Name: Janmejay Mohanty **Course:** Applied Machine Learning **Homework Assignment**: 4
CWID: 20009315 **Course Number:** AAI 695

Solutions

Ans1: Proof of Bayes Theorem:

Using Conditional Probability,

$P(A \cap B) = P(A|B) * P(B)$

Rearranging above equation,

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ -------(1)

Similarly, again using Conditional Probability,

$P(B|A) = \frac{P(A \cap B)}{P(A)}$ -------(2)

Now combining both equations (1) and (2),

$P(A|B) * P(B) = P(B|A) * P(A)$ -------(3)

Now, dividing both sides by $P(B)$ for solving $P(A|B)$, we get:

$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$ -------(4)

Hence Proved. This above equation is knowns as Bayes Theorem which relates the conditional

Probability of B given A.

Bayes theorem is a powerful statistical tool that helps us to update our beliefs about an event or hypothesis based on new evidence or information. This makes it particularly useful for machine learning problems, where we often need to update our models as we receive new data.
In machine learning, Bayes theorem is used in various ways, including:

1. **Bayesian Inference:** This method is employed to calculate the model parameters' probability distribution. Bayesian inference offers a logical method for learning from data by revising prior assumptions about model parameters in light of fresh information.
2. **Bayesian Networks:** These are graphical representations of the probabilistic connections between variables. Bayesian networks can be used to represent complicated systems, and they are especially helpful when there are lots of variables and little available data.
3. **Bayesian Optimization:** This method is used to improve a machine learning model's hyperparameters. We can quickly search the hyperparameter space and identify the best values for our model by using Bayesian inference to update our views about the ideal hyperparameters.

Overall, Bayes theorem provides a principled and flexible framework for reasoning under uncertainty, which makes it a useful tool for many machine learning problems.

Ans2: $P(Cancer|+) = \frac{P(+|Cancer) * P(Cancer)}{P(+|Cancer) * P(Cancer) + P(+|\neg Cancer) * P(\neg Cancer)}$

$$= \frac{0.98 * 0.008}{(0.98 * 0.008) + (0.03 * 0.992)}$$

$$= \frac{0.0078}{0.0078 + 0.0298}$$

$$= \frac{0.0078}{0.0376}$$

$$= 0.2074$$

$$= 0.21$$

And,

$$P(Cancer|+ +) = \frac{P(+|Cancer) * P(Cancer|+)}{P(+|Cancer) * P(Cancer|+) + P(+|\neg Cancer) * P(\neg Cancer|+)}$$

We know that,

$$P(Cancer|+) = 0.21$$

$$P(\neg Cancer|+) = 1 - P(Cancer|+) = 1 - 0.21 = 0.79$$

Hence,

$$P(Cancer|+ +) = \frac{0.98 * 0.21}{0.98 * 0.21 + 0.03 * 0.79}$$

$$= \frac{0.2058}{0.2058 + 0.0237}$$

$$= \frac{0.2058}{0.2295}$$

$$= 0.8967$$

$$P(\neg Cancer|+ +) = 1 - P(Cancer|+ +) = 1 - 0.8967 = 0.103$$

Note: Rounding off may cause different results.

Ans3:

$$P(PlayTennis = Yes) = \frac{8}{12}$$

$$P(PlayTennis = No) = \frac{4}{12}$$

$$P(Outlook = Sun \mid PlayTennis = Yes) = \frac{2}{8}$$

$$P(Outlook = Sun \mid PlayTennis = No) = \frac{3}{4}$$

$$P(Temperature = Cool \mid PlayTennis = Yes) = \frac{3}{8}$$

$$P(Temperature = Cool \mid PlayTennis = No) = \frac{1}{4}$$

$$P(Humidity = High \mid PlayTennis = Yes) = \frac{3}{8}$$

$$P(Humidity = High \mid PlayTennis = No) = \frac{3}{4}$$

$$P(Wind = Strong \mid PlayTennis = Yes) = \frac{3}{8}$$

$$P(Wind = Strong \mid PlayTennis = No) = \frac{2}{4}$$

$P(PlayTennis = Yes \mid Outlook = Sun, Temperature = Cool, Humidity = High, Wind = Strong)$
$= P(PlayTennis = Yes) * P(Outlook = Sun \mid PlayTennis = Yes)$
$* P(Temperature = Cool \mid PlayTennis = Yes) * P(Humidity = High \mid PlayTennis = Yes)$
$* P(Wind = Strong \mid PlayTennis = Yes)$

$$= \frac{8}{12} * \frac{2}{8} * \frac{3}{8} * \frac{3}{8} * \frac{3}{8}$$

$$= 0.6667 * 0.25 * 0.375 * 0.375 * 0.375$$

$$= 0.0088$$

$P(PlayTennis = No \mid Outlook = Sun, Temperature = Cool, Humidity = High, Wind = Strong)$
$= P(PlayTennis = Yes) * P(Outlook = Sun \mid PlayTennis = Yes)$
$* P(Temperature = Cool \mid PlayTennis = Yes) * P(Humidity = High \mid PlayTennis = Yes)$
$* P(Wind = Strong \mid PlayTennis = Yes)$

$$= \frac{4}{12} * \frac{3}{4} * \frac{1}{4} * \frac{3}{4} * \frac{2}{4}$$

$$= 0.3333 * 0.75 * 0.25 * 0.75 * 0.5$$

$$= 0.0234$$

$$= 0.023$$

As $0.023 > 0.0088$

Therefore, based on the Naive Bayes algorithm, the predicted value for PlayTennis for the same new instance $< Outlook = Sun, Temperature = Cool, Humidity = High, Wind = Strong >$ is No.

