

EVALUACIÓN COMPARATIVA DE TRES MÉTODOS DE CLASIFICACIÓN APLICADOS AL PROBLEMA DE LA LOCALIZACIÓN DE FALLAS DE CORTOCIRCUITO EN SISTEMAS DE DISTRIBUCIÓN DE ENERGÍA ELÉCTRICA

Comparative evaluation of three classification techniques applied to fault location problem in power distribution systems

RESUMEN

En este artículo se analiza y compara el desempeño de las Máquinas de Soporte Vectorial (SVM), el Algoritmo de Aprendizaje para Análisis de Datos Multivariable (LAMDA) y las Mezclas Finitas (MF), técnicas del campo de la computación suave, aplicadas al problema de la localización de fallas en sistemas de distribución.

Metodológicamente, el sistema de distribución es subdividido en varias zonas, y mediante la aplicación de cada una de las técnicas propuestas, se identifica la zona en falla. Como resultado, se obtienen los índices de precisión para cada técnica, alcanzando valores superiores a 90% en el caso de las SVM.

PALABRAS CLAVES: Localización de fallas, Sistemas de distribución, Máquinas de Soporte Vectorial, Algoritmo de Aprendizaje para Análisis de Datos Multivariable, Mezclas Finitas.

ABSTRACT

In this paper, the performance of Support Vector Machines (SVM), Learning Algorithm for Multivariable Data Analysis (LAMDA), and Finite Mixtures (MF), three classification techniques from the soft computing field tested and compared in an application on power distribution systems. In this case the proposed techniques are used to solve the fault location problem.

The power distribution system is then subdivided in several zones and by applying each technique one of these is identified as the faulted zone. As results, values of the precision indexes higher than 90% are obtained in case of the application of SVM.

KEYWORDS: *Fault location, power distribution systems, classification techniques, Support Vector Machines (SVM), Learning Algorithm for Multivariable Data Analysis (LAMDA), Finite Mixtures (FM).*

1. INTRODUCCIÓN

En los últimos años, el estudio de la calidad y continuidad del suministro de energía eléctrica ha tomado gran auge. En este marco, el problema de la localización de fallas se posiciona como uno de los más importantes, dado que las interrupciones significan pérdidas de dinero básicamente por dos razones: a) El tiempo de no operación o suspensión de la producción del cliente, y b) Las multas impuestas a las empresas de energía eléctrica por incumplimiento a las regulaciones establecidas [1].

Lo anterior incrementa la importancia que tiene el hecho de localizar de manera confiable y rápida el lugar donde ha ocurrido la falla. Esto permite establecer estrategias que solucionen o mitiguen su impacto en la calidad y continuidad del servicio, para cumplir con las normas especificadas por las comisiones de regulación [2],[3].

Los métodos más eficaces hasta ahora existentes para localización de fallas de cortocircuito en sistemas de

distribución son costosos en términos económicos, y además tienen falencias como la estimación de múltiples sitios de falla y la necesidad de un buen modelo del sistema de distribución [4], [5], [6].

En este documento se propone una metodología económica en términos de la inversión y de fácil implementación que evalúa tres métodos de aprendizaje que tienen enfoques distintos, las Máquinas de Soporte Vectorial (SVM) [7], el Algoritmo de Aprendizaje para Datos Multivariables (LAMDA) [8] y las Mezclas Finitas (MF) [9]. Las SVM trabajan con teoría de generalización y técnicas de optimización, LAMDA es una metodología de base radial, que utiliza el concepto de clases prototipo. Cada conjunto de descriptores de faltas se compara con el prototipo a partir de la medición de la distancia, para determinar su clase. Finalmente, MF es una herramienta con enfoque estadístico. Los tres métodos tienen en su fase de aprendizaje una estrategia de entrenamiento supervisado, que permite evaluarlos en igualdad de

GERMÁN ANDRÉS MORALES ESPAÑA *

Ingeniero Electricista, M.Sc (c)
Universidad Industrial de Santander
German.morales.e@gmail.com

