

Parking garage management using classification techniques, fuzzy logic and evolutionary algorithms

Mariluna Apaza¹, Juan Manotupa¹, Fidel Salas Flores y José Sullato Torres, Mgter
¹Universidad Católica de Santa María, Arequipa, mariluna@hotmial.com, jmanotupa@gmail.com,
fk.salasf@gmail.com, jsullato@ucsm.edu.pe

Abstract– *This article presents the analysis of the data of a garage for light and heavy vehicles to extract enough information that allows us to manage the parking and can predict the number of free spaces in a certain day, as well as to find the potential customers. This was done using the KDD methodology, for data mining we use the WEKA, KEEL and Matlab tools that provide us with different algorithms such as J48, Bayes BayesNet, KNN-C, DT_GA-C and GANN-C to do the evaluation. As a result, the possible potential customers of the company were classified and, in addition, the best result is obtained with the J48 algorithm, which obtained 90.43% accuracy and an error percentage of 0.4553 using the GANN-C algorithm.*

Keywords- *Parking, Space management, potential clients Data Mining, Fuzzy Logic, Evolutionary Algorithms*

Digital Object Identifier (DOI):<http://dx.doi.org/10.18687/LACCEI2018.1.1.209>
ISBN: 978-0-9993443-1-6
ISSN: 2414-6390

Gestión de aparcamientos de cocheras utilizando técnicas de clasificación, lógica difusa y algoritmos evolutivos

Mariluna Apaza-Luque¹, Juan Manotupa-Gil¹, Fidel Salas-Flores¹ y José Sullata-Torres, Mgter¹

¹Universidad Católica de Santa María, Arequipa-Perú, mariluna_lulu@hotmail.com, jmanotupa@gmail.com, fk.salasf@gmail.com, jsullata@ucsm.edu.pe

Abstract— This article presents the analysis of the data of a garage for light and heavy vehicles to extract enough information that allows us to manage the parking and can predict the number of free spaces in a certain day, as well as to find the potential customers. This was done using the KDD methodology, for data mining we use the WEKA, KEEL and Matlab tools that provide us with different algorithms such as J48, Bayes BayesNet, KNN-C, DT_GA-C and GANN-C to do the evaluation. As a result, the possible potential customers of the company were classified and, in addition, the best result is obtained with the J48 algorithm, which obtained 90.43% accuracy and an error percentage of 0.4553 using the GANN-C algorithm.

Keywords- Parking, Space management, potential clients, Data Mining, Fuzzy Logic, Evolutionary Algorithms.

Resumen— En este artículo se presenta el análisis de los datos de una cochera para vehículos ligeros y pesados con el fin de extraer la información suficiente que nos permita gestionar el aparcamiento y sea capaz de predecir el número de plazas libres en un determinado día, así como encontrar los clientes potenciales. Esto se realizó utilizando la metodología de KDD, para la minería de datos usamos las herramientas WEKA, KEEL y Matlab que nos proporcionan diferentes algoritmos como J48, Bayes BayesNet, KNN-C, DT_GA-C y GANN-C para hacer la evaluación. Como resultado se clasificó los posibles clientes potenciales que tiene la empresa y además se obtiene mejor resultado con el algoritmo J48 que se obtuvo 90.43% de exactitud y con un porcentaje de error del 0.4553 utilizando el algoritmo GANN-C.

Palabras clave— Estacionamiento, Gestión de espacio, clientes potenciales, Minería de datos, Lógica difusa, Algoritmos evolutivos.

I. INTRODUCCIÓN

En la actualidad varias empresas de estacionamiento de vehículos no gestionan bien el espacio con el que cuentan para el almacenamiento temporal de vehículos por lo cual la distribución de espacio de los vehículos no se lleva de manera correcta, esto implica que los espacios de estacionamiento no se organicen bien y se pierda el cupo donde otros vehículos se pudieron haber estacionado [1].

Dentro del mismo estacionamiento, el almacenamiento de datos no se registra de manera correcta esto ocasiona que no se cuente con un registro adecuado de las empresas que almacenan sus vehículos, así mismo, no se conoce a los clientes potenciales y esto ocasiona que la empresa genere pérdidas; por tal motivo, es una de las cosas más importantes

para los gerentes de estacionamiento y los planificadores o investigadores correspondientes[2].

Se han realizado estudios sobre esta problemática de aparcamiento que necesita pasar mucho tiempo para inspeccionar todas las posibilidades [3]. En el trabajo de Calot, Maluf y Neffa (2017), describen un sistema de bajo costo para la gestión eficiente de espacios de estacionamiento, con un sistema que detectan el ingreso de vehículos, permitiendo así conocer en tiempo real el número de patente, calcular los niveles de ocupación del estacionamiento y recolectar información sobre servicios adicionales solicitados por el cliente [4]. Vera-Gomez et al. (2016) indican que existen varios estacionamientos inteligentes en el mundo y que gracias a la ayuda de diferentes dispositivos pueden observar los estacionamientos que se encuentran libres [5]. En el trabajo de Cabrera-Cósetl, Mora-Álvarez y Rubén Alejos-Palomares (2009) con el apoyo de la compañía Volvo realizó un estudio para obtener un sistema que estaciona automáticamente un automóvil que controla el volante y los frenos de un Volvo S60 obteniendo una resistencia satisfactoria mediante lógica difusa. En este punto para obtener que un vehículo se estacione solo, ayuda a mejorar la distribución del espacio que se tiene [6].

