

## Part2.R

Cheng Jun

2021-03-26

```
library(caTools)
## Warning: package 'caTools' was built under R version 4.0.4
library(ggplot2)
## Warning: package 'ggplot2' was built under R version 4.0.4
library(lattice)
## Warning: package 'lattice' was built under R version 4.0.4
library(DAAG)
## Warning: package 'DAAG' was built under R version 4.0.4
setwd("C:/Users/Cheng Jun/Desktop/SIM/Year 2/Machine Learning/Coursework")

data1 <- read.table("student-mat.csv",header=1,sep = ';')
data2 <- read.table("student-por.csv",header=1,sep = ';')

#math
sum(is.na(data1))

## [1] 0

columns = names(data1)
columns

## [1] "school"      "sex"         "age"         "address"     "famsize"
## [6] "Pstatus"    "Medu"        "Fedu"        "Mjob"        "Fjob"
## [11] "reason"     "guardian"    "traveltime"  "studytime"   "failures"
## [16] "schoolsup"  "famsup"      "paid"        "activities"  "nursery"
## [21] "higher"     "internet"    "romantic"    "famrel"      "freetime"
## [26] "goout"      "Dalc"        "Walc"        "health"      "absences"
## [31] "G1"         "G2"          "G3"
```

```
data1 <- transform(
  data1,
  school = as.integer(factor(school)),
  sex = as.integer(factor(sex)),
  address = as.integer(factor(address)),
```

```

famsize = as.integer(factor(famsize)),
Pstatus = as.integer(factor(Pstatus)),
Medu = as.integer(factor(Medu)),
Fedu = as.integer(factor(Fedu)),
Mjob = as.integer(factor(Mjob)),
Fjob = as.integer(factor(Fjob)),
reason = as.integer(factor(reason)),
guardian = as.integer(factor(guardian)),
schoolsup= as.integer(factor(schoolsup)),
famsup = as.integer(factor(famsup)),
paid = as.integer(factor(paid)),
activities = as.integer(factor(activities)),
nursery = as.integer(factor(nursery)),
higher = as.integer(factor(higher)),
internet = as.integer(factor(internet)),
romantic = as.integer(factor(romantic))
)
sapply(data1, class)

##      school      sex      age      address      famsize      Pstatus
Medu
## "integer" "integer" "integer" "integer" "integer" "integer"
"integer"
##      Fedu      Mjob      Fjob      reason      guardian      traveltime
studytime
## "integer" "integer" "integer" "integer" "integer" "integer"
"integer"
##      failures      schoolsup      famsup      paid      activities      nursery
higher
## "integer" "integer" "integer" "integer" "integer" "integer"
"integer"
##      internet      romantic      famrel      freetime      goout      Dalc
Walc
## "integer" "integer" "integer" "integer" "integer" "integer"
"integer"
##      health      absences      G1      G2      G3
## "integer" "integer" "integer" "integer" "integer"

data1 = data1[,-(31:32)]
data1[,c(1:30)] = scale(data1[,c(1:30)])
data1

##      school      sex      age      address      famsize      Pstatus
Medu
## 1 -0.362590 -0.9469754 1.0217506 0.5347142 -0.6361343 -2.9346700
1.1424068
## 2 -0.362590 -0.9469754 0.2380778 0.5347142 -0.6361343 0.3398912 -
1.5979820
## 3 -0.362590 -0.9469754 -1.3292678 0.5347142 1.5680153 0.3398912 -

```

```

set.seed(16)
#train test split
split.math = sample.split(data1$G3,SplitRatio=0.7)
training.math = subset(data1,split.math==TRUE)
test.math = subset(data1,split.math==FALSE)

#Linear Regression for math
linearMod.math <- lm(G3 ~., data = training.math)
modelSum.math <- summary(linearMod.math) # capture model summary as an
object
modelCoef.math <- modelSum.math$coefficients # model coefficients
modelCoef.math

