Fuzzy Rule-Based System for Predicting Daily Case in COVID-19 Outbreak

Pınar Cihan
Dept. of Computer Engineering
Tekirdağ Namık Kemal University
Tekirdağ, Turkey
pkaya@nku.edu.tr

***Abstract*—The Covid-19 outbreak appeared in Wuhan in December 2019 and spread rapidly all over the world. The Covid-19 disease does not yet have a clinically proven vaccine and drug for treatment. The most important physical factors in reducing the spread of the epidemic are washing hands, reducing social distance and using a mask. Today in addition to clinical studies, computer-aided studies are also widely carried out for Covid-19 outbreak. Artificial intelligence methods are successfully applied in epidemic studies. In this study, fuzzy rule basing system (FRBS) used to predict the number of Covid-19 daily cases. As a result of the study, the number of daily cases was successfully estimated with FRBS (R^2 = 0.96, MAE = 186 and RMSE = 254).**

Keywords—Covid-19, artificial intelligence, fuzzy rule basing system, prediction, daily case

I. Introduction

The Covid-19 virus first appeared in Wuhan, China, in December 2019, passing from bats to humans. The virus spread all over the world in a very short time and turned into a pandemic and collapsed the health systems of many countries. Covid-19 is a severe acute respiratory syndrome and disease progression can have fatal consequences. The most obvious symptoms of the disease; It is known as fever, dry cough, sore throat, headache, weakness, muscle pain, diarrhea, and shortness of breath. In more advanced cases, it causes severe pneumonia, causing inflammation in the lungs due to oxygen difference and multiple organ failure. Especially this disease has much more dangerous effects for those with chronic diseases, weak resistance or immune system, smokers and the elderly [1].

Artificial intelligence methods are used successfully in the solution of many problems. In the COVID-19 epidemic, many studies have been carried out using artificial intelligence methods. Some of those are; diagnosis using radiology images [2-4], tracking disease [5], estimating the patient's health outcome [6, 7], early detection and diagnosis of infection [8], drug discovery [9], monitoring treatment [10, 11].

Fuzzy Rule-Based Systems (FRBS) are classified as computational intelligence and these methods based on fuzzy concepts. Methods aim at representing knowledge in a set of fuzzy rules. FRBS is proposed for solving complex real-world problems [12]. The performance of the FRBS depends on membership functions and its rule base. Wang and Mendel's (WM) fuzzy rule learning method proposed by Li-Xin Wang and Jerry M. Mendel in 1992 [13]. WM combines linguistic and numerical information for use in FRBS.

In this paper, the WM method has been implemented to forecast the number of verified cases in the Covid-19 outbreak. Dataset used in the study contains 27 March to 28 July 2020 between the dates announced by the ministry of health Turkey verified data. In the used dataset, total number of intensive care patients, total number of intubated patients, number of daily tests, number of daily recovered patients, number of daily mortality were used as input variables, and the number of cases was estimated. To evaluate the prediction performance of the FRBS method, R-square, mean absolute error (MAE) and root mean square error (RMSE) statistical criteria were used.

II. Material and Method

The used dataset contains verified values that have been announced by the Turkish republic health ministry. The dataset consists of samples for 124 days between 27 March to 28 July 2020 and includes 6 variables [14]. These; total number of intensive care patients (IC), the total number of intubated patients (I), number of daily tests, number of daily recovered patients (R), number of daily death (D) are input variables and the number of cases is the output variable. The data containing these variables were announced regularly from 27 March to 28 July 2020. Therefore, samples belonging to 124 days were used in the study.

By processing these real data with the Wang-Mendel Neuro-Fuzzy Rule Inference Method, a model that estimates the number of daily cases was produced. After the model was trained with samples of 86 days (around 70% of the dataset), the number of cases for 38 days (around 30% of the dataset) was estimated.

Modeling is the defined of input-output relationships of a system in mathematical terms. It is difficult to model physical systems precisely with a mathematical formula or equation. This is because real-world data is often non-linear, random or complex. Therefore, approximate modeling is more suitable for real-life applications. Zadeh was the first to point out that not everything in the real world should conform to certain patterns in mathematics. According to Zadeh, there are no clear lines in real life and everything should not be confined with clear patterns such as 1 and 0 (0 means not belonging to the cluster and 1 means being a member of that cluster). According to Zadeh, given certain membership degrees, a little bit of everything, a certain amount of control can be achieved, and this is a more appropriate approach to real life [12].

