

Análisis del impacto social de los infocentros comunitarios en el cantón La Maná-

Ecuador

Analysis of the social impact of the community info-centers in the canton La Maná-

Ecuador

*Yoan Martínez-López*¹

*Wilmer A. Soria Paula*²

*Jahiro Sutherland*³

Resumen

La predicción del impacto social de los infocentros comunitarios es relevante tanto para el responsable, los usuarios y los administradores, al generar información precisa acerca de la situación educativa, la inclusión digital, la apropiación tecnológica y la participación social, a partir del reconocimiento de los beneficios y oportunidades de los servicios. Además, ayuda a la elección de herramientas de comunicación más eficientes para gestionar información y conocimientos según necesidades, intereses y motivaciones personales y de la comunidad. Para los usuarios de la comunidad es relevante al momento de comunicarse con otras personas, familiares y con los servicios comunitarios. En este trabajo se realizó el análisis del impacto social que causará la implementación de infocentros comunitarios en el Cantón La Maná, mediante el uso de técnicas de inteligencia artificial. Para ello se utilizaron los algoritmos de inteligencia artificial y minería de datos. Estos algoritmos están implementados en una

¹ Licenciado en Ciencias de la Computación, Doctor en Ciencias Técnicas. Profesor Titular. Facultad de Informática y Ciencias Exactas, Universidad de Camagüey, Cuba, E-mail: yoan.martinez@reduc.edu.cu. ORCID: <https://orcid.org/0000-0002-1950-567X>.

² Ingeniero en Diseño gráfico computarizado. Máster en Informática Aplicada. Facultad de Informática y Ciencias Exactas, Universidad de Camagüey, Cuba. E-mail: wisorp387@gmail.com. ORCID: <https://orcid.org/0009-0002-7248-2902>

³ Licenciado en Matemática, Máster en Matemática Aplicada, Centro Regional Universitario de Colón. Universidad de Panamá, Panamá. E-mail: jahiro.sutherland@up.ac.pa. ORCID: <https://orcid.org/0000-0003-4700-3818>



Fecha de recepción: 29-9-2025

Fecha de aceptación: 15-12-2025

Creative Commons Atribución 4.0

herramienta computacional para el aprendizaje automático llamada Weka. Además, las variedades de técnicas y algoritmos, y sus respectivas ventajas, analizadas en este estudio acerca de la herramienta Weka permiten identificar sus potencialidades para ser utilizada en la predicción de transformaciones en organizaciones que ofrecen servicio a la sociedad en general. Los resultados experimentales demostraron que estos algoritmos tienen un buen desempeño para la predicción del impacto social. Basado en este caso de estudio se concluye que el algoritmo Kstar es el que mejor desempeño muestra con nuestro conjunto de datos.

Palabras clave: aprendizaje automático, infocentro, algoritmo de clasificación, weka, minería de datos.

Abstract

The prediction of the social impact of the community infocenters is relevant for the person in charge, the users and the administrators, when generating precise information about the educational situation, the digital inclusion, the technological appropriation, social participation, from the recognition of the benefits and opportunities for services, as well as helping to choose more efficient communication tools to manage information and knowledge according to personal and community needs, interest and motivation. For the users of the community, it is relevant when communicating with other people, relatives, with community services. In this work the analysis of the social impact caused by the implementation of community infocenters in the Canton La Maná using artificial intelligence techniques was carried out. For this purpose, artificial intelligence and data mining algorithms are used. These algorithms are implemented in a computational tool for machine learning, called Weka. In addition, the varieties of techniques and algorithms, and their respective advantages, analyzed in this study, about the Weka tool allow identifying the potentials to be used in the determination of predictions of transformations

in organizations that offer services to society in general. The experimental results showed that these algorithms have a good performance for the prediction of social impact. Based on this case study, it is concluded that the Kstar algorithm shows the best performance with our dataset.

Keywords: machine learning, infocenter, classification algorithm, weka, data mining

Introducción

El hombre siempre ha tratado de predecir su futuro, como rasgo esencial del desarrollo ontogenético que lo identifica como ser social que percibe los cambios y transformaciones sociales y tecnológicas introducidas a partir de los avances tecnológicos en la dirección de la información y la comunicación de la actual sociedad de principio del siglo XXI (Jover, 2013).

