

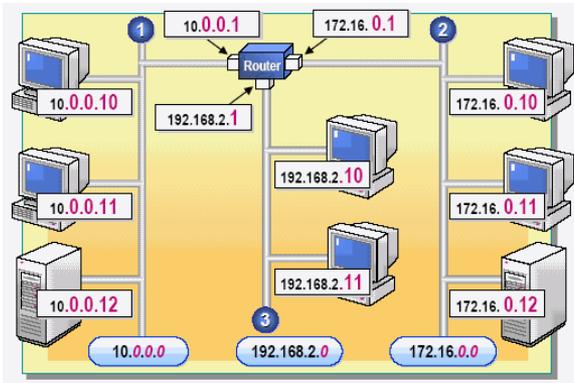
Classification of Hosts in a WLAN using a system based on Support Vector Machine classifier and a Bayesian classifier



M. F. Jojoa



Diferencias entre: Anfitrión – Cliente - Usuario



¿ Cual es la necesidad de clasificar los dispositivos dentro de una LAN – WLAN?

- Organización
- Gestión de anchos de banda
- Congestión
- Escalabilidad



Clasificador Bayesiano

Se asume dos clases, w_1 y w_2 y conocidas las probabilidades condicionales $p(X/w_i)$ y a priori $p(w_i)$. La regla de decisión se establece entre probabilidades a posteriori como sigue [1]:

$$p(W_1|X) \underset{<}{\overset{>}{}} p(W_2|X) \rightarrow X \in \begin{matrix} W_1 \\ W_2 \end{matrix} \text{ con } p(W_i|X) = \frac{p(X|W_i)p(W_i)}{p(X)}$$

$$p(X|W_1)p(W_1) \underset{<}{\overset{>}{}} p(X|W_2)p(W_2) \rightarrow X \in \begin{matrix} W_1 \\ W_2 \end{matrix}$$

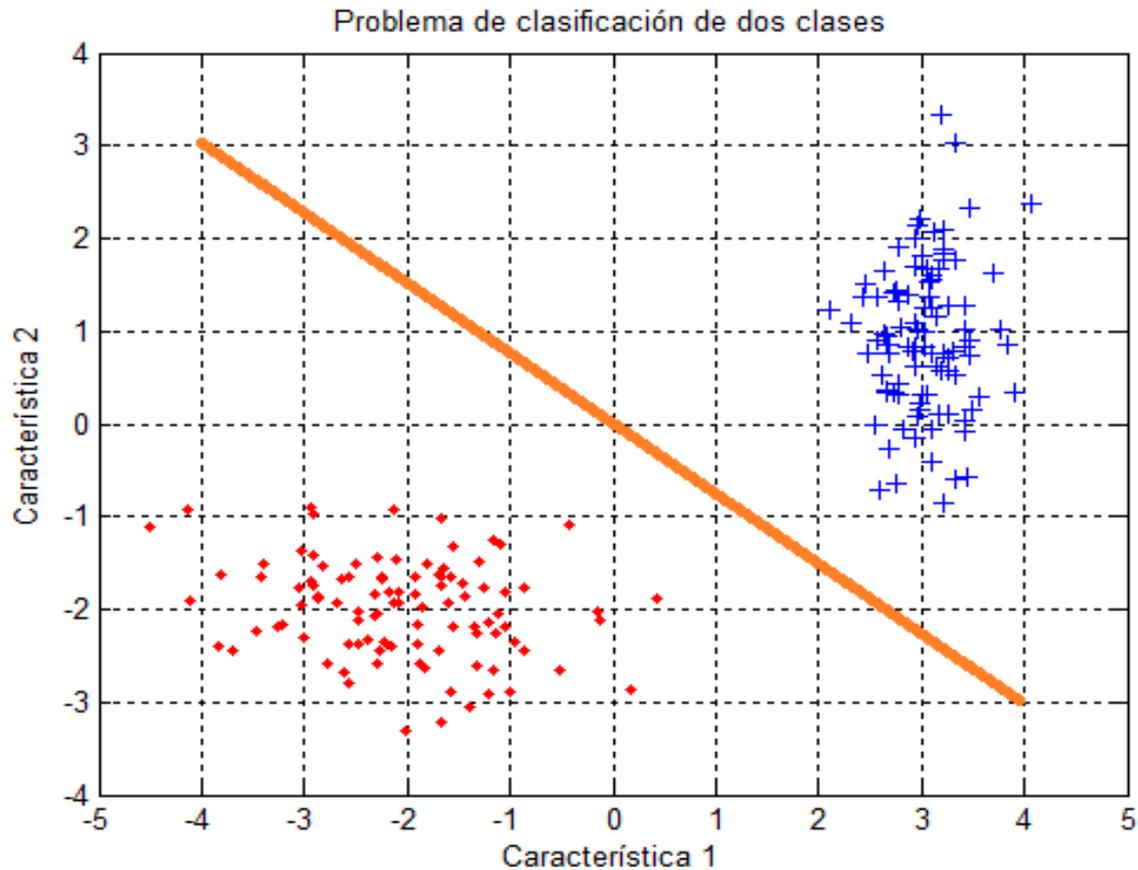
[1] Test de Bayes de mínimo error. [Probabilidad y procesos estocásticos, Harold Romo 2012]

