What Predicts Whether a Person Will Be Unemployed Next Year?

Robert Winslow

2024

Details About the Task

- Binary prediction about whether each person will be unemployed in one year's time.

- Unbalanced data: Only 5 percent of individuals will be unemployed in one year's time.

- The competition's scoring metric placed equal weight on accurate predictions of unemployment and accurate predictions of non-unemployment:

  - $GF \equiv \frac{\text{\# Correctly Predicted Unemployed}}{\text{\# Unemployed}} \cdot \frac{1}{2} + \frac{\text{\# Correctly Predicted Not Unemployed}}{\text{\# Not Unemployed}} \cdot \frac{1}{2}$

- Data is drawn from the CPS outgoing rotation groups
  - people aged 20-64
  - years 2008-2014

Score Comparison for Different Models
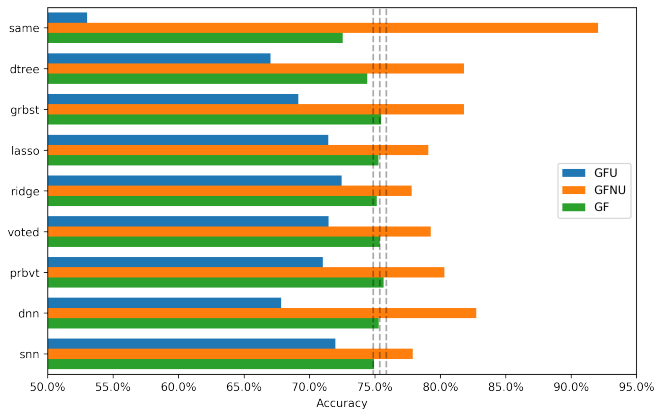
Scores for Simple Heuristic: Assume Empstat Doesn't Change

Accuracy in Predicting Unemployment: 53.0%

Accuracy in Predicting Non-Unemployment: 92.1%

Balanced Accuracy: 72.5%

Scores for Decision Tree

Accuracy in Predicting Unemployment: 67.0%

Accuracy in Predicting Non-Unemployment: 81.8%

Balanced Accuracy: 74.4%

Scores for Gradient Boosted Decision Tree

Accuracy in Predicting Unemployment: 69.2%

Accuracy in Predicting Non-Unemployment: 81.8%

Balanced Accuracy: 75.5%

Scores for Regularized Regression (Lasso)

Accuracy in Predicting Unemployment: 71.5%

Accuracy in Predicting Non-Unemployment: 79.1%

Balanced Accuracy: 75.3%

Scores for Regularized Regression (Ridge)

Accuracy in Predicting Unemployment: 72.5%

Accuracy in Predicting Non-Unemployment: 77.8%

Balanced Accuracy: 75.1%

Scores for 2-out-of-3 Vote

Accuracy in Predicting Unemployment: 71.5%

Accuracy in Predicting Non-Unemployment: 79.3%

Balanced Accuracy: 75.4%

Scores for 2-out-of-3 Vote (With Gradient Boosting)

Accuracy in Predicting Unemployment: 71.0%

Accuracy in Predicting Non-Unemployment: 80.3%

Balanced Accuracy: 75.7%

Scores for Shallow Neural Net

Accuracy in Predicting Unemployment: 72.0%

Accuracy in Predicting Non-Unemployment: 77.9%

Balanced Accuracy: 74.9%

Scores for Deep Neural Net
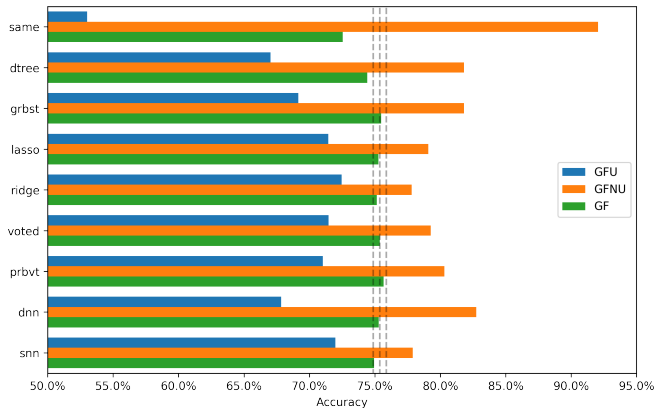
Accuracy in Predicting Unemployment: 67.8%

