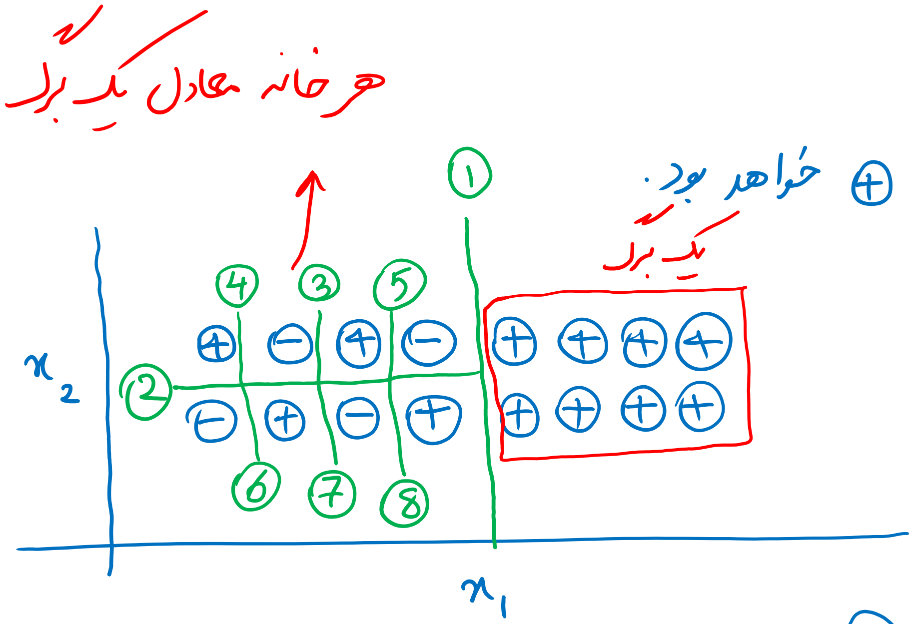
سوال ۴

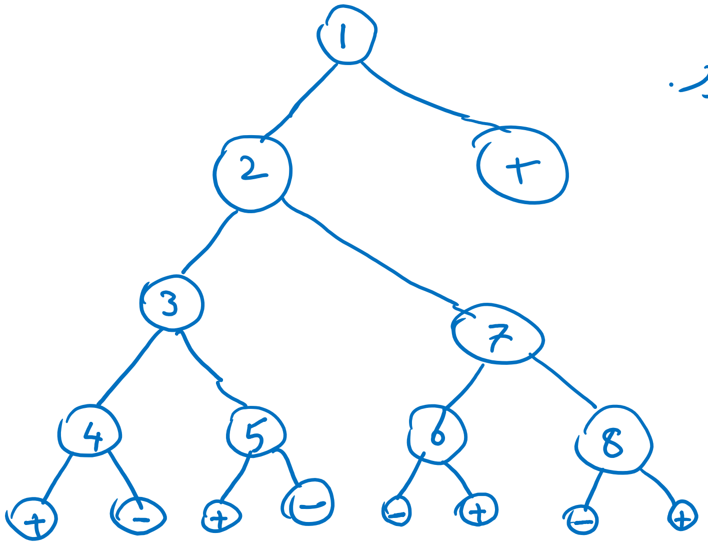
الف) درخت underfit : +

در واقع درخت underfit یک برگ با برچسب + خواهد بود.



درخت overfit :

مطابق شکل روبرو، این درخت ۹ برگ خواهد داشت و 8 رأس داخلی دارد.



ب) تمام نمونه هایی که در نیمه سمت چپ قرار دارند، اشتباه دسته بندی می شوند. بنابراین در مجموع 8 مورد misclassification داریم.

روش leave-one-out به این صورت است که در هر مرتبه یک نمونه را به عنوان تست و سایر نمونه ها را به عنوان آموزش در نظر می گیریم و این کار را به تعداد نمونه های داده شده باید انجام داد. بنابراین، در این مسأله 16 بار باید این کار را انجام دهیم. زمانی که داده های نیمه سمت راست به عنوان تست قرار می گیرند، خطا نداریم ولی برای داده هایی که در نیمه سمت چپ قرار دارند به ازای تک تکشان خطا داریم.

ج) واضح است که نمونه های مثبت همواره در اکثریت قرار دارند و به درستی طبقه بندی می شوند ولی نمونه های منفی همگی اشتباه دسته بندی می شوند.

جواب: 4

سوال ۵

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2$$

subject to: $\begin{cases} y_i - \omega^T x_i - b \leq \epsilon \\ \omega^T x_i + b - y_i \leq \epsilon \end{cases}$

الف)

$$L(\omega, b, \lambda, \mu) = \frac{1}{2}\|\omega\|^2 + \sum_{i=1}^{n} \lambda_i \left(y_i - \omega^T x_i - b - \epsilon \right) + \sum_{i=1}^{n} \mu_i \left(\omega^T x_i + b - y_i - \epsilon \right)$$

ب)

$$\nabla_\omega L = \frac{\partial L}{\partial \omega} = \omega + \sum_{i=1}^{n} -\lambda_i x_i + \sum_{i=1}^{n} \mu_i x_i = 0 \Rightarrow \boxed{\omega = \sum_{i=1}^{n} (\lambda_i - \mu_i) x_i} \quad (1)$$

$$\nabla_b L = \frac{\partial L}{\partial \omega} = \sum_{i=1}^{n} -\lambda_i + \sum_{i=1}^{n} \mu_i = 0 \Rightarrow \boxed{\sum_{i=1}^{n} (\lambda_i - \mu_i) = 0} \quad (2)$$

حال باید روابط ① و ② را در لاگرانژین جایگذاری کنیم :

$$g(\lambda,\mu) = \frac{1}{2}\left(\sum_{i=1}^{n}(\lambda_i-\mu_i)x_i\right)^T\left(\sum_{i=1}^{n}(\lambda_i-\mu_i)x_i\right) + \sum_{i=1}^{n}(\lambda_i-\mu_i)y_i + b\overbrace{\sum_{i=1}^{n}(\mu_i-\lambda_i)}^{0}$$

$$-\epsilon\sum_{i=1}^{n}(\lambda_i+\mu_i) - \sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{n}(\lambda_j-\mu_j)x_j\right)^T x_i + \sum_{i=1}^{n}\mu_i\left(\sum_{j=1}^{n}(\lambda_j-\mu_j)x_j\right)^T x_i$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i-\mu_i)(\lambda_j-\mu_j)x_i^T x_j + \sum_{i=1}^{n}(\lambda_i-\mu_i)y_i - \epsilon\sum_{i=1}^{n}(\lambda_i+\mu_i)$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i-\mu_i)(\lambda_j-\mu_j)x_i^T x_j$$

$$= \frac{-1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i-\mu_i)(\lambda_j-\mu_j)x_i^T x_j + \sum_{i=1}^{n}(\lambda_i-\mu_i)y_i - \epsilon\sum_{i=1}^{n}(\lambda_i+\mu_i)$$

Dual Problem:

$$\max_{\lambda, \mu} \; g(\lambda, \mu)$$

s.t. $$\sum_{i=1}^{n} (\lambda_i - \mu_i) = 0$$

$$g(\lambda, \mu) = \frac{-1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\lambda_i - \mu_i)(\lambda_j - \mu_j) x_i^T x_j + \sum_{i=1}^{n} (\lambda_i - \mu_i) y_i - \epsilon \sum_{i=1}^{n} (\lambda_i + \mu_i)$$