

## Performance in Class Data

### Model Section with all Subsets

```
> library(leaps)
> performance.data <- read.table("/home/rrandles/Desktop/performance.txt",
col.names=c("points", "att", "anote", "aprob", "aread", "arow", "astudy", "asleep", "atv"))

> attach(performance.data)

> performance.data

> subs <- regsubsets(points~., data=performance.data, nbest=5, method=c("exhaustive"))
> summary(subs)
Subset selection object
Call: regsubsets.formula(points ~ ., data = performance.data, nbest = 5,
  method = c("exhaustive"))
8 Variables (and intercept)
5 subsets of each size up to 8
Selection Algorithm: exhaustive

      att  anote  aprob  aread  arow  astudy  asleep  atv
1 ( 1) "*" " " " " " " " " " " " "
1 ( 2) " " "*" " " " " " " " " " "
1 ( 3) " " " " " " " " " " "*" " "
1 ( 4) " " " " " " " " " " " " "*"
1 ( 5) " " " " " " " " " " " " "*" "
2 ( 1) "*" " " "*" " " " " " " " "
2 ( 2) "*" " " " " " " " " " " "*"
2 ( 3) "*" "*" " " " " " " " " " "
2 ( 4) "*" " " " " " " " " "*" " "
2 ( 5) "*" " " " " " " " " " " "*"
3 ( 1) "*" " " "*" " " " " "*" " "
3 ( 2) "*" " " "*" " " " " " " "*"
3 ( 3) "*" "*" " " " " " " " " "*"
3 ( 4) "*" "*" " " " " " " "*" " "
3 ( 5) "*" "*" "*" " " " " " " "
4 ( 1) "*" " " "*" " " " " "*" " "*"
4 ( 2) "*" "*" " " " " " " "*" " "*"
4 ( 3) "*" "*" "*" " " " " "*" " "
4 ( 4) "*" " " "*" " " " " "*" "*" "
4 ( 5) "*" " " "*" "*" " " "*" " "
5 ( 1) "*" "*" "*" " " " " "*" " "*"
5 ( 2) "*" " " "*" "*" " " "*" " "*"
5 ( 3) "*" " " "*" " " " " "*" "*" "*"
5 ( 4) "*" " " "*" " " "*" "*" " " "*"
5 ( 5) "*" "*" "*" "*" " " "*" " " "
6 ( 1) "*" "*" "*" "*" " " "*" " "*"
6 ( 2) "*" "*" "*" " " " " "*" "*" "*"
6 ( 3) "*" "*" "*" " " "*" "*" " " "*"
6 ( 4) "*" " " "*" "*" " " "*" "*" "*"
6 ( 5) "*" " " "*" "*" "*" "*" " " "*"
7 ( 1) "*" "*" "*" "*" " " "*" "*" "*"
7 ( 2) "*" "*" "*" "*" "*" "*" " " "*"
7 ( 3) "*" "*" "*" " " "*" "*" "*" "*"
7 ( 4) "*" " " "*" "*" " " "*" "*" "*"
7 ( 5) "*" "*" "*" "*" "*" "*" "*" "
8 ( 1) "*" "*" "*" "*" "*" "*" "*" "*"

