Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi
Afyon Kocatepe University – Journal of Science and Engineering
https://dergipark.org.tr/tr/pub/akufemubid

Araştırma Makalesi / Research Article
e-ISSN: 2149-3367
DOI: https://doi.org/10.35414/akufemubid.1550027
AKÜ FEMÜBİD **25** (2025) 035102 (510-521)
AKU J. Sci. Eng. **25** (2025) 035102 (510-521)

Analysis of the Impact of Lifestyle Habits on the Spread of COVID-19 Using Artificial Intelligence

Makale Bilgisi / Article Info
Alındı/Received: 14.09.2024
Kabul/Accepted: 09.12.2024
Yayımlandı/Published: 10.06.2025

COVID-19'un Yayılma Sürecinde Yaşam Alışkanlıklarının Etkisinin Yapay Zeka ile Analizi

İrem Sena TEKİN, Erkan ÖZHAN*

Tekirdağ Namık Kemal University, Çorlu Faculty of Engineering, Department of Computer Engineering, Tekirdağ, Türkiye

Öz

Pandemiler, toplumların sosyal ve ekonomik yapıları üzerinde önemli etkiler yaratan olaylardır ve bu süreçlerin yayılmasını etkileyen faktörlerin belirlenmesi, krizlerin yönetimi için önemli bilgiler sunar. Bu çalışmada, Türkiye'de COVID-19 pandemisi sırasında enfekte olan ve olmayan bireylerin davranışsal özellikleri ve alışkanlıkları yapay zeka teknikleri, özellikle sınıflandırma ve birliktelik kuralı metodolojileri kullanılarak analiz edilmiştir. Bulgular, Random Committee algoritmasının başarılı bir performans sergilediğini ortaya koymuştur. Ayrıca, öznitelik indirgeme uygulanmış ve RRF algoritmasının %81 doğruluk oranı ve 0.3663 kappa değeri ile daha yüksek performans gösterdiği gözlemlenmiştir. Birliktelik kuralları analizi sonucunda, "Evet" sınıfı (COVID-19 enfekte olan) için 21 kural, "Hayır" sınıfı için ise 2805 kural tespit edilmiştir. Sonuçlar, yalnız yaşayan, günde 1-3 yakın temas yaşayan ve ev dışında 4-6 saat yakın temasta bulunan bireylerin "Evet" sınıfında güçlü birliktelikler sergilediğini göstermektedir. "Hayır" sınıfında ise sık sık seyahat etmekten kaçınan, pandeminin başından beri önemli bir kilo değişikliği yaşamayan, evden çalışan ve bazen açık alanlarda küçük sosyal etkinliklerden kaçınan bireylerin güçlü birliktelikler gösterdiği ortaya çıkmıştır. Sonuç olarak, bu çalışma yaşam tarzı alışkanlıklarının pandemilerin yayılımı üzerindeki etkisini bilimsel bulgularla göstermiş ve bu faktörlerin yapay zeka teknikleri kullanılarak modellenebileceğini ortaya koymuştur.

Anahtar Kelimeler: *Pandemi; Yaşam Alışkanlıkları; Yapay Zeka; Makine Öğrenmesi; Birliktelik Kuralları; Sınıflandırma.*

Abstract

Pandemics are events that significantly impact the social and economic structures of societies, and identifying the factors influencing their spread provides valuable insights for managing these crises. This study analyzed the behavioral characteristics and habits of individuals infected and not infected during the COVID-19 pandemic in Türkiye using artificial intelligence techniques, specifically classification and association rule methodologies. The findings revealed that the Random Committee algorithm performed effectively. Additionally, feature reduction was applied, and the RRF algorithm achieved higher performance with 81% accuracy and a kappa value of 0.3663. In the subsequent analysis of association rules, 21 rules were identified for the "Yes" class (infected with COVID-19), while 2805 rules were found for the "No" class. The results indicated that individuals who live alone, have 1-3 close contacts per day, and spend 4-6 hours in close contact outside the home exhibited strong associations in the "Yes" class. For the "No" class, individuals who frequently avoid travel, have had no significant weight change since the start of the pandemic, work from home, and sometimes avoid small social gatherings in open spaces showed strong associations. In conclusion, the study scientifically demonstrated that lifestyle habits impact the transmission and spread of pandemics and that these factors can be modeled using artificial intelligence techniques.

