دانشگاه صنعتی شریف
دانشکده‌ی مهندسی کامپیوتر

درس یادگیری ماشین

استاد سید عباس حسینی

تمرین ششم | Decision Tree and Ensemble Learning | مصطفی قدیمی

سؤال ۱. ماشین بردار پشتیبان

- الف)
  - یک)
  - دو)
  - سه)
  - چهار)
- ب)
- پ)
  - یک)
  - دو)
  - سه)
  - چهار)
  - پنج)

سؤال ۲. هسته

- الف)
  - یک) از آن‌جایی که می‌دانیم $k_1(x, x')$ و $k_2(x, x')$ دو هسته‌ی معتبر هستند؛ بنابراین ماتریس هسته K_1 و K_2 هر دو مثبت نیمه‌معین هستند. پس اگر فرض کنیم که ورودی‌های x و x' هر دو، دو بعدی باشند، داریم:

$$k_3(x,x') = (x^T x')^2 = (x_1 x_1' + x_2 x_2')^2 = x_1^2 x_1'^2 + 2x_1 x_1' x_2 x_2' + x_2^2 x_2'^2$$
$$\implies k_3(x,x') = (x_1^2, \sqrt{2}x_1x_2, x_2^2)(x_1'^2, \sqrt{2}x_1'x_2', x_2'^2) = \phi(x)^T\phi(x')$$

  با توجه به نتیجه‌ی بالا داریم:

$$k_3 = k_1 + k_2$$

  پس k_3 نیز مثبت نیمه‌معین بوده و ثابت شد که معتبر نیز هست.
  - دو) با فرض این‌که تابع نگاشت هسته‌ی k_1، $\phi^{(1)}(x)$ با ابعاد M و تابع نگاشت هسته‌ی k_2، $\phi^{(2)}(x)$ با ابعاد N است، داریم:

$$k_4(x,x') = k_1(x,x')k_2(x,x') = \phi^{(1)}(x)^T\phi^{(1)}(x')\phi^{(2)}(x)^T\phi^{(2)}(x') = \Sigma_{i=1}^M \phi_i^{(1)}(x)\phi_i^{(1)}(x')\Sigma_{j=1}^N \phi_j^{(2)}(x)\phi_j^{(2)}(x')$$
$$\implies k_4(x,x') = \Sigma_{i=1}^M\Sigma_{j=1}^N[\phi_i^{(1)}(x)\phi_j^{(2)}(x)][\phi_i^{(1)}(x')\phi_j^{(2)}(x')] = \Sigma_{k=1}^M N\phi_k(x)\phi_k(x') = \phi(x)^T\phi(x')$$

  که در آن $\phi_i^{(1)}(x)$، عنصر iام $\phi^{(1)}(x)$ و $\phi_j^{(2)}(x)$ عنصر jام $\phi^{(2)}(x)$ است.
  - سه)
  از آن‌جایی که k_1 یک هسته‌ی معتبر است، بنابراین آن را به صورت $k_1(x,x') = \phi(x)^T\phi(x')$ می‌توان نوشت. پس:

$$k_5(x,x') = ck_1(x,x') = [\sqrt{a}\phi(x)]^T[\sqrt{a}\phi(x')]$$

  که با توجه به فرض $a \geq 0$ ثابت می‌شود.
  - چهار) اگر بسط تیلور را بنویسیم داریم:

$$k_6(x,x') = a_n k_1(x,x')^n + a_{n-1}k_1(x,x')^{n-1} + \cdots + a_1 k_1(x,x') + a_0$$

  حال با استفاده از نتایج سه قسمت قبل و با توجه به این‌که همه‌ی ضرایب بسط تیلور مثبت است، پس این هسته نیز یک هسته‌ی معتبر است.

- ب) با توجه به این‌که تابع هسته را می‌توان آن را به شکل ضرب داخلی در فضای ویژگی نوشت، پس هسته‌ی معتبر است. برای اثبات باید به رابطه‌ی $k(A,B) = 2^{|A\cap B|} = \phi(A)^T\phi(B)$ برسیم.

اگر

$$\phi_U(X) = \begin{cases} 1 & if\, U \subseteq X \\ 0 & otherwise \end{cases}$$

باشد، داریم:

$$\phi(A)^T\phi(B) = \Sigma_{U\subseteq|A\cap B|}\phi_U(A)\phi_U(B)$$

با استفاده از سیگما (جمع کردن) در رابطه‌ی بالا، همه‌ی زیرمجموعه‌های ممکن $|A \cap B|$ را اگر و تنها اگر هم زیرمجموعه‌ی A و B باشد (مقدار برابر با ۱) داریم. با این کار تعداد زیرمجموعه‌های اشتراک A و B در فضای S را محاسبه کرده‌ایم. علاوه‌براین هم A و هم B به عنوان زیرمجموعه‌ی فضای S معرفی شده‌اند، بنابراین:

$$\phi(A)^T\phi(B) = 2^{|A\cap B|}$$

- پ)
  - یک)

$$k(x,x') = (x^T.x'+c)^2 = k(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}, \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_d \end{pmatrix}) = (c + x_1x'_1 + x_2x'_2 + \cdots + x_dx'_d)^2$$

$$\implies k(x,x') = c^2\Sigma_{i=1}^{d} x_i^2 x_i'^2 + \Sigma_{i=1}^{d} 2cx_ix'_i + \Sigma_{i=1}^{d-1}\Sigma_{j=i+1}^{d} 2x_ix'_ix_jx'_j$$

$$\implies k(x,x') = \begin{pmatrix} c & x_1^2,\dots,x_d^2, & \dots & \sqrt{2}cx_d \end{pmatrix} \begin{pmatrix} c \\ x'_1,\dots,x'_d \\ \vdots \\ \sqrt{2}cx'_d \end{pmatrix}$$

$$\implies k(x,x') = \phi(x)^T\phi(x')$$

  - دو) اگر $c = 0$ باشد، آن‌گاه فضای تبدیل $d+1$ بعد کاهش می‌یابد؛ زیرا تعداد جملاتی که در آن c ضرب شده است، $d+1$ است که با صفر شدن آن حذف می‌شوند.
  - سه)

$$k(x,x') = (x^T.x'+c)^M$$

با توجه به توضیحاتی که در این لینک داده شده است، تعداد جملات برابر با:

$$\#\ of\ expressions = \binom{M+d+1-1}{d+1-1} = \binom{M+d}{d}$$

سؤال ۳. درخت تصمیم

سؤال ۴. یادگیری جمعی

- الف)
- ب)

سؤال ۵. Adaboost