica [9][10], experimentos controlados, o estudios de usabilidad [11].

Lam et. al. [12] definen diferentes tipos de escenarios de evaluación, categorizados en aquellos para comprender los procesos de análisis de datos y aquellos que evalúan las visualizaciones en sí mismas. Basan su categorización en preguntas y objetivos, animando así a considerar el contexto antes de elegir un método de evaluación.

Tory y Moller [13] proponen que la opinión de los expertos en el dominio podría ser un complemento útil para los experimentos controlados, tanto para la evaluación heurística de la usabilidad como para comprender las actividades cognitivas de alto nivel. Sin embargo, no solo las opiniones de los expertos en el dominio, sino también las

opiniones de los especialistas visuales, como artistas, diseñadores gráficos e ilustradores, pueden ser beneficiosas, como se demuestra en [14].

2.2 El Proceso Analítico Jerárquico

El Proceso Analítico Jerárquico (AHP) fue propuesto por Saaty [4] en 1980 y permite determinar los pesos de criterios estructurados jerárquicamente. Los diferentes criterios se comparan por pares y los resultados de las comparaciones se organizan en una matriz. La comparación es cualitativa e indica si un criterio es más significativo que el otro y a qué nivel pertenece la prioridad. La técnica utilizada permite convertir las estimaciones cualitativas obtenidas de los expertos en valores cuantitativos.

Dado que la matriz de evaluación se construye mediante puntajes de expertos, para verificar el sesgo de subjetividad, se colocan los pesos relativos en la fórmula de relación de consistencia (RC) para la prueba de consistencia. Cuando esta es menor que 0,1, la matriz de evaluación es aceptable. El AHP es útil cuando hay muchos intereses involucrados y varias personas participan en el proceso de evaluación [4].

2.3 Redes neuronales de propagación hacia atrás

La red neuronal BP es un modelo de red neuronal de retroalimentación multicapa compuesto por una capa de entrada, una capa oculta y una capa de salida. La estructura del modelo y los pesos de la red neuronal BP se obtienen a través del proceso de aprendizaje. El mismo se divide en dos etapas: una etapa de feed forward que calcula la entrada y salida reales de cada nodo y la etapa de retro propagación, donde el peso de cada conexión se corrige a lo largo de la ruta inversa para reducir el error [15]. La idea básica de la red neuronal BP es lograr el error cuadrático mínimo de la capa de salida, ajustando los pesos y umbrales de la red, de modo que el valor de salida sea lo más cercano posible al valor objetivo. La estructura de la red se muestra en la Fig. 1.

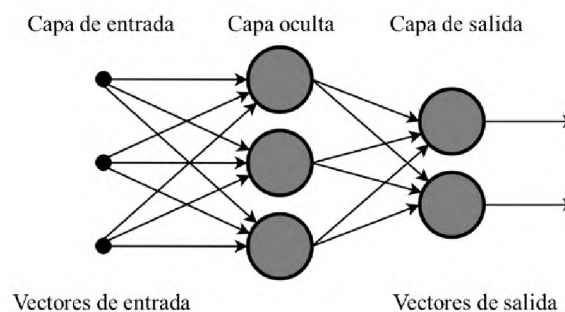


Fig. 1. Diagrama de la estructura básica de una red neuronal BP.

2.4 Combinación de AHP y redes neuronales BP

Dado que el principal inconveniente del AHP es la evaluación subjetiva, muchos autores han intentado mejorarlo combinándolo con otros métodos como la lógica difusa

[16] o máquinas de vector soporte [17]. En cuanto a las redes neuronales BP, si los datos brutos se utilizan directamente para entrenar el modelo, con el aumento de datos, la estructura de la red neuronal BP se vuelve más compleja y el tiempo de entrenamiento es más largo, lo que resulta en una reducción del rendimiento [18].

Atendiendo a estos problemas, Li et. al. [19] proponen el método AHP mejorado, en el cual se tienen en cuenta aspectos de la personalidad de los expertos para corregir los pesos de las evaluaciones una primera vez y luego se combinan los resultados con una red neuronal BP. En [20] se considera la perspectiva o intereses de los expertos.

Esta corrección puede evitar la disminución del rendimiento de la red neuronal BP debido a la redundancia de datos, y la combinación de métodos puede corregir los errores causados por la subjetividad de los expertos.

