Forecasting mental health crises: lowering healthcare costs using prevention and predictive analytics.

Final report

Colleen Bobbie

November 25, 2017

Contents

Introduction 2
Literature Review 2
Dataset 3
Approach 3
Step one: Data Scrubbing and Preprocessing 4
Step two: Classification Algorithm 4
Step three: Interpretation of results 5
Step four: Recommendation system 5
Results 5
Final Dataset 5
Model evaluation 6
Recommendation System 8
Conclusion 8
References 10
Appendix 1: Final Mental Health Crises C4.5 Decision Tree 12
Appendix 2: Truncated tree for recommendation system 15

Introduction

Annual health care expenditures in the United States currently exceed 3.35 trillion dollars, averaging $10 345 per capita (Yoon et al. 2014). These costs are associated with rapid growth in medical prices and an aging population, suggesting that unless there are dramatic shifts in medical practices, these expenditures could increase rapidly in the coming years.

To understand these ballooning costs, several large scale epidemiological studies are being conducted to provide information on the health of United States' citizens. One such study, the Behavioural Risk Factor Surveillance System (BRFSS), conducts surveys to collect uniform data on health risk behaviours, chronic diseases, access to health care, and the use of preventative health services in the United States. This survey provides valuable information on behavioural patterns which, if coupled with current Big Data and Machine Learning techniques, may help to provide valuable insights into persons at risk of mental health crises. By targeting and understanding these populations, preventative health measures could be put into place to ultimately help lower healthcare costs in the United States.

Literature Review

Adults with depression and/or anxiety are significantly more likely to smoke, to be obese, to be physically inactive, to binge drink, and to drink more heavily than those who do not display any symptoms of depression and/or anxiety (Strine et al., 2008). Additionally, a dose-dependent relationship exists between depression severity and smoking level, obesity, and physical inactivity, in which individuals who are more depressed are more likely to engage heavily in these activities (Strine et al., 2008).

In a study of the 2012 BRFSS data, Yoon et al. found that there are significant relationships between depression and childhood mental illness, limited usual activity, and childhood sexual abuse

(2014). The authors showcased a J48 classification tree to predict depression with an 82% accuracy, using these predictive attributes (Yoon et al., 2014). While this paper created a solid foundation of the use of machine learning in helping to predict mental crises, I will build upon this idea by including several other attributes available from the dataset.

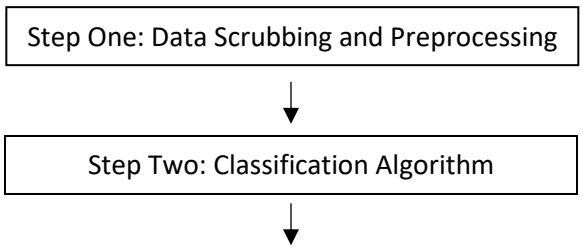
Dataset

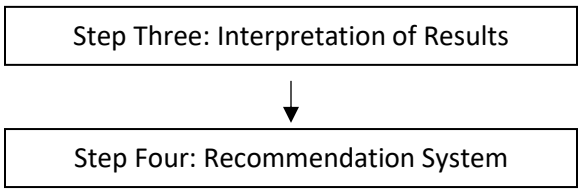
The Behavioural Risk Factor Surveillance System (BRFSS) is a random annual phone-based survey which tracks health risk behaviours, chronic diseases, access to health care, and the use of preventative health services in the United States, available freely for download (https://www.cdc.gov/brfss/annual_data/annual_2016.html). The most current data year (2016) was used for this project, which contained 450 attributes and 486,303 records. All questions asked in the survey (attributes) are available here (https://www.cdc.gov/brfss/questionnaires/pdf-ques/2016brfss_questionnaire_10_14_15.pdf).

Mental illness was characterized by individuals who had current depression/anxiety, a lifetime diagnosis of depression, and/or a lifetime diagnosis of anxiety and the class attribute ("Mental Crises") was compiled based on these answers.

Approach

This project was completed in four main steps. Unless noted, all steps were completed in R.





