

Rodrigo Maciel Bobadilla
Licenciatura en Actuaría, Facultad de Ciencias Actuariales
rmacielbobadilla@gmail.com



Carlos Cuevas Covarrubias
CIEMA, Facultad de Ciencias Actuariales
ccuevas@anahuac.mx

INTRODUCCIÓN

Un problema de clasificación supervisada consiste en la observación de un vector aleatorio $X: \Omega \rightarrow \mathbb{R}^p$ en donde $\Omega = \Omega_0 \cup \Omega_1$; $\Omega_0 \cap \Omega_1 = \emptyset$. Dado un valor x realización de X , el objetivo es inferir si $w \in \Omega_0$ o $w \in \Omega_1$. En este trabajo implementaremos un algoritmo genético que permite construir una función lineal discriminante $S(x) = \alpha'x$ y aplicar la siguiente regla de decisión:

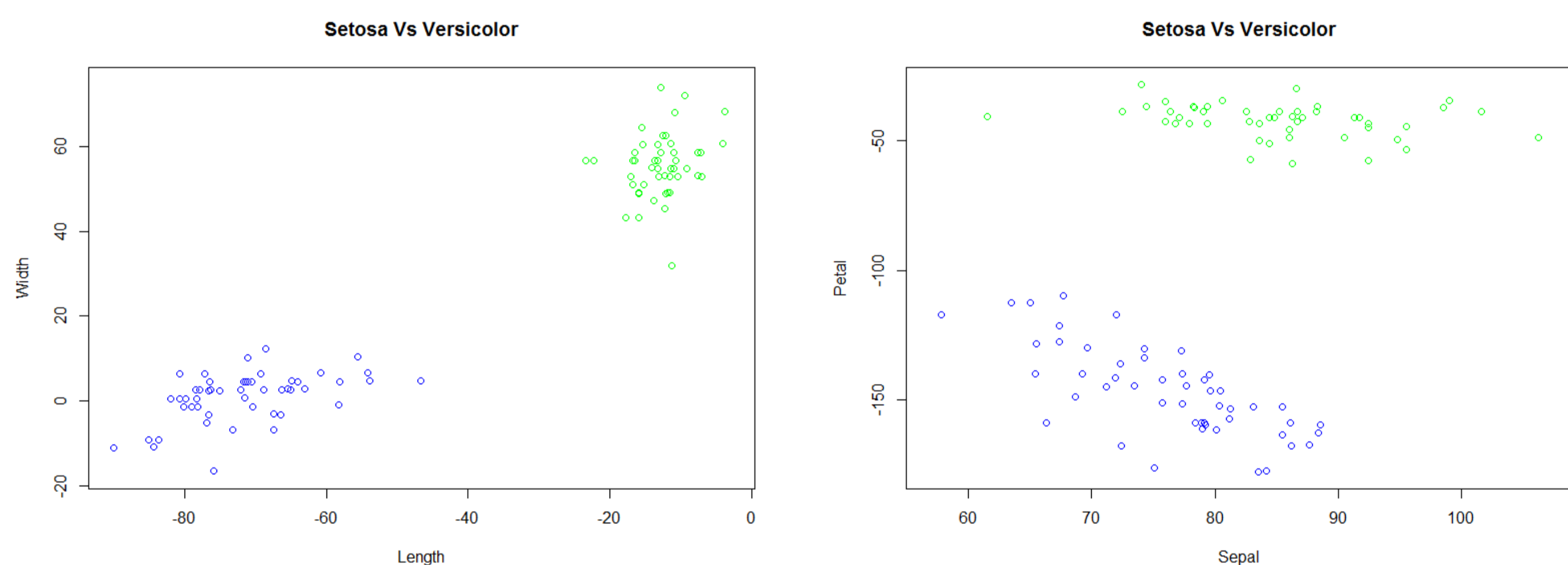
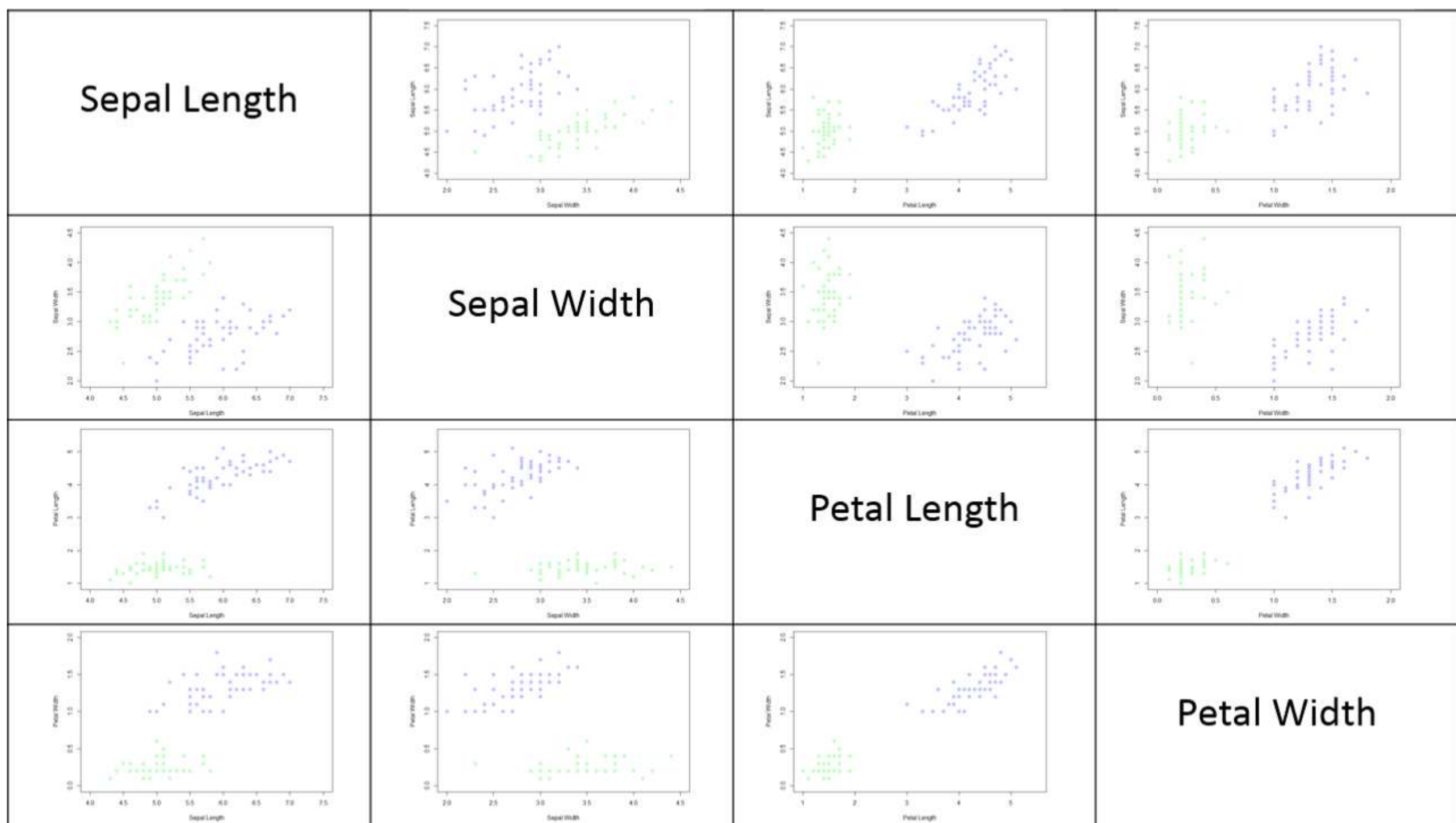
$$\text{Clasifica en } \begin{cases} \Omega_0 & \text{si } s \leq t \\ \Omega_1 & \text{si } s > t \end{cases}$$

