Ty Limpasuvan

CS 155 Final

Multiple Choice Questions

A True

B B

C Left image: K_1 Center: K_2 Right: K_3

D 2

E C

F True

G B

H True

I True

J False

K B

L False

M True

N False

Naive Bayes

1

Grade	Year	P(Happy)
A	Senior	$2/3$
A	Freshman	1
C	Senior	$1/2$
C	Freshman	0

2 1

3 $x = P((\text{grade}, \text{year}) \rightarrow \text{happy})$

prob = random()

if prob $\leq x$ then happy

else not happy

Data Transformations

1 $\bar{w} = Aw$

2 $\text{argmin}_{Aw} \frac{\lambda}{2}|Aw|^2 + \sum_i (y_i - (Aw)^T(Ax))^2$

3 The transformation matrix is used.

Latent Markov Embedding

1 The value of $|U(s')-V(s)|_2^2$ is strictly $\equiv |X(s')-X(s)|_2^2$ because the vectors U, V, and X are equal. This causes the P(s) for the dual-point model to never be less than that of the single point model.

2 The vectors U, V, and X are equivalent.

Neural Net Backprop Gradient Derivation

1 $$\frac{\partial}{\partial w_{11}} L(y,f(x)) = \frac{\partial}{\partial w_{11}}(y-f(x))^2 = -2(y-f(x))\cdot\frac{\partial f(x)}{\partial w_{11}} = -2(y-f(x))\cdot\frac{\partial}{\partial w_{11}}\sigma\left(\sum_{i=1}^{2} u_i\left(\sigma\left(\sum_{j=1}^{2} w_{ji}x_j\right)\right)\right)$$

$$= -2(y-f(x))\cdot\sigma\left(\sum_{i=1}^{2} u_i\left(\sigma\left(\sum_{j=1}^{2} w_{ji}x_j\right)\right)\right)\left(1-\sigma\left(\sum_{i=1}^{2} u_i\left(\sigma\left(\sum_{j=1}^{2} w_{ji}x_j\right)\right)\right)\right)$$

2 $$f(0.1, 0.5) = \sigma\left(\sum_{i=1}^{2} u_i\left(\sigma\left(\sum_{j=1}^{2} w_{ji}x_j\right)\right)\right)$$

$$0.5\cdot\sigma(0.25\cdot 0.1 + 0.05\cdot 0.5) = 0.2562$$
$$+$$
$$-0.1\cdot\sigma(0.1\cdot 0.1 - 0.25\cdot 0.5) = -0.0471$$

$$= \sigma(0.2091) = 0.5521 = f(x)$$

Plugging in: $-2(0.75-0.5521)\cdot 0.5521\cdot(1-0.5521) = \boxed{-0.0979}$

3 The term $\sigma(s)$ causes the vanishing gradient problem, and having more layers causes the weights to decrease, making us reach a gradient of zero more quickly.