

## Analysis and Prediction of 2021 Election Results for The Constituencies in Coimbatore

<sup>1</sup>, Dr. Velvadivu.P, <sup>2</sup>, Dr.Sathya.C, <sup>3</sup>,Shree Varshini.R.P, <sup>4</sup>,Uthra.S

<sup>1</sup>-Faculty, Department of Computing,

<sup>2</sup>-Student, Department of Computing,

Coimbatore Institute of Technology, Coimbatore.

---

**ABSTRACT:** In this study we have analysed the results of the 14th (2011) and 15th (2016) Legislative Assembly Election to predict the result of the 16th Legislative Assembly Election-2021. The main focus of our study is to predict the results for the 2021 election. The data have been collected from official government websites and various authentic websites for 2011, 2016 and 2021 elections. The factors considered in the analysis and prediction are Party name, Age, Constituency, Assets, Liabilities, Criminal Cases etc. The Prediction is done using the k-nearest neighbors and Random Forest Classifier algorithms. The data of the 2011 and 2016 election was trained and tested using KNN and Random Forest classifier model, then the best model is used for the prediction of the 2021 election. The final result of the study would be whether the party would win or lose based on the data given for the election 2021. In this study, the predictions have been done for the Constituencies in Coimbatore, Tamil Nadu.

**KEYWORDS:** Election, Random Forest, KNN, Tamil nadu, Coimbatore.

---

### I. INTRODUCTION:

India is a constitutional democracy with a parliamentary system of government, and at the heart of the system is a commitment to hold regular, free and fair elections. The sixteenth legislative assembly election of Tamil Nadu was held on 6 April 2021, to elect representatives from the 234 constituencies in the Indian state of Tamil Nadu. The state of Tamil Nadu is divided into 234 assembly constituencies, each of which elects a member (called an MLA) to represent it at the state's Legislative Assembly. Tamil Nadu's partisan politics have been dominated by its two leading parties, Dravida Munnetra Kazhagam (DMK) and All India Anna Dravida Munnetra Kazhagam (AIADMK), for the last 50 years.<sup>[1]</sup> Coimbatore is a city in the Indian state of Tamil Nadu. Coimbatore is the second largest city in Tamil Nadu after Chennai and the 16th largest urban agglomeration in India as per the census 2011. Coimbatore elects ten members to the Tamil Nadu Legislative Assembly<sup>[14]</sup>.

### II. LITERATURE REVIEW:

According to T.Sivagnanasambandan<sup>[11]</sup>, though the Tamil Nadu ruling parties do not give enough opportunities for the opposition parties to expose their weaknesses, since there is effective media and most of the people in Tamil Nadu are educated, the exposure through media has brought about changes in the Government. According to Anand Agashe<sup>[1]</sup>, The social media, and the youth demographic becoming so central in electoral politics, good governance will become the central issue in political discourse. This will lead to future elections being fought on 'real' and substantive issues like food security, social security, employment, development, opportunities for growth and so on, unlike in the past. This bodes well for the future of India. According to Dr. Vipin Kumar Singhal<sup>[13]</sup>, Women have demonstrated considerable leadership in community and informal organisations, as well as in public office. However, socialisation and negative stereotyping of women and men, including stereotyping through the media, reinforces the tendency for political decision-making to remain the domain of men. According to Dr.V.S.Prasanth<sup>[10]</sup>, while there is developing cooperation of women in grass-root political developments, a great deal is required to be accomplished for an important portion of women in the formal political structure of the nation. According to Subramaniam Chandran<sup>[12]</sup>, Emerging leaders and parties are forced to accommodate their ideologies in coping with competitive politics. Electoral politics and voting behaviour reflect series of influences mooted by the politicians. According to A.Mathew<sup>[8]</sup>, The very terrain of education i.e., the pre-eminent position of TN in the educational landscape of India, as the image of the government, is a powerful and ideal political arena for the full expression of competitive politics. According to G.Palanithurai<sup>[9]</sup>, Globalisation has toppled many regimes in India. Also, in the 15th Lok sabha Election, the strong money background of the candidates and the distribution of money to the voters worked very well. According to Gongde Guo et.al<sup>[6]</sup>, The k-Nearest-Neighbours (kNN) is a non-parametric classification method, which is simple but effective in many cases. For a

