

CLASIFICACIÓN SVM-DIFUSO

Norma Verónica Ramírez Pérez

Tecnológico Nacional de México en Celaya

norma.ramirez@itcelaya.edu.mx

Martín Laguna Estrada

Tecnológico Nacional de México en Celaya

martin.laguna@itcelaya.edu.mx

Norma Natalia Rubín Ramírez

Tecnológico Nacional de México/Instituto Tecnológico de Tepic

nrubin@ittepic.edu.mx

Resumen

En la actualidad las grandes cantidades de información generadas a través de sistemas informáticos puede representar un gran conocimiento, este alto crecimiento de información ha generado la creación de nuevos métodos que ayuden a manipular ese gran volumen de datos recolectados por las organizaciones dedicadas a la medicina, mercadotecnia y a la automatización de procesos entre otras. Si bien esta información cuando es obtenida no tiene sentido por no estar clasificada para la toma de decisiones, es importante analizarla para hacer un buen estudio de esta información. En este artículo se presenta la implementación del clasificador de Máquina de Soporte Vectorial (Support Vector Machine SVM) con salida difusa, lo cual permitirá analizar información contenida dentro del intervalo $[0,1]$, a diferencia de un SVM básico que toma un conjunto de datos de entrada y salida de 1 y -1. Por la complejidad de datos con incertidumbre, no se puede utilizar un clasificador común para hacer una predicción mucho más exacta y ayudar a dar un diagnóstico más real. Un modelo de SVM difuso, le dará un plus y un grado de verdad a los resultados al realizar predicciones mucho más exactas y la posibilidad de poder tomar decisiones con mayor probabilidad de

éxito. Para las pruebas de funcionalidad de este método, se utilizó una base de datos de un grupo de personas con dos tipos de enfermedades, una de vías urinarias de Inflamación de la vejiga urinaria y la otra de Nefritis de origen pelvis renal.

Palabras Clave: Lógica Difusa, Máquina de Vector Soporte, SVM, SVM Difuso, Vejiga urinaria, Nefritis.

Abstract

In the current, large amounts of information generated through computer systems can represent a great deal of knowledge, this high growth of information has generated methods that they help manipulate that large volume of data, collected by organizations every day in areas such as medicine, marketing, process automation, etc. If this information when it is obtained does not make sense because it is not classified for decisions taking. It is important to analyze to make a good study of this information. This paper present the implementation of classifier support vector machine (Support Vector Machine SVM) with fuzzy output, This classifier to allows the analyzed information contained inside of interval $[0, 1]$, to a basic SVM that takes in the input data set and give the output of 1 and -1. Due to the complexity of the uncertainty data set, a common classifier cannot be used to make a very accurate prediction and to help give a more real diagnosis. A Fuzzy SVM model, gives a certain degree of truth the results obtained in more accurate predictions and the possibility of making decisions with more success, for the tests it was used an database the a group of people with two types of diseases in the urinary tract as inflammation of the bladder urinary and nephritis origin pelvis renal.

Keywords: *Fuzzy Logic, Vector Support Machine, SVM, Fuzzy SVM, Urinary Bladder, Nephritis.*

1. Introducción

Las tecnologías de información en la actualidad han provocado un aumento de métodos para clasificar datos, sobre todo los volúmenes tan altos que

acumulamos a través de códigos de barras, tarjetas electrónicas, sensores remotos, transacciones bancarias, bases de datos históricas, o algunos otros como satélites espaciales o el descifrado del código genético humano, etc. Hoy día los humanos se ven en la necesidad de filtrar, analizar e interpretar grandes cantidades de datos que se miden en terabytes [4], lo que hace muchas veces complicada la labor de clasificar datos y llevar a cabo una decisión inteligente. A partir de las dos últimas décadas se vienen desarrollando y utilizando algoritmos complejos para la extracción de patrones útiles en grandes conjuntos de datos, todo esto ha llevado a utilizar términos como el Bigdata, el cual se ha utilizado para hacer referencias a la obtención de grandes cantidades de datos que se generan día a día en negocios o en grandes empresas [4]. Gran parte de esta información generada representa transacciones y situaciones que se han producido a lo largo del tiempo, que son útiles para entender el presente y predecir información a futuro.

El método tradicional de convertir los datos en conocimiento, consiste en un análisis e interpretación de forma básica por especialistas en la materia que se esté estudiando. Esta forma de actuar muchas veces no es tan rápido y preciso como uno quisiera, y en ocasiones puede tener un coste alto por su lento procesamiento además de subjetivo; de hecho, la enorme cantidad de datos desborda la capacidad humana, pues es difícil comprenderlos o analizarlos de una manera manual y muchas veces no se toman decisiones acertadas sino más bien intuitivas.

Los grandes avances en la informática han motivado a los investigadores a desarrollar algoritmos que ayuden a automatizar esta información a través de métodos de aprendizaje [3] y lleven a cabo tareas de clasificación de datos, que por sus grandes almacenamientos, puedan servir para hacer predicciones acertadas. Los métodos de SVM son clasificadores derivados de la teoría de aprendizaje estadístico postulado por Vladimir [9]. El SVM básico toma un conjunto de datos de entrada y predice, para cada entrada dada, a cuál de las dos clases de salida pertenece, por lo que es un clasificador no-probabilístico lineal binario (solo escoge entre dos opciones). SVM se basa en que cada nuevo dato

puede ser clasificado dentro de la categoría que corresponde basado en el aprendizaje de los datos analizados.