JUAN JOSÉ MORA FLÓREZ **

Ingeniero Electricista, Ph.D.
Profesor asistente
Universidad Tecnológica de Pereira
jjmora@utp.edu.co

GILBERTO CARRILLO CAICEDO *

Ingeniero Electricista, Ph.D.
Profesor titular laureado
Universidad Industrial de Santander
gilberto@uis.edu.co

* **Grupo de Investigación en Sistemas de Energía Eléctrica - GISEL**

** **Grupo de Investigación en Calidad de Energía Eléctrica y Estabilidad - ICE³**
Grupo de Planeamiento de Sistemas Eléctricos – GPE

condiciones. La comparación se realiza con el fin de tener un criterio de selección de la herramienta clasificadora que mejor se comporta ante el problema de localización de fallas en sistemas de distribución.

Como contenido de este artículo, en la sección dos se presentan los fundamentos básicos de cada una de las técnicas de aprendizaje utilizadas. En la sección tres se muestra la manera de adquirir los datos necesarios para entrenar una herramienta clasificadora a partir de registros de tensión y de corriente, tomados cuando el sistema de distribución se encuentra en estado de prefalla y falla. En la sección cuatro se presenta de forma detallada la metodología a seguir para la localización de la zona de falla. En la sección cinco se muestran resultados comparativos de las pruebas de las SVM, LAMDA y las MF para la localización de fallas en un sistema prototipo. Finalmente en la sección seis se presentan las conclusiones más importantes de la investigación.

2. FUNDAMENTOS BÁSICOS DE LOS MÉTODOS DE APRENDIZAJE

Los métodos de aprendizaje, propios del campo de la inteligencia artificial, han servido como modelos de caja negra para resolver muchos problemas de aplicación práctica [2]. En esta sección se presentan brevemente los aspectos fundamentales de cada uno de los métodos usados para la localización de fallas en sistemas de distribución de energía eléctrica.

2.1. Máquinas de Soporte Vectorial

Los clasificadores de soporte vectorial están basados en hiperplanos que separan los datos de entrenamiento en dos subgrupos que poseen cada uno una etiqueta propia. En medio de todos los posibles planos de separación entre las dos clases, existe un único hiperplano de separación óptimo (OSH), de forma que la distancia entre el hiperplano óptimo y el patrón de entrenamiento más cercano sea máxima, con la intención de forzar la generalización de la máquina de aprendizaje, como se muestra en la Figura 1 [7] [10].

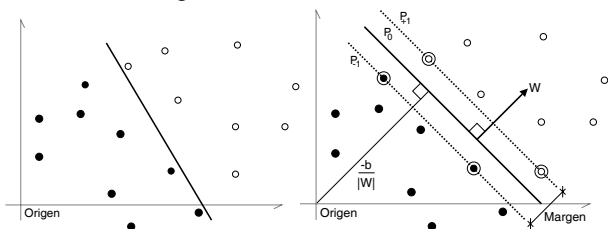


Figura 1. Hiperplanos que separan correctamente los datos. El OSH de la derecha tiene un mayor margen de separación entre clases, por lo tanto se espera una mejor generalización.

Si existen datos erróneos, ruido o alto solapamiento de clases en los datos de entrenamiento, puede afectar el hiperplano clasificador óptimo. Por esta razón se cambia un poco la perspectiva y se busca el mejor hiperplano clasificador que pueda tolerar ruido en los datos de entrenamiento introduciendo una variable de relajación.

La generalización de la SVM a funciones de decisión no lineales consiste en mapear el espacio de entrada sobre un espacio de representación de dimensión alta a través de una función no lineal elegida a priori [10]. Por medio de una función se trazan los datos de entrada a algún espacio de mayor dimensión.

Para resolver el problema de más de dos clases, se construye una función clasificadora global a partir de un conjunto de funciones biclasificadoras. Existen técnicas de descomposición y reconstrucción que permiten a las SVM biclasificadoras manejar problemas de multclasificación con mayor simplicidad y/o menor tiempo de respuesta que una SVM generalizada para multclasificación [10]. Los resultados aquí presentados son obtenidos con el método de descomposición uno contra uno (1-v-1), por sus excelentes resultados, robustez y menor exigencia computacional [2].