The modeling in which confidence intervals and fuzzy membership functions are used in approximate estimation is called fuzzy modeling [15]. Modeling with fuzzy logic consists of three stages: Fuzzifier, rule extraction, and defuzzifier (Fig. 1) [16].

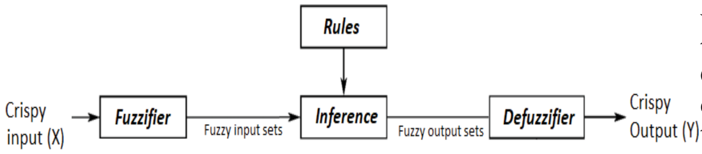

Fig. 1. Basic structure of fuzzy inference system

Fuzzifier: The process of converting crispy input values into fuzzy values. Conversion to fuzzy values is demonstrated by membership functions. Membership function is the function that defines the status of an element belonging to a set. The inputs and outputs used in fuzzy modeling can be defined with standard membership functions such as triangle, trapezoid, and gaussian curve, as well as created with different membership functions.

Rule extraction: Fuzzy inference stage consists of associating the fuzzy set with the other, using membership functions and "if-then" rules. Rules allow controlling the system by logically creating the input-output relationship. Includes a large number of If-then rules. Fuzzy rule bases have proven to be effective tools for modeling complex systems and approximating functions. Mamdani, Takagi-Sugeno, Wang-Mendel type inference mechanisms are frequently used in the literature. In this study, the Wang-Mendel neuro-fuzzy rule inference mechanism was used while creating the rule base.

Can efficiently generate fuzzy rules from sample data using the Wang-Mendel (WM) method proposed by Wang and Mendel [13, 17]. The method divides the input space into fuzzy regions. It then draws rules using a lookup table for each fuzzy subspace. The WM method is simple, convenient and does not require constant repetitive learning. Because of these advantages, the WM method has become a classic method for fuzzy rule generation and has been widely used [18].

Defuzzifier: It is also known as the rounding method. In defuzzifier, fuzzy to crispy conversion is performed. That is, the fuzzy output value is transformed into a crispy output value. There are more than 30 defuzzifier methods such as the Largest Membership Principle Method, the Center of Gravity Method, and the Weighted Average Method. For example, the Largest Membership Policy Method considers the value with the highest membership as the output value.

In this study, to analyze the estimation accuracy, mean absolute error (MAE), root mean square error (RMSE) and R-squared (R^2) statistical metric was used. These statistical measures are defined as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y}| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y})^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum(y_t - \hat{y})^2}{\sum(y_t - \bar{y})^2} \quad (3)$$

Where N is the sample size, $\hat{y}$ is the predicted value of y and $\bar{y}$ is the mean value of y.

III. Results

The daily number of cases is shown in Figure 1. Between 27 March and 28 July 2020, the maximum number of cases is 5138 and the minimum number of cases is 786. After the weekend curfews in April, the number of cases was decreased. Later, with the effect of weather and remote work in workplaces, the spread of the virus slowed.