data record  $t$  to be classified, its  $k$  nearest neighbours are retrieved, and this forms a neighbourhood of  $t$ . However, to apply kNN we need to choose an appropriate value for  $k$ , and the success of classification is very much dependent on this value. In a sense, the kNN method is biased by  $k$ . According to Jehad Ali et.al[7], Random Forest is a generic principle of classifier combination that uses  $L$  tree-structured base classifiers. The model interpretability and prediction accuracy provided by Random Forest is very unique among popular machine learning methods. Accurate predictions and better generalizations are achieved due to utilization of ensemble strategies and random sampling.

### III. RESEARCH METHODOLOGY:

The data for the analysis was collected from the official government website [2] and other authentic websites[3][4]. Data preprocessing was done using various methods. We started with the basic descriptive statistics and the count of null values in each column. The null values for the Criminal case column were filled '0' and the null values for the Assets and the Liabilities column were filled using Interpolation. The Education column's null values were filled using a Random forest classifier model where the factors were Age, Constituency, Assets, Liabilities etc. The Nominal categorical variables like the Candidate sex, Party name, Constituency name were encoded into numerical columns using binary encoder. And the ordinal categorical variable (i.e., Education) was encoded into a numerical column using Labelencoder. After the preprocessing, we trained and tested the 2011 and 2016 data in both kNN(k-nearest neighbor) and Random Forest Classifier Model to find which model is more effective. The Random Forest model was found to be more effective. Finally, we fitted the 2021 sample data to predict the results using the Random Forest Model.

#### DATA ANALYSIS AND INTERPRETATIONS:

Year	Con_name	District_Name	Candidate_Name	C_Sex	C_Age	Party_Name	Tot_Electors	Criminal_Cases	Education	Assets	Liabilities
2016	Mettupalayam	Coimbatore	CHINARAJ.O.K	M	62	AJDMK	279110	0	Graduate	6.321254e+06	0.0
2016	Mettupalayam	Coimbatore	SURENDRAN S.	M	45	DMK	279110	0	8th	3.345470e+07	5823889.0
2016	Mettupalayam	Coimbatore	SHANMUGA SUNDARAM.T.R	M	54	TMC(M)	279110	0	Others	1.791235e+07	3011889.5
2016	Mettupalayam	Coimbatore	JAGANATHAN P.	M	45	BJP	279110	0	Graduate	2.370000e+06	200000.0
2016	Mettupalayam	Coimbatore	None of the Above	NOTA	-1	NOTA	279110	0	Graduate	4.920302e+06	374891.0
2016	Mettupalayam	Coimbatore	MCHAMMED RAFI S	M	39	SDPI	279110	0	Graduate	7.498005e+06	548782.0
2016	Mettupalayam	Coimbatore	MOORTHIL K	M	46	PMK	279110	2	Graduate	1.004401e+07	724673.0
2016	Mettupalayam	Coimbatore	VELLINGRIL K	M	54	KNCK	279110	0	5th	2.328905e+06	0.0
2016	Mettupalayam	Coimbatore	ABOULVAHAS A	M	39	NTK	279110	1	Graduate	4.680000e+05	15000.0
2016	Mettupalayam	Coimbatore	VEENKATESH K	M	45	IND	279110	0	Others	3.330000e+06	0.0

Figure 1: Tamilnadu general Legislative Assembly 2011 and 2016 dataset[2][3]