3 Metodología

Los pasos seguidos para la evaluación de la calidad de las visualizaciones fueron los siguientes [5][15][19][20]:

1. Construcción del sistema de índices de evaluación de la calidad de visualizaciones;
2. Cálculo de los pesos relativos de los índices en todos los niveles mediante AHP. Se utilizó un sistema en línea [21] para garantizar la ejecución sin errores.
3. Cálculo del peso de la evaluación de cada experto;
4. Ponderación y promedio global de los valores de los índices junto con las evaluaciones de los expertos de los pasos 2 y 3;
5. Entrenamiento del modelo de red neuronal de BP tomando el resultado del paso anterior como entrada.
6. Verificación y análisis de la aplicación.

3.1 Selección de índices de evaluación

Se identificaron los criterios más importantes que influyen en la calidad de las visualizaciones de datos a partir de los resultados de estudios relacionados con narrativa basada en datos y buenas prácticas de visualización [22]-[30]. Los mismos se clasificaron en cinco categorías: Narrativa, Diseño, Interacción, No Manipulación y Gráficos Apropriados. La Narrativa hace referencia a cómo se cuenta la historia e incluye el orden de la secuencia de eventos (N1) y la incorporación de los elementos básicos de una historia (N2). El Diseño, se compone de prácticas relativas a la visualización propiamente dicha: usar colores consistentes (D1), resaltar lo importante (D2), eliminar el desorden visual (D3), y usar texto, etiquetas y anotaciones para facilitar la comprensión (D4). La categoría Interacción incluye el incentivo a la exploración (I1), y la estimulación de la curiosidad del usuario (I2). La No Manipulación se refiere a la ética al crear visualizaciones. En este sentido, abarca las prácticas de ser fiel a la información (M1), no citar información fuera de contexto (M2), variar los datos, no el diseño (M3) y no distorsionar los gráficos (M4). Finalmente, la categoría Gráficos Apropriados apunta a la elección de un gráfico simple (G1) y a eliminar la complejidad innecesaria (G2). Las categorías y subcategorías constituyen los índices de primer y segundo nivel

respectivamente. El sistema de índices resultante se muestra en la Tabla 1, e incluye cinco índices de primer nivel y 14 índices de segundo nivel.

Tabla 1. Índices de evaluación de la calidad de visualizaciones de datos.

Objetivo	Primer Nivel	Segundo Nivel
Calidad	Narrativa	N1 - Ordenar la secuencia de eventos N2 - Incorporar elementos básicos de la historia D1 - Usar colores consistentes
	Diseño	D2 - Resaltar lo importante D3 - Eliminar el desorden visual D4 - Usar texto, etiquetas y anotaciones
	Interacción	I1 - Incentivar la exploración I2 - Estimular la curiosidad del usuario
	No Manipulación	M1 - Ser fiel a la información M2 - No citar información fuera de contexto M3 - Variar los datos, no el diseño M4 - No distorsionar los gráficos
	Gráfico Apropiado	G1 - Elegir un gráfico simple G2 - Eliminar complejidad innecesaria

3.2 Cálculo del peso en cada nivel utilizando AHP

Se invitó a 10 usuarios (expertos) a calificar los índices de evaluación, y se comparó la puntuación de los expertos en pares. Se calculó la matriz de evaluación de cada experto en cada nivel. Tomando como ejemplo a uno de los expertos, la matriz de evaluación de los cinco criterios de primer nivel se muestra en la Tabla 2. Se calculó la razón de consistencia (RC) de la matriz de evaluación. El valor fue de 0.079, por lo que la matriz tiene una consistencia aceptable. Este procedimiento se repitió con cada experto.

Tabla 2. Matriz de evaluación

	Narrativa	Diseño	Interacción	No Manipulación	Gráfico Apropiado
Narrativa	1	1	1/2	1/3	1
Diseño	1	1	3	1/2	1
Interacción	2	1/3	1	1/2	1
No Manipulación	3	2	2	1	1
Gráfico Apropiado	1	1	1	1	1

3.3 Cálculo del peso de cada experto

Los expertos fueron divididos en tres grupos, de acuerdo con sus conocimientos, competencias y experiencias con visualizaciones de datos: usuarios avanzados, usuarios intermedios y usuarios ocasionales. De acuerdo con su categoría, se le asignó a cada experto un peso de evaluación y luego se ponderaron y promediaron las estimaciones realizadas en el paso anterior utilizando sus pesos de evaluación.

3.4 Peso ajustado de los índices

El peso global en todos los niveles del sistema de índices se obtiene mediante el método de ponderación y promedio de los valores obtenidos mediante AHP junto con el vector de ponderación de los expertos. Los resultados se muestran en la Tabla 3.

Tabla 3. Peso global de los criterios de calidad para visualizaciones.

