POPULATION DYNAMICS: COMPARATIVE ANALYSIS OF MATHEMATICAL MODELS AND MACHINE LEARNING TECHNIQUES FOR PREDICTING FUTURE POPULATION GROWTH

Abstract

The global population is increasing at a very fast rate, and it is likely to increase another two billion by 2050. India has also seen a prominent increase in demographical numbers across the states. India has already surpassed China in the Population this year. Keeping in view of the current situation, population dynamics has a high scope and importance. It deals with changes in population growth rates, age structures, distribution of people and these factors are closely linked to national and global development challenges and their solutions. Many population growth models have been developed in the past by demographers and scientists to study and predict the future population. Exponential, Logistic and Fibonacci population models are few among them. Fibonacci models are based on Fibonacci numbers in Mathematics and its research study was based on Rabbit population. Later it is generalized and extended to human population. Exponential model created a significant place among the population models and predicted that the world population will increase exponentially. Contradicting to this, Logistic population model said that the population can grow either increasing or decreasing based on the carrying capacity- a new term introduced by Verhulst. Recently, machine learning techniques are introduced to predict the population growth and it gives a more approximate and reliable predictions as compared to differential equation based mathematical techniques. This chapter goes deeply into modeling the human population growth of Fibonacci, Exponential and Logistic population models, and machine

Authors

Sheena K J
Research Scholar
VIT Bhopal University
Madhya Pradesh, India.
sheena.kj2021@vitbhopal.ac.in

Jyoti Badge
Assistant Professor
VIT Bhopal University
Madhya Pradesh, India.
jyoti.badge@gmail.com

learning techniques. The data set is collected from 2011 census of 5 major Indian states, analyzed, compared, and predicted the future population till 2031 using the machine learning techniques and Mathematical population models. Among the machine learning techniques, Random Forest method has been used here to predict the population.

Keywords: Population growth, Exponential model, Logistic model, Fibonacci model, Random Forest method, Future population

I. INTRODUCTION

Many research studies have been conducted using Exponential, Logistic and Fibonacci population models mathematically. These mathematical predictions are accurate to some extent. This research study is an attempt to attain a more accurate and reliable prediction by comparing mathematical and machine learning methods. Machine learning models are gaining utmost importance in the modern world. There are several machine learning methods available for predicting the human population – Linear regression, decision trees, artificial neural network, k-Nearest neighbors – to name a few. There have been studies in the recent past that talks about the comparison of 17 machine learning models[3], forecasting small area populations with Long Short- Term memory networks[2,5], study in Nepal that estimates the post-earthquake human migration within Nepal by tracing the human movement using their mobile phone data[4], another paper that explores supervised ML algorithms to identify the most suitable algorithm for predicting the population growth rate of a particular area[1], population prediction of Turkey with six different machine learning algorithms were done in 2017[6], population forecasting system using machine learning algorithm in Nigeria[8], a model that is used to predict the disease severity and outcome in Covid-19 patients[11] etc.

The data set is taken from 2011 census Govt. of India that deals with the population of first 5 most populated Indian States – namely Uttar Pradesh, Bihar, Maharashtra, West Bengal and Madhya Pradesh. As per the 2011 census, it is observed that 50% of India's demographic growth is expected to take place in these 5 states and hence the projections of these particular states are of great importance. Hence the data sets of these Indian states have been taken in this research study. Here Python programming and MATLAB programming are used to calculate the predictions of these 5 states from 2012 to 2031 using mathematical methods as well as machine learning methods.

II. RESEARCH METHODOLOGY

1. **Exponential Growth Model:** Exponential growth model developed by Malthus [16] consider the rapid growth of human populations against the backdrop of limited resources. In fact, Malthus predicted a society plagued by misery and hunger.

   - This growth model is described by the differential equation $\frac{dNd}{dt}$ = rN, where r = b – d (Fertility – Mortality)

   - On integrating, we get $N(t) = N_0 e^{r(t-t_0)}$, r is the Malthusian factor which is a constant.

2. **Logistic Population Growth Model:** Formulated by Pierre Verhulst, this model [17] introduced a term called carrying capacity. Verhulst found that population growth not only depends on the population size, but also on how far this size is from the carrying capacity. As per Verhulst, population cannot always increase because there will be scarcity of resources like water, food etc. So, he predicted that population will decrease as and when the resources will run out of short. To summarize, the population growth rate decreases as the number of individuals increases. He formulated the equation as follows.

$$\frac{dN}{dt} = \frac{a\,N(t)[a - b\,N(t)]}{a} \quad \text{---------- (1)}$$

a and b are vital coefficients.

On integration and applying the initial values, we get the following equations and with the use of these equations, we have created a MATLAB code to calculate the numbers.

$$N_{max} = \frac{N_1(N_0N_1 - 2N_0N_2 + N_1N_2)}{N_1{}^2 - N_0N_2}$$

$$N = \frac{a/b}{1 + (\frac{\frac{a}{b}}{N_0} - 1)e^{-at}}$$

$$e^{-a} = \frac{N_0(N_2 - N_1)}{N_2(N_1 - N_0)}$$

III. DATA TABLES AND DIAGRAMS

My first study was based on the Exponential and Logistic population models of 2011 census data set of 20 Indian states. The population of the year 2011 has been taken as the base year and predicted the population for the years 2012 to 2031 using both the Exponential and Logistic population models. Table 1 and table 2 depict the comparison of actual and predicted population using both the models for the said years whereas Table 3 tells us the population predictions for the years 2023 to 2031of all the 20 Indian states. The graphs have been plotted to show the comparison between Exponential and Logistic model and we can observe that the exponential model is more accurate and almost matches with the actual data in all the cases. Figure 1, Figure 2 and Figure 3 portray the actual, exponential and logistic curves of each state for the years 2012 to 2022. Figure 4 shows the population predictions of all 20 states for the upcoming years from 2023 to 2031.

