STORYTELLING MODEL INTERPRETATION

1 NOVEMBER 2019
BADM HACKATHON TRAINING DAY

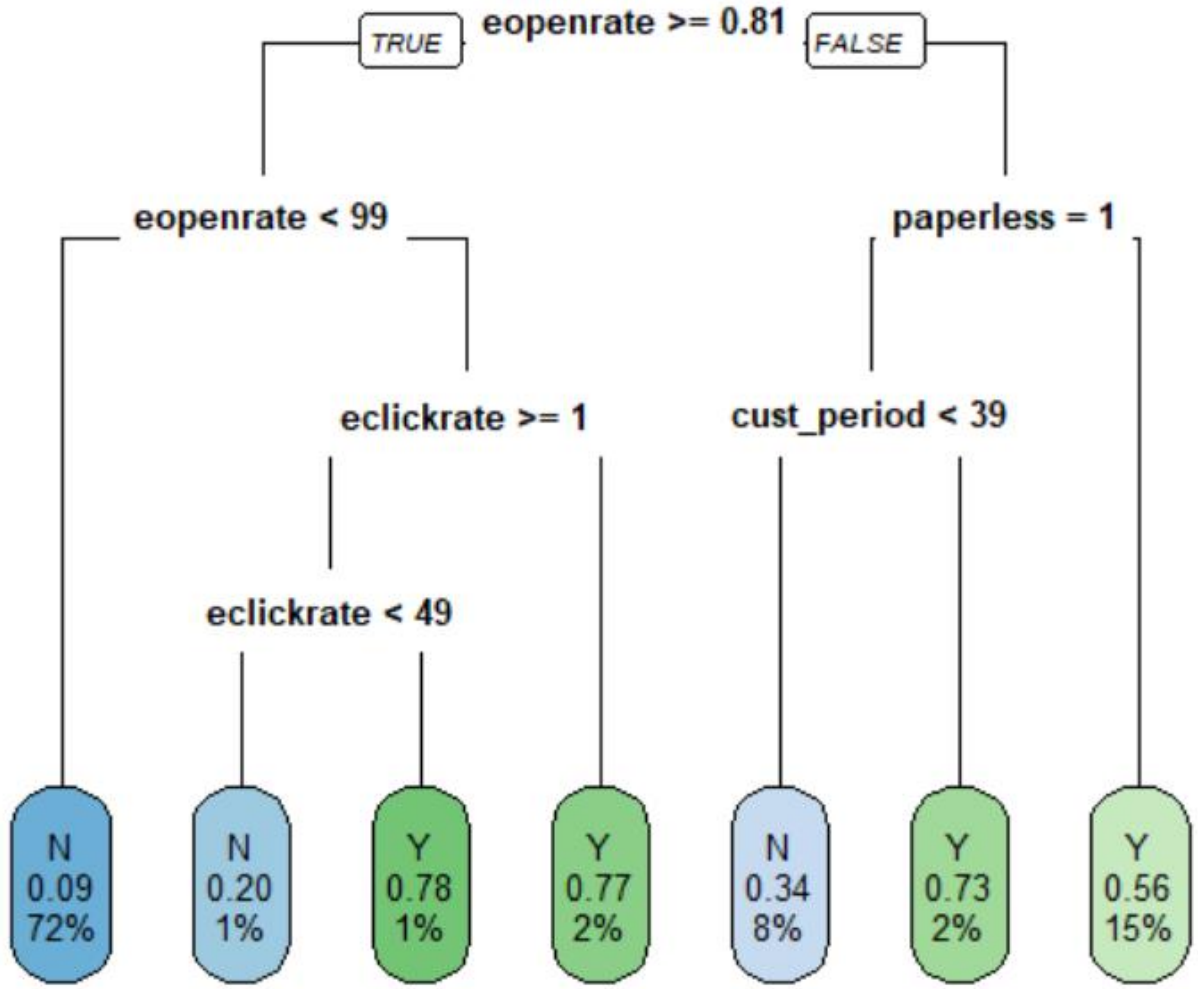
SFU
BEEDIE
SCHOOL OF BUSINESS

MODEL OUTPUTS

SFU BEEDIE SCHOOL OF BUSINESS

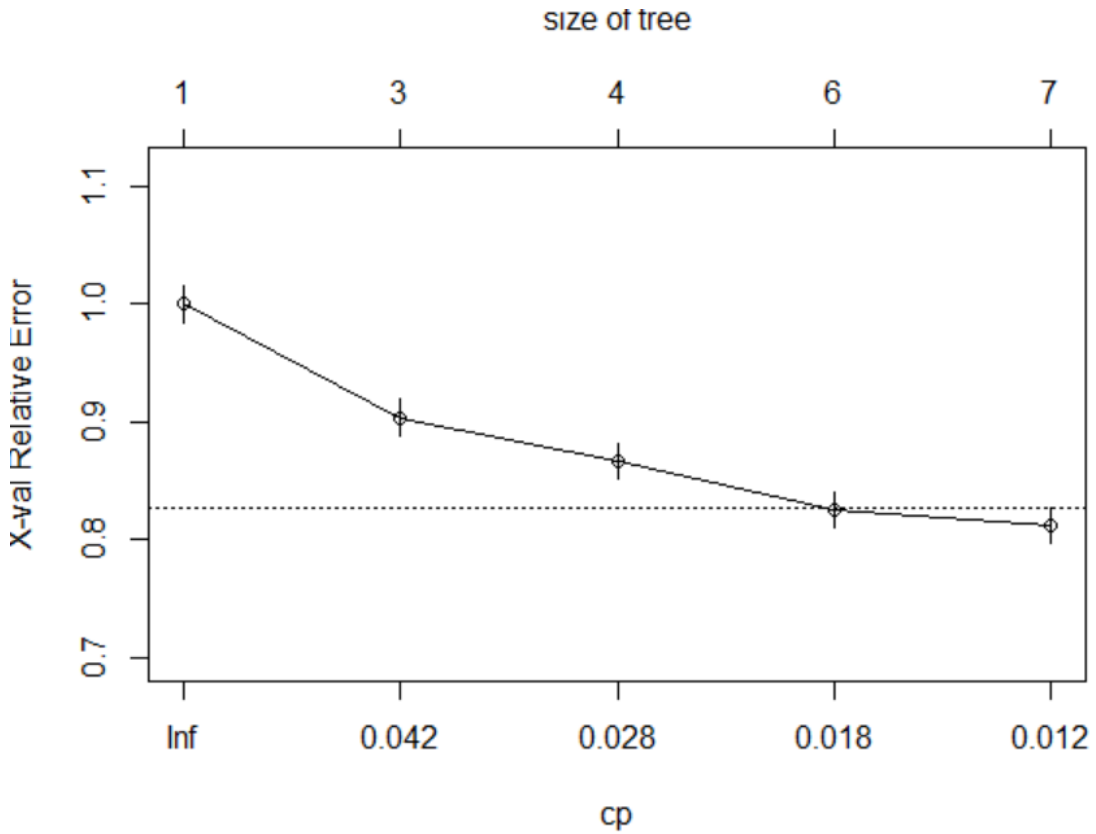
DECISION TREE

Model1.RPart



DECISION TREE

0.01
Complexity
Parameter




```
Variables actually used in tree construction:
[1] cust_period eclickrate eopenrate paperless
Root node error: 3147/15364 = 0.20483 n= 15364
CP nsplit rel error xerror xstd 1 0.048141 0
1.00000 1.00000 0.015896 2 0.037496 2 0.90372
0.90372 0.015297 3 0.020813 3 0.86622 0.86718
0.015054 4 0.015570 5 0.82459 0.82555 0.014764 5
0.010000 6 0.80902 0.81220 0.014668
```

