Larry On Ballots

Digging into Swiss Referendum Voting Patterns

1. Introduction

Switzerland has the most prevalent system of nationwide referendums as a form of direct democracy legislation in the world. Four times each year, Swiss citizens vote on referendums to approve or reject laws passed by the national government, to amend or change the constitution, and to ratify international treaties. A referendum is adopted if it receives both a majority of votes nationally and a majority of votes in a majority of the 26 member states.

Most of the literature that uses machine learning models to predict electoral behavior focuses on elections driven by party identifications: multi-party elections such as national elections, and referendums where political parties have a previously defined preferred outcome (e.g., Brexit). The Swiss referendums present a unique opportunity to bring machine learning modeling to elections that are not necessarily heavily influenced by party identification.

The goal of the project is to use Switzerland's municipality-level census data and apply different machine learning models to identify demographic feature trends in referendum voting outcomes and make referendum results predictions. This means that we are not only interested in developing a model with high accuracy, but also in identifying the more relevant predictive features to see if machine learning models pick up certain phenomena. Most notable, we expect to observe the rural-urban divide as well as the language barrier as germanophone cantons are usually more concervative and franco- and italophone ones.

In this project, we focus on the controversial face-covering ban referendum that took place on March 7th in 2021. The ban was accepted with 51.42% of the total votes, with 83.10% of municipalities voting in favor. We test if it is possible to predict how a municipality voted based on its demographic characteristics, and also determine what features are the best predictors.

2. Datasets

We worked with two datasets from the Swiss Federal Department of Statistics[1][2]. The first one contained valuable socio-demographic information at the municipality-level. More precisely, it contains information about the population (size, age, household, etc.), land use (housing, agriculture, forest, etc.), local economy (employment by sector, unemployment, etc.), and local politics (seats by party).

[1] Canton demographic data: https://www.bfs.admin.ch/bfs/fr/home/statistiques/statistique-regions/portraits-regionaux-chiffres-cles/communes/donnees-explications.html

[2] Voting data - Referendum on Face covering: https://www.bfs.admin.ch/bfs/fr/home/statistiques/politique/votations/annee-2021/2021-03-07/interdiction-dissimuler.assetdetail.15924501.html

We merged this dataset with the voting results by municipality. While we did have access to information such as the number of ballots cast, and the exact number of votes for and against, we decided to use a simple outcome variable that indicates whether a municipality accepted or rejected the face-covering ban (i.e. more or less than 50% votes).

We have a total of 2179 observations, one per municipality. Regarding the voting outcome for the face covering referendum, the results follow the distribution shown in Figure 1, with a mean of 57.33% 'Yes' and a standard deviation of 7.67 percentage points. This reflects 83.1% of municipalities with a 'Yes' majority of votes. The summary statistics for the individual features considered can be found in the Appendix I.

We also found missing values. Most notably, the percentage of the population on social aid would have been a relevant feature, but 25% of the data had missing values for that feature. We decided to remove it. The employment by sector variables also had a small amount of missing values. We imputed data with the mean value of the 5 nearest neighbors, which we believe should be relatively close to the real value.

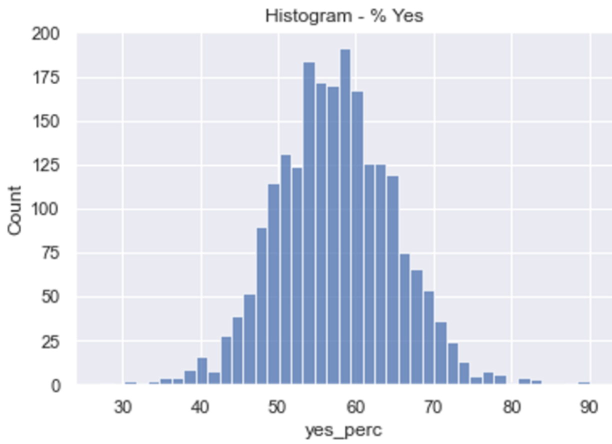


Figure 1 - Histogram of Referendum Outcome at Municipality Level

For further exploratory analysis, we created a correlation matrix between all the features. Due to the large dimensionality, we created an excel file with the results, which can be found in the file *data/correlations.xlsx.* The figure in Appendix II depicts a sample of the correlation matrix highlighting possible collinearity concerns. The figure in Appendix III depicts the results of a Variance Inflation Factor (VIF) test performed as an additional method to identify collinearity. Many of the collinearity issues were due to the values of features being correlated to population levels (e.g., registered voters) or surface levels (e.g. agriculture surface). We adjusted these variables per capita/per acre. The existence of collinearity can incorrectly increase independent variables' standard error reducing the reliability of the model and increasing uncertainty on the statistical significance of each variable.

To better understand the relationship of the features with the outcome variable, we have plotted a scatterplot for each (Appendix IV). We find no significant outliers that need to be treated.

3. Experiment Setup

Our first step before implementing the models was to transform our data. First, we merged the demographics and the election results data, matching by municipality. We then removed irrelevant attributes: those that would either raise collinearity issues, are redundant, or reflect outcomes (blank votes, etc.). Many of the collinearity concerns were associated with attributes being highly correlated with canton population (registered voters, cast ballots, etc.) or multiple attributes combining to equal a different attribute (total employment equaling the sum of primary, secondary, and tertiary employment). We also created dummy variables for the canton a municipality is in (26 cantons, 1 dummy per canton).

Our dataset then contained relevant features and the percentage of "yes" votes. We were interested in the binary classification, so we created a binary outcome variable that indicates whether a municipality voted in favor of the face-covering ban.

We split our data with an 80/10/10 split. 80% of the data in the train set, 10% in the development set, and the remaining 10% in the test set. This choice reflects the small amount of data we could work with. With so few observations, we had to balance getting as much information as possible in the training process with keeping enough data in the development and testing sets.

The last step before running our models was to normalize the features so that we could analyze their relative importance. We opted for a min-max-scaler that would scale all data points of a feature to be between 0 and 1.

4. Methodology

We proceed to experiment with 5 different machine learning models; **logistic regression, KNN, decision tree, random forest, and neural networks.**

- **Logistic regression:** a good accuracy would suggest log-linear relationships between the features and the voting outcome.
- **KNN:** a high accuracy would indicate that similar municipalities have the same voting patterns, although it assumes equal importance for all features.
- **Decision Tree:** yields good insights into which features are good predictors for voting outcomes.
- **Random Forest:** more robust than a single decision tree so we will be able to extract the most important features with more certainty than with the decision tree.
- **Neural Network:** if it turns out to be the best model, we could conclude that voting patterns are complex.

We also want to see how the models behave with and without information on the percentages of seats the political parties have in the municipalities. We want to see if the models primarily rely on information about the political orientation of a municipality or if we can build strong models only on demographic data.

We used sklearn to train the models and tuned a variety of hyperparameters to find the best settings for each. For every model type, we trained a model on the training set and adjusted the hyperparameters based on the accuracy of the development set, keeping the test set completely out of the training process. Key training points by model:

- **Logistic Regression:** We adjust the following hyper parameters.
    - Penalty: Penalty imposed on the number of variables. We choose L2 (penalty equal to the square magnitude of coefficients), to test for multiple solvers.