Desde esa perspectiva, la minería de datos aporta con sus técnicas para poder clasificar y predecir mediante diferentes algoritmos los problemas de servicios públicos [7]. Dentro de los intereses del estudio está la optimización y gestión de los datos de los espacios ocupados [8] y minimización de cruces de vehículos en el estacionamiento [9].

En ese sentido, existe la necesidad imperiosa de contar con un sistema seguro, inteligente, eficiente y confiable que pueda usarse para buscar en el estacionamiento desocupado, orientarlo hacia el estacionamiento, negociar la tarifa de estacionamiento, junto con la administración adecuada del estacionamiento. El objetivo es mejorar la gestión de aparcamiento de las cocheras, así que formulamos las siguientes preguntas ¿Cómo se puede organizar mejor los estacionamientos de los vehículos? ¿Es posible separar lugares para clientes con mayor uso del estacionamiento? La respuesta para las preguntas realizadas se obtendrá realizando el debido procesamiento de los datos obtenidos por la empresa para llevar a su mejoramiento en cuanto al almacenamiento de los vehículos, esto quiere decir que una vez procesado sus datos

con las diferentes herramientas de minería de datos, se obtendrán datos de calidad donde se podrán realizar diferentes métodos para la obtención de información que ayuda a la empresa a tomar mejores decisiones en cuanto a su distribución y separación de lugares para estacionarse.

El artículo estará dividido por las siguientes partes: Se revisa los materiales y métodos para el estudio, la base de datos del aparcamiento, la recopilación y preparación de datos, el modelado y la evaluación, la interpretación de resultados y por último las conclusiones sobre la evaluación de los algoritmos utilizados.

II. MATERIALES Y MÉTODOS

A. Metodología KDD

La metodología utilizada fue la del KDD (Knowledge Discovery in Databases). KDD es el proceso de encontrar información útil y patrones dentro de los datos [10]. El KDD está compuesto de 5 fases que se muestran a continuación en la Fig. 1.

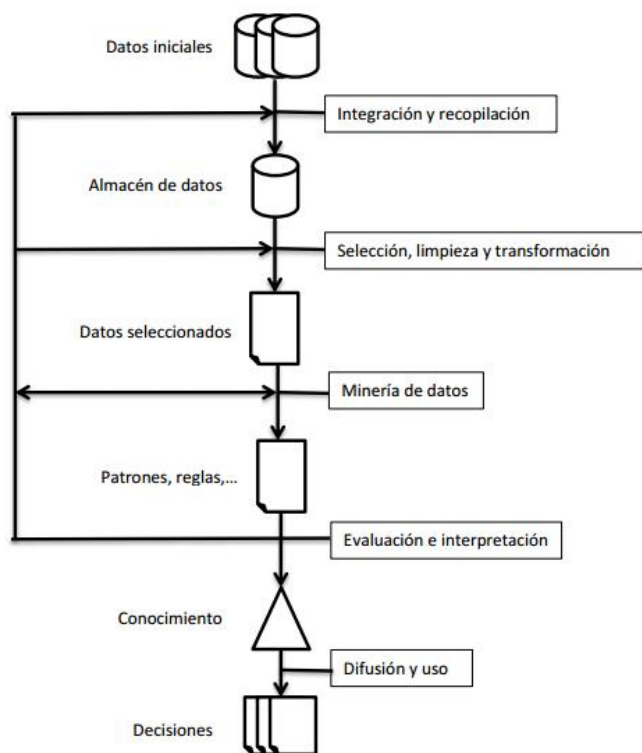


Fig. 1 Fases de KDD.

B. Recolección y preparación de datos

1) Recopilación de datos:

La recopilación de los datos se obtuvieron de la empresa de estacionamiento de vehículos ubicado en

la ciudad de Arequipa. Los datos se almacenan en los 2048 registros dentro de diferentes hojas de Excel, con los datos obtenidos se realizaron los diversos tipos de preprocesamiento para después utilizar las herramientas de minería de datos para obtener información de importancia en la mejora de toma de decisiones de la empresa de estacionamiento.

2) Descripción de los atributos:

Los atributos se muestran en la Tabla I, almacenan varios tipos de registros de una manera no muy ordenada lo cual impedía que la empresa pudiera hacer una buena gestión de los mismos; estos son los siguientes atributos.

TABLA I
ATRIBUTOS DE LOS REGISTROS DE VEHÍCULOS

Nombre de atributo	Tipo	Descripción
Categoría	Texto	Indica el tipo de vehículo que ingresaba a la cochera.
Placa	Texto	Es la placa de cada vehículo que ingresa a la cochera.
Fecha ingreso	Fecha	Es la fecha en la que se ingresó el vehículo la cochera
Fecha salida	Fecha	Es la fecha en la que se retiró el vehículo la cochera
Hora salida	Hora	Es la hora en la que el vehículo dejó la cochera
Días	Numérico	Es el número de días que el vehículo estuvo en la cochera
Importe	Moneda	Es el precio que pagó por los días que guardó el vehículo en la cochera

3) Preparación de datos

Después de obtener los datos, se realizaron el preprocesamiento con el fin de obtener datos de calidad [11].