```

	Estimate	Std. Error	t value	Pr(> t )
## (Intercept)	10.521711376	0.2554291	41.19228954	2.205002e-112
## school	0.086787725	0.3006868	0.28863165	7.731063e-01
## sex	0.593859727	0.3018485	1.96740999	5.025957e-02
## age	-0.495244733	0.3256420	-1.52082562	1.295879e-01
## address	0.474315867	0.2822384	1.68055067	9.411951e-02
## famsize	0.207622122	0.2832284	0.73305550	4.642225e-01
## Pstatus	-0.504362709	0.2789108	-1.80832997	7.177626e-02
## Medu	0.410734750	0.3826972	1.07326299	2.842048e-01
## Fedu	0.142255141	0.3452627	0.41202005	6.806835e-01
## Mjob	-0.558928755	0.2985635	-1.87205996	6.238481e-02
## Fjob	0.079236777	0.2585967	0.30641057	7.595512e-01
## reason	0.412550936	0.2692188	1.53239996	1.267088e-01
## guardian	0.262041665	0.2876055	0.91111478	3.631268e-01
## traveltime	0.010681583	0.2722257	0.03923797	9.687325e-01
## studytime	0.165438406	0.2869226	0.57659584	5.647397e-01
## failures	-1.140604810	0.2838132	-4.01885733	7.777868e-05
## schoolsup	-0.446498741	0.2858829	-1.56182389	1.196150e-01
## famsup	-0.338890709	0.2811037	-1.20557170	2.291410e-01
## paid	0.315804634	0.2887744	1.09360327	2.751987e-01
## activities	0.138529674	0.2702588	0.51258160	6.087040e-01
## nursery	-0.009249268	0.2629081	-0.03518061	9.719643e-01
## higher	0.438221054	0.2491118	1.75913395	7.979794e-02
## internet	-0.009936859	0.2756299	-0.03605145	9.712706e-01
## romantic	-0.482355022	0.2765607	-1.74412009	8.238721e-02
## famrel	0.021937602	0.2718533	0.08069646	9.357490e-01
## freetime	0.342519919	0.2736725	1.25156883	2.119158e-01
## goout	-0.574626804	0.2944829	-1.95130797	5.215549e-02
## Dalc	-0.031952422	0.3519540	-0.09078580	9.277367e-01
## Walc	0.375664109	0.3710454	1.01244781	3.123188e-01
## health	-0.158835756	0.2677130	-0.59330613	5.535214e-01
## absences	0.312878919	0.2735074	1.14395049	2.537561e-01

```
summary(modelCoef.math)
```

	Estimate	Std. Error	t value	Pr(> t )
## Min.	:-1.14061	Min. :0.2491	Min. : -4.0189	Min. :0.00000

```
## 1st Qu.: -0.24886 1st Qu.: 0.2720 1st Qu.: -0.8994 1st Qu.: 0.08825
## Median : 0.08679 Median : 0.2822 Median : 0.3064 Median : 0.27520
## Mean : 0.34069 Mean : 0.2906 Mean : 1.3267 Mean : 0.38640
## 3rd Qu.: 0.32916 3rd Qu.: 0.2965 3rd Qu.: 1.0834 3rd Qu.: 0.64469
## Max. : 10.52171 Max. : 0.3827 Max. : 41.1923 Max. : 0.97196
```

```
Stepwise.complete.math <- step(linearMod.math)
```

```
## Start: AIC=822.19
```

```
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
## Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
## failures + schoolsup + famsup + paid + activities + nursery +
## higher + internet + romantic + famrel + freetime + goout +
## Dalc + Walc + health + absences
```

```
##
##           Df Sum of Sq  RSS   AIC
## - nursery    1    0.022 4308.8 820.20
## - internet    1    0.023 4308.8 820.20
## - traveltime  1    0.027 4308.8 820.20
## - famrel      1    0.114 4308.9 820.20
## - Dalc        1    0.144 4308.9 820.20
## - school      1    1.459 4310.2 820.29
## - Fjob        1    1.644 4310.4 820.30
## - Fedu        1    2.973 4311.7 820.39
## - activities  1    4.602 4313.3 820.49
## - studytime   1    5.823 4314.6 820.57
## - health      1    6.166 4314.9 820.59
## - famsize     1    9.412 4318.2 820.80
## - guardian    1   14.540 4323.3 821.13
## - Walc        1   17.954 4326.7 821.35
## - Medu        1   20.176 4328.9 821.49
## - paid        1   20.948 4329.7 821.54
## - absences    1   22.921 4331.7 821.66
## - famsup      1   25.457 4334.2 821.83
## - freetime    1   27.436 4336.2 821.95
## <none>                4308.7 822.19
## - age          1   40.511 4349.3 822.79
## - reason       1   41.130 4349.9 822.83
## - schoolsup    1   42.725 4351.5 822.93
## - address      1   49.467 4358.2 823.36
## - romantic     1   53.281 4362.0 823.60
## - higher       1   54.202 4362.9 823.66
## - Pstatus      1   57.276 4366.0 823.85
## - Mjob         1   61.384 4370.1 824.11
## - goout        1   66.691 4375.4 824.45
## - sex          1   67.796 4376.5 824.52
## - failures     1  282.892 4591.6 837.81
```

```
##
```

```
## Step: AIC=820.2
```

```
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
```

```

##      Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
##      failures + schoolsup + famsup + paid + activities + higher +
##      internet + romantic + famrel + freetime + goout + Dalc +
##      Walc + health + absences
##
##              Df Sum of Sq      RSS      AIC
## - internet    1      0.021 4308.8 818.20
## - traveltime  1      0.024 4308.8 818.20
## - famrel       1      0.119 4308.9 818.20
## - Dalc         1      0.138 4308.9 818.20
## - school       1      1.472 4310.2 818.29
## - Fjob         1      1.653 4310.4 818.30
## - Fedu         1      2.955 4311.7 818.39
## - activities   1      4.599 4313.4 818.49
## - studytime    1      5.802 4314.6 818.57
## - health       1      6.167 4314.9 818.59
## - famsize      1      9.404 4318.2 818.80
## - guardian     1     14.760 4323.5 819.14
## - Walc         1     18.116 4326.9 819.36
## - Medu         1     20.164 4328.9 819.49
## - paid         1     20.956 4329.7 819.54
## - absences     1     22.899 4331.7 819.66
## - famsup       1     25.442 4334.2 819.83
## - freetime     1     27.420 4336.2 819.95
## <none>                4308.8 820.20
## - age          1     40.490 4349.3 820.79
## - reason       1     41.114 4349.9 820.83
## - schoolsup    1     42.703 4351.5 820.93
## - address      1     49.467 4358.2 821.36
## - romantic     1     53.658 4362.4 821.62
## - higher       1     54.237 4363.0 821.66
## - Pstatus      1     57.515 4366.3 821.87
## - Mjob         1     61.567 4370.3 822.13
## - goout        1     66.868 4375.6 822.46
## - sex          1     67.817 4376.6 822.52
## - failures     1    283.254 4592.0 835.83
##
## Step:  AIC=818.2
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
##      failures + schoolsup + famsup + paid + activities + higher +
##      romantic + famrel + freetime + goout + Dalc + Walc + health +
##      absences
##
##              Df Sum of Sq      RSS      AIC
## - traveltime  1      0.026 4308.8 816.20
## - famrel       1      0.122 4308.9 816.21
## - Dalc         1      0.141 4308.9 816.21
## - school       1      1.510 4310.3 816.29
## - Fjob         1      1.652 4310.4 816.30

