**NEW GLM and Tree Results**

Call:

glm(formula = retweets ~ ., data = covidsent)

Deviance Residuals:

Min 1Q Median 3Q Max

-121276 -10236 -430 6783 164582

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.433e+07 9.058e+05 15.825 < 2e-16 \*\*\*

isretweetedTRUE 6.127e+03 4.739e+02 12.929 < 2e-16 \*\*\*

length -1.065e+02 1.440e+01 -7.396 1.45e-13 \*\*\*

favorites -1.531e+00 5.346e+00 -0.286 0.774610

ishealth 1.832e+03 6.311e+02 2.902 0.003707 \*\*

ispandemic -8.385e+03 1.121e+03 -7.479 7.75e-14 \*\*\*

isvirus -2.543e+02 5.779e+02 -0.440 0.659908

isemergency -7.527e+03 1.597e+03 -4.714 2.44e-06 \*\*\*

isdeaths -2.557e+03 1.135e+03 -2.252 0.024314 \*

iswho 2.245e+04 6.550e+02 34.268 < 2e-16 \*\*\*

iscdc -3.678e+03 1.365e+03 -2.695 0.007046 \*\*

isnih -1.592e+03 5.978e+03 -0.266 0.790023

isdisease -3.306e+03 1.429e+03 -2.314 0.020679 \*

isquarantine 1.264e+02 1.023e+03 0.124 0.901691

isrecover -4.093e+03 2.208e+03 -1.854 0.063764 .

isban -3.992e+03 1.134e+03 -3.521 0.000430 \*\*\*

iscoronavirus 1.308e+04 6.422e+02 20.369 < 2e-16 \*\*\*

iscovid19 -3.839e+03 3.280e+02 -11.705 < 2e-16 \*\*\*

iswash -1.796e+03 1.211e+03 -1.483 0.138200

isracist -1.072e+04 4.412e+03 -2.429 0.015127 \*

isasian -1.852e+03 3.579e+03 -0.517 0.604825

ischinese 3.839e+03 9.399e+02 4.084 4.44e-05 \*\*\*

isinfectious -2.837e+03 2.731e+03 -1.039 0.298938

ch.confirmed -1.986e+02 1.274e+01 -15.584 < 2e-16 \*\*\*

it.confirmed -2.086e+01 1.010e+00 -20.657 < 2e-16 \*\*\*

ot.confirmed 1.480e+01 5.757e-01 25.704 < 2e-16 \*\*\*

us.confirmed -8.863e+01 6.316e+00 -14.031 < 2e-16 \*\*\*

ch.deaths 5.220e+02 4.268e+01 12.228 < 2e-16 \*\*\*

it.deaths 9.116e+01 8.199e+00 11.118 < 2e-16 \*\*\*

ot.deaths -4.888e+01 1.998e+01 -2.447 0.014411 \*

us.deaths -8.108e+02 2.353e+02 -3.447 0.000568 \*\*\*

fear -4.905e+03 3.140e+02 -15.623 < 2e-16 \*\*\*

sadness 9.461e+03 3.729e+02 25.375 < 2e-16 \*\*\*

trust -1.245e+03 2.679e+02 -4.649 3.36e-06 \*\*\*

disgust -5.998e+03 4.806e+02 -12.479 < 2e-16 \*\*\*

negative 2.294e+03 2.759e+02 8.314 < 2e-16 \*\*\*

anticipation 8.628e+02 2.832e+02 3.047 0.002314 \*\*

joy -4.901e+02 4.442e+02 -1.103 0.269935

positive -2.129e+03 2.275e+02 -9.357 < 2e-16 \*\*\*

anger -1.836e+03 4.424e+02 -4.150 3.34e-05 \*\*\*

surprise -3.712e+03 3.984e+02 -9.317 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 606168624)

Null deviance: 2.1208e+13 on 23515 degrees of freedom

Residual deviance: 1.4230e+13 on 23475 degrees of freedom

AIC: 542335

Number of Fisher Scoring iterations: 2

Regression tree:

tree(formula = retweets ~ ., data = covidsent)

Variables actually used in tree construction:

[1] "ch.confirmed" "length" "sadness" "anticipation" "iscoronavirus"