Table 1: Population Predictions – Exponential Model vs Logistic Model (2012- 2016)

Indian States	2012			2013			2014			2015			2016		
	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P
Jammu & Kashmir	12,393	12,367	12,459	12,519	12,467	12,647	12,645	12,569	12,836	12,771	12,671	13,024	12,897	12,774	13,212
Himachal Pradesh	6,923	6,922	7,888	6,982	6,979	7,967	7,041	7,036	8,048	7,100	7,094	8,132	7,158	7,152	8,213
Punjab	28,023	28,016	28,161	28,302	28,292	28,521	28,581	28,571	28,880	28,861	28,852	29,236	29,140	29,136	29,590
Haryana	25,772	25,757	25,784	26,193	26,170	26,176	26,614	26,590	26,568	27,034	27,016	26,960	27,455	27,499	27,352
NCT of Delhi	17,166	17,149	17,180	17,544	17,519	17,530	17,921	17,896	17,880	18,299	18,281	18,230	18,677	18,675	18,578
Rajasthan	69,687	69,647	69,630	70,825	70,763	70,663	71,963	71,897	71,695	73,102	73,049	72,725	74,240	74,220	73,753
Uttar Pradesh	2,03,067	2,02,970	2,03,160	2,06,322	2,06,190	2,06,250	2,09,577	2,09,450	2,09,340	2,12,832	2,12,760	2,12,430	2,16,087	2,16,130	2,15,510
Bihar	1,06,115	1,06,040	1,06,380	1,08,130	1,08,020	1,08,520	1,10,145	1,10,040	1,10,650	1,12,161	1,12,090	1,12,780	1,14,176	1,14,190	1,14,900
Assam	31,598	31,510	31,512	31,991	31,817	31,868	32,383	32,127	32,224	32,776	32,441	32,582	33,168	32,757	32,940
west Bengal	92,037	92,028	92,053	92,797	92,785	92,914	93,558	93,549	93,773	94,318	94,319	94,630	95,079	95,096	95,486
Jharkhand	33,551	33,533	33,416	34,113	34,088	33,901	34,675	34,652	34,385	35,238	35,225	34,868	35,800	35,807	35,350
Odisha	42,373	42,366	42,268	42,771	42,762	42,646	43,170	43,162	43,022	43,568	43,565	43,398	43,966	43,972	43,772
Chhattisgarh	25,950	25,939	25,998	26,356	26,339	26,427	26,761	26,745	26,858	27,166	27,157	27,289	27,571	27,576	27,721
Madhya Pradesh	73,863	73,821	73,896	75,099	75,034	75,021	76,334	76,267	76,145	77,570	77,520	77,269	78,806	78,794	78,390
Gujrat	61,383	61,353	61,304	62,327	62,281	62,273	63,271	63,222	63,242	64,214	64,177	64,211	65,158	65,147	65,180
Maharashtra	1,13,645	1,13,620	1,13,260	1,14,915	1,14,870	1,14,370	1,16,186	1,16,140	1,15,480	1,17,456	1,17,430	1,16,590	1,18,727	1,18,730	1,17,690
Andhra Pradesh	49,936	50,105	49,938	50,295	50,639	50,189	50,654	51,179	50,440	51,012	51,724	50,691	51,371	52,275	50,942
Karnataka	61,722	61,709	61,734	62,349	62,329	62,358	62,975	62,956	62,981	63,602	63,588	63,604	64,229	64,227	64,227
Kerala	33,641	33,637	33,704	33,875	33,870	33,963	34,109	34,105	34,221	34,344	34,341	34,478	34,578	34,579	34,733
Tamil Nadu	72,645	72,639	72,602	73,142	73,135	73,241	73,640	73,634	73,878	74,137	74,136	74,514	74,635	74,642	75,147

Actual (A) vs Exponential Predicted (EXP-P), Logistic Predicted (LOGI-P) (In thousands):

In the exponential model, the growth rate of each state is calculated using the 2011 census data and performed the MATLAB code to get the predicted population for each year for 20 different states. When we look at the growth rate, Tamil Nadu and Kerala are at the lowest growth rate of 0.0068 and 0.0069 respectively whereas the highest growth rate of 0.0213 and 0.0185 are recorded by NCT of Delhi and Bihar respectively.

Table 2: Population Predictions – Exponential Model vs Logistic Model (2017- 2022)

Indian States	2017			2018			2019			2020			2021			2022		
	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P	A	EXP-P	LOGI-P
Jammu & Kashmir	12,999	12,878	13,400	13,101	12,983	13,588	13,203	13,088	13,744	13,305	13,195	13,961	13,408	13,302	14,146	13,505	13,410	14,330
Himachal Pradesh	7,206	7,211	8,299	7,253	7,271	8,387	7,300	7,330	8,476	7,347	7,391	8,567	7,394	7,452	8,660	7,431	7,513	8,756
Punjab	29,380	29,423	29,942	29,619	29,713	30,292	29,859	30,006	30,638	30,099	30,301	30,983	30,339	30,599	31,324	30,535	30,901	31,662
Haryana	27,861	27,889	27,742	28,266	28,336	28,132	28,672	28,790	28,521	29,077	29,251	28,908	29,483	29,720	29,294	29,846	30,196	29,678
NCT of Delhi	19,056	19,077	18,925	19,435	19,487	19,271	19,814	19,907	19,616	20,193	20,335	19,958	20,571	20,733	20,298	20,965	21,220	20,635
Rajasthan	75,248	75,410	74,778	76,256	76,618	75,800	77,264	77,846	76,818	78,273	79,094	77,833	79,281	80,361	78,842	80,153	81,649	79,847
Uttar Pradesh	2,19,051	2,19,550	2,18,590	2,22,015	2,23,020	2,21,660	2,24,979	2,26,550	2,24,720	2,27,943	2,30,140	2,27,780	2,30,907	2,33,780	2,30,820	2,33,297	2,37,480	2,33,850
Bihar	1,15,957	1,16,320	1,17,020	1,17,739	1,18,490	1,19,120	1,19,520	1,20,700	1,21,210	1,21,302	1,22,960	1,23,290	1,23,083	1,25,250	1,25,360	1,24,919	1,27,590	1,27,400
Assam	33,543	33,076	33,298	33,918	33,398	33,657	34,293	33,724	34,016	34,668	34,053	34,376	35,043	34,385	34,736	35,378	34,720	35,096
west Bengal	95,688	95,879	96,339	96,297	96,669	97,191	96,906	97,464	98,040	97,516	98,267	98,887	98,125	99,076	99,731	98,604	99,892	1,00,570
Jharkhand	36,334	36,339	35,830	36,869	37,001	36,309	37,403	37,613	36,785	37,937	38,235	37,259	38,471	38,867	37,731	38,969	39,510	38,200
Odisha	44,312	44,385	44,145	44,658	44,800	44,517	45,004	45,218	44,888	45,350	45,641	45,257	45,696	46,067	45,625	45,987	46,497	45,991
Chhattisgarh	27,956	28,001	28,153	28,340	28,433	28,585	28,724	28,871	29,017	29,109	29,316	29,449	29,493	29,768	29,881	29,836	30,227	30,311
Madhya Pradesh	79,948	80,089	79,510	81,090	81,405	80,627	82,232	82,743	81,741	83,374	84,103	82,851	84,516	85,485	83,957	85,548	86,889	85,058
Gujrat	66,084	66,132	66,149	67,010	67,131	67,116	67,936	68,146	68,082	68,862	69,176	69,046	69,788	70,221	70,007	70,648	71,283	70,965
Maharashtra	1,19,869	1,20,040	1,18,800	1,21,011	1,21,370	1,19,900	1,22,153	1,22,710	1,21,000	1,23,295	1,24,070	1,22,100	1,24,437	1,25,440	1,23,190	1,25,411	1,26,830	1,24,280
Andhra Pradesh	51,655	52,833	51,193	51,938	53,396	51,444	52,221	53,965	51,694	52,504	54,540	51,945	52,787	55,121	52,195	52,972	55,708	52,446
Karnataka	64,752	64,873	64,849	65,275	65,525	65,471	65,798	66,183	66,091	66,322	66,849	66,711	66,845	67,520	64,211	67,268	68,199	67,948
Kerala	34,761	34,818	34,987	34,943	35,059	35,240	35,125	35,302	35,491	35,307	35,546	35,740	35,489	35,792	35,988	35,633	36,040	36,234
Tamil Nadu	74,989	75,151	75,779	75,342	75,664	76,408	75,695	76,181	77,035	76,049	76,700	77,659	76,402	77,224	78,281	76,631	77,751	78,900