Year	Constituency_Name	District_Name	Candidate_Name	C_Sex	C_Age	Party_Name	Total_Electors	Criminal_Cases	Education	Assets	Liabilities
2021	Mettupalayam	Coimbatore	SELVARAJ	M	62.0	AJDMK	279005	0.0	Graduate	1.047096e+07	0.0
2021	Mettupalayam	Coimbatore	SHANMUGASUNDARAM P	M	38.0	IND	279005	0.0	Graduate	2.500000e+04	0.0
2021	Mettupalayam	Coimbatore	SHANMUGA SUNDARAM T.R	M	58.0	DMK	279005	1.0	8th	1.570378e+09	15786749.0
2021	Mettupalayam	Coimbatore	SRINIVASAN	M	47.0	IND	279005	0.0	PG	9.955000e+06	3900000.0
2021	Mettupalayam	Coimbatore	MAHESHVARAN	M	47.0	VTVTK	279005	0.0	5th	2.800000e+04	0.0
2021	Mettupalayam	Coimbatore	YASMIN	F	40.0	NTK	279005	0.0	Literate	5.220775e+06	700000.0
2021	Mettupalayam	Coimbatore	LAKATHALU	M	46.0	IND	279005	0.0	8th	1.106000e+05	262421.0
2021	Mettupalayam	Coimbatore	SARAVANAN	M	51.0	AJMK	279005	0.0	10th	8.977000e+06	5000000.0
2021	Mettupalayam	Coimbatore	NARAPPAN	M	55.0	IND	279005	0.0	8th	2.029000e+04	0.0
2021	Mettupalayam	Coimbatore	SHANMUGASUNDARAM K	M	46.0	IND	279005	0.0	12th	1.113574e+06	300000.0

Figure 2: Coimbatore general Legislative Assembly 2021 dataset[3][4]

The above tables contain details of the Nominees in 2011 and 2016 Coimbatore Election and 2021 Coimbatore Election. It contains Year of Election, Constituency Name, Gender, Age, Party Name, Criminal Cases, Assets, Liabilities and Win or Loss. The details dataset is used to train the model and predict 16th Legislative Assembly election in Tamil Nadu.

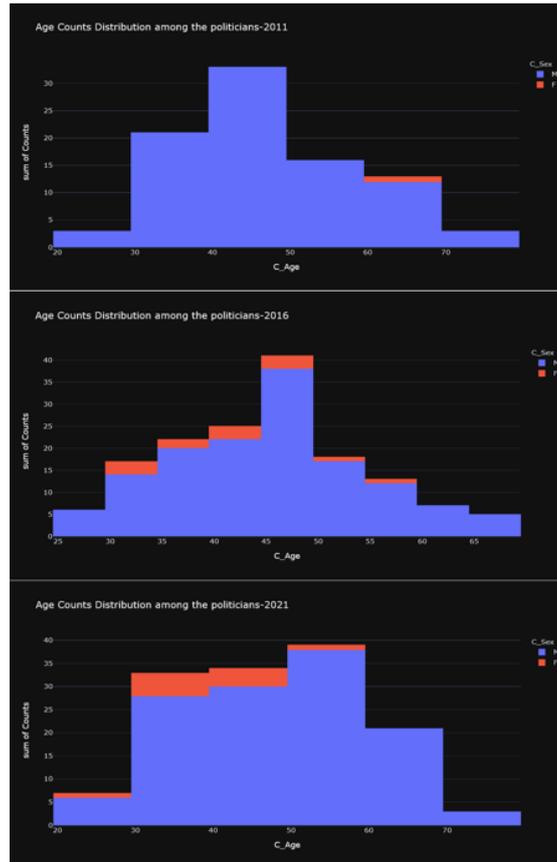


Figure 3: Age Counts Distribution among the politicians - 2011, 2016 and 2021

The above histogram plot shows the age distribution among the candidates. For 2011 there are very few female candidates and they belong to the 60-70 age range. In the 2016 election most of the candidates belong to the 40-50 age range and for the 2021 election the most of the candidates age is distributed between 30-60 range.

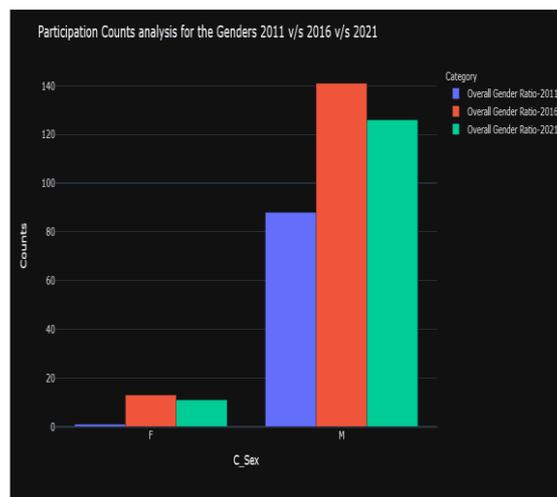


Figure 4: Comparison of Participation Counts for the Genders in Coimbatore in 2011,2016 and 2021

The above bar graph compares the Participation of male and female Candidates in 2011,2016 and 2021. This shows that the participation of candidates has been relatively increased in the 2016 election compared to the 2011 election but the participation of candidates has decreased in 2021 compared to 2016.