Aunque originariamente los SVM fueron pensados para resolver problemas de clasificación binaria, actualmente se utilizan para resolver problemas de regresión, agrupamiento y multclasificación [3]. También son diversos los campos en los que han sido utilizados con éxito, tales como visión artificial, reconocimiento de caracteres, categorización de texto e hipertexto, clasificación de proteínas, procesamiento de lenguaje natural, análisis de series temporales. De hecho, desde su introducción, han ido ganando un merecido reconocimiento gracias a sus sólidos fundamentos teóricos [3].

Dentro de la tarea de clasificación, los SVM pertenecen a la categoría de los clasificadores lineales, puesto que inducen separadores lineales o hiperplanos, ya sea en el espacio original de los ejemplos de entrada, si éstos son separables o cuasi-separables (ruido), o en un espacio transformado (espacio de características), si los ejemplos no son separables linealmente en el espacio original.

Actualmente existen técnicas que permiten analizar patrones de conducta, muchas de ellas comerciales, y algunas otras son tipos de información no trivial mediante la utilización de sofisticados modelos que combinan métodos estadísticos, aprendizaje de máquinas y optimización. Estas técnicas se engloban bajo el concepto de minería de datos (data mining) [4]. La investigación en estos modelos ha sido un tema relevante en estas últimas dos décadas, habiéndose logrado avances significativos en términos de eficiencia y desempeño predictivo. En esta misma línea, el artículo presenta algoritmos que combinan la clasificación mediante SVM y la lógica difusa [6], mediante una prueba con una base de datos de un grupo de personas con dos tipos de enfermedades en las vías urinarias como inflamación de la vejiga urinaria o Nefritis de origen pelvis renal.

Lógica Difusa

En este apartado se realiza una breve descripción de algunos conceptos claves que ayudarán al lector a conocer sobre las técnicas de aprendizaje automático a

través de SVM difuso. El primer concepto a entender son los tipos de sistemas de lógica difusa que existen [6], menciona que éstos tienen una estrecha relación con los conceptos difusos tales como conjuntos difusos y variables lingüísticas. Los sistemas de lógica difusa más populares son los sistemas difusos tipo Mamdani (con fuzzificador y defuzzificador) [6], y los tipo Takagi-Sugeno [6]. Los de tipo Mamdani se basan en realizar reglas del tipo IF-THEN que se ilustra en la figura 1, y que permite ser de mucha ayuda para el conocimiento lingüístico que auxilia al lector a hacer una interpretación más natural pues tiene una aproximación al lenguaje humano.

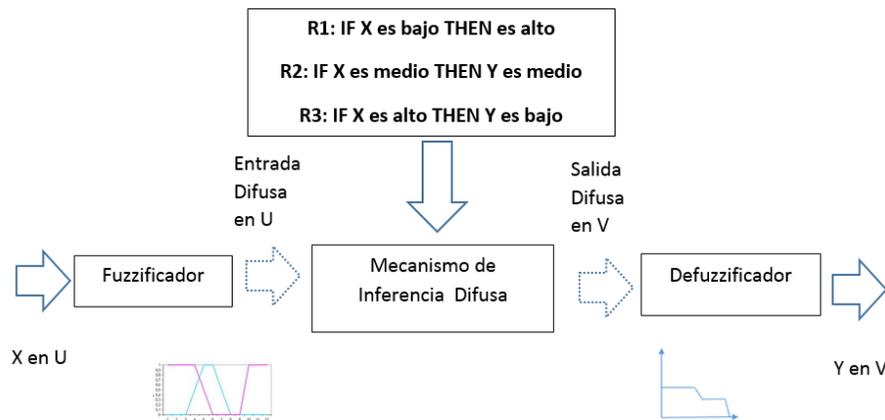


Figura 1 Estructura de un sistema difuso tipo Mamdani.

Las reglas se basan en dos partes, el antecedente y el consecuente, esto quiere decir que si el antecedente se cumple, la conclusión es verdadera. Hay que tomar en consideración que tanto el antecedente como el consecuente de las reglas están dados por expresiones lingüísticas. Como se mencionó anteriormente, los sistemas difusos tipo Mamdani utilizan un fuzzificador donde se ingresa un valor numérico que puede ser de un sensor y para que este valor pueda ser procesado por el sistema difuso, es necesario convertirlo a un “lenguaje” que el mecanismo de inferencia pueda procesar [6].

La función del fuzzificador, es tomar los valores numéricos provenientes del exterior y convertirlos en valores “difusos” que puedan ser procesados por el mecanismo de inferencia. Estos valores difusos son los niveles de pertenencia de

los valores de entrada a los diferentes conjuntos difusos en los cuales se ha dividido el universo de discurso de las diferentes variables de entrada al sistema; por otra parte, el defuzzificador se encarga de convertir la salida difusa del mecanismos de inferencia para que pueda ser interpretada por un elemento externo (por ejemplo un controlador) [6].

