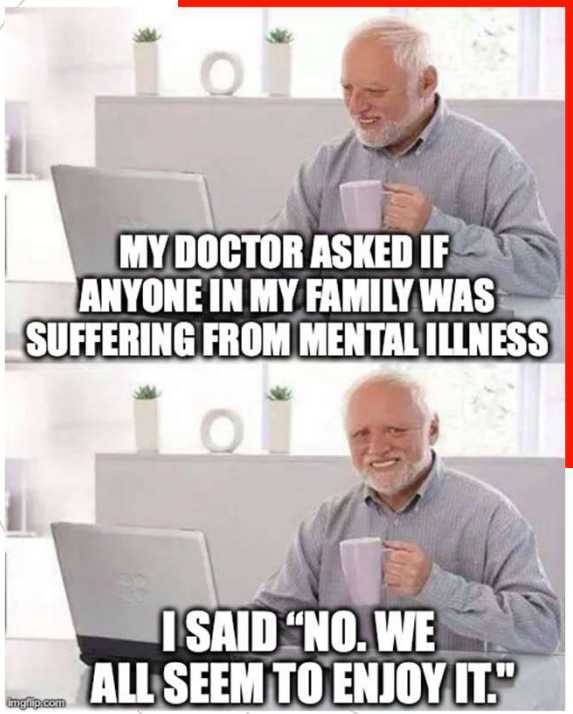
Mental Health in the Workforce: What should you look for?



Ryan McLean

Kevin Nelson

Tyler Wandry

Introduce Problem

- Mental Health rates are on the rise
- Many companies are losing a lot of work every year because of mental health issues
  - Including time at work and employee efficiency
  - That means a lot of companies are worried about it
- Don't lose the job you want because of mental health issues
- Learn what you should be looking for in a company that will support you.
- **What are the attributes of the employee and the company that predict whether mental illness interferes with your work?**

Data Source

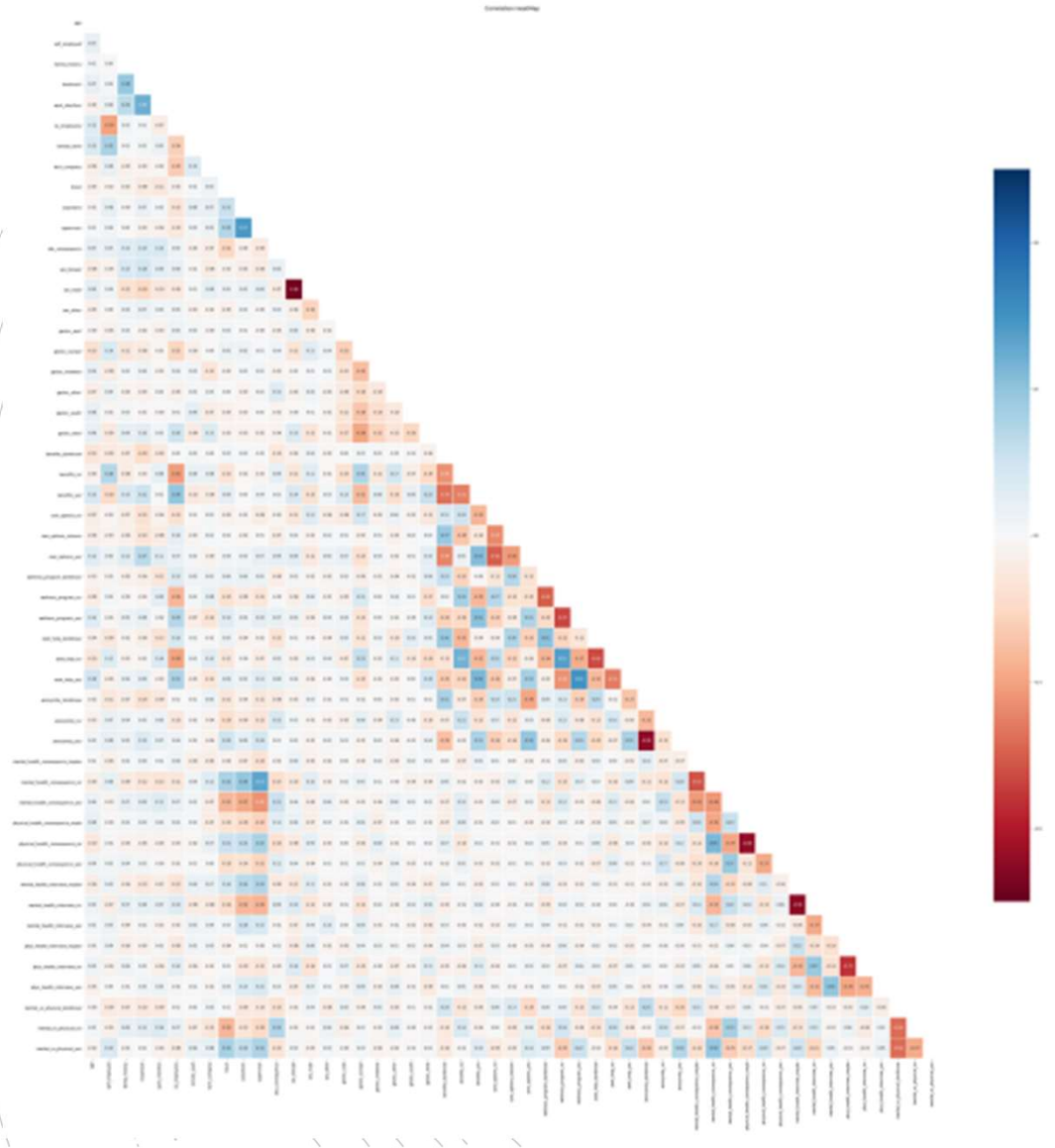
- Data from *Open Sourcing Mental Illness*. A survey collected in 2014 which measures the attitudes toward mental health and frequency of mental illness in technical employees

