Voting Behavior in Italy: A Comprehensive Analysis and Efficient Prediction

Irene Cozzi, Laura Guarnaschelli, Davide Magliano, Alessia Vanzella

May 9, 2025

1 Introduction

Elections are a cornerstone of democratic processes, and understanding voter behavior is crucial for political science, economics, and social policy. In every country, voter choice is influenced by various factors, including socioeconomic status, political beliefs, and demographic characteristics. However, the interdependence of these variables make it challenging to predict voting behavior. Traditional political science models often rely on surveys or econometric analyses to assess the drivers of electoral preferences, but machine learning techniques offer a promising approach to predict voting patterns using large-scale datasets and non-linear, high-dimensional relationships to identify hidden patterns. Furthermore, to the authors' knowledge, no paper analyzes voters' characteristics in the last Italian national elections. We address this gap by applying four methodological approaches: decision tree classification, random forest, gradient boosting, and multinomial logistic regression. The findings reveal that the most predictive features across models are attitudes towards immigration, satisfaction with democracy, age and media consumption.

2 Literature Review

The academic literature on political cleavages and voter characteristics offers extensive insights into the factors that shape electoral choices. Demographic variables such as age and gender, along with socioeconomic indicators like employment status, income level, and education, significantly influence party preferences across various contexts. Tilley and Evans (2014) show that older generations tend to be more conservative within any given election period. In Norway, Geys et al. (2022) suggest that individuals shift towards the right when they are over 55 years old, while individuals under 40 diverge from their parents' political choices, showing a greater propensity to support left-wing parties. While there is substantial literature examining how age affects political participation and voter turnout, particularly the tendency of older citizens to view voting as a civic duty (Carlsson & Johansson-Stenman, 2010), little attention has been paid to the specific party preferences of young voters in Italy. Existing studies often focus more on youth political activism (Genova & Tazzara, 2025) or attitudes toward the European Union (Pendenza & Verderame, 2019), leaving a gap in our understanding of how young Italians align with domestic party platforms.

A growing body of research highlights a persistent gender gap in political preferences. Over the past decades, women's preferences have shifted increasingly leftward in the OECD. One widely accepted explanation is that working women are more inclined to support policies such as public spending on childcare and eldercare, making them more likely to vote for left-leaning parties (Iversen et al., 2005). Additionally, following the implications of the socialization hypothesis outlined by Inglehart and Norris (2000), contemporary women increasingly adopt expectations of financial independence from a young age, shaping long-term political preferences that are deeply ingrained in society and less susceptible to short-term economic or situational factors, as it can be for divorce. Nonetheless, labour market risks, particularly rising employment insecurity driven by globalization and deindustrialization, represent a big factor determining increased support for redistribution and

traditionally left-wing parties, whose intensity and direction often depend on the nature of an individual's employment status; for instance, temporary workers may be more inclined toward "new" left parties (green and other left-libertarian parties), while permanent employees tend to favor moderate left-leaning parties (Marx, 2014).

A similar line of reasoning applies to the concept of material welfare. As McCarty et al. (2006) point out, there are sharp differences in partisan voting by income. All else equal, greater income inequality, particularly in Italy, where economic disparities have a strong geographical component tied to the long-standing issue of *La Questione Meridionale*, should intensify the divergence of economic interests between the rich and the poor. Consequently, this divergence often manifests in distinct party alignments, with lower-income voters more inclined to support left-leaning redistributive platforms, and higher-income individuals gravitating toward market liberalism.

Additionally, following the basic economic theory of positive returns to education, we should expect that income and education are significantly positively correlated. However, while the quality of education (i.e., field of study or institution) is often more predictive of party choice than the quantity of it (i.e. number of years of schooling), a large body of empirical research has found that individuals with higher levels of education are generally less prejudiced toward ethnic out-groups (Vogt, 1997; Schuman et al., 1998). This suggests that higher educational attainment may act as a barrier against support for extreme right parties, regardless of individuals' positions in the labor market. Yet, these findings are not uniform across contexts. In France and Germany, for example, some scholars have identified an intellectual elite within the extreme right, particularly among disaffected academic white men who express resentment toward feminist progress and believe in a left-wing media conspiracy (Müller, 2003; Rynkiewicz et al., 2016). However, it is also important to note that the perception of media conspiracy may also depend on the type and frequency of media consumed. For example, in the USA, the initial spread of the TV in the 1950s crowded out time previously devoted to consuming political information via more established media, such as radio and local newspapers, reducing voter knowledge and turnout (Gentzkow, 2006). Conversely, the most recent influence of partisan networks like Fox News (U.S.) or Mediaset (Italy) - the country's largest private broadcaster, majority-owned by the Berlusconi family - raised voters' political engagement and support for the respective right-leaning parties (DellaVigna & Kaplan, 2007; Durante et al., 2019). Importantly, the effect of media consumption on preferences is not unilateral, but it tends to reinforce previous political positions (Barberá, 2020).

Given our focus on the Italian context, it is crucial to consider cultural attitudes, particularly individuals' levels of religiosity and their views on immigration. The relationship between politics and Catholicism in Italy has historically been robust, exemplified by the dominance of *Democrazia Cristiana* (DC), the largest Catholic party in Europe, which played a central role in Italian politics throughout much of the post-war period until 1994 (Di Battista et al., 2018). To overlook religion would be to neglect a potentially significant factor in shaping moral reasoning and political preferences. More recently, since the early 2000s, Italy has become a major entry point for migrants from the Mediterranean. This shift has heightened the salience of immigration in the political discourse. Far-right parties, particularly Lega Nord, have capitalized on local concerns about immigration, security, and national identity to expand their electoral base. Immigration still continues to be a polarizing issue: for instance, in early June 2025, Italian citizens will vote in a referendum proposing to reduce the minimum legal residency requirement for migrants seeking Italian citizenship, offering insight into the prevailing public attitudes toward migrants.

Lastly, recent polling data by European Movement International from January 2025 reveals a troubling picture of democratic engagement in Italy. Only 34% of Italian citizens express consistent support for democracy, compared to 68% in Denmark. Most strikingly, in Italy a majority of respondents does not consider the European Union important for the country's future. These trends echo the findings of our analysis. In the final section of this paper, we propose a set of policy recommendations aimed at restoring public trust in democratic institutions and in European integration.

3 Data

For our analysis we employ the 11th release of the **European Social Survey (ESS)**, which was conducted in 2023-2024 and is publicly available. The ESS is an academically driven, cross-national survey, founded in 2001. To date, the one-hour face-to-face interviews have been conducted in 40 countries, covering topics such as social conditions, behaviors and attitudes, general health, political views, cultural, religious and national identity, as well as media usage. Our focus is set on Italy. The outcome variable in the dataset is represented by the National Party voted at the last election (the national elections of September 2022) by the individual, defined by the variable *prtvteit (=VotedParty)*. The outcome of this variable was then re-classified in four main coalitions in which parties presented themselves at the elections, going from relatively left-wing to right-wing platforms:

- **Centre-Left (CSX)**: Democratic Party, Alleanza Verdi e Sinistra, Unione Popolare, +Europa
- **Movimento 5 Stelle (M5S)**
- **Terzo Polo (3POLO)**: Azione, Italia Viva
- **Centre-Right (CDX)**: Fratelli d'Italia, Lega, Forza Italia

This classification aims to reduce heterogeneity in the party voted, provided that parties of the same coalitions have very similar characteristics and the voting decision among them reflects less coherent and meaningful motivations than the core vote intention and political preference, which is better captured by the coalition selected. The comprehensive survey presents 640 variables, of which we retained 16, based on a thorough analysis of the literature on voting predictors. The selected variables can be divided in four main categories:

- Demographic factors: Respondent's age, Respondent's gender, Region (five macro-areas).
- Socio-Economic factors: Employment Status, Trade Union Membership, Education Level, Household Income.
- Political Engagement: Interest in Politics, Left-Right Political Orientation, Satisfaction with Democracy, Trust in European Parliament, Media Consumption.
- Social/Cultural Attitudes: Religion, Attitudes Toward Immigration, Gender equality attitudes.

