
Machine Learning, 2021 Fall

Assignment 3

Notice

Due 23:59 (CST), Nov 9, 2021

Plagiarizer will get 0 points.

L^AT_EX is highly recommended. Otherwise you should write as legibly as possible.

1 L1-SVM

In the lecture, we discussed ℓ_1 -regularized SVM problem:

$$\text{minimize} \quad \|\mathbf{w}\|_1 + C \sum_{i=1}^n \max \{1 - y^i (\mathbf{w}^T \mathbf{x}^i + b), 0\}, \quad (1)$$

In this question you have to follow lecture 12 and derive the dual form of ℓ_1 -regularized SVM. Notice that for a real number $p \geq 1$, the p -norm of x is defined by $\|x\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{1/p}$

Hint: You may first derive the equivalent linear programming formulation of 1

Ans

First derive the equivalent linear programming formulation of (1)

$$\min_{\substack{\xi \in \mathbb{R}^n, b \in \mathbb{R} \\ \mathbf{w}^+, \mathbf{w}^- \in \mathbb{R}^p}} C \sum_{i=1}^n \xi_i + \sum_{j=1}^p w_j^+ + \sum_{j=1}^p w_j^- \quad (2)$$

$$\begin{aligned} \text{s.t.} \quad & \xi_i + y_i \mathbf{x}_i^T \mathbf{w}^+ - y_i \mathbf{x}_i^T \mathbf{w}^- + y_i b \geq 1 \quad i \in [n] \\ & \xi \geq 0, \mathbf{w}^+ \geq 0, \mathbf{w}^- \geq 0 \end{aligned} \quad (3)$$

The positive and negative parts of \mathbf{w} are denoted as $w_i^+ = \max\{w_i, 0\}$ and $w_i^- = \max\{-w_i, 0\}$

A dual of (2) is the following LP:

$$\begin{aligned} \max_{\boldsymbol{\alpha} \in \mathbb{R}^n} \quad & \sum_{i=1}^n \alpha_i \\ \text{s.t.} \quad & -1 \leq \sum_{i=1}^n y_i x_{ij} \alpha_i \leq 1 \quad j \in [p] \\ & \mathbf{y}^T \boldsymbol{\alpha} = 0 \\ & 0 \leq \alpha_i \leq C \quad i \in [n] \end{aligned} \quad (4)$$

2 Decision Tree

In this problem you are asked to induce a decision tree from the training data set and evaluate your decision tree on the test set. Table 2 provides the information pieces whether TAs sleep late and attributes may be related to it.

training set			
season	stay up late	w/ CS282	sleep late
summer	False	False	False
summer	False	False	False
summer	True	True	True
summer	False	True	True
winter	True	True	False
test set			
season	stay up late	w/ CS282	sleep late
winter	False	True	True
summer	True	False	False

1. Based on information gain, determine which attribute is the root of the decision tree and explain your reason.

$$H(\text{sleep late} \mid \text{season}) = 0.8$$

$$H(\text{sleep late} \mid \text{stay up late}) = 0.95$$

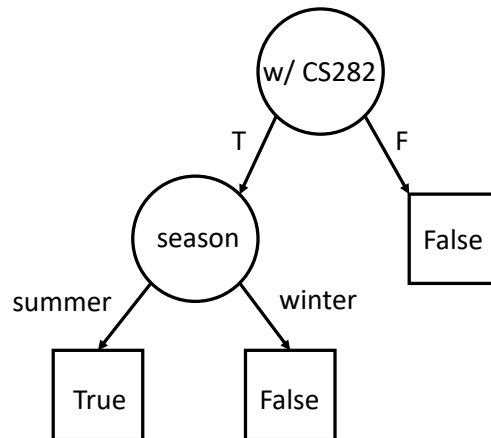
$$H(\text{sleep late} \mid \text{w/ CS282}) = 0.55$$

w/ CS282 will be the root.

2. According to your result above, what's your accuracy on test set if we only have the root node and leaf with majority label.

1.

3. Complete your decision tree and what's your accuracy on test set now.



0.5