Assignment 3 in Artificial Intelligence
Spring 2018

Daniel Aaron Salwerowicz

April 30, 2018

Task 1.1

Rules

```
1. IF   TS = Unchanged  AND  DJI = Up    THEN  Action = Buy

2. IF  DJI = Unchanged  AND   TS = Down  THEN  Action = Sell

3. IF  DJI = Down                        THEN  Action = Sell

4. IF   TS = Up                          THEN  Action = Buy

5. IF  DJI = Up                          THEN  Action = Wait

6. IF  DJI = Unchanged                   THEN  Action = Wait
```

Procedure

Sell:
    DJI = Down (2+, 0-); New rule!
Buy:
    TS = Up (2+, 0-); New rule!
    TS = Unchanged (1+, 1-)
    TS = Unchanged, DJI = Up (1+, 0-); New rule!
Sell:
    DJI = Unchanged (1+, 1-)
    DJI = Unchanged, TS = Down (1+, 0-); New rule!
Wait:
    DJI = Up (1+, 0-); New rule!
    DJI = Unchanged (1+, 0-); New rule!

Task 1.2

Entropy for any set is defined as:

$$E(S) = \sum -p(i) \log_2 p(i)$$

Using that I calculate entropy for Action set to be:

$$\begin{aligned}
E(Action) &= \sum -p(a) \log_2 p(a) \\
&= -\frac{3}{8} \log_2 \frac{3}{8} - \frac{3}{8} \log_2 \frac{3}{8} - \frac{2}{8} \log_2 \frac{2}{8} \\
&= 0.53064 + 0.53064 + 0.50000 \\
&= \underline{1.5613}
\end{aligned}$$

Task 1.3

Information gain for a feature is defines as:

$$Gain(S, A) = E(S) - \sum \left(\frac{|S_v|}{S} \cdot E\left(S_v\right) \right)$$

Using that I can calculate information gain for DJI, CR, and TS.

$$\begin{aligned}
Gain(S, DJI) &= E(S) - \frac{4}{8} E(S, Up) - \frac{2}{8} E(S, Down) - \frac{2}{8} E(S, Unchanged) \\
&= 1.5613 - \frac{1}{2} 0.81128 - \frac{1}{4} 0 - \frac{1}{4} 1 \\
&= \underline{\underline{0.90564}} \\
E(S, Up) &= -\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} - 0 \log_2 0 \\
&= 0.31128 + 0.50000 + 0 = \underline{0.81128} \\
E(S, Down) &= -\frac{2}{2} \log_2 \frac{2}{2} - 0 \log_2 0 - 0 \log_2 0 \\
&= 0 + 0 + 0 = \underline{0} \\
E(S, Unchanged) &= -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} - 0 \log_2 0 \\
&= 0.50000 + 0.50000 + 0 = \underline{1.0000}
\end{aligned}$$

$$Gain(S, CR) = E(S) - \frac{4}{8}E(S, High) - \frac{4}{8}E(S, Low)$$

$$= 1.5613 - \frac{1}{2}\frac{3}{2} - \frac{1}{2}\frac{3}{2}$$

$$= \underline{0.061278}$$

$$E(S, High) = -\frac{2}{4}\log_2\frac{2}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4}$$

$$= 0.50000 + 0.50000 + 0.50000 = \underline{1.5000}$$

$$E(S, Low) = -\frac{2}{4}\log_2\frac{2}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4}$$

$$= 0.50000 + 0.50000 + 0.50000 = \underline{1.5000}$$

$$Gain(S, TS) = E(S) - \frac{4}{8}E(S, Down) - \frac{2}{8}E(S, Up) - \frac{2}{8}E(S, Unchanged)$$

$$= 1.5613 - \frac{1}{2}0.81128 - \frac{1}{4}0 - \frac{1}{4}1$$

$$= \underline{0.90564}$$

$$E(S, Down) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} - 0\log_2 0$$

$$= 0.31128 + 0.50000 + 0 = \underline{0.81128}$$

$$E(S, Up) = -\frac{2}{2}\log_2\frac{2}{2} - 0\log_2 0 - 0\log_2 0$$

$$= 0 + 0 + 0 = \underline{0}$$

$$E(S, Unchanged) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} - 0\log_2 0$$

$$= 0.50000 + 0.50000 + 0 = \underline{1.0000}$$

Both DJI and TS have the same value for information gain, so one of them will be chosen to be the first layer of the decision tree that we will build for using the ID3 algorithm