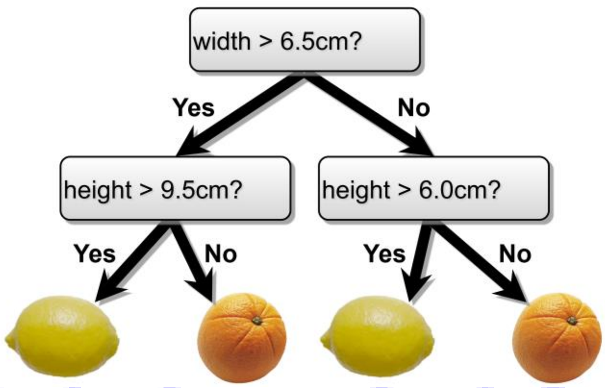
Live Interaction #6:

18th February 2024

E-masters Next Generation Wireless Technologies

EE902 Advanced ML Techniques for Wireless Technology

- Decision Tree Classifiers: (DTC)



- **DTC Advantage:** Intuitive and easy to interpret.
- **How to choose the feature?**
- **Entropy**
- **Information theory.**
- Symbols x_i
- Probabilities of symbols $p(x_i)$
- M symbols $i = 1,2, \dots, M$
- **Entropy** is defined as

$$\sum_{i=1}^{M} p(x_i) \log_2 \frac{1}{p(x_i)}$$

$$= -\sum_{i=1}^{M} p(x_i)\log_2 p(x_i)$$

- Example

	$\mathbf{IC}$	$\overline{\mathbf{IC}}$
$\mathbf{CHOC}$	$\frac{1}{2}$	$\frac{1}{8}$
$\overline{\mathbf{CHOC}}$	$\frac{1}{4}$	$\frac{1}{8}$

- $X = \{\mathbf{IC}, \overline{\mathbf{IC}}\}$

$$p(\mathbf{IC}) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$p(\overline{\mathbf{IC}}) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$H(X) = \frac{3}{4} \times \log_2 \frac{4}{3} + \frac{1}{4}\log_2 4 = 0.811$$

$$p(\mathbf{CHOC}) = \frac{5}{8}$$

$$p(\overline{\mathbf{CHOC}}) = \frac{3}{8}$$

$$H(Y) = \frac{3}{8} \times \log_2 \frac{8}{3} + \frac{5}{8} \times \log_2 \frac{8}{5} = 0.954$$

- **Conditional Entropy:**

$$H(X|Y) = \sum_{j} p(Y = y_j) H(X|Y = y_j)$$

$$H(Y|X) = \sum_{i} p(X = x_i) H(Y|X = x_i)$$

- What is $H(Y|X)$?

	IC	$\overline{\textbf{IC}}$
CHOC	$\frac{1}{2}$	$\frac{1}{8}$
$\overline{\textbf{CHOC}}$	$\frac{1}{4}$	$\frac{1}{8}$

$$H(Y|X=\textbf{IC}) = ?$$

$$p(Y=\textbf{CHOC}|X=\textbf{IC}) = \frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{4}} = \frac{2}{3}$$

$$p(Y=\overline{\textbf{CHOC}}|X=\textbf{IC}) = \frac{\frac{1}{4}}{\frac{1}{2}+\frac{1}{4}} = \frac{1}{3}$$

$$H(Y|X=\textbf{IC}) = \frac{2}{3}\times\log_2\frac{3}{2} + \frac{1}{3}\times\log_2 3$$

$$= 0.918$$

$$H(Y|X=\overline{\textbf{IC}}) = H\left\{\frac{1}{2},\frac{1}{2}\right\}$$

$$= \frac{1}{2}\times\log_2 2 + \frac{1}{2}\times\log_2 2 = 1$$

$$p(X=\textbf{IC})\times H(Y|X=\textbf{IC}) + p(X=\overline{\textbf{IC}})H(Y|X=\overline{\textbf{IC}})$$