Final Sample

- 1,251 Employees
- 79% Male, 19.7% Female, 1.3% other
- 82% worked at a Tech company, 18% were technical employees at non-tech companies
- Average age was 32.08 (*SD* = 7.3), ranged from 18 to 72 years old
- 39% reported a family history of mental illness
- Those who have reported seeking treatment
  - Females: 68.8%
  - Males: 45.4%
- Those who feel mental illness interferes with their work at least sometimes
  - Females: 67.3%
  - Males: 59.1%

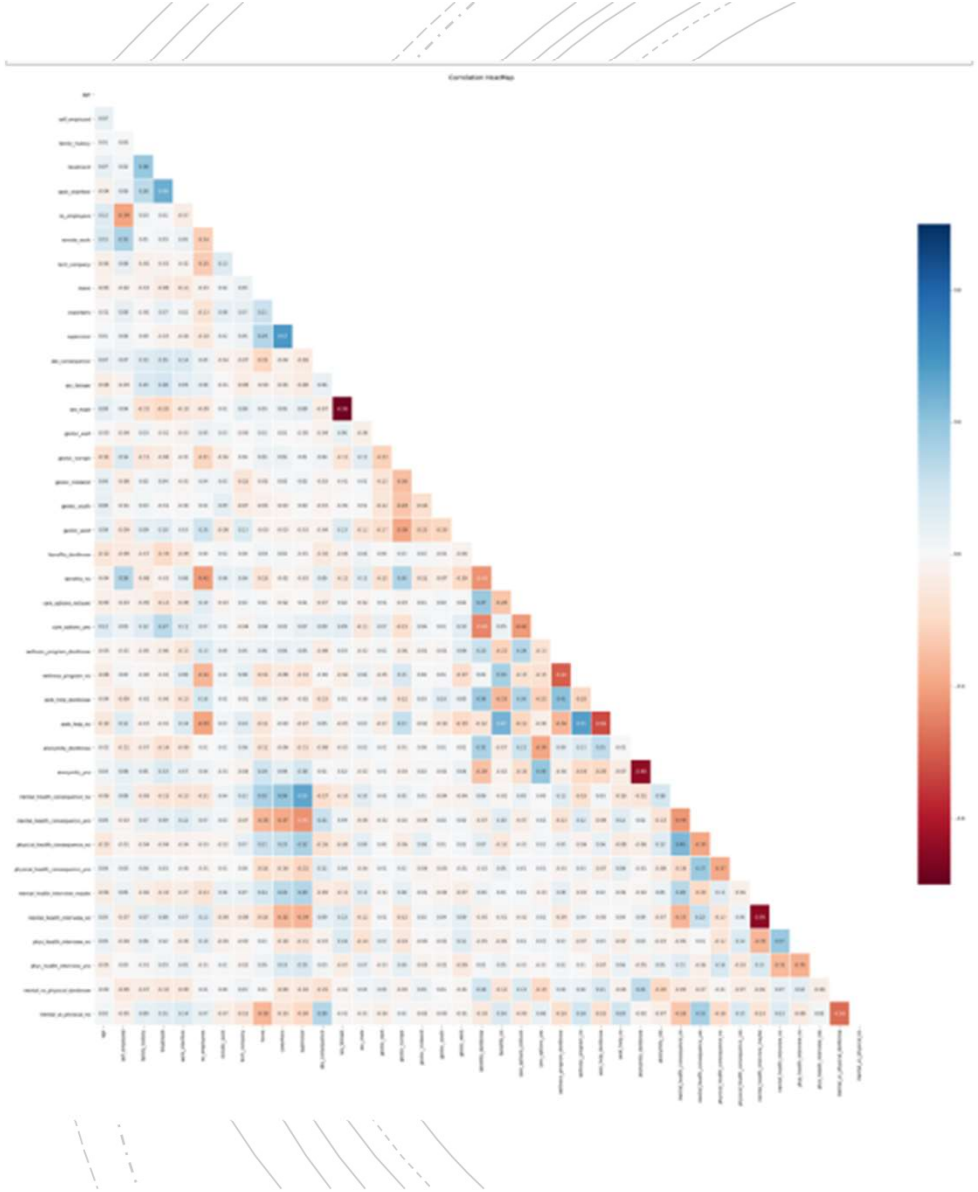
Data Cleaning

- Creation of Dummy Variables
  - Intentionally deleted the one with the smallest correlation with the outcome variable
- Remaining Correlations
  - A few high correlations existed between dummy variables from same variable (e.g., 'I don't know' with 'no')
  - No other significant collinearity issues existed
    - R = 0.57: Would you talk to coworkers about mental illness? What about supervisors?
    - R = 0.53: Would you talk to your supervisor about mental illness? Are there consequences for mental illness at your company?
    - R = 0.56: Has your advisor ever talked about a company wellness program? Does your employer provide resources of where you can seek help?



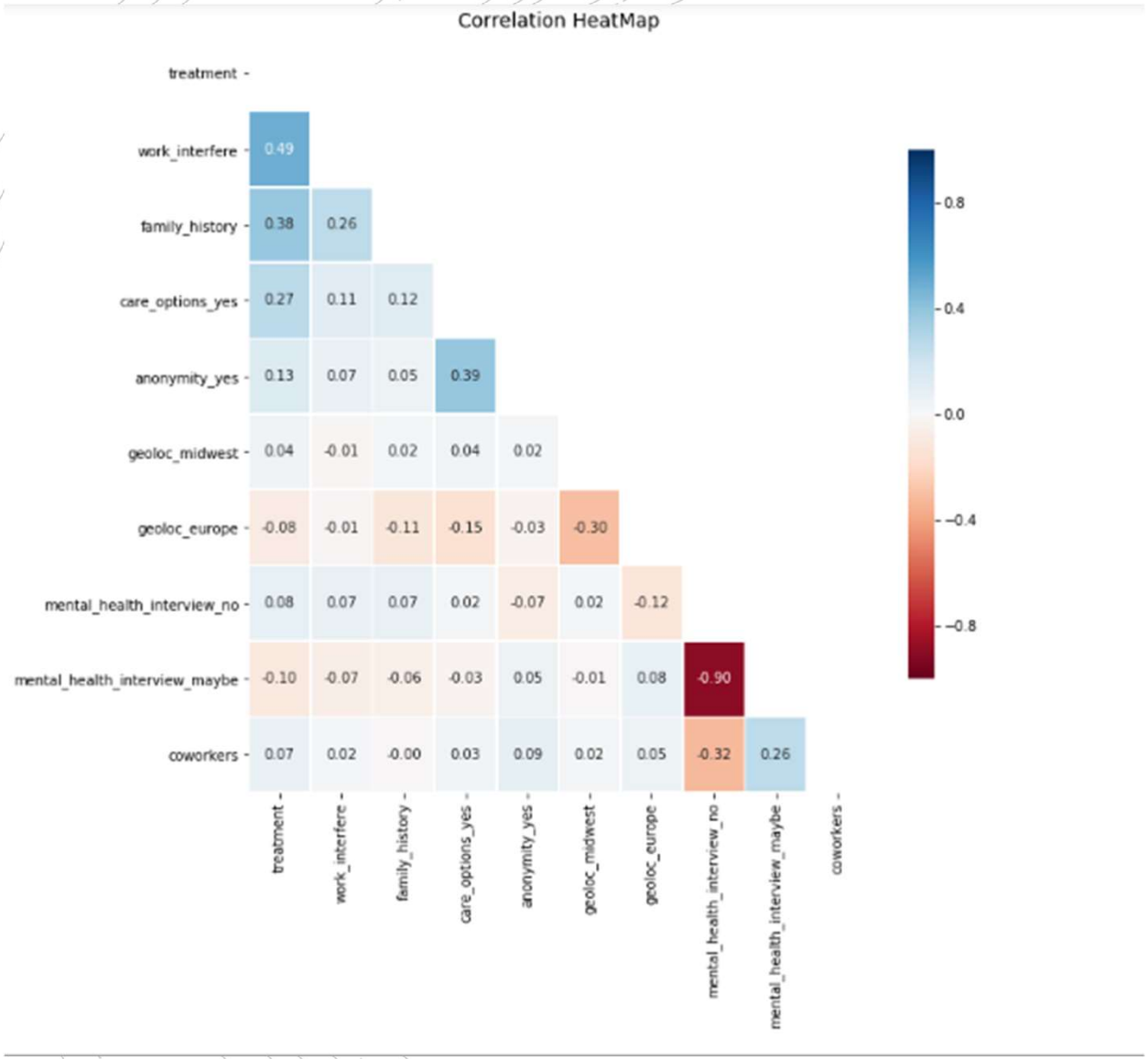
All variables with dummy variables

51 Total Variables



Cleaned Data

39 Total Variables



Significant Predictors from Logistic Regression

10 Total Variables

Main Predictors of mental illness interfering at work

- Have you sought out treatment for mental illness?
- Do you have a family history of mental illness?
- Does your company have benefits for mental illness?
- Would you feel comfortable talking to an employer about mental health issues?
- Would you feel comfortable talking to a colleague about mental health issues?
- Where do you live?

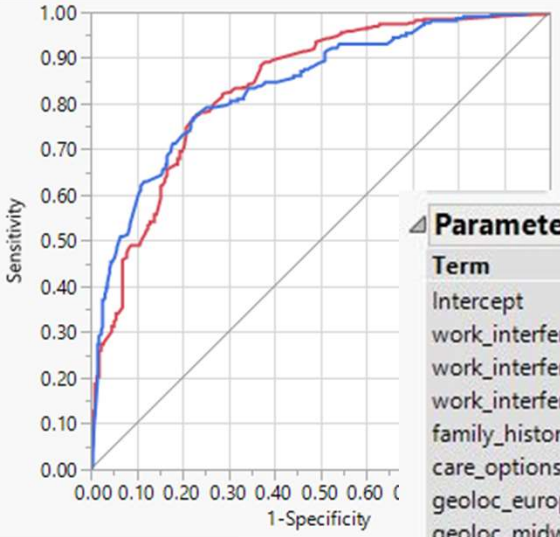
Method

- Preliminary Analyses
  - LogReg
  - CART
  - KNN
  - ANN
  - Boosted Tree
  - Random Forest
  - Cluster Analysis

	actual = yes			actual = no									
Algorithm Name	**TP**	**FN**	**Total**	**TN**	**FP**	**Total**	Total	**Recall**	**Specifity**	**Precision**	**F-meas**	**%Error**	**AUC**
Logistic Regression - Continuous	202	39	241	95	47	142	383	83.8%	66.9%	81.1%	82.4%	22.5%	0.8338
Logistic Regression - Ordinal	207	34	241	3	49	52	293	85.9%	5.8%	80.9%	83.3%	28.3%	0.8379
Logistic Regression w/ sig	224	20	244	99	46	145	389	91.8%	68.3%	83.0%	87.2%	17.0%	0.8443
CART	233	14	247	190	63	253	500	94.3%	75.1%	78.7%	85.8%	15.4%	0.8944
CART w/ sig	233	14	247	190	63	253	500	94.3%	75.1%	78.7%	85.8%	15.4%	0.8955
KNN	138	109	247	208	45	253	500	55.9%	82.2%	75.4%	64.2%	30.8%	--
KNN w/ sig	188	59	247	199	54	253	500	76.1%	78.7%	77.7%	76.9%	22.6%	--
ANN	214	27	241	92	50	142	383	88.8%	64.8%	81.1%	84.8%	20.1%	0.8512
ANN	209	32	241	97	45	142	383	86.7%	68.3%	82.3%	84.4%	20.1%	0.8645
ANN	229	12	241	84	58	142	383	95.0%	59.2%	79.8%	86.7%	18.3%	0.8385
ANN w/ sig	222	22	244	99	46	145	389	91.0%	68.3%	82.8%	86.7%	17.5%	0.8505
ANN w/ sig	225	19	244	95	50	145	389	92.2%	65.5%	81.8%	86.7%	17.7%	0.8476
ANN w/ sig	221	23	244	99	46	145	389	90.6%	68.3%	82.8%	86.5%	17.7%	0.8537
Boosted Tree	229	18	247	195	58	253	500	92.7%	77.1%	79.8%	85.8%	15.2%	0.8996
Boosted Tree w/ sig	229	18	247	199	54	253	500	92.7%	78.7%	80.9%	86.4%	14.4%	0.8939
Random Forest	214	33	247	199	54	253	500	86.6%	78.7%	79.9%	83.1%	17.4%	0.8954
Random Forest w/ sig	233	14	247	188	65	253	500	94.3%	74.3%	78.2%	85.5%	15.8%	0.9026
Cluster SOM 1X3 # 1	72	4	76	17	15	32	108	94.7%	53.1%	82.8%	88.3%	17.6%	0.7882
Cluster SOM 1X3 # 2	115	23	138	68	25	93	231	83.3%	73.1%	82.1%	82.7%	20.8%	0.8540
Cluster SOM 1X3 # 3	14	13	27	8	9	17	44	51.9%	47.1%	60.9%	56.0%	50.0%	0.6242
Cluster SOM 2X2 # 1	49	2	51	11	10	21	72	96.1%	52.4%	83.1%	89.1%	16.7%	0.7586
Cluster SOM 2X2 # 2	69	6	75	12	11	23	98	92.0%	52.2%	86.3%	89.0%	17.3%	0.8119
Cluster SOM 2X2 # 3	14	13	27	8	9	17	44	51.9%	47.1%	60.9%	56.0%	50.0%	0.6438
Cluster SOM 2X2 # 4	70	18	88	57	24	81	169	79.5%	70.4%	74.5%	76.9%	24.9%	0.8266

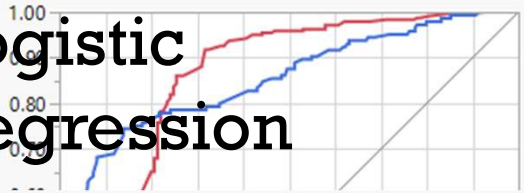
Logistic Regression

Receiver Operating Characteristic



treatment	Area
1	0.8359
0	0.8359

Receiver Operating Characteristic on Validation Data



Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.7196239	0.4946612	12.09	0.0005*
work_interfere[2-1]	2.32766804	0.3535783	43.34	<.0001*
work_interfere[3-2]	0.29916597	0.2749266	1.18	0.2765
work_interfere[4-3]	0.61417882	0.3415963	3.23	0.0722
family_history[0]	-0.5866294	0.1103264	28.27	<.0001*
care_options_yes[0]	-0.4178517	0.1209892	11.93	0.0006*
geoloc_europe[0]	-0.1037796	0.1188672	0.76	0.3826
geoloc_midwest[0]	-0.2872848	0.1631274	3.10	0.0782
coworkers	0.35577947	0.1767026	4.05	0.0441*
anonymity_yes[0]	-0.2734903	0.1331273	4.22	0.0399*
mental_health_interview_maybe[0]	0.55812033	0.3760872	2.20	0.1378
mental_health_interview_no[0]	0.45793812	0.3577321	1.64	0.2005