SFU BEEDIE SCHOOL OF BUSINESS

LOGISTIC REGRESSION

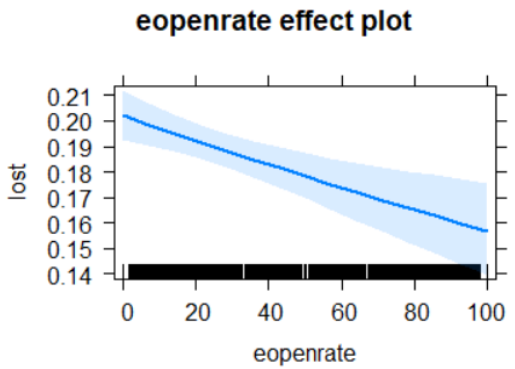
Summary

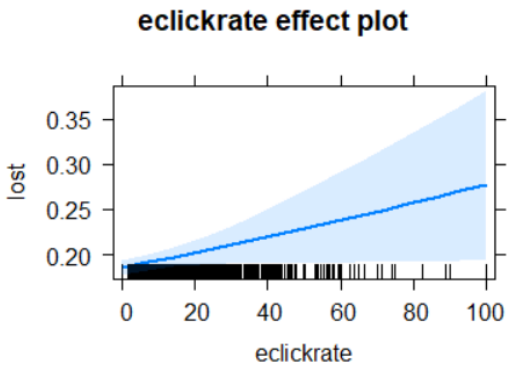
```
Call:
glm(formula = lost ~ eopenrate + eclickrate + avgorder + ordfreq +
    paperless + refill + doorstep + favday + city + cust_period,
    family = binomial(logit), data = filter(Retention2017, Sample ==
        "Estimation"))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1665  -0.7025  -0.5891  -0.4095   2.7587

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)     -1.1791340  0.1385997  -8.507  < 2e-16 ***
eopenrate       -0.0030877  0.0008689  -3.554  0.00038 ***
eclickrate       0.0052301  0.0025311   2.066  0.03880 *
avgorder        -0.0017974  0.0005858  -3.068  0.00215 **
ordfreq         -0.2589222  0.2093322  -1.237  0.21613
paperless       -0.7461929  0.0463274 -16.107  < 2e-16 ***
refill          -0.7789828  0.0988066  -7.884 3.17e-15 ***
doorstep        -0.6818040  0.1682071  -4.053 5.05e-05 ***
favdayMonday    -0.0289822  0.0670705  -0.432  0.66566
favdaySaturday   0.0942395  0.1335878   0.705  0.48053
favdaySunday     0.4632740  0.1454314   3.186  0.00144 **
favdayThursday   0.0455955  0.0703201   0.648  0.51673
favdayTuesday    0.0959347  0.0664393   1.444  0.14875
favdayWednesday  0.0939152  0.0707280   1.328  0.18423
cityCHO          0.7440737  0.1228154   6.058 1.37e-09 ***
cityDCX          0.3260985  0.1185809   2.750  0.00596 **
cityRIC          0.4890922  0.1213393   4.031 5.56e-05 ***
cust_period     -0.0005735  0.0001011  -5.674 1.39e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 15580  on 15363  degrees of freedom
Residual deviance: 14838  on 15346  degrees of freedom
AIC: 14874

Number of Fisher Scoring iterations: 5
```