Primer nivel	Segundo Nivel	Tercer nivel	Peso
Calidad	Narrativa	N1 - Ordenar la secuencia de eventos	0,102
		N2 - Incorporar elementos básicos de la historia	0,059
		D1 - Usar colores consistentes	0,035
	Diseño	D2 - Resaltar lo importante	0,056
		D3 - Eliminar el desorden visual	0,068
		D4 - Usar texto, etiquetas y anotaciones	0,053
		I1 - Incentivar la exploración	0,083
	Interacción	I2 - Estimular la curiosidad del usuario	0,083
		M1 - Ser fiel a la información	0,046
	No Manipulación	M2 - No citar información fuera de contexto	0,06
		M3 - Variar los datos, no el diseño	0,071
		M4 - No distorsionar los gráficos	0,092
	Gráfico Apropiado	G1 - Elegir un gráfico simple	0,096
		G2 - Eliminar complejidad innecesaria	0,096

3.5 Red neuronal BP

Para determinar el peso de los factores de calidad los valores obtenidos por el AHP mejorado se utilizan como muestras para entrenar y probar la red neuronal BP, de modo

que los pesos fijos de cada factor se pueden determinar una vez que la red alcanza estabilidad. Los pasos específicos, siguiendo las definiciones dadas por [20] son:

- Inicializar los pesos y umbrales de la red. El vector peso de la capa de entrada a la capa oculta es $V_k = (v_{qj,k})_{Q \times J}$, en el que $v_{qj,k}$ representa el peso desde la j -ésima capa de entrada al q -ésimo nodo oculto y θ_{qk} es el umbral del q -ésimo nodo de la capa oculta. El vector peso de la capa de entrada a la capa implícita es $W_k = (w_{mq,k})_{M \times Q}$, en el que $w_{mq,k}$ representa el peso desde el q -ésimo nodo de la capa oculta al m -ésimo nodo de salida y θ_{mk} es el umbral del m -ésimo nodo de la capa de salida.
- Ingresar la muestra de aprendizaje y calcular los valores de salida de cada capa. El vector de valores esperados de la capa de salida $T_i = (t_{i1}, t_{i2}, \dots, t_{iM})$ es generado en base a los resultados de evaluación de los expertos, cuando el vector de datos normalizados X_i de la i -ésima muestra es ingresado. El vector salida de la capa de salida es $H_{i,k} = (h_{i1,k}, h_{i2,k}, \dots, h_{iQ,k})$ y el vector de salida de la capa de salida es $O_{i,k} = (o_{i1,k}, o_{i2,k}, \dots, o_{iM,k})$.

4 Resultados

4.1 Preprocesamiento de los datos

Se seleccionaron veinte visualizaciones de datos cuya calidad fue evaluada por los diez expertos de acuerdo con el sistema de índices de evaluación propuesto en la Sección 3.1. El vector de evaluación es X , el vector de ponderación de evaluación es W y el vector de resultado de evaluación Z de las 20 visualizaciones se calcula mediante la fórmula $Z = WX$.

$Z = [0.4120, 0.6950, 0.7188, 0.3601, 0.4237, 0.7112, 0.1720, 0.3239, 0.9107, 0.2914, 0.3527, 0.1251, 0.6548, 0.2101, 0.6511, 0.1056, 0.6119, 0.8621, 0.5651, 0.4402]$

4.2 Modelado de la red neuronal BP.

Para el diseño, entrenamiento y pruebas de la red neuronal BP se utilizó la librería de Scikit-learn [31] para redes neuronales de Python. El número de nodos de la capa de entrada de la red se estableció en $n=14$, el número de nodos de salida fue $M=14$ de acuerdo con el vector de valores esperados y el número de nodos de la capa oculta se estableció en $j = 10$. La tasa de aprendizaje se estableció en 0.05, con un paso de entrenamiento de 20.000 veces y la precisión de convergencia del error de entrenamiento en 10^{-2} .

Seleccionando los primeros 10 elementos del vector de evaluación como el conjunto de datos de entrenamiento, la variable de entrada es $X_t = [X_1, X_2, \dots, X_{10}]^T$. Los valores de la evaluación de los criterios de calidad ($Z_T = [Z_1, Z_2, \dots, Z_{10}]^T$) son usados como valor de salida esperado de la red, y el vector de salida esperado es $t_T = [t_1, t_2, \dots, t_{10}]^T$.

Seleccionando los últimos 10 elementos del vector de evaluación como el conjunto de datos de entrenamiento, la variable de entrada es $X_v = [X_{11}, X_{12}, \dots, X_{20}]^T$. Los

valores de la evaluación de los criterios de calidad ($Z_v = [Z_{11}, Z_{12}, \dots, Z_{20}]^T$) son usados como valor de salida esperado de la red, y el vector de salida esperado es $t_v = [t_{11}, t_{12}, \dots, t_{20}]^T$.