La salida del mecanismo de inferencia es un conjunto difuso resultante que genera la salida numérica a partir de este conjunto para lo cual existen varias opciones como el Centro de Gravedad (ecuación 1), y los Centros Promediados (ecuación 2). En la figura 1 se muestra la configuración básica de un sistema tipo Mamdani. Otro elemento que lo conforma son los mecanismos de inferencia difusa, siendo su tarea el tomar los niveles de pertenencia y apoyarse en la base de reglas para generar la salida del sistema difuso.

$$y = \frac{\sum_i b_i \int \mu(i)}{\sum_i \int \mu(i)} \quad (1)$$

$$y = \frac{\sum_i b_i \mu_{premisas}(i)}{\sum_i \mu_{premisas}(i)} \quad (2)$$

Donde y es la salida del sistema difuso, resultante la defuzzificación de los conjuntos difusos, de acuerdo a su nivel de pertenencia [7].

Máquinas de Soporte Vectorial (SVM)

El uso de Máquinas de Soporte Vectorial o SVM es un método de aprendizaje supervisado usado básicamente para clasificación. El método fue presentado por Schölkopf [8], su teoría está basada en la idea de minimizar el riesgo estructural conocido como (SRM), aunque en la mayoría de las aplicaciones el método ha sido introducido como herramienta para resolver problemas de clasificación. Una SVM [1], lo que hace es mapear los puntos de entrada a un espacio de características de una dimensión mayor, es decir, si los puntos de entrada están en R^2 , entonces son mapeados por la SVM a R^3 a través de un hiperplano que los separe y maximizar el margen (m) de la distancia entre las clases, como se

muestra en la figura 2, en donde la frontera de decisión deberá estar lejos de los datos de ambas clases.

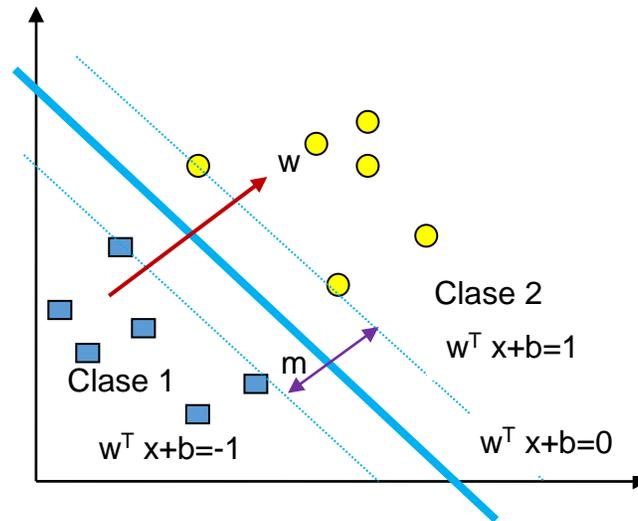


Figura 2 Representación gráfica del método de clasificación SVM.

Cuando se decide maximizar el margen m , se considera como un problema de programación cuadrática (QP) y puede ser resuelto por su problema dual introduciendo multiplicadores de Lagrange. Sin ningún conocimiento del mapeo, la SVM encuentra el hiperplano óptimo utilizando el producto escalar con funciones en el espacio de características que son llamadas kernels, y con lo anteriormente expuesto, es posible encontrar la solución del hiperplano óptimo a través de la combinación de algunos puntos de entrada llamados “vectores soporte” [2].

De acuerdo a recientes publicaciones en diferentes revistas y foros sobre este tema, se ha podido observar que el método SVM ha tenido excelente aceptación en aplicaciones como reconocimiento de patrones, de rostros, de caracteres, en la clasificación de documentos y en aplicaciones físico- matemáticas.

Caso linealmente separable

Sea un sistema coordenado con un conjunto X de puntos dados etiquetados para entrenamiento según se muestra en ecuación 3 y figura 3.

$$(x_1, y_1), \dots, \dots, (x_i, y_i) \quad (3)$$

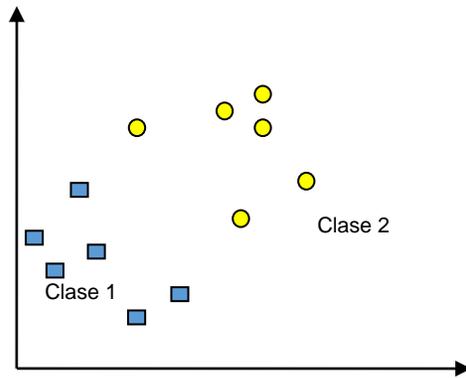


Figura 3 Caso linealmente separable.

