

error count and jadud regression

XXX

2023-07-08

Error count

use error count to predict midterm 1

```
# predict midterm1 with only compiler errors
summary(rfit(Midterm1 ~
             C_totalError_HW3+C_totalError_HW4,
             data = allError))

## Call:
## rfit.default(formula = Midterm1 ~ C_totalError_HW3 + C_totalError_HW4,
##             data = allError)
##
## Coefficients:
##             Estimate Std. Error t.value    p.value
## (Intercept)   92.374659    1.857970  49.7181 < 2.2e-16 ***
## C_totalError_HW3 -0.069547    0.021659  -3.2109  0.001532 **
## C_totalError_HW4 -0.115070    0.057076  -2.0161  0.045070 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.08583972
## Reduction in Dispersion Test: 9.81256 p-value: 8e-05

# predict midterm1 with both compiler and runtime errors
summary(rfit(Midterm1 ~
             R_totalError_HW3+R_totalError_HW4+
             C_totalError_HW3+C_totalError_HW4,
             data = allError))

## Call:
## rfit.default(formula = Midterm1 ~ R_totalError_HW3 + R_totalError_HW4 +
##             C_totalError_HW3 + C_totalError_HW4, data = allError)
##
## Coefficients:
##             Estimate Std. Error t.value    p.value
## (Intercept)   94.689568    2.352654  40.2480 < 2.2e-16 ***
## R_totalError_HW3 -0.106252    0.168496  -0.6306  0.529003
## R_totalError_HW4 -0.259517    0.186306  -1.3930  0.165125
## C_totalError_HW3 -0.063307    0.021734  -2.9128  0.003975 **
## C_totalError_HW4 -0.114421    0.056884  -2.0115  0.045574 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Multiple R-squared (Robust): 0.09816688
## Reduction in Dispersion Test: 5.63312 p-value: 0.00025
```

use error count to predict midterm 2

```
# predict midterm2 with compiler errors
```

```
summary(rfit(Midterm2 ~
  C_totalError_HW3+C_totalError_HW4+
  C_totalError_HW5+C_totalError_HW6+
  C_totalError_HW7++C_totalError_HW8,
  data = allError))
```

```
## Call:
## rfit.default(formula = Midterm2 ~ C_totalError_HW3 + C_totalError_HW4 +
##   C_totalError_HW5 + C_totalError_HW6 + C_totalError_HW7 +
##   +C_totalError_HW8, data = allError)
##
## Coefficients:
##              Estimate Std. Error t.value p.value
## (Intercept)    91.1956187   3.6129660  25.2412 < 2e-16 ***
## C_totalError_HW3  0.0162098   0.0273934   0.5917 0.55500
## C_totalError_HW4 -0.1275316   0.0708203  -1.8008 0.07394 .
## C_totalError_HW5 -0.0051756   0.0163715  -0.3161 0.75238
## C_totalError_HW6 -0.0469343   0.0338544  -1.3864 0.16789
## C_totalError_HW7 -0.0569617   0.0537324  -1.0601 0.29096
## C_totalError_HW8 -0.0546422   0.0309355  -1.7663 0.07957 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.09851068
## Reduction in Dispersion Test: 2.49512 p-value: 0.02535
```

```
# predict midterm2 with both compiler and runtime errors
```

```
summary(rfit(Midterm2 ~
  R_totalError_HW3+R_totalError_HW4 +
  R_totalError_HW5+R_totalError_HW6+
  R_totalError_HW7+R_totalError_HW8+
  C_totalError_HW3+C_totalError_HW4+
  C_totalError_HW5+C_totalError_HW6+
  C_totalError_HW7++C_totalError_HW8,
  data = allError))
```

```
## Call:
## rfit.default(formula = Midterm2 ~ R_totalError_HW3 + R_totalError_HW4 +
##   R_totalError_HW5 + R_totalError_HW6 + R_totalError_HW7 +
##   R_totalError_HW8 + C_totalError_HW3 + C_totalError_HW4 +
##   C_totalError_HW5 + C_totalError_HW6 + C_totalError_HW7 +
##   +C_totalError_HW8, data = allError)
##
## Coefficients:
##              Estimate Std. Error t.value  p.value
## (Intercept)    101.8506519   3.9411250  25.8430 < 2.2e-16 ***
## R_totalError_HW3 -0.0394384   0.2130705  -0.1851 0.853440
## R_totalError_HW4 -0.3636906   0.2230270  -1.6307 0.105355
```

```

## R_totalError_HW5 -0.0619963 0.0208509 -2.9733 0.003508 **
## R_totalError_HW6 -0.0129530 0.0182109 -0.7113 0.478177
## R_totalError_HW7 0.0179124 0.0402279 0.4453 0.656857
## R_totalError_HW8 -0.1251998 0.0455425 -2.7491 0.006821 **
## C_totalError_HW3 0.0430217 0.0289737 1.4849 0.139985
## C_totalError_HW4 -0.0994089 0.0741028 -1.3415 0.182079
## C_totalError_HW5 0.0048874 0.0167103 0.2925 0.770385
## C_totalError_HW6 -0.0712719 0.0342985 -2.0780 0.039663 *
## C_totalError_HW7 -0.0628101 0.0552431 -1.1370 0.257624
## C_totalError_HW8 -0.0460193 0.0314240 -1.4645 0.145464
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.2346442
## Reduction in Dispersion Test: 3.34685 p-value: 0.00029

