Explication de l'algorithme pour la méthode EC

February 23, 2024

Le but de cette algorithme et de calculer T_{EC} pour que la confiance moyenne corresponde à la précision moyenne sur l'ensemble de validation. Pour ce faire on dipose d'un ensemble de validation $\{(x_i, y_i)\}$ pour $i = 1, \ldots, n_{\mathrm{val}}$, et d'un classifieur $\hat{f} : X \to \mathbb{R}^K$ où les $x_i \in \mathbb{R}^p$ sont les caractéristqies (les variables) et les $y_i \in [\![1, K]\!]$ sont les étiquettes de classes associées à ces caractéristqies. L'algorithme calcule les logits $z_i = f(x_i) \in \mathbb{R}^K$ et produit la sortie $\hat{y}_i = \arg\max_k z_{ik}$. Le classifieur prend en entrée des données (ici les caractéristiques x_i extraites d'une observation) et y attribue des logits $z_i = (z_{i1}, ..., z_{iK})$ où pour touts k appartenant à $[\![1, K]\!]$ chaque $z_{ik} \in \mathbb{R}$, pour un k fixé z_{ik} correspond au logit (score) associé à la classe k. Le fait de produire la sortie $\hat{y}_i = \arg\max_k z_{ik}$ signifie que la classe correspondant au logit le plus élevé est choisie comme la prédiction, pour la classe d'appartenance de x_i l'algorithme choisie de prédire que la classe associée à x_i est $\hat{y}_i$.

Ensuite l'algorithme calcul la précision moyenne sur l'ensemble de validation $A_{val} = \frac{1}{n_{\mathrm{val}}} \sum_i \delta(y_i = \hat{y}_i)$

Les logits sont des valeurs brutes, résultant de la dernière couche d'un réseau de neurones avant l'application d'une fonction d'activation. Ces valeurs brutes ne sont pas normalisées et peuvent être n'importe quel nombre réel.

Cependant, avant d'obtenir les probabilités associées à chaque classe, les logits passent généralement par une fonction d'activation softmax. La fonction softmax transforme les logits en probabilités, en produisant une distribution de probabilité sur les classes. Les valeurs résultantes après la fonction softmax seront dans l'intervalle [0,1] et leur somme sera égale à 1. La fonction softmax va transformer le vecteur z_i en un vecteur $z'_i = (\sigma(1)(z_{i1}), ..., \sigma(K)(z_{iK}))$ où pour tous k appartenant à $[\![1, K]\!]$:

$$\sigma(k)(z_{ik}) = \frac{e^{z_{ik}}}{\sum_{j=1}^{K} e^{z_{ij}}}$$

On a que $\sigma(k)(z_{ik})$ est la probablité que x_i appartienne à la classe k, pour prédire qu'elle est la classe associée a x_i on va donc regarder $\hat{y}_i = \max_k \sigma(k)(z_{ik})$ pour k appartenant à $[\![1, K]\!]$, de plus on a bien :

$$\sum_{k=1}^{K} \sigma(k)(z_{ik}) = \frac{\sum_{k=1}^{K} e^{z_{ik}}}{\sum_{j=1}^{K} e^{z_{ij}}} = 1$$

Pour trouver T_{EC} on va enfaite prendre T_{EC} telle que :

$$\frac{1}{n_{\text{val}}} \sum_{i} \max_{k} \sigma(k) \left(\frac{z_{ik}}{T_{\text{EC}}} \right) = A_{\text{val}}$$

De cette manière, T_{EC} permet à ce que la probabilité maximale d'appartenir à une classe aprés l'application de la fonction softmax soit en accord avec la précision moyenne sur l'ensemble de validation.