    - Solvers: We test NewtonGB, LBFGS and LibLinear. The results were relatively similar, with LibLinear performing the best, which tends to work well with high-dimensional datasets.
    - C: The inverse of the regularization strength, we find the optimal value to be 10. Smaller values led to lower accuracy whereas higher values made no difference.

  We also implemented 10-fold cross validation for optimal results.

- **K-Nearest Neighbors (KNN):** We have only one hyperparameter, k (no. of neighbors), for this algorithm.
    - K: We found that the model performs best for a k-value of 15-17 with a maximum accuracy of 86.17% on training data.

  Further increasing the k-value decreases the performance, but it stabilizes at around 84.3% for k-values greater than 25.

- **Decision Tree:** The criterion and the maximum depth were the hyperparameters we had to tune.
    - Criterion: The function that measures the quality of a split. We compare Gini and Entropy accuracy scores for all decision tree depths and found that they performed similarly over the range of depths with Gini averaging slightly higher accuracy.
    - Max depth: We compare decision tree max depths (1-59) and find that a depth of 3 produces the highest accuracy when political parties are included in the attributes and a max depth of 8 when political party attributes are removed.

- **Random Forest:** There are 3 hyperparameters that we need to adjust for random forests.
    - Criterion: Similarly to the decision tree results, we find that Gini performs a bit better than entropy before we move above 100 trees.
    - Number of Trees: The number of trees in the random forest. We find that 40 trees is ideal.
    - Max depth: The maximum depth of the trees in the forest. The parameter actually has very little impact on the accuracy.

  We found that the random forest with the highest development accuracy uses Gini as its criterion and has 40 trees with a maximum depth of 40. See Appendix V for detailed results.

- **Neural Networks:** We test 2 hyperparameters for neural networks for different activation functions.
    - Activation function: Sigmoid, ReLu and Tahn.
    - Number of layers: 2 to 10
    - Number of nodes: 2 to 10

  We do not find combinations with significantly better development accuracies. Appendix VI shows the development accuracies for the different combinations of hyperparameters and activation functions.

5. Results

Figure 2 and 3 show the main performance metrics for the different models that we tested, conditional on excluding or including party political party data. Overall we see an improvement from the base rate (83.10%). We find logistic regression is our best model in terms of predicting power mainly due to its specificity: our dataset is unbalanced and logistic regression predicts 50% of the minority label ('No'). We see that more complex models (random forests and neural networks) show lower performance metrics. This might be due to the relatively small dataset and the sensitivity of the models to different random splits.

	Model	Accuracy	Specificity	recall
0	Logistic Regression	0.921659	0.500000	0.979058
1	KNN	0.898618	0.230769	0.989529
2	Decision Trees	0.894009	0.230769	0.984293
3	Random_Forest	0.884793	0.307692	0.963351
4	Neural Networks	0.898618	0.153846	1.000000

Figure 2. Model results
Including Political Party Data

We also run our models excluding political party data to analyze how using only demographics data would impact the predicting power of our models. We see lower performance metrics across all models but Random Forest. These lower metrics mean that the electoral results are influenced by the electoral behavior in the municipality elections. In other words, party lines seem to play a big role in referendum outcomes.

	Model	Accuracy	Specificity	recall
0	Logistic Regression	0.884793	0.192308	0.979058
1	KNN	0.880184	0.115385	0.984293
2	Decision Trees	0.870968	0.230769	0.958115
3	Random_Forest	0.889401	0.307692	0.968586
4	Neural Networks	0.880184	0.000000	1.000000

Figure 3. Model results
Excluding Political Party Data

In the following section we analyze the performance by model:

- **Logistic Regression:**

Our selected logistic regression model achieves the highest test accuracy of all models tested with 92.01% (compared to the base rate of 83.10%). As expected, given that the majority of municipalities voted 'yes', we have a much higher recall (97.91%) compared to the specificity (50.0%). However, the logistic regression tops all other models in the prediction of ‘no’s, indicating simple log-linear relationships between the features and outcome.

Regarding feature significance, we have plotted the weights obtained for each attribute in Appendix VII. Specific political parties are shown as some of the most relevant features to predict the election results including the far-right party Swiss People’s Party, with the highest positive weight, and the Liberal Green Party and the Green Party at the other side of the spectrum. We also see some cantons with a significant weight, including Neuchatel and Appenzell Ausserrhoden, which is expected as municipalities in the same geographic location are likely to have more similar voting patterns. The significance of the different employment sectors (primary, secondary, tertiary) is also shown as a relevant factor, which also aligns with the expected voting divide between rural and urban areas.

When looking at the results with no political data, our accuracy decreases to 88.48%, mainly driven by a lower specificity of only 19.23%. In terms of feature relevance (Appendix VIII), we observe a few changes. Population becomes the most significant variable, followed by the sector variables. These features are likely to be capturing the urban and rural divide that was being partly captured by the political party variables.

- **K-Nearest Neighbors (KNN):**

For KNN, we get a test accuracy of 89.86% (compared to the base rate of 83.10%), which suggests that similar municipalities have similar voting patterns. Considering that there are mostly yes-municipalities in the dataset, it was unsurprising to get a much higher specificity (99%) compared to the recall (23%).

When removing the political parties, the accuracy decreases (by 89.86% - 88.01% = 1.85 percentage points). Also, the highest accuracy occurs at k = 13, compared to k = 15 -17 in the model with the political parties. This may be due to the decrease in the total number of attributes that need to be checked when comparing similarity between neighbors. See Appendix IX and X for detailed results.

We have little insights on the feature importance as the KNN algorithms rank all features equally. However, the lower accuracy and recall compared to the logistic regression would suggest some features are more important than others.

- **Decision Tree:**

The Decision Tree model produced the second lowest accuracy when political parties were included (89.4%) and the lowest accuracy score of the models when political parties were excluded (87.1%). Despite the lower accuracy scores, the optimum depths and splitting attributes of the decision trees provide interesting insights into Swiss voters characteristics and policy preferences.

The best decision tree for the dataset including political parties has a depth of three and the top two depths of splitting attributes are the Swiss People’s Party (UDC: conservative), The Liberals (PLR: center right), and Green Liberal Party (PLV: centrist). Since the local seats are elected, this result indicates that the inhabitants of a municipality do not change their political opinions frequently and/or drastically. Hence, the decision trees can use the local political environment as an indicator of political opinion. Note that the political parties do publicize their endorsement or opposition to a referendum, which can sway the public opinion.

The best decision tree for the dataset not including political parties has a depth of eight and the top two depths of splitting attributes are population, private households, and the Graubunden canton (dummy variable). Graubunden is the only canton split attribute included in either of the optimal decision trees. One possible reason Graubunden was found to be an effective splitting attribute is because it is one of the few cantons that voted no for the face covering referendum. See Appendix XI and XII for detailed results.

