1. Equation for predicting the probability of win for the Home team:

Ln(p(O=2)/1-p(O=2)) = 3.313 + 0.035\*(POINTS\_H) – 0.035\*(POINTS\_A) + 1.618\*(HTGD) + 0.01\*(TOTAL\_H\_P) – 0.015\*(TOTAL\_A\_P) – 3.320\*([FGS=0]) – 2.473\*([FGS=1])

1. The variables of conceding red cards by home or away teams do not influence the outcomes of the match, as they are not significant as per the model output.

For Match\_O = 1, their p-values are 0.599 and 0.391 which are greater than alpha (0.05)

For Match\_O =2, their p-values are 0.275 and 0.072 which are also greater than alpha.

So, these variables are not considered significant and hence not included in the predicting equation.

1. The variable of Total points earned by the home team and away team in previous season is relevant to the equation for predicting the outcome. The reason being as per the model, its p-value is less than alpha, hence the variable is considered significant.

The variable for home team is in positive direction while the for the away team it is negative direction.

1. Considering the values given:  
   HTGD=2, FGS=1, RED\_H=0, RED\_A=0, POINTS\_H=15, POINTS\_A=18, TOTAL\_H\_P=40, TOTAL\_A\_P=30

Using the equation with significant variables:  
Ln(p/1-p) = 3.313 + 0.035\*15 – 0.035\*18 + 1.618\*2 + 0.01\*40 – 0.015\*30 – 3.320\*0 – 2.473\*1 = 3.921

p/1-p = e^3.921=50.45 => p = 0.9805 = 98.05%

Therefore, the probability that the home team will win the match for the given values is **98.05%**

1. Assuming all other variables to be constant and considering the case of first goal scored by away team, the probabilities can be calculated as:

Log(p1|p0) = (3.535 – 3.521) = 0.014 → y1

Log(p2|p0) = (3.313 – 3.320) = -0.007 → y2

Therefore,

p1|p0 = e^y1 = 1.0141

p2|p0 = e^y2 = 0.9930

To calculate the probabilities of home team win (2), draw (1), loss (0), the below formulae are used:

p0 = 1/(1+ e^y2 + e^y1) = 1/(1 + 0.9930 + 1.0141) = 1/(3.0071) = 0.3325

p1 = e^y1/(1+ e^y2 + e^y1) = 1.0141/(1 + 0.9930 + 1.0141) = 0.3372

p2 = e^y2/(1+ e^y2 + e^y1) = 0.9930/(1 + 0.9930 + 1.0141) = 0.3302

Hence, the probabilities of win, draw, loss when the away team scores the first goal is around 33%, which suggests that it is not possible to decide the result of the match based on the first goal scored by the away team.

1. From the classification table, the correct percentage values for 1, that is for Draw is very low = 25.5%. That means that only 25.5% of predictions are correct for draws from this model. Hence one should not bet on the draws based on this model output.
2. Following are rules that should be used for betting from the CHAID model:

If HTGD = 0 and First Goal scored by = Away team, then **Loss** (Accuracy=49.1%)

If HTGD = 0 and First Goal scored by = Home team, then **Draw** (Accuracy=39.8%)

If HTGD = 0 and First Goal scored by = None scored a goal, then **Draw** (Accuracy=49.7%)

If HTGD = 1 and Total\_H\_P <= 67 then **Win** (Accuracy=73.5%)

If HTGD = 1 and Total\_H\_P > 67 then **Win** (Accuracy=91.7%)

If HTGD = 2,3 or 4, then **Win** (Accuracy=93.1%)

If HTGD = -1 or -4 and Total\_A\_P <=53 then **Loss** (Accuracy=46.1%)  
If HTGD = -1 or -4 and Total\_A\_P >53 then **Loss** (Accuracy=70.9%)  
If HTGD = -2,-3 or -5, then **Loss** (Accuracy=91.3%)

8. (Refer Excel sheet)

9. (Refer Excel sheet)

10. Peter’s choice:

Step 1: Based on the results of the Logistic Regression model, the CHAID tree and comparing these results with the actual outcomes of the 20 matches, the matches whose results are consistent across all three are selected.

Step 2: Since the accuracy/correct percentage for predicting a loss (i.e. 0) is 79.8% according to Exhibit 8, the below matches numbers remain.

|  |  |  |  |
| --- | --- | --- | --- |
| Match Number | HTGD | FGS | RESULT |
| 1 | -2 | 0 | 0 |
| 6 | -1 | 0 | 0 |
| 11 | 0 | 0 | 0 |
| 14 | -2 | 0 | 0 |
| 18 | -2 | 0 | 0 |
| 20 | -2 | 0 | 0 |

Step 3: Based on Exhibit 5, the maximum number of losses for the Home team based on HTGD are for value of HTGD = -1

So, there is a higher chance of Peter making the most out of betting on Match number 6.