Keywords: *Pandemic; Life Habits; Artificial Intelligence; Machine Learning; Association Rules; Classification.*

1. Introduction

The coronavirus is a contagious virus that causes respiratory infections and can spread from person to person. It was first identified in the Wuhan region of China in early December 2019. The COVID-19 pandemic has become a significant public health crisis, severely impacting human health, well-being, and freedom of movement, while also negatively affecting the global economy (Tandan et al., 2021). As of April 30, 2022, a total of 510,270,667 cases and 6,233,526 deaths have been recorded worldwide (Asia, 2022). Scientists from around the world are working intensively and competing to develop vaccines and treatment methods.

One example of numerous studies is the Recovery trial, which demonstrated the benefits of dexamethasone use in hospitalized COVID-19 patients requiring respiratory support (The RECOVERY Collaborative Group, 2021). Similarly, significant progress has been made in the development of vaccines. Scientists are testing 119 vaccines in clinical trials. Vaccines such as Pfizer-BioNTech and Sinovac have demonstrated efficacy against COVID-19, leading to emergency use authorizations in the United

*Sorumlu Yazar / Corresponding Author: Erkan ÖZHAN e-posta / e-mail: erkanozhan@gmail.com

States, Canada, China, and many other countries (Zimmer et al., 2020). However, the safety profile of these vaccines in specific groups such as the elderly and those with chronic conditions remains unclear. Moreover, there is uncertainty regarding whether vaccine manufacturers will be able to meet the global demand and when the entire population will be vaccinated and fully protected against COVID-19.

According to a report from the Center for Infectious Disease Research and Policy, achieving herd immunity-where at least 60-70% of the global population is immune-is essential for ending the COVID-19 pandemic (Moore et al., 2020). For these reasons, it is clear that implementing appropriate health measures, such as testing individuals with symptoms similar to COVID-19, quarantining them, and identifying candidates for hospital care, is crucial. These measures are critical for controlling and managing COVID-19. Numerous studies have emerged that identify the clinical characteristics and determinants of the unprecedented increase in global COVID-19 cases (Älgå et al., 2020; Epsi et al., 2024; Mallah et al., 2021; Sahu et al., 2021; Xie et al., 2023). Cough 59.6%, fever 46.9%, fatigue 27.8% and dyspnea 20.23% were the most common clinical symptoms (Israfil et al., 2021). (Kartsonaki et al., 2023) found age, co-morbidities, smoking and obesity increased the risk of dying from SARS-CoV-2. (Tadie et al., 2024) stated that severe COVID-19 cases are affected by factors such as age, gender, vaccination, knowledge, diabetes, hypertension and insufficient knowledge; advanced age, diabetes and hypertension are important determinants. However, there are few modeling studies related to COVID-19 that address the relationships among various disease determinants. (Sabherwal et al., 2024) used various mathematical methods to model the spread of COVID-19, attempting to simulate the movement of the virus within the population, disease symptoms and spread. (Zheng et al., 2020) developed a hybrid AI model for China's COVID-19 outbreak prediction, enhancing accuracy with pandemi propagation data and natural language processing modules.