- **Random Forest:**

For random forests, we get the lowest test accuracy of all models with 88.48% (compared to the base rate of 83.10%). As expected, the model performs better with yes-municipalities, than with no-municipalities (see confusion matrix in Appendix XIII). However, it is the model with the lowest recall, meaning that it misclassified yes-municipalities as no-municipalities the most out of all models. It performs relatively well when classifying no-municipalities correctly compared to other models, with a specificity of 30.77%.

From the random forest, we can also see which features matter most. This is done by computing the mean decrease in impurity for each feature across the individual trees. Similarly to the models discussed above, in Appendix XIV, we can find that the three most important features are the percentage of seats occupied in the municipality by the Liberal Green Party, the Swiss People's Party, and the Green Party. This indicates that the people mostly vote according to the local political landscape, which makes sense as those seats are directly elected by them.

Interestingly, when removing the political parties, the accuracy actually increases very slightly (by 88.94% - 88.48% = 0.46 percentage points). In Appendix XV we find that the population (absolute and density), employment sectors (secondary and tertiary) and the number of private households are the five most important features in predicting voting patterns. Some other significant features concern the land use, empty housing, and the age of the population.

Interestingly, the states in which the municipalities are located appear to be the least important features. This may be surprising as different voting patterns can often be observed from canton to canton (the francophone ones tend to be more liberal than the germanophone ones). For further analysis, we would suggest calculating the joint significance.

- **Neural Networks:**

We find that neural networks show accuracy rates are slightly better than the base rate (89.86%). However, they tend to predict almost all of the towns as towns that had a 'yes' majority. We hypothesize that this is due to having a relatively small dataset leading to overfitting. The lower performance of neural networks compared to simpler models such as logistic regression suggests that non-linear relationships are not significant for identifying referendum voting patterns. For the purpose of this project, neural networks are less relevant than other models, since they do not allow for an intuitive interpretation of what factors are the most important predictors.

6. Conclusion and Key Lessons

Applying machine learning techniques to a classification problem with a large imbalance in the dataset can be quite challenging. We ran into this as our base rate with 83.1% of yes-municipalities. Our models often predicted the yes-municipalities (majority) correctly, but did much worse on the no-municipalities (minority). However, we were still able to achieve our goal of analyzing which features were the most significant predictors.

Firstly, we found that people are voting along party lines. We found that local politics were the strongest predictor for logistic regression, decision tree, and random forest models. This shows that the referendum is more influenced by party alignment than we originally anticipated.

Additionally, we found that the population and employment sectors were also strong predictors. This is consistent with the urban-rural divide that is often observed across Swiss referendums. Interestingly, the models that exclude political data appear to be using these variables to identify that small, rural towns tend to be more represented by the right wing while large cities by the left wing.

Something that is often observed in Swiss politics is a divide between the different language-regions in terms of political opinions. For instance, francophone cantons tend to be more liberal than germanophone ones. However, the cantons were the weakest features in the random forest, indicating that the language divide might not be as important in this referendum. Exploring joint significance may be interesting to analyze this further.

Taking a step back from the features, we found that the two most complex models, random forests and neural networks, did not perform better than simpler models such as logistic regression. This may be due to overfitting due to the limited amount of observations in the dataset.

Following, we discuss future steps for further analysis. Firstly, we would suggest re-training the models allocating more weight to the no-municipalities, which could help improve the specificity of the models and increase overall accuracy. Secondly, it would also be interesting to include other features related to campaigns such as features from social media and news articles, to further explore voting behavior. Furthermore, this project's methodology could also be expanded to similar referendums in Switzerland, similar referendums in other countries, and other types of votes. We found that this was an interesting exercise for the discovery of the mechanisms underlying voting patterns. In terms of predictive power, it would be interesting to analyze this model's performance to predict future referendums on similar votes. These findings could then be used to inform political science, campaigns, and activism.

Appendices:

I. Summary statistics - Demographics and Referendum Results

	count	mean	std	min	25%	50%	75%	max
population	2172.0	3962.261971	12875.284703	32.0	720.2500	1555.50	3834.250	420217.0
population_variation	2172.0	9.209899	11.342446	-30.3	2.4000	7.95	14.400	92.8
population_density	2172.0	437.707643	792.787000	1.0	80.7500	185.00	467.000	12811.0
foreigner_percentage	2172.0	16.947744	9.702113	0.0	9.6000	15.20	23.125	57.8
age_percentage_less_20	2172.0	20.471455	3.367026	2.1	18.7000	20.60	22.500	37.2
age_percentage_between_20_64	2172.0	60.154880	3.197797	39.5	58.3000	60.25	62.200	81.1
age_percentage_more_64	2172.0	19.374540	4.413891	6.5	16.4000	19.00	21.600	40.3
marriage_rate	2172.0	4.168692	2.590607	0.0	2.8000	4.00	5.200	36.7
divorce_rate	2172.0	1.918692	1.762745	0.0	1.0000	1.80	2.500	38.5
birth_rate	2172.0	9.224401	3.879388	0.0	7.0000	9.30	11.200	47.0
death_rate	2172.0	7.624448	3.952898	0.0	5.5000	7.30	9.400	57.7
private_households	2172.0	1754.744936	6161.564561	14.0	306.7500	673.50	1636.750	204411.0
avg_household_size	2172.0	2.323849	0.193711	1.5	2.2000	2.30	2.400	3.3
total_surface	2172.0	18.415746	33.450123	0.3	4.4000	8.30	16.800	438.6
housing_and_infrastructure_surface	2172.0	14.925046	14.776796	0.1	5.8000	10.00	18.900	97.3
housing_and_infrastructure_surface_variation	2172.0	26.891805	30.481848	-37.0	8.7500	18.00	35.000	395.0
agriculture_surface_perc	2172.0	45.981584	19.259597	0.0	33.0000	47.65	60.800	91.5
agriculture_variation_surface_perc	2172.0	-39.160221	61.395848	-616.0	-44.0000	-20.00	-9.000	95.0
forest_surface_perc	2172.0	32.484484	16.096235	0.0	20.4000	31.00	42.800	88.2
unproductive_surface_perc	2172.0	6.609300	14.006514	0.0	0.3000	1.10	4.700	95.0
employment_total	2076.0	2498.994701	14026.548188	11.0	215.0000	598.00	1620.000	491193.0
employment_primary	2136.0	75.572566	82.941617	0.0	24.0000	47.00	96.000	751.0
employment_secondary	2115.0	516.084161	1417.563517	0.0	40.0000	149.00	476.000	34946.0
employment_tertiary	2166.0	1845.255771	12618.036130	4.0	97.0000	313.50	986.250	462410.0
establishments_total	1875.0	346.909867	1411.799868	9.0	69.0000	133.00	287.500	45057.0
establishments_primary	2022.0	26.276954	28.328050	0.0	9.0000	16.00	32.000	272.0
establishments_secondary	2009.0	47.456944	102.259616	0.0	11.0000	24.00	52.000	2637.0
establishments_tertiary	2151.0	250.039981	1227.898035	4.0	33.0000	74.00	186.000	42368.0
empty_housing_units	2172.0	1.963315	1.694620	0.0	0.7875	1.55	2.680	13.1
new_housing_units_per_capita	2172.0	6.218370	8.603579	0.0	0.7000	3.40	8.300	96.0
social_aid_perc	1692.0	2.223286	1.617768	0.2	1.1000	1.75	2.800	11.2
PLR	2172.0	14.923343	9.179349	0.0	8.7000	13.80	20.000	69.7
PDC	2172.0	12.609162	13.882447	0.0	2.3000	7.90	18.300	79.5
PS	2172.0	13.621731	6.374857	0.0	9.2000	13.30	17.700	49.3
UDC	2172.0	31.226750	13.469845	0.0	21.4000	30.40	39.500	84.1
PEV_PCS	2172.0	2.336096	2.638545	0.0	0.5000	1.60	3.400	30.4
PVL	2172.0	6.255479	4.025103	0.0	3.4000	6.10	8.900	23.8
PBD	2172.0	2.853085	4.862200	0.0	0.0000	1.00	3.200	64.6
PST_Sol	2172.0	0.767495	1.986292	0.0	0.0000	0.00	0.500	32.0
PES	2172.0	10.988168	6.113617	0.0	6.9000	10.00	14.600	38.4
small_right_parties	2172.0	2.209438	4.222997	0.0	0.0000	0.60	2.225	27.2