Step One: Data Scrubbing and Preprocessing

Data Scrubbing and Preprocessing included removing incomplete attributes (i.e. those with >25% unanswered/blank answers) and transforming attributes for downstream processing using R.

Synthetic Minority Over-sampling Technique was used to combat an imbalanced class design. SMOTE generates a random set of minority class observations, using bootstrapping and the datum point's *k*-nearest neighbours. This reduced the bias towards the majority class, while ensuring the 'new' samples in the minority class were representative of the pre-existing values (Chawla et al. 2002).

To further clean the dataset, a Pearson's correlation test was used to determine the correlation between each feature and the class attribute. Attributes with <10% correlation were discarded from downstream analysis.

Step Two: Classification Algorithm

The main decision tree was compiled using 10 fold cross-validation. The chosen algorithm, C4.5 or J48, was built using a multi-step process. First, the single variable was found which best splits the data into two groups. Second, the data were separated, and the process was repeated recursively until the subgroups either reached a maximum size of 5 or no further improvements were made. This strategy employed a splitting criterion known as the 'gain ratio', and was pruned using a bottom-up strategy known as 'error-based' pruning. Finally, accuracy and Area Under the Curve (AUC) were assessed to determine the reliability of the final tree and model.

Step Three: Interpretation of results

The tree selection from the algorithm was visualized using R, and a clinical meaning was teased from the results.

Step Four: Recommendation system

A user-input recommendation system was built in R for use by clinicians to assess patient's potential for developing mental health crises in the near future, based on patient history.

Results

Final Dataset

82.6% of respondents in the raw dataset reported never having been diagnosed with depression or anxiety, while 17.3% of respondents had been previously diagnosed. To alleviate the bias of the disparity in the class attribute, SMOTE was applied to the dataset to equalize the number of respondents who had and had not been diagnosed with a depressive disorder (Figure 1). This SMOTEd datatset included 33532 records and 131 attributes.

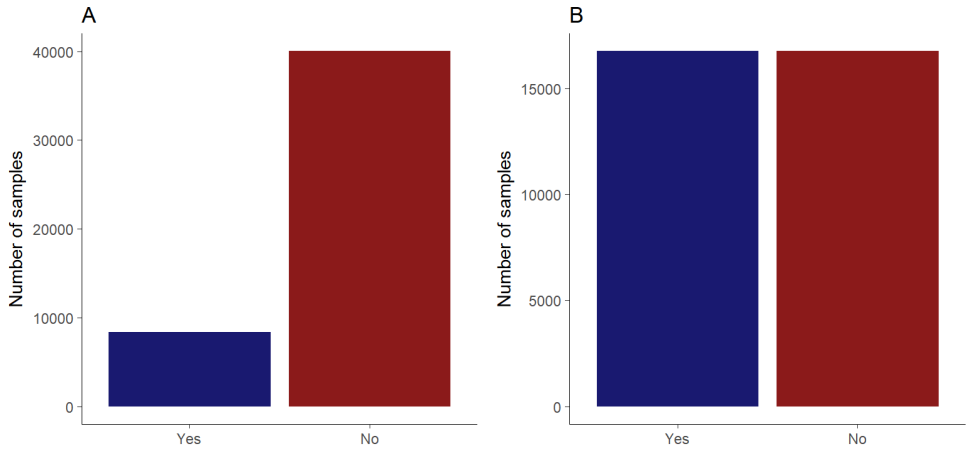


Figure 1: Number of samples before (A) and after (B) applying the Synthetic Minority Over-sampling Technique (SMOTE) to the mental health dataset. This balanced the class attribute to ensure downstream analyses were not influenced by an unbalanced class design.

To further preprocess the dataset, a threshold of 10% correlation of the feature:class attribute was calculated using Peason's correlation. This cleaning step produced a final dataset with 26 attributes (Table 1, page 7).

Model evaluation

Confusion Matrix and Accuracy

A confusion matrix summarizes the performance of the model. The confusion matrix for this specific model is reported in Table 1, and the model accuracy (calculated as (true observations/all observations)) was 81.07%.

