AI theory HW Week1

Lim Jing Yun

May 29, 2019

Vanilla Accuracy

$$\frac{1}{n}\sum_{i=1}^{n} 1[f(x_i) == y_i]$$

Class-wise Averaged Accuracy

$$A = \frac{1}{C}\sum_{c=1}^{C} \frac{1}{\sum_{i=1}^{n} 1[y_i == c]} \sum_{i=1}^{n} 1[y_i == c] 1[f(x_i) == c]$$

labels	frequency
c_1	$p_1 \cdot n$
c_2	$p_2 \cdot n$
c_3	$p_3 \cdot n$

Predict constantly the most frequently class

$f(x_i) = \text{argmax}_c p_c = c\uparrow$
$\max_c p_c = p\uparrow$

$$\begin{aligned}
\text{Vanilla Accuracy} &= \frac{1}{n}\sum_{i=1}^{n} 1[f(x_i) == y_i] \\
&= \frac{1}{n}\left(\sum_{i=1}^{p_1\cdot n} 1[c\uparrow == c_1] + \sum_{i=1}^{p_2\cdot n} 1[c\uparrow == c_2] + \sum_{i=1}^{p_3\cdot n} 1[c\uparrow == c_3]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{p\uparrow\cdot n} 1[c\uparrow == c\uparrow] + \sum_{i=p\uparrow\cdot n+1}^{n} 1[c\uparrow == c_i]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{p\uparrow\cdot n} 1 + \sum_{i=1}^{(1-p\uparrow)\cdot n} 0\right) \\
&= \frac{1}{n}\left(p\uparrow \cdot n\right) \\
&= p\uparrow
\end{aligned}$$

Let c_{rest} denotes values of class c that fulfill the condition $[c \neq c\uparrow]$

$$
\begin{aligned}
&\text{Class-wise Averaged Accuracy} \\
&= \frac{1}{C}\sum_{c=1}^{C}\frac{1}{\sum_{i=1}^{n} 1[y_i == c]}\sum_{i=1}^{n} 1[y_i == c]1[f(x_i) == c] \\
&= \frac{1}{3}\sum_{c=1}^{3}\frac{1}{\sum_{i=1}^{n} 1[y_i == c]}\sum_{i=1}^{n} 1[y_i == c]1[c\uparrow == c] \\
&= \frac{1}{3}\Bigg(\frac{1}{\sum_{i=1}^{n} 1[y_i == c_1]}\sum_{i=1}^{n} 1[y_i == c_1]1[c\uparrow == c_1] + \frac{1}{\sum_{i=1}^{n} 1[y_i == c_2]}\sum_{i=1}^{n} 1[y_i == c_2]1[c\uparrow == c_2] \\
&+ \frac{1}{\sum_{i=1}^{n} 1[y_i == c_3]}\sum_{i=1}^{n} 1[y_i == c_3]1[c\uparrow == c_3]\Bigg) \\
&= \frac{1}{3}\Bigg(\frac{1}{\sum_{i=1}^{p\uparrow\cdot n} 1[c\uparrow == c\uparrow]}\sum_{i=1}^{p\uparrow\cdot n} 1[c\uparrow == c\uparrow]1[c\uparrow == c\uparrow] \\
&+ \frac{1}{\sum_{i=p\uparrow\cdot n+1}^{n} 1[y_i == c_{rest}]}\sum_{i=p\uparrow\cdot n+1}^{n} 1[y_i == c_{rest}]1[c\uparrow == c_{rest}]\Bigg) \\
&= \frac{1}{3}\Bigg(\frac{1}{\sum_{i=1}^{p\uparrow\cdot n} 1}\sum_{i=1}^{p\uparrow\cdot n} 1\cdot 1 + \frac{1}{\sum_{i=1}^{(1-p\uparrow)\cdot n} 1}\sum_{i=1}^{(1-p\uparrow)\cdot n} 1\cdot 0\Bigg) \\
&= \frac{1}{3}\Bigg(\frac{1}{p\uparrow\cdot n}p\uparrow\cdot n + \frac{1}{(1-p\uparrow)\cdot n}(1-p\uparrow)\cdot n\cdot 1\cdot 0\Bigg) \\
&= \frac{1}{3}(1+0) \\
&= \frac{1}{3}
\end{aligned}
$$

Predict constantly the least frequent class

$f(x_i) = \text{argmin}_c p_c = c\downarrow$
$\min_c p_c = p\downarrow$

$$
\begin{aligned}
\text{Vanilla Accuracy} &= \frac{1}{n}\sum_{i=1}^{n} 1[f(x_i) == y_i] \\
&= \frac{1}{n}\left(\sum_{i=1}^{p\downarrow\cdot n} 1[c\downarrow == c\downarrow] + \sum_{i=p\downarrow\cdot n+1}^{n} 1[c\downarrow == c_i]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{p\downarrow\cdot n} 1 + \sum_{i=1}^{(1-p\downarrow)\cdot n} 0\right) \\
&= \frac{1}{n}\left(p\downarrow\cdot n\right) \\
&= p\downarrow
\end{aligned}
$$