```

```

## - Fedu      1      2.935 4311.7 816.39
## - activities 1      4.659 4313.4 816.50
## - studytime 1      5.787 4314.6 816.57
## - health    1      6.147 4314.9 816.59
## - famsize   1      9.397 4318.2 816.80
## - guardian  1     14.805 4323.6 817.15
## - Walc      1     18.097 4326.9 817.36
## - Medu      1     20.186 4329.0 817.49
## - paid      1     20.963 4329.8 817.54
## - absences  1     22.919 4331.7 817.67
## - famsup    1     25.528 4334.3 817.83
## - freetime  1     27.497 4336.3 817.96
## <none>      4308.8 818.20
## - age       1     40.576 4349.4 818.79
## - reason    1     41.179 4350.0 818.83
## - schoolsup  1     42.695 4351.5 818.93
## - address   1     50.301 4359.1 819.41
## - higher    1     54.473 4363.3 819.68
## - romantic  1     54.731 4363.5 819.69
## - Pstatus   1     59.036 4367.8 819.97
## - Mjob      1     63.433 4372.2 820.25
## - goout     1     67.682 4376.5 820.51
## - sex       1     67.927 4376.7 820.53
## - failures  1    283.411 4592.2 833.84
##
## Step: AIC=816.2
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + Fjob + reason + guardian + studytime + failures +
##      schoolsup + famsup + paid + activities + higher + romantic +
##      famrel + freetime + goout + Dalc + Walc + health + absences
##
##           Df Sum of Sq   RSS   AIC
## - famrel   1      0.122 4308.9 814.21
## - Dalc     1      0.134 4308.9 814.21
## - school   1      1.587 4310.4 814.30
## - Fjob     1      1.686 4310.5 814.31
## - Fedu     1      2.918 4311.7 814.39
## - activities 1      4.680 4313.5 814.50
## - studytime 1      5.763 4314.6 814.57
## - health   1      6.181 4315.0 814.60
## - famsize  1      9.551 4318.4 814.81
## - guardian 1     14.780 4323.6 815.15
## - Walc     1     18.231 4327.0 815.37
## - Medu     1     20.161 4329.0 815.49
## - paid     1     21.186 4330.0 815.56
## - absences 1     22.915 4331.7 815.67
## - famsup   1     25.934 4334.7 815.86
## - freetime 1     27.633 4336.4 815.97
## <none>     4308.8 816.20
## - age     1     40.917 4349.7 816.82