Primero se elimina la columna que no tiene atributos filtro. Después se tiene los registros de los clientes que no cuenta con una columna propia y si no estaba dividida por segmentos lo que se hizo es crear una columna para los clientes y de esta manera ordenar los datos. Se reordenaron las categorías, después de ello se puso los datos llenaron los datos en blanco en Weka [12].

Una vez que los datos han sido preprocesados, ya son datos de calidad los cuales pueden ser sometidos a diferentes herramientas con las cuales se podrá obtener la información necesaria para el mejoramiento en cuanto al servicio brindado por la empresa de estacionamientos.

C. Modelado y evaluación

Al ser nuestro objetivo la predicción de un modelo tenemos dos tipos de modelado posibles, modelos de clasificación o modelos de regresión, pero al tratarse el

atributo de salida de una variable texto, tan solo podemos realizar los modelos de clasificación.

Dentro de los modelos de clasificación existen varios algoritmos. Nosotros nos centraremos en dos de ellos: Árbol de decisión y BayesNet por ser los que usan más comúnmente en la literatura [13].

III. RESULTADOS

Se ejecutará cada una de los algoritmos: Árbol de decisión (J48) [14] y BayesNet [15] en la herramienta Weka, para comprobar cuál es el mejor. Con el fin de poder comparar de manera correcta los diferentes resultados el algoritmo utiliza división porcentual del 66%. A continuación, mostramos los resultados por algoritmo.

A. Árbol de decisión (J48)

La evaluación del algoritmo J48 muestra los resultados que se muestran en la Fig. 2:

```

=== Summary ===

Correctly Classified Instances      1852      90.4297 %
Incorrectly Classified Instances    196       9.5703 %
Kappa statistic                    0.8849
Mean absolute error                 0.0166
Root mean squared error             0.0922
Relative absolute error             15.6746 %
Root relative squared error         40.0713 %
Total Number of Instances          2048
    
```

Fig.2 Resultado de J48 con los datos del aparcamiento.

Weka también proporciona detalle exacto de la clasificación por clase del algoritmo J48, que se muestra en la Fig. 3

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
0,000	0,000	0,000	0,000	0,000	0,000	?	?	DEL CAR
0,000	0,000	0,000	0,000	0,000	0,000	0,935	0,015	TRANSP. LIBERTAD
0,167	0,000	1,000	0,167	0,286	0,408	0,946	0,185	ZETRANSA
0,100	0,001	0,250	0,100	0,143	0,156	0,951	0,172	FEMACO
0,800	0,000	1,000	0,800	0,889	0,894	0,987	0,804	ATENUZ
0,875	0,006	0,618	0,875	0,724	0,732	0,998	0,760	MEGAFLOTAS
0,519	0,000	1,000	0,519	0,683	0,718	0,969	0,564	ACOINSA
0,767	0,000	0,971	0,767	0,857	0,861	0,985	0,797	MOTA ENGLI
0,869	0,001	0,991	0,869	0,926	0,924	0,991	0,903	MUR
0,861	0,000	1,000	0,861	0,925	0,923	0,983	0,905	INMARFACS
0,675	0,000	1,000	0,675	0,806	0,810	0,981	0,813	SATURNO
0,953	0,001	0,993	0,953	0,973	0,971	0,997	0,969	TILOPSAC
0,943	0,003	0,973	0,943	0,958	0,954	0,992	0,950	DCR MINERIA Y CONSRU
0,938	0,000	1,000	0,938	0,968	0,964	0,994	0,970	TRITON
0,982	0,000	1,000	0,982	0,991	0,990	0,999	0,991	SOL DEL PACIFICO
0,963	0,118	0,769	0,963	0,855	0,797	0,979	0,940	TERCERO
0,904	0,035	0,919	0,904	0,902	0,885	0,988	0,922	

Fig.3 Detalles de clasificación Weka J48.

La matriz de confusión del desempeño del algoritmo J48 se muestra en la Fig. 4.

```

-- Confusion Matrix ---

a b c d e f g h i j k l m n o p <-- classified as
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | a = DEL CAR
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | b = TRANSP. LIBERTAD
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5 | c = ZETRANSA
0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 8 | d = FEMACO
0 0 0 0 4 0 0 0 0 0 0 0 0 0 0 0 0 1 | e = ATENUZ
0 0 0 0 0 21 0 0 0 0 0 0 0 0 0 0 0 3 | f = MEGAFLOTAS
0 0 0 0 0 0 14 0 0 0 0 0 0 0 0 0 13 | g = ACOINSA
0 0 0 0 0 0 0 33 0 0 0 0 0 0 0 0 10 | h = MOTA ENGLI
0 0 0 0 0 0 0 0 106 0 0 0 0 0 0 0 16 | i = MUR
0 0 0 0 0 0 0 0 0 118 0 0 0 0 0 0 19 | j = INMARFACS
0 0 0 0 0 0 0 0 0 0 110 0 0 0 0 0 53 | k = SATURNO
0 0 0 0 0 0 0 0 0 0 0 142 0 0 0 0 7 | l = TILOPSAC
0 0 0 0 0 0 0 0 0 0 0 0 182 0 0 11 | m = DCR MINERIA Y CONSR
0 0 0 0 0 0 0 0 0 0 0 0 0 1 273 0 17 | n = TRITON
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 275 5 | o = SOL DEL PACIFICO
0 0 0 3 0 12 0 1 1 0 0 1 4 0 0 572 | p = TERCERO
    
```

Fig.4 Matriz de confusión Weka J48

B. Bayes BayesNet

La evaluación del algoritmo BayesNet muestra los resultados que se muestran en la Fig. 5.

```

=== Summary ===

Correctly Classified Instances      1804      88.0859 %
Incorrectly Classified Instances    244      11.9141 %
Kappa statistic                    0.8608
Mean absolute error                 0.0176
Root mean squared error             0.1038
Relative absolute error             16.5411 %
Root relative squared error         45.093 %
Total Number of Instances          2048
    
```

Fig.5 Resultado de BayesNet con los datos del aparcamiento.