Let c_{rest} denotes values of class c that fulfill the condition $[c \neq c\downarrow]$

Class-wise Averaged Accuracy

$$
\begin{aligned}
&= \frac{1}{3}\sum_{c=1}^{3}\frac{1}{\sum_{i=1}^{n}1[y_i == c]}\sum_{i=1}^{n}1[y_i == c]1[c\downarrow== c] \\
&= \frac{1}{3}\bigg(\frac{1}{\sum_{i=1}^{p\downarrow\cdot n}1[c\downarrow== c\downarrow]}\sum_{i=1}^{p\downarrow\cdot n}1[c\downarrow== c\downarrow]1[c\downarrow== c\downarrow] + \frac{1}{\sum_{i=p\downarrow\cdot n+1}^{n}1[y_i == c_{rest}]}\sum_{i=p\downarrow\cdot n+1}^{n}1[y_i == c_{rest}]1[c\downarrow== c_{rest}]\bigg) \\
&= \frac{1}{3}\bigg(\frac{1}{\sum_{i=1}^{p\downarrow\cdot n}1}\sum_{i=1}^{p\downarrow\cdot n}1\cdot 1 + \frac{1}{\sum_{i=1}^{(1-p\downarrow)\cdot n}1}\sum_{i=1}^{(1-p\downarrow)\cdot n}1\cdot 0\bigg) \\
&= \frac{1}{3}\bigg(\frac{1}{p\downarrow\cdot n}p\downarrow\cdot n + \frac{1}{(1-p\downarrow)\cdot n}(1-p\downarrow)\cdot n\cdot 1\cdot 0\bigg) \\
&= \frac{1}{3}(1+0) \\
&= \frac{1}{3}
\end{aligned}
$$

Expected values of the two accuracies when we predict each sample x with the class c with probability q_c

$$P(f(x_i) = c_1) = q_1 \qquad P(f(x_i) = c_2) = q_2 \qquad P(f(x_i) = c_3) = q_3$$

$$q_1 + q_2 + q_3 = 1$$

Given $f(x_i) = c_j$ where $j \in$ [1,2,3]

$$
\begin{aligned}
\text{Vanilla Accuracy given } [f(x_i) = c_j] &= \frac{1}{n}\sum_{i=1}^{n}1[f(x_i) == y_i] \\
&= \frac{1}{n}\left(\sum_{i=1}^{p_j\cdot n}1[c_j == y_i] + \sum_{i=p_j\cdot n+1}^{n}1[c_j == y_i]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{p_j\cdot n}1[c_j == c_j] + \sum_{i=p_j\cdot n+1}^{n}1[c_j == y_i]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{p_j\cdot n}1 + \sum_{i=1}^{(1-p_j)\cdot n}0\right) \\
&= \frac{1}{n}(p_j\cdot n) \\
&= p_j
\end{aligned}
$$

Let c_{rest} denotes values of class c that fulfill the condition $[c \neq c_j]$

$$\begin{aligned}
&\text{Class-wise Averaged Accuracy given } [f(x_i) = c_j] \\
&= \frac{1}{C} \sum_{c=1}^{C} \frac{1}{\sum_{i=1}^{n} 1[y_i == c]} \sum_{i=1}^{n} 1[y_i == c] 1[f(x_i) == c] \\
&= \frac{1}{3} \sum_{c=1}^{3} \frac{1}{\sum_{i=1}^{n} 1[y_i == c]} \sum_{i=1}^{n} 1[y_i == c] 1[c_j == c] \\
&= \frac{1}{3} \Bigg(\frac{1}{\sum_{i=1}^{p_j \cdot n} 1[y_i == c_j]} \sum_{i=1}^{p_j \cdot n} 1[y_i == c_j] 1[c_j == c_j] \\
&+ \frac{1}{\sum_{i=p_j \cdot n+1}^{n} 1[y_i == c_{rest}]} \sum_{i=p_j \cdot n+1}^{n} 1[y_i == c_{rest}] 1[c_j == c_{rest}] \Bigg) \\
&= \frac{1}{3} \Bigg(\frac{1}{\sum_{i=1}^{p_j \cdot n} 1[c_j == c_j]} \sum_{i=1}^{p_j \cdot n} 1[c_j == c_j] 1[c_j == c_j] \\
&+ \frac{1}{\sum_{i=p_j \cdot n+1}^{n} 1[y_i == c_{rest}]} \sum_{i=p_j \cdot n+1}^{n} 1[y_i == c_{rest}] 1[c_j == c_{rest}] \Bigg) \\
&= \frac{1}{3} \Bigg(\frac{1}{\sum_{i=1}^{p_j \cdot n} 1} \sum_{i=1}^{p_j \cdot n} 1 \cdot 1 + \frac{1}{\sum_{i=1}^{(1-p_j) \cdot n} 1} \sum_{i=1}^{(1-p_j) \cdot n} 1 \cdot 0 \Bigg) \\
&= \frac{1}{3} \Bigg(\frac{1}{p_j \cdot n} p_j \cdot n + \frac{1}{(1-p_j) \cdot n} \cdot 0 \Bigg) \\
&= \frac{1}{3}(1+0) \\
&= \frac{1}{3}
\end{aligned}$$