Cada punto entrenamiento $x_i \in R^n$ pertenece a alguna de las dos clases y se le ha dado una etiqueta $y_i \in \{-1,1\}$ para $i= 1, \dots, n$. En casi en todos los casos, encontrar un hiperplano adecuado del espacio de entrada, es poco probable y práctico, por lo que se tienen que buscar alternativas para solucionarlo; una de ellas es mapear el espacio de entrada a un espacio de características de una dimensión mayor y buscar el hiperplano óptimo. Sea $z = \varphi(x)$ la notación del correspondiente vector en el espacio de características con un mapeo a un espacio de características Z . Para encontrar el hiperplano se aplica ecuación 4.

$$w \cdot z + b = 0 \quad (4)$$

Al definirlo por (w, b) , es preciso que podamos separar el punto x_i de acuerdo la función dada por ecuación 5.

$$f(x) = \text{sing}(w \cdot z_i + b) \begin{cases} 1 & y_i = 1 \\ -1 & y_i = -1 \end{cases} \quad (5)$$

Donde $w \in Z$ y $b \in R$, el conjunto X se dice que es linealmente separable si existe (w, b) , tal que las inecuaciones sean válidas para todos los elementos del conjunto X para el caso linealmente separable de X , ecuación 6.

$$\begin{cases} (w \cdot z_i + b) \geq 1, & y_i = 1 \\ (w \cdot z_i + b) \leq -1, & y_i = -1 \end{cases} \quad i = 1, \dots, l \quad (6)$$

De esta forma se encuentra un único hiperplano óptimo, en donde el margen entre dos proyecciones de los puntos de entrenamiento de las dos clases se maximiza.

Caso no linealmente separable

Cuando se quiere tratar datos que no son linealmente separables, un análisis previo puede ser generalizado al introducir algunas variables positivas $\xi_i > 0$ de tal manera que la ecuación 6 es modificada a ecuación 7.

$$y_i(w \cdot z_i + b) \geq 1 - \alpha_i, \quad i = 1, \dots, n \quad (7)$$

Los $\xi_i = 0$ de acuerdo a la ecuación 7, cuando el punto x_i no satisface la ecuación 4, entonces la $\sum_{i=1}^l \xi_i$ puede ser tomada como algún tipo de medida del error en la clasificación (figura 4).

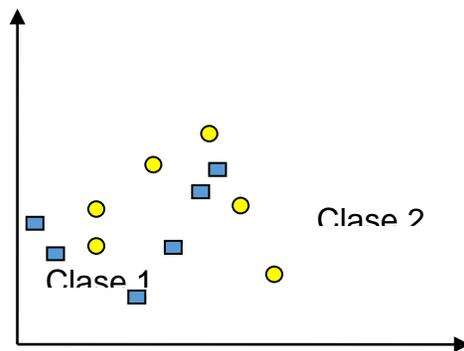


Figura 4 Caso linealmente no separable.

El hiperplano óptimo es redefinido como la solución al problema del caso lineal no separable mostrado en la figura 4 y ecuación 8.

$$\min \left\{ \frac{1}{2} w \cdot w + C \sum_{i=1}^l \alpha_i \right\} \quad (8)$$

Sujeto a $y_i(w \cdot z_i + b) \geq 1 - \alpha_i, \quad i = 1, \dots, l$
 $\alpha_i \geq 0, \quad i = 1, \dots, l$

En la ecuación 8 se considera C como una constante que puede ser definida para la regularización, es libre y puede ser ajustada en el método SVM para hacer un balance entre la maximización del margen y violación a la clasificación [2].

Al buscar el hiperplano óptimo, en la ecuación 8, se obtiene un problema QP que puede ser resuelto con una ecuación de LaGrange y transformado en el dual, ecuación 9.

$$\max W(\alpha) = \sum_{i=1}^l \alpha_i = \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j z_i \cdot z_j \quad (9)$$

Sujeto a: $\sum_{i=1}^l y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1, \dots, l$

Donde $\alpha = (\alpha_1, \dots, \alpha_l)$ es un vector de multiplicadores de Lagrange positivos asociados con las constantes en la ecuación 6.

Por otro lado, el teorema de Khun-Tucker es importante en la SVM [2]. De acuerdo a este teorema, la solución de $\bar{\alpha}_i$ de la ecuación 9 satisface ecuaciones 10 y 11.

$$\bar{\alpha}_i (y_i (\bar{w} \cdot z_i + \bar{b}) - 1 + \alpha_i) = 0 \quad i = 1, \dots, l \quad (10)$$

$$(C - \bar{\alpha}_i) \alpha_i = 0 \quad i = 1, \dots, l \quad (11)$$

De esta igualdad se deduce que los únicos valores $\bar{\alpha}_i \neq 0$ de la ecuación 11, son aquellos en que las constantes en la ecuación 5 son satisfechas en el signo de la igualdad.

El punto x_i correspondiente con $\bar{\alpha}_i > 0$, es llamado vector soporte, sólo que hay dos tipos de vectores soporte en un caso no separable. El primer tipo se presenta en el caso $0 < \bar{\alpha}_i < C$, en donde el correspondiente vector soporte x_i satisface las igualdades $(\bar{w} \cdot z_i + \bar{b}) = 1$ y $\alpha_i = 0$. El segundo tipo se presenta en el caso en que $\bar{\alpha}_i = C$. Con el correspondiente $\alpha_i = 0$ diferente de cero, el correspondiente vector soporte x_i no satisfará la ecuación 6, y nos referiremos a estos vectores soporte como errores. El punto x_i como se aprecia en la figura 5, correspondiente con $\bar{\alpha}_i = 0$ es clasificado correctamente y está definitivamente alejado del margen de la decisión.

