

## 『IT Text 自然言語処理の基礎』 第1版第2刷 正誤表

ページ	行, 式番号等	誤	正
39	下から 10 行目	すなわち, すべてのクラスに関して分類される	すなわち, すべてのクラス <small>に</small> に関して分類される
70	8 行目	単語 $x_i$ と文脈単語 $c_j$ が共起する確率 $P(x_i, c_i)$	単語 $x_i$ と文脈単語 $c_j$ が共起する確率 $P(x_i, c_j)$
	式(3.6)	$\text{PMI}(x_i, c_j) = \log \frac{P(x_i, c_i)}{P(x_i)P(c_j)} = \log \frac{Z \cdot \#(x_i, c_j)}{\#(x_i)\#(c_j)}$	$\text{PMI}(x_i, c_j) = \log \frac{P(x_i, c_j)}{P(x_i)P(c_j)} = \log \frac{Z \cdot \#(x_i, c_j)}{\#(x_i)\#(c_j)}$
	下から 4 行目	式 (3.6) では, $P(x_i, c_i) < P(x_i)P(c_j)$ であるとき,	式 (3.6) では, $P(x_i, c_j) < P(x_i)P(c_j)$ であるとき,
79	式(3.38)	$w_{\text{France}} - w_{\text{Paris}} + w_{\text{Italy}}$	$w_{\text{Paris}} - w_{\text{France}} + w_{\text{Italy}}$
95	式(4.23)	$\mathbf{h}_t = \mathbf{o}_t \odot \tanh \mathbf{c}_t$	$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$
175	式(7.7)右辺	$\text{softmax}_{x_i}(\mathbf{W}_{xh} \mathbf{H}_i^{(L)})$	$\text{softmax}_{x_i}(\mathbf{W}^T \mathbf{H}_i^{(L)})$
176	1~2 行目	ここで, $\mathbf{W}_{xh} \in \mathbb{R}^{ \mathcal{V}  \times d}$ はマスクされた単語を予測するための全結合層のパラメータ, $\mathbf{H}_i^{(L)} \in \mathbb{R}^d$ は…	ここで, $\mathbf{W} \in \mathbb{R}^{d \times  \mathcal{V} }$ は単語埋込み行列, $\mathbf{H}_i^{(L)} \in \mathbb{R}^d$ は…
181	12~13 行目	BART では 0 から 3 個の連続したサブワードを一つのマスクとする	BART では 30% の単語について, $\lambda = 3$ のポアソン分布により決定した長さの <span style="color:red">スパン</span> の連続したサブワードを一つのマスクとする
295	下から 12 行目	112) K. Tadao:	112) T. Kasami:
	下から 10 行目	113) Y. Matsumoto and T. Kudo:	113) T. Kudo and Y. Matsumoto: