_____

A Prediction Study about the Pandemic Era based on Machine Learning Methods

Meltem Eryılmaz
Atılım University
Ankara
Turkey
meltem.eryilmaz@atilim.edu.tr

Önder Ertan
Atılım University
Ankara
Turkey
ertan.onder@student.atilim.edu.tr

Furkan Yalçınkaya
Atılım University
Ankara
Turkey
yalcinkaya.furkan@student.atilim.edu.tr

Ekin Kara
Atılım University
Ankara
Turkey
kara.ekin@student.atilim.edu.tr

Abstract: Coronavirus pandemic has been going on since late 2019, millions of people died worldwide, vaccination has recently started in many countries and new strategies are sought by countries since they are still struggling to defeat the virus. So, this research is made to predict the possible ending time of the coronavirus pandemic in Turkey using data mining and statistical studies. Data mining is a computer science study that processes large amounts of data and produces new useful information. It is especially used to support decision making in companies today. So, this project could support the decision making of authorities in developing an effective strategy against the on-going pandemic. During the research we have practiced on Turkey's coronavirus and vaccination data between 13 January 2021 and 28 May 2021. We used Rapidminer and the Random Forest method for data mining. After all the simulations we have applied and observed during our research, it was clearly seen that vaccination parameters were decreasing the new cases. Also, the stringency index did not affect the new cases. As a conclusion of our research and observations, we think that the government should vaccinate as many people as it can in order to relax restrictions for the last time.

Keywords: Covid-19, data mining, rapid miner, random forest

1. Introduction

New coronavirus disease 2019 (COVID-19) is a type of virus that can be transmitted by respiratory tract and contact. Firstly, COVID-19 appeared in a group of patients who developed respiratory tract symptoms such as fever, coughing, difficulty in breathing in Wuhan, China in late December 2019. After studies that were carried on this group of patients, the new virus, COVID-19, was identified [4, 21].

As it is known, the world has been struggling with the COVID-19 virus for a year. Currently, Turkey has been trying to carefully go through the normalization process. Last time the government went for the partial normalization in late May 2020 [1], unfortunately because of a variety of disregarded reasons, the country went back to where it first started. Number of cases, the positivity rate and deaths soared up. After weeks of healthcare providers' and authorities' battling with the virus and a-17-days-long full lockdown [7], the country managed to maintain control of the pandemic. At the moment, according to the officially given information, cases and deaths are declining. For the present the country is normalizing again after readjusting the pandemic values, the future of Turkey's fight against coronavirus is not certainly known. Medical experts are afraid of living through the previous vicious cycle again.

We think that due to the lack of processed and analyzed data about the virus, today we are struggling to develop a full scale and efficient strategy against it. Therefore, we have decided to do this research on COVID-19 and vaccination data using data mining and statistical studies. At this point giving a brief information about data mining

_____

would be great for the readers. Data mining is the process of sorting, analyzing and generating useful information from large data sets which are retrieved from databases or companies. Data mining has been increasingly used since the 20th century, due to it being productive and helpful in various types of industrial sectors. Thus, as is seen, the goal of this research is to make a prediction for the possible ending time of the pandemic especially based on the number of cases and vaccination progress in Turkey. In the following parts of the paper we are going to explain how we performed our research step by step.

2. Material and Methodology

2.1 Material

As the first step of our study we had to collect data to work on. Daily data have been officially announced by the Ministry of Health in Turkey. But in order to get a more detailed and large amount of data we needed datasets. So we preferred to retrieve data from a repository that belonged to "Our World in Data" (OWID) which is an online publication website that focuses on problems and crises of the world by storing, analyzing and interpreting data. OWID stored its data on GitHub. [3] The GitHub repository includes the COVID-19 data in a spreadsheet which is continuously updated. Stringency index was calculated and added to the excel file by OWID however it is important to mention that the calculation is made by using values that are known as indicators and related to government pandemic policies [13].

