



Northeastern University, Khoury College of Computer Science

# CS 6220 Data Mining | Assignment 5

Due: Some Date in 2024 (100 points)

Wenbo Qian

<https://github.com/wenboqian/CS-6220/tree/main/A5>

## Naive Bayes, Bayes Rules

According to formula  $P(A | B) = P(B | A) * P(A) / P(B)$

$P(\text{Parkinsons} | \text{Positive Detection}) = P(\text{Positive Detection} | \text{Parkinsons}) * P(\text{Parkinsons}) / P(\text{Positive Detection})$

$P(\text{Positive Detection}) = P(\text{Positive Detection} | \text{Parkinsons}) * P(\text{Parkinsons}) + P(\text{Positive Detection} | \text{No Parkinsons}) * P(\text{No Parkinsons})$

So  $P(\text{Parkinsons} | \text{Positive Detection}) = 0.09 / (0.9 * 0.1) + (0.1 * 0.9) = 0.5$

## The Sum of Conditional Probabilities

$P(A_i | B) = P(B | A_i) * P(A_i) / P(B)$

$$\sum_i P(A_i | B) = \sum_i P(B | A_i) * \frac{P(A_i)}{P(B)} = 1/P(B) \sum_i P(B | A_i)P(A_i)$$

Because  $P(B) = \sum_i P(B | A_i)P(A_i)$

So  $\sum_i P(A_i | B) = 1 / \sum_i P(B | A_i)P(A_i) * \sum_i P(B | A_i)P(A_i) = 1$