```

```

## - reason      1      41.235 4350.0 816.84
## - schoolsup   1      42.730 4351.5 816.93
## - address     1      53.962 4362.8 817.65
## - higher      1      54.457 4363.3 817.68
## - romantic    1      54.784 4363.6 817.70
## - Pstatus     1      59.020 4367.8 817.97
## - Mjob        1      63.476 4372.3 818.25
## - goout       1      67.663 4376.5 818.51
## - sex         1      67.901 4376.7 818.53
## - failures    1     284.410 4593.2 831.90
##
## Step:  AIC=814.21
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + Fjob + reason + guardian + studytime + failures +
##      schoolsup + famsup + paid + activities + higher + romantic +
##      freetime + goout + Dalc + Walc + health + absences
##
##              Df Sum of Sq    RSS    AIC
## - Dalc        1      0.127 4309.1 812.21
## - school      1      1.502 4310.4 812.30
## - Fjob        1      1.745 4310.7 812.32
## - Fedu        1      2.990 4311.9 812.40
## - activities  1      4.654 4313.6 812.51
## - studytime   1      5.776 4314.7 812.58
## - health      1      6.065 4315.0 812.60
## - famsize     1      9.689 4318.6 812.83
## - guardian    1     14.917 4323.9 813.16
## - Walc        1     18.113 4327.0 813.37
## - Medu        1     20.039 4329.0 813.49
## - paid        1     21.115 4330.1 813.56
## - absences    1     22.837 4331.8 813.67
## - famsup      1     25.850 4334.8 813.86
## - freetime    1     28.818 4337.8 814.05
## <none>                4308.9 814.21
## - age        1     41.140 4350.1 814.84
## - reason     1     41.309 4350.2 814.85
## - schoolsup   1     42.833 4351.8 814.95
## - address     1     54.326 4363.3 815.68
## - higher      1     54.613 4363.5 815.70
## - romantic    1     56.195 4365.1 815.80
## - Pstatus     1     59.129 4368.1 815.98
## - Mjob        1     64.016 4373.0 816.29
## - goout       1     67.671 4376.6 816.52
## - sex         1     68.174 4377.1 816.56
## - failures    1    289.887 4598.8 830.24
##
## Step:  AIC=812.21
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + Fjob + reason + guardian + studytime + failures +
##      schoolsup + famsup + paid + activities + higher + romantic +

```

```

##      freetime + goout + Walc + health + absences
##
##      Df Sum of Sq    RSS    AIC
## - school      1      1.432 4310.5 810.31
## - Fjob        1      1.700 4310.8 810.32
## - Fedu        1      3.144 4312.2 810.42
## - activities  1      4.894 4314.0 810.53
## - studytime   1      5.869 4314.9 810.59
## - health      1      6.080 4315.1 810.61
## - famsize     1      9.575 4318.6 810.83
## - guardian    1     15.388 4324.5 811.20
## - Medu        1     19.931 4329.0 811.49
## - paid        1     21.007 4330.1 811.56
## - Walc        1     22.701 4331.8 811.67
## - absences    1     22.711 4331.8 811.67
## - famsup      1     26.091 4335.2 811.89
## - freetime    1     29.140 4338.2 812.08
## <none>                4309.1 812.21
## - reason      1     41.418 4350.5 812.86
## - age         1     42.074 4351.1 812.91
## - schoolsup   1     43.867 4352.9 813.02
## - higher      1     54.512 4363.6 813.70
## - address     1     54.751 4363.8 813.71
## - romantic    1     57.411 4366.5 813.88
## - Pstatus     1     59.039 4368.1 813.98
## - Mjob        1     64.784 4373.8 814.35
## - goout       1     67.613 4376.7 814.53
## - sex         1     68.331 4377.4 814.57
## - failures    1    290.000 4599.1 828.26
##
## Step:  AIC=810.31
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Fedu +
##      Mjob + Fjob + reason + guardian + studytime + failures +
##      schoolsup + famsup + paid + activities + higher + romantic +
##      freetime + goout + Walc + health + absences
##
##      Df Sum of Sq    RSS    AIC
## - Fjob      1      1.706 4312.2 808.42
## - Fedu      1      3.281 4313.8 808.52
## - activities 1      4.510 4315.0 808.60
## - studytime  1      5.230 4315.7 808.64
## - health     1      6.173 4316.7 808.70
## - famsize    1     10.280 4320.8 808.97
## - guardian   1     14.604 4325.1 809.24
## - Medu       1     19.699 4330.2 809.57
## - paid       1     21.376 4331.9 809.68
## - absences   1     21.466 4332.0 809.68
## - Walc       1     23.006 4333.5 809.78
## - famsup     1     28.293 4338.8 810.12
## - freetime   1     29.872 4340.4 810.22