In the logistic model, the carrying capacity is calculated for each state based on the census data 2011 and performed the MATLAB code to get the predictions. When we calculate the carrying capacity, we can observe that Tamil Nadu and Kerala are at a rate of 0.02 whereas Jammu & Kashmir has a carrying capacity of 0.03. It varies from 0.02 to 0.04 depending upon the state population of the base year.

The following figures give us a detailed picture of exponential and Logistic data mapped against the actuals. Figure 1 shows the comparison of both the models with the actuals for the years 2012 to 2015 and there we can see the data is almost matching with the actuals in both the models, out of which exponential model is closer to the actual.

On the other hand, fig 2 gives us a comparison of both the models with actuals for the years 2016 to 2019 and there is a noticeable difference in both the models from the actual data, Exponential data being the more accurate. When we observe figure 3, that is the comparison of the models with the actual for the years 2020 to 2022 and it is seen that the data varies significantly from the actual in both the cases, Exponential being the more accurate.

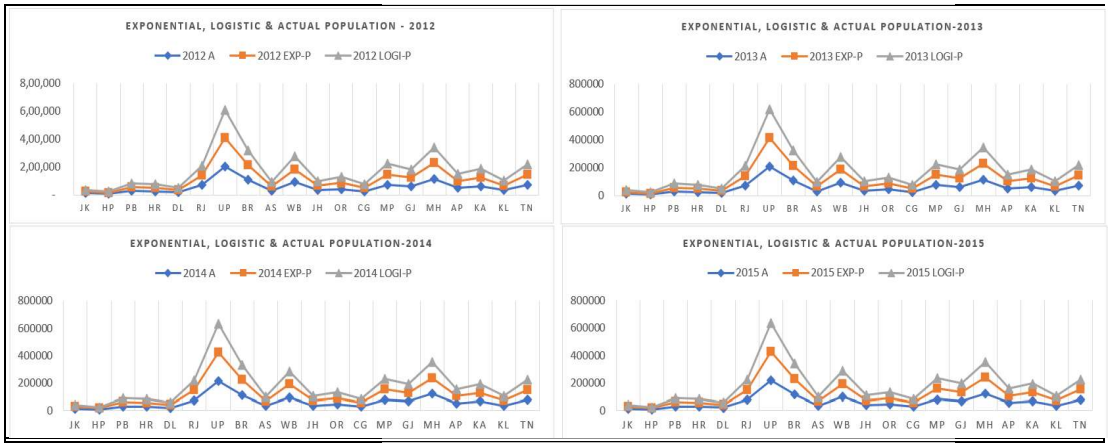


Figure 1: Comparison between Exponential, Logistic and Actual Curves – 2012 -2015

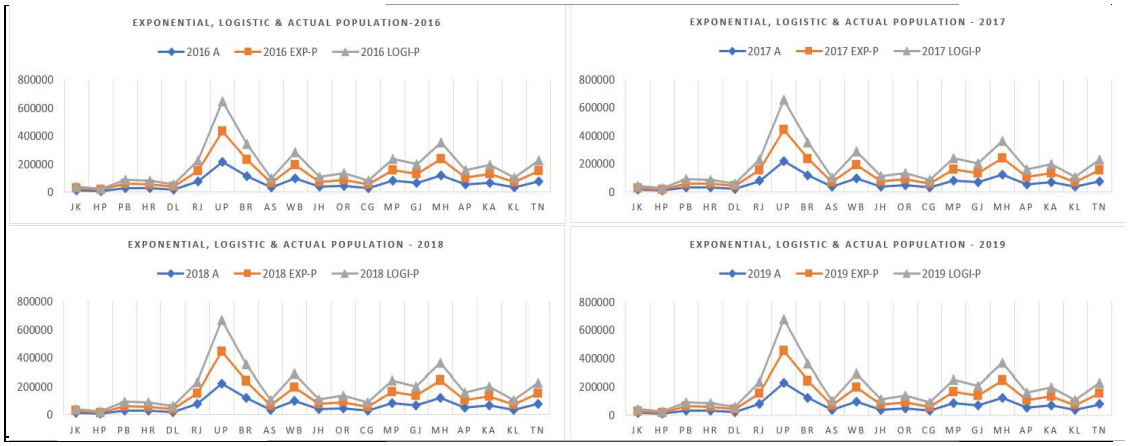


Figure 2: Comparison between Exponential, Logistic and Actual curves –2016-2019

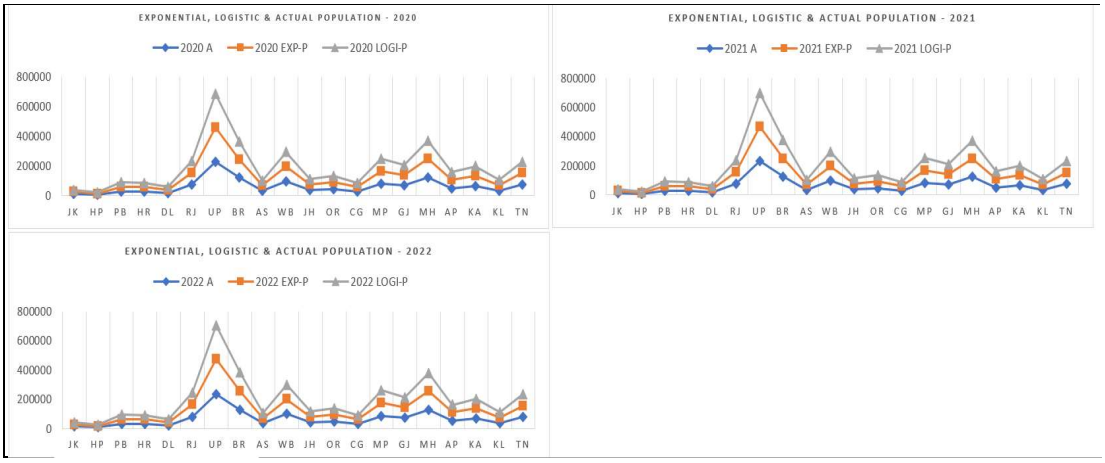


Figure 3: Comparison between Exponential, Logistic and Actual curves –2020-2022

Based on the above observations and comparative study till 2022, the population is predicted for the upcoming years 2023 to 2031 using both the models and is given below in table 3. Figure 4 and 5 give a pictorial representation of the table 3 data. It is very clear that logistic population prediction is on a higher side than the exponential population.