Confusion Matrix

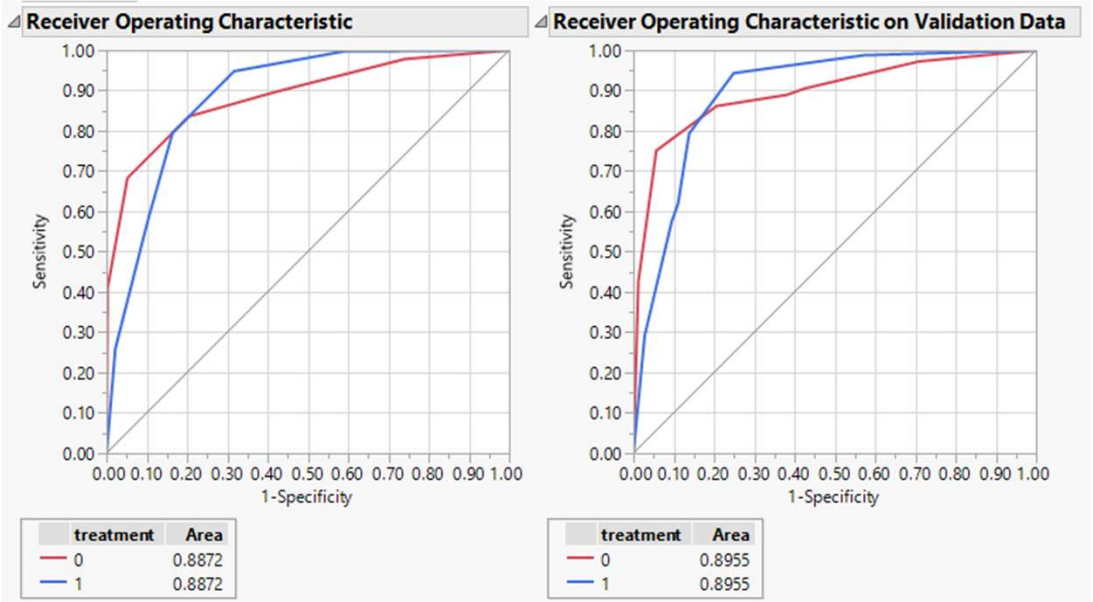
Training