2.2 Methodology

In this part of the article you will be given information about the stages we went through.
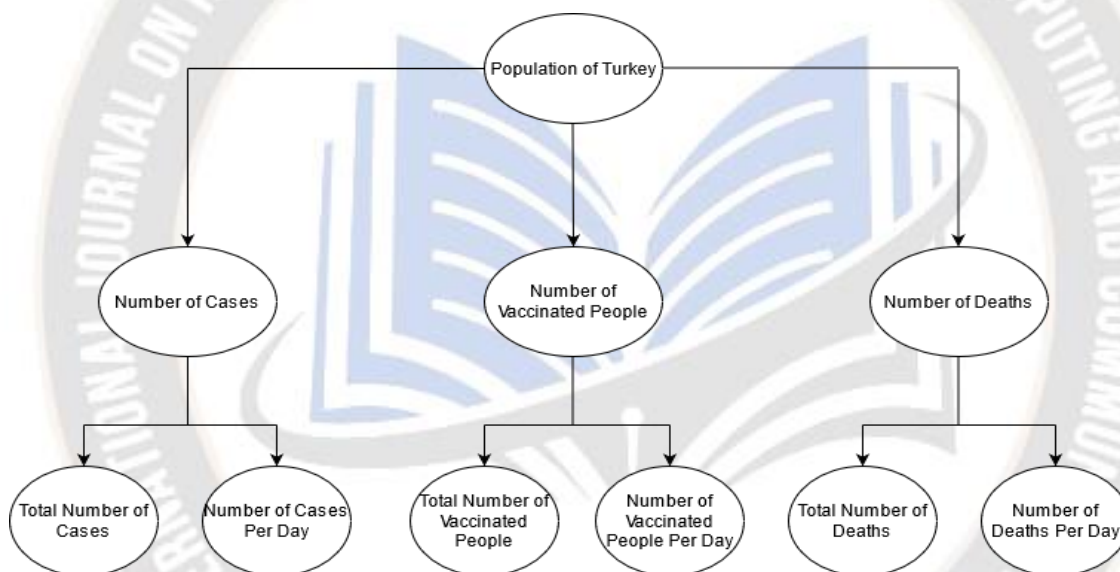
2.2.1 Conceptual Modelling

Figure 1. Schematic conceptual modelling of data warehouse

The process of developing a data model for data that will be stored in a database is known as data modelling. On the other hand conceptual modelling defines what the system includes. The goal of conceptual modelling is to organize, scope and define business concepts and rules.

Here in Figure 1, we designed the schematic conceptual modelling of the data warehouse by dividing the population of Turkey, which can be seen as a universal set, into three different branches that are "Number of Cases", "Number of Vaccinated People", "Number of Deaths". Each branch also includes two sub branches of "Total Number" and "Number … Per Day". The model covers concepts of our project and focuses on representing data to be seen as "real world" data.

2.2.2 Preprocessing

Data preprocessing is one of the crucial stages of data mining. In preprocessing, data are literally filtered by cleaning the dataset from nonessential and noisy data. The OWID excel file included other countries' COVID-19 information, since we did not need them. Since our study focuses only on Turkey and the course of the pandemic as part of our research, we applied the "Extract, Transform, Load" (ETL) [16] process by cleaning unnecessary cells and removing all the other countries and unneeded parts in the dataset. Also in order to make sure the data was correct we have compared the OWID dataset with official publications of the Turkish Ministry of Health, as a result of which we have verified the correctness [17].

date	total_cases	new_cases	total_deaths	new_deaths	reproduction_rate	new_tests	total_tests	positive_rate	tests_per_case	total_vaccinations	people_vaccinated	people_fully_vaccinated	new_vaccinations	stringency_index
2020-03-11	1	1												23,15
2020-03-12	1	0												23,15
2020-03-13	5	4												23,15
2020-03-14	5	0												23,15
2020-03-15	6	1												23,15
2020-03-16	18	12												52,78
2020-03-17	47	29	1	1										52,78
2020-03-18	98	51	1	0										58,33
2020-03-19	192	94	3	2										58,33
2020-03-20	359	167	4	1	5,13									58,33
2020-03-21	670	311	9	5	5,11									62,04
2020-03-22	1236	566	30	21	4,75									62,04
2020-03-23	1529	293	37	7	4,04									62,04
2020-03-24	1872	343	44	7	3,64									67,59
2020-03-25	2433	561	59	15	3,6									67,59
2020-03-26	3629	1196	75	16	3,67									67,59
2020-03-27	5698	2069	92	17	3,52	7533	47823							70,37
2020-03-28	7402	1704	108	16	3,05	7641	55464							75,93
2020-03-29	9217	1815	131	23	2,69	9982	65446							75,93
2020-03-30	10827	1610	168	37	2,54	11535	76981							75,93
2020-03-31	13531	2704	214	46	2,39	15422	92403							75,93
2020-04-01	15679	2148	277	63	2,12	14396	106799							75,93
2020-04-02	18135	2456	356	79	1,88	18757	125556							75,93
2020-04-03	20921	2786	425	69	1,73	16160	141716	0,162	6,2					75,93
2020-04-04	23934	3013	501	76	1,69	19664	161380	0,156	6,4					75,93
2020-04-05	27069	3135	574	73	1,7	20065	181445	0,154	6,5					75,93
2020-04-06	30217	3148	649	75	1,76	21400	202845	0,154	6,5					75,93
2020-04-07	34109	3892	725	76	1,74	20023	222868	0,158	6,3					75,93
2020-04-08	38226	4117	812	87	1,66	24900	247768	0,16	6,3					75,93
2020-04-09	42282	4056	908	96	1,52	28578	276346	0,16	6,2					75,93
2020-04-10	47029	4747	1006	98	1,44	30864	307210	0,158	6,3					75,93
2020-04-11	52167	5138	1101	95	1,41	33170	340380	0,158	6,3					77,78
2020-04-12	56956	4789	1198	97	1,37	35720	376100	0,154	6,5					77,78
2020-04-13	61049	4093	1296	98	1,35	34456	410556	0,148	6,7					75,93
2020-04-14	65111	4062	1403	107	1,28	33070	443626	0,14	7,1					75,93
2020-04-15	69392	4281	1518	115	1,21	34090	477716	0,136	7,4					75,93
2020-04-16	74193	4801	1643	125	1,12	40427	518143	0,132	7,6					75,93
2020-04-17	78546	4353	1769	126	1,03	40270	558413	0,125	8					75,93
2020-04-18	82329	3783	1890	121	1	40520	598933	0,117	8,6					77,78
2020-04-19	86306	3977	2017	127	1,03	35344	634277	0,114	8,8					77,78
2020-04-20	90980	4674	2140	123	1,08	39703	673980	0,114	8,8					75,93
2020-04-21	95591	4611	2259	119	1,03	39429	713409	0,113	8,9					75,93
2020-04-22	98674	3083	2376	117	0,91	37535	750944	0,107	9,3					75,93
2020-04-23	101790	3116	2491	115	0,81	40962	791906	0,101	9,9					75,93
2020-04-24	104912	3122	2600	109	0,75	38351	830257	0,097	10,3					75,93
2020-04-25	107773	2861	2706	106	0,73	38308	868565	0,094	10,6					75,93
2020-04-26	110130	2357	2805	99	0,74	30177	898742	0,09	11,1					75,93