Today, advancements in data processing capabilities, structured data extraction, and data mining have made it possible to perform analyses using multiple data-related techniques, such as data classification, clustering, and similarity analysis, to establish relationships among different determinants (Adamo, 2001). In addition to the efforts made in developing COVID-19 vaccines, it is believed that using machine learning and artificial intelligence techniques can also accelerate the process (Libbrecht & Noble, 2015). Although many studies have been conducted on COVID-19, most approaches have focused on more complex methods for predictions, such as deep neural networks, RNNs, CNNs, and LSTMs (Alakus & Turkoglu, 2020). In recent years, machine learning techniques have been widely used in biomedical studies for prediction and knowledge discovery (Tarca et al., 2007).

This study examines the impact of lifestyle habits in Türkiye on the likelihood of contracting the virus responsible for the COVID-19 pandemic. The first section of the study reviews previous research in this area, followed by a description of the study's methodology. After providing information about the methodology, the experimental design section outlines the methods used in the study and the evaluation metrics, and the analysis results are discussed. Finally, the experimental results section presents and interprets the findings from classification and association rule analysis.

1.1 Literature Review

This section provides an overview of significant studies and current approaches in the literature for the automatic detection of COVID-19. In their study on the impact of COVID-19 and the Dutch 'intelligent lockdown' on daily mobility, de Haas, Faber, and Hamersma (2020) highlight significant changes in work, travel, and outdoor activities. (De Haas et al., 2020) approximately 80% of people reduced their outdoor activities, with older individuals showing the largest decline. Remote work increased substantially, with 39% of employees working almost entirely from home, and 27% expecting to continue this practice after the pandemic. Travel behavior also shifted, with a 55% reduction in trips and a 68% decrease in distance traveled. The study suggests that some of these changes, especially regarding remote work and active modes of transportation like walking and cycling, may become long-term trends (De Haas et al., 2020). Muralidharan et al. (2022) propose an approach for automatic COVID-19 detection using chest X-rays. They decompose a single X-ray image into seven modes, which are classified by a multi-scale deep CNN into normal, pneumonia, or COVID-19. Using two public datasets (1,225 and 9,000 images), the model achieved 96% and 100% accuracy for multi-class and binary classification, and 97.17% and 96.06% accuracy for the second dataset. The model outperformed existing methods (Muralidharan et al., 2022). Linden at al. (2022) developed a machine learning model using data from the Lean European Open Survey on SARS-CoV2-infected patients (LEOSS) to predict COVID-19 patient mortality, achieving 80% AUC. The study found intersections

between dementia-related molecular mechanisms and COVID-19 and suggested that some anti-cancer drugs might have potential efficacy against COVID-19 (Linden et al., 2021). Van Lissa et al. (2022) investigated the determinants of infection prevention behaviors during the early stages of the COVID-19 pandemic using a multinational survey. Conducted from March to May 2020 with 56,072 participants from 28 countries, the study used the Random Forests (RF) algorithm and predicted 52% of the variance in infection prevention behavior. The model highlighted that social beliefs and societal factors are more significant predictors than individual psychological states (Van Lissa et al., 2022). Tandan et al. (2020) analyzed COVID-19 patient data to identify common symptoms and patterns. In a study of 1,560 patients, the most common symptoms were fever (67%), cough (37%), fatigue/body ache (11%), pneumonia (11%), and sore throat (8%). The study highlights that cough is a significant symptom, especially in patients presenting with fever and heart failure (Tandan et al., 2021). Rahimi et al. (2021) reviewed machine learning prediction models for COVID-19. The study focused on classifying these models, evaluating their criteria, and comparing solution approaches. It highlighted COVID-19's global spread and over 36 million cases. Key findings include that medicine, biochemistry, and mathematics are the most discussed fields, with terms such as coronavirus, prediction, and epidemic being prominent. Deep Learning and sir models are noted as the most frequently used. The research aims to help identify gaps and develop new prediction models (Rahimi et al., 2023). Brinati et al. (2020) address the limitations of RRT-PCR tests for COVID-19 detection, such as long turnaround times and high costs, suggesting a need for quicker, cheaper alternatives. Using data from 279 patients at San Raffaele Hospital in Italy, two machine learning models were developed to distinguish between positive and negative cases. An interpretable decision tree model was also created to assist clinicians. The study shows that blood test analysis combined with machine learning can be a valuable alternative, particularly in resource-limited settings. Key indicators like lymphopenia and elevated C-reactive protein (CRP) levels were effectively linked to COVID-19 positivity by the models (Brinati et al., 2020).