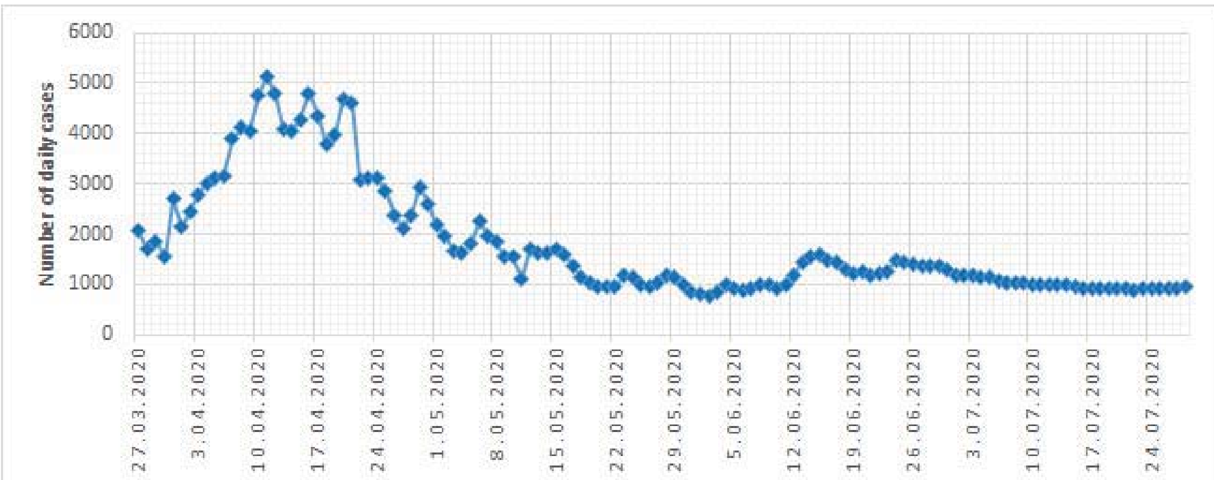

Fig. 2. Distribution of covid-19 daily cases

The number of cases per day was estimated with FRBS, and prediction results are given in Table 1. Also, prediction performance obtained was shown with the Taylor diagram (Fig. 3).

TABLE I. Prediction Performance Results

R^2	*RMSE*	*MAE*
0.96	254	186

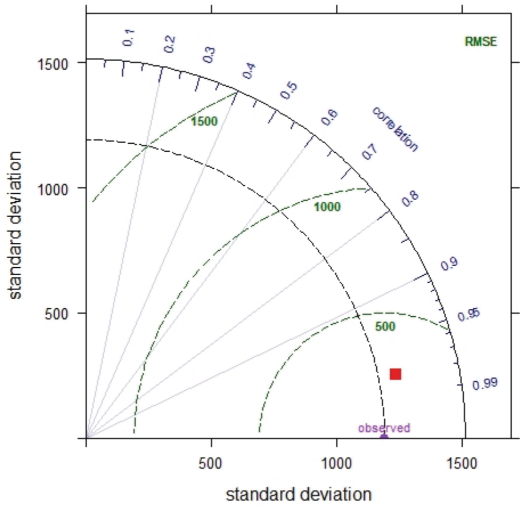

Fig. 3. Prediction results of Covid-19 daily cases

When the performance criteria are examined, it is seen that the model predictions were successful with RMSE = 254 and R = 0.98 of the number of cases with the FRBS method in Table 1 are given.

Fig. 4 shows the number of cases estimated by the model, the actual number of cases and the prediction error of the method. It is seen that the actual and predicted values are close and the errors are low.

As in previous epidemic studies [19, 20], it is seen that artificial intelligence methods can be applied successfully in the Covid-19 outbreak.