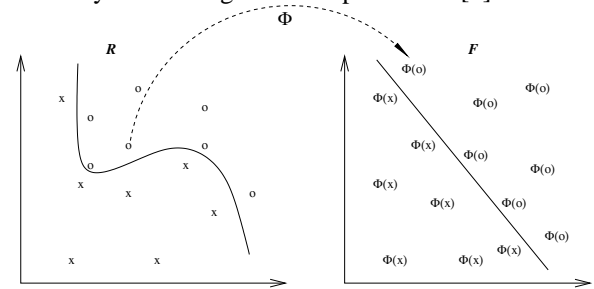


Figura 2. La SVM no lineal mapea el espacio de entrada en otro de representación de dimensión alta y luego construye un OSH sobre este último, cuya representación en el espacio de entrada es una función de separación no lineal.

2.2. Algoritmo de aprendizaje para análisis de datos multivariable

El método LAMDA (Learning Algorithm for Multivariate Data Analysis), o Algoritmo de aprendizaje para el análisis de datos multivariables, es una estrategia de clasificación desarrollada por N. Piera [11] basada en una idea original de J. Aguilar [8]. La formación y el reconocimiento de clases en este método están basados en la atribución de un objeto a una clase a partir de la norma heurística llamada adecuación máxima.

El método LAMDA representa un sistema de clases o de conceptos mediante la conexión lógica de toda la información marginal disponible. De esta manera la adecuación global de un objeto a una clase se calcula a partir de la adecuación marginal de cada atributo, y según una regla heurística de adecuación máxima, el objeto pertenecerá a la clase que presente el mayor grado de adecuación [12].

LAMDA es una metodología de clasificación conceptual, debido a que los objetos no clasificados son confrontados a un prototipo o concepto formado por cada una de las clases existentes. El carácter difuso de los prototipos modela la imprecisión en la formación de conceptos.

La diversidad de problemas abordados con LAMDA es muy amplia e incluye el reconocimiento de formas [8], las aplicaciones médicas y el diagnóstico de procesos industriales [12], entre otros. Esta flexibilidad radica en

varias propiedades del método como la posibilidad de realizar un aprendizaje supervisado y no supervisado; la utilización de atributos cualitativos y/o cuantitativos; y la posibilidad de realizar un aprendizaje secuencial [2].

Una propiedad importante que diferencia a LAMDA consiste en la capacidad de modelar la indistinguibilidad total o la homogeneidad en el universo de descripción, mediante una clase denominada clase no informativa (NIC), que acepta todos los objetos contenidos en el universo de descripción con el mismo grado de adecuación. Según el principio de adecuación máxima, esta clase representa un límite mínimo de asignación de un objeto a una clase. Así, a diferencia de la mayoría de los algoritmos de clasificación, el límite mínimo de aceptación de un objeto a una clase no se fija a priori, sino que se determina por el propio contexto [12].

2.3. Mezclas finitas

Es una técnica de modelado estadístico con la cual se obtiene una estimación de la función densidad de probabilidad de los datos en una muestra aleatoria, la cual se representa como una suma ponderada finita de las componentes de densidad multivariadas [9]. Los datos multivariados son frecuentemente heterogéneos, y el objetivo del análisis de este tipo de información es determinar la homogeneidad de la muestra y observar la existencia de grupos tal como se muestra en la Figura 2.

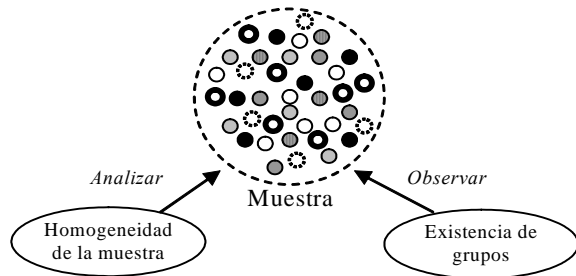


Figura 3. Representación de datos multivariados.