Figure 2. Data set of COVID-19 in Turkey

Hence our final version of excel file had the following attributes left: date, total_cases, new_cases, total_deaths, new_deaths, reproduction_rate, new_tests, total_tests, positive_rate, tests_per_case, total_vaccinations, people_vaccinated, people_fully_vaccinated, new_vaccinations, stringency_index.
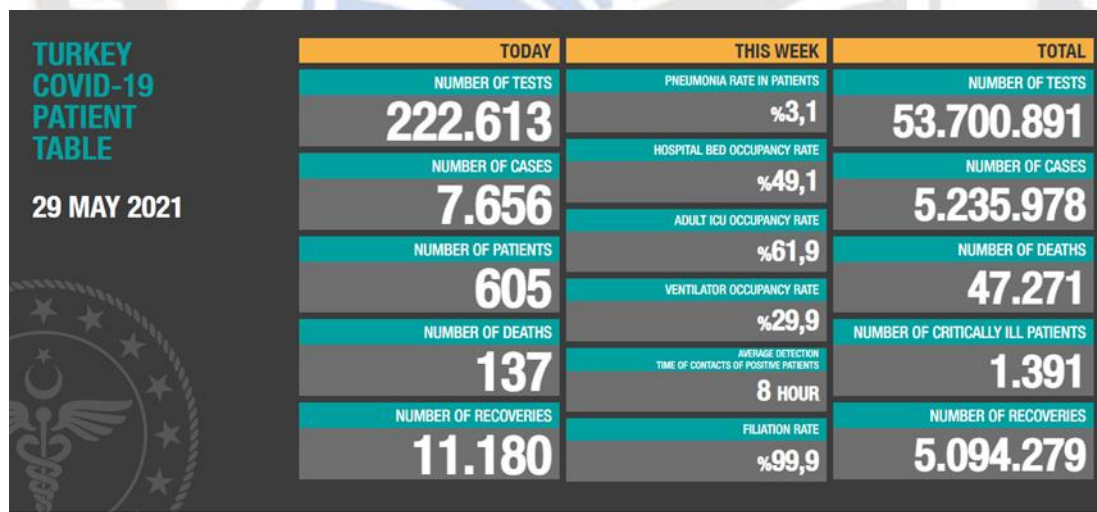

Figure 3. Turkey Covid-19 Patient Table for 29 May 2021

Thus in preprocessing we used the following operations to simplify our dataset:

Data Cleaning: In our data set, we cleaned the attributes and data that we did not need as well as duplicate data such as constant number of population. On the other hand, there was not any incomplete, noisy, inconsistent or intentional data.

Data Reduction: In this technique we got rid of other countries and their data since we focus only on Turkey and we limited our data with the data as of the first time vaccination started in Turkey which is 13 January 2021. Then we focused on the total number of the attributes.