	count	mean	std	min	25%	50%	75%	max
canton_id	2179.0	13.704452	8.458174	1.0	3.0	17.0	22.0	26.0
registered_voters	2179.0	2522.651675	7173.123155	30.0	530.0	1099.0	2563.5	234028.0
cast_ballots	2179.0	1297.025241	3874.621587	17.0	277.5	576.0	1310.0	128959.0
participation_rate	2179.0	53.257458	9.361085	19.5	46.8	52.1	58.9	94.2
blank_votes	2179.0	13.356586	48.254538	0.0	2.0	5.0	12.0	1203.0
invalid_votes	2179.0	4.139514	13.116165	0.0	0.0	0.0	2.0	202.0
valid_ballots	2179.0	1279.529142	3828.651792	16.0	275.5	568.0	1298.5	127806.0
yes	2179.0	655.045434	1393.930347	9.0	163.0	331.0	720.5	39255.0
no	2179.0	624.483708	2488.276033	4.0	110.0	241.0	560.0	88551.0
yes_perc	2179.0	57.332171	7.672309	27.1	52.2	57.2	62.3	90.0

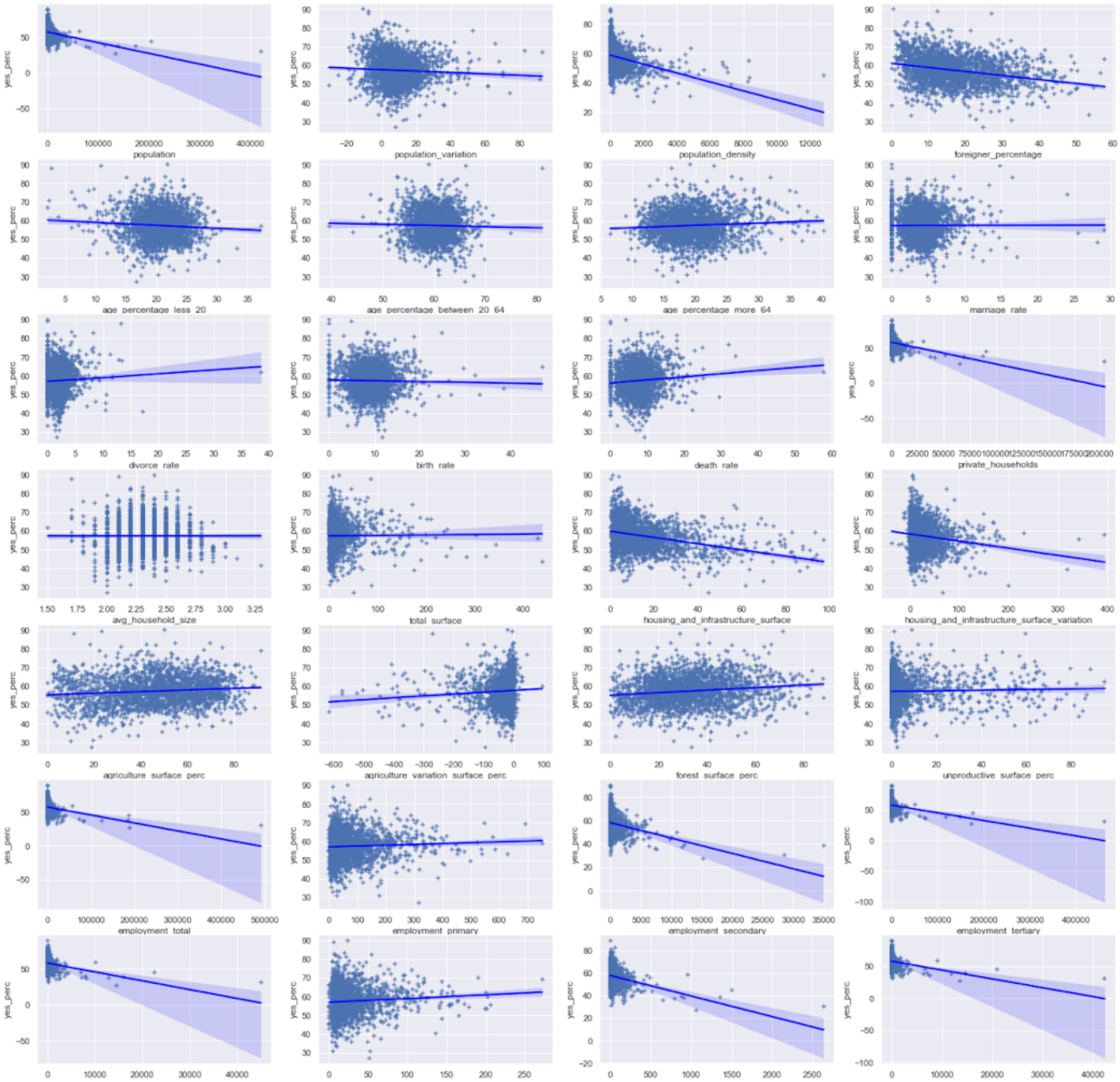
II. Abbreviated Correlation Matrix to identify possible collinearity.