Table 3: Prediction of Population – Exponential vs Logistic – 2023 - 2031

Indian States	2023		2024		2025		2026		2027		2028		2029		2030		2031	
	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P	EXP-P	LOGI-P
Jammu & Kashmir	13519	14,514	13629	14,697	13740	14,879	13852	15,059	13964	15,238	14078	15,417	14192	15,593	14308	15,769	14424	15,943
Himachal Pradesh	7574.9	8,853	7637.2	8,953	7700.1	9,055	7763.5	9,159	7827.4	9,266	7891.9	9,375	7956.9	9,487	8022.4	9,602	8088.4	9,720
Punjab	31205	31,997	31512	32,329	31823	32,658	32136	32,983	32453	33,305	32772	33,623	33095	33,937	33421	34,248	33750	34,555
Haryana	30680	30,061	31172	30,441	31672	30,820	32179	31,196	32695	31,570	33219	31,941	33751	32,310	34292	32,676	34842	33,039
NCT of Delhi	21677	20,970	22144	21,301	22621	21,630	23108	21,955	23605	22,276	24113	22,593	24632	22,907	25163	23,216	25704	23,520
Rajasthan	82958	80,847	84288	81,841	85638	82,828	87011	83,809	88406	84,783	89822	85,750	91262	86,709	92725	87,660	94211	88,602
Uttar Pradesh	241240	2,36,860	245050	2,39,860	248930	2,42,840	252870	2,45,810	256870	2,48,750	260940	2,51,680	265070	2,54,580	269260	2,57,470	273520	2,60,320
Bihar	129980	1,29,430	132400	1,31,440	134870	1,33,430	137390	1,35,400	139960	1,37,340	142570	1,39,260	145230	1,41,150	147950	1,43,020	150710	1,44,860
Assam	35058	35,457	35400	35,817	35745	36,177	36093	36,538	36445	36,898	36800	37,257	37159	37,617	37521	37,976	37887	38,335
West Bangal	100710	1,01,410	101540	1,02,250	102380	1,03,080	103220	1,03,910	104070	1,04,730	104930	1,05,550	105790	1,06,370	106660	1,07,180	107540	1,07,990
Jharkhand	40163	38,666	40827	39,130	41502	39,590	42188	40,047	42886	40,500	43595	40,950	44316	41,396	45049	41,838	45794	42,275
Odisha	46932	46,356	47370	46,719	47813	47,081	48260	47,440	48711	47,798	49166	48,154	49625	48,508	50089	48,860	50557	49,210
Chhattisgarh	30693	30,741	31166	31,170	31647	31,597	32135	32,023	32630	32,447	33133	32,870	33644	33,290	34163	33,708	34690	34,124
Madhya Pradesh	88317	86,155	89769	87,246	91244	88,331	92743	89,409	94268	90,481	95817	91,545	97391	92,601	98992	93,650	100620	94,690
Gujrat	72360	71,920	73453	72,871	74564	73,818	75690	74,760	76834	75,698	77996	76,630	79174	77,556	80371	78,476	81585	79,390
Maharashtra	128230	1,25,370	129650	1,26,460	131080	1,27,540	132530	1,28,620	134000	1,29,700	135480	1,30,770	136980	1,31,840	138490	1,32,900	140030	1,33,960
Andhra Pradesh	56302	52,696	56902	52,946	57508	53,196	58121	53,446	58740	53,695	59366	53,945	59999	54,194	60638	54,443	61285	54,692
Karnataka	68884	68,564	69577	69,179	70276	69,793	70982	70,405	71696	71,015	72416	71,624	73144	72,230	73879	72,834	74622	73,437
Kerala	36290	36,479	36541	36,722	36794	36,963	37409	37,203	37305	37,440	37564	37,676	37824	37,911	38086	38,143	38349	38,373
Tamil Nadu	78281	79,516	78815	80,129	79353	80,740	79894	81,347	80440	81,951	80988	82,551	81541	83,148	82097	83,742	82658	84,332

[Exponential Predicted (EXP-P) vs Logistic Predicted (LOGI-P), Numbers in Thousands]