```



```

## <none>                                4310.5 810.31
## - reason      1      40.439 4350.9 810.89
## - age         1      42.170 4352.7 811.00
## - schoolsup   1      44.098 4354.6 811.13
## - address     1      53.934 4364.4 811.75
## - higher      1      56.240 4366.7 811.90
## - romantic    1      56.811 4367.3 811.93
## - Pstatus     1      58.438 4368.9 812.04
## - Mjob        1      64.066 4374.6 812.39
## - sex         1      67.042 4377.5 812.58
## - goout       1      69.013 4379.5 812.71
## - failures    1     294.200 4604.7 826.60
##
## Step: AIC=808.42
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Fedu +
##      Mjob + reason + guardian + studytime + failures + schoolsup +
##      famsup + paid + activities + higher + romantic + freetime +
##      goout + Walc + health + absences
##
##              Df Sum of Sq    RSS    AIC
## - Fedu        1      3.906 4316.1 806.67
## - activities   1      4.619 4316.8 806.71
## - studytime    1      4.986 4317.2 806.74
## - health       1      6.404 4318.6 806.83
## - famsize      1      9.426 4321.6 807.02
## - guardian     1     14.055 4326.3 807.32
## - Medu         1     19.546 4331.7 807.67
## - paid         1     20.991 4333.2 807.76
## - absences     1     21.202 4333.4 807.78
## - Walc         1     24.295 4336.5 807.97
## - freetime     1     28.832 4341.0 808.26
## - famsup       1     29.256 4341.5 808.29
## <none>                                4312.2 808.42
## - reason      1     40.063 4352.3 808.98
## - age         1     43.524 4355.7 809.20
## - schoolsup    1     44.718 4356.9 809.27
## - address     1     53.321 4365.5 809.82
## - higher      1     55.113 4367.3 809.93
## - romantic    1     56.707 4368.9 810.04
## - Pstatus     1     57.461 4369.7 810.08
## - Mjob        1     62.366 4374.6 810.39
## - sex         1     66.835 4379.0 810.68
## - goout       1     69.222 4381.4 810.83
## - failures    1    293.325 4605.5 824.65
##
## Step: AIC=806.67
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Mjob +
##      reason + guardian + studytime + failures + schoolsup + famsup +
##      paid + activities + higher + romantic + freetime + goout +
##      Walc + health + absences

```

```

##
##          Df Sum of Sq    RSS    AIC
## - studytime  1      4.344 4320.5 804.95
## - activities  1      5.082 4321.2 804.99
## - health     1      5.833 4321.9 805.04
## - famsize    1      9.519 4325.6 805.28
## - guardian   1     11.397 4327.5 805.40
## - paid       1     20.304 4336.4 805.97
## - absences   1     21.135 4337.2 806.02
## - Walc       1     24.504 4340.6 806.24
## - famsup     1     27.198 4343.3 806.41
## - freetime   1     28.866 4345.0 806.51
## <none>                4316.1 806.67
## - reason     1     39.656 4355.8 807.20
## - age        1     41.885 4358.0 807.34
## - Medu       1     42.975 4359.1 807.41
## - schoolsup   1     43.243 4359.3 807.43
## - address    1     53.025 4369.1 808.05
## - romantic   1     54.787 4370.9 808.16
## - higher     1     58.157 4374.3 808.37
## - Pstatus    1     60.875 4377.0 808.55
## - Mjob       1     62.268 4378.4 808.64
## - sex        1     64.342 4380.4 808.77
## - goout      1     69.185 4385.3 809.07
## - failures   1    308.378 4624.5 823.78
##
## Step:  AIC=804.95
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Mjob +
##       reason + guardian + failures + schoolsup + famsup + paid +
##       activities + higher + romantic + freetime + goout + Walc +
##       health + absences
##
##          Df Sum of Sq    RSS    AIC
## - activities  1      6.20 4326.6 803.34
## - health     1      6.43 4326.9 803.36
## - famsize    1      9.16 4329.6 803.53
## - guardian   1     11.41 4331.9 803.68
## - absences   1     20.01 4340.5 804.23
## - Walc       1     21.63 4342.1 804.33
## - paid       1     22.55 4343.0 804.39
## - famsup     1     26.96 4347.4 804.67
## - freetime   1     27.36 4347.8 804.69
## <none>                4320.5 804.95
## - age        1     40.31 4360.8 805.52
## - reason     1     41.20 4361.7 805.57
## - Medu       1     42.54 4363.0 805.66
## - schoolsup   1     42.65 4363.1 805.67
## - address    1     52.02 4372.5 806.26
## - romantic   1     53.03 4373.5 806.33
## - sex        1     60.33 4380.8 806.79

