CSCI 6364 PSET 4 Report

Weigeng Li G44406413

Problem1

First, we assume all variables X_i are independent.
Based on Naïve Bayes,

$$P(Y|X) = P(Y = y|X_1 = x_1, X_2 = x_2, X_3 = x_3 \dots X_n = x_n) = \frac{P(Y) * \prod_{i=1}^{n} P(X_i|Y)}{\prod_{i=1}^{n} P(X_i)}$$

For Naïve Bayes Classifier, the classification is Y with higher $P(Y|X)$

$$\arg\max P(Y = y|X) = \arg\max \frac{P(Y = y) * \prod_{i=1}^{n} P(X_i|Y = y)}{\prod_{i=1}^{n} P(X_i)}$$

$$= \arg\max \prod_{i=1}^{n} P(X_i|Y = y) * P(Y = y)$$

In this problem, the distribution of samples are:
Number of $Y = yes$ sample 9, Number of $Y = no$ sample 5,

$$P(Y = yes) = \frac{9}{14};\ P(Y = no) = \frac{5}{14}$$

Buy computer	Age=young	Income=low	Student=yes	Credit rating=fair
Yes	2	3	6	6
No	3	1	1	2

The probabilities are

	Age=young	Income=low	Student=yes	Credit rating=fair
$P(X_i|Y = yes)$	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$P(X_i|Y = no)$	$\frac{3}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

$$\prod_{i=1}^{n} P(X_i|Y = yes) * P(Y = yes) = \frac{2}{9} * \frac{1}{3} * \frac{2}{3} * \frac{2}{3} * \frac{9}{14} = \frac{4}{189} \approx 0.021$$

$$\prod_{i=1}^{n} P(X_i|Y = no) * P(Y = no) = \frac{3}{5} * \frac{3}{5} * \frac{1}{5} * \frac{2}{5} * \frac{5}{14} = \frac{9}{875} \approx 0.010$$

$$\prod_{i=1}^{n} P(X_i|Y = yes) * P(Y = yes) > \prod_{i=1}^{n} P(X_i|Y = no) * P(Y = no)$$

Therefore, the classification is yes.

Problem2

(a) Give that the model, and the first day is rainy (State 1)

$$P(\{S1, S2, S3, S3, S1\}) = P(S1|S1) * P(S2|S1) * P(S3|S2) * P(S3|S3) * P(S1|S3)$$
$$= 0.4 * 0.3 * 0.2 * 0.8 * 0.1 = 0.00192$$

(b)

Define matrix A is the transition probability matrix, p_{ij} is the probability.

$$p_{ij} = P(S_j|S_i);\ A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix},$$

Define X_i be the observation of day i. Stay in state S_i for exactly d days means it stay in S_i for $d-1$ days and move for other state in the d^{th} day.

The probability is.

$$P = p_{ii}^{d-1} * (1 - p_{ii})$$

Probability of the model stay in rainy (State 1)

$$P = 0.6 * (0.4)^{d-1}$$

Probability of the model stay in rainy (State 2)

$$P = 0.4 * (0.6)^{d-1}$$

Probability of the model stay in rainy (State 3)

$$P = 0.2 * (0.8)^{d-1}$$

(c)

Given the state is S_i. The next observation stays in S_i is a Binomial distribution with probability p_{ii}. Therefore, the state stays in S_i for exactly d days can be seen as a Bernoulli experiment that fails (stay in S_i state) $d-1$ time and success (not in S_i state) at d time. This is a geometric distribution D_{s_i} with $p = 1 - p_{ii}$.

$$D = p_{ii}^{d-1} * (1 - p_{ii})$$

$$E(D_{s_i}) = \sum_{d=1}^{\infty} dp(1-p)^{d-1} = 1 - p_{ii} \sum_{d=1}^{\infty} d(p_{ii})^{d-1}$$
$$= 1 - p_{ii} * \frac{1}{(1 - p_{ii})^2} = \frac{1}{1 - p_{ii}}$$

$E(D_{S_1}) = \frac{1}{1-0.4} \approx 1.6667$

$E(D_{S_2}) = \frac{1}{1-0.6} = 2.5$

$E(D_{S_3}) = \frac{1}{1-0.8} = 5$

Problem 3

Notation:

$$Q = \{q_0, q_1, q_2 \dots q_N\}, A\ set\ of\ posibile\ state$$
$$Z = \{0,1,2,3, \dots M\}, A\ set\ of\ posibile\ observations$$
$$X = \{x_0, x_1, x_2 \dots . x_T\}, Markov\ process\ sequence$$
$$O = \{o_0, o_1, o_2 \dots o_T\}, a\ observation\ sequence$$
$$A: transition\ probability\ matrix$$
$$B: observation\ probability\ matrix$$

$$Forward\ Algorithm: \quad \alpha_t(i) = b_i(O_t) \sum_{j=0}^{N-1} \alpha_{t-1}(j) a_{ij} = P(o_0, o_1, o_2 \dots o_t, x_t = q_i|\lambda)$$

By the forward Algorithm

$\alpha_t(i)$	$P(o_0, o_1, o_2 \dots o_t, x_t = H\|\lambda)$	$P(o_0, o_1, o_2 \dots o_t, x_t = C\|\lambda)$
$O_0 = 0$	0.06	0.28
$O_1 = 2$	0.077	0.0186
$O_2 = 2$	0.03067	0.003426

$$P(O|\lambda) = \sum_{j=0}^{N-1} \alpha_{t-1}(j)$$

	$P(O\|\lambda)$
$O = \{O_0 = 0\}$	0.34
$O = \{O_0 = 0, O_1 = 2\}$	0.0956
$O = \{O_0 = 0, O_1 = 2, O_2 = 2\}$	0.034096

$$Backward\ Algorithm: \beta_t(i) = \sum_{j=0}^{N-1} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j) \ = P(o_{t+1}, o_{t+2}, o_{t+3} \dots o_{t-}, |x_t = q_i, \lambda)$$

By the backward Algorithm

$\beta_t(i)$	$P(o_{t+1}, o_{t+2}, o_{t+3} \dots o_{t-}, \|x_t = H, \lambda)$	$P(o_{t+1}, o_{t+2}, o_{t+3} \dots o_{t-}, \|x_t = C, \lambda)$
$O_0 = 0$	0.1408	0.0916
$O_1 = 2$	0.38	0.26
$O_2 = 2$	1	1

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)} = P(x_t = q_i|O, \lambda)$$

The distribution of probability is:

$\gamma_t(i)$	$P(x_t = H\|O, \lambda)$	$P(x_t = C\|O, \lambda)$
$t = 0$	0.24777	0.75223
$t = 1$	0.85817	0.14183
$t = 2$	0.89952	0.10048

Choosing x_t with highest probability.

The Most likely sequence of X is: $\{\boldsymbol{C}, \boldsymbol{H}. \boldsymbol{H}\}$

Note: Metrics are calculated by code. See Link below:

https://github.com/WeigengLi/6364MLHW4