Maquina de Soporte Vectorial

Este tipo de clasificador basa su funcionamiento en las siguientes características [2]:

- Hace una estimación de las funciones de distribución de probabilidad.
- Configura hiperplanos en el aprendizaje lineal de las máquinas.
- Su configuración es única.

Problema de Clasificación de dos clases

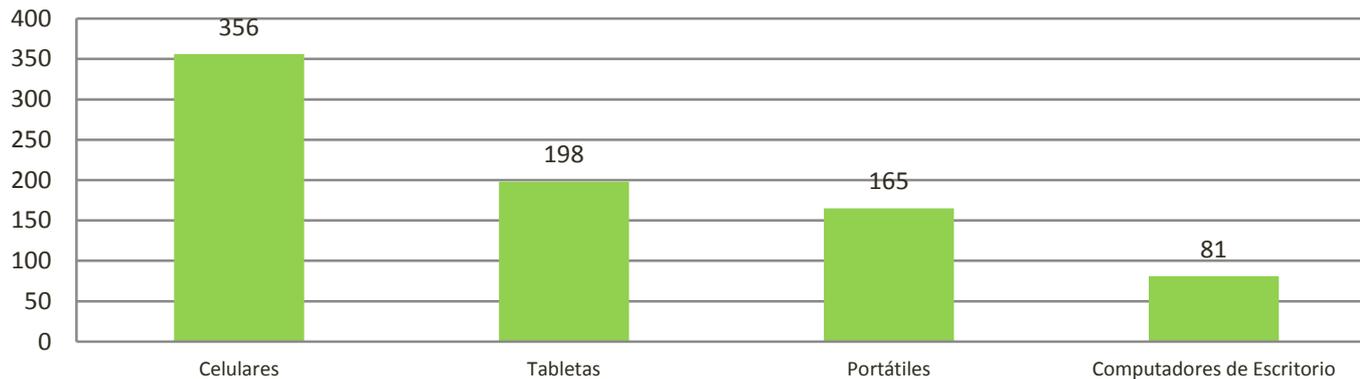


Características de los datos

Clase	Vector de medias [Volumen, Tiempo]	Matriz de covarianza
Celulares	[2.01, 93.47]	$\begin{bmatrix} 0,3149 & -1,3397 \\ -1,3397 & 92,6862 \end{bmatrix}$
Tabletas	[7.14, 45.14]	$\begin{bmatrix} 0,9912 & 1,6217 \\ 1,6217 & 80,58 \end{bmatrix}$
Portátiles	[20.15, 58.49]	$\begin{bmatrix} 7,97 & -1,5384 \\ -1,5384 & 55,6705 \end{bmatrix}$
PC Escritorio	[24,87 119,97]	$\begin{bmatrix} 3,2617 & 0,1869 \\ 0,1869 & 3,5477 \end{bmatrix}$