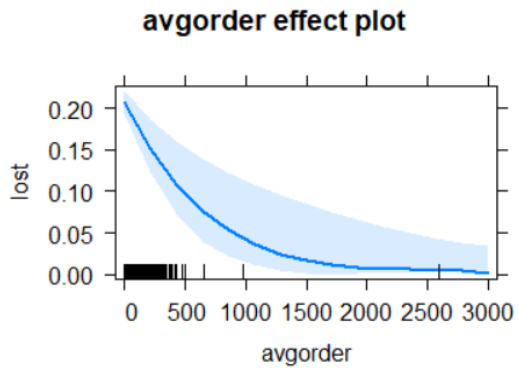
SFU BEEDIE SCHOOL OF BUSINESS

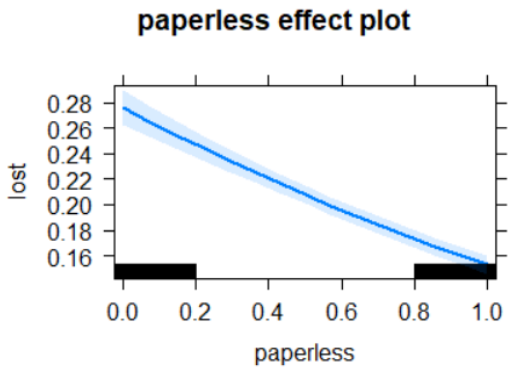
LOGISTIC REGRESSION

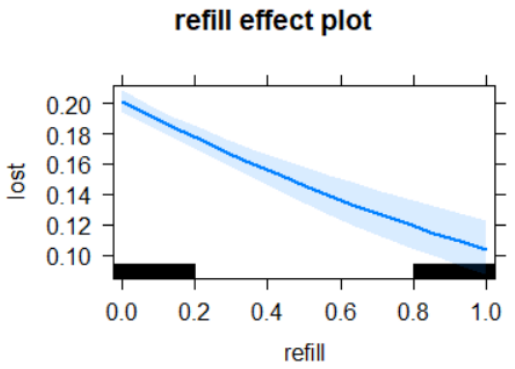
Effect Plots

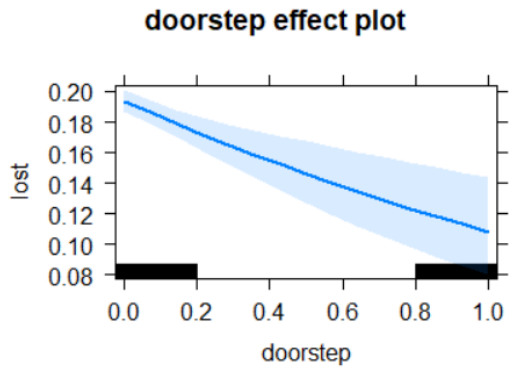


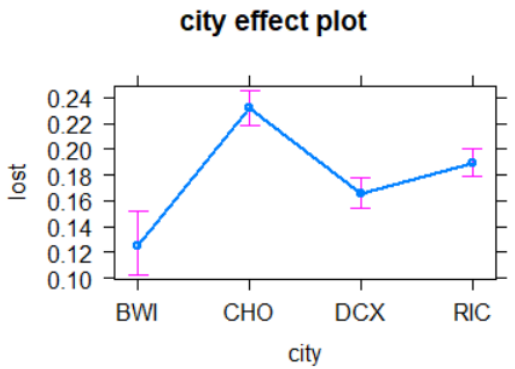


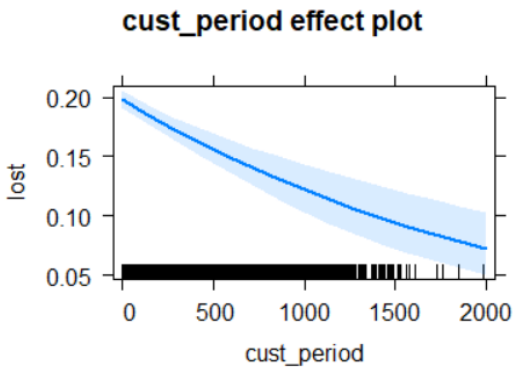


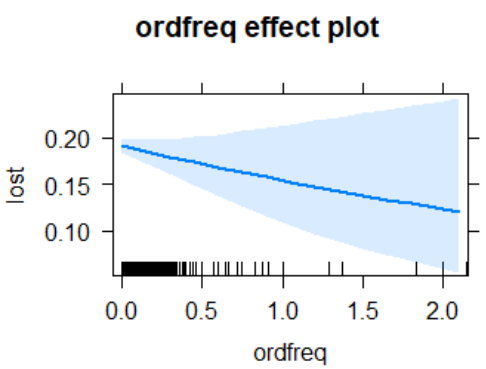








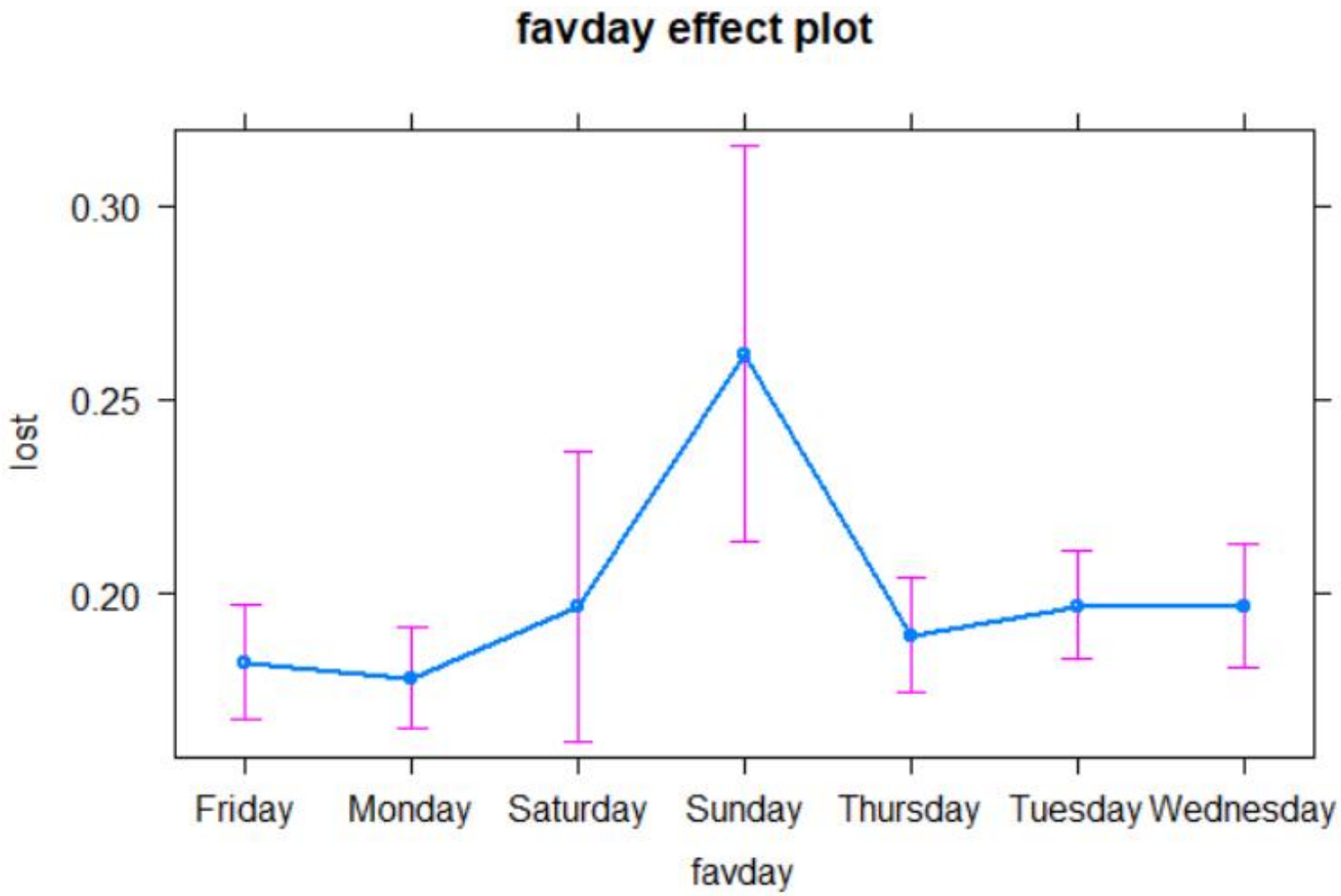




SFU BEEDIE SCHOOL OF BUSINESS

LOGISTIC REGRESSION

Effect Plots
Cont.



SFU BEEDIE SCHOOL OF BUSINESS

STEPWISE REGRESSION

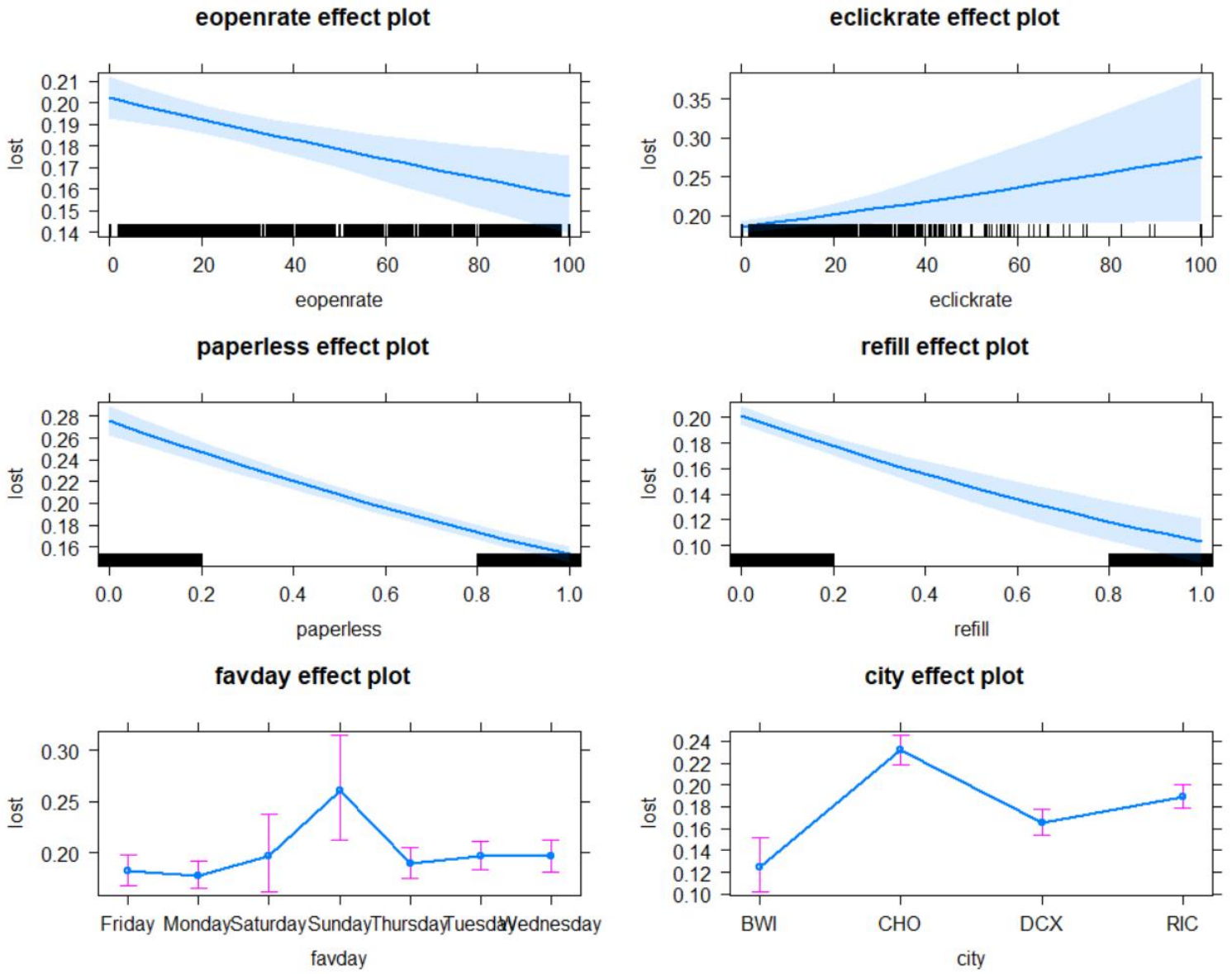
Summary

```
Call:
glm(formula = lost ~ eopenrate + eclickrate + avgorder + paperless +
    refill + doorstep + favday + city + cust_period, family = binomial(logit),
    data = filter(Retention2017, Sample == "Estimation"))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1606  -0.7026  -0.5893  -0.4107   2.7597

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)     -1.1877134  0.1384274  -8.580  < 2e-16 ***
eopenrate       -0.0030927  0.0008690  -3.559 0.000373 ***
eclickrate       0.0051114  0.0025310   2.020 0.043430 *
avgorder        -0.0018368  0.0005860  -3.134 0.001722 **
paperless       -0.7436690  0.0462733 -16.071  < 2e-16 ***
refill          -0.7846520  0.0987058  -7.949 1.87e-15 ***
doorstep        -0.6890675  0.1680878  -4.099 4.14e-05 ***
favdayMonday    -0.0293886  0.0670663  -0.438 0.661240
favdaySaturday   0.0955010  0.1335632   0.715 0.474594
favdaySunday     0.4579039  0.1453359   3.151 0.001629 **
favdayThursday   0.0458790  0.0703158   0.652 0.514097
favdayTuesday    0.0951621  0.0664318   1.432 0.152007
favdayWednesday  0.0922542  0.0707152   1.305 0.192033
cityCHO          0.7453465  0.1227835   6.070 1.28e-09 ***
cityDCX          0.3282712  0.1185503   2.769 0.005622 **
cityRIC          0.4916868  0.1212979   4.054 5.04e-05 ***
cust_period     -0.0005723  0.0001011  -5.662 1.50e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 15580  on 15363  degrees of freedom
Residual deviance: 14839  on 15347  degrees of freedom
AIC: 14873

Number of Fisher Scoring iterations: 5
```

