令和4

問1 ゲームを買うラベルが持つエントロピーを I_0 とすると、

(1-1) $I_0 = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = \log_2 5 - \frac{2}{5} - \frac{3}{5}\log_2 3$

$= 2.32 - 0.4 - 0.6 \times 1.58 = 0.972$ //

(1-2) 評判

良い： Yes 1コ No 1コ

普通： Yes 1コ No 1コ

悪い： Yes 0コ No 1コ

0.6×0.98

$\log 3 - \frac{2}{3}$

時間

有： Yes 2コ No 1コ

無： Yes 0コ No 2コ

お金

有： Yes 2コ No 2コ

無： Yes 0コ No 1コ

(1-3) 評判が「良い」「普通」「悪い」であるときのエントロピーを

I_{11}, I_{12}, I_{13} とすると、

$I_{11} = I_{12} = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

$I_{13} = -\frac{0}{1}\log\frac{0}{1} - \frac{1}{1}\log\frac{0}{1} = 0$

であるから、エントロピーの期待値 $E[I_1]$ は、

$E[I_1] = \frac{2}{5}\times 1 + \frac{2}{5}\times 1 + \frac{1}{5}\times 0 = 0.8$

同様に、時間，お金で分割したときのエントロピーの期待値を

$E[I_2], E[I_3]$ とすると、

$E[I_2] = \frac{3}{5}\times\left(-\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{3}\right) + \frac{2}{5}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$

$= \frac{3}{5}\times(1.58 - 0.667) = 0.548$

$E[I_3] = \frac{4}{5}\times\left(-\frac{2}{4}\log\frac{2}{4} - \frac{2}{4}\log\frac{2}{4}\right) + \frac{1}{5}\left(-\frac{0}{1}\log\frac{0}{1} - \frac{1}{1}\log\frac{1}{1}\right)$

$= 0.8$

各属性の情報利得は

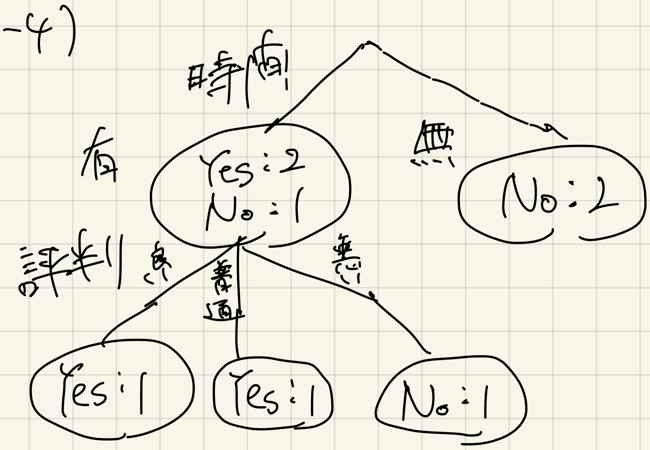
$I_0 - E[I_1] = 0.972 - 0.8 = 0.172$

$I_0 - E[I_2] = 0.972 - 0.548 = 0.424$

$I_0 - E[I_3] = 0.972 - 0.8 = 0.172$

となる。情報利得が最大となるのは、時間であり、この属性で分岐させるのがよい。

(1-4)



問2

(2-1) バックプロパゲーション

出力層にて与えられる教師信号と学習結果との誤差を、出力層側から入力層側へと伝播させていくことによって、損失関数の最小化を可能とするための技術のこと。　(73字)

(2-2)

学習データに過剰に適応し、汎化性能が低くなること。

学習データ数が少なすぎたり、学習データが対象の真の分布を反映していなかったりすることにより、適切に真のデータ分布を推定することができないことにより発生する。

(2-3)

信念分布という

環境の変化にマルコフ性を仮定し、直前の対話状態の分布、状態遷移確率、観測確率から、次の対話状態の分布が得られる。言語理解により、ドメイン、意図、スロット値が得られ、ここでは、スロット値が利用される。

(2-4)

- 物理的な定義：音響信号の信振が一定以下である区間の長さが一定以上あるとき、それを発話の切れ目とみなす。
- 対話行〜　：意味的なまとまりを一発話とみなす。
- 話者交代〜　：相手の発話に区切られた区間を一発話とみなす。

