

# Latino Trump Support: ML Analysis Summary

## CMPS 2016 | Random Forest Classification

### Study Overview

DV: Trump vote (1) vs Non-Trump vote (0) among Latino voters

Data: CMPS 2016, Latino respondents who voted

Method: Random Forest (500 trees, balanced class weights)

### Sample Characteristics

Total Latino Voters:	3,001
Trump Voters (DV=1):	539 (18.0%)
Non-Trump Voters (DV=0):	2,462 (82.0%)
Features (one-hot encoded):	1,517
Train/Test Split:	80/20 stratified

### Model Performance

#### Full Model (all 1,517 features)

Test ROC-AUC:	0.9383
---------------	--------

#### Non-Partisan Model (1,471 features)

Test ROC-AUC:	0.8736
AUC Drop:	0.0647 (6.9% relative)

Note: Non-partisan model excludes 46 variables measuring party ID, ideology, and candidate favorability ratings.

### Key Findings

1. Trump favorability is strongest single predictor
2. Party ID and partisan leanings dominate full model
3. Removing partisan vars reduces AUC by only 6.9%
4. Non-partisan model reveals immigration attitudes:
  - Views on immigration levels (increase/decrease)
  - Citizenship pathway for undocumented
  - Attitudes toward hearing Spanish spoken
  - Support for deportation policies
5. Strong predictive power (AUC 0.87) without partisan vars

## Top 15 Predictors - Full Model (AUC: 0.9383)

Rank	Feature	Importance	Std
1	C4_(4) Very unfavorable	0.017	0.0033
2	C242_HID_(2) Democratic Part	0.0062	0.0021
3	L46_(2) Democrat	0.0047	0.0025
4	C2_(4) Very unfavorable	0.0029	0.002
5	L266_(2) Democrats	0.0022	0.0016
6	C41_(5) Strongly disagree	0.0021	0.0007
7	C9_(4) Very unfavorable	0.002	0.0018
8	C25_(2) Democrat	0.0019	0.0014
9	C337_(2) Increase	0.0016	0.0008
10	C4_(2) Somewhat favorable	0.0015	0.0009
11	C2_(3) Somewhat unfavorabl	0.0015	0.0005
12	C142_(3) Bothers me a lot	0.0009	0.0006
13	C115_(4) Never	0.0009	0.0004
14	C114_(4) Never	0.0008	0.0004
15	C45_(5) Strongly disagree	0.0008	0.0004

## Top 15 Predictors - Non-Partisan Model (AUC: 0.8736)

Rank	Feature	Importance	Std
1	C41_(5) Strongly disagree	0.0062	0.0023
2	C337_(2) Increase	0.0059	0.0025
3	C158_(02) Should not	0.0057	0.0023
4	C142_(3) Bothers me a lot	0.004	0.0016
5	C38_(3) They should be required to leave their jobs and	0.0038	0.0014
6	C228_(5) Strongly oppose	0.0038	0.0025
7	C45_(5) Strongly disagree	0.0036	0.0012
8	B1A207_(4) Strongly disagree	0.0034	0.0022
9	C115_(4) Never	0.0028	0.001
10	C337_(3) Stay the same	0.0028	0.0009
11	C40_(5) Strongly disagree	0.0024	0.001
12	C111_(4) Never	0.0021	0.001
13	C112_(4) Never	0.0021	0.0006
14	L241_(5) Strongly oppose	0.0021	0.001
15	C42_Other	0.0021	0.0013

# Appendix: Variables & Methods

## Partisan Variables Removed (46 columns total)

Variable	Description
C25	Party registration
C26	Strong partisan identification
C31	Ideology (liberal-conservative scale)
L46	Which party better on immigration
L266	Which party better for Latinos
L267	Which party better on values
C2	Hillary Clinton favorability
C3	Bernie Sanders favorability
C4	Donald Trump favorability
C5	Ted Cruz favorability
C8	Bill Clinton favorability
C9	Barack Obama favorability
C242_HID	Party identification (derived)

## Interpretation Notes

Permutation Importance: Measures feature importance by shuffling each feature and measuring AUC decrease.  
(50 repeats for stability)

SHAP Values: Measures average contribution of each feature to individual predictions.  
(Computed on 500 test observations)

Report generated: 2025-11-30 05:30