

Výpočtové učenie

Písomka 6.4.2005

1. Dané sú reálne čísla a, b , $a \leq b$. Intervalový koncept $C_{a,b}$ je definovaný takto:

$$C_{a,b}(y) = 1, \text{ ak } y \text{ je v intervale } [a, b], \text{ inak je rovné } 0.$$

Nech H je hypotézový priestor všetkých intervalov, do ktorého patrí aj prázdny interval. Nasleduje učiaci algoritmus pre H :

```
empty:=true;
for i:=1 to m do
  if  $b_i = 1$  then
    if empty then begin
       $a:=x_i; b:=x_i;$ 
      empty:=false;
    end else begin
      if  $x_i > b$  then  $b:=x_i;$ 
      if  $x_i < a$  then  $a:=x_i;$ 
    end;
  if empty then  $L(s)$ :=prázdny interval else  $L(s)$ := $C_{a,b}$ 
```

Dokážte, že L je PAC s vhodne zvolenou dĺžkou vzorky. Akou?

2. Popíšte rozhodovacie zoznamy a konzistentný algoritmus pre ne. Dokážte, že pre všetky $n \geq k \geq 1$ platí, že $D_{n,k}$ je podmnožinou $DL(M_{n,k})$. Ukážte, že ľubovoľná booleovská funkcia môže byť reprezentovaná rozhodovacím zoznamom.
3. Sformulujte problém potenciálnej naučiteľnosti a dokážte, že ak H je potenciálne naučiteľný hypotézový priestor a L je konzistentný algoritmus, tak L je PAC.
4. Analyzujte čas behu učiacich algoritmov vzhľadom na odstupňovanie podľa n . Sformulujte vetu, ktorá vyjadrí situáciu pre konzistentný algoritmus L v odstupňovaných hypotézových priestoroch.
5. Nech $G = (V, E)$ je graf s vrcholmi $V = \{1, 2, 3, 4, 5, 6\}$ a hranami $E = \{12, 13, 15, 16, 23, 25, 34, 45, 26\}$. Vytvorte odpovedajúcu tréningovú vzorku $s(G)$, ktorá je použiteľná v dôkaze o NP-ťažkosti. Nájdite najmenšiu hodnotu k , pre ktorú je funkcia C_5^k konzistentná s $s(G)$ a uveďte explicitnú formulu pre takú funkciu.
6. Učenie s očíslovaním hypotéz. Predpokladajme, že hypotézy sú očíslované tak, že všetky, ktoré potrebujeme budú v prvej polovici celej postupnosti hypotéz. Vieme skontrolovať milión hypotéz za sekundu. Máme daný príkladový priestor $\{0, 1\}^9$. Ako dlho to bude trvať v hodinách?
7. Napíšte postupnosť hypotéz generovaných tréningovou vzorkou
(111001011,1), (001000111,0), (110010011,1), (000011110,1)
ak chceme dosiahnuť monočlen. Napíšte dodatočné príklady na to, aby sme dosiahli cieľový koncept $u_2u_4u_8u_9$.

Výpočtové učenie

Písomka 18.5.2005

1. VC dimenzia – definícia rastovej funkcie, definícia VC dimenzie, ohraničenie VC dimenzie konečných hypotézových priestorov, VC dimenzia reálneho perceptrónu.
2. VC dimenzia a potenciálna naučiteľnosť – veta o hypotézových priestoroch s nekonečnou VC dimenziou a jej dôkaz.
3. VC dimenzia a zložitosť tréningovej vzorky konzistentných algoritmov – dolné a horné hranice na veľkosť tréningovej vzorky.
4. Hľadanie najmenej konzistentnej hypotézy a Occam algoritmy, čo všetko s tým súvisí, definície, príklady.
5. Nech hypotézový priestor je tvorený všetkými konvexnými päťuholníkmi v rovine v ľubovoľnej polohe. Určte VCdim tohto hypotézového priestoru.