Weka también proporciona detalle exacto de la clasificación por clase del algoritmo BayesNet, que se muestra en la Fig. 6

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
0,000	0,000	0,000	0,000	0,000	0,000	?	?	DEL CAR
1,000	0,000	0,800	1,000	0,889	0,894	1,000	1,000	TRANSP. LIBERTAD
0,500	0,001	0,600	0,500	0,545	0,547	0,965	0,526	ZETRANSA
0,800	0,001	0,800	0,800	0,800	0,799	0,997	0,852	FEMACO
0,600	0,002	0,375	0,600	0,462	0,473	0,991	0,669	ATENUZ
0,750	0,002	0,818	0,750	0,783	0,781	0,994	0,822	MEGAFLOTAS
0,667	0,003	0,750	0,667	0,706	0,703	0,985	0,777	ACOINSA
0,907	0,005	0,796	0,907	0,848	0,846	0,995	0,951	MOTA ENGLI
0,902	0,019	0,753	0,902	0,821	0,812	0,992	0,942	MUR
0,810	0,013	0,822	0,810	0,816	0,803	0,980	0,905	INMARFACS
0,877	0,006	0,923	0,877	0,899	0,891	0,993	0,963	SATURNO
0,906	0,008	0,894	0,906	0,900	0,892	0,995	0,964	TILOPSAC
0,984	0,016	0,868	0,984	0,922	0,916	0,998	0,986	DCR MINERIA Y CONSRU
0,931	0,021	0,880	0,931	0,905	0,889	0,995	0,979	TRITON
0,986	0,020	0,887	0,986	0,934	0,925	0,999	0,995	SOL DEL PACIFICO
0,800	0,017	0,950	0,800	0,868	0,827	0,981	0,959	TERCERO
0,881	0,015	0,886	0,881	0,880	0,861	0,990	0,959	

Fig.6 Detalles de clasificación Weka BayesNet.

La matriz de confusión del desempeño del algoritmo BayesNet se muestra en la Fig. 7.

```
== Confusion Matrix ==
a b c d e f g h i j k l m n o p <-- classified as
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | a = DEL CAR
0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | b = TRANSP. LIBERTAD
0 0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 1 0 0 1 0 0 | c = ZETRAMSA
0 0 0 0 8 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 | d = FEMACO
0 0 0 0 0 3 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 0 0 | e = ATENUZ
0 0 0 0 0 0 0 18 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 6 1 0 0 | f = MEGAFLOTAS
0 0 0 0 0 0 0 0 18 0 0 0 0 3 0 0 0 2 2 2 0 2 1 0 0 3 1 0 | g = ACOINSA
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 3 1 | h = MOTA ENGLI
0 0 1 0 0 0 0 0 0 0 110 2 0 0 0 2 3 3 3 1 1 0 0 0 0 0 0 0 0 | i = MUR
0 0 1 1 0 0 0 0 1 0 11 111 2 0 1 6 1 2 1 0 0 1 0 0 0 0 0 0 | j = INMARFACS
0 0 0 0 0 0 0 0 3 1 2 1 143 1 1 11 0 0 1 0 0 0 0 0 0 0 0 0 | k = SATURNO
0 0 0 0 0 0 0 0 0 1 3 1 3 135 0 6 0 0 0 1 0 0 1 0 0 0 0 0 | l = TILOFSAC
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 190 1 0 1 1 0 0 0 0 0 0 0 | m = DCR MINERIA Y CONSR
0 0 0 0 0 0 0 0 0 0 3 0 1 0 1 271 7 8 1 0 0 0 0 0 0 0 0 0 | n = TRITON
0 0 0 0 0 0 0 0 0 0 0 4 0 0 0 0 0 0 276 0 1 0 0 0 0 0 0 0 | o = SOL DEL PACIFICO
0 0 0 2 5 3 2 8 13 17 5 14 20 7 23 475 1 0 0 0 0 0 0 0 0 0 0 | p = TERCERO
```

Fig.7 Matriz de confusión Weka BayesNet

A continuación, se muestra cuál ha sido el mejor algoritmo para cada uno de los modelos. Para ello utilizaremos las instancias clasificadas correctamente como elemento de comparación. Ver Tabla II.

TABLA II
COMPARACIÓN DE LOS ALGORITMOS

Algoritmos	Instancias clasificadas correctamente
Árbol de decisión (J48)	90.43%
BayesNet	88.09%

D. Lógica Difusa

Según Zadeh (1964), un conjunto difuso es una clase de objetos con grados de membresía. Tal conjunto se caracteriza por una función de membresía (característica) que asigna a cada objeto un grado de membresía que varía entre cero y uno [16]. Las variables lingüísticas nos permitirá representar los términos lingüísticos del modelamiento.