```

```

## - Pstatus      1      61.47 4381.9 806.86
## - higher       1      61.60 4382.1 806.87
## - Mjob         1      62.58 4383.0 806.93
## - goout        1      67.07 4387.5 807.21
## - failures     1     318.69 4639.1 822.66
##
## Step: AIC=803.34
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Mjob +
##      reason + guardian + failures + schoolsup + famsup + paid +
##      higher + romantic + freetime + goout + Walc + health + absences
##
##              Df Sum of Sq    RSS    AIC
## - health      1       7.18 4333.8 801.80
## - famsize     1       9.42 4336.1 801.95
## - guardian    1      10.50 4337.1 802.01
## - absences    1      19.12 4345.8 802.56
## - Walc        1      20.12 4346.8 802.63
## - paid        1      21.48 4348.1 802.71
## - famsup      1      27.14 4353.8 803.08
## - freetime    1      29.17 4355.8 803.20
## <none>                4326.6 803.34
## - age         1      41.83 4368.5 804.01
## - schoolsup    1      42.87 4369.5 804.07
## - Medu        1      43.29 4369.9 804.10
## - reason      1      44.45 4371.1 804.17
## - address     1      48.88 4375.5 804.45
## - romantic    1      51.25 4377.9 804.61
## - Pstatus     1      56.79 4383.4 804.95
## - Mjob        1      60.98 4387.6 805.22
## - sex         1      64.23 4390.9 805.42
## - goout       1      64.78 4391.4 805.46
## - higher      1      66.55 4393.2 805.57
## - failures    1     324.29 4650.9 821.36
##
## Step: AIC=801.8
## G3 ~ sex + age + address + famsize + Pstatus + Medu + Mjob +
##      reason + guardian + failures + schoolsup + famsup + paid +
##      higher + romantic + freetime + goout + Walc + absences
##
##              Df Sum of Sq    RSS    AIC
## - famsize     1      10.50 4344.3 800.47
## - guardian    1      11.51 4345.3 800.54
## - Walc        1      17.50 4351.3 800.92
## - absences    1      20.55 4354.4 801.11
## - paid        1      24.18 4358.0 801.34
## - freetime    1      28.68 4362.5 801.63
## - famsup      1      29.33 4363.2 801.67
## <none>                4333.8 801.80
## - age         1      38.56 4372.4 802.26
## - schoolsup    1      40.57 4374.4 802.38

```

```

## - Medu      1      44.49 4378.3 802.63
## - reason    1      50.45 4384.3 803.01
## - address   1      52.98 4386.8 803.17
## - romantic  1      53.19 4387.0 803.18
## - Pstatus   1      57.55 4391.4 803.46
## - goout     1      62.36 4396.2 803.76
## - sex       1      63.40 4397.2 803.82
## - higher    1      64.83 4398.6 803.91
## - Mjob      1      65.39 4399.2 803.95
## - failures  1     334.95 4668.8 820.42
##
## Step: AIC=800.47
## G3 ~ sex + age + address + Pstatus + Medu + Mjob + reason + guardian +
##      failures + schoolsup + famsup + paid + higher + romantic +
##      freetime + goout + Walc + absences
##
##              Df Sum of Sq    RSS    AIC
## - guardian    1      10.87 4355.2 799.16
## - absences     1      19.91 4364.2 799.74
## - Walc         1      21.52 4365.8 799.84
## - paid         1      26.16 4370.5 800.14
## - freetime     1      26.24 4370.6 800.14
## <none>                    4344.3 800.47
## - famsup       1      31.92 4376.2 800.50
## - age          1      38.41 4382.7 800.91
## - schoolsup    1      41.27 4385.6 801.09
## - Medu         1      43.09 4387.4 801.21
## - romantic     1      48.96 4393.3 801.58
## - reason       1      49.76 4394.1 801.63
## - address      1      56.16 4400.5 802.03
## - Mjob         1      64.24 4408.6 802.54
## - higher       1      64.58 4408.9 802.56
## - goout        1      65.78 4410.1 802.64
## - sex          1      67.56 4411.9 802.75
## - Pstatus      1      73.23 4417.6 803.10
## - failures     1     334.37 4678.7 819.01
##
## Step: AIC=799.16
## G3 ~ sex + age + address + Pstatus + Medu + Mjob + reason + failures +
##      schoolsup + famsup + paid + higher + romantic + freetime +
##      goout + Walc + absences
##
##              Df Sum of Sq    RSS    AIC
## - Walc         1      17.05 4372.2 798.25
## - absences     1      24.31 4379.5 798.71
## - paid         1      27.57 4382.8 798.91
## - freetime     1      28.76 4384.0 798.99
## - famsup       1      30.33 4385.5 799.09
## - age          1      30.77 4386.0 799.12
## <none>                    4355.2 799.16

