CS141 PS4 Q2

Cody Falzone SID: 860929046

May 29, 2017

2. If there are m items (or features), there are 3m − 2m+1 + 1 different association rules possible. Prove this.

A rule is defined in the form:

$X \implies Y$

Given that there are m features, this means that X may be formed by taking a subset of n items chosen from m features. One possible subset is:

X = $\binom{m}{n}$

For all possible subsets, we have:

$$\sum_{n=1}^{m} \binom{m}{n}$$

Given that we picked n items from m, we are left with m - n items to pick from for choosing Y. That is:

Y = $\binom{m-n}{p}$

For all possible subsets, we have:

$$\sum_{p=1}^{m-n} \binom{m-n}{p}$$

For a given set of items sets that are large candidates we must enumerate all possibilities of the rules given by these combinations. For every $\binom{m}{n}$ ways to pick X there are $\binom{m-n}{p}$ ways we can pick Y. This leads to the multiplication:

$$\sum_{n=1}^{m} \binom{m}{n} \cdot \sum_{p=1}^{m-n} \binom{m-n}{p}$$

The summation of n choose m is equivalent to:

$$\sum_{n=0}^{m} \binom{m}{n} = 2^n$$

By starting the summation at n = 1 is the same as subtracting 2^0, which equals 1. Therefore we have:

$$\sum_{n=1}^{m} \binom{m}{n} = 2^n - 1$$

This accounts for invalid null sets not being in our rules.

$(2^m - 1) \cdot (2^{m-n} - 1)$

The multiplication of these summations is equivalent to:

$3^m - 2^{m+1} + 1$