Para construir el hiperplano óptimo $\bar{w} \cdot z + b$, se utiliza ecuación 12.

$$\bar{w} = \sum_{i=1}^l y_i \alpha_i z_i \quad (12)$$

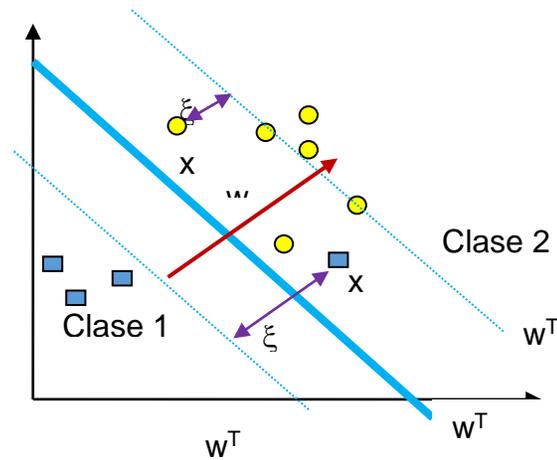


Figura 5 Caso no separable.

Y el escalar b puede ser determinado de las condiciones de la ecuación de Khun y Tucker según se muestra en la ecuación 11.

La función de decisión generalizada determinada por la ecuaciones 5 y 12 es tal que obtiene ecuación 13.

$$f(x) = \text{sign}(w \cdot z + b) \text{sign}\left(\sum_{i=1}^l y_i \alpha_i z_i \cdot z + b\right) \quad (13)$$

Kernel para el caso no linealmente separable

Como no se tiene conocimiento de φ , el cálculo de las ecuaciones 9 y 13 no se pueden llevar a cabo. Una buena propiedad de las SVM es que no es necesario tener ningún conocimiento acerca de φ cómo se aprecia en la figura 6. Para saberlo solo se necesita una función $K(\cdot, \cdot)$ llamada kernel que calcule el producto punto de los puntos de entrada en el espacio de características Z , como se describe mediante ecuación 14.

$$z_i \cdot z_j = \varphi(x_i) \cdot \varphi(x_j) = k(x_i, x_j) \quad (14)$$

En esto, las funciones que satisfagan el teorema de Mercer, pueden ser usadas como productos escalares y pueden ser usadas como kernels y podemos utilizar el kernel polinomial de grado d , ecuación 15.

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^d \quad (15)$$

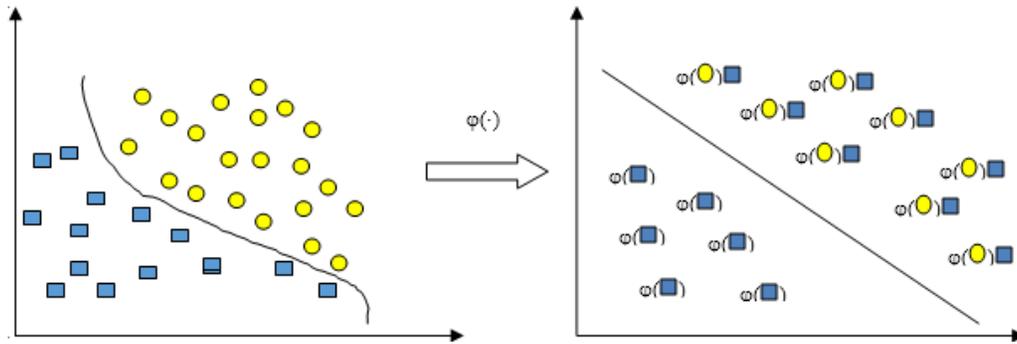


Figura 6 Kernel para la transformación del espacio de datos.

Clasificador SVM

El hiperplano no lineal de separación puede ser encontrado como la solución, ecuación 16.

$$\text{Max } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (16)$$

sujeto a $\sum_{i=1}^l y_i \alpha_i = 0, \quad 0 < \bar{\alpha}_i < C, \quad i = 1, \dots, l$

Y la función de decisión dada por ecuación 17.

$$f(x) = \text{sign}(w \cdot z + b) \text{sign}\left(\sum_{i=1}^l \alpha_i y_i K(x_i, x_j) + b\right) \quad (17)$$

Estos algoritmos computacionales son sólo algunos de la colección de los algoritmos de Inteligencia Artificial, que pueden ser enfocados para la clasificación y el reconocimiento de patrones [2].

2. Métodos

Para el desarrollo de este trabajo, se realizaron las pruebas con una base de datos obtenida de encuestas a personas que padecen enfermedades de vías urinarias en el estado de Guanajuato, y bajo la observación de expertos en la materia como médicos urólogos e internistas. Dicha base de datos cuenta con 685 instancias, con 6 atributos que se ilustran en la tabla 1, quienes a través de esta información determinaron 2 tipos de diagnóstico: Inflamación de la vejiga urinaria o Nefritis de origen pelvis renal.

Tabla 1 Información relacionada con la base de datos.