4.3 Entrenamiento de la red neuronal BP

Después de 50 pasos de cálculo, la varianza máxima de los datos de simulación es 0.0098, el error cuadrático medio de entrenamiento de la red es menor que el valor objetivo 10^{-2} . Este resultado satisface los requisitos de precisión en el aprendizaje. Es decir, los parámetros seleccionados por el modelo cumplen los requisitos. El error relativo del resultado del entrenamiento se muestra en la Tabla 4. Se puede observar que el valor de salida de la red está muy cerca del valor esperado.

Tabla 4. Resultados del entrenamiento de la red neuronal BP

	1	2	3	4	5	6	7	8	9	10
Valor esperado	0.4120	0.6950	0.7188	0.3601	0.4237	0.7112	0.1720	0.3239	0.9107	0.2914
Valor de entrenamiento	0.4152	0.7041	0.7190	0.3700	0.4245	0.7190	0.1790	0.3255	0.9019	0.2944
Error relativo	-0,0032	-0,0091	-0,0002	-0,0099	-0,0008	-0,0078	-0,0070	-0,0016	0,0088	0,0030

4.4 Prueba de la red neuronal BP

El resultado de la evaluación de los datos de la muestra 11-20 se comparó con el valor de salida esperado. El error se muestra en la Tabla 5. Se puede observar que el error entre el valor de salida y el valor esperado es pequeño.

Los resultados de la evaluación de la red neuronal BP muestran que el modelo tiene una alta eficiencia y un error pequeño y puede ser utilizado en la evaluación de la calidad visualizaciones.

Tabla 5. Resultados de la prueba de la red neuronal BP

	51	52	53	54	55	56	57	58	59	60
Valor esperado	0.3501	0.1299	0.6448	0.2050	0.6547	0.1047	0.6203	0.8580	0.5766	0.4393
Valor de entrenamiento	0.3527	0.1251	0.6548	0.2101	0.6511	0.1056	0.6119	0.8621	0.5651	0.4402
Error relativo	-0,0026	0,0048	-0,01	-0,0051	0,0036	-0,0009	0,0084	-0,0041	0,0115	-0,0009

5 Conclusiones

En este trabajo, se propone la aplicación del método AHP mejorado para la evaluación de la calidad de visualizaciones de datos. Se analizaron una serie de estudios existentes referentes a narrativa basada en datos y buenas prácticas de visualización y en base a los resultados del análisis se confeccionó el sistema de índices de evaluación. Se

seleccionaron diferentes grupos de usuarios como expertos evaluadores y se calculó el peso de cada experto en función de sus competencias, experiencia y conocimiento respecto a visualizaciones. Se ajustaron los valores iniciales de los índices con los pesos de los expertos y luego se ponderaron para mitigar la subjetividad de los cálculos. Finalmente, con los resultados de la evaluación de 20 visualizaciones se entrenó un modelo de red neuronal BP.

Se comprobó que el algoritmo AHP mejorado puede reducir la subjetividad de los expertos en la evaluación, y la combinación con una red neuronal BP permite realizar una evaluación más precisa y efectiva.

Referencias

- [1] W. Willett, J. Heer, J. M. Hellerstein, and M. Agrawala, “CommentSpace: Structured support for collaborative visual analysis,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2011, pp. 3131–3140, doi: 10.1145/1978942.1979407.
- [2] S. Carpendale, “Evaluating information visualizations,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4950 LNCS, pp. 19–45, 2008, doi: 10.1007/978-3-540-70956-5_2.
- [3] C. Plaisant, “The challenge of information visualization evaluation,” *Proc. Work. Adv. Vis. Interfaces AVI*, pp. 109–116, 2004, doi: 10.1145/989863.989880.
- [4] T. L. Saaty, “The Analytic Hierarchy Process: Planning, priority setting, resource allocation”, McGraw-Hill, 1980. Doi: [https://doi.org/10.1016/0377-2217\(82\)90022-4](https://doi.org/10.1016/0377-2217(82)90022-4).
- [5] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning internal representation by error propagation. *Neurocomputing: Foundations of research*. Cambridge: MIT Press. 14.
- [6] Bailer-Jones, C. A. L., Gupta, R., Singh, H. P., & Bulsari, A. B. (2001). An introduction to artificial neural networks. *Neural Networks for Chemical Engineers*, 19(12), 36–62.
- [7] Few S. Effectively Communicating Numbers: Selecting the Best Means and Manner of Display [White Paper]. Secondary Effectively Communicating Numbers: Selecting the Best means and manner of Display [White Paper]; 2005. http://www.perceptualedge.com/articles/Whitepapers/Communicating_Numbers.pdf
- [8] Cleveland W, McGill R. Graphical perception: theory, experimentation, and application to the development of graphical methods. *J Am Stat Assoc* 1984; 79 (387): 531–54.
- [9] J. Heer and M. Bostock. Crowdsourcing graphical perception: using Mechanical Turk to assess visualization design. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 203–212, 2010.
- [10] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1303–1312, 2009.
- [11] S. Carpendale. Evaluating information visualizations. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization: Human-Centered Issues and Perspectives*, volume 4950, pages 19–45. Springer LNCS, 2008.
- [12] H. Lam *et al.*, “Empirical Studies in Information Visualization: Seven Scenarios To cite this version: Empirical Studies in Information Visualization: Seven Scenarios,” *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 9, pp. 1520–1536, 2012.