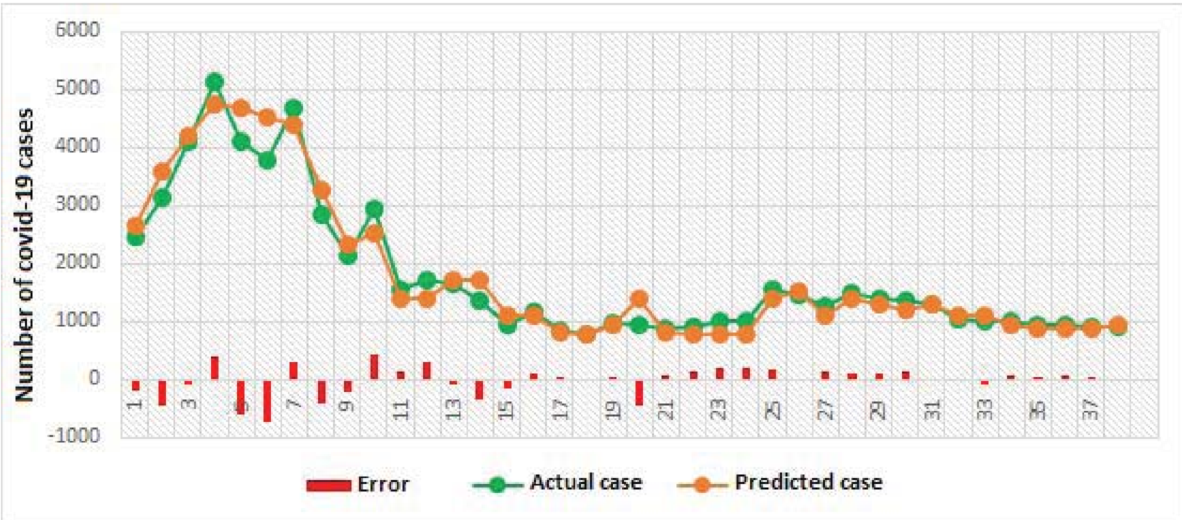

Fig. 4. Actual vs predicted daily case and model error

IV. Conclusion

The Covid-19 outbreak has affected more than 210 countries as of today. The rapid increase in the number of cases caused healthcare industries to collapse. Countries have faced major problems such as medicine, personnel, and hospital capacity. Each country was taken on measures to reduce the spread of the virus. In addition, countries are making predictions for the future situation of the Covid-19 outbreak by using artificial intelligence methods. Thus, the burden of the health sector can be reduced by foreseeing future situations and making strategies and plans. In this study, the number of daily cases in Turkey were estimated by FRBS method. As a result of the study, it was seen that the number of daily cases could be estimated successfully (R^2 = 0.96, RMSE = 254, MAE = 186).

References

[1] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine, p. 103792, 2020.

[2] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, et al., "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)," MedRxiv, 2020.

[3] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," arXiv preprint arXiv:2003.10849, 2020.

[4] L. Wang and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," arXiv preprint arXiv:2003.09871, 2020.

[5] Y. Wang, M. Hu, Q. Li, X.-P. Zhang, G. Zhai, and N. Yao, "Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner," arXiv preprint arXiv:2002.05534, 2020.

[6] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, et al., "A machine learning-based model for survival prediction in patients with severe COVID-19 infection," MedRxiv, 2020.

[7] X. Qi, Z. Jiang, Q. Yu, C. Shao, H. Zhang, H. Yue, et al., "Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: A multicenter study," medRxiv, 2020.

[8] H. Luo, Q.-l. Tang, Y.-x. Shang, S.-b. Liang, M. Yang, N. Robinson, et al., "Can Chinese medicine be used for prevention of corona virus disease 2019 (COVID-19)? A review of historical classics, research evidence and current prevention programs," Chinese journal of integrative medicine, pp. 1-8, 2020.

[9] A. Zhavoronkov, V. Aladinskiy, A. Zhebrak, B. Zagribelnyy, V. Terentiev, D. S. Bezrukov, et al., "Potential COVID-2019 3C-like protease inhibitors designed using generative deep learning approaches," Insilico Medicine Hong Kong Ltd A, vol. 307, p. E1, 2020.

[10] K. Biswas and P. Sen, "Space-time dependence of corona virus (COVID-19) outbreak," arXiv preprint arXiv:2003.03149, 2020.

[11] J. Stebbing, A. Phelan, I. Griffin, C. Tucker, O. Oechsle, D. Smith, et al., "COVID-19: combining antiviral and anti-inflammatory treatments," The Lancet Infectious Diseases, vol. 20, pp. 400-402, 2020.

[12] L. A. Zadeh, "Information and control," Fuzzy sets, vol. 8, pp. 338-353, 1965.

[13] L.-X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," IEEE Transactions on systems, man, and cybernetics, vol. 22, pp. 1414-1427, 1992.

[14] W. H. Organization. (30.09.2020). WHO Coronavirus Disease (COVID-19) Dashboard. Available: https://covid19.who.int/table?tableChartType=heat

[15] G. Chen, T. T. Pham, and N. Boustany, "Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems," Appl. Mech. Rev., vol. 54, pp. B102-B103, 2001.

[16] J. Yen and R. Langari, Fuzzy logic: intelligence, control, and information vol. 1: Prentice Hall Upper Saddle River, NJ, 1999.

[17] L.-X. Wang, "The WM method completed: a flexible fuzzy system approach to data mining," IEEE Transactions on fuzzy systems, vol. 11, pp. 768-782, 2003.

[18] Y.-F. Wang, D. Wang, and T.-Y. Chai, "Extraction of fuzzy rules with completeness and robustness," 2010.

[19] S. S. Chenar and Z. Deng, "Development of genetic programming-based model for predicting oyster norovirus outbreak risks," Water research, vol. 128, pp. 20-37, 2018.

[20] R. Liang, Y. Lu, X. Qu, Q. Su, C. Li, S. Xia, et al., "Prediction for global African swine fever outbreaks based on a combination of random forest algorithms and meteorological data," Transboundary and Emerging Diseases, vol. 67, pp. 935-946, 2020.