HW 1 – Machine Learning

Mathematics behind the decision tree

Let a group of laws: $R = \{r_1, r_2, \dots, r_k\}$. Every law r_i has m_i^{ϕ} don't care bits (ϕ).

The number of times each law appears depends on the number of don't care bits it has: $\boldsymbol{N_{r_i} = 2^{m_i^{\phi}}}$

In total the number of laws (after duplicates) is: $\boldsymbol{N_{tot} = \sum_{i=1}^{k} N_{r_i}}$

Our goal is to compute the Information Gain (IG) - $\boldsymbol{IG(R, b_j) = H(R) - H(R \mid b_j)}$

$$\begin{aligned}
\boldsymbol{H(R)} &= \boldsymbol{-\sum_{i=1}^{k} \frac{N_{r_i}}{N_{tot}} \cdot log_2\left(\frac{N_{r_i}}{N_{tot}}\right)} \\
&= \boldsymbol{\frac{1}{N_{tot}}\left(log_2(N_{tot}) \cdot \sum_{i=1}^{k} N_{r_i} - \sum_{i=1}^{k} N_{r_i}\, log_2(N_{r_i}\right)} \\
&= \boldsymbol{log_2(N_{tot}) - \frac{1}{N_{tot}} \sum_{i=1}^{k} N_{r_i}\, log_2(N_{r_i})}
\end{aligned}$$

To compute $\boldsymbol{H(R \mid b_j)}$ we need to evaluate separately the cases for which $b_j = 0$ and $b_j = 1$.

- $b_j = 0$: Contains the laws for which $b_j = 0$ and the ones where $b_j = \phi$ -> In total N_0 laws
- Same for $b_j = 1$
- $N_{tot} = N_o + N_1$

$$\begin{aligned}
&\boldsymbol{H(R \mid b_j = 0) = -\sum\nolimits_{b_j=0} \frac{N_{r_i}}{N_0} \cdot log_2\left(\frac{N_{r_i}}{N_0}\right) - \sum\nolimits_{b_j=\phi} \frac{\frac{N_{r_i}}{2}}{N_0} \cdot log_2\left(\frac{\frac{N_{r_i}}{2}}{N_0}\right)} \\
&\boldsymbol{= \frac{1}{N_0}\left[N_0 \cdot log_2(N_0) - \sum\nolimits_{b_j=0} N_{r_i}\, log_2(N_{r_1}) - \frac{1}{2}\sum\nolimits_{b_j=\phi} N_{r_i}\left(log_2(N_{r_i}) - log_2(N_0) - 1\right)\right]} \\
&\boldsymbol{= log_2(N_0) - \frac{1}{N_0}\left[\sum\nolimits_{b_j=0} N_{r_i}\, log_2(N_{r_i}) + \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2} log_2(N_{r_i}) - \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2}\right]}
\end{aligned}$$

$$\begin{aligned}
&\boldsymbol{H(R \mid b_j) = \frac{N_0}{N_{tot}} \cdot H(R \mid b_j = 0) + \frac{N_1}{N_{tot}} \cdot H(R \mid b_j = 1)} \\
&\boldsymbol{= \frac{1}{N_{tot}}\Bigg[N_0 \cdot log_2(N_0) - \sum\nolimits_{b_j=0} N_{r_i}\, log_2(N_{r_i}) - \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2} log_2(N_{r_i}) + \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2} + N_1} \\
&\qquad\qquad \boldsymbol{\cdot\, log_2(N_1) - \sum\nolimits_{b_j=1} N_{r_i}\, log_2(N_{r_i}) - \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2} log_2(N_{r_i}) + \sum\nolimits_{b_j=\phi} \frac{N_{r_i}}{2}\Bigg]}
\end{aligned}$$

$$= \frac{1}{N_{tot}}\left[N_0 \cdot log_2(N_0) + N_1 \cdot log_2(N_1) + \sum\nolimits_{b_j=\phi} N_{r_i} - \sum\nolimits_{b_j=1} N_{r_i}\, log_2(N_{r_i}) \right.$$
$$\left. - \sum\nolimits_{b_j=0} N_{r_i}\, log_2(N_{r_i}) - \sum\nolimits_{b_j=\phi} N_{r_i}\, log_2(N_{r_i})\right]$$
$$= \frac{1}{N_{tot}}\left[N_0 \cdot log_2(N_0) + N_1 \cdot log_2(N_1) + \sum\nolimits_{b_j=\phi} N_{r_i} - \sum_{i=1}^{k} N_{r_i}\, log_2(N_{r_i})\right]$$

Hence, we found:

$$IG(R, b_j) = H(R) - H(R \mid b_j) = \; log_2(N_{tot}) - \frac{1}{N_{tot}}\left[N_0 \cdot log_2(N_0) + N_1 \cdot log_2(N_1) + \sum\nolimits_{b_j=\phi} N_{r_i}\right]$$

Where:

$$N_0 = \sum\nolimits_{b_j=0} N_{r_i} + \frac{1}{2} \cdot \sum\nolimits_{b_j=\phi} N_{r_i}, \qquad N_1 = \sum\nolimits_{b_j=1} N_{r_i} + \frac{1}{2} \cdot \sum\nolimits_{b_j=\phi} N_{r_i}$$