After dealing with missing data, the final dataset consists of 837 units of observation. Appendix A shows the summary statistics. Appendix B shows that the Left-Right self-placement variable is an overwhelmingly dominant predictor of the Voted Coalition, in line with the results of several other papers (Kuenzi & Lambright, 2011; Norris, 2002). While this confirms the accuracy of our party classification into coalitions we exclude the variable from the final prediction analysis.

4 Methodology and Empirical analysis

In our analysis, we employ four methodologies:

- Decision Tree Classifier
- Random Forest
- Boosting (XGBoost)
- Multinomial Logistic Regression

The choice of the methods is dictated by the nature of our independent and dependent variables, which are all categorical, mostly discrete ordinal. Before applying these methods to our datasets, we investigated it deeper in order to obtain the most accurate and robust results. First of all, as noted in the univariate analysis (Appendix C), there is a strong imbalance between the outcome variable's classes (Voted Coalition) in terms of votes received. This is a relevant issue given that most classifiers aim to maximize accuracy, and if one class dominates, the model might always predict that class to boost its accuracy. For instance, our preliminary results yield very high accuracy, which is however driven by the two major classes (CDX and CSX), while algorithms almost never predict correctly the two classes with fewer votes (M5S and 3POLO), as they are assigned very low probability and misclassified. Therefore, we tried (i) oversampling methodology for increasing minority classes (SMOTE[1]) as well as (ii) bootstrapping (random resampling with replacement) and

[1]The Synthetic Minority Over-sampling Technique creates new synthetic observations for the minority class by interpolating between existing ones (taking the k-nearest neighbors).

(iii) weighted random forests, applied only to the training set, with the first method yielding the best results. Secondly, to ensure sample representativeness when splitting the dataset into training and test sets, in terms of proportion of votes for each coalition, we employed a stratification method in all of our four models. Finally, since in most methodologies there are several hyperparameters to be selected, we use cross-validation methods to ensure the best combination of them. After having verified a convergence between GridSearch and RandomSearch, we opted for the former given that the dimension of our dataset is limited and it is more exhaustive.

4.1 The Decision Tree Classifier

The Decision Tree is the simplest method in terms of interpretation, as it allows an immediate visualisation of the most relevant variables in defining individuals' voting patterns. When no technique is applied to handle class imbalance, the accuracy is comparably higher, driven by the two biggest classes, as shown by the confusion matrix (Appendix D). We therefore show the ultimate improved Decision Tree with Cross-Validation and SMOTE (Appendix E). We observe that Attitudes Towards Immigration are the most important variable: individuals with more negative views on immigration, as well as those who report higher satisfaction with democracy, are more likely to vote for the Center-Right coalition. Age and Region are also importantly used, though without a clear indication. Overall, the accuracy with cross-validation and oversampling is slightly below 50%, and the selected tree is quite deeper, losing some interpretability.

4.2 Random Forests

A Random Forest is an ensemble learning technique that combines multiple decision trees to improve accuracy and robustness. The algorithm builds multiple decision trees using bootstrap sampling (randomly selecting data with replacement) and each tree is trained on a different subset of the training data. At each split, only a random subset of features is considered (hence, it reduces correlation between trees), and finally predictions from all trees are aggregated to reach a final decision. The Forest represents an improvement to the Decision Tree as it reduces overfitting and it is more accurate and stable. It is however more computationally expensive and less interpretable, given that we cannot plot a single tree. Therefore, results are presented in terms of 1) feature importance and 2) confusion matrices. Results are portrayed in Appendix F. We observe that, in addition to the previously discussed "Attitudes towards immigration" and "Satisfaction Democracy", two others emerge as relevant predictors: Age and Media Consumption. Furthermore, the accuracy is between 60 and 70%, higher than that of the decision tree outcomes, and the model manages to predict correctly also some of the observations from the less represented classes.

4.3 XGBoost

While Random Forests builds many decision trees in parallel (i.e., independently), with each tree seeing a random subset of the data and feature, Boosting builds trees sequentially, each one trying to fix errors in the previous ones provided that the model gives more weight to harder-to-classify observations. The final prediction is a weighted combination of all produced trees. This method is generally more effective at handling imbalanced data. We use XGBoost, a gradient boosting algorithm with performance improvements. Results are presented in terms of confusion matrices and feature importance, where the F-score represents how much each variable affects the model's ability to make optimized predictions. Results (Appendix G) are overall similar to the Random Forest's both in terms of accuracy and variable importance. However, we observe that Media Consumption overcomes Satisfaction Democracy and Attitudes Towards Immigration in terms of importance, as well as Religion, while Household Income becomes another relevant feature.

4.4 Multinomial Logistic Regression

Finally, we implemented a Multinomial Logistic Regression. Firstly, we train the MLR using cross-validation to get the best parameters and observe the prediction capacity of the model, with the outcome portrayed in terms of the confusion matrix (Appendix H). Secondly, we moved to a simple

regression on the overall dataset for statistical inference and coefficient interpretation: the summary results table (Appendix H) presents the standardised effects, displayed through the coefficients in log-odds terms. Therefore, the direction is easier to interpret through the sign, while the magnitude requires a transformation of the coefficients: the Appendix and the Notebook provide deeper insights.

Results show that the Center-Right, M5S and 3POLO coalitions are correlated with younger and more religious voters compared to the Center-Left (CSX, baseline): as age increases, and self-declared religiousness decreases, the log-odds of voting for those coalitions decrease. Compared to CSX, CDX voters are linked with lower trust in the European Parliament, higher satisfaction with democracy and worse attitude toward immigration. As these attitudes improve, the odds of voting for M5S and 3POLO decrease, and fall substantially for CDX, in favor of CSX. No significant results are retrieved on Media Consumption, possibly due to its exposure-polarising (echo-chamber) nature discussed in Literature (Section 2). In interpreting our estimates we should always take into consideration the overall lower representation of M5S and 3POLO classes. For this reason, we chose to retain the Center-Left coalition (CSX) as the baseline category, as it provides a clearer basis for comparison with the other major bloc, the Center-Right (CDX).