En la sociedad de la información, signada por los cambios y avances de las tecnologías de la informática, resulta de gran interés por diversos investigadores de las ciencias, el estudio de la Inteligencia Artificial (IA) quizás, por ser unas de las disciplinas que más despiertan la imaginación de los que oyen hablar de ella por primera vez (Raffio, 2016). Es evidente que la Inteligencia Artificial, cada vez está más presente en la vida cotidiana de las personas, en ocasiones de manera inadvertida, cuando se realiza cualquier acción o actividad social con uso de las TIC (Limón, 2016).

En tal sentido la IA, desde el punto de vista de la ingeniería, utiliza diversas herramientas en la solución de problemas porque sus herramientas se presentan en diferentes técnicas, tales como: robótica, redes neuronales, sistemas de expertos, algoritmos genéticos, búsqueda de soluciones, representación del conocimiento, reconocimiento de patrones y procesamiento de lenguaje natural, las cuales proveen elementos fundamentales en las áreas que pertenecen a la misma (Torres Soler, 2016).

Por otro lado, los infocentros comunitarios, dentro de este contexto y las posibilidades que ofrece la IA, constituyen un espacio importante y de gran validez para las comunidades urbanas y rurales, al dar oportunidad inclusiva y autorregulativa a todos sus miembros. Estos espacios, como parte de la educación y la pedagogía, facilitan el acceso y uso de las posibilidades que ofrecen las Tecnologías de la Información y la Comunicación (TIC), promoviendo nuevos saberes y condiciones de vida.

La predicción del impacto social de los infocentros comunitarios de Ecuador es un desafío presente prácticamente en cualquiera de estas instituciones u organizaciones creadas para ofrecer servicio de uso de las Tecnología de la Información y la Comunicación (TIC) por parte de usuarios de comunidades rurales y urbanas fundamentalmente (Calderón et al., 2017, Gonzalo 2005).

Existen diversos ejemplos recientes; sin embargo, muy pocos abordan la temática (Soria Paula, 2017). Además, los usuarios de infocentros comunitarios se ven beneficiados por la incorporación de nuevas aplicaciones y herramientas de comunicación según sus necesidades, intereses y motivaciones, a partir de la aplicación de técnicas o algoritmos para prever lo que puede ocurrir en lo político-económico-social de la comunidad y; por consiguiente, de recomendaciones que permiten la toma de decisiones oportunas y en correspondencia a las transformaciones del entorno social y de sí mismo, tanto del presente como futura.

Desarrollo

Según Blanco y Hermida (2016), con el rápido crecimiento de información en Internet, la minería de datos se ha convertido en un campo muy importante para aprovechar esa gran cantidad de datos, al facilitar descubrir nuevos patrones o extraer información relevante sobre los datos de forma automática. Además, Hernández et al., (2004) consideran que el aprendizaje

automático (AA) es una serie de técnicas para que una máquina pueda aprender y generalizar comportamientos a partir de información suministrada. Los algoritmos aprenden de los datos para poder hacer predicciones.

Dentro del aprendizaje automático se encuentran distintos métodos. Uno de ellos es el aprendizaje supervisado (Espino, 2017). Este aprendizaje usa técnicas para deducir una función a partir de datos de entrenamiento. El objetivo es predecir el valor correspondiente a cualquier objeto de entrada a partir de los datos de entrenamiento que ha visto previamente.

Un ejemplo de este tipo de algoritmo es el problema de la clasificación, donde el sistema de aprendizaje etiqueta los datos de entrada basándose en los ejemplos ya etiquetados, y suministrados para que aprenda (Medrano, 2006).

Clasificación de los algoritmos

La clasificación: los datos son objetos caracterizados por atributos que pertenecen a diferentes clases (etiquetas discretas). La meta es inducir un modelo para poder predecir una clase dados los valores de los atributos. Se usan, por ejemplo: arboles de decisión, reglas, análisis de discriminantes, árboles de regresión, regresión lineal, redes neuronales, KNN, entre otros. (Álvarez y Álvarez, 2006).

