International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 6 Issue: 5 171 - 173

Naïve Bayes Model for Analysis of Voting Rate Failure in Election System

Gururaj K S
Ph.D. Scholar, Computer Science & Engineering,
Visvesvaraya Technological University,
Belagavi, India
gururaj.k.s79@gmail.com

Dr. K.Thippeswamy.
Professor and Chairman,
Department of Studies in CSE, Regional Center,
Visvesvaraya Technological University
Mysuru, India
thippeswamy@vtu.ac.in

Abstract—These days several exercises are striking the engineering perspectives on non-engineering domain and it is major challenging too for the engineering community. One such non-engineering domain is prediction of the results in election system. So many factors like literacy, rainfall, population and so on may affect the results in election systems. Certainly there are some parameters which are relatively dependent on each other. In this connection, this paper tried to design a probabilistic model to analyze failure voting rates in election system by involving one of the strongest baye's model called as Naïve Bayes Model.

Keywords—*Naïve Bayes Model; Election system; Failure rate in voting*

*****

I. INTRODUCTION

In democratic countries, the selection of right leader on every few years is one of the major responsibility of each and every individual and it is crucial part of the system, in turn affect the development of the country. Due to some reasons, part of the community will reject in the involvement of the selection process. This rejection may lead to selection of person who may not be fit for the position. Certainly that causes many problems and may lead to spoil/corrupt the system. In this connection there are several statistical approaches available to analyze the issue. Hence certain engineering or statistical tools are necessary to analyse the reason for failure in voting rates at the minimal time. Election system involves either the entire city or districts. Indeed the size of the data will be obviously in large size. This paper tried an probabilistic model as a engineering perspective to analyze the big amount of data or information in minimal time for failure voting rate in the selection process of right leader.

II. REVIEW OF LITERATURE

A. Naïve Bayes Classifier

Naïve Bayes Model is called as one of the strongest bayes approach while designing predictive and decision making systems. Equation 1 is used to find the conditional probability when some parameters are given

$$P(Cause,Effect1,..,Effectn) = P(Cause)\ \pi i P(Effecti|Cause) \quad (1)$$

Cause is the actual parameter to estimate the probability when some effects were given.

Figure 1 shows the arrangement of cause and effects in the estimation of probability estimation. Marsono, M.N et. al. proposed a hardware architecture span control by targeting Altera Stratix CPLD device using Naïve bayes classifier [2]. Also they have tried to reduce the computation complexity by including Logarithmic Number System. Thakur, S.K. et. al. proposes a methodology using Naïve Bayes Model for machine classification of Nepali Texts.

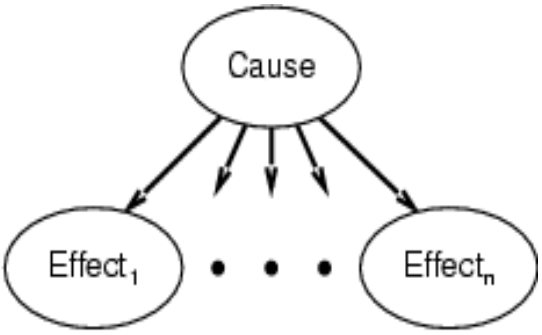

Figure1: Naïve Bayes Model for the Conditional Probability Estimation

In this they have tried to introduce an hybrid approach called as lexicon-pooled Naive Bayes Classifier, found an efficient result during classification [1].

III. METHODOLOGY

Figure 2 shows the proposed methodology of the paper. The proposed methodology is mainly on naïve bayes classifier. In the first phase, data's are collected and normalized then estimate conditional probabilities of all the mentioned random variables when maximum voting is given. Post processing phase includes the classification of the samples as worst, best of average classes. Thus finding out reasons for the failure voting rates during elections.

Apply these estimations in the Naïve Bayes classifier as shown in equation 2.

$$P(V, L, G, R) = P(V) \times P(L/V) \times P(P/V) \times P(G/V) \times P(R/V) \quad (2)$$

Where,

V = Voting, L=Literacy, G=GDP and R=Rainfall

Here probability of voting can be computed, provided P(L/V), P(G/V) and P(R/V) are known through survey.