1) Variables lingüísticas

Se considera las siguientes variables

- Espacios: Se consideran "Espacios" como la unidad mínima de aparcamiento.
- Frecuencia: Se considera la "Frecuencia" como mensual y los días que cada vehículo se queda en la cochera.
- C_Potencial (Cliente Potencial): Aquí podremos determinar a primera instancia cuáles serán nuestros posibles clientes potenciales.

Del análisis de nuestras variables podemos concluir que Espacios y Frecuencia son variables de entrada, mientras que C_potencial es nuestra variable de salida, que dependerá de Espacios y Frecuencia. Ver Fig. 8.

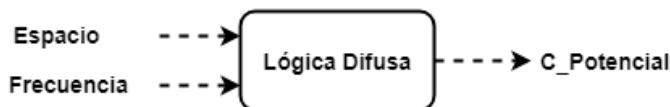


Fig. 8 Variables de entrada y salida para nuestro caso.

2) Funciones de pertenencia y rango

Para la variable Espacio (espacio de aparcamiento): El espacio de aparcamiento variará entre 1 y 12. Este rango ha sido dividido en cuatro regiones dependiendo del tamaño del vehículo. A cada uno de estos tramos o espacios se les ha asignado valores lingüísticos difusos, los mismos que son totalmente diferentes respecto al espacio que estos ocupan en el aparcamiento. A continuación, se describen en detalle los valores y rangos mencionados para la variable espacio. Ver Tabla III.

TABLA III
VALORES Y RANGOS PARA LA VARIABLE ESPACIO

Valores	Rango	Descripción
Pequeño	1,2	Espacio pequeño
Regular	2,4	Espacio Regular
Grande	4,6	Espacio Grande
SGrande	6,12	Espacio Grande Super Grande

La función de pertenencia a utilizar será triangular, configurada en MatLab utilizando su librería difusa [17], la definimos en forma detallada como:

Parámetros:

- [1, 1, 2]
- [2, 3, 4]
- [4, 5, 6]
- [6, 9, 12]

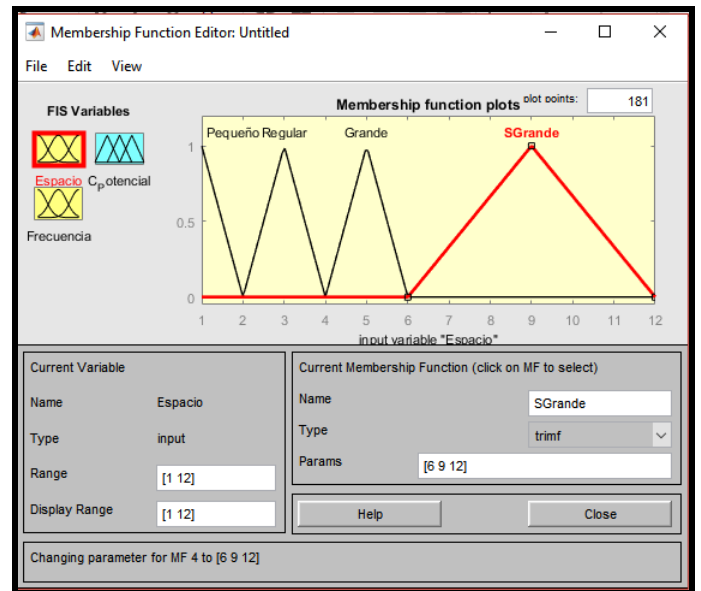


Fig.9 Función de Pertenencia de la variable Espacio.

Para la variable Frecuencia (Días que se queda en la cochera): El espacio de aparcamiento variará entre 1 y 30. Este rango ha sido dividido en cuatro regiones dependiendo del tamaño del vehículo. A cada uno de estos tramos o espacios se les ha asignado valores

lingüísticos difusos, los mismos que son totalmente diferentes respecto a la frecuencia en la que los clientes dejan su vehículo en el aparcamiento. A continuación, se describen en detalle los valores y rangos mencionados para la variable Frecuencia. Ver Tabla IV.

TABLA IV
VALORES Y RANGOS PARA LA VARIABLE FRECUENCIA

Valores	Rango	Descripción
Esporádico	1,4	Frecuencia esporádica
Regular	4,9	Frecuencia regular
Frecuente	9,2	Frecuencia frecuente
M_Frecuente	20,3	Frecuencia muy frecuente

La función de pertenencia a utilizar será Gaussiana, que será cargada en MatLab. La definimos en forma detallada en la Fig. 10.