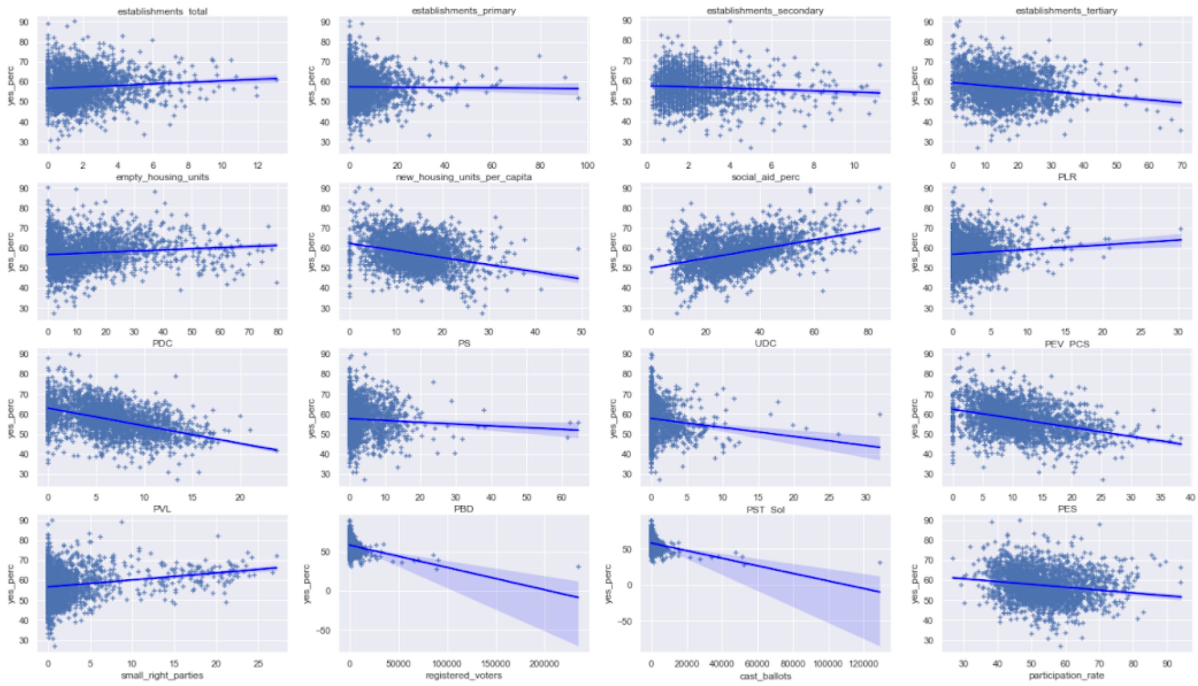
	population	population_density	private_households	total_surface	housing_and_infrastructure_surface	agriculture_variation_surface_perc	employment_total	employment_primary	employment_secondary	employment_tertiary	establishments_total	establishments_primary	establishments_secondary	establishments_tertiary
population	1.000	0.491	0.999	0.086	0.337	-0.194	0.974	0.146	0.837	0.966	0.984	0.113	0.957	0.979
population_density	0.491	1.000	0.475	-0.155	0.861	0.028	0.428	-0.125	0.441	0.398	0.515	-0.145	0.489	0.436
private_households	0.999	0.475	1.000	0.087	0.324	-0.189	0.979	0.136	0.838	0.972	0.985	0.105	0.952	0.982
total_surface	0.086	-0.155	0.087	1.000	-0.269	-0.807	0.075	0.386	0.090	0.069	0.094	0.429	0.138	0.082
housing_and_infrastructure_surface	0.337	0.861	0.324	-0.269	1.000	0.082	0.279	-0.185	0.350	0.255	0.324	-0.221	0.344	0.277
agriculture_variation_surface_perc	-0.194	0.028	-0.189	-0.807	0.082	1.000	-0.153	-0.349	-0.218	-0.138	-0.184	-0.344	-0.280	-0.167
employment_total	0.974	0.428	0.979	0.075	0.279	-0.153	1.000	0.098	0.815	0.998	0.986	0.066	0.915	0.987
employment_primary	0.146	-0.125	0.136	0.386	-0.185	-0.349	0.098	1.000	0.177	0.083	0.127	0.937	0.228	0.105
employment_secondary	0.837	0.441	0.838	0.090	0.350	-0.218	0.815	0.177	1.000	0.775	0.803	0.147	0.891	0.788
employment_tertiary	0.966	0.398	0.972	0.069	0.255	-0.138	0.998	0.083	0.775	1.000	0.984	0.051	0.896	0.986
establishments_total	0.984	0.515	0.985	0.094	0.324	-0.184	0.986	0.127	0.803	0.984	1.000	0.097	0.943	0.999
establishments_primary	0.113	-0.145	0.105	0.429	-0.221	-0.344	0.066	0.937	0.147	0.051	0.097	1.000	0.192	0.074
establishments_secondary	0.957	0.489	0.952	0.138	0.344	-0.280	0.915	0.228	0.891	0.896	0.943	0.192	1.000	0.930
establishments_tertiary	0.979	0.436	0.982	0.082	0.277	-0.167	0.987	0.105	0.788	0.986	0.999	0.074	0.930	1.000

III. Variance Inflation Factor values for each of the demographic variables.

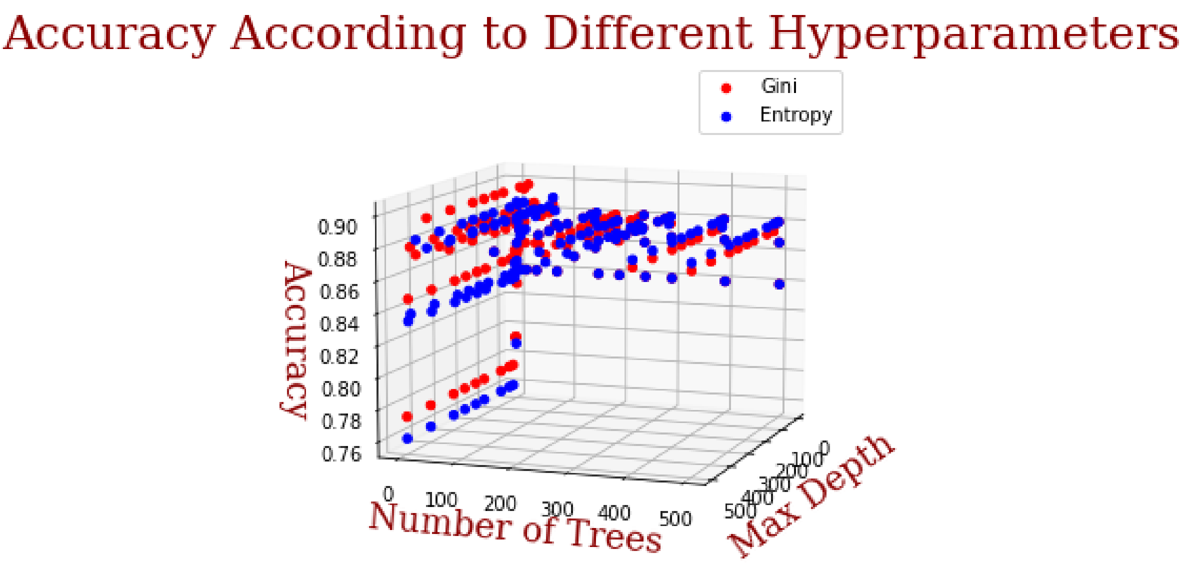
	VIF	features
0	844.381903	population
1	2.903329	population_variation
2	9.833690	population_density
3	10.136636	foreigner_percentage
4	79916.603926	age_percentage_less_20
5	672634.575217	age_percentage_between_20_64
6	73142.485990	age_percentage_more_64
7	4.094369	marriage_rate
8	2.348625	divorce_rate
9	8.411329	birth_rate
10	6.296188	death_rate
11	863.769642	private_households
12	531.921483	avg_household_size
13	7.123426	total_surface
14	81854.349151	housing_and_infrastructure_surface
15	8.026231	housing_and_infrastructure_surface_variation
16	461021.346296	agriculture_surface_perc
17	7.590454	agriculture_variation_surface_perc
18	243763.683947	forest_surface_perc
19	44486.314936	unproductive_surface_perc
20	1200.236090	employment_total
21	13.874228	employment_primary
22	21.635748	employment_secondary
23	1055.144811	employment_tertiary
24	226.719107	establishments_total
25	15.040481	establishments_primary
26	43.236438	establishments_secondary
27	252.781598	establishments_tertiary
28	2.778755	empty_housing_units
29	1.798396	new_housing_units_per_capita
30	5.670716	social_aid_perc
31	15.522057	PLR
32	14.676447	PDC
33	16.336256	PS
34	45.937831	UDC
35	2.815561	PEV_PCS
36	7.243754	PVL
37	3.235819	PBD
38	2.064433	PST_Sol
39	13.155562	PES
40	2.688230	small_right_parties

