UNIVERSIDAD
DE GRANADA

Aprendizaje Automático
grado en ingeniería informática

TRABAJO 3

Cuestiones de Teoría

Autor
Vladislav Nikolov Vasilev

Rama
Computación y Sistemas Inteligentes

Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación

Curso 2018-2019

Índice

Ejercicio 1	**2**
Ejercicio 2	**2**
Ejercicio 3	**2**
Ejercicio 4	**6**
Ejercicio 5	**7**
Ejercicio 6	**7**
Ejercicio 7	**7**
Ejercicio 8	**7**
Ejercicio 9	**8**
Ejercicio 10	**8**
Referencias	**9**

Ejercicio 1

¿Podría considerarse Bagging como una técnica para estimar el error de predicción de un modelo de aprendizaje? Diga si o no con argumentos. En caso afirmativo compárela con validación cruzada.

Solución

Ejercicio 2

Considere que dispone de un conjunto de datos linealmente separable. Recuerde que una vez establecido un orden sobre los datos, el algoritmo perceptron encuentra un hiperplano separador interando sobre los datos y adaptando los pesos de acuerdo al algoritmo

Algorithm 1 Perceptron

Entradas: $(\mathbf{x}_i, y_i) = 1, \ldots, n$, $w = 0$, $k = 0$
repeat
  $k \leftarrow (k+1) \mod n$
  if $\text{sign}(y_i) \neq \text{sign}(\mathbf{w}^T \mathbf{x}_i)$ **then**
    $\mathbf{w} \leftarrow \mathbf{w} + y_i \mathbf{x}_i$
  end if
until todos los puntos bien clasificados

Modificar este pseudo-código para adaptarlo a un algoritmo simple de SVM, considerando que en cada iteración adaptamos los pesos de acuerdo al caso peor clasificado de toda la muestra. Justificar adecuadamente/matematicamente el resultado, mostrando que al final del entrenamiento solo estaremos adaptando los vectores soporte.

Solución

Ejercicio 3

Considerar un modelo SVM y los siguientes datos de entrenamiento: Clase-1:$\{(1,1), (2,2), (2,0)\}$, Clase-2:$\{(0,0), (1,0), (0,1)\}$

a) Dibujar los puntos y construir por inspección el vector de pesos para el hiperplano óptimo y el margen óptimo.

Solución

Primero vamos a dibujar los puntos para ver como se ditribuyen en el espacio:

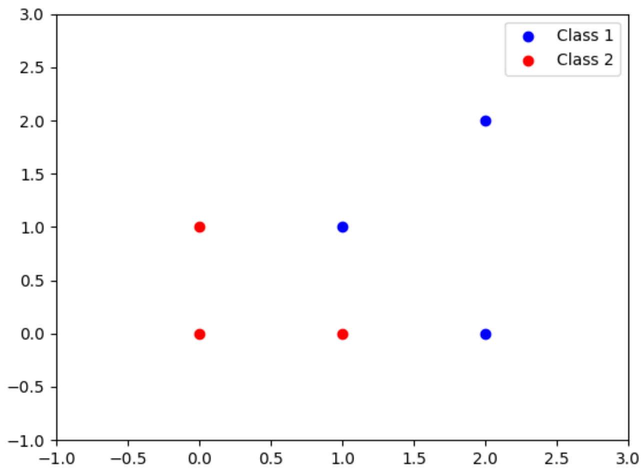

Figura 1: Dibujo con los puntos de las dos clases en el espacio.

Se puede ver claramente que los puntos de los dos clases son linealmente separables, ya que perfectamente se pueden separar mediante un hiperplano que pase en medio de ellas.