Ans4: First Iteration,

$$net_c = w_{co} + (a * w_{ca}) + (b * w_{cb})$$

$$= 0.1 + 0.1 + 0$$

$$= 0.2$$

$$O_c = \frac{1}{1 + e_c^{-net}} = \frac{1}{1 + e^{-0.2}} = 0.55$$

$$net_d = w_{do} + (O_c * w_{dc})$$

$$= 0.1 + (0.55 * 0.1)$$

$$= 0.155$$

Similarly,

$$O_d = \frac{1}{1 + e_d^{-net}} = \frac{1}{1 + e^{-0.155}} = 0.539$$

Using Backpropagation,

$$\delta_d = O_d * (1 - O_d) * (t_a - O_a)$$
$$= 0.539 * (1 - 0.539) * (1 - 0.539)$$
$$= 0.115$$
$$\Delta w_{dc} = \eta * \delta_d * O_c + \alpha * O$$
$$= 0.3 * 0.115 * 0.55$$
$$= 0.019$$
$$\Delta w_{d0} = 0.034$$
$$\therefore w_{dc} = w_{dc} + \Delta w_{dc}$$
$$= 0.1 + 0.019$$
$$= 0.119$$
$$\therefore w_{d0} = w_{d0} + \Delta w_{d0}$$
$$= 0.1 + 0.034$$
$$= 0.134$$
$$\delta_c = O_c * (1 - O_c) * (w_c * \delta_d)$$
$$= 0.55 * (1 - 0.55) * (0.1 * 0.115)$$
$$= 0.003$$
$$\Delta w_{ca} = \eta * \delta_c * x_a + \alpha * 0$$
$$= 0.3 * 0.003 * 1$$
$$= 0.001$$
$$\Delta w_{c0} = 0.001$$
$$\Delta w_{cb} = 0$$
$$w_{c0} = w_{c0} + \Delta w_{c0} = 0.1 + 0.001 = 0.101$$
$$w_{ca} = w_{ca} + \Delta w_{ca} = 0.1 + 0.001 = 0.101$$
$$w_{cb} = w_{cb} + \Delta w_{cb} = 0.1 + 0 = 0.1$$

Second Iteration

$$net_c = w_{c0} + a * w_{ca} + b * w_{cb}$$
$$= 0.101 + 0 + 0.1$$
$$= 0.201$$

$$O_c = \frac{1}{1 + e^{-0.201}} = 0.55$$

$$net_d = w_{d0} + (O_c * w_{dc})$$

$$= 0.134 + (0.55 * 0.119)$$

$$= 0.1994$$

Similarly,

$$O_d = \frac{1}{1 + e^{-0.1994}} = 0.5496$$

Using Backpropagation,

$$\delta_d = O_d * (1 - O_d) * (1 - O_d)$$

$$= 0.5496 * (1 - 0.5496) * (1 - 0.5496)$$

$$= -0.136$$

$$\Delta w_{dc} = \eta * \delta_d * O_c + \alpha * \Delta w_{dc}$$

$$= (0.3 * (-0.136) * 0.55) + (0.9 * 0.019)$$

$$= -0.0053$$

$$\Delta w_{d0} = -0.01$$

$$\therefore w_{dc} = w_{dc} + \Delta w_{dc}$$

$$= 0.119 + (-0.0053)$$

$$= 0.113$$

$$w_{d0} = w_{d0} + \Delta w_{d0}$$

$$= 0.134 + (-0.01)$$

$$= 0.124$$

$$\delta_c = O_c * (1 - O_c) * (w_{dc} * \delta_d)$$

$$= 0.55(1 - 0.55)(0.113 * -0.136)$$

$$= -0.0038$$

$$\Delta w_{ca} = \eta * \delta_c * x_a + \alpha * \Delta w_{ca}$$

$$= 0.3 * (-0.0038) * 0 + 0.9 * 0.101$$

$$= 0.0009$$

$$\Delta w_{c0} = 0$$

$$\Delta w_{cb} = -0.001$$

$$\therefore w_{c0} = w_{c0} + \Delta w_{c0} = 0.101 + 0 = 0.101$$

$$w_{ca} = w_{ca} + \Delta w_{ca} = 0.101 + 0.0009 = 0.102$$

$$w_{cb} = w_{cb} + \Delta w_{cb} = 0.100 + (-0.001) = 0.099$$

Final Results:

$$w_{c0} = 0.101$$

$$w_{ca} = 0.102$$

$$w_{cb} = 0.099$$

$$w_{d0} = 0.124$$

$$w_{dc} = 0.113$$