4/18/19

Thursday, April 18, 2019 1:32 PM

See syllabus

Case Study Credit Card Screening

good credit card declared bad : "false positive"
bad credit card declared good : "false negative"

2 kinds of info:

(truth) → B = (card really is bad)
G = (card is really good)

what (system says)

$\oplus$ = (system says bad)
$\ominus$ = (system says good)

1% = prevalence
= $P(B)$

97% = $P(\ominus \mid G)$ "specificity"

98% = $P(\oplus \mid B)$ "sensitivity"

1st way

(2x2 contingency table)
cross-tabulation of truth vs system

		truth B	truth G	
system says	$\oplus$	98	297	395
	$\ominus$	2	9,603	9,605
		100	9,900	10,000 ← transactions

sensitivity, specificity, prevalence

we want $P(B \mid \oplus) = \frac{98}{395} \doteq 25\%$

(false positive rate) $= P(G \mid \oplus) \doteq 75\%$

(false negative rate) $\approx P(B \mid \ominus) = \frac{2}{9{,}605} \doteq 0.02\%$

approve | deny — fraud score Fs: 0 $\ominus$ c $\oplus$ 100

· Is the card bad? **Fact** (statistical inference)
· Should we deny the transaction? **Choice** (Bayesian decision theory)
↑ system

		truth B	truth G
system says	$\oplus$	ok	
	$\ominus$		ok

agents: {· customer, · merchant, · bank}

bank suffers small loss (bad) but not really bad

(card bad but system says good) (bad) much worse to bank

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n)$$

Simulation-based approximate probabilities

· (1750) Buffon needle
· (1908) W. Gosset
· (1941) Metropolis & Ulam (Monte Carlo Sampling)
John Von Neuman
algorithms for pseudo-random # generation

Fact: All probabilities are conditional, conditional on assumptions

$P(H) =$ undefined $\quad P(H \mid \text{fair coin tossing}) = \frac{1}{2}$

Chain rule for $\cap$ = and

$$P(A_1 \cap \dots \cap A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \dots P(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

hard

$P(A) = P(A \text{ and } B_1$ or A and B_2 or ... or A and $B_K)$

simple addition rule for or
$= P(A \text{ and } B_1) + P(A \text{ and } B_2) + \dots + P(A \text{ and } B_K)$

$= P(B_1) \cdot P(A \mid B_1) + P(B_2) \cdot P(A \mid B_2) + \dots + P(B_K) \cdot P(A \mid B_K)$ (LTP)