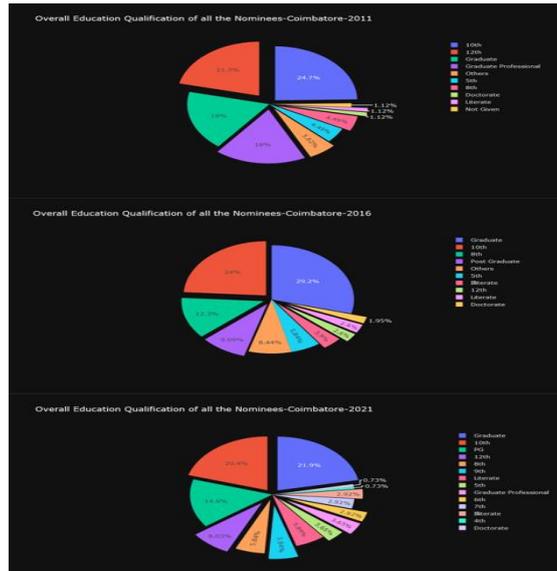


Figure 5: Overall Education of all the Nominees-Coimbatore in 2011, 2016 and 2021  
 The above piechart gives a graphical analysis of Educational Qualification of Nominees in Coimbatore in 2011, 2016 and 2021. In the 2011 election, most of the candidates had only Secondary education followed by Higher Secondary education. But in the 2016 and 2021 elections, nearly half of the candidates were Graduates followed by Graduate Professionals. This shows that most of the candidates in the 2016 and the 2021 election are educated and have the basic educational qualification.

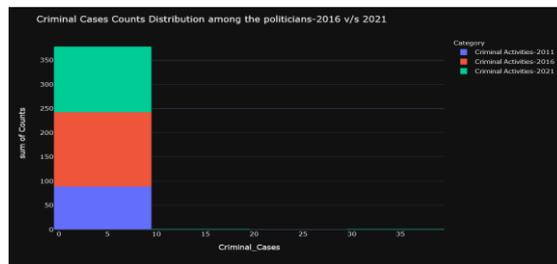


Figure 6: Criminal Cases Distribution among politicians -Coimbatore in 2011, 2016 and 2021.  
 The above graph shows that most of the candidate's criminal cases are between 0-10 but in 2021 candidates there are few candidates who have higher criminal cases.

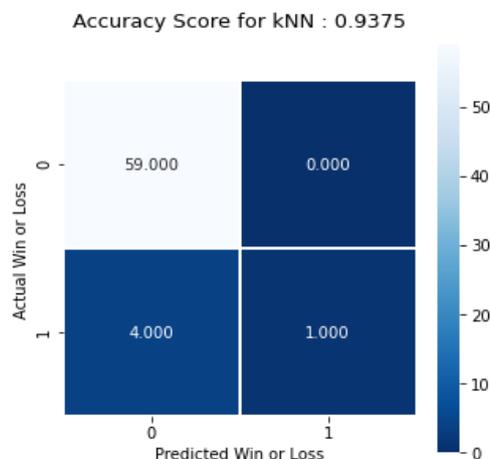


Figure 7: Actual and Predicted Win and Loss of KNN model

The above confusion matrix gives the accuracy and the count of Actual vs Predicted Win Loss of KNN model which is tested with obtained collected Coimbatore Election data. The pre-processed data of the 2011 and the 2016 Coimbatore Election was trained and tested using k-nearest neighbours algorithm which gives an accuracy of 93.75%. But when we look at the confusion matrix we can see that there are errors when predicting the “win” of Candidates. Since only 1 is correctly predicted we fitted the Random Forest Model.

