

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

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.

Answer:

- **Heart Rate**
Heart rate can provide insights into their overall well-being. An elevated heart rate can be a symptom of various conditions, ranging from stress and dehydration to more severe issues like heart diseases or infections.
- **Blood Pressure**
Blood pressure readings offer critical information about the force of blood against artery walls as the heart pumps it around the body.
- **Oxygen Saturation**
Oxygen saturation levels in the blood provide essential insights into respiratory and cardiovascular health. Levels below this range could indicate respiratory issues.

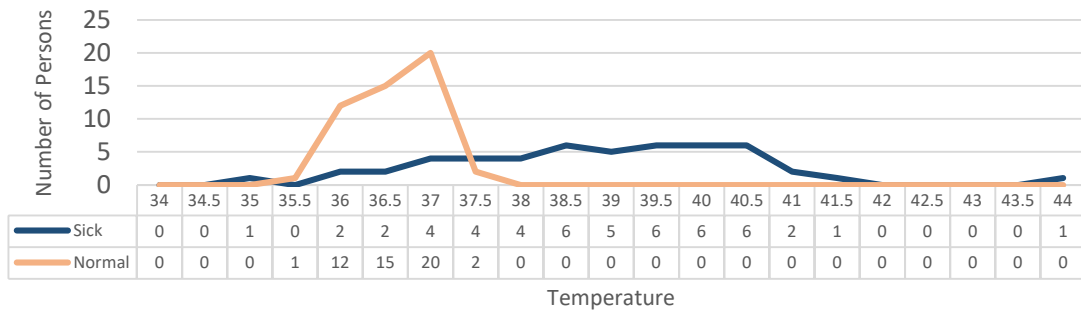
Other reasonable answers are acceptable:

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.

Answer:

SCUT has a large and diverse student population, which exceeds 80,000 students across various schools. The characteristics of students within the School of Medical Engineering may significantly differ from those in other schools. Therefore, a sample size of 50 students from the School of Medical Engineering is not sufficiently representative of the broader SCUT student body.

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



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

Answer:

- Outlier Detection**
 The temperature of a student is very high (44°C), which may be abnormal. The sample may be removed.
- Noise Reduction**
 Some temperatures are relatively low, e.g. 35°C in sick persons, and 36 in normal persons. These values may be adjusted.

Other reasonable answers are acceptable:

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

Answer:

Labeling someone as sick only if their temperature is above 39°C is not effective because many sick people have temperatures lower than 39°C.

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

Answer:

37.5°C

- 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?

Answer:

Since mistakenly marking a healthy person as sick carries a higher cost, this type of error should be avoided. Therefore, based on the provided data, the threshold should be shifted to right to reduce such mistakes, and 38°C would be a more reasonable setting.

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

Answer:

In general, using more features into a model can offer a richer understanding of a health status, potentially improving the accuracy of the system. However, adding more features can also complicate the model, particularly if these features don't contribute significantly to distinguishing between healthy and sick individuals. If the learning model is not able to handle this increased complexity, its performance may decrease. Additionally, the process of collecting more features also increase the cost and effort for sample collection.