La clasificación consiste en encontrar fronteras de separación entre las clases, dentro de las cuales se encuentran: las lineales, donde la frontera es una línea (en dos dimensiones) o un hiperplano (en N-dimensiones); y las no lineales, que se refieren a cualquier otro tipo de frontera. Los conjuntos de clasificadores construyen varios predictores (clasificación o regresión) y después los usan de manera conjunta, siendo más precisos que los algoritmos individuales, siempre que los algoritmos base predigan mejor que el azar. En tal sentido, la esencia es que si

los distintos clasificadores no están correlacionados en los errores, el uso conjunto de todos ellos será mejor que usar cualquiera de ellos por separado (Morales, 2014).

El algoritmo M5' hace uso de un modelo de regresión lineal conocido como model tree (árbol modelo). Es un árbol de decisión usado en predicción numérica que estima el valor de la clase de la instancia cuando se alcanza una de sus hojas. El model tree se usa para predecir el valor de una instancia, el árbol es recorrido hacia abajo hasta llegar a una hoja y se utilizan valores de los atributos de la instancia que permiten tomar decisiones de enrutado en cada nodo.

El Kstar, es un tipo de algoritmo de clasificación basado en instancias, es decir, que la clasificación de una instancia está basada en la clasificación de instancias de entrenamiento similares, determinadas por alguna función de similitud. Su particularidad radica en que usa una función de distancia basada en entropía (Martinez López & Madera Quintana, 2016).

El RPTree es considerado como un método de aprendizaje rápido mediante árboles de decisión, el cual lo construye usando la información de varianza y lo moda usando como criterio la reducción del error. Solo clasifica valores para atributos numéricos una vez (González, 2006).

Weka como herramienta de aprendizaje automático

Se comparte los estudios realizados por García (2011) y García (2012) al señalar ambos autores que la WEKA (Acrónimo de Waikato Environment for Knowledge Analysis) es un entorno para la experimentación de análisis de datos que permite aplicar, analizar y evaluar las técnicas más relevantes de análisis de datos; principalmente, las provenientes del aprendizaje automático, sobre cualquier conjunto de datos del usuario.

Desde un análisis de usabilidad, Blanco y Hermida (2016), considera que la herramienta Weka contiene una colección de algoritmos de aprendizaje automático y tratamiento de textos, tanto la versión gráfica como el código en Java, sustentado de otra parte por Lavin, (2017).

Resultados experimentales

La base de conocimiento se obtiene por los resultados de una encuesta aplicada en la provincia de Cotopaxi en Ecuador, específicamente en el cantón (división administrativa parecida a un municipio) La Maná. Esta encuesta se aplica a una muestra de 212 personas escogidas aleatoriamente en la población. Se procede a procesar los datos para lograr su estandarización, tratando de evitar datos incompletos, con ruido o simplemente inconsistentes.

Después de procesar los datos, se genera una base de conocimiento en Microsoft Excel del paquete de Office y se procede a un filtrado de control de la información para establecer un modelo de datos con el que trabajar. Se ejecuta el guardado de la base de datos en el formato .csv para que sea compatible con la herramienta Weka a utilizar en el experimento.

En esta base de datos se definen para la aplicación 17 atributos, seis se consideran como datos individuales y exclusivos (sexo, edad, nivel de instrucción, estado civil, lugar de residencia) de cada encuestado; es sobre los atributos restantes que se desarrolla la investigación. Para realizar la clasificación y pre-procesamiento de datos se toma en cuenta los siguientes pasos:

- **Filtrado de datos:** Los datos tienen que estar en un formato que entienda la herramienta WEKA, por lo que se procedió a convertir la base de datos que se encontraba en un formato XLSX a un CSV (delimitado por comas), permitiendo así que sea compatible con la herramienta que se está utilizando. Además, es necesario eliminar valores no válidos, vacíos o que sobrepasan un umbral y que en el modelo aportarían ruido. Una vez filtrados los datos, se verifica que no existan datos innecesarios que perjudiquen nuestro análisis. Por ello, es necesario elegir las variables que más influyen en la predicción sin perjudicar el resultado, como en este caso es necesario predecir el impacto se asume como variable al atributo que mide el impacto, el

cual debe ser evaluado con las técnicas Cross-folds Validation(Refaeilzadeh, 2008) y Porcentaje Split 66%(Larvin, 2017).

