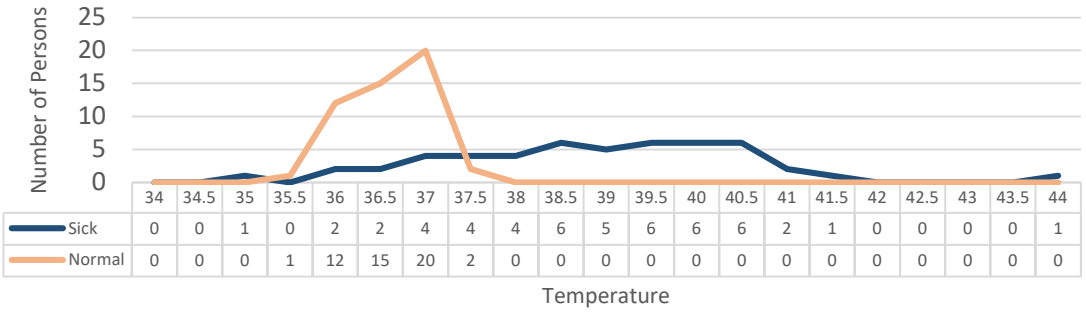
Artificial Intelligence III:
Artificial Intelligence and Deep Learning
Ch02 – Fundamental Principles
Tutorial

Jessica and you have been tasked with developing a machine learning system to determine whether a student at SCUT is sick upon campus entry.

1. Identify three relevant feature that could be used by this diagnostic system apart from body temperature, and explain your choices.

2. Jessica proposes to randomly select 50 individuals from the School of Medical Engineering for the model's development phase. Discuss the effectiveness of this strategy.

3. Consider a scenario where 50 independent and identically distributed (i.i.d.) samples have been gathered:



	34	34.5	35	35.5	36	36.5	37	37.5	38	38.5	39	39.5	40	40.5	41	41.5	42	42.5	43	43.5	44
Sick	0	0	1	0	2	2	4	4	4	6	5	6	6	6	2	1	0	0	0	0	1
Normal	0	0	0	1	12	15	20	2	0	0	0	0	0	0	0	0	0	0	0	0	0

   a. Propose a necessary data cleaning method for this collected data.

   b. Jessica has set the decision threshold for body temperature at 39.0°C. Discuss the effectiveness of this choice.

   c. Assuming the costs of all kinds of error are uniform, determine an appropriate temperature threshold.

   d. If the penalty for falsely identifying a healthy individual as ill is greater, should the threshold established in (c) still apply? If not, what adjustments would you recommend?

   e. Jessica believes incorporating additional features enhance the system's accuracy. Do you agree with her? Are there any potential shortcomings?