Exercise 4

Context:

A dataset of pairs consisting of $\mathbb{T} : (\mathbb{R}^+ \times L \times \mathbb{R})^* \times \{-1, 1\} \to \mathbb{N}$ where L is a finite set of labels for every sample $s = ((vt_1, l_1, r_1) \ldots (vt_n, l_n, r_n), y)$ with $\mathbb{T}(s) > 0$ $vt_1 = 0$ and for every $1 \leq i < n$, $vt_i \leq vt_{i+1}$.

Definition 1. A tree-stump Ψ is a *Sequence-tree stump* $T = (V = V_l \cup V_v \cup V_c, E = E_\top \cup E_\perp, \mathcal{V}_l, \mathcal{V}_v, \mathcal{V}_c)$ where $|\mathcal{V}_e| = 1$. Given a triple $x = (vt, s_x)$ where $s_x \in (\mathbb{R}^+ \times L \times \mathbb{R})^*$ and $vt \in \mathbb{R}^{\geq 0}$ the application of Ψ to x is the following, written $\Psi(x)$, is:

$$\Psi(s_x) = \begin{cases} (s_x[i].vt, s_x[i+1:]) & \text{if there exists } i \text{ such that } i = \min\limits_{\substack{vt - s_x[i].vt \leq \Psi.duration, \\ \Psi.label = s_x[i].l}} i \\ s_x & \text{otherwise} \end{cases}$$

Assignment:

Train a boosting classifier $((\alpha_1, \Psi_1), \ldots, (\alpha_m, \Psi_m))$ where:

- $X_1 = (\overrightarrow{0}; X)$
- $Z_1 = (X_1; Y)$, $W_1[j] = \frac{1}{|X|}$ for each $1 \leq j \leq |X|$;
- Ψ_i is the best sequence-tree stump Ψ_i, i.e., the one that minimizes the weighted error ϵ_i, for the weighted sample set $(W_i, (X_i; Y))$ (more details in the appendix);
- and $X_i = \Psi_{i-1}(X_{i-1})$ for each $i > 1$;
- $$\alpha_i = \frac{1}{2} \ln \left(\frac{1 - \epsilon_i}{\epsilon_i} \right);$$
- $$W_i[j] = \begin{cases} \dfrac{W_{i-1}[j]}{2 \sum\limits_{j', Y[j'] * \Psi_i.predict(X[j']) \geq 0} W_{i-1}[j']} & \text{if } Y[j] * \Psi_{i-1}.predict(X[j]) \geq 0 \\ \dfrac{W_{i-1}[j]}{2 \sum\limits_{j', Y[j'] * \Psi_i.predict(X[j']) \leq 0} W_{i-1}[j']} & \text{otherwise.} \end{cases}$$

  for each $i > 1$.

Notes:

Do not mistake $\Psi(x)$ with $\Psi.predict(x)$, the first returns the sequence consumed according to Definition 1, the latter returns either -1 or 1 that is class assigned by Ψ to the sample x.

Appendix: Compute the best Sequence-tree stump

Algorithm 1 Best-tree

Input: $W \in (\mathbb{R}^+)^n, VT \in (\mathbb{R}^+)^n, X \in (\mathbb{R}^+ \times L \times \mathbb{R})^*)^n, Y \in \{-1, 1\}^n$
$CandidatePairs \leftarrow \{\}$
for $j \in \{1, \ldots, n\}$ **do**
  let $s_x = X[j]$
  let $vt = VT[j]$
  for $i \in \{1, \ldots, |s_x|\}$ **do**
    $CandidatePairs \leftarrow CandidatePairs \cup \{(s_x[i].l, s_x[i].vt - vt) : 0 \leq i < |s_x|\}$
  end for
end for
return $\underset{\Psi \in \{TreePair(W,VT,X,Y,l,vt):(l,vt) \in CandidatePairs\}}{\arg\max} W^T \cdot (\Psi.predict(X) - Y)$

Algorithm 2 TreePair

Input: $W \in (\mathbb{R}^+)^n, VT \in (\mathbb{R}^+)^n, X \in (\mathbb{R}^+ \times L \times \mathbb{R})^*)^n, Y \in \{-1,1\}^n, l \in L, d \in \mathbb{R}^+$

$Tree \leftarrow EventNode(l, d)$

$I_\top \leftarrow \{j : 0 \leq j < n, \exists i(X[j][i] = l, X[j][i] - VT[j] \leq d)\}$

$I_\perp \leftarrow \{j : 0 \leq j < n\} \setminus I_\top$

$Tree.falseChild \leftarrow Leaf\left(\underset{y \in \{1,-1\}}{\arg\max} \sum_{\substack{j \in I_\perp,\\ Y[j]=y}} w_j\right)$

$P_\top \leftarrow \{(j, X[i].v) : j \in I_\top, i = \min_{X[j][i].l=l} i\}$

$values \leftarrow \{-\infty\} + \{v : (i, v) \in P_\top\}.to\_list$

$values.sort()$

$Tree.trueChild \leftarrow ValueNode()$

$cnode \leftarrow Tree.trueChild$

$i \leftarrow 1$

do

    $cnode.value \leftarrow values[i]$

    $cnode.trueChild \leftarrow Leaf\left(\underset{y \in \{1,-1\}}{\arg\max} \sum_{\substack{(j,v) \in P_\top, Y[j]=y,\\ values[i-1] < v \leq values[i]}} w_j\right)$

    if $i < |values| - 2$ **then**

        $cnode.falseChild \leftarrow ValueNode()$

        $cnode \leftarrow cnode.falseChild$

    else

        $cnode.falseChild \leftarrow Leaf\left(\underset{y \in \{1,-1\}}{\arg\max} \sum_{\substack{(j,v) \in P_\top, Y[j]=y,\\ values[i] < v \leq values[i+1]}} w_j\right)$

    end if

    $i \leftarrow i + 1$

while $i < |values| - 1$

return $Tree$