```

Jadud

use jadud to predict midterm 1 grades

```
# predict midterm1 with only compiler jadud measures
summary(rfit(Midterm1 ~ jadud_hw_03_c + jadud_hw_04_c, data = jadud_df))
```

```
## Call:
## rfit.default(formula = Midterm1 ~ jadud_hw_03_c + jadud_hw_04_c,
##      data = jadud_df)
##
## Coefficients:
##              Estimate Std. Error t.value    p.value
## (Intercept)   98.9173     1.6103 61.4297 < 2.2e-16 ***
## jadud_hw_03_c -49.2191     6.4254 -7.6601 3.186e-13 ***
## jadud_hw_04_c  -3.6324     5.7801 -0.6284  0.5302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1810045
## Reduction in Dispersion Test: 30.38859 p-value: 0
```

```
# predict midterm1 with both compiler and runtime jadud measures
```

```
summary(rfit(Midterm1 ~ jadud_hw_03_c + jadud_hw_04_c + jadud_hw_03_r + jadud_hw_04_r, data = jadud_df))
```

```
## Call:
## rfit.default(formula = Midterm1 ~ jadud_hw_03_c + jadud_hw_04_c +
##      jadud_hw_03_r + jadud_hw_04_r, data = jadud_df)
##
## Coefficients:
##              Estimate Std. Error t.value    p.value
## (Intercept)   99.3683     1.6027 62.0022 < 2.2e-16 ***
## jadud_hw_03_c -47.9519     6.5295 -7.3439 2.398e-12 ***
## jadud_hw_04_c  -4.5003     5.8821 -0.7651  0.4449
## jadud_hw_03_r  13.5565    62.2608  0.2177  0.8278
## jadud_hw_04_r  -7.2937     8.9085 -0.8187  0.4137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1847079
## Reduction in Dispersion Test: 15.46232 p-value: 0
```