```

```

## - schoolsup 1      41.65 4396.8 799.80
## - Medu      1      43.28 4398.5 799.90
## - romantic  1      47.08 4402.3 800.14
## - reason    1      49.50 4404.7 800.30
## - address   1      56.09 4411.3 800.71
## - Mjob      1      62.70 4417.9 801.12
## - goout     1      63.98 4419.2 801.20
## - sex       1      67.26 4422.5 801.41
## - higher    1      68.95 4424.1 801.52
## - Pstatus   1      80.79 4436.0 802.26
## - failures  1     324.68 4679.9 817.08
##
## Step:  AIC=798.25
## G3 ~ sex + age + address + Pstatus + Medu + Mjob + reason + failures +
##      schoolsup + famsup + paid + higher + romantic + freetime +
##      goout + absences
##
##           Df Sum of Sq    RSS    AIC
## - freetime  1     28.575 4400.8 798.05
## - age       1     29.260 4401.5 798.09
## - absences  1     31.142 4403.4 798.21
## <none>                        4372.2 798.25
## - famsup    1     34.581 4406.8 798.43
## - paid      1     37.072 4409.3 798.59
## - Medu      1     40.464 4412.7 798.80
## - schoolsup  1     44.957 4417.2 799.08
## - goout     1     47.689 4419.9 799.25
## - reason    1     49.096 4421.3 799.34
## - address   1     49.417 4421.7 799.36
## - romantic  1     49.743 4422.0 799.38
## - Mjob      1     64.571 4436.8 800.31
## - higher    1     67.208 4439.5 800.47
## - Pstatus   1     80.293 4452.5 801.29
## - sex       1     96.472 4468.7 802.29
## - failures  1    312.483 4684.7 815.37
##
## Step:  AIC=798.05
## G3 ~ sex + age + address + Pstatus + Medu + Mjob + reason + failures +
##      schoolsup + famsup + paid + higher + romantic + goout + absences
##
##           Df Sum of Sq    RSS    AIC
## - absences  1     26.527 4427.4 797.72
## - famsup    1     31.638 4432.5 798.04
## <none>                        4400.8 798.05
## - age       1     32.151 4433.0 798.07
## - paid      1     33.456 4434.3 798.15
## - goout     1     35.552 4436.4 798.28
## - Medu      1     36.087 4436.9 798.31
## - romantic  1     46.123 4446.9 798.94
## - reason    1     46.605 4447.4 798.97

```

```

## - schoolsup 1 46.804 4447.6 798.98
## - address 1 49.478 4450.3 799.15
## - Mjob 1 56.252 4457.1 799.57
## - higher 1 66.993 4467.8 800.24
## - Pstatus 1 80.352 4481.2 801.06
## - sex 1 117.084 4517.9 803.33
## - failures 1 302.469 4703.3 814.46
##
## Step: AIC=797.72
## G3 ~ sex + age + address + Pstatus + Medu + Mjob + reason + failures +
## schoolsup + famsup + paid + higher + romantic + goout
##
## Df Sum of Sq RSS AIC
## - age 1 23.024 4450.4 797.15
## - paid 1 30.119 4457.5 797.59
## - famsup 1 30.491 4457.8 797.62
## <none> 4427.4 797.72
## - goout 1 34.013 4461.4 797.84
## - romantic 1 38.136 4465.5 798.09
## - schoolsup 1 44.988 4472.3 798.52
## - Medu 1 45.097 4472.4 798.52
## - address 1 46.717 4474.1 798.62
## - Mjob 1 53.409 4480.8 799.04
## - reason 1 56.707 4484.1 799.24
## - higher 1 64.146 4491.5 799.70
## - Pstatus 1 93.181 4520.5 801.49
## - sex 1 109.120 4536.5 802.46
## - failures 1 294.403 4721.8 813.55
##
## Step: AIC=797.15
## G3 ~ sex + address + Pstatus + Medu + Mjob + reason + failures +
## schoolsup + famsup + paid + higher + romantic + goout
##
## Df Sum of Sq RSS AIC
## - famsup 1 25.19 4475.6 796.72
## - paid 1 29.13 4479.5 796.96
## - schoolsup 1 30.21 4480.6 797.03
## <none> 4450.4 797.15
## - goout 1 37.78 4488.2 797.50
## - romantic 1 44.08 4494.5 797.88
## - reason 1 54.04 4504.4 798.50
## - Medu 1 55.81 4506.2 798.60
## - address 1 56.81 4507.2 798.67
## - Mjob 1 57.66 4508.0 798.72
## - higher 1 75.10 4525.5 799.79
## - Pstatus 1 92.65 4543.0 800.86
## - sex 1 120.82 4571.2 802.57
## - failures 1 338.60 4789.0 815.47
##
## Step: AIC=796.72

```

```
## G3 ~ sex + address + Pstatus + Medu + Mjob + reason + failures +  
## schoolsup + paid + higher + romantic + goout  
##
```

	Df	Sum of Sq	RSS	AIC
## - paid	1	18.45	4494.0	795.86
## <none>			4475.6	796.72
## - schoolsup	1	36.52	4512.1	796.97
## - goout	1	39.19	4514.8	797.13
## - romantic	1	46.56	4522.1	797.58
## - reason	1	48.66	4524.2	797.71
## - Medu	1	48.76	4524.3	797.72
## - Mjob	1	59.17	4534.7	798.35
## - address	1	60.84	4536.4	798.46
## - higher	1	72.21	4547.8	799.15
## - Pstatus	1	94.73	4570.3	800.52
## - sex	1	129.30	4604.9	802.61
## - failures	1	339.17	4814.7	814.95

```
##
```

```
## Step: AIC=795.86
```

```
## G3 ~ sex + address + Pstatus + Medu + Mjob + reason + failures +  
## schoolsup + higher + romantic + goout  
##
```