Table 1: Confusion matrix for Mental Health model.

	Yes	**No**
Yes	34.9	3.8
No	15.1	46.2

ROC curve and AUC

The Area under the Curve (AUC) of the Receiver Operating Characteristic (ROC) is a good general indication of the performance of a model. The AUC value can range from 0.5 (the model performs no better than random chance) to 1 (model perfectly explains the response within the test set). For this model, the AUC was 0.83 (Figure 2).

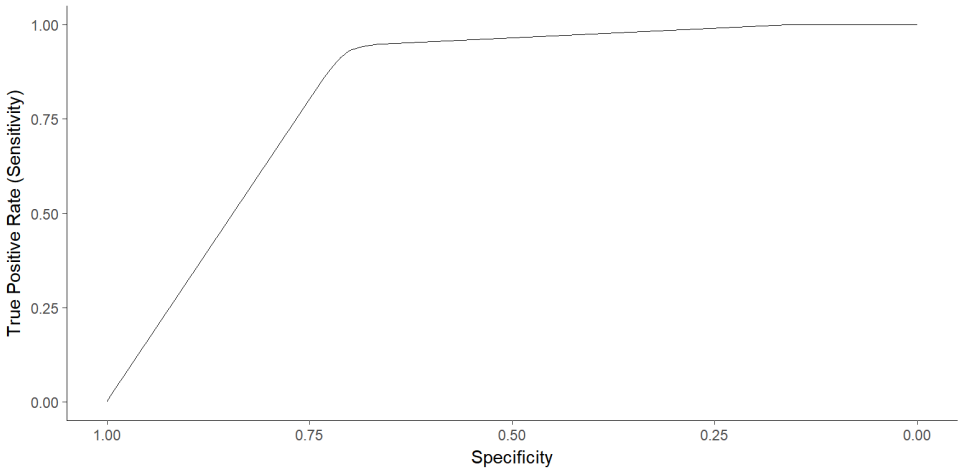


Figure 2: Area under the Curve (AUC) of the Receiver Operating Characteristic (ROC). In this model, the AUC is 0.83.

Table 2: Class and explanation of final attributes in the mental health dataset.

Attribute	Values	Correlation with class attribute
General Health	1: Excellent, 2: Very Good, 3: Good, 4: Fair, 5: Poor	-0.295607016
Multiple Healthcare Professionals?	1: Only one, 2: More than one; 3: None	0.103815018
Cost prohibiting seeing a doctor	1: Yes, 2: No	0.165653515
Participate in physical activities or exercise in past month	1: Yes, 2: No	-0.158701513
Ever told have asthma	1: Yes, 2: No	0.100176029
Ever told have COPD	1: Yes, 2: No	0.186787737
Ever told have arthritis	1: Yes, 2: No	0.237111272
Time of last visit to dentist or dental clinic	1: within the year, 2: within the past 2 years; 3: within the past 5 years; 4: 5 or more years ago	-0.149358224
Number of permanent teeth removed	1: 1-5; 2: 6 or more; 3: All; 8: None	0.139458736
Sex of Respondent	1: Male; 2: Female	-0.248546199
Marital Status	1: Married; 2: Divorced, 3: Widowed, 4: Separated, 5: Never married	-0.136709092
Education Level	1: Never attended; 2: Elementary; 3: Some high school, 4: High school graduate, 5: Some college or technical school, 6: College graduate	0.118860168
Own or rent home?	1: Own, 2: Rent; 3: Other arrangement	-0.251743782
Employment Status	1: Employed; 2: Self-employed; 3: Out of work for >1 year; 4: Out of work for <1 year; 5: Homemaker; 6: Student	-0.38432588
Blind or difficulty seeing	1: Yes, 2: No	0.251269243
Difficulty concentrating or remembering	1: Yes, 2: No	0.442148749
Difficulty walking or climbing stairs	1: Yes, 2: No	0.215583219
Difficulty dressing or bathing	1: Yes, 2: No	0.124759584
Difficulty doing errands alone	1: Yes, 2: No	0.245702066
Smoked at least 100 cigarettes in entire life	1: Yes, 2: No	0.109147887
Frequency of days currently smoking in a month	1: Every day, 2: Some days, 3: Not at all	0.133169638
Have delayed getting medical care	1: Yes, 2: No	0.16235243
Been without healthcare coverage in the past 12 months	1: Yes, 2: No	0.184359727
Activity has been limited due to health problems	1: Yes, 2: No	0.191887956
Having health problems that require special equipment	1: Yes, 2: No	0.165581842
Been diagnosed with depressive disorder (CLASS attribute)	1: Yes, 2: No	1