```

```

> result1 <-
with(performance.data, leaps(cbind(att, anote, apro, aread, arow, astudy, asleep, atv), points, metho
d="r2", nbest=5))
> result1$r2
 0.236954735 0.025244857 0.020646532 0.017912376 0.006492919
 0.256081848 0.248988757 0.248379527 0.244765876 0.237815609
 0.270822642 0.265954134 0.263602187 0.261099754 0.259610711
 0.281615182 0.278721964 0.276719138 0.272832043 0.272116346
 0.290739874 0.282414751 0.282024497 0.281706647 0.279951519
 0.293524833 0.290976415 0.290747991 0.282906669 0.282509173
 0.293863233 0.293542998 0.290976502 0.283112545 0.282348477
 0.293863525

> result2 <-
with(performance.data, leaps(cbind(att, anote, apro, aread, arow, astudy, asleep, atv), points, metho
d="adjr2", nbest=5))
> result2$adjr2
 [1] 0.225733481 0.010910222 0.006244275 0.003469911 -0.008117480
 [6] 0.233875336 0.226570511 0.225943095 0.222221574 0.215063836
[11] 0.237678217 0.232588413 0.230129559 0.227513379 0.225956653
[16] 0.237406886 0.234335623 0.232209547 0.228083246 0.227323505
[21] 0.235328926 0.226353404 0.225932660 0.225589979 0.223697732
[26] 0.226241483 0.223450360 0.223200181 0.214612066 0.214176713
[31] 0.214138114 0.213781723 0.210925462 0.202173638 0.201323305
[36] 0.201255463

> result3 <-
with(performance.data, leaps(cbind(att, anote, apro, aread, arow, astudy, asleep, atv), points, metho
d="Cp", nbest=5))
> result3$Cp
-0.08390181 18.20477610 18.60200498 18.83819657 19.82467291
 0.26379164 0.87653236 0.92916101 1.24132829 1.84173105
 0.99039980 1.41096858 1.61414288 1.83031697 1.95894879
 2.05808009 2.30801238 2.48102760 2.81681655 2.87864247
 3.26983842 3.98900884 4.02272122 4.05017880 4.20179653
 5.02925816 5.24940465 5.26913721 5.94651423 5.98085218
 7.00002525 7.02768897 7.24939713 7.92872958 7.99473396
 9.00000000

> n <- nrow(performance.data)
> n
[1] 70
> SSE <- (1-result1$r2)*var(points)*(n-1)
> AIC <- n*log(SSE) - n*log(n) + 2*result1$size
> AIC
334.6810 351.8218 352.1513 352.3464 353.1557
334.9039 335.5682 335.6250 335.9607 336.6020
335.5030 335.9688 336.1927 336.4302 336.5711
336.4591 336.7405 336.9346 337.3098 337.3787
337.5643 338.3812 338.4193 338.4502 338.6211
339.2889 339.5410 339.5635 340.3332 340.3720
341.2554 341.2871 341.5410 342.3131 342.3877
343.2554

> SBC <- n*log(SSE) - n*log(n) + log(n)*result1$size
> SBC
339.1780 356.3188 356.6482 356.8434 357.6526
341.6494 342.3137 342.3705 342.7062 343.3475
344.4969 344.9628 345.1867 345.4242 345.5651
347.7016 347.9830 348.1771 348.5523 348.6211
351.0553 351.8722 351.9102 351.9412 352.1120
355.0284 355.2805 355.3030 356.0727 356.1114
359.2434 359.2751 359.5289 360.3011 360.3756
363.4918

```

```

> reg1 <- lm(points~att)
> summary(reg1)
Call:
lm(formula = points ~ att)

Residuals:
    Min       1Q   Median       3Q      Max
-31.495  -8.151   2.099   6.787  22.693

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  58.931     4.276  13.781 < 2e-16 ***
att           6.188     1.347   4.595 1.93e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 10.77 on 68 degrees of freedom
Multiple R-squared: 0.237, Adjusted R-squared: 0.2257
F-statistic: 21.12 on 1 and 68 DF, p-value: 1.929e-05
-----

```

```

> reg2 <- lm(points~att+aprob)
> summary(reg2)

Call:
lm(formula = points ~ att + aprob)

Residuals:
    Min       1Q   Median       3Q      Max
-32.853  -7.185   1.038   6.816  21.647

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  59.353     4.266  13.913 < 2e-16 ***
att           6.500     1.360   4.778 1.00e-05 ***
aprob        -5.342     4.070  -1.313  0.194
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 10.71 on 67 degrees of freedom
Multiple R-squared: 0.2561, Adjusted R-squared: 0.2339
F-statistic: 11.53 on 2 and 67 DF, p-value: 4.967e-05
-----