- [13] M. Tory and T. Moller. Evaluating visualizations: Do expert reviews work? *IEEE Computer Graphics and Applications*, 25(5):8–11, Sept./Oct. 2005. doi> 10.1109/MCG.2005.102
- [14] D. Acevedo, C. D. Jackson, F. Drury, and D. H. Laidlaw. Using visual design experts in critique-based evaluation of 2D vector visualization methods. *IEEE Transactions on Visualization and Computer Graphics*, 14(4):877–884, July/Aug. 2008. doi> 10.1109/TVCG.2008.29
- [15] Fu, J., & Zhang, L. (2012). Green marketing performance evaluation of circular economy based on AHP and BP neural network model. *Science and Technology Management Research*, 20, 222–227.
- [16] Akbar, M. A., Shameem, M., Khan, A. A., Nadeem, M., Alsanad, A., & Gumaei, A. (2020). A fuzzy analytical hierarchy process to prioritize the success factors of requirement change management in global software development. *Journal of Software: Evolution and Process*. doi:10.1002/smr.2292
- [17] Wei, Wang. (2011). Network security assessment of AHP and SVM combination. *Computer Simulation*, 28(3), 182–185
- [18] Wang, X., Liang, D., Song, W., & Zhou, Y. (2018). Distributed intelligent pension system based on bp neural network. *Wireless Personal Communications*, 1–10, 1–12
- [19] Li, W., Xu, G., Xing, Q. et al. Application of Improved AHP-BP Neural Network in CSR Performance Evaluation Model. *Wireless Pers Commun* 111, 2215–2230 (2020). <https://doi.org/10.1007/s11277-019-06981-z>
- [20] Liu, Y. (2018) ‘An improved AHP and BP neural network method for service quality evaluation of city bus’, *Int. J. Computer Applications in Technology*, Vol.58,No.1, p.37–44
- [21] Goepel, K.D. (2018). Implementation of an Online Software Tool for the Analytic Hierarchy Process (AHP-OS). *International Journal of the Analytic Hierarchy Process*, Vol. 10 Issue 3 2018, pp 469-487, <https://doi.org/10.13033/ijahp.v10i3.590>
- [22] C. Nussbaumer Knafloc, *Storytelling with Data: A data visualization guide for business professionals*. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2015.
- [23] R. Kosara, J. Mackinlay, and T. Software, “Storytelling: The Next Step for Visualization.”
- [24] R. Kosara, “An Argument Structure for Data Stories.”
- [25] E. Segel and J. Heer, “Narrative Visualization: Telling Stories with Data,” 2010.
- [26] C. Tong et al., “Storytelling and visualization: A survey,” *VISIGRAPP 2018 - Proc. 13th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.*, vol. 3, pp. 212–224, 2018, doi: 10.5220/0006601102120224.
- [27] C. Tong et al., “Storytelling and visualization: An extended survey,” *Inf.*, vol. 9, no. 3, Mar. 2018, doi: 10.3390/info9030065.
- [28] J. Boy, F. Detienne, and J. D. Fekete, “Storytelling in information visualizations: Does it engage users to explore data?,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015-April, pp. 1449–1458, 2015, doi: 10.1145/2702123.2702452.
- [29] C. Nussbaumer Knaflic, “Data Stories.”
- [30] E. R. Tufte, *The Visual Display of Quantitative Information*.
- [31] Scikit-learn: machine learning in Python — scikit-learn 0.24.2 documentation <https://scikit-learn.org/stable/index.html>. Accessed: 2021-07-16