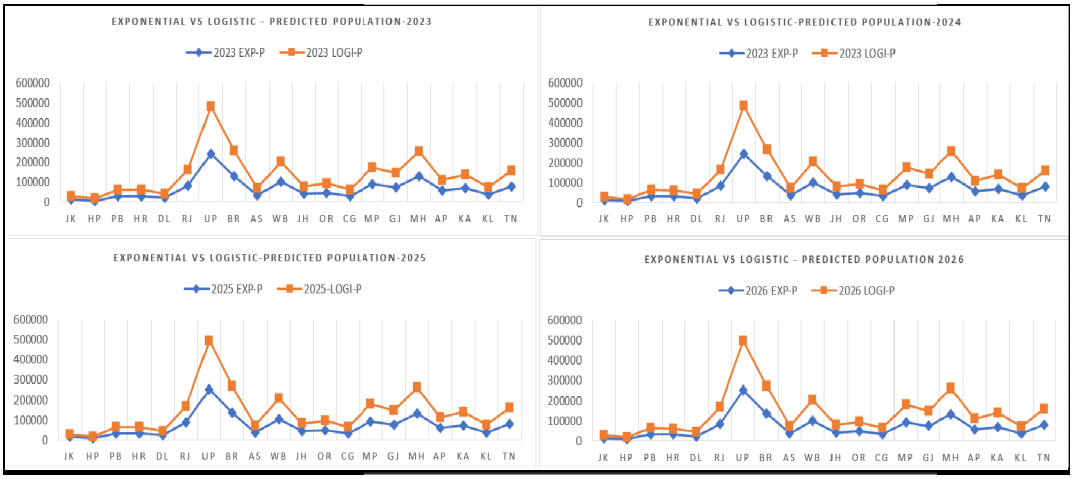


Figure 4: Prediction of Exponential and Logistic Population Curves – 2023 - 2026

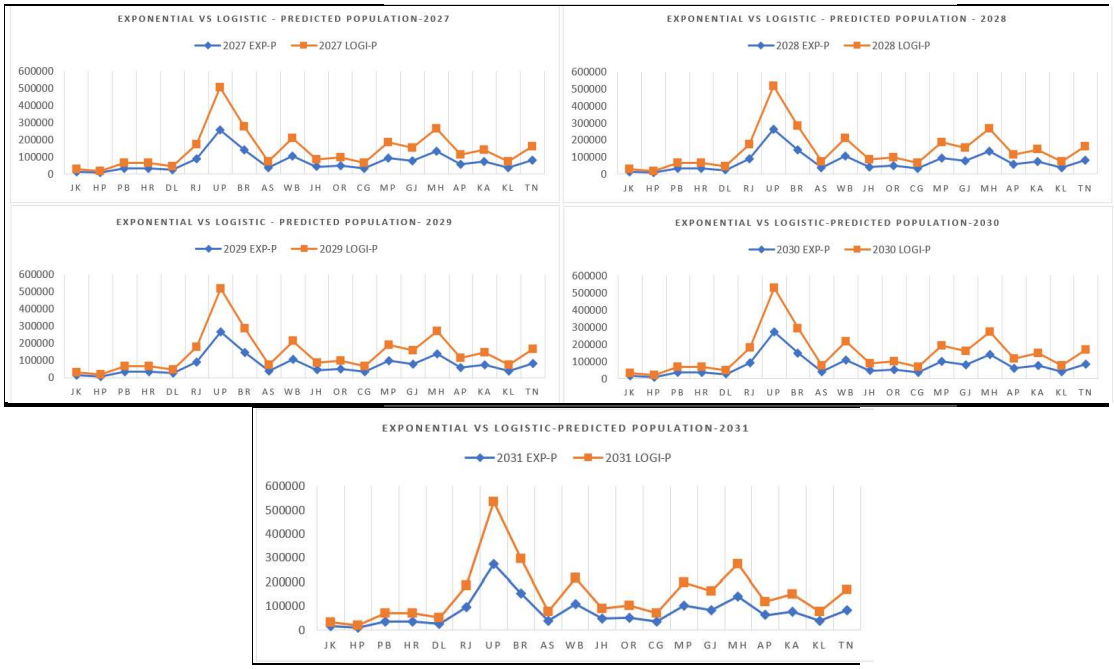


Figure 5: Prediction of Exponential and Logistic Population Curves – 2027 - 2031

Indian States Abbreviations in the Graphs

Jammu &Kashmir	Himachal Pradesh	Punjab	Haryana	Delhi	Rajasthan	Uttar Pradesh	Bihar	Assam	West Bengal
JK	HP	PB	HR	DL	RJ	UP	BR	AS	WB

Jharkhand	Orissa	Chhattisgarh	Madhya Pradesh	Gujrat	Maharashtra	Andhra Pradesh	Karnataka	Kerala	Tamil Nadu
JH	OR	CG	MP	GJ	MH	AP	KA	KL	TN

IV. STATISTICAL TESTS

1. **Chi-Square Test:** Chi-Square test of independence [21] has been conducted on the exponential and logistic models separately for each year starting from 2012 till 2022. The null hypothesis for the test - "There is no significant difference in the predicted and the actual values" and the alternate hypothesis - "There is a significant difference in the predicted and actual values". The level of significance has been taken as 0.05 and degree of freedom is calculated as (c-1) (r-1) and that is 19. The critical value for the degree of freedom 19 is found to be 30.14 from the chi-square table. The following table gives us the overview of the chi-square values calculated per year for both the models. This test shows that the predicted population is very close to the actual population from 2012 to 2016, and then from 2016 to 2022 it starts to vary significantly from the actual in the case of exponential model. Whereas, in the case of logistic model, from 2012 to 2014 it shows a match with the actual population but starts to vary apart from the actual as we move from 2015 to 2022. The logistic model shows more fluctuation from the actual as compared to the exponential model in this research study.

Table 4: Chi-Square Values – Exponential vs Logistic Models – 2012 - 2022

Year	Chi Square value (Exponential model)	Chi-square value (Logistic model)	Critical Value
2012	1.09	4.50	30.14
2013	3.97	10.62	
2014	8.37	21.14	
2015	14.36	35.77	
2016	22.27	54.57	
2017	39	79.47	
2018	69.59	112.19	
2019	116.13	149.33	
2020	182.14	198.19	
2021	269.41	350.99	
2022	437.67	317.92	

2. **Mean Absolute Percentage Error Test:** The Mean Absolute Percentage Error (MAPE) test [1] is calculated to understand the error percentage for each model. It is calculated using the formula (Actual – Predicted)/Actual * 100 for the years 2012 to 2022 in the case of both the models. The data calculated is displayed in Table 5 state wise. This table tells a fair idea about the correctness of the model. Exponential model produces less error as compared to the Logistic model.

Table 5: MAPE % Values – State Wise - Exponential vs Logistic Models

Indian States	MAPE % (Exponential)	MAPE % (Logistic)
Jammu & Kashmir	0.73	3.18
Himachal Pradesh	0.32	15.46
Punjab	0.35	1.98
Haryana	0.35	4.15
NCT of Delhi	0.37	0.72
Rajasthan	0.55	0.47
Uttar Pradesh	0.50	0.13
Bihar	0.69	1.02
Assam	1.27	0.66
west Bengal	0.38	0.81
Jharkhand	0.39	1.31
Odisha	0.33	0.28
Chhattisgarh	0.38	0.77
Madhya Pradesh	0.46	0.44
Gujrat	0.26	0.16
Maharashtra	0.33	0.79
Andhra Pradesh	2.46	0.74
Karnataka	0.40	0.59
Kerala	0.34	0.77
Tamil Nadu	0.43	1.23

V. FIBONACCI POPULATION GROWTH MODEL

In 1202, Leonardo Fibonacci researched the breeding of a pair of rabbits under specific circumstances and he recorded his observations. According to him, the predictions can be made based on the previous year's population of rabbits and it follows Fibonacci number series in Mathematics and hence the name Fibonacci Population Growth [18]. Later it is observed that the human population also can be predicted using this population model.

Let X(n) and Y(n) be the no. of adults and juveniles of a hypothetical population with the dynamics is given by the recursive equations, $X(n+1) = X(n) + Y(n)$, $Y(n+1) = X(n)$ with $X(0) = 1$ and $Y(0) = 0$. The system of equations can be written in the form, $X(n+1) = X(n) + X(n-1)$

VI. RANDOM FOREST MODEL

There are variety of Machine Learning methods that is been used in predicting the data recently. Out of that, Linear regression is the most commonly used method predicting the population growth that has been used very often in almost all research papers. Based on rigorous study of research papers and comparison of machine learning methods, it has been observed that Random Forest method gives 95% accuracy in population prediction as compared to the other methods.