METODOLOGÍA

El algoritmo genético consiste en:

- i. Definición de una población inicial
- ii. Evaluación de la aptitud mediante una función objetivo
- iii. Selección de individuos con probabilidad proporcional a la medida de aptitud
- iv. Cruzamiento y construcción de una nueva generación
- v. Mutación y volver al paso ii

Este algoritmo se programó en R

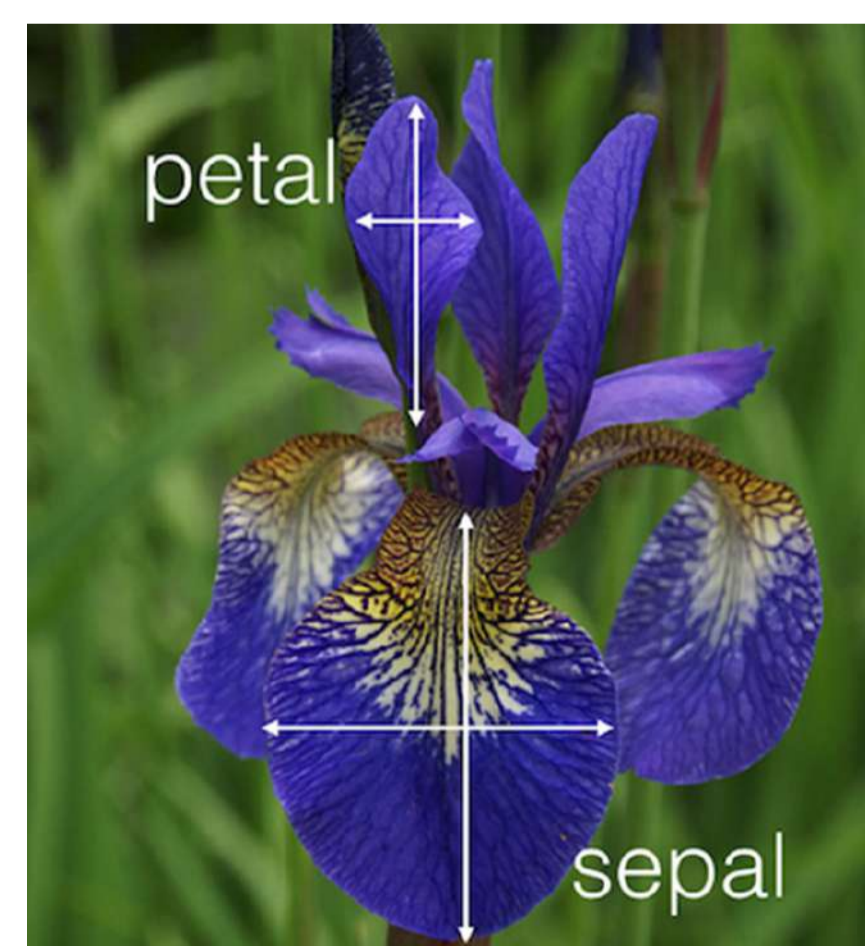
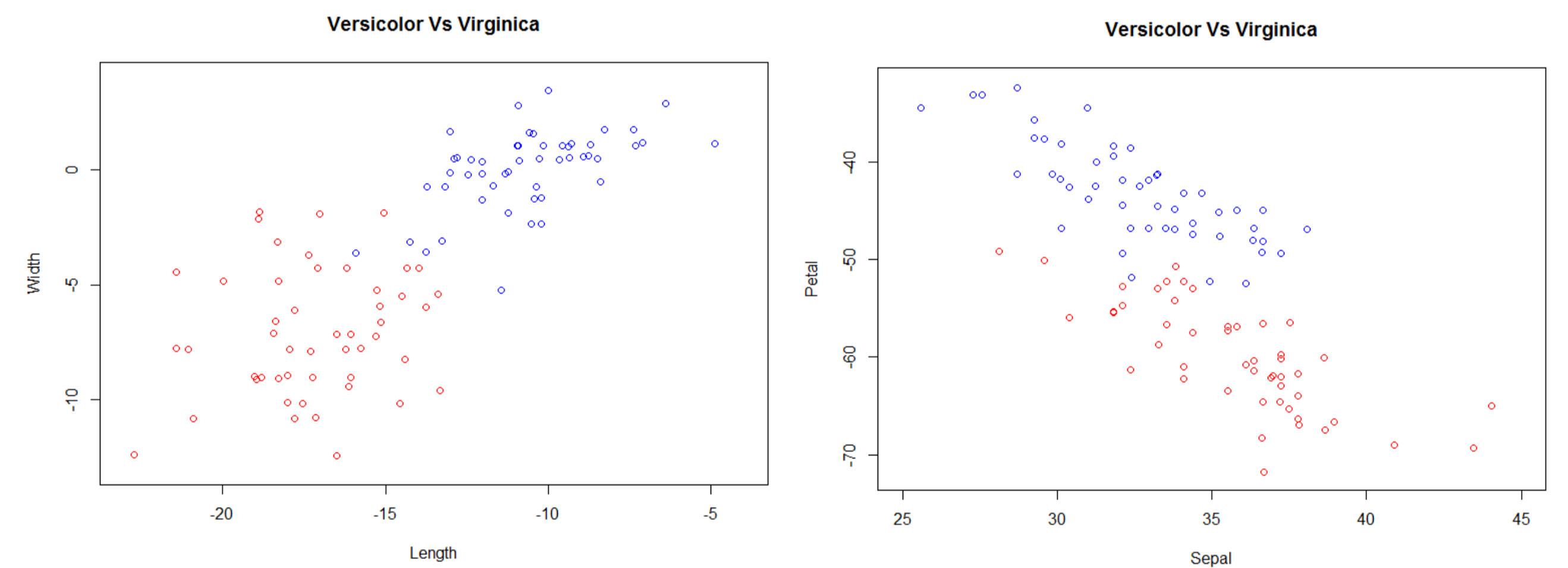
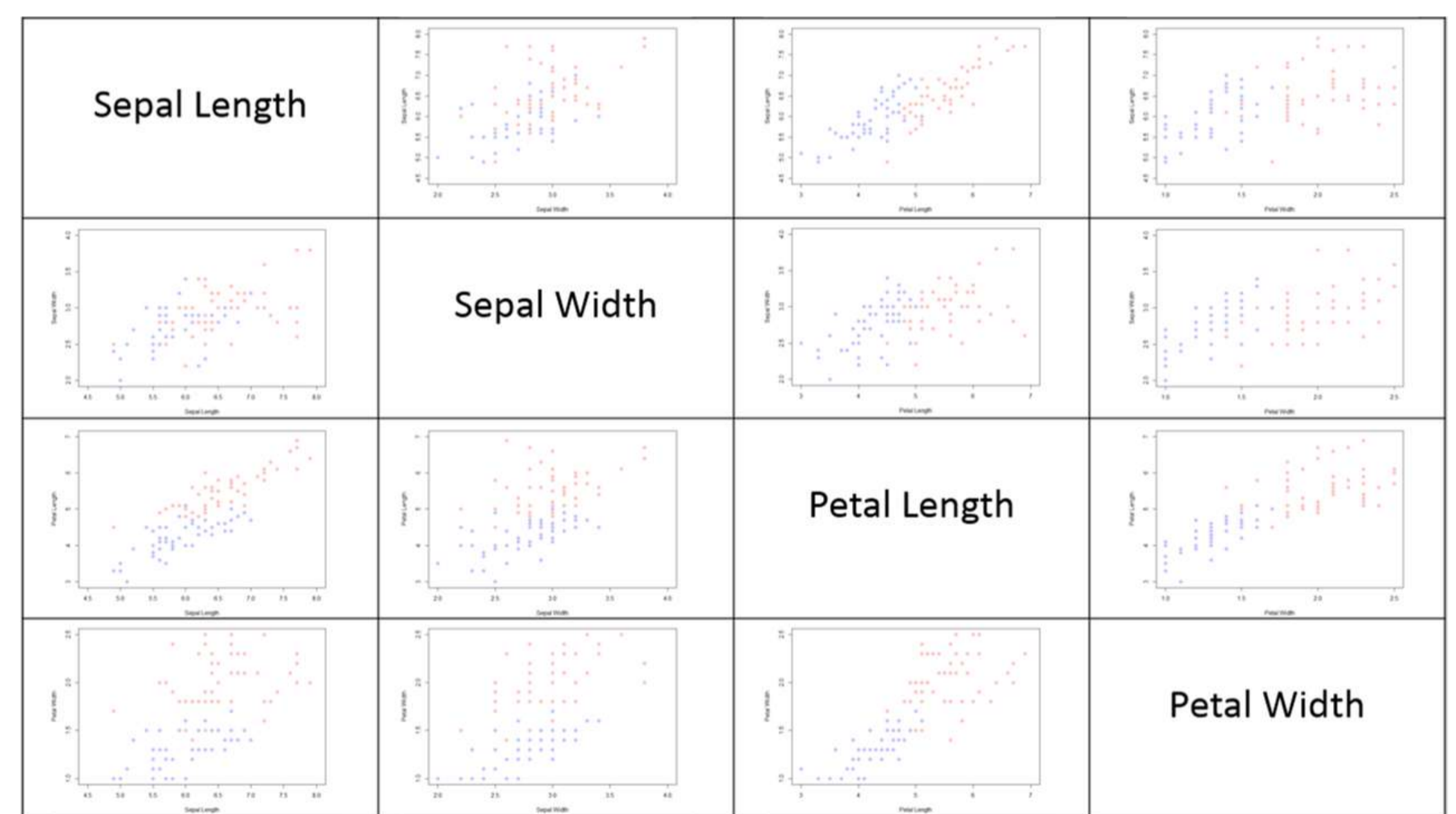


