

Tarea 05 - Reconocimiento de patrones

Giovanni Gamaliel López Padilla

Problema 01

Explora en <https://colab.research.google.com/drive/1pwCqLdvxeqChzDG3MoFigR7m6lsInKs5?usp=sharing> el efecto de cambiar los parámetros en una SVM y convéncete que es de acuerdo a (congruente con) el funcional de costo de una SVM.

Experimento 1

¿Cuál es el efecto de cambiar el parámetro λ/γ (cost)?

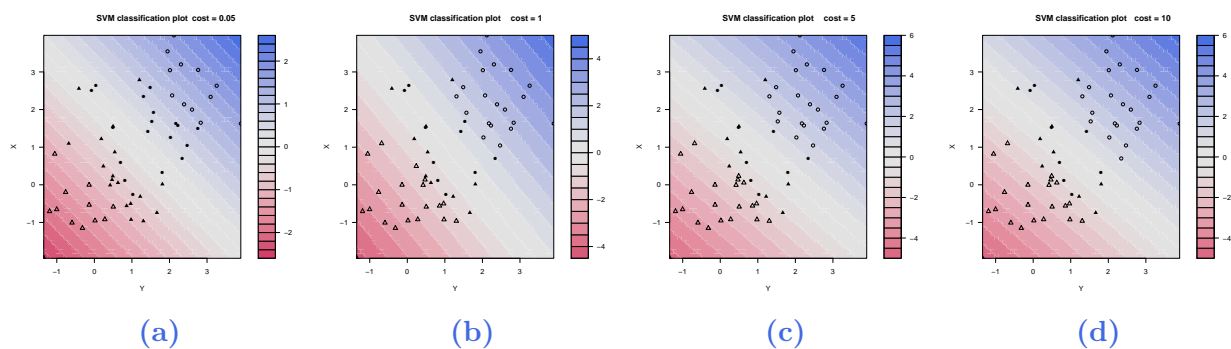


Figura 1

Experimento 2

¿Cuál es el efecto de aumentar el grado de polinomio?

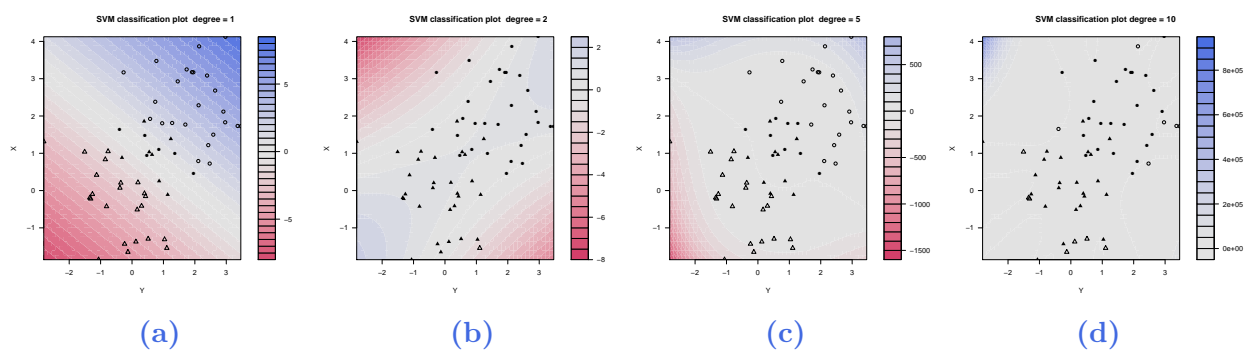


Figura 2

Experimento 3

¿Cuál es el efecto de cambiar el parámetro de kernel de base radial?

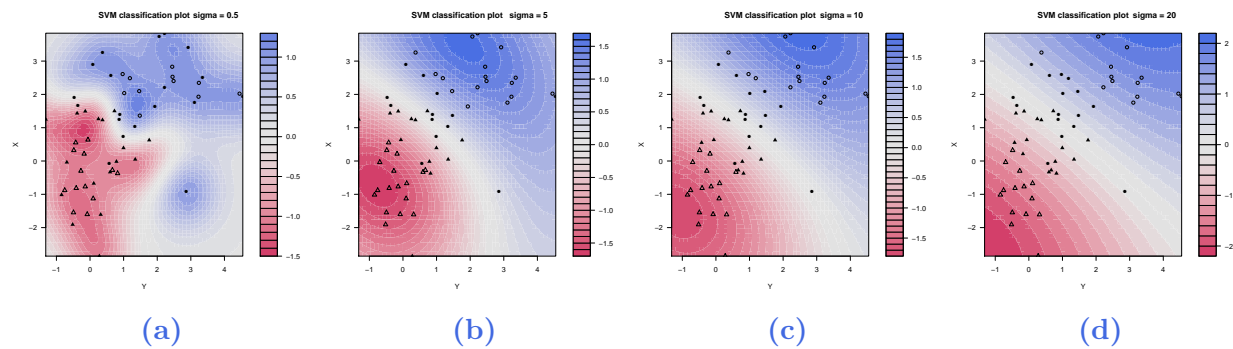


Figura 3

Experimento 4

¿Cuál es el efecto de cambiar el parámetro σ del kernel de base radial?