$$\begin{aligned}
E(\text{Vanilla Accuracy}) &= P(f(x_i) = c_1) \cdot \text{ Vanilla Accuracy given } [f(x_i) = c_1] \\
&+ P(f(x_i) = c_2) \cdot \text{ Vanilla Accuracy given } [f(x_i) = c_2] \\
&+ P(f(x_i) = c_3) \cdot \text{ Vanilla Accuracy given } [f(x_i) = c_3] \\
&= q_1 \cdot p_1 + q_2 \cdot p_2 + q_3 \cdot p_3
\end{aligned}$$

$$\begin{aligned}
E(\text{Class-wise Average Accuracy}) &= P(f(x_i) = c_1) \cdot \text{ Class-wise Average Accuracy given } [f(x_i) = c_1] \\
&+ P(f(x_i) = c_2) \cdot \text{ Class-wise Average Accuracy given } [f(x_i) = c_2] \\
&+ P(f(x_i) = c_3) \cdot \text{ Class-wise Average Accuracy given } [f(x_i) = c_3] \\
&= q_1 \cdot \frac{1}{3} + q_2 \cdot \frac{1}{3} + q_3 \cdot \frac{1}{3} \\
&= (q_1 + q_2 + q_3) \cdot \frac{1}{3} \\
&= \frac{1}{3}
\end{aligned}$$

Expected values of the two accuracies when we predict each sample x with the class c with probability p_c (i.e. $q_c = p_c$)

When $q_c = p_c$

$$\begin{aligned} E(\text{Vanilla Accuracy}) &= q_1 \cdot p_1 + q_2 \cdot p_2 + q_3 \cdot p_3 \\ &= p_1 \cdot p_1 + p_2 \cdot p_2 + p_3 \cdot p_3 \\ &= (p_1)^2 + (p_2)^2 + (p_3)^2 \end{aligned}$$

$$\begin{aligned} E(\text{Class-wise averaged accuracy}) &= (q_1 + q_2 + q_3) \cdot \frac{1}{3} \\ &= (p_1 + p_2 + p_3) \cdot \frac{1}{3} \\ &= \frac{1}{3} \end{aligned}$$

Suppose you predict constantly the most frequent class $f(x_i) = \text{argmax}_c p_c$. Find the set of test set frequencies p_1; p_2; p_3 such that (A) the vanilla accuracy will be equal to the class-wise averaged accu racy.

When $f(x_i) = \text{argmax}_c p_c$: Vanilla Accuracy $= \max_c p_c$ and Class-wise averaged accuracy $= \frac{1}{3}$

for vanilla accuracy = class-wise averaged accuracy,

$$\max_c p_c = \frac{1}{3}$$

$$\max(p_1, p_2, p_3) = \frac{1}{3}$$

Let $\max_c p_c = p_1$,for the above equation to be fulfilled, $p_1 = \frac{1}{3}$. Since $\sum_{i=1}^{3} p_i = 1$,

$$p_1 + p_2 + p_3 = 1$$

$$\frac{1}{3} + p_2 + p_3 = 1$$

$$p_2 + p_3 = \frac{2}{3} \tag{1}$$

For $\max_c p_c = p_1 = \frac{1}{3}$,

$$p_2 \leq \frac{1}{3} \text{and } p_3 \leq \frac{1}{3} \tag{2}$$

Satisfying equations 1 and 2 yields,

$$p_2 = p_3 = \frac{1}{3}$$

Therefore the test set required is ,

$$p_1 = p_2 = p_3 = \frac{1}{3}$$

Suppose you predict constantly the most frequent class $f(x_i) = \text{argmax}_c p_c$. Find the set of test set frequencies p_1; p_2; p_3 such that(B) the vanilla accuracy will be much higher than the class-wise averaged accuracy (for example, provide a p_1; p_2; p_3, such that the vanilla accuracy is 0.95, but the class-wise accuracy is $\frac{1}{3}$ only).

When $f(x_i) = \text{argmax}_c p_c$: Vanilla Accuracy $= \max_c p_c$ and Class-wise averaged accuracy $= \frac{1}{3}$

$$\max(p_1, p_2, p_3) >> \frac{1}{3}$$

From our calculations above, we can see that with $f(x_i) = \text{argmax}_c p_c$ the class-wise accuracy is a constant independent of p_1, p_2 and p_3.
Therefore to achieve the example stated in the question,

$$p_1 = 0.95$$

$$p_2 = 0.025$$

$$p_3 = 0.025$$