T	Temperatura del paciente {35C-42C}
N	Náuseas [poca, moderada, mucha]
D	Dolor lumbar [poca, moderada, mucha]
M	Molestia al orinar (necesidad continua de orinar) [poca, moderada, mucha]
DI	Disuria [poca, moderada, mucha]
A	Ardor de la uretra, picazón, hinchazón de la salida de la uretra [poca,
DIAG1	Decisión: Inflamación de la vejiga urinaria
DIAG2	Decisión: Nefritis de origen de la pelvis renal origen

Para trabajar con la base de datos, se utilizó el modelo SVM-DIFUSO que se ilustra en la figura 7, y que describe paso a paso el proceso de analizar los datos.

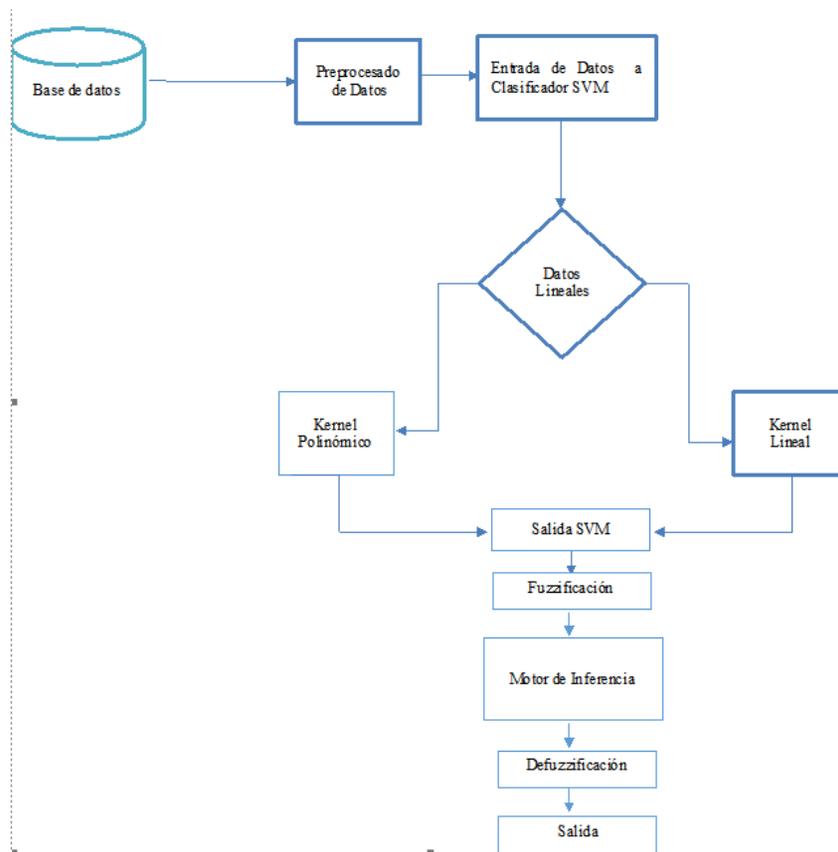


Figura 7 Diagrama del Modelo SVM-DIFUSO.

Para el análisis de datos se empieza desde el preprocesado de los mismos y hacer los ajustes necesarios, esto con relación a la entrada de SVM, ya que al utilizar aprendizaje supervisado [5], necesitamos que su salida sea de 1 y -1. A

continuación se tuvo que determinar qué tipo de kernel utilizar, pues muchas veces los datos no son linealmente separables, por lo que se debe de probar cuál de los kernels es el más apropiado para procesar los datos. El kernel escogido a emplear fue el RBF de grado 1. Se utilizó el 70% de los datos para el entrenamiento y el 30% para la validación. Para la clasificación se utilizó el aprendizaje supervisado y se introdujeron los datos objetivo, es decir, las etiquetas para las instancias que nos permitan determinar cuáles son los vectores soporte para ser introducidos a los conjuntos difusos.

Se da entrada a la fuzzificación de los datos para realizar los conjuntos difusos de cada uno de los atributos para determinar sus etiquetas lingüísticas como se muestra en la figura 8.

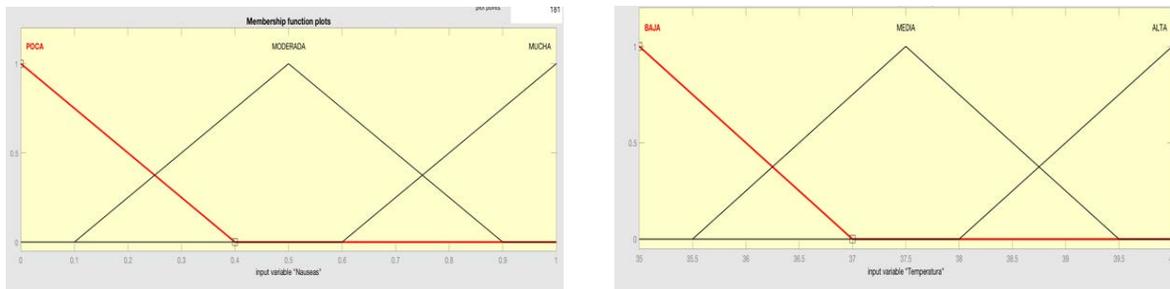


Figura 8 Conjuntos difusos.