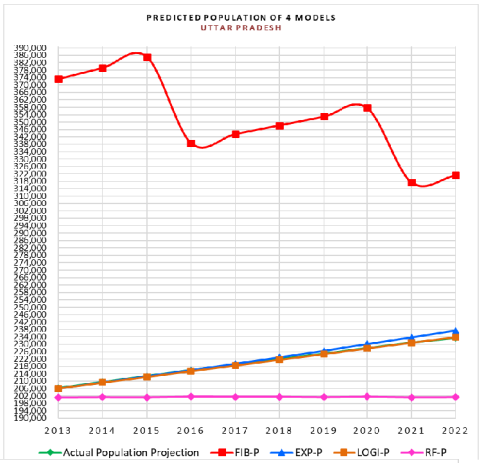
Random Forest [5] is an ensemble machine learning algorithm. Ensemble learning is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. It belongs to the supervised learning technique. Random Forest is a classifier that contains several decision trees [7] on various subsets of the given data set and takes the average to improve the predictive accuracy of the data set. The advantages of Random Forest are it takes less training time as compared to other algorithms, it predicts the output with high accuracy, and it can also maintain accuracy when a large proportion of data is missing.
https://www.javatpoint.com/machine-learning-random-forest-algorithm

VII. COMPARISON OF ALL THE MODELS

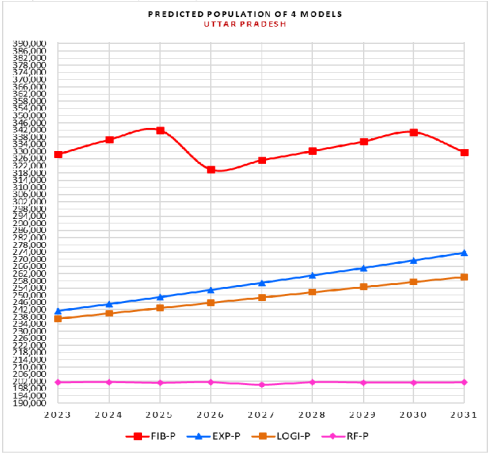
The exponential, Logistic, Fibonacci and Random Forest models have been used here to predict the human population from 2012 to 2031. Here we used mathematical methods and calculated the projections using MATLAB programming language for the first three models and Python programming language for the machine learning based Random Forest model. Tables given below shows the projections using three models and the actual projection by Govt. of India census of all the 5 states.
(https://nhm.gov.in/New_Updates_2018/Report_Population_Projection_2019.pdf)

The comparison between all four models and actual projections can be clearly seen in the below figures. It is observed that the actual projections and exponential model predictions are very close to each other and in fact in some cases they completely overlap with each other. Hence the future projections made are more accurate in exponential model than the other models.

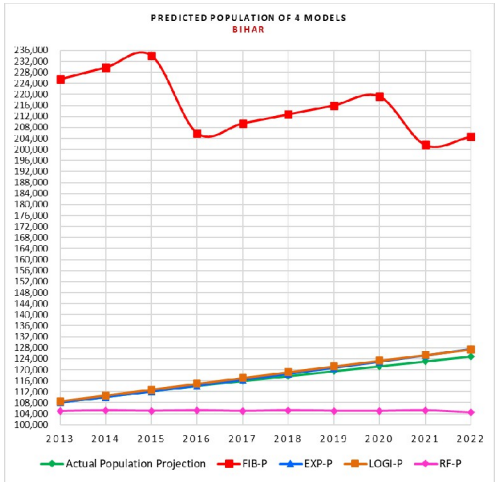
PREDICTED POPULATION OF 4 MODELS UTTAR PRADESH					
Year	Actual Population Projection	FIB-P	EXP-P	LOGI-P	RF-P
2013	206,322	373,290	206,190	206,250	201,196
2014	209,577	379,240	209,450	209,340	201,350
2015	212,832	385,190	212,760	212,430	201,291
2016	216,087	338,626	216,130	215,510	201,676
2017	219,051	343,576	219,550	218,590	201,600
2018	222,015	348,262	223,020	221,660	201,602
2019	224,979	352,947	226,550	224,720	201,375
2020	227,943	357,633	230,140	227,780	201,616
2021	230,907	317,548	233,780	230,820	201,275
2022	233,297	321,248	237,480	233,850	201,373



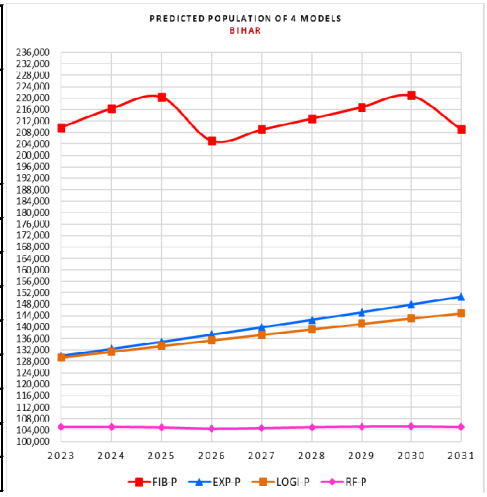
PREDICTED POPULATION OF 4 MODELS UTTAR PRADESH				
Year	FIB-P	EXP-P	LOGI-P	RF-P
2023	328,443	241,240	236,860	201,423
2024	336,543	245,050	239,860	201,585
2025	341,865	248,930	242,840	201,275
2026	320,073	252,870	245,810	201,547
2027	325,138	256,870	248,750	200,262
2028	330,286	260,940	251,680	201,519
2029	335,516	265,070	254,580	201,386
2030	340,822	269,260	257,470	201,352
2031	329,673	273,520	260,320	201,433