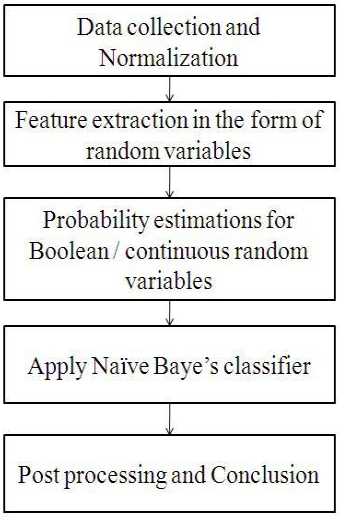

Figure 2: Methodology for the analysis of Voting Rate Failure.

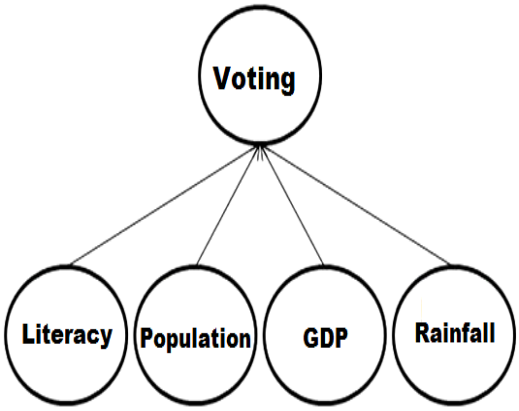

Figure 3: Naïve bayes model for the analysis of Voting Rate Failure.

Figure 3 shows the possible causes and effect in the analysis of failure voting rate. As per the analysis literacy, population count, Gross Domestic Product and Rainfall rate would be the parameters which may effect on voting. Certainly literacy will educate the people to understand the importance and need of voting. Similarly the increase in population may create the need of food and shelter, which may leads to corruption / other unwanted demands. Sometimes the developments in the districts may also trigger the people to vote. India is rich and mainly depends on cultivation. Hence proper rain certainly fulfills the demands of the people thereby may increase the interest in voting.

Equation 2 yields the probability estimation of Voting when literacy, GDP and Rainfall rates of the location is given. Post processing classifies the worst, average and best voting

IV. RESULTS AND DISCUSSION

A. Voting rates [6][7]

Table 1: Voting rates of various loka sabha elections

General Election	Year	Male	Female	Total Voting
1st	1952 [4]	-	-	61.2
2nd	1957 [5]	-	-	62.2
3rd	1962	63.31	46.63	55.42
4th	1967	66.73	55.48	61.33
5th	1971	60.90	49.11	55.29
6th	1977	65.63	54.91	60.49
7th	1980	62.16	51.22	56.92
8th	1984	68.18	58.60	63.56
9th	1989	66.13	57.32	61.95
10th	1991	61.58	51.35	56.93
11th	1996	62.06	53.41	57.94
12th	1998	65.72	57.88	61.97
13th	1999	63.97	55.64	59.99
14th	2004	52.65	44.65	48.74
15th	2009	60.24	55.82	58.19
16th	2014	67.09	65.63	66.4

Table 1 is showing the voting rates in percentage at all loka sabha elections. Total voting is the ratio of number of people voted and the total number of people eligible to vote. 16th election has made new record in voting, i.e., the highest voting when compared to all other elections. Hence the conditional probabilities namely P(L/V), P(P/V), P(G/V) and P(R/V) at this year can be called as corresponding probabilities when maximum voting is given.

That is Probability of Literacy when voting is given: P(L/V) = 0.67 Probability of GDP when voting is given: P(G/V)=0.14 Probability of Rainfall when voting is given: P(R/V)=0.32 Using the above prior probabilities, classification of all the elections as per the naïve bayes is given in the table 2.