Data Transformation and Data Discretization: Since there was not any incomplete, noisy, inconsistent or intentional data as we mentioned before, we have not transformed our data. There was no need for it. But there were missing data in early vaccination.

	date	total_cases	total_deaths	reproduction_rate	total_tests	positive_rate	tests_per_case	total_vaccinations	people_vaccinated	people_fully_vaccinated	stringency_index
2	2021-03-01	2711479	28638	1,18	33305552	0,075	13,4	8774718	6975558	1799160	66,67
3	2021-03-02	2723316	28706	1,2	33440843	0,077	13	9073640	7123722	1949918	64,81
4	2021-03-03	2734836	28771	1,19	33578861	0,078	12,8	9317546	7240515	2077031	64,81
5	2021-03-04	2746158	28839	1,19	33716746	0,079	12,7	9560038	7353536	2206502	64,81
6	2021-03-05	2757460	28901	1,19	33854960	0,08	12,5	9756289	7446337	2309952	64,81
7	2021-03-06	2769230	28965	1,21	33993552	0,082	12,2	9883607	7509923	2373684	64,81
8	2021-03-07	2780417	29030	1,22	34122851	0,083	12	9917348	7524943	2392405	64,81
9	2021-03-08	2793632	29094	1,23	34262280	0,086	11,6	10001980	7565023	2436957	64,81
10	2021-03-09	2807387	29160	1,22	34403108	0,087	11,4	10222994	7676519	2546475	72,22
11	2021-03-10	2821943	29227	1,22	34548238	0,09	11,1	10360283	7745480	2614803	72,22
12	2021-03-11	2835989	29290	1,21	34694624	0,092	10,9	10502494	7807838	2694656	72,22
13	2021-03-12	2850930	29356	1,22	34843480	0,095	10,6	10742277	7858565	2883712	72,22
14	2021-03-13	2866012	29421	1,23	34993578	0,097	10,3	10923284	7887544	3035740	72,22
15	2021-03-14	2879390	29489	1,23	35126003	0,099	10,1	10992095	7892795	3099300	72,22
16	2021-03-15	2894893	29552	1,24	35277116	0,1	10	11261612	7911946	3349666	72,22
17	2021-03-16	2911642	29623	1,25	35435502	0,101	9,9	11696296	7948320	3747976	72,22
18	2021-03-17	2930554	29696	1,28	35603028	0,103	9,7	12142245	7977222	4165023	72,22
19	2021-03-18	2950603	29777	1,31	35787480	0,105	9,5	12431482	7991510	4439972	72,22
20	2021-03-19	2971633	29864	1,33	35975339	0,107	9,4	12844577	8007734	4836843	72,22
21	2021-03-20	2992694	29959	1,34	36165468	0,108	9,3	13029754	8016078	5013676	72,22
22	2021-03-21	3013122	30061	1,34	36355374	0,109	9,2	13057722	8017698	5040024	72,22
23	2021-03-22	3035338	30178	1,35	36556589	0,11	9,1	13265952	8045789	5220163	72,22
24	2021-03-23	3061520	30316	1,37	36768437	0,112	8,9	13589167	8091685	5497482	72,22
25	2021-03-24	3091282	30462	1,38	36990175	0,116	8,6	13969867	8133043	5836884	72,22
26	2021-03-25	3120013	30619	1,36	37212928	0,119	8,4	14227410	8163156	6064254	72,22
27	2021-03-26	3149094	30772	1,35	37433913	0,122	8,2	14491247	8195554	6295693	72,22
28	2021-03-27	3179115	30923	1,35	37657127	0,125	8	14652659	8217124	6435535	72,22
29	2021-03-28	3208173	31076	1,36	37877236	0,128	7,8				72,22
30	2021-03-29	3240577	31230	1,36	38102747	0,133	7,5	14949260	8370872	6578388	72,22
31	2021-03-30	3277880	31385	1,36	38338045	0,138	7,3	15301462	8614368	6687094	72,22
32	2021-03-31	3317182	31537	1,36	38578057	0,142	7	15834787	8994454	6840333	72,22
33	2021-04-01	3357988	31713	1,36	38821795	0,148	6,8	16148683	9206500	6942183	72,22
34	2021-04-02	3400296	31892	1,36	39070763	0,153	6,5	16317170	9313746	7003424	72,22
35	2021-04-03	3445052	32078	1,36	39319889	0,16	6,3	16602919	9478896	7124023	72,22
36	2021-04-04	3487050	32263	1,34	39566099	0,165	6,1	16683067	9546643	7136424	72,22
37	2021-04-05	3529601	32456	1,32	39811595	0,169	5,9	16993574	9796506	7197068	72,22
38	2021-04-06	3579185	32667	1,32	40083142	0,173	5,8	17367875	10088238	7279637	
39	2021-04-07	3633925	32943	1,32	40385250	0,175	5,7	17704888	10359390	7345498	
40	2021-04-08	3689866	33201		40689742	0,178	5,6	18089783	10627037	7462746	
41	2021-04-09	3745657	33454					18320384	10777451	7542933	