Parámetros:

- [3 , 4]
- [3 , 9]
- [3 , 20]
- [3 . 30]

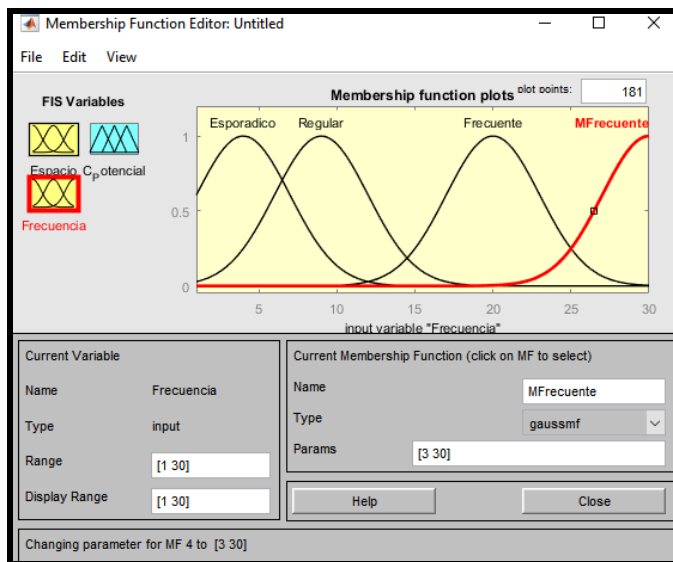


Fig.10 Función de Pertenencia de la variable Frecuencia.

Para la variable C_Potencial (Posible Cliente Potencial): El cliente Potencial variará entre 0 a 10. Este rango ha sido dividido en tres regiones para identificar la probabilidad del cliente potencial. A cada uno de estos tramos o espacios se les ha asignado valores Lingüísticos difusos, los mismos que son totalmente diferentes respecto a la evaluación de las reglas preestablecidas. A continuación, se describen en detalle los valores y rangos mencionados para la variable C_Potencial. Ver Tabla V.

TABLA V
VALORES Y RANGOS PARA LA VARIABLE C_POTENCIAL

Valores	Rango	Descripción
Bajo	0,3	Cliente potencial bajo
Medio	4,7	Cliente potencial medio
Alto	8,1	Cliente potencial alto

La función de pertenencia a utilizar será triangular, que será cargada en MatLab. La definimos en forma detallada en la figura N.º 14.

Parámetros:

- [0, 2, 3]
- [3, 5.5 , 7]
- [7, 9 ,10]

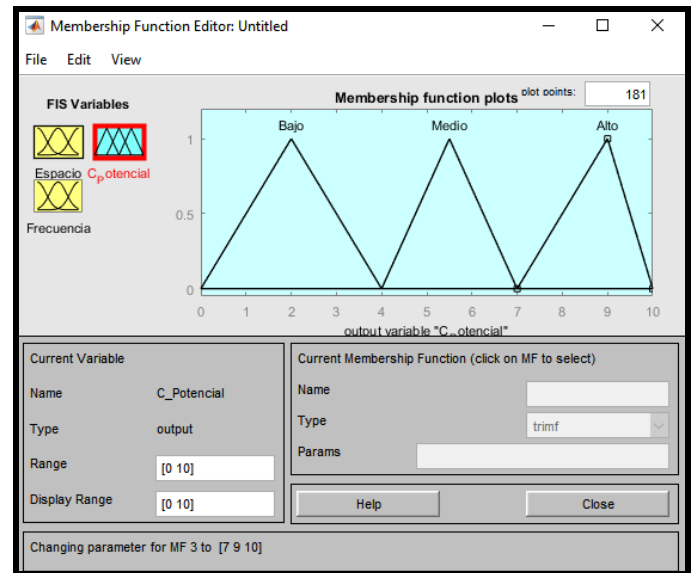


Fig.11 Función de Pertenencia de la variable C_Potencial.

3) Reglas

- Si Espacios (Pequeño) y Frecuencia (Esporádico) entonces C_Potencial (Bajo)
- Si Espacios (Pequeño) y Frecuencia (M_Frecuente) entonces C_Potencial (Medio)
- Si Espacios (Pequeño) o Frecuencia (Regular) entonces C_Potencial (Bajo)
- Si Espacios (Regular) y Frecuencia (Esporádico) entonces C_Potencial (Medio)
- Si Espacios (Regular) y Frecuencia (Frecuente) entonces C_Potencial (Medio)
- Si Espacios (Grande) y Frecuencia (Esporádico) entonces C_Potencial (Bajo)
- Si Espacios (Grande) y Frecuencia (Frecuente) entonces C_Potencial (Alto)
- Si Espacios (SGrande) y Frecuencia (M_Frecuente) entonces C_Potencial (Alto)
- Si Espacios (SGrande) y Frecuencia (Esporádico) entonces C_Potencial (Medio)

Digital Object Identifier: (to be inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).

- Si Espacios (SGrande) y Frecuencia (Regular) entonces C_Potencial (Medio)

E. Algoritmos evolutivos

Los algoritmos evolutivos se utilizan en modelos computacionales de imitan la evolución de los individuos que forman una población, y que a través de operadores de selección y búsqueda tienen el objetivo de optimizar los resultados [18]. Uno de los más utilizados son los Algoritmo Genéticos (GA, por sus siglas en inglés).

Utilizando la herramienta Keel [19], se eligió el tipo de problema clasificación y se cargó la data de aparcamiento, anteriormente ya procesada, se eligió el algoritmo de selección de conjunto de entrenamiento evolucionario CHC-TSS, y luego se escogió tres algoritmos evolutivos populares para su análisis como son:

- Lazy Learning (kNN-C) [20]
- Decision Trees (DT_GA-C) [21]
- Evolutionary Neutral Networks (GANN-C)[22]

Se eligió el algoritmo Clas-FiedmanAlignedST para las pruebas de clasificación. El modelo experimental construido se muestra en la Fig. 12.