5 Conclusion and Policy Recommendations

Overall, the best performances in terms of accuracy appear from Random Forests and Boosting models, using cross-validation, stratification and oversampling (results provided by the three different techniques to handle imbalance are similar). Although models yield slightly different results based on the specific methods and enhancements applied, we can draw broad conclusions on the most relevant explanatory variables. Whereas factors such as employment status and gender appear to have limited impact at the ballot box, political engagement and cultural attitudes, such as Attitudes Towards Immigration and Satisfaction with Democracy, systematically explain Italians' voting behavior in the 2022 elections. For instance, people with stricter views on immigration tend to support more Center-Right and Movimento 5 stelle. Similarly, people with low trust in the European Union and more religious tend to support the Center-Right. Interestingly, Media Consumption is a strong predictor but it is not linked with one specific coalition, as media likely polarizes and creates strong echo chambers on both sides of the political spectrum (Barberá, 2020). Finally, older people are more likely to vote for Center-Right and Terzo Polo.

Considering the main predictors of voting behaviour in Italy, key and impartial policy recommendations are brought forward in order to strengthen democratic engagement and mitigate polarization of political preferences. Firstly, it would be essential to moderate media narrative through objective reporting on governmental communication channels (such as the RAI) and strengthening fact checking protocols on media outlets, given the importance of the variable for voters. These interventions are especially relevant for Immigration and Democracy issues and European themes as they could objectivise perception of migrations, and increase trust in National and European institutions, possibly resulting in less polarized voting behaviour. Secondly, given the important role played by age, it is crucial to involve young people through civic engagement initiatives in schools in order to increase their turnout in the elections. It is recommended to promote local intergenerational dialogue, especially at the municipal levels through elder's councils as the examples of France with Conseils des Sages (Villages et Villes Sages, nd). The final objective would be to foster dialogues between generations, bridging ideological divides linked to age. Lastly, in order to address Euroscepticism, a key area of intervention would be advocating for participatory mechanisms such as European Citizens' Initiative which allow citizens to propose European legislations. Citizens' involvement could enhance trust in European institutions at the national level (European Commission, 2025).

Overall, these evidence-based policy interventions regarding the main factors in predicting voting behaviour aim to foster a more informed Italian society and reduce unreasoned divides and biased views at the polls.

References

Barberá, Pablo (2020). "Social Media, Echo Chambers, and Political Polarization". In: *Social Media and Democracy*. Ed. by Joshua A. Tucker and Nathaniel Persily. SSRC Anxieties of Democracy. Cambridge: Cambridge University Press, pp. 34–55. ISBN: 978-1-108-83555-8. URL: https://www.cambridge.org/core/books/social-media-and-democracy/social-media-echo-chambers-and-political-polarization/333A5B4DE1B67EFF7876261118CCFE19 (visited on 05/09/2025).

Campbell, Angus et al. (Sept. 1980). *The American Voter*. en. Chicago, IL: University of Chicago Press. ISBN: 978-0-226-09254-6. URL: https://press.uchicago.edu/ucp/books/book/chicago/A/bo24047989.html (visited on 05/08/2025).

Carlsson, Fredrik and Olof Johansson-Stenman (2010). "Why Do You Vote and Vote as You Do?" en. In: *Kyklos* 63.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6435.2010.00475.x, pp. 495–516. ISSN: 1467-6435. DOI: 10.1111/j.1467-6435.2010.00475.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6435.2010.00475.x (visited on 05/08/2025).

DellaVigna, Stefano and Ethan Kaplan (Aug. 2007). "The Fox News Effect: Media Bias and Voting". In: *The Quarterly Journal of Economics* 122.3, pp. 1187–1234. ISSN: 0033-5533. DOI: 10.1162/qjec.122.3.1187. URL: https://doi.org/10.1162/qjec.122.3.1187 (visited on 05/09/2025).

Di Battista, Silvia, Monica Pivetti, and Chiara Berti (Apr. 2018). "Moral Foundations, Political Orientation and Religiosity In Italy". en. In: *The Open Psychology Journal* 11.1, pp. 46–58. ISSN: 1874-3501. DOI: 10.2174/1874350101811010046. URL: https://openpsychologyjournal.com/VOLUME/11/PAGE/46/ (visited on 05/08/2025).

Durante, Ruben, Paolo Pinotti, and Andrea Tesei (July 2019). "The Political Legacy of Entertainment TV". en. In: *American Economic Review* 109.7, pp. 2497–2530. ISSN: 0002-8282. DOI: 10.1257/aer.20150958. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20150958 (visited on 05/08/2025).

Finseraas, H., N. Jakobsson, and A. Kotsadam (June 2012). "The Gender Gap in Political Preferences: An Empirical Test of a Political Economy Explanation". en. In: *Social Politics: International Studies in Gender, State & Society* 19.2, pp. 219–242. ISSN: 1072-4745, 1468-2893. DOI: 10.1093/sp/jxs005. URL: https://academic.oup.com/sp/article-lookup/doi/10.1093/sp/jxs005 (visited on 05/08/2025).

Genova, Carlo and Martina Tazzara (2025). *Youth Political Activism in Italy: Participation in Radical-Left Groups*. en. SpringerBriefs in Sociology. Cham: Springer Nature Switzerland. ISBN: 978-3-031-88671-3 978-3-031-88672-0. DOI: 10.1007/978-3-031-88672-0. URL: https://link.springer.com/10.1007/978-3-031-88672-0 (visited on 05/08/2025).

Gentzkow, Matthew (Aug. 2006). "Television and Voter Turnout". en. In: *Quarterly Journal of Economics* 121.3, pp. 931–972. ISSN: 0033-5533, 1531-4650. DOI: 10.1162/qjec.121.3.931. URL: https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.121.3.931 (visited on 05/08/2025).

Geys, Benny, Tom-Reiel Heggedal, and Rune J. Sørensen (Aug. 2022). "Age and vote choice: Is there a conservative shift among older voters?" en. In: *Electoral Studies* 78, p. 102485. ISSN: 02613794. DOI: 10.1016/j.electstud.2022.102485. URL: https://linkinghub.elsevier.com/retrieve/pii/S0261379422000452 (visited on 05/08/2025).

Home | European Citizens' Initiative (2025). Accessed on 2025-05-09. URL: https://citizens-initiative.europa.eu/_en (visited on 05/09/2025).

Inglehart, Ronald and Pippa Norris (Oct. 2000). "The Developmental Theory of the Gender Gap: Women's and Men's Voting Behavior in Global Perspective". en_US. In: ISSN: 0192-5121. URL: http://nrs.harvard.edu/urn-3:HUL.InstRepos:37971075 (visited on 05/08/2025).

Iversen, Torben, Frances Rosenbluth, and David Soskice (July 2005). "Divorce and the Gender Division of Labor in Comparative Perspective". In: *Social Politics: International Studies in Gender, State & Society* 12.2, pp. 216–242. ISSN: 1072-4745. DOI: 10.1093/sp/jxi012. URL: https://doi.org/10.1093/sp/jxi012 (visited on 05/08/2025).

Kuenzi, Michelle and Gina M. S. Lambright (Nov. 2011). "Who votes in Africa? An examination of electoral participation in 10 African countries". EN. In: *Party Politics* 17.6. Publisher: SAGE Publications Ltd, pp. 767–799. ISSN: 1354-0688. DOI: 10.1177/1354068810376779. URL: https://doi.org/10.1177/1354068810376779 (visited on 05/08/2025).