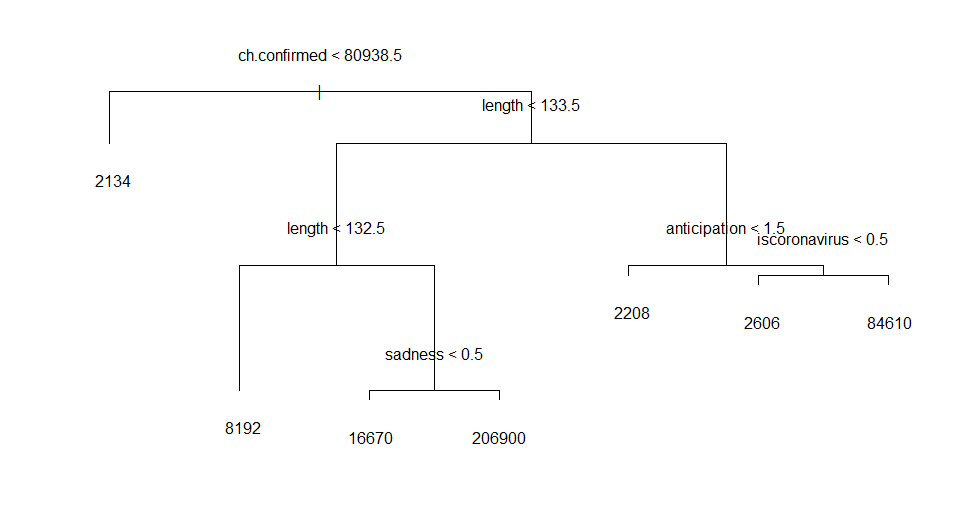
Number of terminal nodes: 7

Residual mean deviance: 43530000 = 1.023e+12 / 23510

Distribution of residuals:

Min. 1st Qu. Median Mean 3rd Qu. Max.

-206900 -2126 -1965 0 -820 56460



**WITH Sentiment only**

Call:

glm(formula = retweets ~ ., data = sentonly)

Deviance Residuals:

Min 1Q Median 3Q Max

-75126 -9085 -2201 4675 172598

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.428e+07 9.547e+05 14.955 < 2e-16 \*\*\*

isretweetedTRUE 7.190e+03 5.013e+02 14.343 < 2e-16 \*\*\*

length -7.825e+01 1.514e+01 -5.167 2.39e-07 \*\*\*

favorites 1.707e+00 5.683e+00 0.300 0.7639

ch.confirmed -1.974e+02 1.344e+01 -14.689 < 2e-16 \*\*\*

it.confirmed -2.559e+01 1.069e+00 -23.945 < 2e-16 \*\*\*

ot.confirmed 1.627e+01 6.070e-01 26.803 < 2e-16 \*\*\*

us.confirmed -1.097e+02 6.675e+00 -16.441 < 2e-16 \*\*\*

ch.deaths 5.051e+02 4.517e+01 11.181 < 2e-16 \*\*\*

it.deaths 1.250e+02 8.603e+00 14.529 < 2e-16 \*\*\*

ot.deaths 2.672e+00 2.116e+01 0.126 0.8995

us.deaths -1.073e+03 2.489e+02 -4.313 1.62e-05 \*\*\*

fear -6.551e+03 3.220e+02 -20.342 < 2e-16 \*\*\*

sadness 9.829e+03 3.771e+02 26.062 < 2e-16 \*\*\*

trust -1.438e+03 2.822e+02 -5.097 3.48e-07 \*\*\*

disgust -6.686e+03 4.898e+02 -13.651 < 2e-16 \*\*\*

negative 2.671e+03 2.782e+02 9.603 < 2e-16 \*\*\*

anticipation 1.320e+03 2.988e+02 4.417 1.01e-05 \*\*\*

joy 2.112e+02 4.658e+02 0.453 0.6503

positive -2.678e+03 2.387e+02 -11.220 < 2e-16 \*\*\*

anger -1.088e+03 4.596e+02 -2.367 0.0179 \*

surprise -5.310e+03 4.034e+02 -13.161 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 685338877)

Null deviance: 2.1208e+13 on 23515 degrees of freedom

Residual deviance: 1.6101e+13 on 23494 degrees of freedom

AIC: 545202

Number of Fisher Scoring iterations: 2

Regression tree:

tree(formula = retweets ~ ., data = sentonly)

Variables actually used in tree construction:

[1] "ch.confirmed" "length" "sadness" "anticipation" "anger"

Number of terminal nodes: 8

Residual mean deviance: 44720000 = 1.051e+12 / 23510

Distribution of residuals:

Min. 1st Qu. Median Mean 3rd Qu. Max.

-206900 -2126 -1964 0 -820 56460