(189, 22) (64, 22) (189,) (64,)  
 RMSE on train data: 0.0  
 RMSE on test data: 0.21650635094610965  
 Accuracy: 0.953125

Figure 8: Statistical data of Random Forest Model fitted with all Features

From the above statistical data, we infer that the Root Mean Squared Error of training data is 0% while for testing data it is 21% which is comparatively less. Also, accuracy of the model fitted with all the features is 95.3% which is appreciable.

Accuracy Score of Random Forest Model fitted with all features: 0.953125

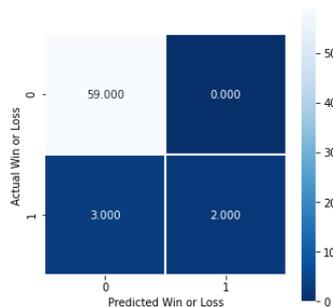


Figure 12: Actual and Predicted Win and Loss of Random Forest model fitted with all Features

The above confusion matrix gives the accuracy and the count of Actual vs Predicted Win Loss of Random Forest model fitted with all the Features which is tested with the collected Election dataset. We infer that the fitted Random forest model is better in terms of Accuracy and Prediction compared to the kNN model. To improve the performance of the model we try to fit the model with important features only.

	feature	importance
1	C_Age	0.196230
4	Assets	0.148354
20	Party_Name_6	0.100534
17	Party_Name_3	0.085523
5	Liabilities	0.084513
2	Tot_Electors	0.065440
18	Party_Name_4	0.065363
19	Party_Name_5	0.056764
21	Edu_ordinal	0.044710
0	Year	0.034318
16	Party_Name_2	0.023467
8	Con_name_2	0.021034

Figure 9: 12 Important Features of Random Forest Model and its Importance Values. The above table gives the 12 most Important Features of the above fitted Random Forest Model and its Importance Values. We infer that Age and Assets of a Candidate plays a major role in predicting the Win or Loss of a candidate in the election.

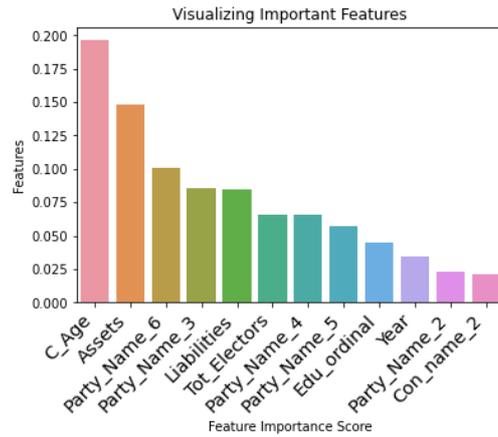


Figure 10: 12 most Important Features and its value.

The above graph shows the 12 most important features of the Random Forest Model and its importance value. Features are taken along X-axis and Feature Importance Score along Y-axis.

Accuracy Score of Random Forest Model fitted with 9 most important Features : 0.984375

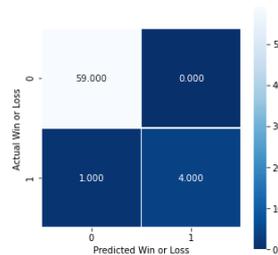


Figure 11: Actual and Predicted Win and Loss of Random Forest model fitted with 9 most Important Features

The above confusion matrix gives the accuracy and the count of Actual vs Predicted Win Loss of Random Forest model fitted with 9 most Important Features which is trained and tested with collected Election data. This shows that the Random Forest model fitted with 9 most important features gives an accuracy of 98.43% which is comparatively higher than the model fitted with all the features. So using this model for prediction would give us a better result with less time compared to the model with all the features.