Marx, Paul (2014). "Labour market risks and political preferences: The case of temporary employment". en. In: *European Journal of Political Research* 53.1. _eprint: https://onlinelibrary.wiley.com/doi/pd 6765.12027, pp. 136–159. ISSN: 1475-6765. DOI: 10.1111/1475-6765.12027. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6765.12027 (visited on 05/08/2025).

McCarty, Nolan, Keith T Poole, and Howard Rosenthal (2006). "Polarized America: The Dance of Ideology and Unequal Riches Chapter 2". en. In:

Müller, Jan-Werner (2003). "From National Identity to National Interest: The Rise (and Fall) of Germany's New Right". en. In: *ResearchGate*. DOI: 10.1057/9781403982544_10. URL: https://www.researchgate.net/publication/305251956_From_National_Identity_to_National_Interest_The_Rise_and_Fall_of_Germany's_New_Right (visited on 05/08/2025).

Norris, Pippa (Sept. 2002). *Democratic Phoenix: Reinventing Political Activism*. 1st ed. Cambridge University Press. ISBN: 978-0-521-81177-4 978-0-521-01053-5 978-0-511-61007-3. DOI: 10.1017/CBO9780511610073. URL: https://www.cambridge.org/core/product/identifier/9780511610073/type/book (visited on 05/08/2025).

Pendenza, Massimo and Dario Verderame (May 2019). "Young People in the Age of Crisis. Discovering Patterns of Change in Cosmopolitanism and Attitudes towards Europe/European Union among University Students in Southern Italy". en. In: Publisher: Brill. DOI: 10.1163/25895745-00101007. URL: https://brill.com/view/journals/yogo/1/1/article-p138_138.xml (visited on 05/08/2025).

Rynkiewicz, Joseph, Mohamed Raouf Benmakrelouf, and Wafa Karouche (2016). "Assessment of the Influence of Education Level on Voting Intention for the Extreme Right in France." en. In: 37.3, pp. 205–215.

Schuman, Howard et al. (1998). *Racial Attitudes in America*. en. URL: https://www.hup.harvard.edu/books/9780674745698 (visited on 05/08/2025).

Terreni, Federico (Jan. 2025). *Democracy in Italy at Stake – Listen To Europe*. en-US. URL: https://listentopeople.eu/democracy-in-italy-at-stake/ (visited on 05/08/2025).

Tilley, James and Geoffrey Evans (Mar. 2014). "Ageing and generational effects on vote choice: Combining cross-sectional and panel data to estimate APC effects". en. In: *Electoral Studies* 33, pp. 19–27. ISSN: 02613794. DOI: 10.1016/j.electstud.2013.06.007. URL: https://linkinghub.elsevier.com/retrieve/pii/S0261379413000875 (visited on 05/08/2025).

Villages et Villes Sages (2025). fr-FR. URL: https://villagesetvillessages.fr/ (visited on 05/09/2025).

Vogt, W. Paul (1997). *Tolerance & education : learning to live with diversity and difference*. eng. Thousand Oaks : Sage Publications. ISBN: 978-0-7619-0216-4 978-0-7619-0217-1. URL: http://archive.org/details/toleranceeducati0000vogt_e0k2 (visited on 05/08/2025).

The following appendices report the most relevant results. Full outputs from all models are available in the replication package at: `https://github.com/davi-mag/votingbehavior`.

Appendix A Summary Statistics

Table 1: Summary Statistics (Part 1)

Statistic	Age	Education Level	Household Income	Political Interest	Left-Right Scale
Count	837	837	837	837	837
Mean	53.98	3.78	5.72	2.66	5.25
Std Dev	17.18	1.75	2.33	0.79	2.42
Min	18	1	1	1	0
25%	41	2	4	2	3
50%	55	4	6	3	5
75%	68	4	8	3	7
Max	90	7	10	4	10

Table 2: Summary Statistics (Part 2)

Statistic	Satisfaction with Democracy	Trust in EU Parliament	Attitude Toward Immigration	Attitude Toward LGBT	Media Consumption
Count	837	837	837	837	837
Mean	5.37	5.09	4.57	1.85	153.39
Std Dev	1.97	1.99	2.28	0.87	217.66
Min	0	0	0	1	0
25%	4	4	3	1	60
50%	6	5	5	2	90
75%	7	6	6	2	140
Max	10	10	10	5	1147