Top Attributes

Information gain measures how much 'information' a feature contributes knowledge about the class attribute. Therefore, attributes with higher information gain provide more information than attributes with lower information gain. The three attributes with the highest information gain for this model were:

1. Difficulty concentrating or remembering
2. Being without health care coverage in the past 12 months
3. Time since last visit to the dentist or dental clinic

The complete tree can be found in Appendix 1 of this document.

Recommendation System

A recommendation system was compiled to provide a user-interface program for use by doctors when their patients are in the examination room. This system can be found in the github folder of this report, and, while for brevity's sake the tree was truncated to a max depth of 3, a recommendation system based on the final tree should be programmed in the future.

Conclusion

Overall, the C4.5 model performed fairly well, with an accuracy of 81%. The model predicted 34.9% of mental crises correctly, and only misclassified 3.8% of mental crises as non-crises. This low FN rate is essential in a working model, as the 'cost' of misclassifying a mental crises is much higher than the cost of misclassifying a non-mental crises.

The features with the highest information gain provide interesting insights into the respondents' behaviours in this study. Depression and anxiety have been previously linked to early cognitive decline in the elderly (Sinoff & Werner, 2003, Blazer et al., 1987) and the decreased ability to concentrate is a

common symptom of depression (Nestler et al., 2002). Individuals with depression are also significantly less likely to continue with antidepressant therapy if they have no health insurance than individuals who have private insurance (Olfson et al., 2006), perhaps due to the higher healthcare costs associated with depression (Simon et al., 1995). Finally, the findings of this model support the results of previous studies, positively linking depressive disorders and levels of periodontal disease, and suggesting a negative correlation with tooth brushing and dental checkups to depression may exist (Antilla et al., 2006). The oral microbiome also may play an important role in human mental health, as highlighted through evidence of differences in bacterial community composition in individuals with schizophrenia than controls (Castro-Nallar et al., 2015), although the directionality of this relationship remains unresolved.

The results of the current study can be implemented to help highlight patents who are at risk of developing a mental health crises. The recommendation system based off this model can assist doctors in quickly identifying at-risk patients, leading to both higher rates of preventative healthcare and early intervention, ultimately lowering healthcare costs associated with treating depression and anxiety in the United States.

Future projects should focus on increasing overall accuracy of the model to ensure reliability when providing doctors direction in regards to their patients' mental health.

References

Antilla, S., Knuuttila, M, Ylostalo, P, & Joukamaa M. (2006). Symptoms of depression and anxiety in relation to dental health behavior and self-perceived dental treatment need. European Journal of Oral Science. 114; 109-14.

Behavioural Risk Factor Surveillance System – Centres for Disease Control and Prevention. (2015). 2016 Behavioural Risk Factor Surveillance System Questionnaire. Access to pdf here: https://www.cdc.gov/brfss/questionnaires/pdf-ques/2016brfss_questionnaire_10_14_15.pdf

Blazer, D., Hughes, DC, & George, LK. (1987). The epidemiology of depression in an elderly community population. Gerontologist. 27(3); 281-287.

Castro-Nallar, E, Bendall, ML, Perez-Losada, M, …, & Crandall, KA. (2015). Composition, taxonomy and functional diversity of the oropharynx microbiome in individuals with schizophrenia and controls. PeerJ. 3; e1140.

Chawla, NV, Bowyer, KW, Hall, LO, Kegelmeyer, WP. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. 16; 321-357.

Diepgen, TL, & Mahler, V. (2002). The epidemiology of skin cancer. British Journal of Dermatology. 146(s61); 1-6.