En general, si una población donde se ha definido una variable aleatoria x , puede subdividirse en G grupos más homogéneos asociados a la proporción de elementos y una función de densidad, entonces la función densidad en toda la población está dada por la mezcla de densidades. La muestra para el análisis se representa en la Figura 3.

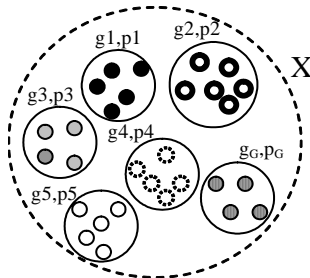


Figura 4. Representación del análisis de la muestra

Finalmente, en la ecuación (1) se muestra la expresión que representa la Función Distribución de Probabilidad (FDP) de la muestra, considerando el vector de parámetros correspondiente a la función de densidad de

cada grupo. La correcta estimación de los parámetros se convierte en un paso muy importante, debido a que involucra la representación de la información almacenada en la muestra, tal como se presenta en (1).

$$f_{MF}(\bar{x}) = \sum_{i=1}^G p_i f_i(\bar{x}, \theta_i) \tag{1}$$

Donde, G es el número de grupos, p el coeficiente de mezclado, x los datos de la muestra, f es la FDP del grupo y θ los parámetros de la FDP.

3. CARACTERIZACIÓN DEL SISTEMA DE DISTRIBUCIÓN EN FALLA

La metodología de caracterización propuesta utiliza los valores eficaces de tensión y corriente de fase y de línea durante la falla, para obtener seis valores eficaces de tensión y seis valores eficaces de corriente.

3.1. Definición de descriptor

Un descriptor es una característica tomada de una señal (en este caso señales de tensión y corriente), y que se caracteriza por su aporte de información sobre algún evento. Los descriptores proporcionan toda la información necesaria para el ajuste de los métodos de clasificación utilizados para la localización de fallas en sistemas de distribución de energía eléctrica. En esta investigación se busca caracterizar una determinada zona de falla, con descriptores obtenidos de las medidas tomadas únicamente en un terminal de línea, que en este caso corresponde a las subestaciones de distribución.

La caracterización de huecos de tensión causados por la falla de cortocircuito en el sistema de distribución, a partir de los valores eficaces, se describe en [2] y [13]. Por ejemplo, a partir del hueco de tensión se adquiere su profundidad o variación en la tensión denotada como (dV), definida como se muestra en la Figura 5.

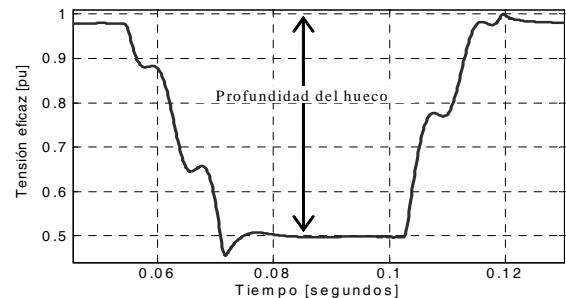


Figura 5. Hueco de tensión

Es importante determinar el estado estable de la falla, ya que la parte transitoria de la señal varía con respecto al ángulo de incidencia de la falla. Por ello los descriptores se deben obtener en la parte estable del hueco de tensión.

Los descriptores de las señales de corriente se obtienen de manera análoga a los de tensión, mencionados anteriormente.

3.2. Descriptores propuestos

Debido a que dos fallas ubicadas en dos lugares distintos dentro del sistema de distribución y vistas desde la

subestación pueden producir huecos similares en las fases falladas, se concluye que la información de las fases falladas no es suficiente para determinar la ubicación del evento. El comportamiento de las fases no falladas muestra que a partir de las tensiones de fase y de línea se obtiene una marcada diferencia entre fallas localizadas en

sitios diferentes del sistema de distribución. A manera de ejemplo de la situación propuesta y para el circuito de la figura 6, se muestra en la figura 7 el comportamiento de las tensiones eficaces ante una falla monofásica en la fase A, ubicada en dos lugares distintos (Nodo 2 y 12 que distan 34,6 km).