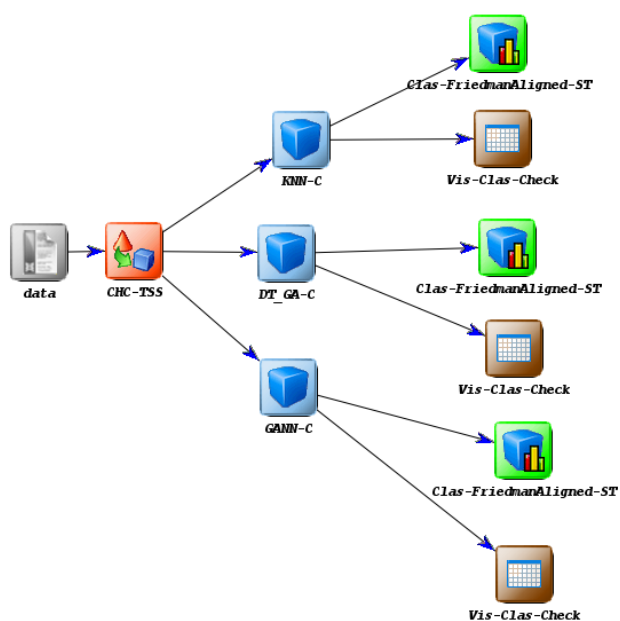


Fig.12 Diagrama de experimento en la herramienta Keel.

1) Lazy Learning (kNN-C)

En la Tabla VI se muestran los resultados de la evaluación del algoritmo KNN-C.

TABLA VI
RESULTADOS DEL ALGORITMO KNN-C EN KEEL

Global Classification Error + N/C	stddev Global Classification Error + N/C	Calsificación Correcta	Global B/C
0.4110	0.0620	0.5890	0.0000

2) Árbol de decisión (DT-GA-C)

En la Tabla VII se muestran los resultados de la evaluación del algoritmo DT-GA-C.

TABLA VII
RESULTADOS DEL ALGORITMO DT-GA-C EN KEEL

Global Classification Error + N/C	stddev Global Classification Error + N/C	Calsificación Correcta	Global B/C
0.4477	0.0790	0.5523	0.0000

3) Redes Neuronales Evolutivas (GANN-C)

En la Tabla VIII se muestran los resultados de la evaluación del algoritmo GANN-C.

TABLA VIII
RESULTADOS DEL ALGORITMO GANN-C EN KEEL

Global Classification Error + N/C	stddev Global Classification Error + N/C	Calsificación Correcta	Global B/C
0.4553	0.0553	0.5447	0.0000

En la Tabla IX se muestra el cuadro comparativo entre los tres algoritmos evaluados con sus resultados obtenidos. Para ello se utilizó la métrica de *Global classification Error* como elemento de comparación.

TABLA IX
COMPARACIÓN DE LOS ALGORITMOS POR PORCENTAJE DE ERROR

Algoritmo	Resultado
Lazy Learning (kNN-C)	0.4110
Decision Trees (DT_GA-C)	0.4477
Evolutionary Neutral Networks (GANN-C)	0.4553

IV. INTERPRETACIÓN DE RESULTADOS

Al haber realizado el tratamiento de los datos, limpieza y el adecuado preprocesamiento, se pudo lograr los resultados esperados aplicando la clasificación de los datos.

Se aplico las técnicas de árboles de decisión como es el J48, y nos ayudó a identificar la periodicidad de la concurrencia de diferentes clientes en distintas fechas del año, de este modo sabremos con técnicas de predicción cuando reservar algunos cupos en la cochera para clientes frecuentes o captar clientes potenciales según la inteligencia de negocios que decidamos usar en determinado momento.

Aplicamos también lógica difusa para poder identificar entradas que nos sirvieron para el análisis de las reglas y de esta manera determinar los posibles clientes potenciales según el cruce de posibilidades de nuestras entradas. Como se aprecia en la Fig. 13, los clientes potenciales son aquellos que evidencian una mayor Frecuencia y mayor Espacio de uso del aparcamiento

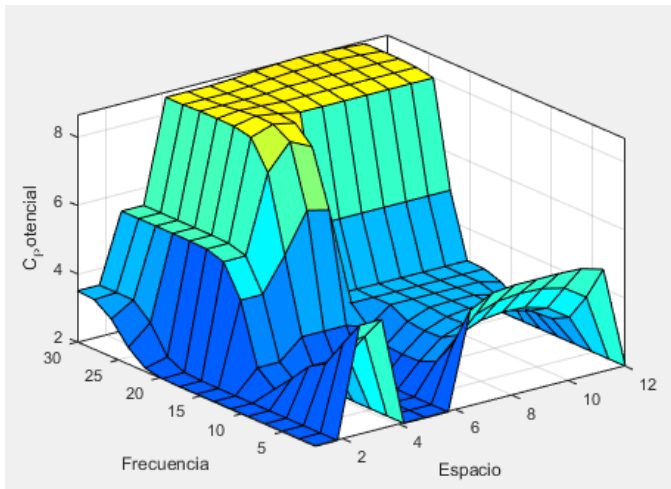


Fig.13 Surface de Resultados de reglas propuestas

V. CONCLUSIONES

Haciendo uso de las técnicas de minería de datos se puede modelar las soluciones para los problemas de exactitud en la predicción con datos históricos para determinar cupos disponibles a clientes potenciales.