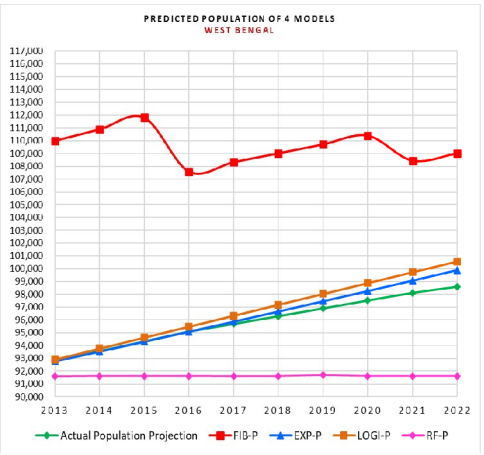
PREDICTED POPULATION OF 4 MODELS BIHAR					
Year	Actual Population Projection	FIB-P	EXP-P	LOGI-P	RF-P
2013	108,130	225,516	108,020	108,520	104,948
2014	110,145	229,776	110,040	110,650	105,231
2015	112,161	234,036	112,090	112,780	105,046
2016	114,176	205,974	114,190	114,900	105,251
2017	115,957	209,505	116,320	117,020	105,002
2018	117,739	212,756	118,490	119,120	105,248
2019	119,520	216,007	120,700	121,210	105,054
2020	121,302	219,259	122,960	123,290	105,006
2021	123,083	201,663	125,250	125,360	105,214
2022	124,919	204,643	127,590	127,400	104,476



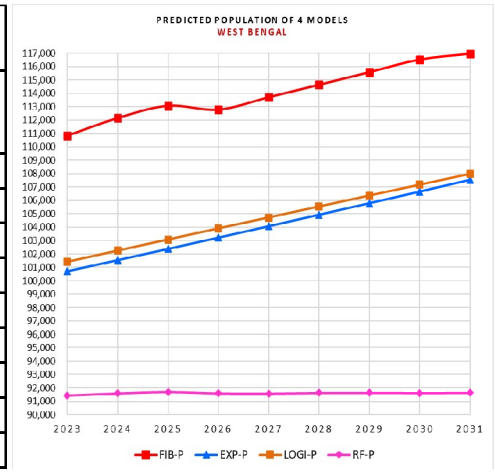
PREDICTED POPULATION OF 4 MODELS BIHAR				
Year	FIB-P	EXP-P	LOGI-P	RF-P
2023	209,806	129,980	129,430	105,054
2024	216,428	132,400	131,440	105,092
2025	220,461	134,870	133,430	104,931
2026	205,177	137,390	135,400	104,481
2027	209,012	139,960	137,340	104,678
2028	212,917	142,570	139,260	105,008
2029	216,888	145,230	141,150	105,204
2030	220,942	147,950	143,020	105,229
2031	209,119	150,710	144,860	105,035



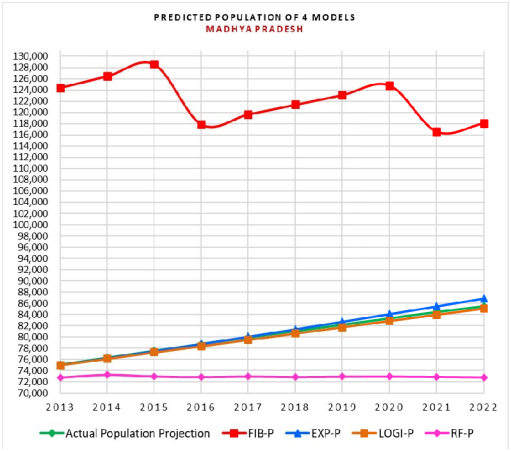
PREDICTED POPULATION OF 4 MODELS WEST BENGAL					
Year	Actual Population Projection	FIB-P	EXP-P	LOGI-P	RF-P
2013	92,797	109,987	92,785	92,914	91,597
2014	93,558	110,891	93,549	93,773	91,632
2015	94,318	111,795	94,319	94,630	91,635
2016	95,079	107,573	95,096	95,486	91,630
2017	95,688	108,320	95,879	96,339	91,615
2018	96,297	109,011	96,669	97,191	91,623
2019	96,906	109,702	97,464	98,040	91,692
2020	97,516	110,394	98,267	98,887	91,628
2021	98,125	108,441	99,076	99,731	91,630
2022	98,604	109,013	99,892	100,570	91,624



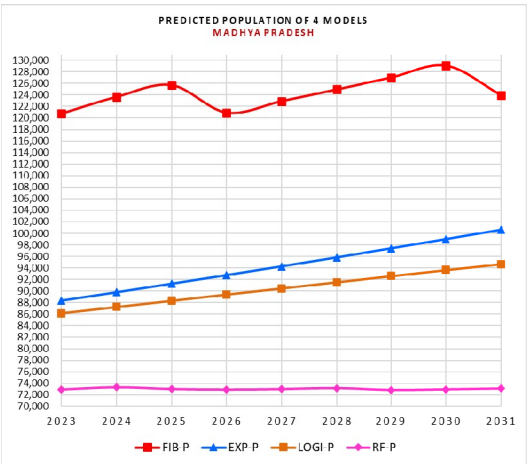
PREDICTED POPULATION OF 4 MODELS WEST BENGAL				
Year	FIB-P	EXP-P	LOGI-P	RF-P
2023	110,828	100,710	101,410	91,413
2024	112,152	101,540	102,250	91,582
2025	113,079	102,380	103,080	91,707
2026	112,787	103,220	103,910	91,584
2027	113,714	104,070	104,730	91,568
2028	114,653	104,930	105,550	91,624
2029	115,595	105,790	106,370	91,635
2030	116,545	106,660	107,180	91,617
2031	116,981	107,540	107,990	91,630



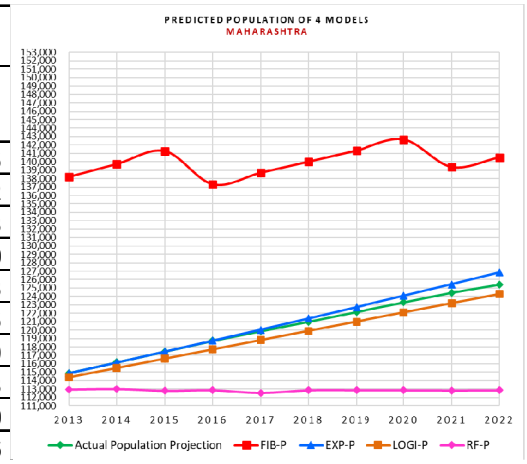
PREDICTED POPULATION OF 4 MODELS MADHYAPRADESH					
Year	Actual Population Projection	FIB-P	EXP-P	LOGI-P	RF-P
2013	75,099	124,415	75,034	75,021	72,779
2014	76,334	126,483	76,267	76,145	73,327
2015	77,570	128,550	77,520	77,269	72,988
2016	78,806	117,859	78,794	78,390	72,871
2017	79,948	119,659	80,089	79,510	72,991
2018	81,090	121,382	81,405	80,627	72,881
2019	82,232	123,104	82,743	81,741	72,967
2020	83,374	124,827	84,103	82,851	72,978
2021	84,516	116,584	85,485	83,957	72,914
2022	85,548	118,092	86,889	85,058	72,816



