analysis

November 17, 2021

0.1 Overview

0.2 Data Cleaning

0.3 Feature Engineering

First, we create **y**, the feature we would like to predict, as the normalized average of responses to questions 1-15 on the survey. Those questions include, for example, "Secret organizations communicate with extraterrestrials, but keep this fact from the public" and "The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of some organization", and participants are asked to respond with their level of agreement from 1 to 7.

0.4 Results

Dep. Variable:	GCB	**R-squared:**	0.139
Model:	OLS	**Adj. R-squared:**	0.133
Method:	Least Squares	**F-statistic:**	22.29
Date:	Wed, 17 Nov 2021	**Prob (F-statistic):**	3.82e-68
Time:	12:41:16	**Log-Likelihood:**	579.98
No. Observations:	2495	**AIC:**	-1122.
Df Residuals:	2476	**BIC:**	-1011.
Df Model:	18		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
TIPI2	0.0142	0.002	6.565	0.000	0.010	0.018
TIPI5	0.0056	0.003	2.008	0.045	0.000	0.011
TIPI6	-0.0014	0.002	-0.670	0.503	-0.006	0.003
vocabulary_misclassification	0.0706	0.017	4.051	0.000	0.036	0.105
STEM	-0.0383	0.010	-3.774	0.000	-0.058	-0.018
education_2	0.0380	0.009	4.120	0.000	0.020	0.056
education_3	0.0088	0.011	0.812	0.417	-0.012	0.030
urban_3	0.0242	0.008	2.882	0.004	0.008	0.041
gender_2	0.0219	0.008	2.707	0.007	0.006	0.038
engnat_1	0.0061	0.009	0.653	0.514	-0.012	0.025
religion_2	-0.0753	0.009	-8.266	0.000	-0.093	-0.057
religion_3	0.0873	0.028	3.119	0.002	0.032	0.142
religion_7	0.0603	0.014	4.179	0.000	0.032	0.089
religion_12	0.0996	0.013	7.947	0.000	0.075	0.124
orientation_2	-0.0120	0.011	-1.130	0.259	-0.033	0.009
orientation_5	0.0431	0.017	2.525	0.012	0.010	0.077
voted_2	0.0260	0.009	3.022	0.003	0.009	0.043
married_1	-0.0284	0.010	-2.842	0.005	-0.048	-0.009
constant	0.2304	0.013	17.307	0.000	0.204	0.257
const	0.2304	0.013	17.307	0.000	0.204	0.257

Omnibus:	85.135	**Durbin-Watson:**	1.925
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	42.486
Skew:	0.107	**Prob(JB):**	5.95e-10
Kurtosis:	2.398	**Cond. No.**	5.64e+17

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 6.51e-31. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.

0.5 Conclusion

R^2 is small, but we found some variables that are slightly correlated with belief in conspiracy theories, blah blah blah.

[ ]: