Bayesian Learning Algorithm

What is Bayesian Algorithm?

- Bayesian learning algorithm is a method of calculating probabilities for hypothesis

- One of the most practical approaches to certain type of learning problems

Use of Bayesian analysis

- Used to justify a design choice in neural network algorithm

- Provides perspective for understanding other learning algorithms

- Outperforms other methods

Bayesian Theorem:

$$P(a \mid T) = \frac{P(T \mid a) \times P(a)}{P(T)}$$

P (T | a)   -   Conditional Probability

P (a)   -   Prior Probability

P (T)   -   Estimated Probability

Prior Probability

P (Asymptomatic)   = 142 / 302 = 0.4702 = 47%

P (Abnormal angina) =   50 / 302 = 0.1656 = 17%

P (Angina)   =   23 / 302 = 0.0761 =   8%

P (No tang)   =   87 / 302 = 0.2881 = 28%

| Attribute | Value | Count (Chest pain type) | | | |
		Asymptomatic	Abnormal Angina	No Tang	Angina
Gender	M	104	32	19	52
	F	38	18	4	35
Heart Rate	0 - 100	6	0	0	2
	100 - 130	39	3	3	6
	130 - 170	84	28	12	57
	170 - inf.	13	19	8	22

Attribute	Value	Probabilities (Chest pain type)			
		Asymptomatic	Abnormal Angina	No tang	Angina
Gender	M	104 / 142	32 / 50	19 / 23	52 / 87
	F	38 / 142	18 / 50	4 / 23	35 / 87
Heart Rate	0 - 100	6 / 142	0	0	2 / 87
	100 - 130	39 / 142	3 / 50	3 / 23	6 / 87
	130 - 170	84 / 142	28 / 50	12 / 23	57 / 87
	170 - inf.	13 / 142	19 / 50	8 / 23	22 / 87

Conditional Probability

P (T | Asymptomatic)   =   104/142 x 84/142 = 0.4332

P (T | Abnormal Angina) =   32/50 x 28/50   = 0.3584

P (T | Angina)   =   19/23 x 12/23   = 0.4309

P (T | No Tang)   =   52/87 x 57/87   = 0.3916

Combining the conditional and prior probabilities, we estimate a likelihood of each chest pain type:

Likelihood of Asymptomatic   =   0.4702 x 0.4332 = 0.2037

Likelihood of Abnormal Angina   =   0.1656 x 0.3584 = 0.0594

Likelihood of Angina   =   0.0761 x 0.4309 = 0.0328

Likelihood of No Tang   =   0.2881 x 0.3916 = 0.1128

Estimated Probability

- The estimated probability P(T) is a sum of likelihood values of each class

P (T) = 0.2037 + 0.0594 + 0.0328 + 0.1128 = 0.4087

Actual or Final Probability

$$P\,(\text{Asymptomatic}) = \frac{0.4332 \times 0.4702}{0.4087} = 0.50 = 50\%$$

$$P\,(\text{Abnormal Angina}) = \frac{0.3584 \times 0.1656}{0.4087} = 0.15 = 14\%$$

$$P\,(\text{Angina}) = \frac{0.4309 \times 0.0761}{0.4087} = 0.08 = 8\%$$

$$P\,(\text{No Tang}) = \frac{0.3916 \times 0.2881}{0.4087} = 0.28 = 28\%$$

Advantages of Bayesian Method

- Really easy to use

- It requires one scan of training data

- New instances can be classified by combining the predictions of multiple hypothesis

Disadvantages of Bayesian Method

- It does not always give us results that are satisfied enough to do our classification

- The attributes that we would use are not always independent

- Division of the ranges can effect the results