Midterm Review 10/15/2015

Q3-5

$P(y=1 \mid \vec{x}=(1,0,1))$

"dog cat"
"cat dog"

$P(w_1, w_2 .. w_n \mid y=1) = \prod_{i=1}^{n} P(w_i \mid y=1)$

Token-level Likelihood

$= P(\text{dog} \mid y=1)\, P(\text{cat} \mid y=1)$

$= \frac{1}{Z} P(y=1)\, P(x=(1,0,1) \mid y=1)$

$= \frac{1}{Z} \frac{4}{10} P(w=\text{"word1"} \mid y=1)\, P(w=\text{"word3"} \mid y=1)$

$= \frac{1}{Z} \frac{4}{10} \frac{1}{10} \frac{7}{10}$

$= \frac{1}{Z} \frac{28}{1000}$

$P(y=0 \mid \vec{x})$

$= \frac{1}{Z} P(y=0)\, P(\text{"word1"} \mid y=0)\, P(\text{"word3"} \mid y=0)$

$= \frac{1}{Z} \frac{6}{10} \frac{5}{10} \frac{3}{10}$

$= \frac{1}{Z} \frac{90}{1000}$

$$P(y \mid \vec{x}) = \frac{P(y)\,P(x \mid y)}{\sum_{y' \in \{0,1\}} P(y')\,P(x \mid y')}$$

$$\vec{x} = (2, 0, 1)$$

dog cat

$$P(x \mid y) = P(dog \mid y)\,P(dog \mid y)\,P(cat \mid y)$$

$$= [P(dog \mid y)]^2\,P(cat \mid y)$$

$$P(x \mid y) = \prod_{v=1}^{V} [P(w = v \mid y)]^{x_v}$$

$$\log P(x \mid y) = \sum_v x_v \log P(w = v \mid y)$$

Word-Type-Level NB Likelihood

Q5

Tokens vs. Types

the dog dog was happy

Num Tokens = 5

Num Types = 4

"word" is ambig!

Types = Unique Words