システム側によって「発話」を始めないとユーザの発話が終わらないため、話者交代〜は、発話終端の検出には使用できない。

(3-1)

$$J_t = |y_t - d_t|^2$$
$$= \{ (\sum_{k=0}^{N-1} h_k s_{t-k} + \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} m_{k,l} s_{t-k} s_{t-l}) - d_t \}^2$$

(3-2)

$$\frac{\partial J_t}{\partial h_k} = 2(y_t - d_t)(s_{t-k})$$
$$= 2 \{ (\sum_{k=0}^{N-1} h_k s_{t-k} + \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} m_{k,l} s_{t-k} s_{t-l}) - d_t \} s_{t-k}$$

$$\frac{\partial J_t}{\partial m_{k,l}} = 2 \{ (\sum_{k=0}^{N-1} h_k s_{t-k} + \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} m_{k,l} s_{t-k} s_{t-l}) - d_t \} s_{t-k} s_{t-l}.$$

$$\frac{\partial J_t}{\partial m_{k,l}} =$$

$$\frac{\partial}{\partial m_{k,l}} \sum_k (m_{k,0} s_{t-k} s_{t-0} + m_{k,1} s_{t-k} s_{t-1} + \cdots + m_{k,l} s_{t-k} s_{t-l} + \cdots + m_{k,N-1} s_{t-k} s_{t-(N-1)})$$
$$= \frac{\partial}{\partial m_{k,l}} \sum_k m_{k,l} s_{t-k} s_{t-l}$$
$$= \frac{\partial}{\partial m_{k,l}} (m_{0,l} s_{t-0} s_{t-l} + \cdots + m_{k,l} s_{t-k} s_{t-l} + \cdots + m_{N-1,l} s_{t-(N-1)} s_{t-l})$$
$$= s_{t-k} s_{t-l}.$$

(4-1)

$$P(z_2=s_1|x_1=a, x_2=b)$$

$$=\frac{P(x_2=b|z_2=s_1)P(z_2=s_1|x_1=a)}{P(x_2=b|z_2=s_1)P(z_2=s_1|x_1=a)+P(x_2=b|z_2=s_2)P(z_2=s_2|x_1=a)}$$

$$=\frac{0.8\times P(z_2=s_1|z_1=s_1)P(z_1=s_1|x_1=a)}{0.8\times P(z_2=s_1|z_1=s_1)P(z_1=s_1|x_1=a)+0.3P(z_2=s_2|z_1=s_1)P(z_1=s_1|x_1=a)}$$

$$=\frac{0.8\times 0.6\times P(x_1=a|z_1=s_1)P(z_1=s_1)/(P(x_1=a|z_1=s_1)P(z_1=s_1))}{(0.8\times 0.6+0.3\times 0.4)P(x_1=a|z_1=s_1)P(z_1=s_1)/(P(x_1=a|z_1=s_1)P(z_1=s_1))}$$

$$=\frac{0.42}{0.42+0.12}=\frac{42}{54}=\frac{7}{9}$$

(4-2)

記号列[a,b,b]を出力するのは$(z_1,z_2,z_3,z_4)=(s_1,s_1,s_2,s_3),(s_1,s_2,s_2,s_3)$の2通りのみである。よって$z_2=s_1$, $z_2=s_2$のときの2通りの確率を求め、その和によって求める解が得られる。

$$\begin{array}{r} 32 \\ \times\ 6 \\ \hline 192 \end{array}$$

・$z_2=s_1$のとき

$$0.8\times0.3\times0.4\times0.8\times0.6\times0.2\times1$$

・$z_2=s_2$のとき

$$0.8\times0.3\times0.2\times0.3\times0.4\times0.2\times1$$

$$\begin{array}{r} 0.216 \\ \times\ 0.48 \\ \hline 1728 \\ 864 \\ \hline 0.10368 \end{array}$$

$$0.8\times0.3\times0.4\times0.8\times0.6\times0.2\times1+0.8\times0.3\times0.2\times0.3\times0.4\times0.2\times1$$

$$=0.8\times0.3\times0.2(0.4\times0.8\times0.6+0.2\times0.3\times0.4)$$

$$=0.48\times(0.192+0.024)$$

$$=0.48\times0.216=0.10368 \ //$$

(4-3)

ある連続するN単語が与えられたとき、最もらしい次の1単語を推定するときに用いられる(N-gram)

このとき、xは、それまでに生成された単語列に、状態zは、直前のN単語列に対応する。

観測変数