IV. Scatterplots - Features and output

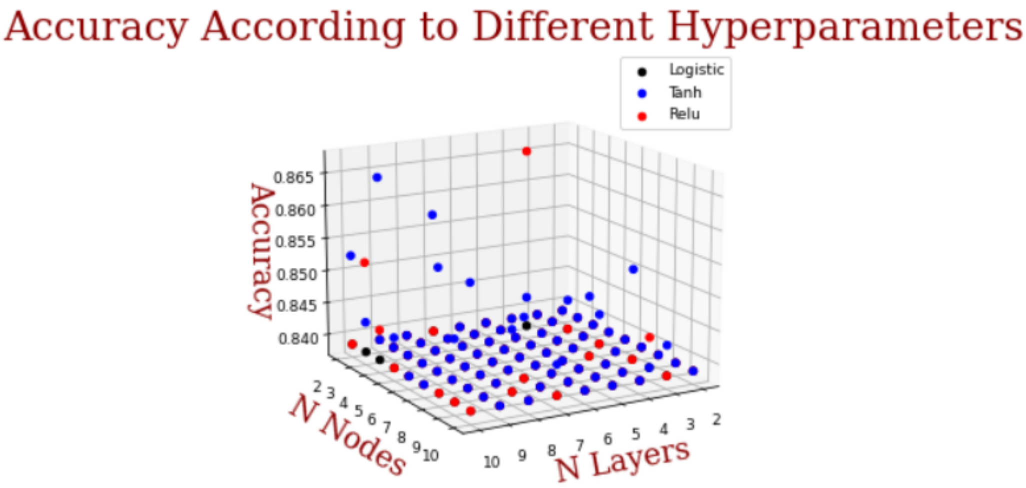




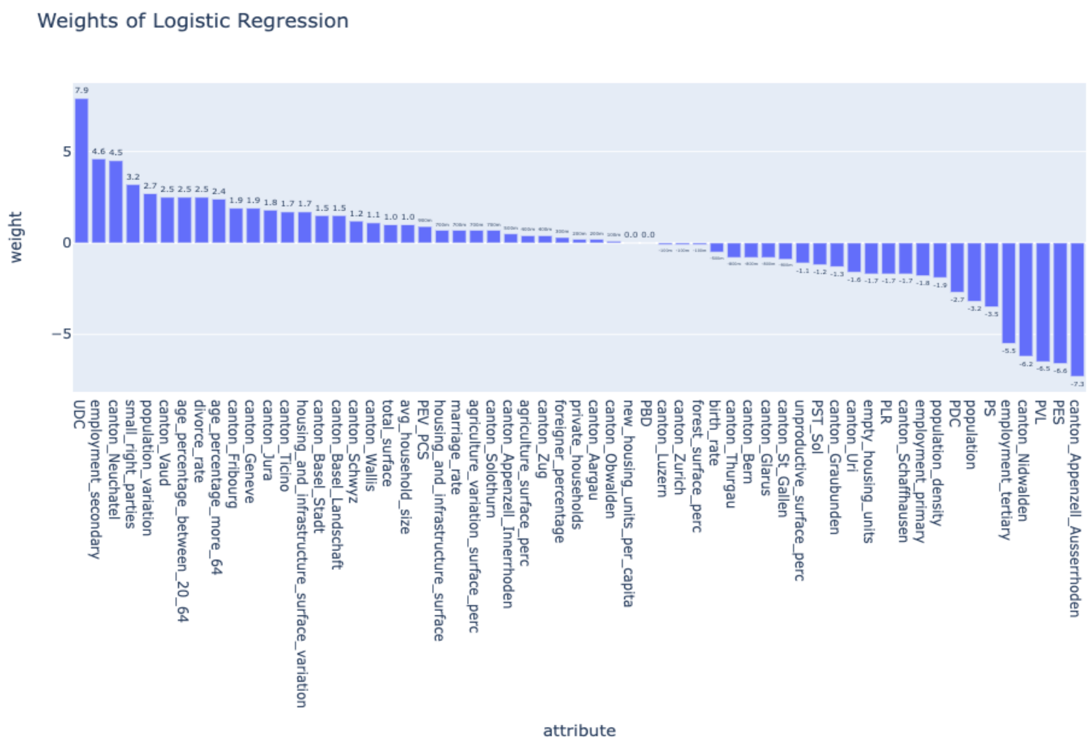
V. Random Forest Training Accuracies



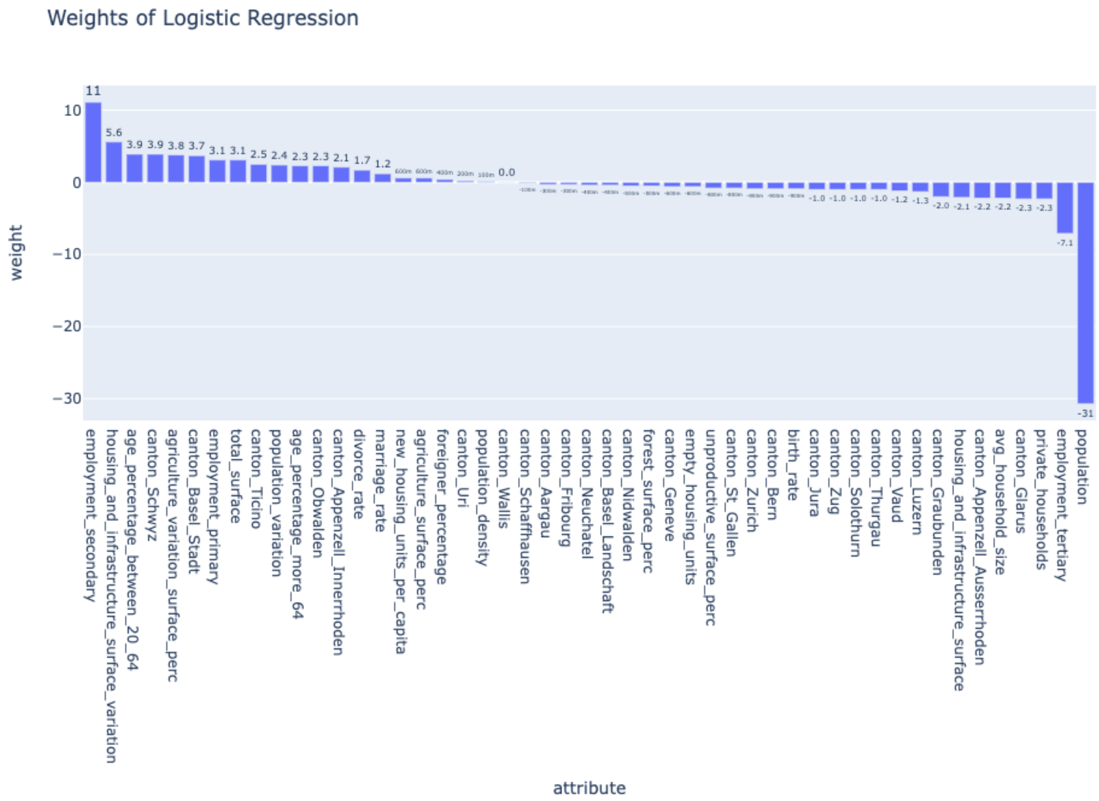
VI. Neural Networks Training Accuracies