Hayes, DK, Greenlund, KJ, Denny CH, …, & Keenan, NL. (2005). Racial/Ethnic and Socioeconomic Disparities in Multiple Risk Factors for Heart Disease and Stroke – United States, 2003. Morbidity and Mortality Weekly Report – Centers for Disease Control and Prevention. 54(05); 113-117.

Nestler, EJ, Barrot, M, DiLeone, RJ, Eisch, AL, Gold, SJ, & Monteggia, LM. (2002). Neurobiology of depression. Neuron. 34(1); 13-25.

Olfson, M, Marcus, SC, Tedeshi, M, Wan, GJ. (2006). Continuity of Antidepressant Treatment for adults with depression in the United States. American Journal of Psychiatry. 163(1); 101-108.

Simon, GE, VonKorff, M, & Barlow, W. (1995). Health care costs of primary care patients with recognized depression. JAMA Psychiatry. 52(10); 850-856.

Sinoff, G, & Werner, P. (2003). Anxiety disorder and accompanying subjective memory loss in the elderly as a predictor of future cognitive decline. Geriatric Psychiatry. 18(10); 951-959.

Strine, TW, Mokdad, AH, Ballux LS, Gonzalez, O, Crider, R, Berry, JT, & Kroenke, K. (2008). Depression and anxiety from the United States: findings from the 2006 behavioural risk factor surveillance system. Psychiatric services. 59(12); 1383-1390.

Yoon, S, Taha B, & Bakken, S. (2014). Using a Data Mining Approach to discover Behaviour Correlates of Chronic Disease: A Case Study of Depression. Studies in Health Technology and Informatics 201; 71-78.