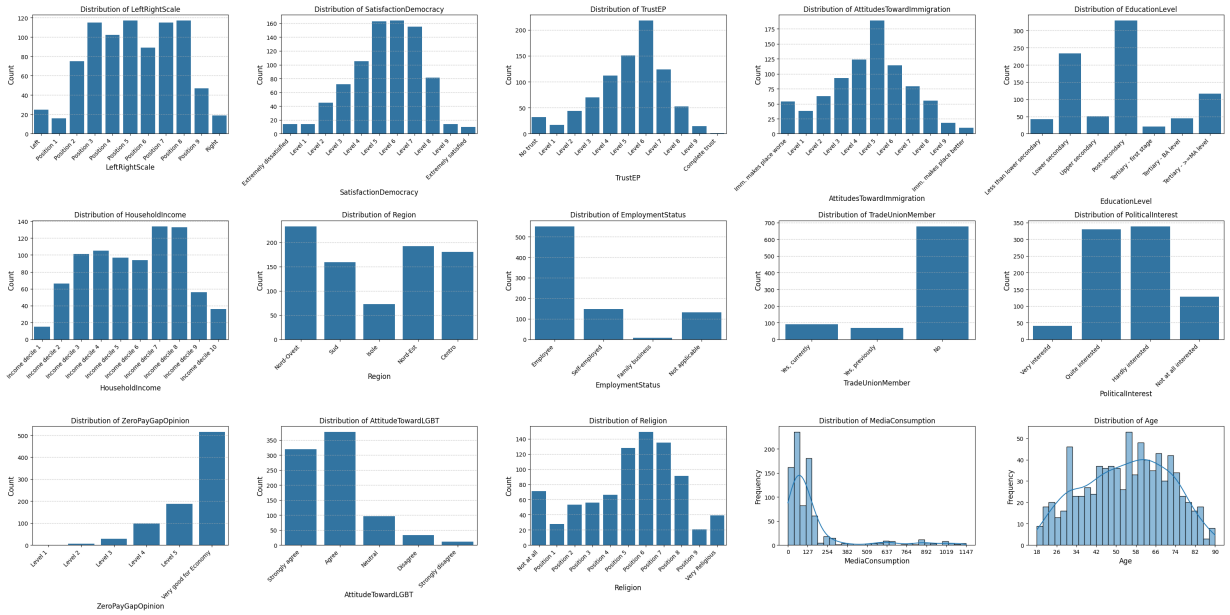
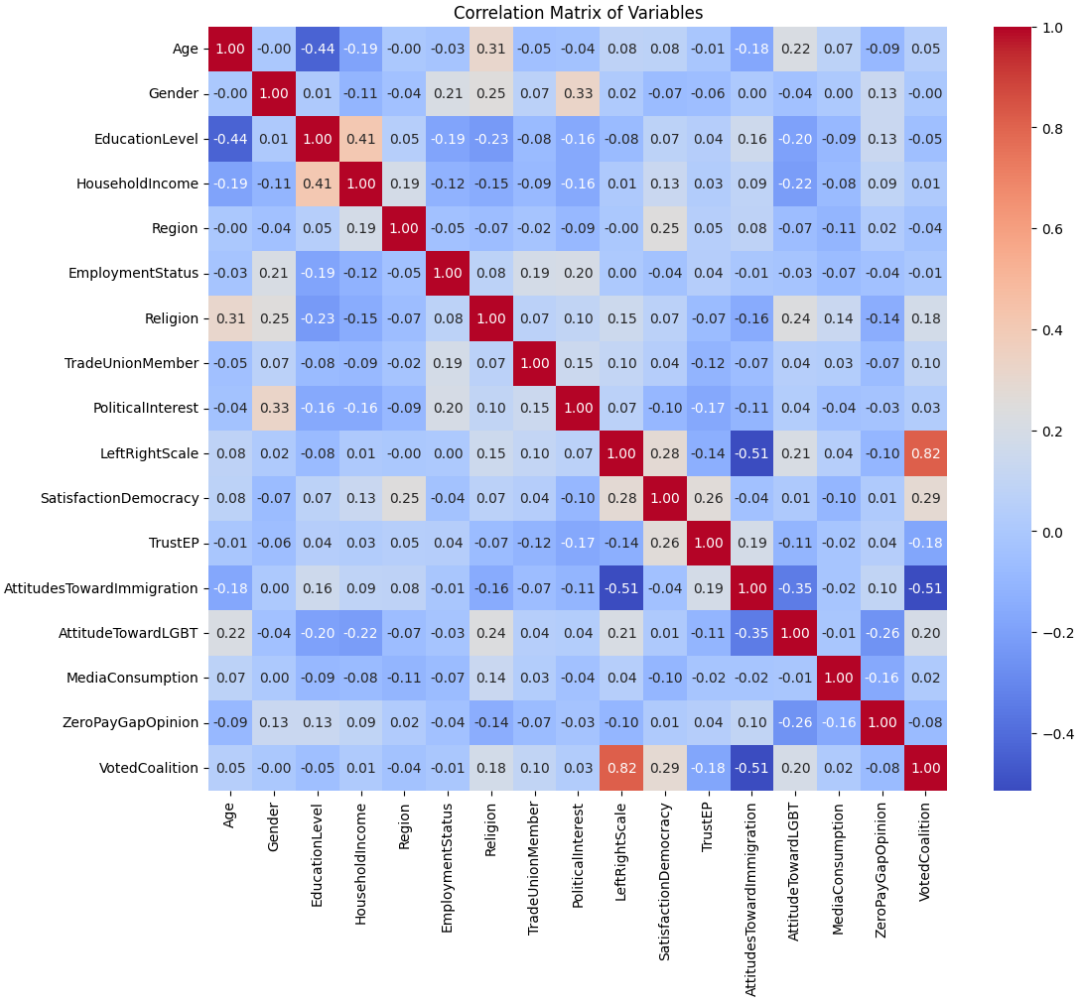
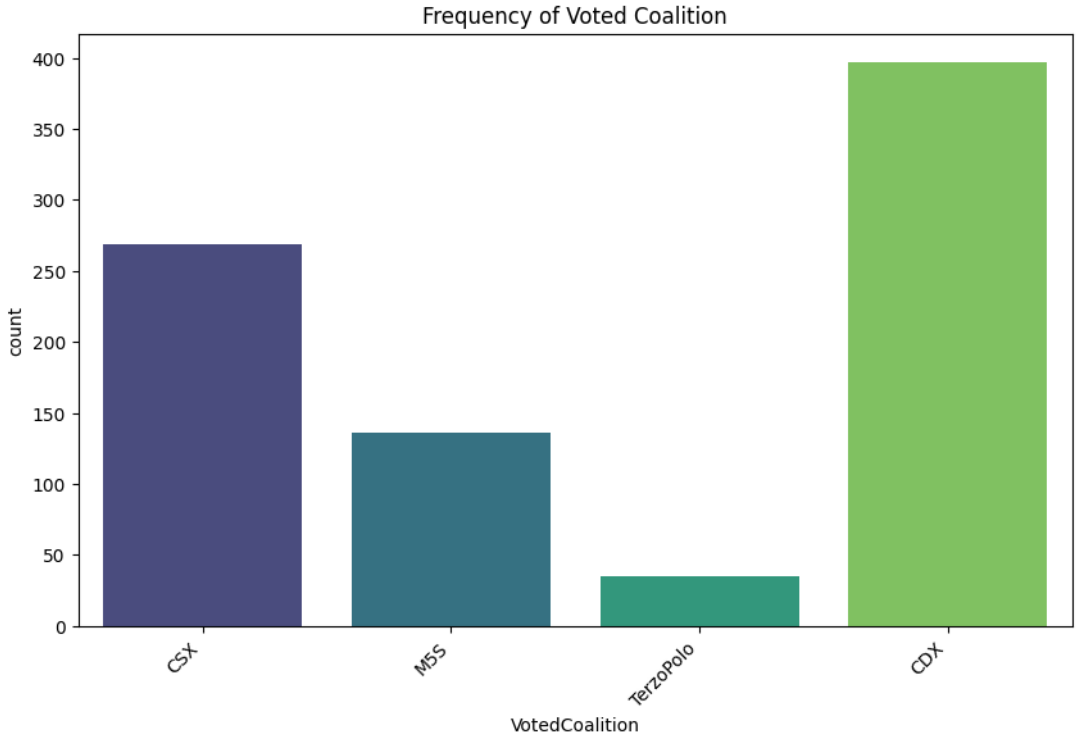