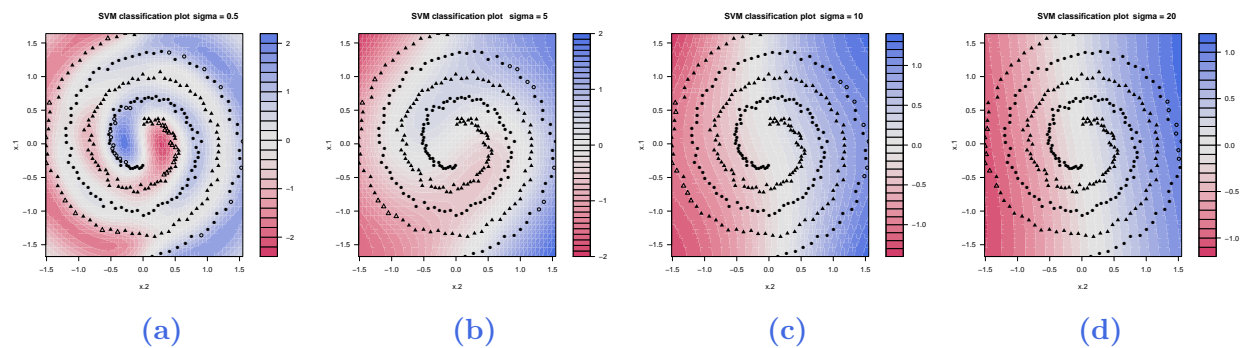


Figura 4

Problema 2

Vimos que minimizar $E(1 - Y(g(X)))_+$ sobre g conduce al clasificador óptimo $\hat{y}(x) = \text{sgn}(g(x))$. Usando el mismo camino, muestra que se obtiene el mismo resultado para $E(\exp(-Yg(X)))$

Problema 3

Supongamos que (X, Y) cumplen los supuestos del clasificador binario LDA. Sin embargo, a partir de una muestra de (X, Y) , alguien decida usar QDA (el clasificador bayesiano óptimo para el caso donde $X|Y = Y \sim \mathcal{N}(\mu_y, \Sigma_y)$, o sea no aprovechar que las covarianzas son iguales) y no LDA.

Experimento 1

¿Cómo se comparan el error de entrenamiento de QDA con el de LDA para este caso? No hay que hacer cálculos formales sino dar argumentos intuitivos

Problema 4

Este ejercicio es sobre el uso de métodos de clasificación para detectar billetes falsos:

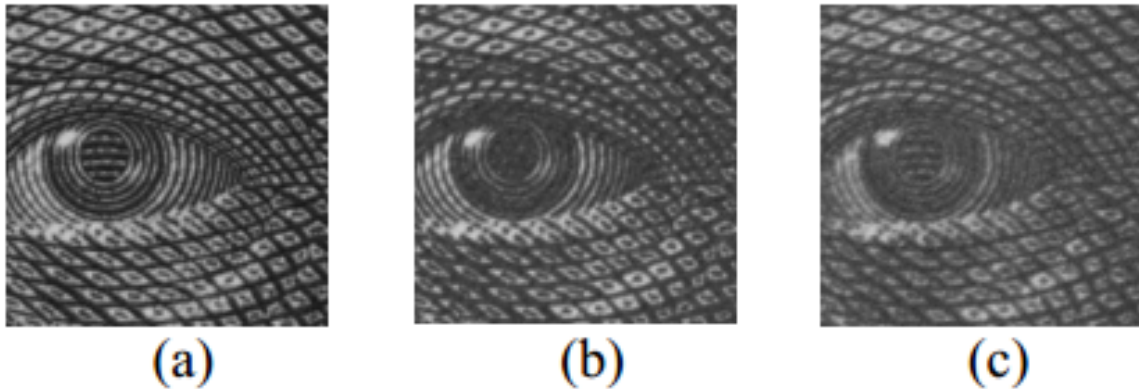


Figura 5: (a) (parte de) un billete de verdad, (b) billete falso de alta calidad, (c) billete falso de baja calidad.

En el paper que se anexa a la tarea se resume cada billete con cuatro características (varianza, skewness, curtosis y entropía) extraídas de la forma del histograma de los coeficientes de la transformación de Wavelet. Los histogramas a continuación muestran como cambia la forma cuando el billete ya no es auténtico.

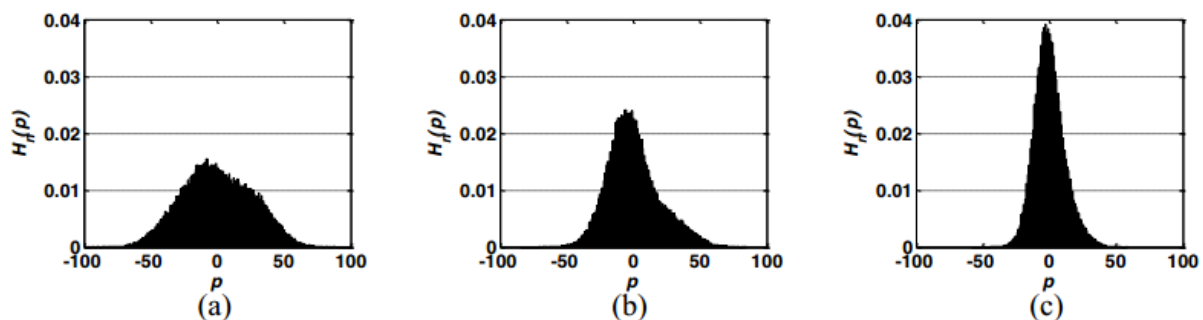


Figura 6: (a) histograma de los coeficientes de un billete de verdad, (b) billete falso de alta calidad, (c) billete falso de baja calidad.

Se anexo el conjunto de datos. La última columna indica si el billete es falso o no.

- Resume, visualiza y analiza los datos
- Construye algunos clasificadores interesantes basados en SVM (explora diferentes kernels). Estima su poder predictivo, para eso divide muchas veces los datos en conjunto de prueba y de entrenamiento y cuenta falsos positivos y falsos negativos. Las instrucciones básicas de SVM para R y Python están al fianl de [recpat4b.pdf](#)