Figure 4. Simplified dataset of Turkey for last month

2.2.3 Choosing Data Mining Method

After processing and preparing our data in a spreadsheet file, we used the Random Forest method to model and analyze our data in Rapidminer which is an advanced data mining tool that comes with many widgets and tools to make a prediction [20]. Random Forest is a popular machine learning model because it produces decent results without the need for hyperparameter estimation and can be used to solve both regression and classification issues [15]. We tried to use the traditional Decision Tree method too however, since one of the biggest problems of this method is excessive learning-data overfitting [18]. Also after using Random Forest we realised that our results were more realistic and logical, in addition this method performed the best among the others.
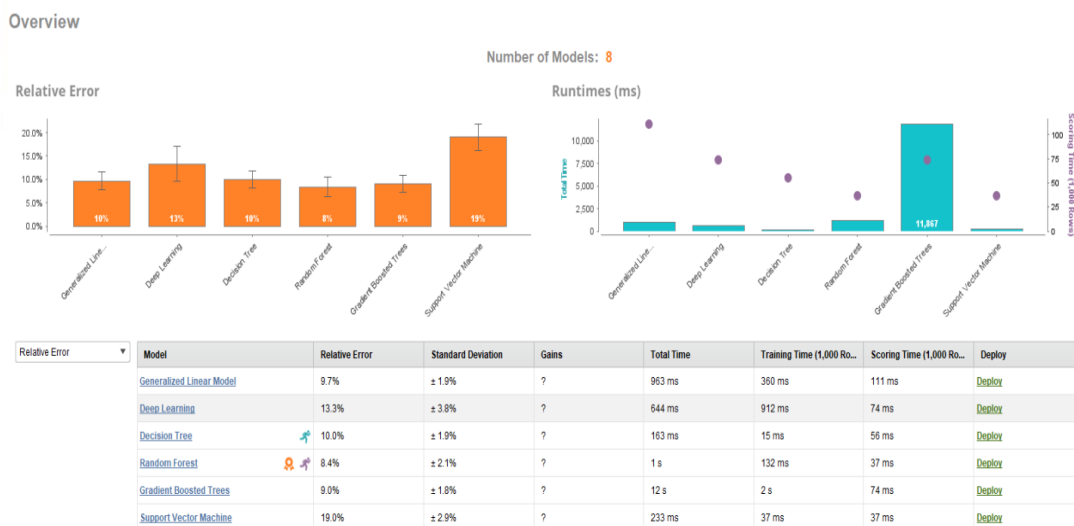

Figure 5. Relative error results overview of models

3. Results

In this part of the research we will be talking about the prediction stage that we have done using Rapidminer. We have encountered 8.4% of relative error and 2.1% of standard deviation in Random Forest results. Rapidminer took more time for analyzing the data in the Random Forest method compared to other methods in modelling. Green and tones of it support the prediction, while red and tones of it contradict the prediction of Rapidminer (as it can be seen in Figure 8). The data mining tool completed empty spaces in the input data table (such as data related to vaccination which was partially missing until 11 February) using its deep learning and prediction methods, but these assumptions of the program did not affect the results.

Random Forest - Predictions

Row No. ↑	new_cases	prediction(new_cases)	total_cases	total_deaths	new_deaths	reproductio...	new_tests	total_tests	positive_rate	total_vaccinations	people_vaccinated	people_fully_vaccinated	new_vaccinati...	stringency_i
27	18912	14658.868	2930554	29696	73	1.280	167526	35603028	0.103	12142245	7977222	4165023	445949	72.220
28	20049	19453.949	2950603	29777	81	1.310	184452	35787480	0.105	12431482	7991510	4439972	289237	72.220
29	21030	20062.869	2971633	29864	87	1.330	187859	35975339	0.107	12844577	8007734	4836843	413095	72.220
30	21061	21129.726	2992694	29959	95	1.340	190129	36165468	0.108	13029754	8016078	5013676	185177	72.220
31	26182	23610.527	3061520	30316	138	1.370	211848	36768437	0.112	13589167	8091685	5497482	323215	72.220
32	37303	30842.995	3277880	31385	155	1.360	235298	38338045	0.138	15301462	8614368	6687094	352202	72.220
33	39302	35403.372	3317182	31537	152	1.360	240012	38578057	0.142	15834787	8994454	6840333	533325	72.220
34	42308	40630.661	3400296	31892	179	1.360	248968	39070763	0.153	16317170	9313746	7003424	168487	72.220
35	42551	42272.112	3529601	32456	193	1.320	245496	39811595	0.169	16993574	9796506	7197068	310507	72.220
36	54562	51577.776	3903573	34182	243	1.210	301068	41892922	0.180	18728978	11108847	7620131	234182	72.220
37	59187	57485.885	3962760	34455	273	1.190	306563	42199485	0.181	19062758	11388406	7674352	333780	83.330
38	61028	56449.248	4384624	36613	346	1.020	322128	44409756	0.191	20283227	12426580	7856647	207723	83.330
39	61967	55777.475	4446591	36975	362	0.980	318839	44728595	0.190	20477685	12590582	7887103	194458	83.330
40	43301	39800.816	4710582	39057	346	0.790	282192	46435343	0.161	21678461	13383037	8295424	283676	87.040
41	40444	39044.952	4751026	39398	341	0.760	283261	46718604	0.153	22132803	13512810	8619993	454342	87.040
42	31891	32971.190	4820591	40131	394	0.720	265287	47261999	0.141	22816891	13712254	9104637	248719	87.040
43	28997	26134.660	4929118	41527	336	0.670	241747	48229733	0.122	23729113	14132161	9596952	434636	87.040
44	20107	23769.714	4998089	42465	278	0.620	251194	48963800	0.104	24691044	14476881	10214163	223325	87.040
45	18052	19519.177	5016141	42746	281	0.600	228873	49192673	0.099	24853353	14545524	10307829	162309	87.040
46	15191	14572.831	5031332	43029	283	0.580	197468	49390141	0.095	24918773	14585980	10332793	65420	87.040