In a study conducted by Zhang et al. (2020) in Hong Kong, the effects of the COVID-19 pandemic on social behaviors were examined. Their findings revealed that during the pandemic, daily close contacts decreased by 59%, while the total close contact time was reduced by 10%. These behavioral changes contributed to a 63.1% reduction in the risk of influenza transmission. The study highlighted that limiting close contact played a critical role in curbing the spread of the virus (Zhang et al., 2021). Similarly, in this study conducted in Türkiye, individuals who live alone and engage in 4-6 hours of close contact per day were found to have a higher likelihood of contracting COVID-19. These results underscore the significant impact of human behaviors on the transmission dynamics of pandemics.

This study examines the impact of factors, generally defined as personal characteristics and behaviors, on the spread of the COVID-19 pandemic using data mining and artificial intelligence techniques. The analysis was conducted on data collected from survey forms completed voluntarily. Initially, a data preprocessing phase was applied using basic data mining techniques. Then, classification algorithms were employed to predict whether individuals were infected with the virus based on their personality and behavioral characteristics. Finally, the behavior patterns of infected and non-infected individuals were extracted using association rules. This study provides several significant contributions to the existing literature on COVID-19:

- This research focuses on individuals' daily behavioral patterns and habits to identify precursor behaviors influencing individual transmission risk, contrasting previous studies that primarily focused on symptoms and medical data.
- This study uses association rule mining to analyze behavioral and habit-based data, expanding our understanding of social and individual factors influencing disease spread.
- Feature reduction techniques were applied to enhance the efficiency and performance of the data mining process.
- An AI-based model has been developed to automate the prediction of COVID-19 transmission by utilizing behavioral data. This model can serve as a valuable tool for decision-making and policy development.

This study addresses a gap in the existing literature on the transmission of the COVID-19 virus by exploring the role of behavioral factors, including the avoidance of travel, time spent in enclosed areas, and habits within and outside the home environment. The findings offer insights that can inform predictions regarding the spread of disease.

2. Materials and Methods

2.1 Data Preparation and Analysis

The data used in this study were collected through a survey created via Google Forms. The survey consists of 43 questions compiled from two main sources. The first source is the “ICARE Study Survey” (Duval & Collaborator, 2020). The second source comes from the survey

questions used in Zhang et al.'s study (Zhang et al., 2021). 515 survey responses formed the dataset for this study. Preprocessing involved converting data to nominal values and identifying outliers using WEKA (Waikato Environment for Knowledge Analysis) and R-Studio. The processed data were then formatted as .arff for WEKA and .csv for R-Studio. The dataset contains 397 examples classified as "Yes" and 118 examples classified as "No". The data were transferred to WEKA in. arff format and to R-Studio in .csv format. All classification methods provided by WEKA and the apriori algorithm in R-Studio were used for analysis. The results obtained are detailed in the experimental results section.

2.2 Experimental Design

In this study, the data was analyzed using 54 nominal classification algorithms in WEKA with a 10-fold cross-validation method. The dataset was divided into 10 folds; each fold was used as a test set while the remaining 9 fold were used for training the model. In R-Studio, the apriori algorithm was employed for association rule mining with R programming language for each attribute in the dataset. Rules were derived based on the presence of COVID-19 virus in the test. The dataset contains two classes: “cov_enf_ol = Yes” and “cov_enf_ol” = No. To generate more reliable and robust rules, parameter constraints on confidence and support were applied to the algorithm. The performance of the methods was evaluated using metrics such as correctly classified instances, Kappa, precision, recall, f-measure, and root mean square error (RMSE). These metrics were compared to determine which methods yielded better results.