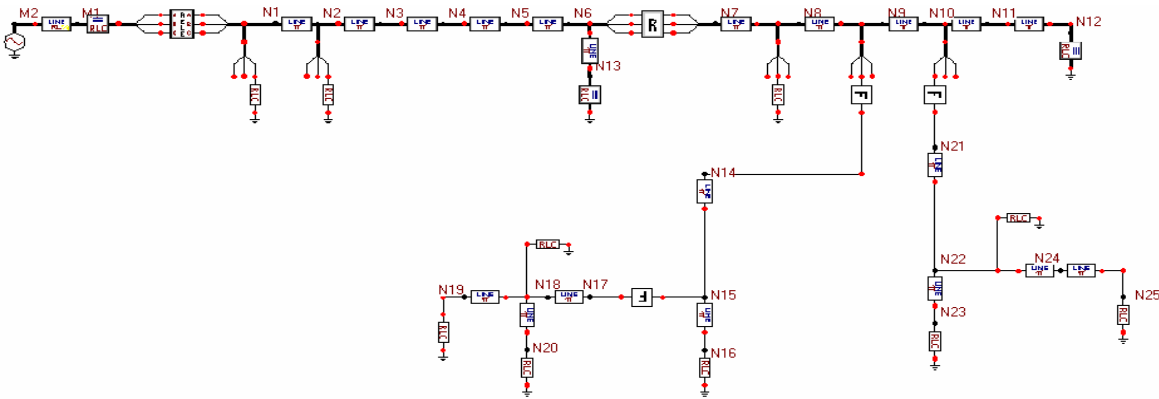


Figura 6. Circuito típico de un sistema de distribución de 25 kV implementado en ATP Draw

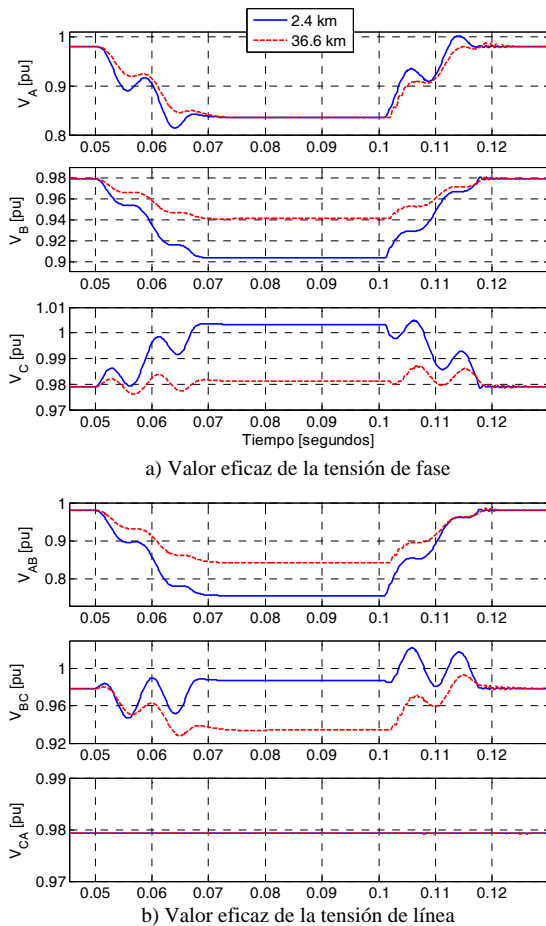


Figura 7: Tensiones de fase (a), y de línea (b), en falla monofásica de la fase A, en el nodo 2 y en el nodo 12 del sistema de la Figura 6

Debido a que el hueco de tensión sólo brinda información de la magnitud de la señal, y con el fin de obtener la mayor información posible a partir de este descriptor, se

utilizan las tensiones de línea y de fase. Del análisis de los tres valores eficaces de las tensiones de fase y de línea, se puede determinar la información de magnitudes, ángulos y acoples mutuos entre las fases.