Figure 6. Random Forest Prediction

Rapidminer used new_cases, total_cases, total_deaths, reproduction_rate, new_tests, total_tests, positive_rate, total_vaccinations, people_vaccinated, people_fully_vaccinated, new_vaccinated, stringency_index and made predictions for new_cases. people_vaccinated represents the people who are vaccinated with the first dose, people_fully_vaccinated represents people who are vaccinated with the second dose as well. As it can be seen in Figure 6 new cases are predicted to be growing.
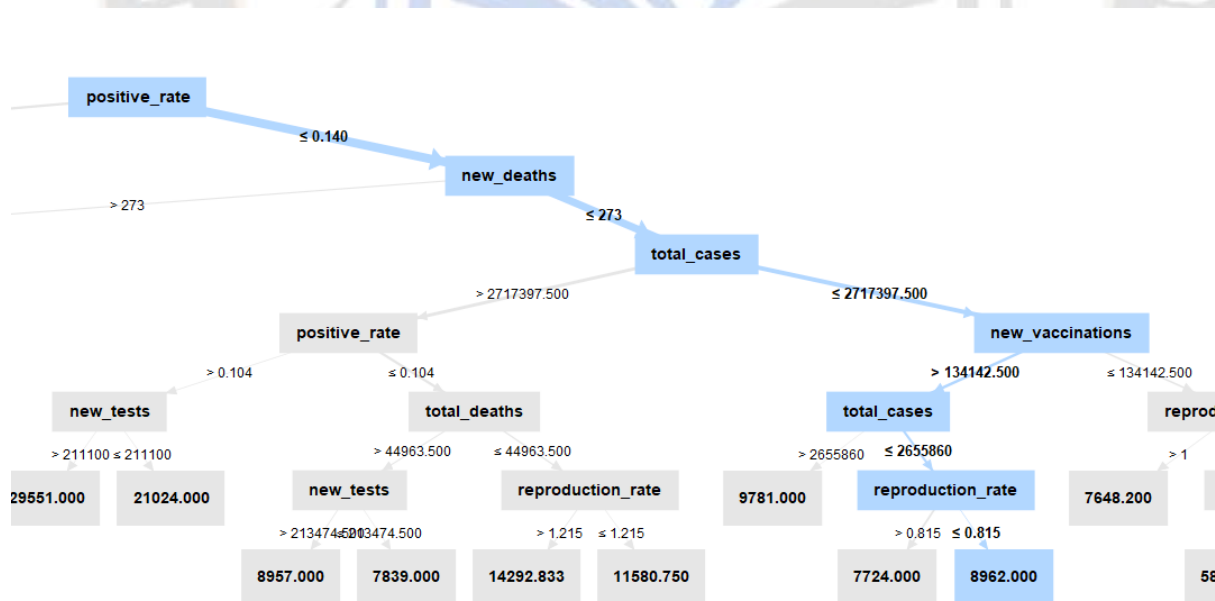

Figure 7. Random Forest Tree

Optimal Parameters for Random Forest modelling was also calculated by Rapidminer. As maximal depth decreases and number of trees increase, error rate increases too. As it can be seen in Figure 8 the tool chose the parameters with least error rate which had 60 trees and maximal depth of 7.

_____

Random Forest - Optimal Parameters

Optimal Parameters
Number Of Trees: **60**
Maximal Depth: **7**

Error Rates for Parameters

Number of Trees	Maximal Depth	Error Rate
140	4	12.6%
20	7	11.0%
60	7	10.6%
100	7	10.9%

Figure 8. *Optimal Parameters*

According to our experiments in the simulator, increasing the stringency index when the number of cases is high does not help the course of Covid-19 pandemic. As we increased the positive_rate in the simulation, the new cases proportionally increased as well. Also if new_tests and total_tests are increased new_cases are increased in proportion to these. On the other hand when we simulated vaccination data, this decreased the new cases a little bit even though this has a big importance in our observation. But with the combination of vaccination data and some other related data, new cases decreased proportionally.