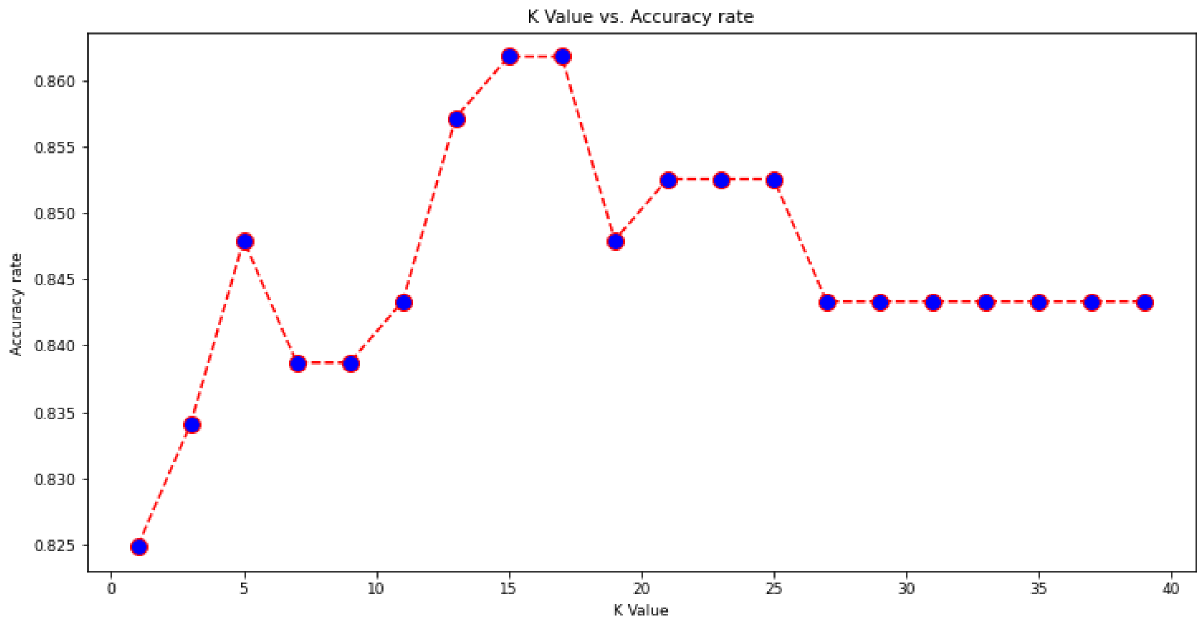
VII. Attributes and weights for logistic regression model



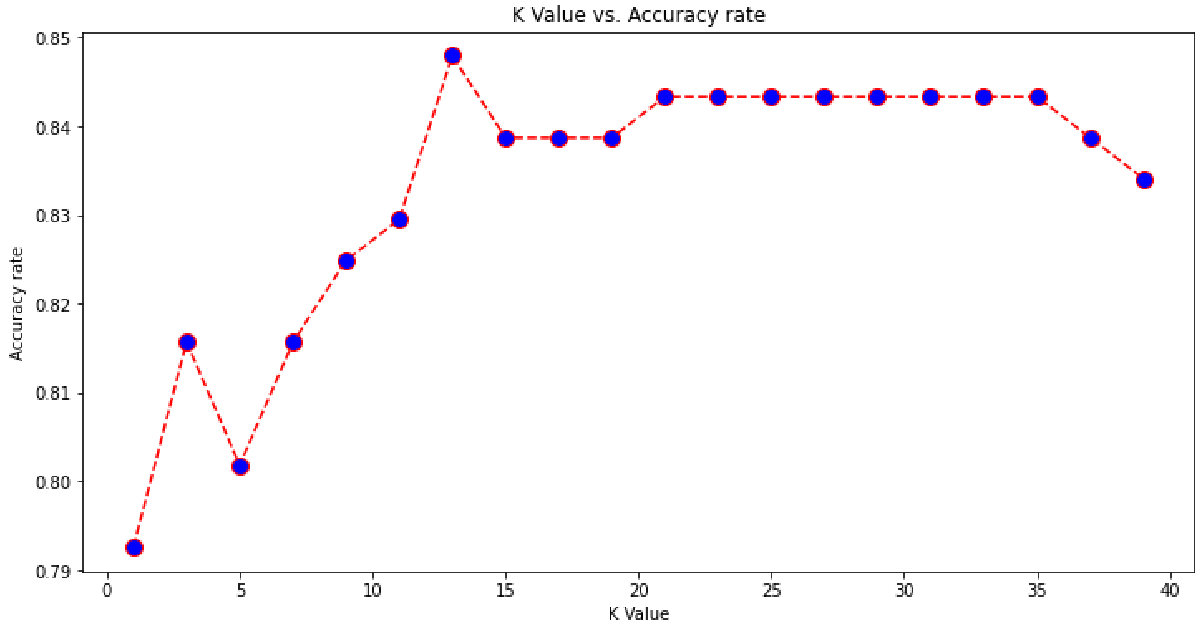
VIII. Attributes and weights for logistic regression model with no political data