Los descriptores definidos corresponden a las variaciones en los valores eficaces de la tensión de fase (dV), y de la tensión de línea (dVL), para un total de 6 descriptores de tensión. De manera análoga los descriptores de corriente de línea (dIL), se calculan como la resta de las señales de cada una de las fases (dI). Estos últimos descriptores asociados a la corriente de línea no tienen sentido físico, ya que debido que el generador está en estrella, las corrientes de línea y de fase, son iguales.

4. METODOLOGÍA DE LOCALIZACIÓN DE FALLAS

4.1. Zonificación de la red

Se define como zona a la región donde se desea discriminar la ubicación de la falla. El criterio de zonificación depende del conocimiento de la topología, protecciones, longitudes de las redes, tipo de usuarios del sistema, entre otros.

Se recomienda utilizar una zona por cada circuito ramal, debido a que una de las mayores falencias de los algoritmos de distancia ya existentes es la múltiple estimación [2], [4].

4.2. Registro de datos de falla

Los datos de falla son necesarios para la creación de una base de datos que contenga las señales (tres tensiones y corrientes de fase en el tiempo), vistas en la subestación cuando ha ocurrido una falla en el sistema de distribución. Estos datos se pueden obtener del historial de fallas existente en la subestación o con la simulación

del sistema bajo falla con algún software especializado como el *Alternative Transients Program* ATP [14].

4.3. Adquisición de descriptores

Para la adquisición de descriptores se obtiene el valor eficaz de las tensiones y corrientes de fase y de línea, a partir del filtrado a las señales, para obtener la fundamental. Estos descriptores se agrupan y etiquetan dependiendo de la zona a la cual pertenece la falla.

Adicionalmente, se debe realizar un escalamiento simple a los descriptores (entre [-1,1] o [0,1]). Este paso es muy importante para evitar dificultades numéricas durante los cálculos que hacen los procesos más lentos.

Los descriptores se deben dividir en dos grupos, uno se utilizará en la etapa de entrenamiento y el otro grupo se utilizará para probar la capacidad de aprendizaje del método de clasificación ya entrenado. Tanto en la etapa de entrenamiento como de prueba, se utilizan todas las barras, pero en la selección del conjunto de entrenamiento se utilizan 5 de las 21 resistencias de falla linealmente distribuidas en el rango de análisis.

4.4. Entrenamiento de los métodos

El entrenamiento de cada uno de los métodos propuestos en este artículo (SVM, MF y LAMDA), consiste en presentar una serie de datos (entradas y salidas), para que el método “aprenda” a partir de ejemplos. Como entradas se tienen los conjuntos de descriptores obtenidos de las señales de tensión y de corriente, y cómo salidas se tiene la zona en la cual ocurrió la falla.

Los datos de entrenamiento son diferentes a los datos de prueba y corresponden aproximadamente al 25% de los datos totales de fallas. Para información más detallada acerca de la etapa de entrenamiento ver [2], [13].

4.5. Prueba

Con el fin de obtener la precisión del modelo, se realiza una prueba final con datos que no fueron utilizados en el entrenamiento, y que para este caso corresponden al 75% del total de los datos. A partir de la prueba se obtiene un índice de precisión dado por la ecuación (2).

$$\% \text{ Precisión} = \frac{\text{Datos clasificados correctamente}}{\text{Total de datos}} \times 100\% \quad (2)$$

5. PRUEBAS Y RESULTADOS

Como circuito de prueba, se utiliza el modelo de un sistema real de Saskpower and Ligth de Saskatoon, Canadá, que es un sistema de distribución de 25 kV utilizado en [2] y [5] para pruebas del método de localización de fallas y se presenta en la Figura 6.

5.1. Zonificación del circuito

El circuito se divide en zonas, para ubicar la localización de falla más probable. Básicamente se busca discriminar el ramal donde se encuentra la falla. Considerando este criterio, las zonas que se definen para localizar las fallas en esta prueba son: Zona 1(barras 1, 2, 3, 4, 5, 6, 13),

Zona 2 (barras 8, 9, 10, 11, 12), Zona 3 (barras 15, 16, 18, 19, 20), Zona 4 (barras 22, 23, 24, 25).

La simulación del circuito se realiza usando ATP para diferentes tipos de falla (monofásicas a tierra, bifásicas, bifásicas a tierra, y trifásicas), en todas las barras posibles del circuito de la figura 6 y con 21 resistencias de falla que varían entre 0,05 y 40 ohms.

El desempeño de las herramientas clasificadoras, se mide a partir de sus aciertos definidos como se muestra en la ecuación (2) sobre los datos de prueba. Los datos de prueba no se utilizan en la etapa de entrenamiento y por tanto este porcentaje es una medida del comportamiento de los métodos ante datos desconocidos.

5.2. Estructuras de los clasificadores entrenados

En este artículo se analizan los resultados de dos tipos de clasificadores, uno para todas las fallas y otro sólo para fallas monofásicas (Éste se hace por ser el tipo de falla más común en sistemas de distribución, cerca al 90% [13]). La cantidad de datos de entrenamiento y prueba utilizados se muestran en la Tabla 1.

Tipo de localizador	Entrenamiento		Prueba	
	Registros	%	Registros	%
Todas las fallas	225	23.81	720	76.19
Monofásicas	705	23.81	2256	76.19

Tabla 1. Distribución de datos de entrenamiento y prueba

Para la prueba se utiliza más del 75% del total de datos (se entrenan todos los clasificadores con menos del 25 %, la cantidad restante del total de datos), como se presenta en la Tabla 1. La gran cantidad de datos de prueba en relación a los datos de entrenamiento aumenta la confiabilidad de los porcentajes de exactitud proporcionados en la etapa de prueba.

5.3. Análisis de resultados

En las Tablas 2 y 3 se muestran los resultados con diferentes conjuntos de descriptores, para el análisis del clasificador de todos los tipos de fallas y para el clasificador monofásico, respectivamente.

Combinación de descriptores	Índice de precisión [%]		
	SVM	LAMDA	MF
dV	91,62	66,84	62,41
dVL	92,20	66,93	62,41
dI	99,47	67,77	64,98
dIL	98,80	73,54	63,12
dV_dVL	92,02	66,93	62,41
dV_dI	100,00	68,22	62,06
dV_dIL	100,00	71,50	62,01
dVL_dI	100,00	68,22	61,97
dVL_dIL	99,51	73,40	61,66
dI_dIL	100,00	74,16	64,45
dV_dVL_dI	100,00	68,48	61,70
dV_dVL_dIL	100,00	72,74	61,66
dV_dI_dIL	100,00	68,79	62,37

dVL_dI_dIL	100,00	70,08	62,19
dV_dVL_dI_dIL	100,00	68,88	62,06

Tabla 2. Resultados de precisión de los métodos de aprendizaje en arquitectura que considera todos los tipos de fallas.

Combinación de descriptores	Índice de precisión [%]		
	SVM	LAMDA	MF
dV	95,97	62,36	53,33
dVL	95,69	64,86	53,47
dI	97,78	60,83	61,94
dIL	96,11	65,28	61,94
dV_dVL	100,00	63,75	57,92
dV_dI	100,00	61,81	74,86
dV_dIL	100,00	64,44	75,00
dVL_dI	100,00	63,33	72,64
dVL_dIL	98,33	65,14	72,22
dI_dIL	100,00	64,17	71,81
dV_dVL_dI	100,00	63,06	83,19
dV_dVL_dIL	100,00	64,72	85,42
dV_dI_dIL	100,00	63,19	84,72
dVL_dI_dIL	100,00	64,58	83,06
dV_dVL_dI_dIL	100,00	64,17	85,28

Tabla 3. Resultados de precisión de los métodos de aprendizaje en arquitectura que considera la falla monofásica.

Se aprecia a partir de la comparación de los resultados obtenidos, que tanto para el caso monofásico como el que considera todas las fallas, la precisión de los tres métodos es superior al 53%.