Accuracy in Predicting Non-Unemployment: 82.7%
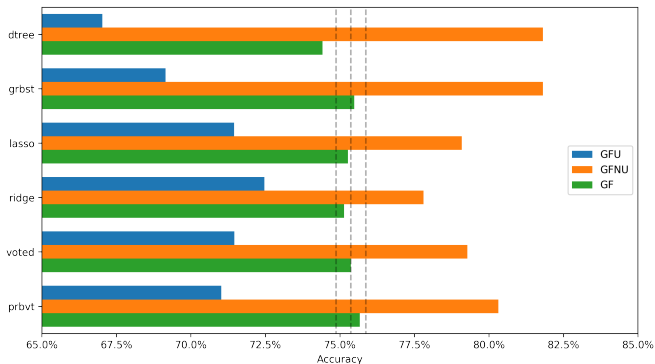
Balanced Accuracy: 75.3%

Score Comparison for Different Models
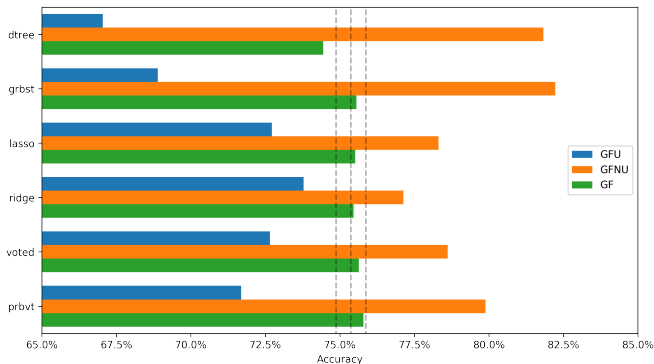
Does Adding Extra Features from CPS Help?

Does Adding Extra Features from CPS Help?

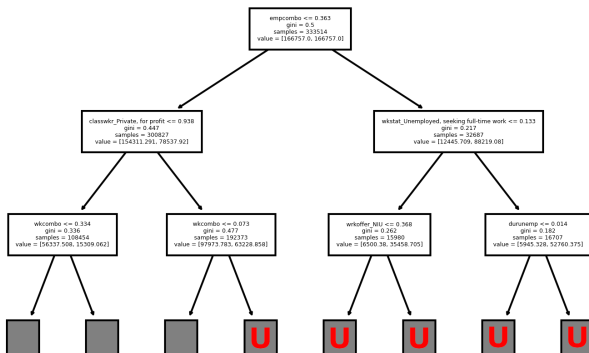**Using only variables from the MEBDI sample:**

Using additional variables from IPUMS CPS:

An Example Small Decision Tree

Some Large Coefficients in Lasso

Coefficient	Variable Description
0.092	Unemployed and Seeking Full Time Work
0.064	NILF or Unemployed
0.050	Last worked full-time over a year ago
0.040	Family Income Under $5000
0.030	Industry: Personnel supply services
...	
-0.013	High School Diploma or higher
-0.014	Race: Asian
-0.015	Race: White
-0.025	Re-entered labor force last week
-0.025	In the Armed Forces

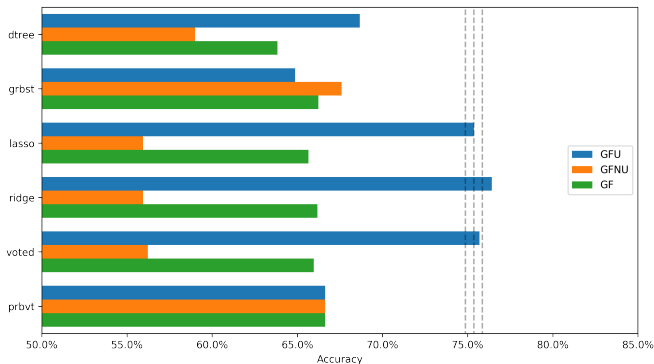# Restricting Analysis to those who are Employed

Using the entire sample:

Restricting the sample to only the employed:

Big Lasso Coefficients (When Looking at Only the Employed)

Coefficient	Variable Description
0.031	Industry: Personnel supply services
0.021	Works Part-time for Economic Reasons
0.018	Family Income Under $5000
0.015	Occ: Construction, except supervisors
0.013	Family Income $5000-$8000
...	
-0.008	Industry: Hospitals
-0.008	Married with Spouse Present
-0.009	Race: Asian
-0.009	Race: White
-0.010	Same Employer as Last Month