

1 anomaly detection が適しているのはどれ?

- ① 病院の患者のたくさんデータから、(インフルエンザなどの)特定の病気を患っているか判断する \times
- ② primary care 患者のたくさんデータから、異常な状態の個人を見つける \bigcirc
- ③ コンピュータチップを製造する工場で、欠陥チップを見つける \bigcirc
- ④ クレジットカードの利用記録から、食料、洋服などの購入物のタイプを分類する \times

$$y = \begin{cases} 1 & \text{if } p(x) < \epsilon \\ 0 & \text{if } p(x) \geq \epsilon \end{cases}$$

CV に対し false positive が多い。 \rightarrow 多くは 1 と予想し、実際は 0 だった。
 $\epsilon \rightarrow$ 小

$$3 \quad p(x) = \prod_{j=1}^n p(x_j; \mu_j, \sigma_j^2)$$

$x_1 =$ vibration intensity (振動の強度) $0 \leq x_1 \leq 1$

$x_2 =$ heat generated (発熱量) $0 \leq x_2 \leq 1$

normal $x_1 \approx x_2$, $x_1 > 大$ and $x_2 < 小 \rightarrow$ abnormal

$$x_3 = \frac{x_1}{x_2} \quad \begin{array}{l} \text{normal だと } x_3 \approx 1 \\ \text{これだと } x_3 \gg 1 \\ \text{逆だと } x_3 \approx 0 \end{array}$$

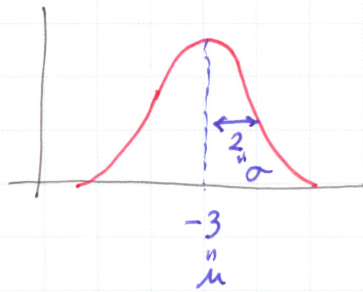
\times 4 ① anomaly detection system を作る時、数値による評価は役に立つ \bigcirc

② たくさん abnormal 事例が与えられる。 normal 事例は少ない。 \times 逆

③ $p(x)$ を negative 例で学習する \times negative 例は少ない

④ CV で (normal & abnormal 例を用いて) 評価する時、分類の accuracy はよい指標ではない \times
 skewed data

5 1D dataset $\{x^{(1)}, \dots, x^{(m)}\}$



$\therefore \mu_1 = -3, \sigma^2 = 4$

Quiz 1 2回目

2 cvに付して、anomalyがあるところを立上げて。 $v \rightarrow 大 \wedge$

4 ① モデル $p(x)$ を negative example ($y=0$) に fit させる。 positive は使わない。 ○

② 数値評価基準は 変なやつ ○

③ 以上の anomalous 存在がある。 normal 事例は少ない X

④ cv で 評価している時 (positive と negative 1), accuracy は よい基準だが X skewed