Table 2: Naïve bayes estimation for including all the parameters

General Election	P(V)	P(V, L, G, R) * 10	Class
1st	0.612	0.183698	Best
2nd	0.622	0.1867	Best
3rd	0.5542	0.166349	Worst
4th	0.6133	0.184088	Best
5th	0.5529	0.165958	Worst
6th	0.6049	0.181567	Best
7th	0.5692	0.170851	Average
8th	0.6356	0.190782	Best
9th	0.6195	0.185949	Best

10th	0.5693	0.170881	Average
11th	0.5794	0.173913	Average
12th	0.6197	0.186009	Best
13th	0.5999	0.180066	Best
14th	0.4874	0.146298	Worst
15th	0.5819	0.174663	Average
16th	0.664	0.199306	Best

The third column of the table 2 is processed under k-means algorithm for the classification yields the fourth column classification.

IV. Conclusion

This paper has tried to predict the reason for the failure in voting rates using Naïve Bayes Probabilistic model by considering four major parameters. As per the analysis these major factors are used and found the rated the voting in elections. This outcome can be further enhanced to predict the forthcoming election voting.

ACKNOWLEDGMENT

We would like to thank Election Commission of India and Census of India for providing data and relevant information through web media. We would like to thank Regional Center, VTU, Mysuru and GSSS Institute of Engineering and Technology for Women, Mysuru for their kind support

References

[1] Thakur, S.K.; Singh, V.K., "A lexicon pool augmented Naive Bayes Classifier for Nepali Text," in Contemporary Computing (IC3), 2014 Seventh International Conference on ,vol., no., pp.542-546, 7-9 Aug. 2014, doi: 10.1109/IC3.2014.6897231

[2] Marsono, M.N.; El-Kharashi, M.W.; Gebali, F., "Binary LNS-based naive Bayes hardware classifier for spam control," in Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on , vol., no., pp.4 pp.-3677, 21-24 May 2006, doi: 10.1109/ISCAS.2006.1693424

[3] Eugene Lukacs (1942). "A Characterization of the Normal Distribution". The Annals of Mathematical Statistics 13 (1): 91–93. doi:10.1214/aoms/1177731647.

[4] "Statistical Report on Lok Sabha Elections 1951-52" (PDF). Election Commission of India. Retrieved 23 November 2014.

[5] "Statistical Report on Lok Sabha Elections 1957". Election Commission of India.

[6] http://ceo.uk.gov.in/pages/view/165/169-voting-percentage-in-various-lok-sabha-elections

[7] http://eci.nic.in/eci_main1/SVEEP/VoterTurnoutHighlightsLokSabha2014.pdf

About the Authors

Dr. K Thippeswamy:

Received his Ph.D. degree from the Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Ananthapur, Andra Pradesh in the year 2012, M.E degree in Computer Science and Engineering from University Visvesvaraya College of Engineering (UVCE), Bangalore in 2004 and Bachelors Degree in Computer Science and Engineering from University B.D.T College of Engineering (UBDTCE), Davangere in the year 1998. He is currently heading the Department of Computer Science and Engineering, Visvesvaraya Technological University, PG Center, Mysore, Karnataka, where he is involved in research and teaching activities. His major areas of research are Data Mining & Knowledge Discovery, Big data, Information Retrieval, and Cloud Computing. He is having 18 years of Teaching and 3 years Research experience. He has published around 40 papers which include International Journals, International Conferences and Notional Conferences, he has conducted two national and one International conference successfully. He is a Life member of India Society for Technical Education (LMISTE), Computer Society of India (CSI) & International Association of Engineers (IAENG). He is Reviewer Committee Member for the international Journal Bioinformatics and Data Mining. He has organized one International Conference and two National Conference and many workshops. Currently he is working as Professor and HOD Dept. of Computer Science & Engg. VTU, PG Centre, Mysuru.

Gururaj K S:

Pursuing his Ph.D. degree in Visvesvaraya Technological University, Belagavi, India and Working as a Associate Professor in GSSS Institute of Engineering and Technology for Women, Mysore, India