



UNIVERSIDAD DE GRANADA

APRENDIZAJE AUTOMÁTICO
GRADO EN INGENIERÍA INFORMÁTICA

TRABAJO 2

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 7	3
Referencias	4

Ejercicio 1

Identificar de forma precisa dos condiciones imprescindibles para que un problema de predicción puede ser aproximado por inducción desde una muestra de datos. Justificar la respuesta usando los resultados teóricos estudiados.

Solución

Ejercicio 2

El jefe de investigación de una empresa con mucha experiencia en problemas de predicción de datos tras analizar los resultados de los muchos algoritmos de aprendizaje usados sobre todos los problemas en los que la empresa ha trabajado a lo largo de su muy dilatada existencia, decide que para facilitar el mantenimiento del código de la empresa van a seleccionar un único algoritmo y una única clase de funciones con la que aproximar todas las soluciones a sus problemas presentes y futuros. ¿Considera que dicha decisión es correcta y beneficiará a la empresa? Argumentar la respuesta usando los resultados teóricos estudiados.

Solución

Ejercicio 3

¿Que se entiende por una solución PAC a un problema de aprendizaje? Identificar el porqué de la incertidumbre e imprecisión.

Solución

En el ámbito del aprendizaje, una solución PAC significa que es *Probably Approximately Correct*, lo cuál traducido al español vendría a ser algo así como “correcta probablemente aproximada”. Veamos qué significa todo esto sobre la desigualdad de Hoeffding aplicada al problema de aprendizaje:

$$\mathbb{P}(\mathcal{D} : |E_{in} - E_{out}| > \varepsilon) \leq 2e^{-2\varepsilon^2 N} \quad (1)$$

- La parte de “probablemente” hace referencia a una alta probabilidad. En la expresión mostrada en (1), se puede ver una probabilidad de que algo malo suceda, es decir, que la diferencia entre los valores de E_{in} y E_{out} sea mayor que

un ε dado, o lo que es lo mismo, que los errores disten mucho entre sí. Como en la expresión de la parte derecha nos encontramos con un exponencial negativo, con los valores adecuados de ε y N podemos hacer que esa probabilidad de que algo malo pase sea pequeña. Por tanto, la probabilidad de que la diferencia sea menor que ε vendría dada por:

$$\mathbb{P}(\mathcal{D} : |E_{in} - E_{out}| < \varepsilon) \geq 1 - 2e^{-2\varepsilon^2 N}$$

la cuál sí que tendría un valor muy alto, siendo por tanto más “probable” que esa diferencia sea menor que ε .

- La parte de “aproximada” indica que E_{in} no es exactamente igual que E_{out} , pero que ambos valores están muy próximos. Esta aproximación viene dada por el valor de ε .

La **incertidumbre** viene dada por la probabilidad. Nunca se puede tener la certeza de que el resultado sea 100 % correcto, pero se puede afirmar con una alta probabilidad de que así sea (por eso es PAC). La **imprecisión**, por otro lado, viene dada por el valor de ε . Es decir, los valores de E_{in} y E_{out} , al estar aprendiendo de una muestra la cuál puede tener un tamaño no lo suficientemente grande o no ser muy representativa de la población, van a ser diferentes. Si pudiésemos aprender de toda la población directamente, en ese caso ε sería 0, ya que los dos errores serían iguales, pero habría que pagar muchos costes de tiempo, potencia de cómputo y almacenamiento. Por tanto, al estar siempre aprendiendo de una muestra y no de la población entera nos vamos a encontrar con estos dos problemas.

Ejercicio 7

¿Por qué la desigualdad de Hoeffding definida para clases \mathcal{H} de una única función no es aplicable de forma directa cuando el número de hipótesis de \mathcal{H} es mayor de 1? Justificar la respuesta.

Solución

Referencias

- [1] Texto referencia
<https://url.referencia.com>