use jadud to predict midterm 2 grades

```
# predict midterm2 with only compiler jadud measures
```

```
summary(rfit(Midterm2 ~
              jadud_hw_03_c + jadud_hw_04_c + jadud_hw_05_c +
              jadud_hw_06_c + jadud_hw_07_c + jadud_hw_08_c, data = jadud_df))
```

```
## Call:
## rfit.default(formula = Midterm2 ~ jadud_hw_03_c + jadud_hw_04_c +
##      jadud_hw_05_c + jadud_hw_06_c + jadud_hw_07_c + jadud_hw_08_c,
##      data = jadud_df)
##
## Coefficients:
```

```
##               Estimate Std. Error t.value    p.value
## (Intercept)    97.4574     1.3090 74.4538 < 2.2e-16 ***
## jadud_hw_03_c -39.9251     7.4531 -5.3568 1.834e-07 ***
## jadud_hw_04_c   4.3990     6.1913  0.7105  0.47801
## jadud_hw_05_c  -5.9471    11.0519 -0.5381  0.59096
## jadud_hw_06_c -40.9062    17.1230 -2.3890  0.01759 *
## jadud_hw_07_c -15.6508     8.6737 -1.8044  0.07230 .
## jadud_hw_08_c  -9.0145    10.9670 -0.8220  0.41183
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1900261
## Reduction in Dispersion Test: 10.40094 p-value: 0
# predict midterm2 with both compiler and runtime jadud measures
summary(rfit(Midterm2 ~
              jadud_hw_03_c + jadud_hw_04_c + jadud_hw_05_c +
              jadud_hw_06_c + jadud_hw_07_c + jadud_hw_08_c +
              jadud_hw_03_r + jadud_hw_04_r + jadud_hw_05_r +
              jadud_hw_06_r + jadud_hw_07_r + jadud_hw_08_r, data = jadud_df))

## Call:
## rfit.default(formula = Midterm2 ~ jadud_hw_03_c + jadud_hw_04_c +
##      jadud_hw_05_c + jadud_hw_06_c + jadud_hw_07_c + jadud_hw_08_c +
##      jadud_hw_03_r + jadud_hw_04_r + jadud_hw_05_r + jadud_hw_06_r +
##      jadud_hw_07_r + jadud_hw_08_r, data = jadud_df)
##
## Coefficients:
##               Estimate Std. Error t.value    p.value
## (Intercept)   103.2925     2.0754 49.7695 < 2.2e-16 ***
## jadud_hw_03_c -33.8606     7.5561 -4.4812 1.113e-05 ***
## jadud_hw_04_c   1.3468     6.3772  0.2112  0.83291
## jadud_hw_05_c   1.0471    11.2094  0.0934  0.92565
## jadud_hw_06_c -41.3119    17.4119 -2.3726  0.01839 *
## jadud_hw_07_c -16.9318     8.7138 -1.9431  0.05308 .
## jadud_hw_08_c -12.6171    10.9961 -1.1474  0.25227
## jadud_hw_03_r   2.3828    63.7165  0.0374  0.97020
## jadud_hw_04_r -10.7597     9.2594 -1.1620  0.24629
## jadud_hw_05_r -29.8552    16.6065 -1.7978  0.07337 .
## jadud_hw_06_r -37.8256    21.5767 -1.7531  0.08077 .
## jadud_hw_07_r  -5.4219     9.6043 -0.5645  0.57288
## jadud_hw_08_r -36.3899     8.2510 -4.4104 1.512e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.2635331
## Reduction in Dispersion Test: 7.75308 p-value: 0
```

RED

use RED to predict midterm 1 grades

```
# predict midterm1 with only compiler RED measures
summary(rfit(Midterm1 ~ RED_hw_03_c + RED_hw_04_c, data = RED_df))

## Call:
## rfit.default(formula = Midterm1 ~ RED_hw_03_c + RED_hw_04_c,
## data = RED_df)
##
## Coefficients:
## Estimate Std. Error t.value p.value
## (Intercept) 94.000000 1.624354 57.8692 < 2.2e-16 ***
## RED_hw_03_c -0.256098 0.059171 -4.3281 2.121e-05 ***
## RED_hw_04_c -0.367219 0.180493 -2.0345 0.04287 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1054053
## Reduction in Dispersion Test: 15.90633 p-value: 0

# predict midterm1 with both compiler and runtime RED measures
summary(rfit(Midterm1 ~ RED_hw_03_c + RED_hw_04_c + RED_hw_03_r + RED_hw_04_r, data = RED_df))

## Call:
## rfit.default(formula = Midterm1 ~ RED_hw_03_c + RED_hw_04_c +
## RED_hw_03_r + RED_hw_04_r, data = RED_df)
##
## Coefficients:
## Estimate Std. Error t.value p.value
## (Intercept) 94.500000 1.516400 62.3187 < 2.2e-16 ***
## RED_hw_03_c -0.243905 0.058601 -4.1621 4.252e-05 ***
## RED_hw_04_c -0.395723 0.178310 -2.2193 0.0273 *
## RED_hw_03_r 2.067071 10.532009 0.1963 0.8446
## RED_hw_04_r -0.895147 0.652994 -1.3708 0.1716
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1138511
## Reduction in Dispersion Test: 8.60806 p-value: 0
```