Una vez ya definidos los conjuntos, que por espacio solo se ilustran algunos de ellos, se determinó el conjunto de salida, que cuenta con los dos diagnósticos, Inflamación de la vejiga urinaria y la de Nefritis de origen pelvis renal para determinar las reglas de producción ilustrados en la figura 9.

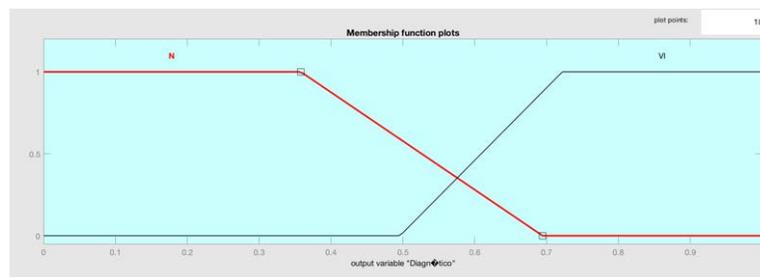


Figura 9 Conjunto de salida.

Ya una vez determinados los conjuntos tanto de entrada como de salida, se utilizaron las reglas de producción IF-ELSE-THEN como se puede ver en la figura 10.

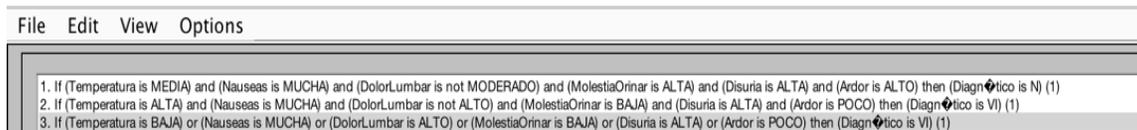


Figura 10 Reglas de Producción IF-ELSE-THEN.

3. Resultados

Durante la fase de experimentación se clasificaron 658 datos con seis atributos y dos diagnósticos fueron obtenidos de la base de datos. Cabe aclarar que en la fase del SVM la clasificación nos llevó a una predicción del 90%, al tener detectados los vectores soporte que de acuerdo al hiperplano, se pueden ver claramente en la figura 11.

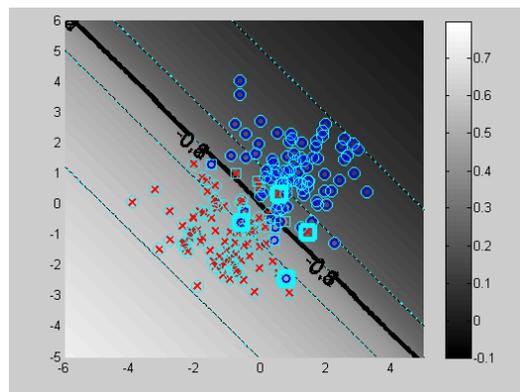


Figura 11 Salida de vectores soporte.

Los resultados del algoritmo SMV arrojaron una clasificación sólo por diagnóstico, lo cual quiere decir de que no sabemos cuál sería el diagnóstico real, o que efectivamente se haga un razonamiento aproximado, por lo que al dar ese tipo de datos, nos da un resultado determinista, y lo que se pretendió es que nos dé un resultado estocástico. Del análisis se obtuvo que el diagnóstico de Nefritis fue del 85.8% del total de los datos de entrada, mientras que la del diagnóstico de

inflamación de la vejiga fue del 85.9 % del total de datos de entrada, dando como un porcentaje total de la variable a clasificar (variable difusa de salida) del 80% de los datos clasificados en alguna de las dos clasificaciones propuestas.

Se puede observar que del experimento realizado primeramente mediante un SVM básico, no fue posible reportar este alto grado de clasificación como el realizado utilizando un SVM con lógica difusa. Por otra parte, los resultados aquí mostrados son alentadores, ya que el objetivo de crear un SVM difuso fue la de mejorar los resultados de clasificación para combatir la incertidumbre y dar una respuesta mucho más exacta, de acuerdo a los conceptos de la lógica difusa.

Como una demostración de los resultados obtenidos, en la figura 12 se ilustra solo algunos de ellos, además el toolbox permite construir una gráfica que muestra en que regiones se encuentra un mayor número de resultados obtenidos, todo esto para ayudar hacer más visible al investigador el análisis de los resultados.