Appendix 1: Final Mental Health Crises C4.5 Decision Tree

```
DIFFICULTY.CONCENTRATING.OR.REMEMBERING <= 1.8973: 1 (9259.0/846.0)

DIFFICULTY.CONCENTRATING.OR.REMEMBERING > 1.8973

|   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON <= 1.949175

|   |   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON <= 1

|   |   |   RESPONDENTS.SEX <= 1.227905: 2 (34.0/8.0)

|   |   |   RESPONDENTS.SEX > 1.227905

|   |   |   |   EMPLOYMENT.STATUS <= 6.999589

|   |   |   |   |   MULTIPLE.HEALTH.CARE.PROFESSIONALS <= 2.910701

|   |   |   |   |   |   GENERAL.HEALTH <= 1.497093: 2 (6.0)

|   |   |   |   |   |   GENERAL.HEALTH > 1.497093: 1 (41.0/10.0)

|   |   |   |   |   MULTIPLE.HEALTH.CARE.PROFESSIONALS > 2.910701: 2 (14.0/4.0)

|   |   |   |   EMPLOYMENT.STATUS > 6.999589: 2 (7.0/1.0)

|   |   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON > 1: 1 (1180.0)

|   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON > 1.949175

|   |   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC <= 4

|   |   |   BLIND.OR.DIFFICULTY.SEEING <= 1.947215

|   |   |   |   BLIND.OR.DIFFICULTY.SEEING <= 1.000196

|   |   |   |   |   RESPONDENTS.SEX <= 1.030521: 2 (310.0/88.0)

|   |   |   |   |   RESPONDENTS.SEX > 1.030521

|   |   |   |   |   |   DIFFICULTY.WALKING.OR.CLIMBING.STAIRS <= 1.91232

|   |   |   |   |   |   |   DIFFICULTY.WALKING.OR.CLIMBING.STAIRS <= 1.030688

|   |   |   |   |   |   |   |   MARITAL.STATUS <= 2.620818: 1 (145.0/43.0)

|   |   |   |   |   |   |   |   MARITAL.STATUS > 2.620818

|   |   |   |   |   |   |   |   |   COULD.NOT.SEE.DR..BECAUSE.OF.COST <= 1.52357: 1 (23.0/5.0)

|   |   |   |   |   |   |   |   |   COULD.NOT.SEE.DR..BECAUSE.OF.COST > 1.52357

|   |   |   |   |   |   |   |   |   |   NUMBER.OF.PERMANENT.TEETH.REMOVED <= 6.33735: 2
(85.0/32.0)

|   |   |   |   |   |   |   |   |   |   NUMBER.OF.PERMANENT.TEETH.REMOVED > 6.33735

|   |   |   |   |   |   |   |   |   |   |   DIFFICULTY.DOING.ERRANDS.ALONE <= 1.499765: 2
(2.0)

|   |   |   |   |   |   |   |   |   |   |   DIFFICULTY.DOING.ERRANDS.ALONE > 1.499765: 1 (5.0)

|   |   |   |   |   |   |   DIFFICULTY.WALKING.OR.CLIMBING.STAIRS > 1.030688: 1 (15.0/1.0)

|   |   |   |   |   |   DIFFICULTY.WALKING.OR.CLIMBING.STAIRS > 1.91232: 2 (250.0/105.0)
```

```
|   |   |   |   BLIND.OR.DIFFICULTY.SEEING > 1.000196: 1 (363.0)
|   |   |   BLIND.OR.DIFFICULTY.SEEING > 1.947215
|   |   |   |   EMPLOYMENT.STATUS <= 6.017024
|   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING <= 1.999682
|   |   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING <= 1.897369: 2 (720.0/191.0)
|   |   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING > 1.897369: 1 (239.0)
|   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING > 1.999682
|   |   |   |   |   |   BLIND.OR.DIFFICULTY.SEEING <= 1.998253
|   |   |   |   |   |   |   BLIND.OR.DIFFICULTY.SEEING <= 1.948813: 2 (37.0/7.0)
|   |   |   |   |   |   |   BLIND.OR.DIFFICULTY.SEEING > 1.948813: 1 (45.0)
|   |   |   |   |   |   BLIND.OR.DIFFICULTY.SEEING > 1.998253: 2 (18648.0/4439.0)
|   |   |   |   EMPLOYMENT.STATUS > 6.017024
|   |   |   |   |   ACTIVITY.LIMITATION.DUE.TO.HEALTH.PROBLE <= 1.998707
|   |   |   |   |   |   EVER.TOLD.HAD.ASTHMA <= 1.479631
|   |   |   |   |   |   |   COULD.NOT.SEE.DR..BECAUSE.OF.COST <= 1.020158
|   |   |   |   |   |   |   |   FREQUENCY.OF.DAYS.NOW.SMOKING <= 1.499135: 1 (8.0)
|   |   |   |   |   |   |   |   FREQUENCY.OF.DAYS.NOW.SMOKING > 1.499135
|   |   |   |   |   |   |   |   |   MULTIPLE.HEALTH.CARE.PROFESSIONALS <= 2.495795
|   |   |   |   |   |   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING <= 1.948625: 1
(5.0/1.0)