3. Methods Used in the Analysis

This section provides a detailed explanation of the classification and association rule mining algorithms used in the study. This study utilized the open-source software WEKA for applying classification models. Various machine learning algorithms were applied to the dataset created from the survey responses using WEKA, and their performance was evaluated. A total of 54 classification models were tested in WEKA, and the top 5 models (Random Committee, RRF, Random Forest, RDA, Boosting) were selected based on the Kappa metric and are discussed in this section. Additionally, the apriori algorithm was used for association rule mining in the R-Studio application, and details of the algorithm are provided in this section.

3.1 Random Committee

In the Random Committee algorithm, a series of base classifiers are created by using different random number nodes on the same data, and the final classification result is obtained by averaging the predictions made by these base classifiers (Niranjan et al., 2018). When analyzing algorithm performance, it was found that the best results were achieved by providing the Random Committee parameter to the MLR algorithm included in the R package.

3.2 RRF (Regularized Random Forest)

The RRF algorithm is a prediction model that works by creating numerous copies of a decision tree prediction model and using all these trees together. It operates similarly to the Random Forest algorithm but with a few differences. One major difference is that during the training of trees in the RRF algorithm, regularization is applied to prevent the overfitting problem. Another difference is the random rotation of attributes during the training process, giving each tree a unique attribute configuration, which helps improve the model's generalization. Finally, while the RRF algorithm uses a regularization parameter added to the loss function to prevent overfitting during training, this is not used in the Random Forest algorithm (Jana et al., 2021).

3.3 Random Forest

Random Forest is a machine learning method used for classification and regression analysis. It involves creating numerous copies of a decision tree prediction model and using all of these trees together. Each tree is trained on randomly selected samples from the dataset, and then makes predictions for all samples in the dataset (Fernández et al., 2020). The final prediction is made by averaging the predictions of all trees to achieve the most accurate result. A key feature that distinguishes Random Forest from other decision tree algorithms is that each tree works with a unique subset of data, leading to different predictions. This process reduces the influence of dominant features in the dataset and helps build a more robust model (Grekousis et al., 2022). Important parameters of the Random Forest algorithm include the number of trees, sampling method, feature selection, and node splitting criterion (Breiman, 2001).

3.4 Boosting

Adaboost is a boosting algorithm in machine learning designed to enhance accurate predictions. The core concept of this algorithm is to build an ensemble model to improve weak models and increase model accuracy. A weak model is defined as one with low performance or slightly better than a random classifier. Adaboost improves these weak classifiers by adjusting their weights and then constructs the final model (Walker, 2021).

3.5 Apriori

The Apriori algorithm is used in data mining and statistics. It was first introduced by Agrawal and Srikant in 1994 (Agrawal & Srikant, 1994). This algorithm helps to find frequently occurring items in a dataset and to identify item co-occurrences. It uses a frequency counter to calculate the occurrence of items, which increments for each row in the dataset where an item set appears (Papi et al., 2022).

4. Metrics Used in Evaluating Analysis Results

In this study, the open-source WEKA program was used to evaluate the performance of classification models. WEKA was employed to test 54 different classification algorithms on the dataset created from survey responses. The performance evaluations were based on the kappa metric as well as other important metrics, including Correctly Classified Instances (CCI), f-measure, precision, and recall. The top-performing algorithms, namely Random Committee, RRF, Random Forest, RDA, and Boosting, were identified and their details are discussed in this section. Additionally, the apriori algorithm was used for association rule mining in R-Studio, and details of this algorithm are also provided in this section.

4.1 Confusion Matrix

The confusion matrix is a table that includes all positive and negative examples of the model, showing how well the model has performed in classifying the test data. It is an important concept for performance evaluation. The matrix is calculated using four quantities, which are the entries of the confusion matrix. Table 1 presents the confusion matrix.