RESULTADOS

A manera de prueba, se realizó un análisis discriminante con la base de los iris de Fisher. Los resultados obtenidos fueron los siguientes:

Iris Setosa Vs Iris Versicolor		
Modelo	Función lineal discriminante	Porcentaje de clasificación correcta
Discriminante de Fisher	$2.85 \text{ SL} + 18.42 \text{ SW} - 21.21 \text{ PL} - 32.59 \text{ PW} + 13.81$	100%
Algoritmo genético	$3.95 \text{ SL} + 19.01 \text{ SW} - 22.29 \text{ PL} - 38.94 \text{ PW} + 13.82$	100%

Iris Setosa Vs Iris Versicolor		
Modelo	Función lineal discriminante	Porcentaje de clasificación correcta
Discriminante de Fisher	$3.55 \text{ SL} + 5.57 \text{ SW} - 6.97 \text{ PL} - 12.38 \text{ PW} + 16.66$	97%
Algoritmo genético	$2.86 \text{ SL} + 5.63 \text{ SW} - 6.48 \text{ PL} - 11.76 \text{ PW} + 17.47$	98%



DISCUSIÓN

Los algoritmos genéticos son un mecanismo efectivo de optimización en estadística multivariada. Su aplicación permite construir funciones lineales discriminantes óptimas con respecto a diversos criterios.

REFERENCIAS

1. Duda RO, Hart PE, Stork DG. Pattern Classification. 2a ed. Nueva Jersey: John Wiley & Sons; 2000.
2. Montano A. Una aplicación de los algoritmos genéticos en la discriminación. Revista Investigación Operacional. 2008;29.