- Uso de algoritmos para obtener un modelo de conocimiento: a continuación, se aplicaron diversos algoritmos de conocimiento que mediante patrones de comportamiento que observan en las variables realizan asociaciones de las mismas para finalmente llegar a determinar con que algoritmo se realizaran los respectivos análisis.
- Interpretación y evaluación: finalmente es necesario comprobar que los resultados del modelo son válidos. Por ello es importante realizar comparaciones entre los distintos algoritmos para procurar escoger el mejor, es decir, el que menor error cometa en la predicción.

Técnicas e instrumentos de investigación

Para el desarrollo de esta investigación se utilizarán los siguientes parámetros: algoritmos clasificación, herramienta Weka e indicadores de criterios de investigación, ver tabla 1.

Tabla 1.

Técnicas e instrumento de experimentación.

Algoritmos de clasificación	KStar IBk RandomForest BayesNet NaiveBayes NaiveBayes MultinomialText NaiveBayesUpdateable Logistic MultilayerPerceptron SimpleLogistic SMO LWL ClassificationViaRegression DecisionStump J48 LMT RandomT ree REPTree
-----------------------------	--

Indicador o criterios de medición	Correctly Classified Instances Relative absolute error ROC Area
-----------------------------------	---

Instrumento Weka 3.7

Fuente: Elaboración del autor

Resultados de la técnica 10-fold Cross-Validation

Para estimar cuán precisa es la predicción de este modelo se aplica la técnica del *k-fold Cross Validation*. En este caso se empleó una validación cruzada con 10 iteraciones para evaluar los resultados de los algoritmos.

Tabla 2.

Resultados para la Técnica Cross-Validation.

ALGORITMO	CORRECTLY CLASSIFIED INSTANCES	RELATIVE ABSOLUTE ERROR	ROC AREA
<i>KStar</i>	67,453%	50,382%	0,8720
<i>IBk</i>	65,094%	54,082%	0,7350
<i>RandomForest</i>	64,623%	73,032%	0,7810
<i>BayesNet</i>	57,076%	75,824%	0,7280
<i>NaiveBayes</i>	59,434%	76,285%	0,7240
<i>NaiveBayesMultinomialText</i>	52,830%	100,000%	0,3320
<i>NaiveBayesUpdateable</i>	59,434%	76,285%	0,7240
<i>Logistic</i>	56,132%	66,327%	0,6980
<i>MultilayerPerceptron</i>	60,849%	60,825%	0,7630
<i>SimpleLogistic</i>	57,547%	82,2186%	0,7190
<i>SMO</i>	52,830%	103,055%	0,6830
<i>LWL</i>	54,245%	91,259%	0,7260
<i>Classification Via Regression</i>	60,849%	76,658%	0,7630
<i>DecisionStump</i>	52,830%	95,385%	0,5480
<i>J48</i>	57,547%	74,255%	0,6760
<i>LMT</i>	60,377%	66,648%	0,7300
<i>RandomTree</i>	60,849%	58,850%	0,7340
<i>REPTree</i>	50,472%	88,794%	0,6470

Fuente: Datos de los instrumentos aplicados por el autor.

Los resultados obtenidos en los procesos se resumen en la Tabla 2 en la que se aprecia que el *KStar* logra un mayor porcentaje de clasificación con un valor de 67,453 % con respecto al resto de los algoritmos. Así mismo, logra un menor error relativo al proyectar una estimación 50,382 % en este acápite y presenta un valor mayor del resto de algoritmos en el área bajo la curva al exhibir un valor de 0,8720. Es así que se determina que el *KStar* es el algoritmo con mejor resultado de manera global.

Otros algoritmos que presentan valores significativos de instancias correctamente clasificadas son el *IBk* y el *RandomForest* con proporciones poco mayores del 65 y 64 porciento. Para el valor de error absoluto relativo destacan, nuevamente, el *IBk*, con un 54,082 % y el *RandomTree* con un 58,850 %.