SFU BEEDIE SCHOOL OF BUSINESS

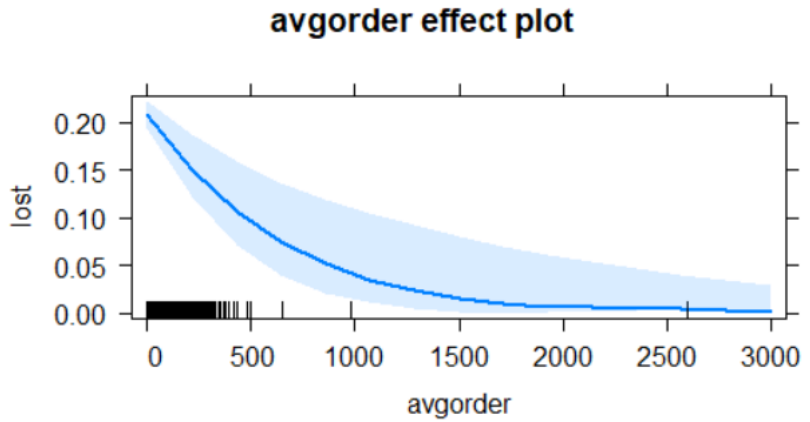
STEPWISE REGRESSION

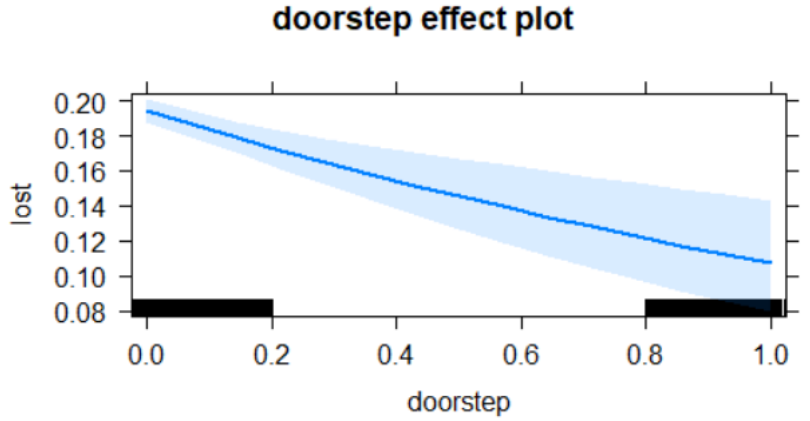
Effect Plots

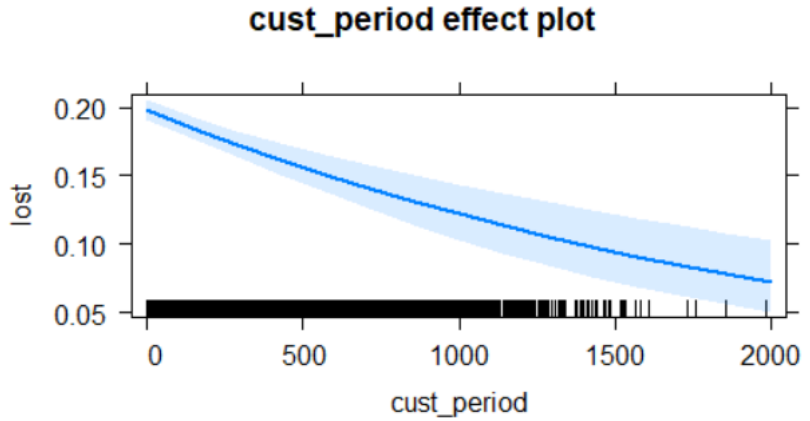


STEPWISE REGRESSION

Effect Plots Cont.







SFU BEEDIE SCHOOL OF BUSINESS

RANDOM FOREST

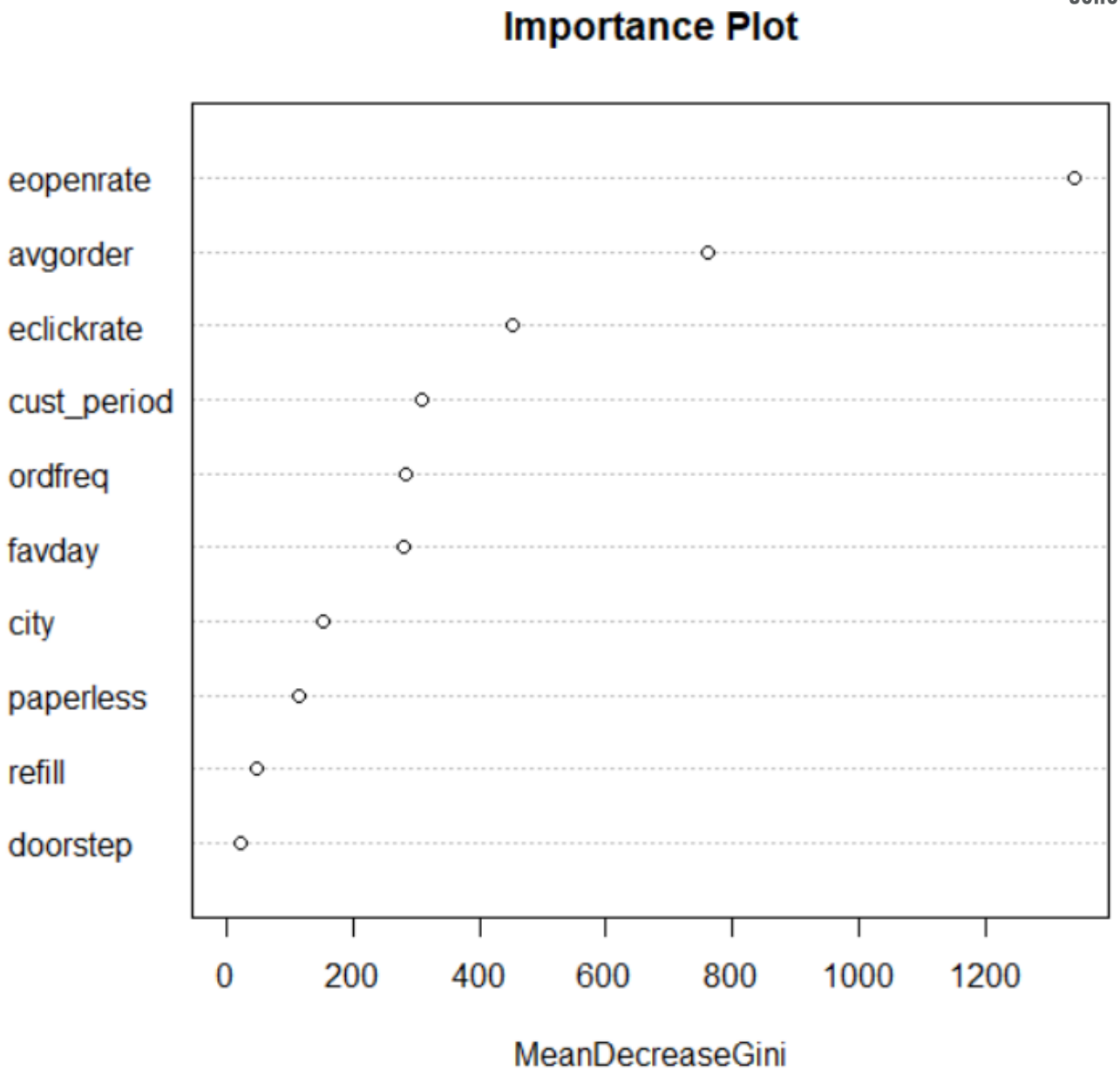
Summary

```
Call:
 randomForest(formula = lost ~ eopenrate + eclickrate + avgorder +      ordfreq + paperless + refill + door
step + favday + city +      cust_period, data = filter(Retention2017, Sample == "Estimation"),      importa
nce = TRUE, ntree = 500, mtry = 4)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 4

        OOB estimate of  error rate: 17.08%
Confusion matrix:
      N    Y class.error
N 11184 1033  0.08455431
Y  1591 1556  0.50556085
```