Table 1. Confusion Matrix.

	Actual Positive	**Actual Negative**
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

- *TP (True Positive):* Represents the number of positive examples correctly classified as positive. TP is the count of examples that are actually positive and are classified as positive by the model.
- *TN (True Negative):* Represents the number of negative examples correctly classified as negative. TN is the count of examples that are actually negative and are classified as negative by the model.
- *FP (False Positive):* Represents the number of negative examples incorrectly classified as positive. FP is the count of examples that are actually negative but are classified as positive by the model.
- *FN (False Negative):* Represents the number of positive examples incorrectly classified as negative. FN is the count of examples that are actually positive but are classified as negative by the model.

4.2 Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) is a measure of precision commonly used to assess the accuracy of a prediction model. It quantifies how much predicted values deviate from actual values. RMSE is particularly useful in evaluating the effects of different parameters during model development. The primary advantage of RMSE is that it provides a clear indication of the magnitude of errors in predictions. The RMSE formula is expressed in Equation (1) (Liemohn et al., 2021).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2]} \qquad (1)$$

4.3 Precision

Precision measures the proportion of truly positive instances among those classified as positive. It indicates how many of the positively classified examples are actually positive. Precision is calculated as the ratio of correctly classified examples to all examples assigned to the positive class. Precision ranges from 0 to 1, where 1 means all examples in the class are correctly predicted, and 0 means none of the examples in the class are correctly predicted (Işik & Kapan Ulusoy, 2021). Precision is formulated by the Equation (2):

$$\mathrm{Precision} = \frac{TP}{TP + FP} \qquad (2)$$

4.4 Recall

Recall, also known as True Positive Rate, is defined as the ratio of correctly classified positive examples to all examples assigned to the positive class. In short, it measures the proportion of positive examples that are correctly classified. Recall ranges between 0 and 1, where 1 represents a perfect positive class and 0 represents a complete miss of positive class examples. Recall is a crucial metric because it focuses on minimizing the number of missed positive examples. It provides the proportion of positive examples that the model successfully identifies or classifies. The formula for recall is given in Equation (3).

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \qquad (3)$$

4.5 F-Measure (F-score)

When evaluated individually, precision and recall metrics may not yield meaningful and definitive results. In particular, with datasets that have imbalanced class distributions, these metrics can lead to misleading evaluations. Evaluating both metrics together provides a

more accurate assessment. The F-Measure, also known as the F-score in the literature, is the harmonic mean of precision and recall. The harmonic mean is used to avoid ignoring extreme values (Işik & Kapan Ulusoy, 2021). A high F1 score indicates that both metrics have high values, as shown in Equation (4).

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

4.6 Accuracy (CCI-Correctly Classified Instances)

Accuracy is the ratio of the number of correctly classified instances by the model to the total number of instances in the dataset. While accuracy is a commonly used evaluation metric, it can be misleading, especially in datasets with imbalanced class distributions. For example, assigning all instances to the dominant class can achieve high accuracy, regardless of the performance on the minority class. Therefore, accuracy should be interpreted considering the actual class distribution. It generally represents the percentage of instances in the training data that were correctly predicted by the model. The formula for accuracy is given in Equation (5):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

4.7 Kappa

Kappa is a metric that ranges between -1 and +1 and quantitatively measures the agreement between expected and observed values. A value close to +1 indicates a high level of agreement between the model's predictions and the actual values. Kappa is calculated using the formula in Equation (6) and represents the harmonic mean of precision and recall. Here, Pr(a) represents the expected agreement, Pr(e) represents the observed agreement, and Kappa (K) is calculated according to the chi-square table.

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \quad (6)$$

5. Experimental Results

This section details the results of applying various artificial intelligence-based classification algorithms to our dataset. We first identify the most influential features for classification and then explore association rules within the data to understand relationships between variables and COVID-19 infection status.