From the obtained result we can go a bit further by replacing the values of $N_{r_i} = 2^{m_i^\phi}$:

$$N_0 = \sum\nolimits_{b_j=0} 2^{m_i^\phi} + \sum\nolimits_{b_j=\phi} 2^{m_i^\phi - 1} = N_0^j + N_\phi^j$$

$$N_1 = \sum\nolimits_{b_j=1} 2^{m_i^\phi} + \sum\nolimits_{b_j=\phi} 2^{m_i^\phi - 1} = N_1^j + N_\phi^j$$

$$N_{tot} = N_0 + N_1 = N_0^j + N_1^j + 2N_\phi^j$$

$$IG(R, b_j) = log_2(N_{tot})$$
$$- \frac{1}{N_{tot}}\left[\left(N_0^j + N_\phi^j\right) log_2\left(N_0^j + N_\phi^j\right) + \left(N_1^j + N_\phi^j\right) log_2\left(N_1^j + N_\phi^j\right) + 2N_\phi^j\right]$$

$$log_2\left(N_{r_i}\right) = log_2\left(2^{m_i^\phi}\right) = m_i^\phi \Rightarrow H(R) = log_2(N_{tot}) - \frac{1}{N_{tot}}\sum_{i=1}^{k} m_i^\phi \cdot 2^{m_i^\phi}$$

$$N_0 = \sum\nolimits_{b_j=0} 2^{m_i^\phi} + \sum\nolimits_{b_j=\phi} 2^{m_i^\phi - 1} = N_0^j + N_\phi^j, \qquad N_1 = \sum\nolimits_{b_j=1} 2^{m_i^\phi} + \sum\nolimits_{b_j=\phi} 2^{m_i^\phi - 1} = N_1^j + N_\phi^j$$

$$\begin{aligned} H(R \mid b_j = 0) &= -\sum\nolimits_{b_j=0} \frac{2^{m_i^\phi}}{N_0^j + N_\phi^j} \cdot \left[m_i^\phi - log_2\left(N_0^j + N_\phi^j\right)\right] - \sum\nolimits_{b_j=\phi} \frac{2^{m_i^\phi - 1}}{N_0^j + N_\phi^j} \\ &\quad \cdot \left[m_i^\phi - 1 - log_2\left(N_0^j + N_\phi^j\right)\right] \\ &= -\frac{1}{N_0^j + N_\phi^j}\left[\sum\nolimits_{b_j=0}\left[m_i^\phi - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi} + \sum\nolimits_{b_j=\phi}\left[m_i^\phi - 1 - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi - 1}\right] \end{aligned}$$

$$IG\left(R, b_j\right) = H(R) - \frac{N_0^j + N_\phi^j}{N_{tot}} H(R \mid b_j = 0) - \frac{N_1^j + N_\phi^j}{N_{tot}} H(R \mid b_j = 1)$$

$$\begin{aligned} &= log_2(N_{tot}) - \frac{1}{N_{tot}}\sum_{i=1}^{k} m_i^\phi \cdot 2^{m_i^\phi} \\ &\qquad + \frac{1}{N_{tot}}\left[\sum\nolimits_{b_j=0}\left[m_i^\phi - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi}\right. \\ &\qquad \left. + \sum\nolimits_{b_j=\phi}\left[m_i^\phi - 1 - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi - 1}\right] \\ &\qquad + \frac{1}{N_{tot}}\left[\sum\nolimits_{b_j=0}\left[m_i^\phi - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi}\right. \\ &\qquad \left. + \sum\nolimits_{b_j=\phi}\left[m_i^\phi - 1 - log_2\left(N_0^j + N_\phi^j\right)\right]2^{m_i^\phi - 1}\right] \end{aligned}$$

$$\begin{aligned} &= log_2(N_{tot}) - \frac{1}{N_{tot}} log_2\left(N_0^j + N_\phi^j\right)N_0^j - \frac{1}{N_{tot}} log_2\left(N_1^j + N_\phi^j\right)N_1^j - \frac{1}{N_{tot}} log_2\left[2\left(N_0^j + N_\phi^j\right)\right]N_\phi^j \\ &\qquad - \frac{1}{N_{tot}} log_2\left[2\left(N_1^j + N_\phi^j\right)\right]N_\phi^j \\ &= log_2(N_{tot}) - \frac{1}{N_{tot}}\left[\left(N_0^j + N_\phi^j\right)\ log_2\left(N_0^j + N_\phi^j\right) + \left(N_1^j + N_\phi^j\right)\ log_2\left(N_1^j + N_\phi^j\right) + 2N_\phi^j\right] \end{aligned}$$

Conditional entropy by dismissing the ϕ (wild cards)

BEFORE:

$$H(R \mid b_j) = \frac{1}{N_{tot}}\left[(N_0^j + N_\phi^j) \cdot log_2\left(N_0^j + N_\phi^j\right) + (N_1^j + N_\phi^j) \cdot log_2\left(N_1^j + N_\phi^j\right) + 2N_\phi^j - \sum_{i=1}^{k} m_i^\phi \cdot 2^{m_i^\phi}\right]$$

AFTER:

$$H(R \mid b_j) = \frac{1}{N_0^j + N_1^j}\left[N_0^j \cdot log_2\left(N_0^j\right) + N_1^j \cdot log_2\left(N_1^j\right)\right]$$