CSE440

Friday, July 12, 2024 4:14 PM

Adam Optimizer

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (\beta_1 \to 0.9)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (\beta_2 \to 0.999)$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t$$

$\epsilon \to \approx 0$ [Extremely small]

BoW, Naive Bayes

Training:-

	call	can	have	I	later	me	now	sorry	U	will	won	!
+ Sorry I'll call later	1	0	0	1	1	0	0	1	0	1	0	0
+ U can call me now	1	1	0	0	0	1	1	0	1	0	0	0
− U have won call now!!	1	0	1	0	0	0	1	0	1	0	1	2

Testing

Sorry! U can not unsubscribe	0	1	0	0	0	0	0	1	1	0	0	1
	✓	✓	✓		✓							

$$\therefore P(+ \mid \text{Sonny! U can}) = P(\text{Sonny} \mid +) \cdot P(! \mid +) \cdot P(U \mid +) \cdot P(\text{can} \mid +) \cdot P(+)$$

$$= \frac{1+1}{10+12} \cdot \frac{0+1}{10+12} \cdot \frac{1+1}{10+12} \cdot \frac{1+1}{10+12} \cdot \frac{2}{3} = 0.000023$$

$$\therefore P(- \mid \text{sonny! U can}) = P(\text{sonny} \mid -) \cdot P(! \mid -) \cdot P(U \mid -) \cdot P(\text{can} \mid -) \cdot P(-)$$

$$= \frac{0+1}{7+12} \cdot \frac{2+1}{7+12} \cdot \frac{1+1}{7+12} \cdot \frac{0+1}{7+12} \cdot \frac{1}{3} = 0.000015$$

Cross Entropy loss :-

$y \rightarrow$ Actual

$\hat{y} \rightarrow$ Predicted

$$\text{Gain} = \hat{y}^{y} \cdot (1-\hat{y})^{(1-y)}$$

$$\text{Loss} = -\log_e \left\{ \hat{y}^{y} \cdot (1-\hat{y})^{(1-y)} \right\}$$

Performance Metrics

TP	FP
FN	TN

✓ Precision :- (No Failure Case)

$$P_+ = \frac{TP}{TP + FP} \qquad P_- = \frac{TN}{TN + FN}$$

✓ Recall :- (Medical)

$$P_+ = \frac{TP}{TP + FN} \qquad P_- = \frac{TN}{FP + TN}$$

✓ F-1 score :- (Maximizing Precision & Recall)

$$\therefore f\text{-}1 = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2}{\frac{R+P}{PR}} = \frac{2PR}{P+R}$$

Term - Term Co-occurance Matrix:-

["any big cat", "big cat", "cat dog cat"]

	any	big	cat	dog
any	1	1	1	0
big	1	2	2	0
cat	1	2	3	1
dog	0	0	1	1

	any	big	cat	dog
	1	1	1	0
	0	1	1	0
	0	0	1	1

BoW = W

$\Rightarrow W^T \cdot W$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

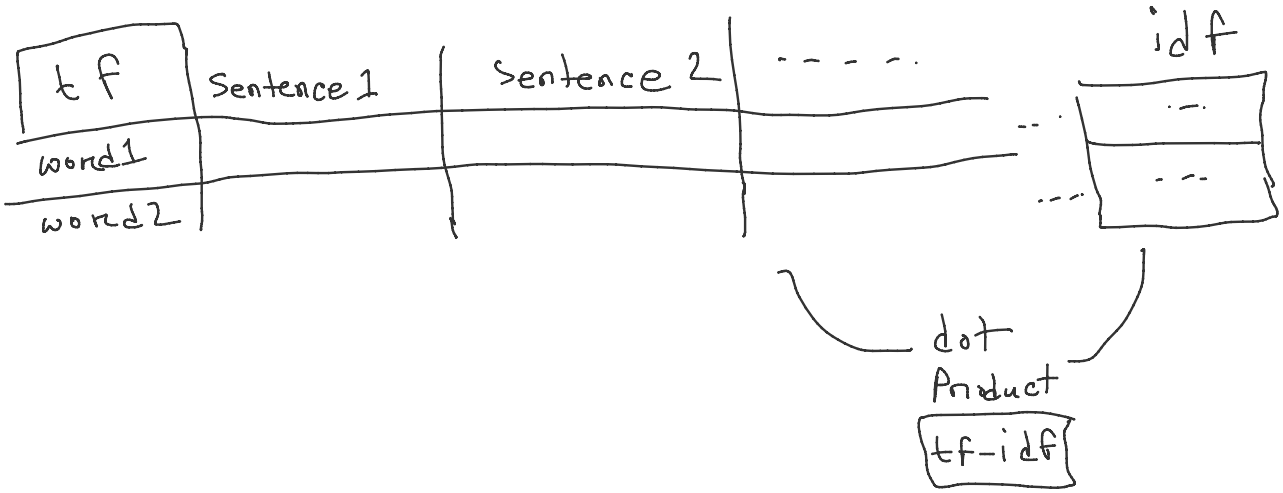
$$= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 2 & 0 \\ 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

tf-idf (term frequency - inverse document frequency):-

$$tf = \log_{10}(1 + f(w, d))$$

$$idf = \log_{10}\left(\frac{N}{f(w, D)}\right)$$

tf x idf = vector values



Positive Pointwise Mutual Information (PPMI):-

$$\max\left(\log_2\left(\frac{P(w,c)}{P(w)\cdot P(c)}\right), 0\right) = PPMI(w,c)$$