4.1 Performance Evaluation of Classification Algorithms

In this study, the performance of classification algorithms was evaluated using the open-source software WEKA. WEKA is a comprehensive tool that encompasses machine learning and data mining techniques. Among WEKA's classification algorithms, 54 different models were tested. Based on the kappa value, the top 5 algorithms demonstrating the highest performance were identified, and the evaluation metrics for these algorithms are presented in Table 2.

Table 2. Top 10 Algorithms with the Best Evaluation Results.

Algorithm	Evaluation Metrics				
	CCI	Kappa	Precision	Recall	F-Measure
Random Committee	420	0.3172	0.812	0.816	0.777
RRF	417	0.2983	0.8	0.81	0.771
Random Forest	422	0.293	0.854	0.819	0.77
RDA	421	0.2883	0.845	0.817	0.769
Boosting	410	0.279	0.772	0.796	0.764
Ranger	416	0.2411	0.828	0.808	0.753
Lazy.IBk	388	0.2102	0.726	0.753	0.735
Evtree	407	0.2083	0.765	0.79	0.742
Ada	409	0.204	0.779	0.794	0.741
SMO	375	0.2019	0.718	0.728	0.723

Table 2 presents the top 10 algorithms based on various evaluation metrics. Among these, Random Committee exhibits the highest F-Measure score of 0.777. This indicates that Random Committee provides a well-balanced performance in terms of precision and recall. Although Random Forest leads in precision (0.854) and Recall (0.819), the Random Committee's F-Measure demonstrates superior overall classification effectiveness, highlighting its capability to harmonize precision and recall. Therefore, despite other algorithms showing higher precision or recall individually, Random Committee's robust F-Measure makes it particularly effective for classification tasks.

4.2 Feature Extraction and Reduction

Feature extraction and reduction play a crucial role in enhancing model performance. By using reduced features, models can achieve the same classification success with less data, which shortens the data collection process. Additionally, models with fewer features utilize computational resources more efficiently, leading to reduced hardware usage and shorter time to obtain

results. The most important features were identified using the SignificanceAttributeEval function in the WEKA software, as shown in Table 3.

Table 3. Top 10 Important Features.

Ranking Value	Attribute Name	Full Attribute Name
0.2714	kap_al_mas_tak	Mask Wearing in Indoor Spaces
0.2412	ger_olm_sey_kac	Avoiding Unnecessary Travel
0.2299	kac_coc?	Number of Children
0.2272	sab_suy_el_yik	Hand Washing with Soap
0.2128	int_haberden*	Internet News Consumption*
0.2107	yas_ara	Age Range
0.2072	han_kac_kisi	Number of People in Household
0.2040	siz_vir_var_ken_kar_alm	Taking Precautions During Pandemic
0.1781	elt_sig*	Electronic Cigarette Use*
0.1571	aci_al_buy_sos_top_kac*	Avoiding Large Social Gatherings*
0.1505	dis_mas_kul	Mask Wearing Outdoors
0.1403	kil_ara	Age Interval
0.1368	yem_icm_mek_kac_der*	Eating/Drinking Outside Frequency*
0.1311	soylenti_agizdan_agiza	Word of Mouth Rumors
0.1264	aci_al_kuc_sos_top_kac*	Avoiding Small Social Gatherings*
0.1224	sag_diy_yem	Healthy Diet
0.121	sig_icm	Smoking
0.1196	sos_med_hab*	Social Media News Consumption*
0.1178	radyo_haber	Radio News Consumption
0.1175	sal_gen_hay_kal	Overall Life Quality During Pandemic
0.1505	dis_mas_kul	Mask Wearing Outdoors

The top feature in Table 3, "Mask Wearing in Indoor Spaces" (kap_al_mas_tak), has the highest importance value of 0.2714. This suggests that mask usage in enclosed areas plays a significant role in the classification model. Following closely, "Avoiding Unnecessary Travel" (ger_olm_sey_kac) and "Number of Children" (kac_coc?) are also highly influential with values of 0.2412 and 0.2299, respectively. These features indicate that both travel behavior and family size contribute significantly to the model's predictions. "Hand Washing with Soap" (sab_suy_el_yik) and "Internet News Consumption*" (int_haberden*) further highlight critical factors such as hygiene practices and information sources. The importance of these features underscores the role of personal health practices and news consumption in the classification process.