En el área bajo la curva sobresalen el *RandomForest* con 0,7810 y el *MultilayerPerceptron* con 0,7630. Se puede apreciar claramente el porcentaje de precisión de las instancias que fueron analizadas, obteniendo mejores resultados el algoritmo *lazy.Kstar* con un 67,453% seguido por el algoritmo *lazy.Ibk* con un 65,094 % y posterior a este el *trees.RandomForest* con un 64,623 %.

Resultados de la Técnica Percentage Split 66 %

En la Tabla 3 se exponen los resultados de los algoritmos para las instancias correctamente clasificadas, el error relativo absoluto y el área bajo la curva, para analizar cuál de ellos ofrece un mejor resultado en el experimento del Percentage Split 66%.

Tabla 3.

Resultados para la técnica de clasificación Percentage Split 66%.

ALGORITMO	CORRECTLY CLASSIFIED INSTANCES	RELATIVE ABSOLUTE ERROR	ROC AREA
KStar	63,8889 %	57,1312 %	0,8470
IBk	63,8889 %	58,2150 %	0,7450

RandomForest	62,5000 %	75,7858 %	0,7560
BayesNet	52,7778 %	86,5378 %	0,6570
NaiveBayes	54,1667 %	85,1169 %	0,6480
NaiveBayes MultinomialText	48,6111 %	100,0000 %	0,5000
NaiveBayesUpdateable	54,1667 %	85,1169 %	0,6480
Logistic	54,1667 %	79,3435 %	0,6960
MultilayerPerceptron	58,3333 %	73,5841 %	0,6820
SimpleLogistic	50,0000 %	81,6175 %	0,6760
SMO	58,3333 %	102,6866 %	0,6290
LWL	55,5556 %	92,1588 %	0,6650
ClassificationViaRegression	55,5556 %	88,9882 %	0,6350
DecisionStump	48,6111 %	97,2808 %	0,5080
J48	47,2222 %	89,4656 %	0,5760
LMT	58,3333 %	72,4132 %	0,6820
RandomTree	58,3333 %	71,7204 %	0,6500
REPTree	51,3889 %	93,3782 %	0,5680

Fuente: Datos de los instrumentos aplicados por el autor

Para este test se destaca nuevamente el algoritmo *KStar* al presentar valores superiores en los métodos analizados en la investigación. En las instancias correctamente clasificadas arroja valor de 63,8889 %, para el error relativo absoluto una estimación de 57,1312 % y en el área bajo la curva un valor de 0,8470. Aunque; en este caso, el algoritmo *IBk* tiene resultados muy cercanos al del *KStar* al presentar valores de 63,8889 % en las instancias correctamente clasificadas, un 58,2150 % en el error relativo absoluto y 0,7450 para el área bajo la curva. Como muestran estos resultados el *KStar* es el algoritmo que mejor comportamiento tuvo teniendo en cuenta que es el que obtuvo resultados superiores durante los experimentos.

Es posible establecer analogías y comparaciones, a decir de Méndez et al., (2024), las alternativas comunitarias para un desarrollo sostenible presentan principios teóricos y conceptuales hacia un futuro justo y equitativo, por ser las alternativas comunitarias de desarrollo sostenible (ACDS) una respuesta a desafíos sociales, ambientales y económicos que enfrentan los países. Estas iniciativas, surgidas desde las propias comunidades, proponen modelos de

desarrollo basados en la cooperación, la autogestión y la responsabilidad ambiental, en contraposición al modelo de desarrollo tradicional.

Una revisión crítica de la literatura y de experiencias comunitarias, examinando los principios centrales de las propuestas teóricas alternativas y los conceptos más relevantes para el análisis de las ACDS, destacan enfoques como la ecología social, la economía del decrecimiento y el enfoque de derechos humanos, que convergen en la necesidad de priorizar justicia social, equidad económica y protección ambiental.

Conclusiones

El estudio realizado acerca de los algoritmos de predicción y del aprendizaje automático, refleja información suficiente, que confirman sus potencialidades de aplicabilidad como vía para la predicción del impacto social de infocentros comunitarios en el Cantón La Maná.