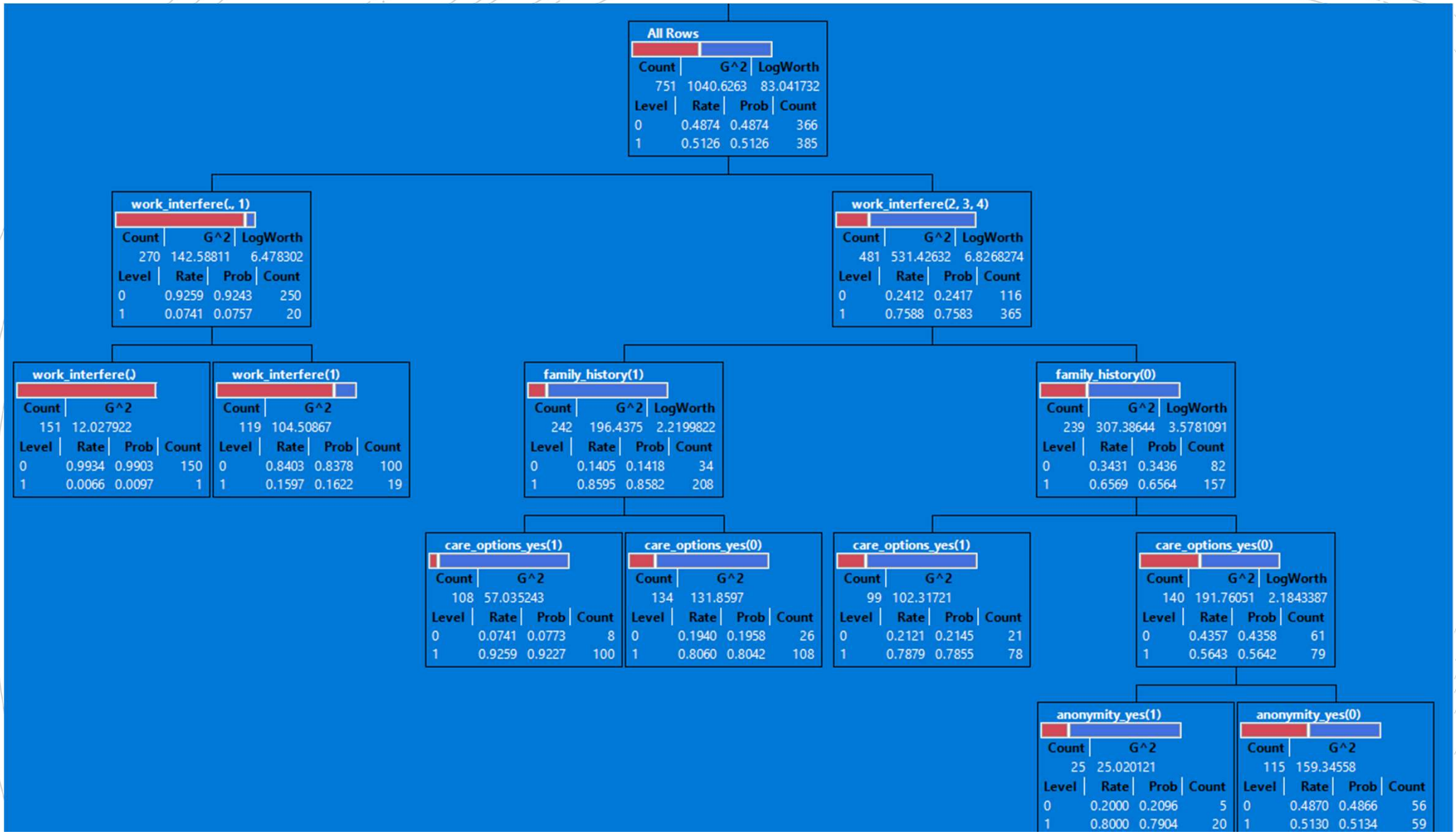
Actual	Predicted Count	
treatment	1	0
1	351	33
0	100	116