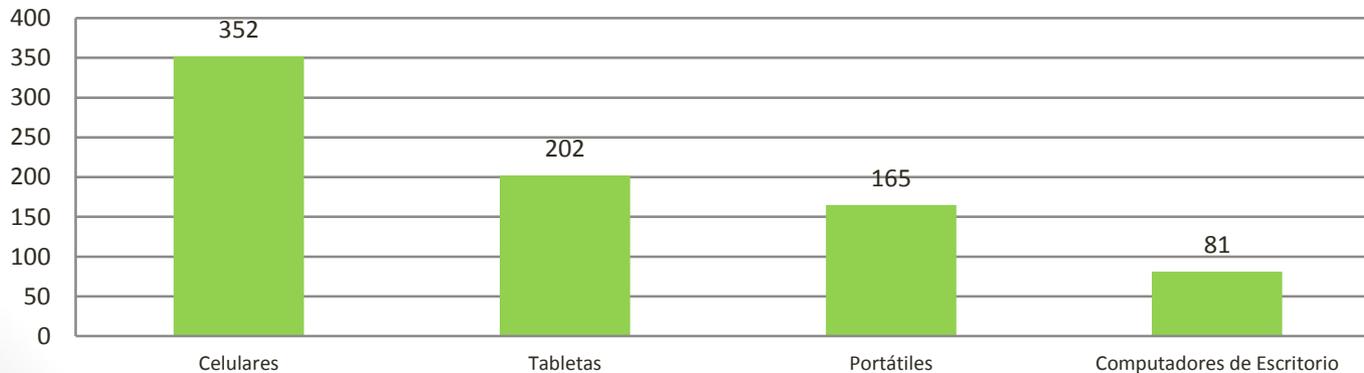
Valores de Entrada Vs Clasificación Bayesiano

Valores de entrada para la prueba de los clasificadores

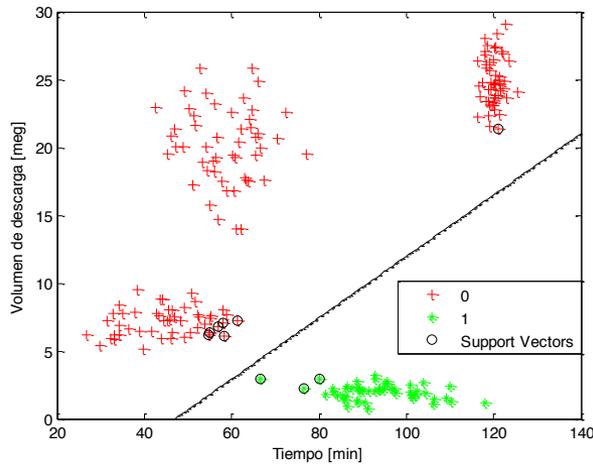


Error Promedio de Clasificación 0,99%

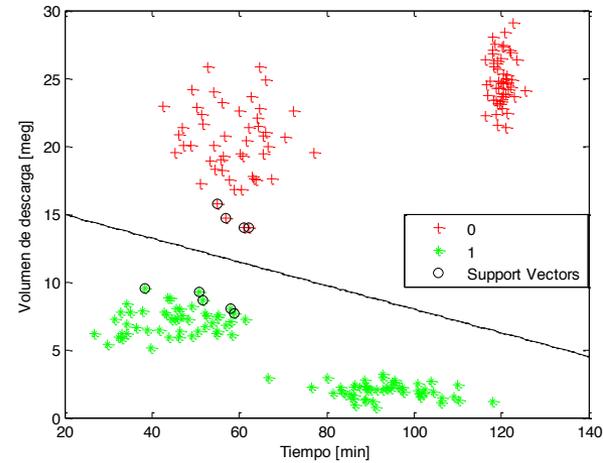
Valores de salida del clasificador Bayesiano