El aprendizaje automático, como regularidad, se caracteriza por una serie de técnicas, algoritmos y métodos que permiten, a partir de datos e información suministrada hacer predicciones de cambios y transformaciones del ámbito social, provocados por los servicios que brindan los infocentros comunitarios.

La Herramienta ofrece variedades de técnicas y algoritmos necesarios que hacen considerarla una herramienta eficaz para ser utilizada en la predicción del impacto social de infocentros comunitarios en el Cantón La Maná, además de facilitar la aplicación de los pasos para la generalización e interpretación de los algoritmos de predicción.

Del análisis de los resultados y la interpretación de los mismos, reflejan que el algoritmo Lazy KStar es el más aceptado con respecto al resto de algoritmos de clasificación/regresión (predicción), para valorar el impacto social de infocentros comunitarios en el Cantón La Maná de Ecuador, lo que corrobora la hipótesis de trabajo de la investigación desarrollada.

Referencias

- Álvarez-Díaz, M. & Álvarez, A. (2006). Predicción No-Lineal de Tipos de Cambio: Algoritmos Genéticos, Redes Neuronales y Fusión de Datos. ISME-DSEA Department of Electrical Engineering, University of Pisa.
- <http://webx06.webs.uvigo.es/sites/default/files/wp0205.pdf>
- Blanco, E. J., & Hermida, H. (2016). Algoritmos de clustering y aprendizaje automático aplicados a Twitter. Upcommons.
- Calderón, M. J. , & Acuña, J.. (2017). Conectividad rural y cambio social: los Infocentros Comunitarios en el Ecuador. Revista Publicando, 4(11), 190-207.
- Espino Timón, C. (2017). Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso (Trabajo de Fin de Grado). Universidad Oberta de Catalunya, España.
- <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/59565/6/caresptimTFG0117mem%C3%B3ria.pdf>
- García González, F. J. (2011). Aplicación de técnicas de Minería de Datos del Centro Andaluz del Medio Ambiente. Masteres.
- García Serrano, A. (2012). Inteligencia Artificial. Fundamentos, prácticas y aplicaciones. RC Libros.
- González, P. (2006). Desarrollo de técnicas de minería de datos en procesos industriales: Modelización en líneas de producción de acero. Universidad de la Rioja.

Gonzalo Vega, M. (2005). Análisis de acceso y uso de los Infocentros venezolanos (Trabajo de grado presentado para optar al título de Magíster en Ciencias de la Comunicación). Universidad de Zulia, República Bolivariana de Venezuela.

Hernández, J., Ramírez, M. J., & Ferri, C. (2004). Introducción a la Minería de Datos. Madrid: Pearson Prentice Hall.

Jover, J. N. (2013). OEI - Programación - CTS+I - Sala de lectura.
<https://www.oei.es/historico/salactsi/nunez06.htm>

Lavin, F. (2017). Weka. Percentage split option: <http://weka.8497.n7.nabble.com/Percentage-split-option-td39126.html>

Limón Peréz, M. (2016). Construcción de un prototipo de un programa personalizado de tipo Chatbot en ambiente Java con un Lenguaje Natural. Construcción de un prototipo de un programa personalizado de tipo Chatbot en ambiente Java con un Lenguaje Natural. Mexico.

Martínez López, Y., & Madera Quintana, J. (2016). Estudio del comportamiento del Algoritmo K *. ResearchGate.

Morales Hernández, A. (2014). Construcción de Sistemas Multiclasificadores usando Algoritmos Genéticos y Medidas de Diversidad (Tesis en Opción al Título de Licenciado en Ciencia de la Computación). Universidad Central «Marta Abreu» de Las Villas, Santa Clara.

Méndez López, A. J., Cabrera Ruiz, I. I., & Pajón Naranjo, A. (2024). Alternativas comunitarias para un desarrollo sostenible en Colombia: Principios teóricos y conceptuales hacia un futuro justo y equitativo. *Didáctica y Educación*, 15(2), 318–347.
<https://revistas.ult.edu.cu/index.php/didascalia/article/view/2002>