```

```

> reg3 <- lm(points~att+aprob+astudy)
> summary(reg3)

Call:
lm(formula = points ~ att + aprob + astudy)

Residuals:
    Min       1Q   Median       3Q      Max
-31.5302  -6.5144  -0.1905   7.3234  23.9641

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  56.9069     4.7530  11.973 < 2e-16 ***
att           6.3785     1.3611   4.686 1.44e-05 ***
aprob        -6.3883     4.1598  -1.536  0.129
astudy        0.7439     0.6440   1.155  0.252
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 10.68 on 66 degrees of freedom
Multiple R-squared: 0.2708, Adjusted R-squared: 0.2377
F-statistic: 8.171 on 3 and 66 DF, p-value: 0.0001054

```

## Stepwise Selection

```
> regg <- lm(points~att+anote+aprob+aread+arow+astudy+asleep+atv)
> step(regg)
Start: AIC=343.26
points ~ att + anote + aprob + aread + arow + astudy + asleep +
      atv
```

	Df	Sum of Sq	RSS	AIC
- arow	1	0.00	7295.4	341.26
- asleep	1	3.31	7298.7	341.29
- aread	1	29.83	7325.2	341.54
- anote	1	111.07	7406.5	342.31
- atv	1	118.97	7414.4	342.39
- aprob	1	154.86	7450.3	342.73
- astudy	1	170.23	7465.6	342.87
<none>			7295.4	343.26
- att	1	2100.43	9395.8	358.97

```
Step: AIC=341.26
points ~ att + anote + aprob + aread + astudy + asleep + atv
```

	Df	Sum of Sq	RSS	AIC
- asleep	1	3.50	7298.9	339.29
- aread	1	29.82	7325.2	339.54
- anote	1	113.20	7408.6	340.33
- atv	1	121.03	7416.4	340.41
- aprob	1	155.21	7450.6	340.73
- astudy	1	170.31	7465.7	340.87
<none>			7295.4	341.26
- att	1	2116.79	9412.2	357.09

```
Step: AIC=339.29
points ~ att + anote + aprob + aread + astudy + atv
```

	Df	Sum of Sq	RSS	AIC
- aread	1	28.77	7327.7	337.56
- anote	1	114.78	7413.7	338.38
- atv	1	140.23	7439.1	338.62
- aprob	1	152.48	7451.4	338.74
- astudy	1	167.21	7466.1	338.87
<none>			7298.9	339.29
- att	1	2141.08	9440.0	355.30

```
Step: AIC=337.56
points ~ att + anote + aprob + astudy + atv
```

	Df	Sum of Sq	RSS	AIC
- anote	1	94.27	7422.0	336.46
- aprob	1	124.16	7451.8	336.74
- atv	1	144.85	7472.5	336.93
- astudy	1	195.84	7523.5	337.41
<none>			7327.7	337.56
- att	1	2133.75	9461.4	353.45

```
Step: AIC=336.46
points ~ att + aprob + astudy + atv
```

	Df	Sum of Sq	RSS	AIC
- atv	1	111.50	7533.5	335.50
- astudy	1	161.80	7583.8	335.97
<none>			7422.0	336.46
- aprob	1	245.27	7667.2	336.74
- att	1	2427.46	9849.4	354.27

Step: AIC=335.5  
points ~ att + aprob + astudy

	Df	Sum of Sq	RSS	AIC
- astudy	1	152.29	7685.7	334.90
<none>			7533.5	335.50
- aprob	1	269.20	7802.7	335.96
- att	1	2506.64	10040.1	353.61

Step: AIC=334.9  
points ~ att + aprob

	Df	Sum of Sq	RSS	AIC
- aprob	1	197.61	7883.4	334.68
<none>			7685.7	334.90
- att	1	2618.79	10304.5	353.43

Step: AIC=334.68  
points ~ att

	Df	Sum of Sq	RSS	AIC
<none>			7883.4	334.68
- att	1	2448.1	10331.4	351.61

Call:  
lm(formula = points ~ att)

Coefficients:  
(Intercept)           att  
      58.931            6.188

### Validation

```
> model <- glm(points~att)
> MSPRESS <- cv.glm(performance.data, model)$delta[1]
> MSPRESS
      1
118.338
> n <- nrow(performance.data)
> n
[1] 70
> PRESS <- n*MSPRESS
> PRESS
      1
8283.642
```