La técnica de clasificación por árbol de decisión de J48 mostró la mejor exactitud con un valor de 90.43%; esto se traduce en datos más precisos que generan información suficiente para la toma de decisiones en la empresa.

La librería difusa de MatLab ayudó a calificar los valores difusos que se encuentra en la base de datos de aparcamiento para el análisis de las reglas y de esta manera determinar que los posibles clientes potenciales son aquellos que evidencian una mayor Frecuencia y un mayor Espacio de uso.

El uso de los algoritmos evolutivos ha permitido obtener un mejor resultado utilizando el algoritmo Evolutionary Neural Networks (GANN-C) con un porcentaje de error del 0.4553 frente a otros algoritmos que generaron valores menores.

Utilizando las herramientas se pudo predecir quiénes son nuestros clientes potenciales a los cuales se podrá separar y asegurar su espacio dentro del estacionamiento de los vehículos, además de conocer el tiempo que se quedaría dentro de la cochera, esto llevaría a tener mejores ofertas y/o promociones para estos clientes, también se podrá organizar mejor el espacio y cuáles son los vehículos que podrían llegar durante la semana y así poder separar el espacio para los clientes potenciales.

REFERENCIAS

[1] Faheem, S. A. Mahmud, G. M. Khan, M. Rahman, and H. Zafar, "A survey of intelligent car parking system," *J. Appl. Res. Technol.*, vol. 11, no. 5, pp. 714–726, 2013.

[2] N. Chen, L. Wang, L. Jia, H. Dong, and H. Li, "Parking Survey Made Efficient in Intelligent Parking Systems," *Procedia Eng.*, vol. 137, pp. 487–495, 2016.

[3] T. Lin, H. Rivano, and F. Le Mouel, "A Survey of Smart Parking Solutions," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 12, pp. 3229–3253, 2017.

[4] E. Calot, M. Maluf, and M. Neffa, "Estacionamiento Inteligente con IoT," 2017.

[5] J. A. Vera-Gómez, A. Quesada-Arencibia, C. R. García, R. S. Moreno, and F. G. Hernández, "An intelligent parking management system for Urban areas," *Sensors (Switzerland)*, vol. 16, no. 6, 2016.

[6] R. Cabrera-Cosetl, M. Z. Mora-Alvarez, and R. Alejos-Palomares, "Self-Parking System Based in a Fuzzy Logic Approach," *2009 Int. Conf. Electr. Commun. Comput.*, pp. 119–124, 2009.

[7] E. Al Nuaimi, S. Al Marzooqi, and N. Zaki, "Predicting the decision for the provision of municipal services using data mining approaches," in *2016 3rd MEC International Conference on Big Data and Smart City, ICBDSOC 2016*, 2016, pp. 312–317.

[8] P. R. Adki and J. Agarkhed, "Cloud assisted time-efficient vehicle parking services," in *Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016*, 2017, vol. 1.

[9] C. Tsiaras, L. Hobi, F. Hofstetter, S. Liniger, and B. Stiller, "ParkITsmart: Minimization of cruising for parking," in *Proceedings - International Conference on Computer Communications and Networks, ICCCN*, 2015, vol. 2015–Octob.

[10] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The KDD process for extracting useful knowledge from volumes of data," *Commun. ACM*, vol. 39, no. 11, pp. 27–34, 1996.

[11] S. García, J. Luengo, and F. Herrera, "Data Preprocessing in Data Mining," *Intell. Syst. Ref. Libr.*, vol. 72, 2015.

[12] M. A. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor.*, vol. 11, no. 1, pp. 10–18, 2009.

[13] C. Xie, J. Lu, and E. Parkany, "Work travel mode choice modeling with data mining: decision trees and neural networks," *Transp. Res. Rec. J. Transp. Res. Board*, no. 610, 2003.

[14] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision Tree Analysis on J48 Algorithm for Data Mining," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 6, pp. 2277–128, 2013.

[15] R. Bouckaert, "Bayesian network classifiers in Weka for version 3-5-7," *Dep. pf Comput. Sci. Univ. Waikato, New Zeal.*, p. 47, 2008.

[16] L. a. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[17] J. S. R. Jang, *MATLAB: Fuzzy logic toolbox user's guide: Version 1*. Math Works, 1997.

[18] A. Villagra, D. Pandolfi, M. G. Lasso, M. E. de San Pedro, and G. Leguizamón, "Algoritmos evolutivos y su aplicabilidad en la tarea de clasificación," in *VIII Workshop de Investigadores en Ciencias de la Computación*, 2006.

[19] J. Alcalá-Fdez, M. J. del Jesus, J. M. Garrell, F. Herrera, C. Herbás, and L. Sánchez, "Proyecto KEEL: Desarrollo de una herramienta para el análisis e implementación de algoritmos de extracción de conocimiento evolutivos," *Tendencias la Minería Datos en España, Red Española Minería Datos y Aprendiz.*, pp. 413–424, 2004.

[20] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[21] D. R. Carvalho and A. A. Freitas, "A hybrid decision tree/genetic algorithm method for data mining," *Inf. Sci. (Nj)*, vol. 163, no. 1, pp. 13–35, 2004.

[22] G. F. Miller, P. M. Todd, and S. U. Hegde, "Designing Neural Networks using Genetic Algorithms," in *Proceedings of the Third International Conference on Genetic Algorithms*, 1989, pp. 379–384.