	Df	Sum of Sq	RSS	AIC
## <none>			4494.0	795.86
## - goout	1	36.54	4530.6	796.10
## - schoolsup	1	36.72	4530.7	796.11
## - romantic	1	45.11	4539.1	796.62
## - Medu	1	52.78	4546.8	797.09
## - reason	1	54.75	4548.8	797.21
## - Mjob	1	58.88	4552.9	797.46
## - address	1	62.68	4556.7	797.69
## - higher	1	84.82	4578.8	799.04
## - Pstatus	1	89.35	4583.4	799.31
## - sex	1	118.84	4612.9	801.09
## - failures	1	356.39	4850.4	815.00

```
#stepwise regression
```

```
summary(Stepwise.complete.math)
```

```
##
```

```
## Call:
```

```
## lm(formula = G3 ~ sex + address + Pstatus + Medu + Mjob + reason +  
## failures + schoolsup + higher + romantic + goout, data =  
training.math)
```

```
##
```

```
## Residuals:
```

	Min	1Q	Median	3Q	Max
##	-11.1255	-1.8651	0.4852	2.6638	8.7366

```
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.5096    0.2487  42.253 < 2e-16 ***
## sex         0.6970    0.2633   2.647  0.0086 **
## address     0.4713    0.2451   1.923  0.0556 .
## Pstatus    -0.5823    0.2537  -2.295  0.0225 *
## Medu       0.5338    0.3026   1.764  0.0789 .
## Mjob      -0.5234    0.2809  -1.863  0.0635 .
## reason     0.4577    0.2548   1.797  0.0735 .
## failures   -1.1939    0.2604  -4.584 7.03e-06 ***
## schoolsup  -0.3853    0.2618  -1.471  0.1424
## higher     0.5229    0.2338   2.236  0.0262 *
## romantic   -0.4203    0.2577  -1.631  0.1041
## goout      -0.3701    0.2521  -1.468  0.1433
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.118 on 265 degrees of freedom
## Multiple R-squared:  0.2333, Adjusted R-squared:  0.2014
## F-statistic: 7.329 on 11 and 265 DF, p-value: 6.186e-11

#random forest for math grade
library(randomForest)

rf <- randomForest(
  G3 ~ .,
  data=training.math
)

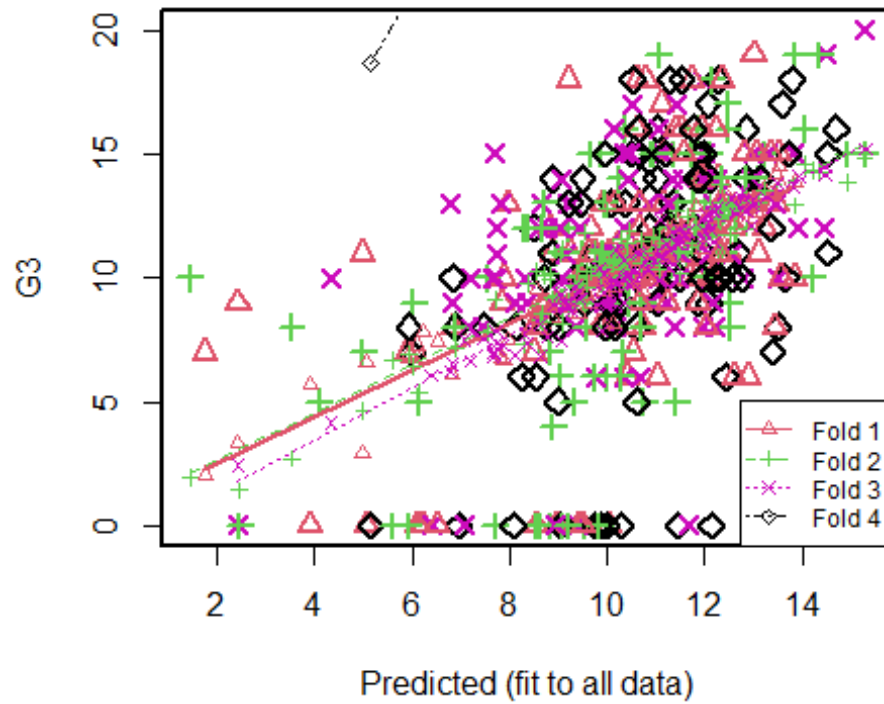
x1 = test.math[,31]
x2 = predict(rf, newdata=test.math[,c(1:30)])
rss <- sum((x2 - x1) ^ 2) ## residual sum of squares
tss <- sum((x1 - mean(x1)) ^ 2) ## total sum of squares
rsq <- 1 - rss/tss
print(rsq)

## [1] 0.2972521

Results <- suppressWarnings(CVlm(data1, form.lm= G3 ~., m=4, dots=FALSE,
seed=2, legend.pos="bottomright", printit=FALSE, main="Mathematics:Small
symbols=predicted values, Big symbols=actual values."))
```



tics: Small symbols = predicted values, Big symbols = actual values



```
attr(Results, 'ms')
```

```
## [1] 215.8877
```