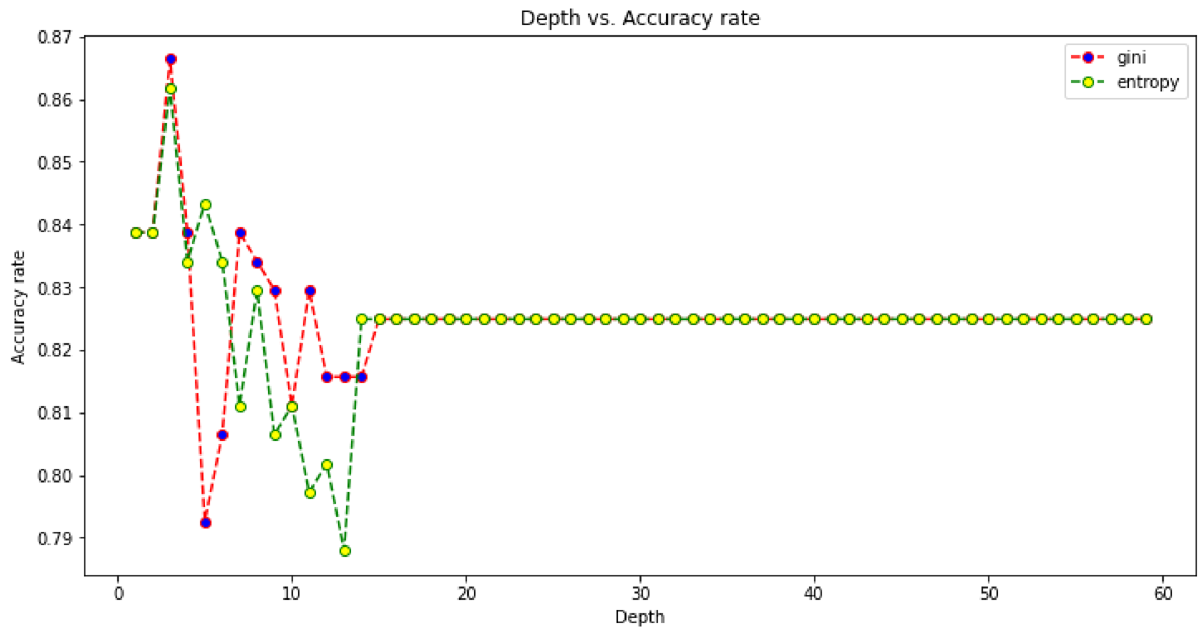
IX. KNN model with political parties



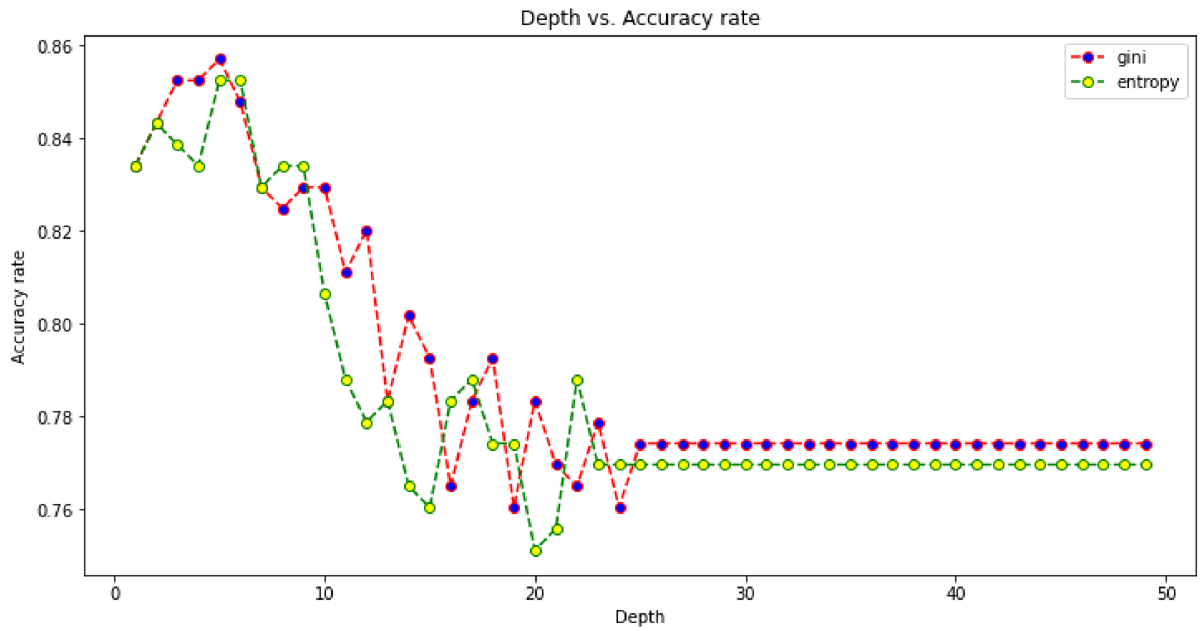
X. KNN model without political parties



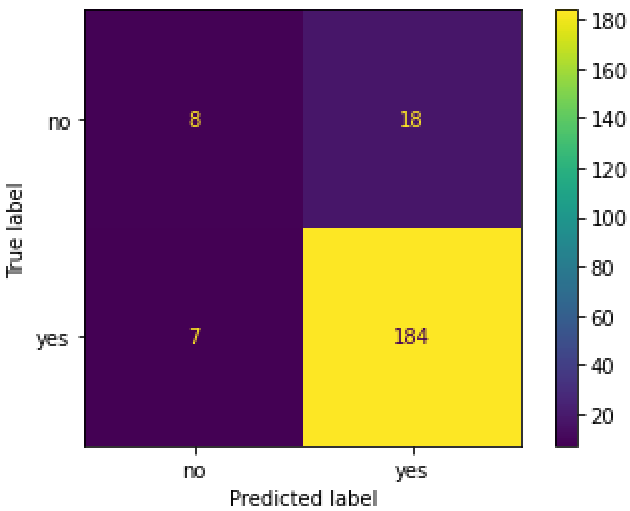
XI. Decision trees - With political parties



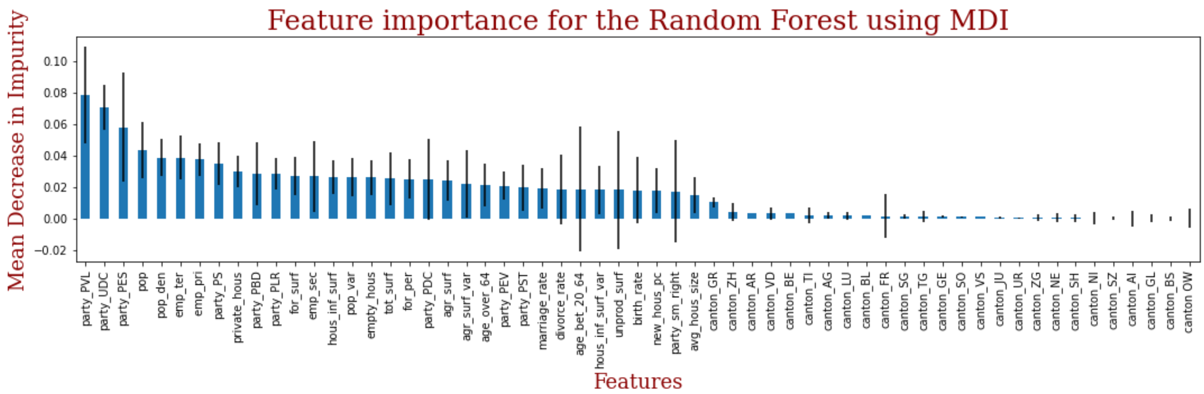
XII. Decision trees - Without political parties



XIII. Random Forest Confusion Matrix



XIV. Random forest feature importance with party variables



XV. Random forest feature importance without party variables