Para el caso de LAMDA, los resultados son muy similares, tanto en el caso del clasificador monofásico, como el que considera todas las fallas. En el caso del MF, se aprecia que para el caso monofásico hay un notable incremento en la precisión en la medida que se aumenta la cantidad de información disponible o el número de descriptores, tal como se muestra en la tabla 3. La metodología basada en MF alcanza resultados promedio de precisión del 83%, para el caso de un conjunto de entrada de tres y cuatro descriptores. Finalmente, para el caso más satisfactorio corresponde a las SVM, los resultados de precisión en el caso de la arquitectura monofásica, como aquella diseñada para cualquier tipo de falla son superiores al 90% y muchos de ellos iguales al 100%.

6. CONCLUSIONES

En este artículo se presenta la metodología y los resultados de aplicación de tres métodos de aprendizaje y clasificación, al problema de localización de fallas en sistemas de distribución de energía eléctrica.

Los métodos basados en técnicas de aprendizaje y clasificación, son modelos de caja negra que se adaptan perfectamente al problema de localización de fallas. Como entrada tienen descriptores o características

tomadas de las señales de tensión y corriente registradas en la subestación de distribución. Como salidas, los métodos de aprendizaje propuestos muestran la zona más probable en la cual ocurrió la falla.

Como resultado se aprecia que el método basado en Máquinas de Soporte Vectorial – SVM, es el que mejor se comporta ante las dos arquitecturas propuestas y para los diferentes conjuntos de descriptores de entrada, con resultados promedio mayores al 90% en el índice de precisión. Estos resultados se obtuvieron en el circuito utilizado como prototipo de prueba.

7. BIBLIOGRAFÍA

- [1] M. Bollen. "Understanding Power Quality Problems: Voltage Sags and Interruptions". IEEE Press. 2000.
- [2] J. Mora. "Localización de Faltas en Sistemas de Distribución de Energía Eléctrica usando Métodos Basados en el Modelo y Métodos Basados en el Conocimiento", Tesis Doctoral. Universitat de Girona, España. 2006.
- [3] IEEE Std C37.114. "IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines" Power System Relaying Committee 2004.
- [4] J. Mora, J. Meléndez, G. Carrillo. "Comparison of impedance based fault location methods for power distribution system". Electric Power System Research. 2006 (accepted Paper).
- [5] R. Das. "Determining the locations of faults in distribution systems". Canada, 1998, 206 p. Doctoral dissertation. University of Saskatchewan Saskatoon.
- [6] M. Saha, E. Rosolowski. "Method and device of fault location for distribution networks. 2002". US Patent number 6,483,435 B2.
- [7] V. Vapnik "The nature of Statistical Learning Theory" Second Edition, Springer Verlag, 2000.
- [8] J. Aguilar, M. Balssa, R López. "Estimation recursive d'une partition. Exemples d'apprentissage et auto-apprentissage dans Rn et In". Rapport technique 880I39, LAAS/CNRS, 1980.
- [9] G. McLachlan Geoffrey, D. Peel. "Finite Mixture Models". John Wiley & Sons, New York, 2000.
- [10] B. Smola. "Learning with Kernels Support Vector Machines, Regularization, Optimization and Beyond". The MIT Press, Cambridge. 2002.
- [11] N. Piera, P. Desroches, J. Aguilar. "An incremental conceptual clustering system. Report technique 89420, LAAS/CNRS, 1989.
- [12] J. Waissman. "Construction d'un modele comportemental pour la supervision de procedes: Application a une station de traitement des eaux" Doctoral thesis - Institut National Polytechnique de Toulouse, Pgs 17-28. 2000
- [13] G. Morales, A. Gómez. "Estudio e implementación de una herramienta basada en máquinas de soporte vectorial aplicada a la localización de fallas en sistemas de distribución". Trabajo de grado. Universidad Industrial de Santander. 2005.
- [14] J. Mora, J Meléndez, J. Bedoya. "Extensive Events Database Development using ATP and Matlab to Fault Location in Power Distribution Systems". In proceedings of 2006 IEEE PES Transmission and Distribution Conference and Exposition, August, Caracas, 2006.