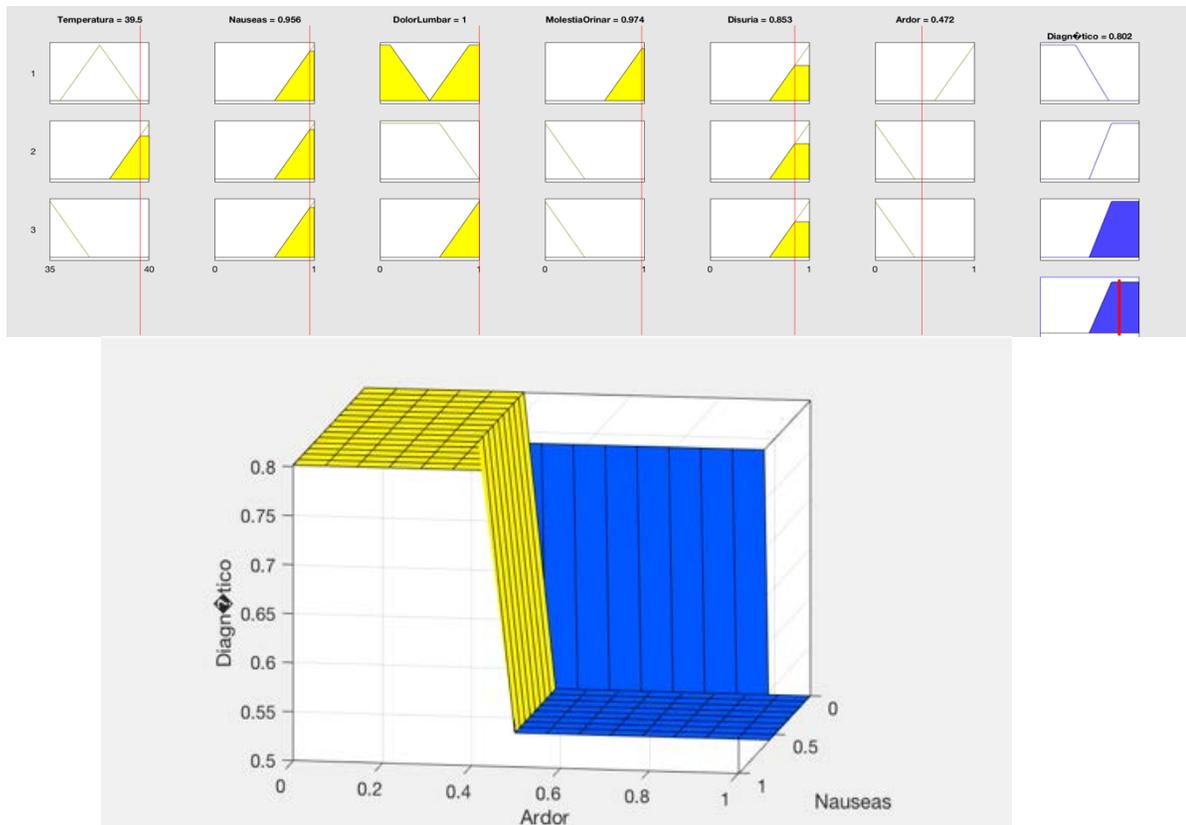


Figura 12 Resumen de resultados.

4. Discusión

El trabajo realizado representa una primera etapa del desarrollo de un SVM difuso genérico, que puede ser aplicado en cualquier ámbito para un proceso mucho más rápido y con un alto grado de certeza en la clasificación de cualquier tipo de información. Durante esta primera etapa se demuestra que un SVM puede ser muy útil en el proceso de clasificación cuando la complejidad de la información no representa problemas, es decir que se puede homogenizar en un intervalo en que los valores determinen un falso o un verdadero, pero cuando el grado de dificultad que presenta la información se incrementa por contener valores que pueden generar incertidumbre, es decir que no se tenga la certeza de ser verdadero o falso, lo mejor es utilizar herramientas que nos arrojen información estocástica y no determinista.

De los resultados obtenidos en este trabajo se puede evidenciar como los algoritmos de clasificación ayudan a trabajar con grandes bases de datos generadas por cualquier disciplina y que hoy en día, muchos de los algoritmos se han vuelto obsoletos y obligan a la generación de nuevas herramientas que ayuden a procesar mucho más rápido y emitir un verdadero diagnóstico sin ambigüedades y con un alto razonamiento aproximado a la realidad, que es el caso de la lógica difusa aquí presentado.

5. Bibliografía y Referencias

- [1.] C.J.C. Burges (1998). A tutorial on support vector machine for pattern recognition. Data.
- [2.] Carmona, E. (2014). "Tutorial sobre Máquinas de Vectores Soporte (SVM)" en UNED, Consultada en [http://www.ia.uned.es/~ejcarmona/publicaciones/\[2013-Carmona\]%20SVM.pdf](http://www.ia.uned.es/~ejcarmona/publicaciones/[2013-Carmona]%20SVM.pdf) (fecha de consulta 01-07-2017).
- [3.] Carbonell, J., Michalski, R. Mitchell, T., (1983) Machine Learning an Artificial Intelligence Approach, an overview of machine learning. Chapter 1
- [4.] Hall, M., Witten I. y Frank E. (2011). Data Mining: Practical Machine Learning Tools and Techniques. 3a ed. Massachusetts: Morgan Kaufmann. mining and knowledge discovery, 2(2):121_167.

- [5.] Maldonado, S. (2012). “Modelos de selección de atributos para Support Vector Machines”, <http://www.dii.uchile.cl/~ris/RISXXVI/maldonado1.pdf> (fecha de consulta 05-07-2017).
- [6.] Mendel, J., (1995) Fuzzy Logic Systems for Engineering: A tutorial, Proceedings of IEEE, Vol. 83, No 3 marzo,
- [7.] Mendel, J., (2001). Uncertain Rule based Fuzzy Logic System, Prentice Hall, New York, 2001.
- [8.] Scholkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., And William Son, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Compute.* 13, 7, 1443–1471.
- [9.] Vladimir Vapnik y Chervonenkis (2011). *The Nature of Statistical Learning Theory*, Springer.