Figure 1: Distributions of variables

Appendix B Correlation Matrix of All Variables



Appendix C Frequency of Voted Coalition



Appendix D Confusion Matrix of Decision Trees

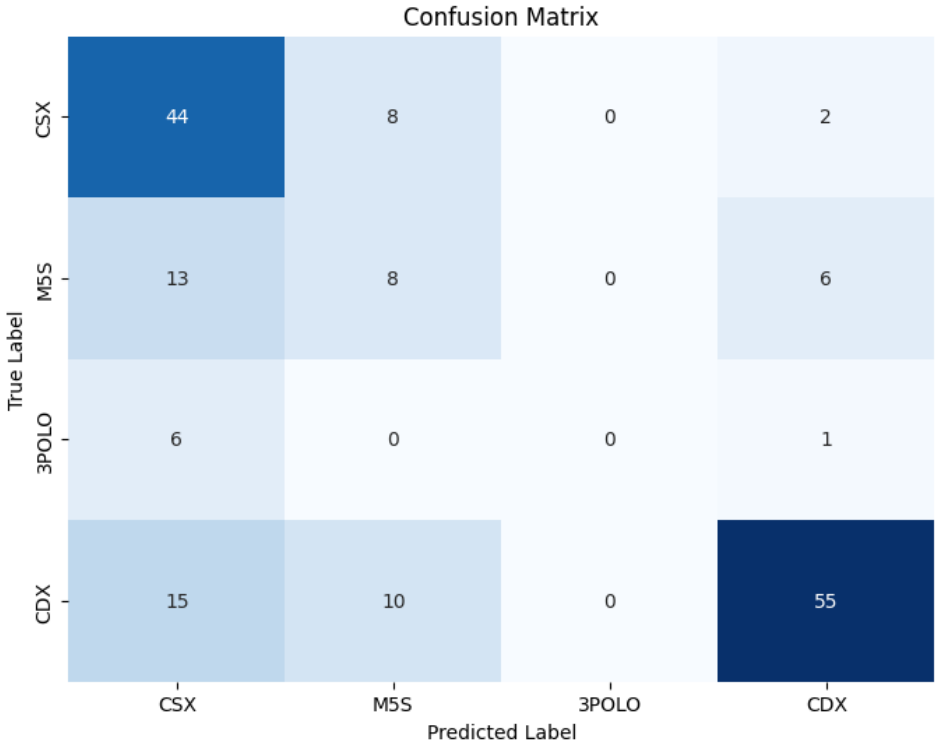


Figure 2: Confusion Matrix of Baseline Decision Tree. Accuracy: 64%

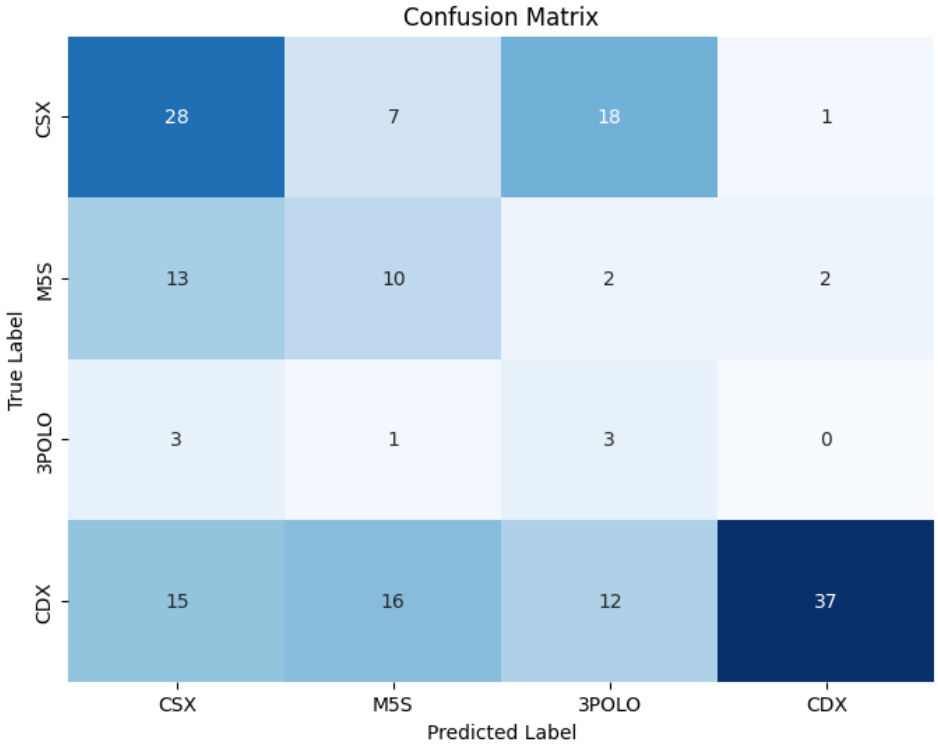


Figure 3: Confusion Matrix of Decision Tree with GridSearchCV and SMOTE on training set only. Accuracy: 46%