$$= \frac{3}{4}\times 0.918 + \frac{1}{4}\times 1 = 0.938$$

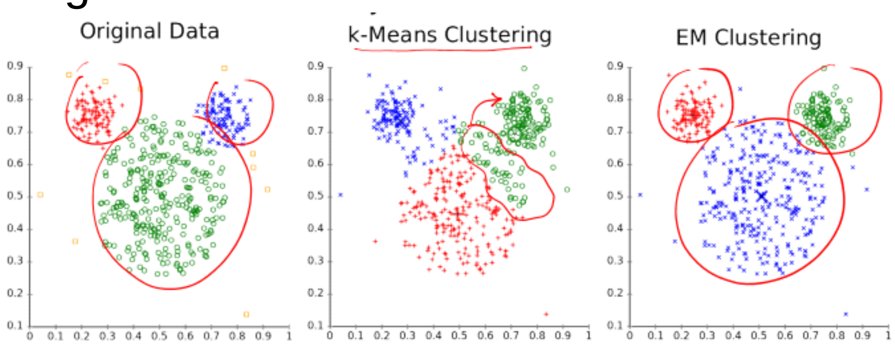
$$H(X|Y) = 0.7955$$

- Information Gain:

$$I(X;Y) = H(X) - H(X|Y)$$

$$= 0.811 - 0.7955 = 0.0155$$

- Choose the feature that has the ***maximum information gain***.
- **EM Algorithm:**
- Clustering.



- K-Means $\alpha_i(j) = 0{,}1$: **HARD CLUSTERING**.
- EM Algorithm performs **SOFT CLUSTERING**.

$$0 \leq \alpha_i(j) \leq 1$$

$$\sum_i \alpha_i(j) = 1$$

- $\alpha_i(j)$ is the Probability $\bar{\mathbf{x}}(j) \in \mathcal{C}_i$
- M-Step: **Centroid computation.**

$$\bar{\boldsymbol{\mu}}_i^{(l)} = \frac{\sum_{j=1}^{M} \alpha_i^{(l)}(j)\bar{\mathbf{x}}(j)}{\sum_{j=1}^{M} \alpha_i^{(l)}(j)}$$

- E-step: **Soft cluster assignment**.

$$\alpha_i^{(l)}(j) = \Pr(\bar{\mathbf{x}}(j) \in \mathcal{C}_i)$$

$$= \underbrace{\frac{p_i \times \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\left\|\bar{\mathbf{x}}(j)-\bar{\boldsymbol{\mu}}_i^{(l-1)}\right\|^2}}{\sum_k p_k \times \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\left\|\bar{\mathbf{x}}(j)-\bar{\boldsymbol{\mu}}_k^{(l-1)}\right\|^2}}}_{\text{Bayes rule}}$$

- LDA with prior probabilities:
- Choose $\mathcal{C}_0$ if

$$\bar{\mathbf{h}}^T(\bar{\mathbf{x}} - \widetilde{\boldsymbol{\mu}}) \geq \ln\frac{p_1}{p_0}$$

- Simplification:

$$(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0) - (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1}(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1) \leq 2\ln\frac{p_0}{p_1}$$

$$\Rightarrow 2(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\bar{\mathbf{x}} + \bar{\boldsymbol{\mu}}_0^T \mathbf{R}^{-1}\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1 \mathbf{R}^{-1}\bar{\boldsymbol{\mu}}_1 \leq 2\ln\frac{p_0}{p_1}$$

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0) \leq \ln\frac{p_0}{p_1}$$

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\left(\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)\right) \leq \ln\frac{p_0}{p_1}$$

$$\Rightarrow (\overline{\boldsymbol{\mu}}_0 - \overline{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} \left(\bar{\mathbf{x}} - \frac{1}{2} (\overline{\boldsymbol{\mu}}_1 + \overline{\boldsymbol{\mu}}_0) \right) \geq \ln \frac{p_1}{p_0}$$

- **Assignment #6 Deadline: 23rd Feb Friday 11:59 PM.**
- **Assignment #5, 6 Discussion: 24th Feb Saturday 2:00 PM – 3:00 PM.**
- **Quiz #3: 24th February Saturday 3:30 - 4:30 PM.**
- **Live interaction #7: 25th February Sunday 2:00 – 3:00 PM.**
- **Assignment #7 Deadline: 1st March Friday 11:59 PM.**
- **Live interaction #8: 3rd March Sunday 2:00 – 3:00 PM.**
- **Assignment #8 Deadline: 7st March Thursday 11:59 PM.**
- **Assignment #7, 8 Discussion: 8th March Friday 8:00 PM – 8:30 PM.**
- **Quiz #4: 8th March Friday 9:00 - 9:45 PM.**
- **Final Exam: 10th March Sunday 9:00 AM – 12:00 PM. (Please check!!)**