Validation

Actual	Predicted Count	
treatment	1	0
1	224	20
0	46	99

CART


Receiver Operating Characteristic
Sensitivity
1-Specificity
0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00
treatment Area
0 0.8872
1 0.8872
Receiver Operating Characteristic on Validation Data
Sensitivity
1-Specificity
0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00
treatment Area
0 0.8955
1 0.8955


All Rows
Count G^2 LogWorth
751 1040.6263 83.041732
Level Rate Prob Count
0 0.4874 0.4874 366
1 0.5126 0.5126 385
work_interfere(., 1)
Count G^2 LogWorth
270 142.58811 6.478302
Level Rate Prob Count
0 0.9259 0.9243 250
1 0.0741 0.0757 20
work_interfere(2, 3, 4)
Count G^2 LogWorth
481 531.42632 6.8268274
Level Rate Prob Count
0 0.2412 0.2417 116
1 0.7588 0.7583 365
work_interfere(.)
Count G^2
151 12.027922
Level Rate Prob Count
0 0.9934 0.9903 150
1 0.0066 0.0097 1
work_interfere(1)
Count G^2
119 104.50867
Level Rate Prob Count
0 0.8403 0.8378 100
1 0.1597 0.1622 19
family_history(1)
Count G^2 LogWorth
242 196.4375 2.2199822
Level Rate Prob Count
0 0.1405 0.1418 34
1 0.8595 0.8582 208
family_history(0)
Count G^2 LogWorth
239 307.38644 3.5781091
Level Rate Prob Count
0 0.3431 0.3436 82
1 0.6569 0.6564 157
care_options_yes(1)
Count G^2
108 57.035243
Level Rate Prob Count
0 0.0741 0.0773 8
1 0.9259 0.9227 100
care_options_yes(0)
Count G^2
134 131.8597
Level Rate Prob Count
0 0.1940 0.1958 26
1 0.8060 0.8042 108
care_options_yes(1)
Count G^2
99 102.31721
Level Rate Prob Count
0 0.2121 0.2145 21
1 0.7879 0.7855 78
care_options_yes(0)
Count G^2 LogWorth
140 191.76051 2.1843387
Level Rate Prob Count
0 0.4357 0.4358 61
1 0.5643 0.5642 79
anonymity_yes(1)
Count G^2
25 25.020121
Level Rate Prob Count
0 0.2000 0.2096 5
1 0.8000 0.7904 20
anonymity_yes(0)
Count G^2
115 159.34558
Level Rate Prob Count
0 0.4870 0.4866 56
1 0.5130 0.5134 59

Discussion

- We discovered that CART, LogReg, Boosted Tree, and Random Forest did really well
- Most important predictors:
  - Individuals
    - Work_interfere[Rarely-Never]
    - Do you have a family history?
  - Companies
    - Availability of Care options
    - Is your anonymity protected if you seek out mental health services