Appendix E Decision Tree

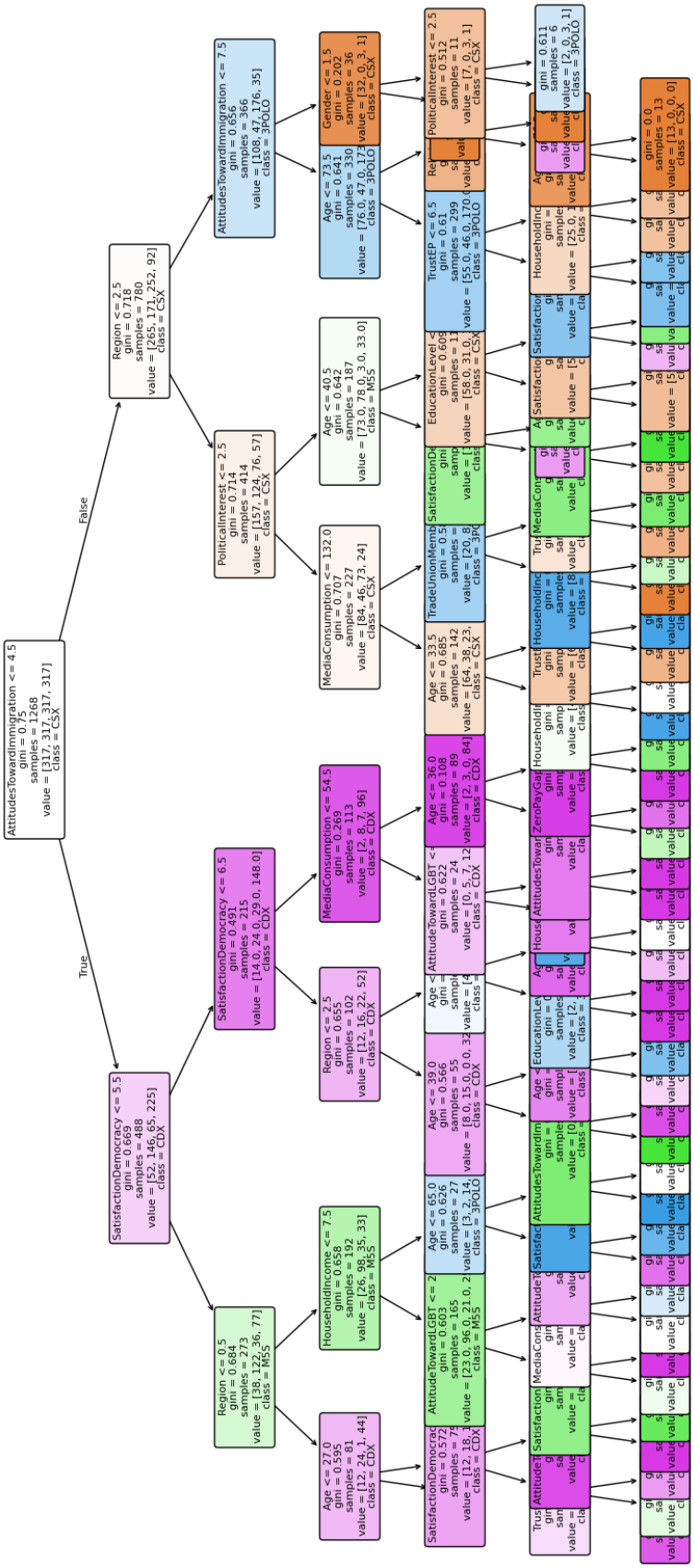


Figure 4: Decision Tree with GridSearch Cross-validation + SMOTE

Appendix F Random Forest results with GridSearchCV and SMOTE

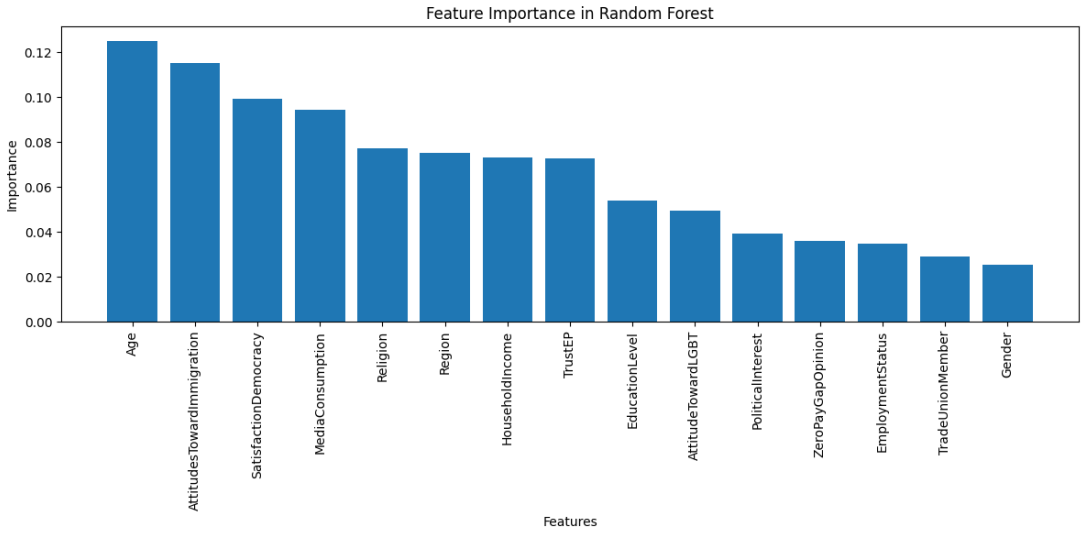


Figure 5: Feature importance. Accuracy: 64%

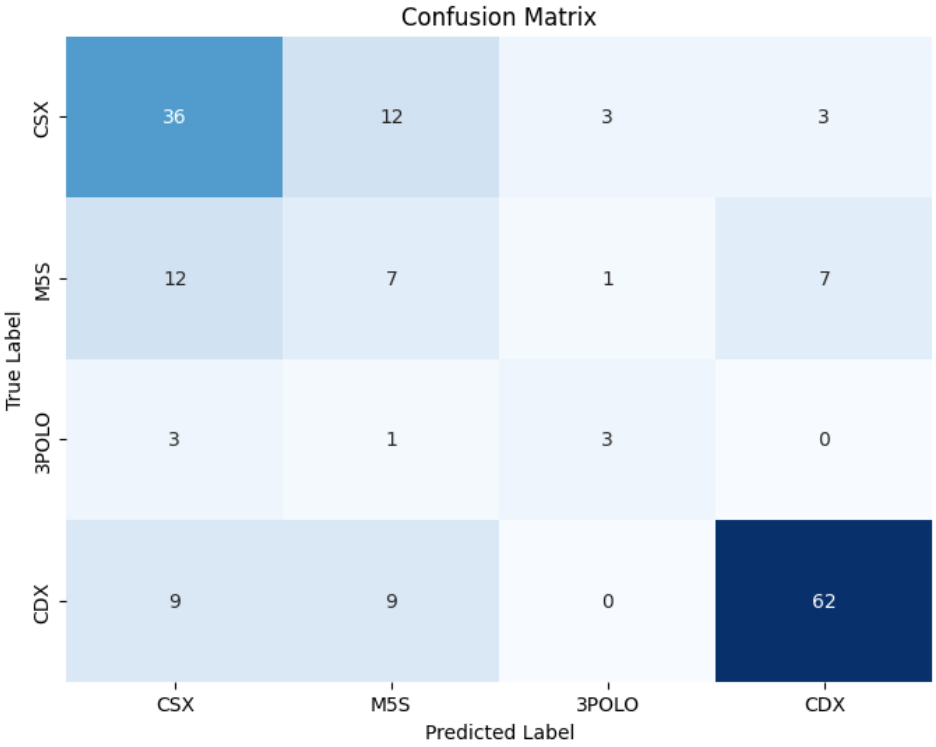


Figure 6: Confusion Matrix

Appendix G XGBoost results with GridSearchCV and SMOTE

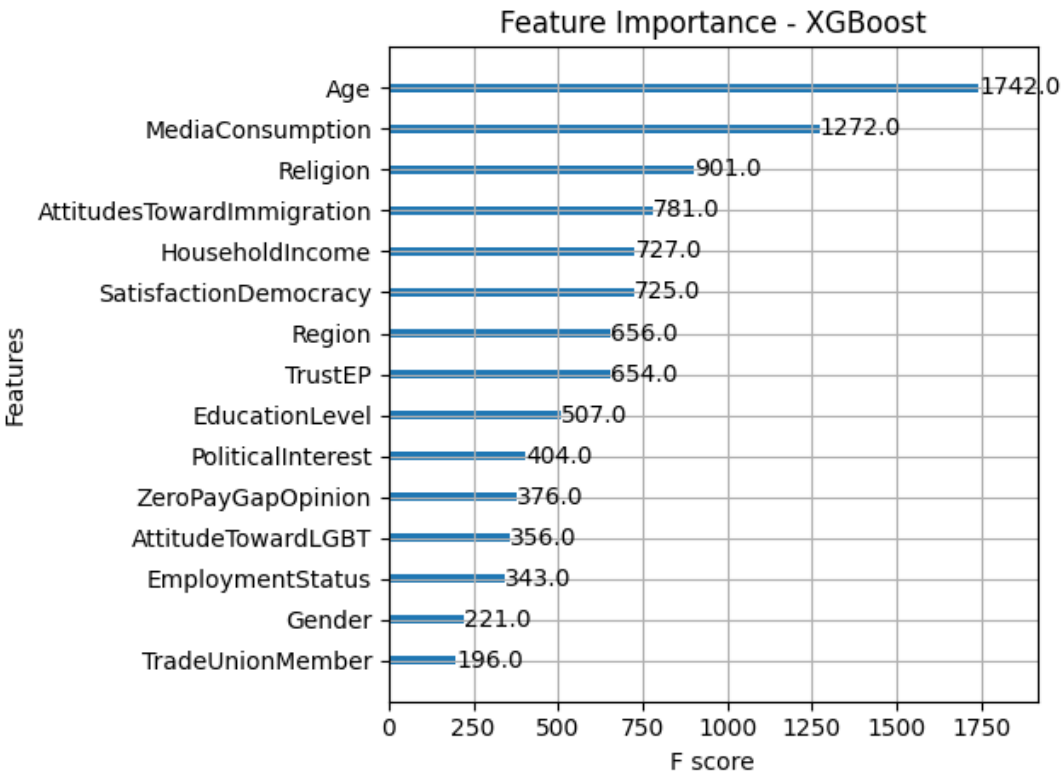


Figure 7: Feature importance. Accuracy: 65%

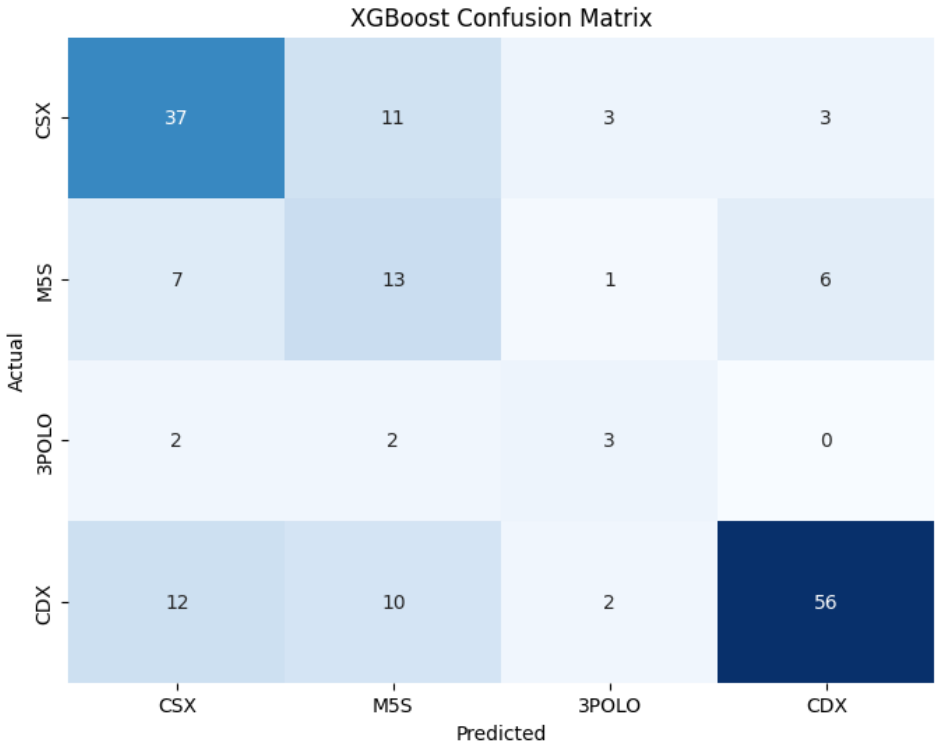


Figure 8: Confusion Matrix

Appendix H Multinomial Logistic Regression Results

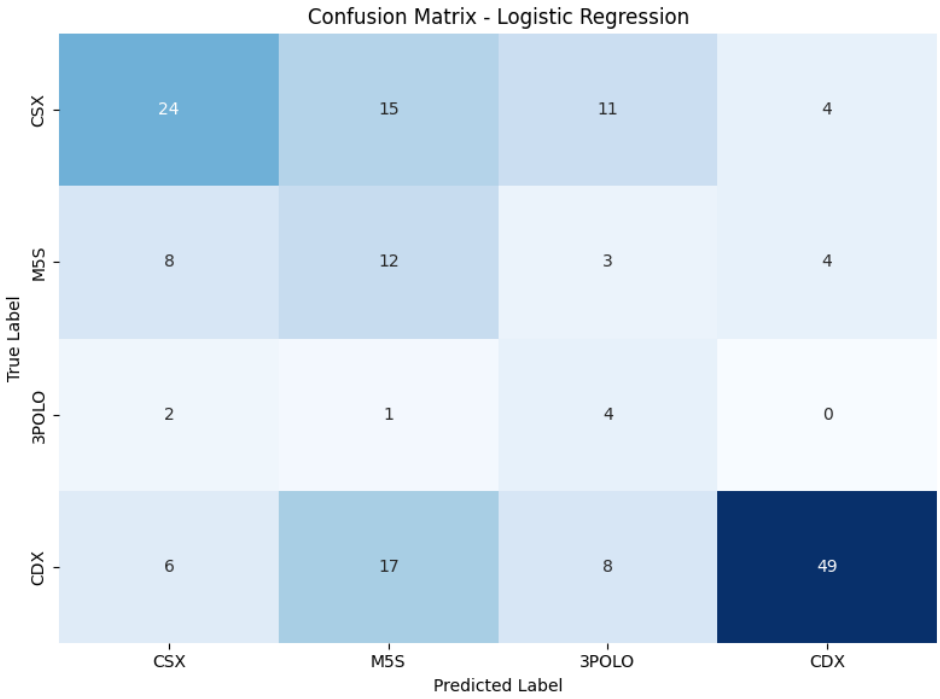


Figure 9: Confusion Matrix. Accuracy: 53%

Table 3: Summary Output Table from MLR Model (Baseline: CSX)

Baseline: CSX	**M5S**			**3POLO**			**CDX**		
Variable	**Coef**	**SE**	**p**	**Coef**	**SE**	**p**	**Coef**	**SE**	**p**
Intercept	-0.571	0.146	0.000	-2.363	0.321	0.000	0.506	0.112	0.000
Age	-0.820	0.141	0.000	-2.364	0.243	0.097	-0.465	0.125	0.000
Gender	-0.241	0.130	0.064	-0.164	0.184	0.446	0.050	0.116	0.771
Education	-0.216	0.150	0.150	-0.325	0.234	0.165	-0.117	0.126	0.354
Income	-0.293	0.129	0.023	0.239	0.237	0.313	0.058	0.119	0.628
Region	0.033	0.125	0.789	0.721	0.245	0.003	-0.255	0.102	0.011
Employment	0.106	0.114	0.353	-0.444	0.208	0.032	0.008	0.110	0.944
Religion	0.322	0.127	0.011	0.481	0.182	0.008	0.390	0.112	0.000
Union Member	-0.143	0.114	0.211	-0.013	0.182	0.942	0.068	0.108	0.531