Random Forest - Predictions

Row No. ↑	new_cases	prediction(new_cases)	total_cases	total_deaths	new_deaths	reproduction_rate	new_tests	total_tests	positive_rate	total_vaccinations	people_vaccinated	people_fully_vaccinated	new_vaccinat
35	42551	42272.112	3529601	32456	193	1.320	245496	39811595	0.169	16993574	9796506	7197068	310507
36	54562	51577.776	3903573	34182	243	1.210	301068	41892922	0.180	18728978	11108847	7620131	234182
37	59187	57485.885	3962760	34455	273	1.190	306563	42199485	0.181	19062758	11388406	7674352	333780
38	61028	56449.248	4384624	36613	346	1.020	322128	44409756	0.191	20283227	12426580	7856647	207723
39	61967	55777.475	4446591	36975	362	0.980	318839	44728595	0.190	20477685	12590582	7887103	194458
40	43301	39800.816	4710582	39057	346	0.790	282192	46435343	0.161	21678461	13383037	8295424	283676
41	40444	39044.952	4751026	39398	341	0.760	283261	46718604	0.153	22132803	13512810	8619993	454342
42	31891	32971.190	4820591	40131	394	0.720	265287	47261999	0.141	22816891	13712254	9104637	248719
43	28997	26134.660	4929118	41527	336	0.670	241747	48229733	0.122	23729113	14132161	9596952	434636
44	20107	23769.714	4998089	42465	278	0.620	251194	48963800	0.104	24691044	14476881	10214163	223325
45	18052	19519.177	5016141	42746	281	0.600	228873	49192673	0.099	24853353	14545524	10307829	162309
46	15191	14572.831	5031332	43029	283	0.580	197468	49390141	0.095	24918773	14585980	10332793	65420
47	13604	14346.353	5044936	43311	282	0.560	213863	49604004	0.090	25018078	14618167	10399911	99305
48	10512	11361.450	5117374	44760	223	0.620	202243	50869896	0.058	25772743	14947740	10825003	85316
49	10174	11164.779	5127548	44983	223	0.640	206836	51076732	0.056	25950934	14986982	10963952	178191
50	11553	10086.042	5151038	45419	233	0.650	218683	51512691	0.054	26821460	15366071	11455389	284297
51	7523	9156.524	5194010	46446	178	0.700	216655	52594333	0.044	28161235	16115588	12045647	112508
52	9375	10445.326	5203385	46621	175	0.710	223104	52817437	0.042	28397847	16255095	12142752	236612
53	8426	9102.584	5220549	46970	183	0.730	219465	53258116	0.039	28605453	16377100	12228353	94120
54	7773	12008.341	5228322	47134	164	0.999	202041.646	38405250.110	0.099	28730177	16452864	12277313	124724

Figure 9. *Random Forest Prediction*

4. Conclusion

To conclude our research we can say that related to the results and state of worldwide pandemic, parallel to the speed of the vaccination process the pandemic can be taken under control. Even if the people get infected when they are vaccinated they can survive it. The vaccination speed should be increased in Turkey in order to take the pandemic under control and slow down the rise of new cases and deaths in the short term. As we have experienced in the simulations, the stringency index did not

_____

affect the rate of cases or the numbers vitally. So the stringency index which represents the physical measures, especially lockdowns, is not the strategy we should be applying when coping with the pandemic. Also, while lockdowns are pointless, they affect the economy and social life negatively [11]. In fact, although measures are taken most of the time, the ability of flu-like viruses to spread rapidly, the insufficient amount of vaccines and, as the most important reason, the weakness of the health systems in countries where the disease continues, cause these pandemics to carry on. However, if we take cognizance of countries who vaccinated most of their population such as Israel and USA, the restrictions are reduced and people are returning to their old normal lives [10, 19]. Therefore, as mentioned the only weapon we have against the virus is vaccines right now and if we manage to vaccinate most of Turkey, we may have a one way ticket to normalization. If Turkey succeeds in vaccinating most of its population, considering the possibility of vaccinated people being less mortal against variants of the virus, as the number of cases and the existence of infection decrease, we expect that the pandemic will end completely by the end of 2022 in Turkey.