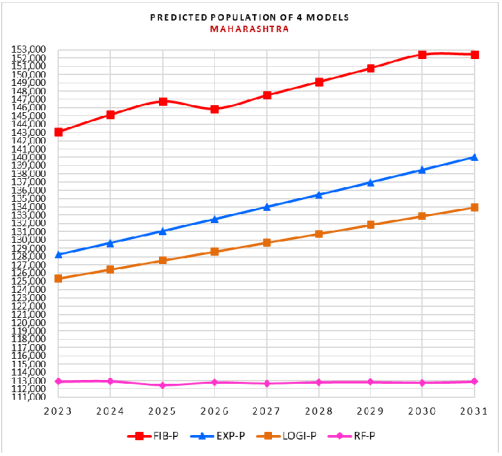
PREDICTED POPULATION OF 4 MODELS MADHYAPRADESH				
Year	FIB-P	EXP-P	LOGI-P	RF-P
2023	120,758	88,317	86,155	72,891
2024	123,668	89,769	87,246	73,308
2025	125,700	91,244	88,331	72,985
2026	120,937	92,743	89,409	72,878
2027	122,925	94,268	90,481	72,975
2028	124,946	95,817	91,545	73,161
2029	126,998	97,391	92,601	72,814
2030	129,086	98,992	93,650	72,926
2031	123,971	100,620	94,690	73,084



PREDICTED POPULATION OF 4 MODELS MAHARASHTRA					
Year	Actual Population Projection	FIB-P	EXP-P	LOGI-P	RF-P
2013	114,915	138,199	114,870	114,370	112,935
2014	116,186	139,734	116,140	115,480	112,992
2015	117,456	141,268	117,430	116,590	112,793
2016	118,727	137,309	118,730	117,690	112,860
2017	119,869	138,686	120,040	118,800	112,553
2018	121,001	140,004	121,370	119,900	112,855
2019	122,153	141,333	122,710	121,000	112,859
2020	123,295	142,663	124,070	122,100	112,853
2021	124,437	139,376	125,440	123,190	112,830
2022	125,410	140,527	126,830	124,280	112,846



PREDICTED POPULATION OF 4 MODELS MAHARASHTRA				
Year	FIB-P	EXP-P	LOGI-P	RF-P
2023	143,060	128,230	125,370	112,910
2024	145,136	129,650	126,460	112,953
2025	146,738	131,080	127,540	112,442
2026	145,872	132,530	128,620	112,796
2027	147,489	134,000	129,700	112,651
2028	149,120	135,480	130,770	112,826
2029	150,770	136,980	131,840	112,848
2030	152,434	138,490	132,900	112,758
2031	152,466	140,030	133,960	112,904



When we compare all the above models, we observe that exponential model gives a more accurate prediction as compared to the other models. The below Table shows a comparison based on the behaviour of all the models.

Sl.no	Properties	Exponential	Logistic	Fibonacci	Random Forest (ML model)
1	Model Behavior	Population will show an increasing trend	It shows increasing and decreasing trends	It depends on the previous year population	It is a random generated method and so behavior changes randomly.
2	Accuracy	Very good	Pretty good	Not so good	It is good
3	Reliability	Reliable	Reliable	Not reliable	Reliable to an extent

Based on my two studies as described above, exponential models always give a more accurate and reliable prediction and hence can be considered as the best model.

REFERENCES

[1] Brintha Rajkumari S, P. C. (2020). Prediction of Population Growth Using Machine Learning Techniques. European Journal of Molecular & Clinical Medicine, 1871-1879.
[2] Chian -Yue Wang, S. J. (2021). Regional Population Forecast and Analysis Based on Machine Learning Strategy. Entropy, 1-12.
[3] Otoom, M. M. (2021). Comparing the Performance of 17 Machine Learning Models in Predicting Human Population Growth of countries. International Journal of Computer Science and Network Security, 220-225.
[4] R Mahesh. R.Priyanga, S. (2019). Population Prediction Using MAchine Learning. Journal of Emerging Technologies and Innovative Research, 189-192.
[5] J. Brownlee, "Random Forest for Time Series Forecasting," November 2, 2020, https://machinelearningmastery.co5]m/random-forest-for-time-series-forecasting/
[6] F.V. Şahinarslan, A.T. Tekin, and F. Çebi, "Application of machine learning algorithms for population forecasting," International Journal of Data Science, vol. 6, no. 4, pp. 257–270, 2021.
[7] E. A. Rady, H. Fawzy, and A.M. Abdel Fattah, "Time Series Forecasting Using Tree Based Methods," Journal of Statistics Applications & Probability, vol.10, no. 1, pp. 229-244, 2021
[8] Dr. N. Ashioba and N. N. Daniel, "Population Forecasting System Using Machine Learning Algorithm,"

International Journal of Computer Trends and Technology, vol. 68, no. 12, pp. 40-43, December 2020.
[9] F.V. Şahinarslan, A.T. Tekin, and F. Çebi, "Application of machine learning algorithms for population forecasting," International Journal of Data Science, vol. 6, no. 4, pp. 257–270, 2021
[10] E. A. Rady, H. Fawzy, and A.M. Abdel Fattah, "Time Series Forecasting Using Tree Based Methods," Journal of Statistics Applications & Probability, vol.10, no. 1, pp. 229-244, 2021
[11] F.V. Şahinarslan, A.T. Tekin, and F. Çebi, "Application of machine learning algorithms for population forecasting," International Journal of Data Science, vol. 6, no. 4, pp. 257–270, 2021
[12] C.Y. Wang and S.J. Lee, "Regional Population Forecast and Analysis Based on Machine Learning Strategy," Entropy, vol. 23, no. 656, pp. 1-12, 2021
[13] O. Folorunso, A. Akinwale, O. Asiribo, and T. Adeyemo, "Population prediction using artificial neural network," African Journal of Mathematics and Computer Science Research, vol. 3, no. 8, pp. 155–162, 2010.
[14] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B.-Won On, W. Aslam, and G. S. Choi, "COVID-19 Future Forecasting Using Supervised Machine Learning Models," IEEE Access, vol.8, pp.101489-101499, 2020
[15] M. Sh. Hajirahimova, L. R. Yusifova, "Experimental Study of Machine Learning Methods in Anomaly Detection," Problems of Information Technology, vol. 13, no. 1, pp. 9-19, 2022.
[16] Hironmoy Mondal, U. K. (2018). Mathematical Modeling and Predicting the Current Trends of Human Population Growth in Bangladesh. Modelling, Measurement and Control D, 1 -7.
[17] Bacaer, N. (2011). Verhulst and the logistic equation (1838). London: Springer - Verlag London Limited.
[18] Sigler, L. (2002). Fibonacci's Liber Abaci. Lewisburg: Springer, Newyork.