	Constituency_Name	Party_Name
253	Mettupalayam	AIADMK
276	Sulur	AIADMK
287	Kavundampalayam	AIADMK
300	Coimbatore (North)	AIADMK
317	Thondamuthur	AIADMK
330	Coimbatore(South)	AMMK
359	Singanallur	MNM
363	Kinathukadavu	AIADMK
381	Pollachi	AIADMK
385	Valparai	AIADMK

Figure 12: Constituency Name and Predicted Winning Party

The above table gives the predicted result using the Random Forest Model fitted with 9 most Important Features. The Prediction is made for the Constituencies in Coimbatore and according to our prediction AIADMK will win in most of the constituencies i.e., Mettupalayam, Sulur, Kavundampalayam, Coimbatore(North), Thondamuthur, Kinathukadavu, Pollachi and Valparai while a win for AMMK in Coimbatore(South) and for MNM in Singanallur.

#### IV. CONCLUSIONS:

Coimbatore recorded a voter turnout of 68.72% which was marginally higher than the previous assembly elections on April 6, 2021. However, the turnout was lower than many other districts in Tamil Nadu. According to election officials, 21 lakh of the total 30 lakh voters spread across the 10 assembly constituencies in the district exercised their right to franchise on the day. Unlike the 2016 election, female voters outnumbered their male counterparts this time<sup>[15]</sup>. And based on the above analysis and prediction, we can see that AIADMK will win in 8 constituencies while AMMK and MNM has chance to win 1 constituency each in the district. This study was conducted based on the 14th and 15th Legislative Assembly election results.

#### V. ABBREVIATION OF POLITICAL PARTIES:

AIADMK: All India Anna Dravida Munnetra Kazhagam

DMK : Dravida Munnetra Kazhagam

MNM : Makkal Needhi Maiam

AMMK : Amma Makkal Munnetra Kazhagam

#### REFERENCES:

- [1] Anand Agashe, 2014, The 16th General Election - A Sizeable Section of Indian Voters in a Mood for 'High risk, High returns', FES India Paper, 8
- [2] Anon, 2021 Tamil Nadu Legislative Assembly Election, Accessed on 5th April 2021, [https://en.wikipedia.org/wiki/2021\\_Tamil\\_Nadu\\_Legislative\\_Assembly\\_election](https://en.wikipedia.org/wiki/2021_Tamil_Nadu_Legislative_Assembly_election)
- [3] Anon, Tamil Nadu General Legislative Election 2016, Accessed on 23rd March 2021, <https://eci.gov.in/files/file/3473-tamil-nadu-general-legislative-election-2016/>
- [4] Anon, Tamil Nadu 2016, Accessed on 23rd March 2021, <https://myneta.info/tamilnadu2016/>
- [5] Anon, TN Election Promises 2021, Accessed on 23rd March 2021, <https://en.electionpromises.in/>
- [6] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, Kieran Greer, 2004, KNN Model - Based Approach in Classification
- [7] Jihad Ali , Rehanullah Khan, Nasir Ahmad, Imran Maqsood, 2012, Random Forests and Decision Trees, International Journal of Computer Science, 9, 5(3), 273
- [8] A.Mathew, 2016, Competitive Politics in Tamil Nadu Higher Education Policy, The Indian Journal of Technical Education, 39, 2, 17
- [9] G.Palanithurai, 2009, An analysis of 15th Lok Sabha Election in Tamil Nadu, Madhya Pradesh Journal of Social Sciences, 14, 2, 45
- [10] Dr.V.S.Prasanth, 2020, Political Status of Women in Tamilnadu, Dogo Rangsang Research Journal, 10, 6(7), 280
- [11] T.Sivagnanasambandan, 2012, The Role played by the opposition parties in the Tamil Nadu Legislative Assembly 1992-2006 A Study, Shodhganga,
- [12] Subramaniam Chandran, Elections in Tamilnadu: Who wins, Why, How?
- [13] Dr.Vipin Kumar Singhal, 2015, Political Empowerment of Women - An Analytical Analysis ,Sunrise Publications, 3
- [14] Anon, 2021, Coimbatore, Accessed on 10th April 2021, [https://en.wikipedia.org/wiki/Coimbatore#Administration\\_and\\_politics](https://en.wikipedia.org/wiki/Coimbatore#Administration_and_politics)
- [15] Mayilvaganam, 2021, Dist records 68.32% voter turnout, The Times of India, Accessed on 10th April 2021, <https://timesofindia.indiatimes.com/city/coimbatore/dist-records-68-32-voter-turnout/articleshow/81939982.cms>