|   |   |   |   |   |   |   |   |   |   DIFFICULTY.CONCENTRATING.OR.REMEMBERING > 1.948625: 2
(43.0/13.0)
|   |   |   |   |   |   |   |   |   MULTIPLE.HEALTH.CARE.PROFESSIONALS > 2.495795: 1 (4.0)
|   |   |   |   |   |   |   COULD.NOT.SEE.DR..BECAUSE.OF.COST > 1.020158: 1 (206.0/45.0)
|   |   |   |   |   |   EVER.TOLD.HAD.ASTHMA > 1.479631
|   |   |   |   |   |   |   SMOKED.AT.LEAST.100.CIGARETTES <= 1.33024
|   |   |   |   |   |   |   |   EDUCATION.LEVEL <= 3.498361
|   |   |   |   |   |   |   |   |   OWN.OR.RENT.HOME <= 2.4999
|   |   |   |   |   |   |   |   |   |   DELAYED.GETTING.MEDICAL.CARE <= 6.087: 1 (4.0)
|   |   |   |   |   |   |   |   |   |   DELAYED.GETTING.MEDICAL.CARE > 6.087
|   |   |   |   |   |   |   |   |   |   |   DELAYED.GETTING.MEDICAL.CARE <= 7.587271: 2
(63.0/23.0)
|   |   |   |   |   |   |   |   |   |   |   DELAYED.GETTING.MEDICAL.CARE > 7.587271
|   |   |   |   |   |   |   |   |   |   |   |   FREQUENCY.OF.DAYS.NOW.SMOKING <= 1.499135: 2
(2.0)
|   |   |   |   |   |   |   |   |   |   |   |   FREQUENCY.OF.DAYS.NOW.SMOKING > 1.499135: 1
(5.0)
```

```
|   |   |   |   |   |   |   |   |   OWN.OR.RENT.HOME > 2.4999: 1 (4.0)
|   |   |   |   |   |   |   |   EDUCATION.LEVEL > 3.498361: 1 (351.0/123.0)
|   |   |   |   |   |   |   SMOKED.AT.LEAST.100.CIGARETTES > 1.33024
|   |   |   |   |   |   |   |   GENERAL.HEALTH <= 3.042237
|   |   |   |   |   |   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST <= 1.97594: 1
(7.0/2.0)
|   |   |   |   |   |   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST > 1.97594: 2
(125.0/37.0)
|   |   |   |   |   |   |   |   GENERAL.HEALTH > 3.042237
|   |   |   |   |   |   |   |   |   EDUCATION.LEVEL <= 5.434887
|   |   |   |   |   |   |   |   |   |   NUMBER.OF.PERMANENT.TEETH.REMOVED <= 1.499735
|   |   |   |   |   |   |   |   |   |   |   MARITAL.STATUS <= 3.499295: 1 (43.0/12.0)
|   |   |   |   |   |   |   |   |   |   |   MARITAL.STATUS > 3.499295: 2 (8.0/1.0)
|   |   |   |   |   |   |   |   |   |   NUMBER.OF.PERMANENT.TEETH.REMOVED > 1.499735: 2
(119.0/53.0)
|   |   |   |   |   |   |   |   |   EDUCATION.LEVEL > 5.434887
|   |   |   |   |   |   |   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST <= 1.47823:
2 (3.0)
|   |   |   |   |   |   |   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST > 1.47823: 1
(22.0/2.0)
|   |   |   |   |   ACTIVITY.LIMITATION.DUE.TO.HEALTH.PROBLE > 1.998707: 2 (119.0/41.0)
|   |   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC > 4
|   |   |   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC <= 7.997777: 1 (800.0)
|   |   |   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC > 7.997777
|   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST <= 1.802312: 1 (13.0/2.0)
|   |   |   |   X.EVER.TOLD..YOU.HAVE..COPD..CHRONIC.OBST > 1.802312: 2 (150.0/28.0)


Number of Leaves  :          43


Size of the tree :     85
```

Appendix 2: Truncated tree for recommendation system

```
DIFFICULTY.CONCENTRATING.OR.REMEMBERING = 1
|   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON < 2
|   |   BLIND.OR.DIFFICULTY.SEEING < 1.95 : 1 (2472/122)
|   |   BLIND.OR.DIFFICULTY.SEEING >= 1.95 : 1 (2687/786)
|   WITHOUT.HEALTH.CARE.COVERAGE.PAST.12.MON >= 2
|   |   TOLD.HAVE.ARTHRITIS < 2 : 1 (4319/452)
DIFFICULTY.CONCENTRATING.OR.REMEMBERING = 2
|   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC < 1
|   |   LENGTH.OF.TIME.SINCE.LAST.ROUTINE.CHECKU < 1 : 2 (12737/4202)
|   |   LENGTH.OF.TIME.SINCE.LAST.ROUTINE.CHECKU >= 1 : 2 (4219/1904)
|   LAST.VISITED.DENTIST.OR.DENTAL.CLINIC >= 1
|   |   X.EVER.TOLD..YOU.HAVE.DIABETES = 2 : 1 (3033/742)
|   |   X.EVER.TOLD..YOU.HAVE.DIABETES = 1 : 2 (7907/3903)
```