Para intentar obtener un hiperplano óptimo, vamos a suponer que éste tiene que pasar entre los puntos de las dos clases que estén más cerca entre sí (es decir, que tiene que pasar entre los vectores soporte). Estos puntos son, para la Clase-1, el $(1, 1)$ y el $(2, 0)$, y para la Clase-2 son el $(0, 1)$ y el $(1, 0)$. Por tanto, sabiendo que el hiperplano óptimo tiene que pasar entre estos puntos, dejando la mayor cantidad de margen a cada lado, podemos suponer que pasará justo en el punto medio para cada par de puntos que están a la misma altura y son de clases diferentes. Es decir, que para los puntos $(0, 1)$ y $(1, 1)$ (los cuáles son de diferente clase), sabemos que seguramente ese hiperplano pasará por el $(0.5, 1)$. Para los dos puntos de abajo, sabiendo que tiene que pasar entre los puntos $(0, 0)$ y $(1, 0)$, seguramente pasará por el punto $(1.5, 0)$, el cuál está justo en medio de los dos anteriores. Esto se puede ver mejor en la siguiente imagen:

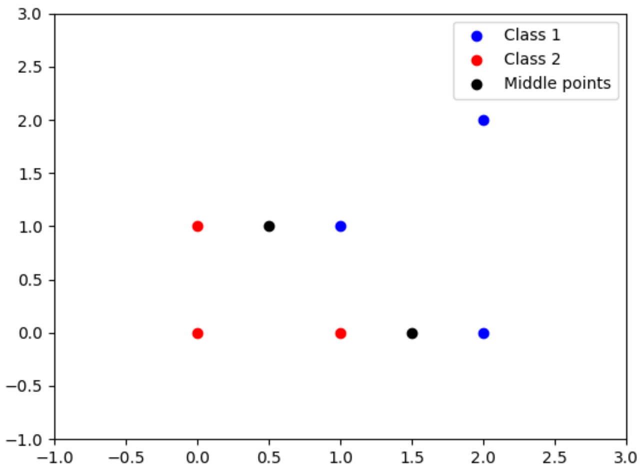

Figura 2: Dibujo con los puntos medios entre los vectores soporte de las dos clases, representados en negro.

Ahora lo único que nos queda es obtener el hiperplano que separa las dos clases. Como ya sabemos los puntos por los que puede pasar, lo único que tenemos que obtener es la recta que pasa por esos dos puntos. Para eso podemos partir de la ecuación de la recta, la cuál viene dada por la forma:

$$y = ax + b \tag{1}$$

donde a es la pendiente de la recta y b el término independiente. Sustituyendo los valores de los puntos por x e y en la expresión dada por (1), obtenemos el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} 1 &= 0.5a + b \\ 0 &= 1.5a + b \end{aligned}\right\} \tag{2}$$

Ahora resolvemos el sistema de ecuaciones para obtener la solución:

$$\left.\begin{aligned} 1 - 0.5a &= b \\ -1.5a &= b \end{aligned}\right\} \tag{3}$$

$$\begin{aligned} 1 - 0.5a &= -1.5a \\ 1 &= -a \end{aligned} \tag{4}$$

De aquí, obtenemos que:

$$\left.\begin{aligned} a &= -1 \\ b &= 1.5 \end{aligned}\right\} \tag{5}$$

Y finalmente, con los resultados obtenidos en (5), sustityendo en la expresión dada en (1), obtenemos que la ecuación de la recta es la siguiente:

$$y = -x + 1.5 \tag{6}$$

Esta recta es, en un principio, el hiperplano óptimo que separa las dos clases. Para obtener los márgenes, lo único que tenemos que hacer es obtener rectas paralelas a las del hiperplano que pasen por los vectores soporte. Para ello, lo único que tenemos que modificar es el valor de b que hemos obtenido con tal de obtener cada margen (el coeficiente libre indica el desplazamiento en el eje X que hace la recta para cortar con este eje en $x = 0$). En los dos casos es muy fácil obtener estos valores de b, ya que algunos de los vectores soporte están sobre el eje X. Por tanto, tenemos que para la Clase-1, la recta que representa el margen es la siguiente:

$$y = -x + 2 \tag{7}$$

Para la Clase-2, la recta es la siguiente:

$$y = -x + 1 \tag{8}$$

Por tanto, veamos como quedaría gráficamente el resultado de pintar el hiperplano óptimo y los márgenes:

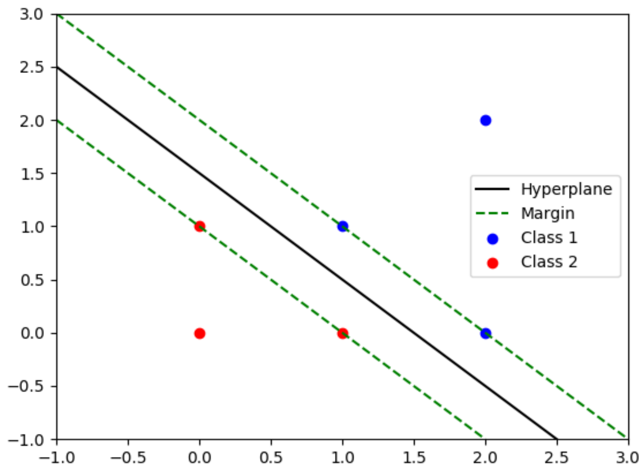

Figura 3: Dibujo del hiperplano óptimo con las dos clases y los márgenes para cada clase.

b) ¿Cuáles son los vectores soporte?

Solución

Los vectores soporte son los siguientes:

- Para la **Clase-1**, los vectores soporte son $(1, 1)$ y $(2, 0)$.
- Para la **Clase-2**, los vectores soporte son $(0, 1)$ y $(1, 0)$.

c) Construir la solución en el espacio dual. Comparar la solución con la del apartado (a)

Solución

Ejercicio 4

¿Cúal es el criterio de optimalidad en la construcción de un árbol? Analice un clasificador en árbol en términos de sesgo y varianza. ¿Que estrategia de mejora propondría?

Solución

Ejercicio 5

¿Cómo influye la dimensión del vector de entrada en los modelos: SVM, RF, Boosting and NN?

Solución

Ejercicio 6

El método de Boosting representa una forma alternativa en la búsqueda del mejor clasificador respecto del enfoque tradicional implementado por los algoritmos PLA, SVM, NN, etc. a) Identifique de forma clara y concisa las novedades del enfoque; b) Diga las razones profundas por las que la técnica funciona produciendo buenos ajustes (no ponga el algoritmo); c) Identifique sus principales debilidades; d) ¿Cuál es su capacidad de generalización comparado con SVM?

Solución

Ejercicio 7

Discuta pros y contras de los clasificadores SVM y Random Forest (RF). Considera que SVM por su construcción a través de un problema de optimización debería ser un mejor clasificador que RF. Justificar las respuestas.

Solución

Ejercicio 8

¿Cuál es a su criterio lo que permite a clasificadores como Random Forest basados en un conjunto de clasificadores simples aprender de forma más eficiente? ¿Cuales son las mejoras que introduce frente a los clasificadores simples? ¿Es Random Forest óptimo en algún sentido? Justifique con precisión las contestaciones.

Solución

Ejercicio 9

En un experimento para determinar la distribución del tamaño de los peces en un lago, se decide echar una red para capturar una muestra representativa. Así se hace y se obtiene una muestra suficientemente grande de la que se pueden obtener conclusiones estadísticas sobre los peces del lago. Se obtiene la distribución de peces por tamaño y se entregan las conclusiones. Discuta si las conclusiones obtenidas servirán para el objetivo que se persigue e identifique si hay algo que lo impida.

Solución

Ejercicio 10

Identifique que pasos daría y en que orden para conseguir con el menor esfuerzo posible un buen modelo de red neuronal a partir una muestra de datos. Justifique los pasos propuestos, el orden de los mismos y argumente que son adecuados para conseguir un buen óptimo. Considere que tiene suficientes datos tanto para el ajuste como para el test.

Solución

Referencias

[1] Texto referencia
`https://url.referencia.com`