SFU BEEDIE SCHOOL OF BUSINESS

RANDOM FOREST

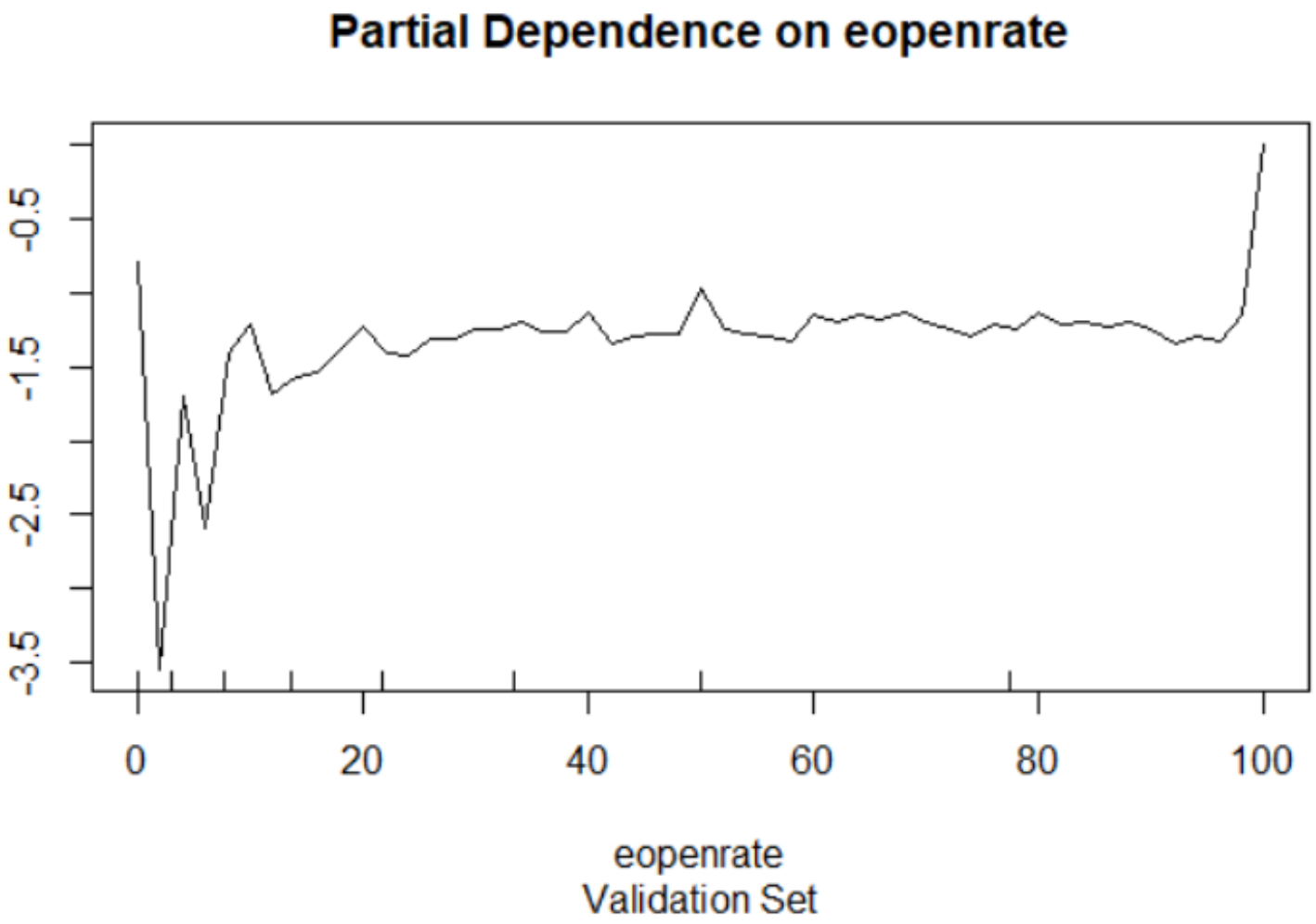
Importance Plot



SFU
BEEDIE
SCHOOL OF BUSINESS

RANDOM FOREST

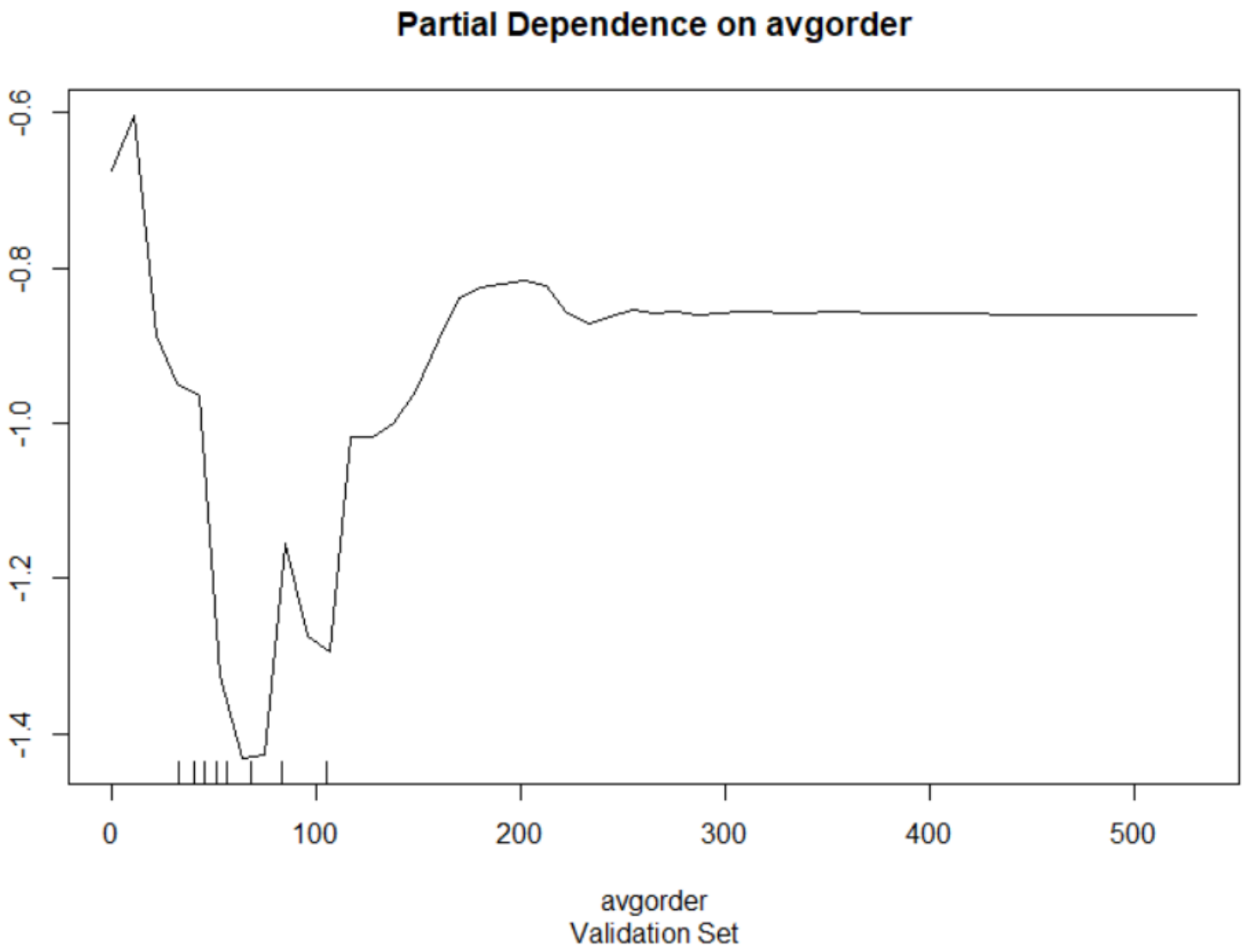
Partial Dependence Plot eopenrate



SFU
BEEDIE
SCHOOL OF BUSINESS

RANDOM FOREST

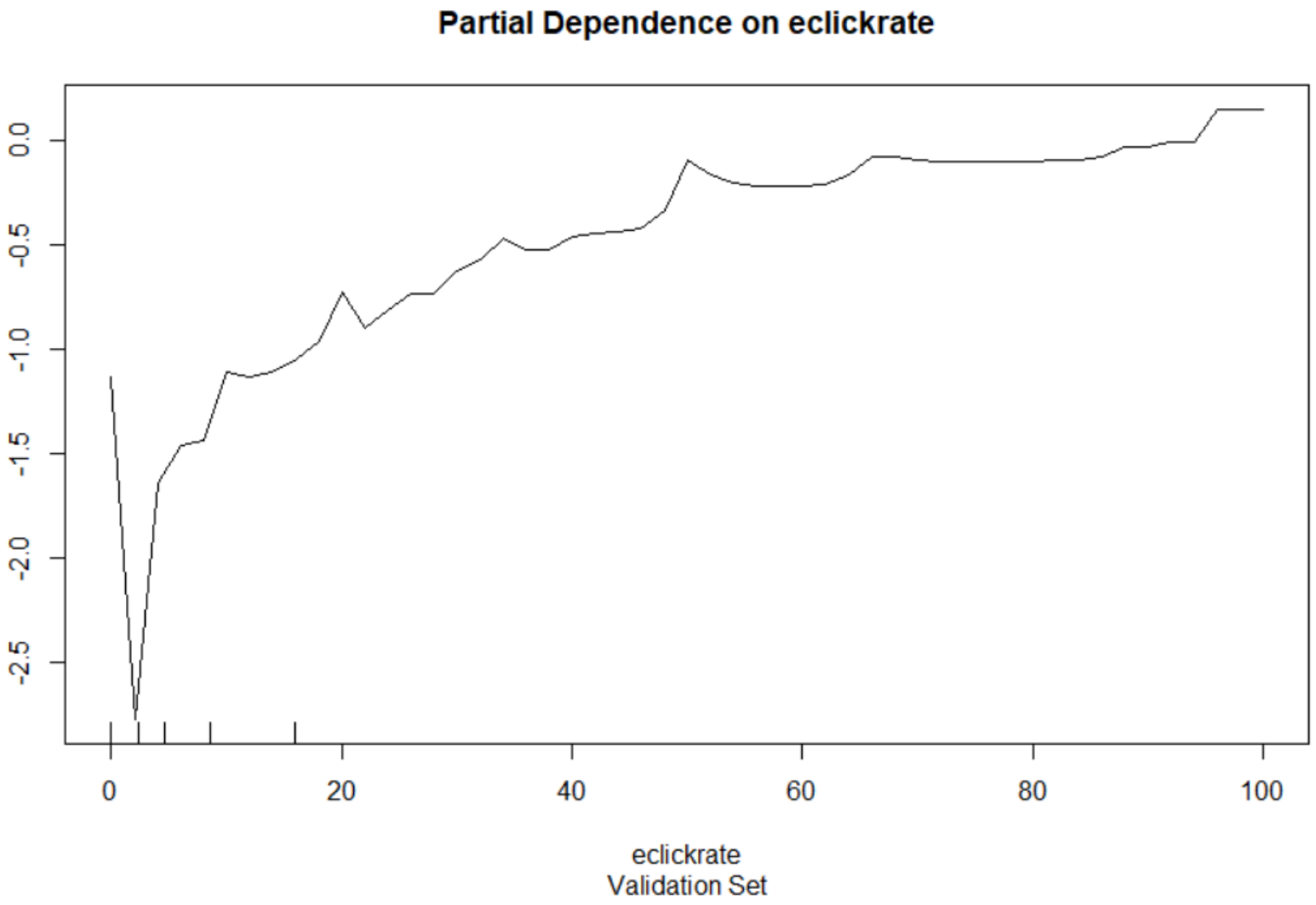
Partial Dependence Plot avgorder



SFU BEEDIE SCHOOL OF BUSINESS

RANDOM FOREST

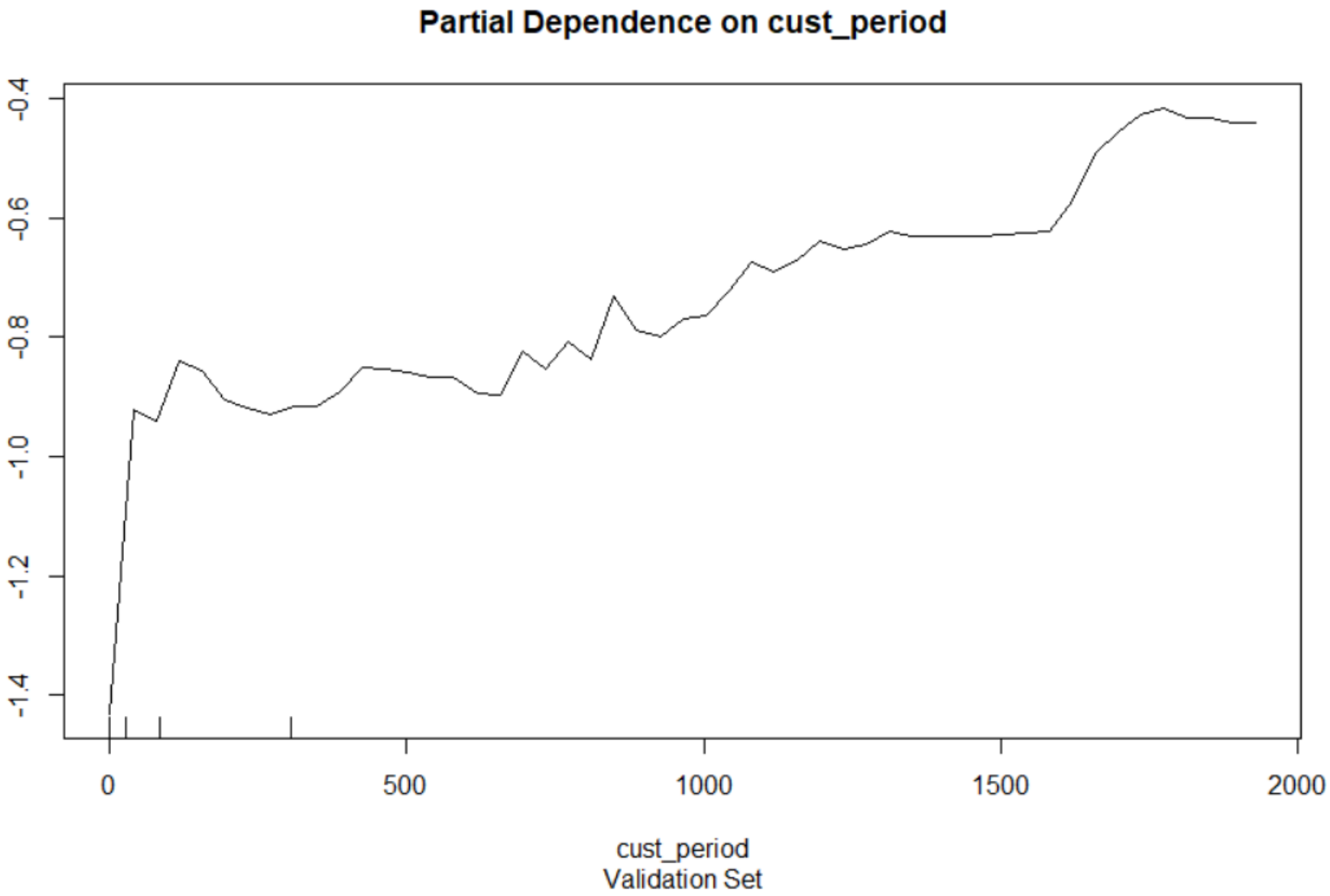
Partial Dependence Plot eclickrate



SFU
BEEDIE
SCHOOL OF BUSINESS

RANDOM FOREST