Pol. Interest	0.247	0.129	0.055	-0.060	0.222	0.788	-0.095	0.104	0.370
Satisfaction	-0.145	0.130	0.263	0.483	0.220	0.028	1.162	0.135	0.000
Trust EP	-0.223	0.123	0.075	-0.468	0.202	0.022	-0.304	0.122	0.012
Immigration	-0.744	0.149	0.000	-0.641	0.252	0.014	-1.754	0.148	0.000
LGBT	0.068	0.138	0.618	-0.747	0.307	0.013	-0.105	0.116	0.642
Media	-0.053	0.118	0.655	-0.350	0.336	0.298	0.103	0.180	0.571
Pay Gap	-0.026	0.126	0.834	0.183	0.250	0.466	-0.079	0.109	0.466

Pseudo R^2 = 0.2819 Log-Likelihood = -689.18 LLR p-value = 2.91e-86 N = 837

Interpretation of Coefficients

To interpret the MLR coefficients, we exponentiate them to obtain odds ratios:

$$\text{Odds Ratio} = \exp(-1.754) \approx 0.173$$

$$\text{Percent change in odds} = (\text{Odds Ratio} - 1) \times 100 \approx -82.7\%$$

Focusing on Center-Right (CDX) coalition and Immigration Attitudes variable (standardized):

- A one-standard-deviation increase in pro-immigration attitudes is associated with a 1.754 decrease in the log-odds of voting for CDX relative to the Center-Left (CSX), holding other variables constant.
- This translates to an 83% reduction in the odds of supporting CDX versus CSX.
- In practical terms, respondents with more favorable views toward immigration are significantly less likely to vote for the CDX coalition.
- This effect is highly statistically significant, confirming the strong predictive power of immigration attitudes in shaping voting behavior.