References

[1] "Prof. Serap Şimşek Yavuz: Salgın yönetimi Bilim Kurulu'nda olsaydı bunları yaşamazdık" *Gazete Duvar*. [Online]. Available: https://www.gazeteduvar.com.tr/prof-serap-simsek-yavuz-salgin-yonetimi-bilim-kurulunda-olsaydi-bunlari-yasamazdik-makale-1518082

[2] F. BUDAK and Ş. KORKMAZ, "COVID-19 PANDEMİ SÜRECİNE YÖNELİK GENEL BİR DEĞERLENDİRME: TÜRKİYE ÖRNEĞİ," *Sosyal Araştırmalar ve Yönetim Dergisi*, no. 1, pp. 62–79, 2020. https://dergipark.org.tr/tr/download/article-file/1107760

[3] Owid, "owid/covid-19-data," *GitHub*. [Online]. Available: https://github.com/owid/covid-19-data

[4] N. Rezaei, *Coronavirus Disease - COVID-19.* Cham, Switzerland: Springer International Publishing, 2021.

[5] P. G. Auwaerter, "Coronavirus COVID-19 (SARS-CoV-2): Johns Hopkins ABX Guide," *Coronavirus COVID-19 (SARS-CoV-2) | Johns Hopkins ABX Guide*. [Online]. Available: https://www.hopkinsguides.com/hopkins/view/Johns_Hopkins_ABX_Guide/540747/all/Coronavirus_COVID_19__SARS_CoV_2_.

[6] H. Ritchie, E. Ortiz-Ospina, D. Beltekian, E. Mathieu, J. Hasell, B. Macdonald, C. Giattino, C. Appel, L. Rodés-Guirao, and M. Roser, "Coronavirus Pandemic (COVID-19) - Statistics and Research," *Our World in Data*, 05-Mar-2020. [Online]. Available: https://ourworldindata.org/coronavirus

[7] T.C. İçişleri Bakanlığı Bilgi İşlem Dairesi Başkanlığı, "81 İl Valiliğine Tam Kapanma Tedbirleri Genelgesi Gönderildi," *T.C. İçişleri Bakanlığı*, 26-Apr-2021.

[Online]. Available: https://www.icisleri.gov.tr/81-il-valiligine-tam-kapanma-tedbirleri-genelgesi-gonderildi

[8] Mathieu, E., Ritchie, H., Ortiz-Ospina, E. *et al.* A global database of COVID-19 vaccinations. *Nat Hum Behav* (2021). https://doi.org/10.1038/s41562-021-01122-8

[9] Hasell, J., Mathieu, E., Beltekian, D. *et al.* A cross-country database of COVID-19 testing. *Sci Data* 7, 345 (2020). https://doi.org/10.1038/s41597-020-00688-8

[10] E. Leshem and A. Wilder-Smith, "COVID-19 vaccine impact in Israel and a way out of the pandemic," *The Lancet*, vol. 397, no. 10287, pp. 1783–1785, 2021. https://doi.org/10.1016/S0140-6736(21)01018-7

[11] Kochhar, A. S., Bhasin, R., Kochhar, G. K., Dadlani, H., Mehta, V. V., Kaur, R., & Bhasin, C. K. (2020). Lockdown of 1.3 billion people in India during Covid-19 pandemic: A survey of its impact on mental health. *Asian journal of psychiatry*, 54, 102213. https://doi.org/10.1016/j.ajp.2020.102213

[12] D. Witteveen and E. Velthorst, "Economic hardship and mental health complaints during COVID-19," *Proceedings of the National Academy of Sciences*, vol. 117, no. 44, pp. 27277–27284, 2020. https://doi.org/10.1073/pnas.2009609117

[13] Oxford University, *Calculation and presentation of the Stringency Index 4.0*, pp. 1–3, Apr. 2020. https://www.bsg.ox.ac.uk/sites/default/files/Calculation%20and%20presentation%20of%20the%20Stringency%20Index.pdf

[14] H. N. Nawazish and V. K. Shukla, "Integrating "Random Forest" with Indexing and Query Processing for Personalized Search," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 280-283, doi: 10.1109/ICRITO48877.2020.9197888.

[15] Y. M. C. Zhang,Ensemble Machine Learning, Springer, New York, NY, USA, 2012

[16] P. Vassiliadis and A. Simitsis, "Extraction, Transformation, and Loading," *Encyclopedia of Database Systems*, pp. 1095–1101, 2009. https://doi.org/10.1007/978-0-387-39940-9_158

[17] *T.C. Sağlık Bakanlığı*. [Online]. Available: https://covid19asi.saglik.gov.tr/

[18] Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, 27(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

[19] Bagcchi S. (2021). The world's largest COVID-19 vaccination campaign. *The Lancet. Infectious diseases*, 21(3), 323. https://doi.org/10.1016/S1473-3099(21)00081-5

[20] M. Karimi, "Data Analysis using RapidMiner," 2017.

[21] T.C. Sağlık Bakanlığı, "COVID-19 Nedir ?," *COVID-19 Nedir?* [Online]. Available: https://covid19.saglik.gov.tr/TR-66300/covid-19-nedir-.html