Partial Dependence Plot cust_period



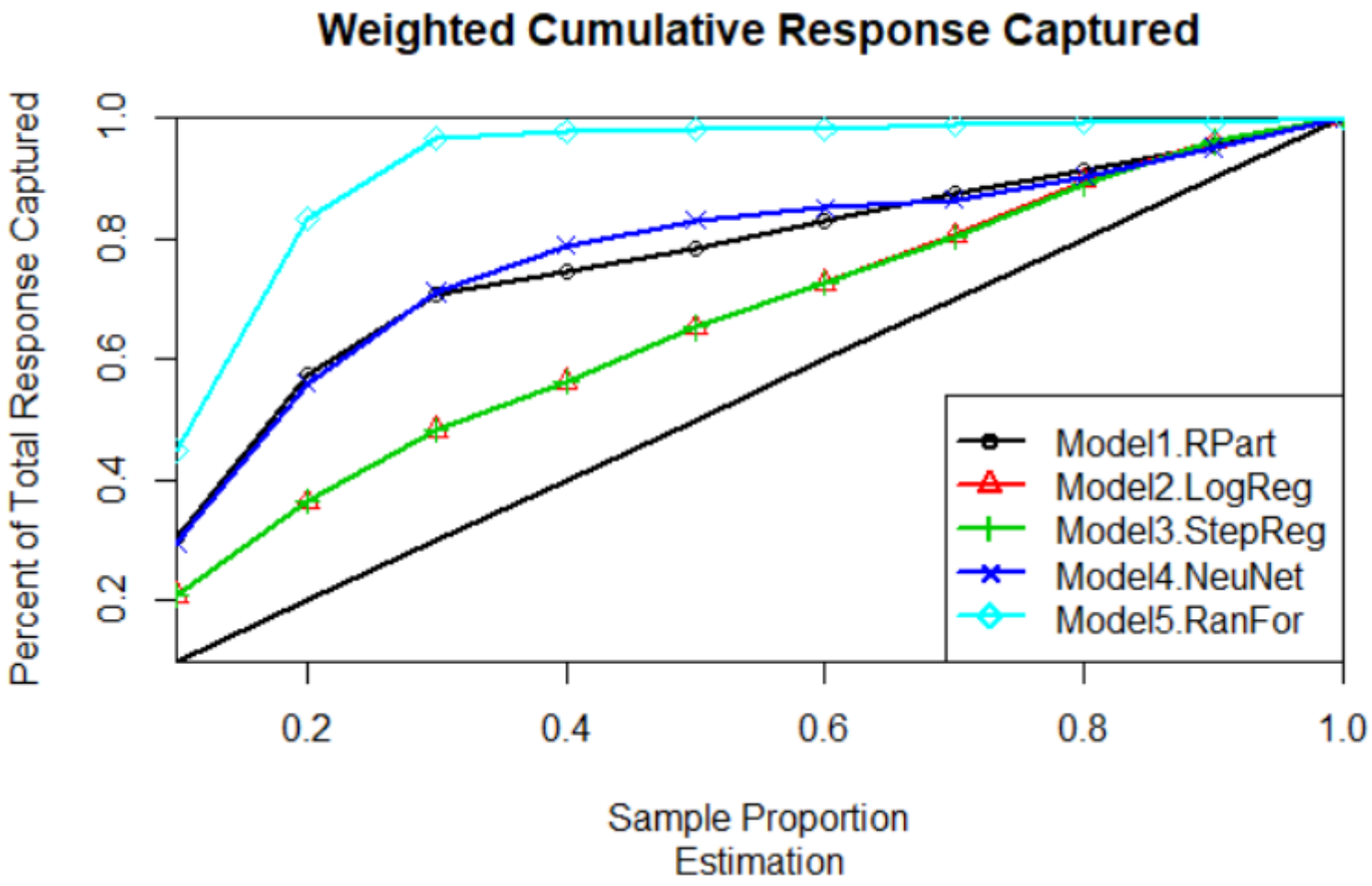
SFU
BEEDIE
SCHOOL OF BUSINESS

LIFT CHARTS

SFU BEEDIE SCHOOL OF BUSINESS

CUMULATIVE LIFT

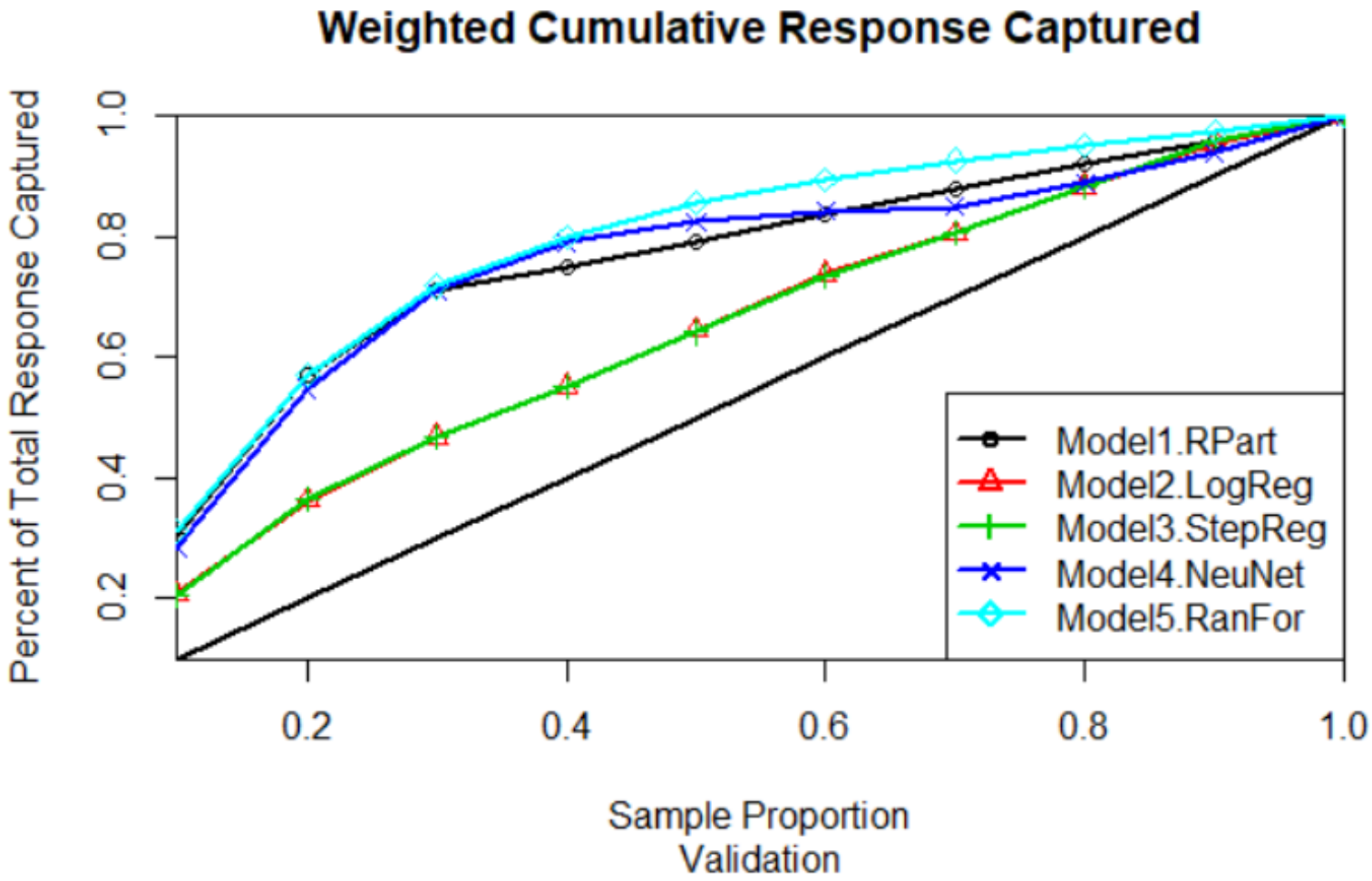
Estimation



SFU BEEDIE SCHOOL OF BUSINESS

CUMULATIVE LIFT

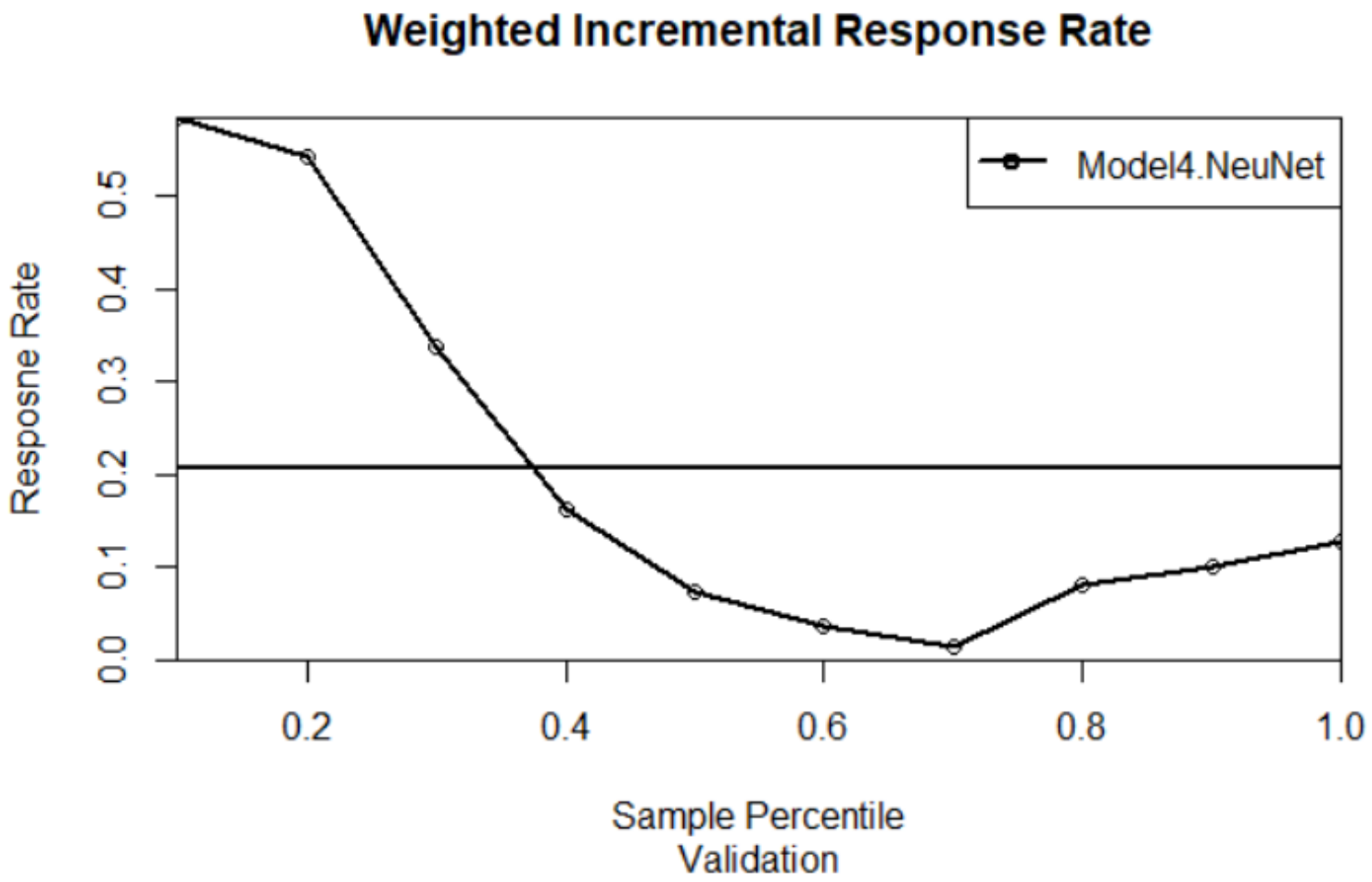
Validation



SFU BEEDIE SCHOOL OF BUSINESS

INCREMENTAL LIFT

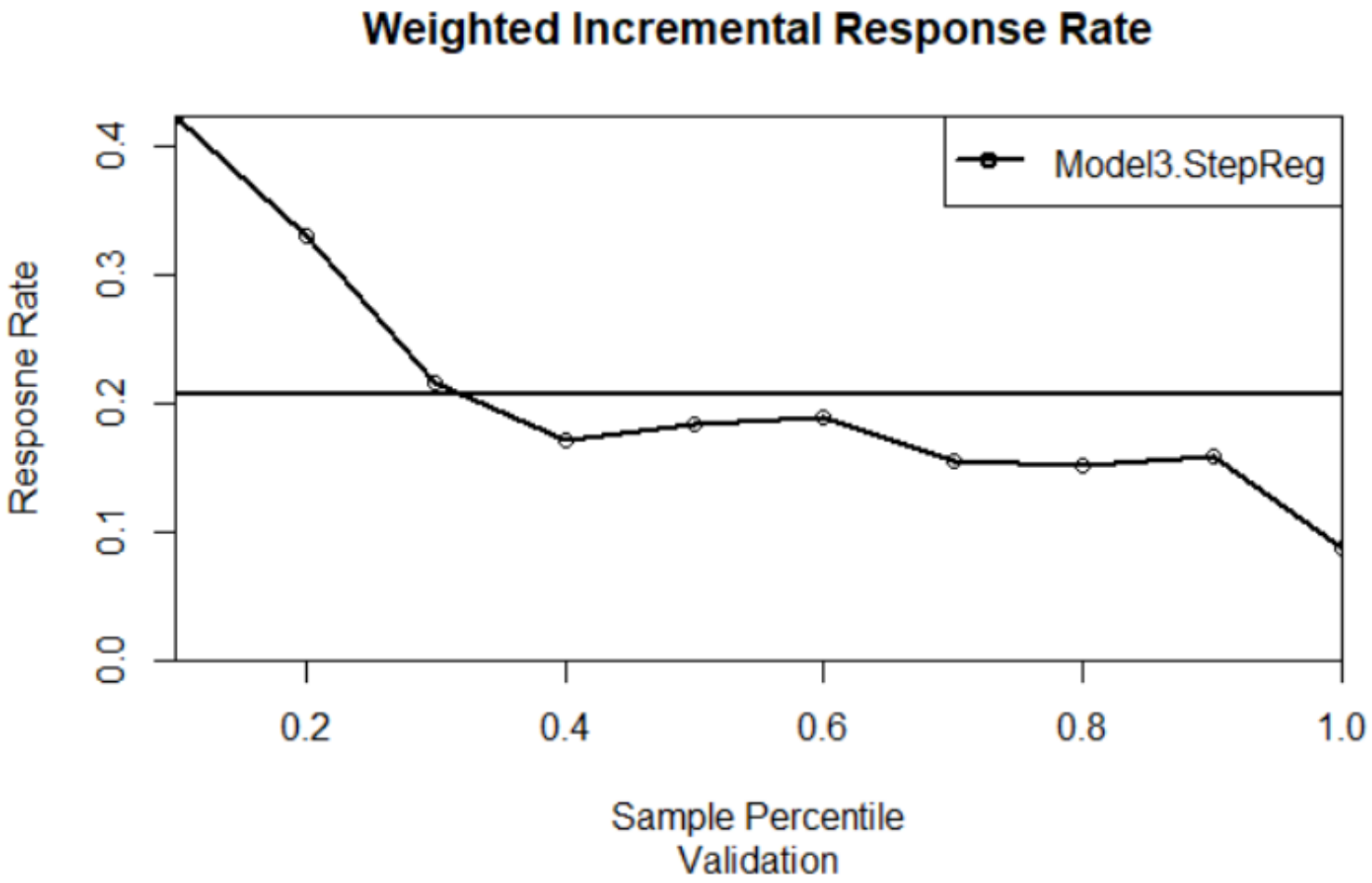
Neural Network Validation



SFU BEEDIE SCHOOL OF BUSINESS

INCREMENTAL LIFT

Stepwise Regression Validation



SFU
BEEDIE
SCHOOL OF BUSINESS