use RED to predict midterm 2 grades

```
# predict midterm2 with only compiler RED measures
summary(rfit(Midterm2 ~
              RED_hw_03_c + RED_hw_04_c + RED_hw_05_c +
              RED_hw_06_c + RED_hw_07_c + RED_hw_08_c, data = RED_df))

## Call:
## rfit.default(formula = Midterm2 ~ RED_hw_03_c + RED_hw_04_c +
## RED_hw_05_c + RED_hw_06_c + RED_hw_07_c + RED_hw_08_c, data = RED_df)
##
## Coefficients:
## Estimate Std. Error t.value p.value
```

```

## (Intercept) 93.026505    1.283148 72.4986 < 2e-16 ***
## RED_hw_03_c -0.072810    0.058704 -1.2403 0.21596
## RED_hw_04_c -0.435490    0.178999 -2.4329 0.01564 *
## RED_hw_05_c -0.079192    0.070768 -1.1190 0.26414
## RED_hw_06_c -0.033521    0.075138 -0.4461 0.65587
## RED_hw_07_c -0.241766    0.153092 -1.5792 0.11547
## RED_hw_08_c -0.053010    0.050379 -1.0522 0.29365
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.07207637
## Reduction in Dispersion Test: 3.44359 p-value: 0.00272
# predict midterm2 with both compiler and runtime RED measures
summary(rfit(Midterm2 ~
              RED_hw_03_c + RED_hw_04_c + RED_hw_05_c +
              RED_hw_06_c + RED_hw_07_c + RED_hw_08_c +
              RED_hw_03_r + RED_hw_04_r + RED_hw_05_r +
              RED_hw_06_r + RED_hw_07_r + RED_hw_08_r, data = RED_df))

## Call:
## rfit.default(formula = Midterm2 ~ RED_hw_03_c + RED_hw_04_c +
##   RED_hw_05_c + RED_hw_06_c + RED_hw_07_c + RED_hw_08_c + RED_hw_03_r +
##   RED_hw_04_r + RED_hw_05_r + RED_hw_06_r + RED_hw_07_r + RED_hw_08_r,
##   data = RED_df)
##
## Coefficients:
##               Estimate Std. Error t.value    p.value
## (Intercept) 96.4316084   1.7272406 55.8299 < 2.2e-16 ***
## RED_hw_03_c -0.0547927   0.0588409 -0.9312 0.3526140
## RED_hw_04_c -0.3346032   0.1853340 -1.8054 0.0721678 .
## RED_hw_05_c -0.0035622   0.0706550 -0.0504 0.9598295
## RED_hw_06_c -0.0363201   0.0743600 -0.4884 0.6256523
## RED_hw_07_c -0.2566736   0.1541638 -1.6649 0.0971291 .
## RED_hw_08_c -0.0573922   0.0497040 -1.1547 0.2492822
## RED_hw_03_r  6.7381570 10.3097964  0.6536 0.5139675
## RED_hw_04_r -0.6718638   0.6497619 -1.0340 0.3020898
## RED_hw_05_r -0.4396455   0.1312511 -3.3497 0.0009291 ***
## RED_hw_06_r -0.4010379   0.5336166 -0.7515 0.4530033
## RED_hw_07_r -0.3714000   0.3534333 -1.0508 0.2943103
## RED_hw_08_r -1.3180180   0.2643633 -4.9856 1.129e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.1903674
## Reduction in Dispersion Test: 5.09444 p-value: 0

```

HW grades

use HW grades to predict midterm 1

```
summary(rfit(Midterm1 ~ HW3 + HW4, data = grade))

## Call:
## rfit.default(formula = Midterm1 ~ HW3 + HW4, data = grade)
##
## Coefficients:
##              Estimate Std. Error t.value    p.value
## (Intercept) 43.445376   7.792011   5.5756 5.96e-08 ***
## HW3          0.577252   0.053811  10.7274 < 2.2e-16 ***
## HW4         -0.055839   0.099049  -0.5638   0.5734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.2566742
## Reduction in Dispersion Test: 46.78884 p-value: 0
```

use HW grades to predict midterm 2

```
## Call:
## rfit.default(formula = Midterm2 ~ HW3 + HW4 + HW5 + HW6 + HW7 +
##              HW8, data = grade)
##
## Coefficients:
##              Estimate Std. Error t.value    p.value
## (Intercept) 50.079876   8.378750   5.9770 7.250e-09 ***
## HW3          0.272271   0.068555   3.9716 9.188e-05 ***
## HW4         -0.015550   0.117729  -0.1321   0.89502
## HW5         -0.284973   0.121679  -2.3420   0.01992 *
## HW6          0.328977   0.059966   5.4860 9.541e-08 ***
## HW7          0.090036   0.064289   1.4005   0.16253
## HW8          0.072989   0.056842   1.2841   0.20023
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Multiple R-squared (Robust): 0.3298909
## Reduction in Dispersion Test: 21.90709 p-value: 0
```