Entrenamiento del SVM

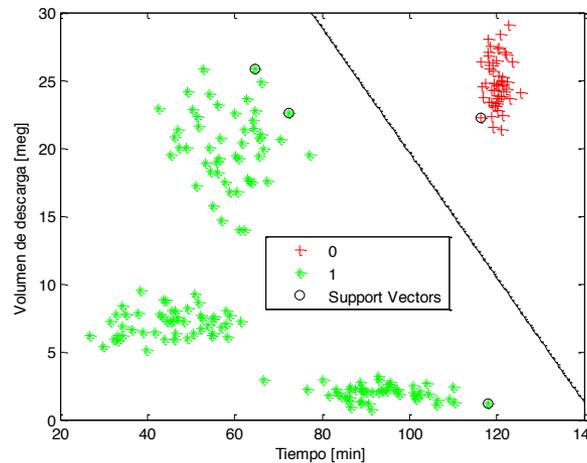


Celulares



Celulares y Tabletas

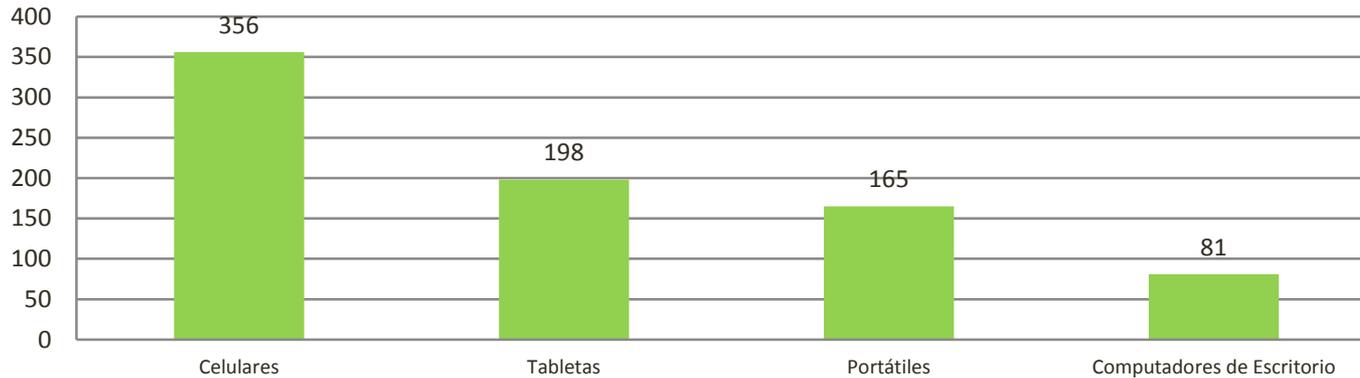
Entrenamiento con
50 valores conocidos



Celulares, Tabletas y
Portátiles

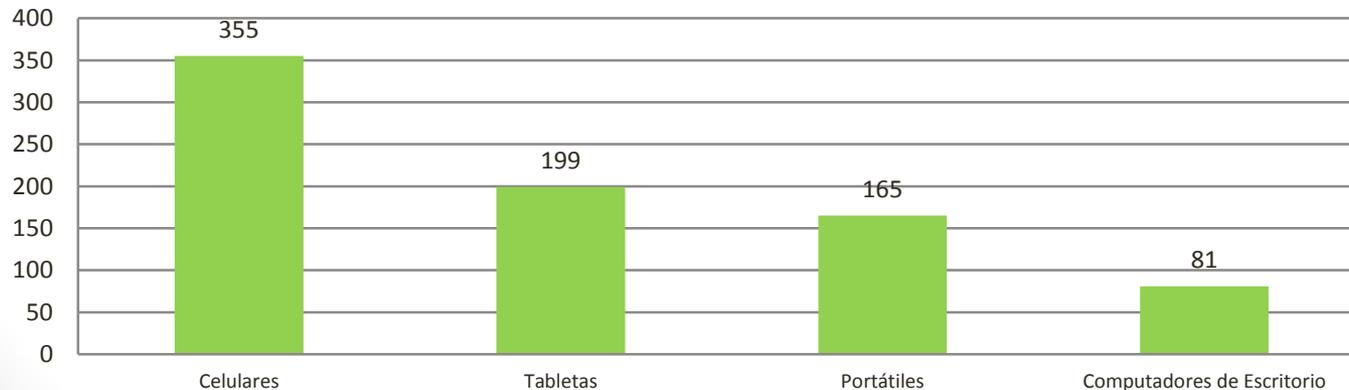
Valores de Entrada Vs Clasificación SVM

Valores de entrada para la prueba de los clasificadores



Error Promedio de Clasificación 0,24%

Valores de salida del clasificador SVM



Conclusiones

- En la administración de una red de datos se requiere conocer además del número de clientes conectados, el tipo de dispositivo y el volumen de descarga demandada; parámetros con los cuales se puede optimizar el uso de los recursos de una organización.
- Para aplicaciones prácticas, por eficiencia, se prefiere implementar un clasificador de baja complejidad computacional pero igualmente eficaz, lo cual apunta al clasificador SVM como la mejor opción de los dos.
- El clasificador Bayesiano, para su implementación requiere de un previo análisis estadístico de los datos, cuyos parámetros como vector de medias y covarianzas serán más confiables, cuanto mayor es el número de muestras o datos de entrenamiento, lo cual no siempre es posible de lograr.

Muchas Gracias

- ¿?