PUTTING IT ALL TOGETHER
FINANCIAL IMPACT

SFU BEEDIE SCHOOL OF BUSINESS

FINANCIAL IMPACT

Loss per order if churn

custid	avgorder	ordfreq	cust_period	Sample	score	Potential Loss
NF85M6	40.02	0	0	Holdout	0.98	$ 40.02
KL7E2H	40.02	0	0	Holdout	0.98	$ 40.02
C56P7B	40.02	0	0	Holdout	0.98	$ 40.02
9JMNAX	40.02	0	0	Holdout	0.98	$ 40.02
2C8YQT	40.02	0	0	Holdout	0.98	$ 40.02
7XP45K	14.63	0	0	Holdout	0.974	$ 14.63
P5MSYK	74.66	0.018480493	487	Holdout	0.864	$ 671.94
46V8DD	35.29	0.004413063	1133	Holdout	0.862	$ 176.45
VC3DLQ	15.95	0.578947368	19	Holdout	0.86	$ 175.45
JS7YEU	103.28	0.012820513	546	Holdout	0.858	$ 722.96
QXZWPL	203.55	0.032085561	187	Holdout	0.858	$ 1,221.30
NW87UP	58.38	0.081632653	49	Holdout	0.848	$ 233.52
8NAKCE	117.66	0.011644833	687	Holdout	0.844	$ 941.28
SAC5P6	93.87	0.022288262	673	Holdout	0.844	$ 1,408.05
FAS9W3	25.63	0.013793103	145	Holdout	0.844	$ 51.26
HTSXKU	88.04	0.033613445	119	Holdout	0.842	$ 352.16
					Total	**$ 6,169.10**

Potential Loss = Average Order x Order Frequency x Customer Period
Or if order frequency is 0
Potential Loss = Average Order