For feature reduction, the Particle Swarm Optimization (PSO) package developed by Moraglio et al. was used in the Weka software (Moraglio et al., 2007). The parameters were adjusted to "iterations=200, populationSize=40," different from the default values. Results obtained using the "Use full training set" option are shown in Figure 2. Upon examining the reduced features and their importance ranking, it was observed that the features marked with * in Table 3 were not included among the reduced features. When the reduced features were used to reclassify the algorithms from Table 2, the results shown in Table 4 were obtained.

Table 4. Performance of classification algorithms using reduced features.

	Evaluation Metrics				
Algorithm	**CCI**	**Kappa**	**Precision**	**Recall**	**F-Measure**
RRF	421	0.3663	0.804	0.817	0.792
RDA	409	0.324	0.774	0.796	0.775
Random Forest	421	0.3164	0.820	0.817	0.777
Random Committee	413	0.3124	0.780	0.802	0.774
Boosting	407	0.3073	0.767	0.790	0.769
Lazy.IBk	402	0.302	0.759	0.781	0.765
Ranger	411	0.2437	0.780	0.798	0.754
Evtree	409	0.2351	0.771	0.794	0.751
Ada	405	0.2124	0.756	0.786	0.743
SMO	399	0.1564	0.732	0.775	0.725

In Table 4, the RRF algorithm demonstrates the highest performance in terms of accuracy (CCI: 421), which is comparable to the best-performing Random Forest algorithm from the Table 2. The RRF algorithm also shows high values in Kappa and F-Measure metrics, indicating a strong overall performance. Although the Random Committe algorithm maintained the highest F-Measure value in the Table 2, its performance has declined in some metrics. Specifically, the F-Measure value for Random Committe decreased from 0.777 to 0.774. In summary, feature reduction has improved the performance of some algorithms, with RRF achieving the highest results. However, the performance of some algorithms has declined compared to the Table 2. The comparison of Table 2 and Table 4 reveals that feature reduction has positively impacted the performance of certain algorithms, while it has led to a decrease in others. Notably, significant increases in F-Measure values are observed in the Lazy.IBk and RRF algorithms, indicating substantial improvements in classification performance after feature reduction. However, the Random Committee algorithm experienced a decline in its F-Measure, suggesting that reduced features resulted in a performance loss for this algorithm. Other algorithms

exhibited more limited changes in their F-Measure values (see Figure 1).

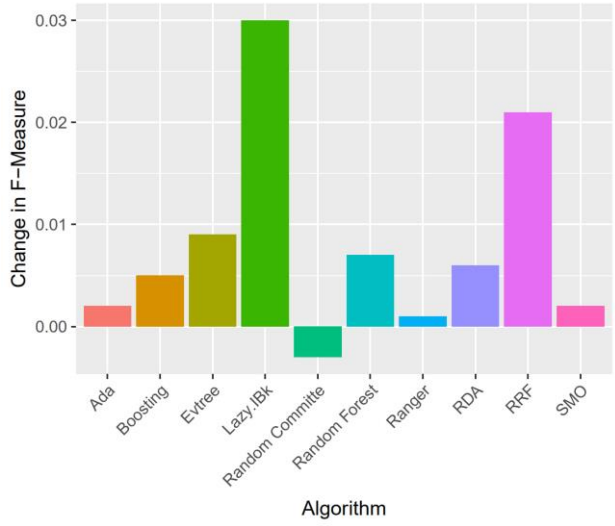

Figure 1. F-Measure Change After Feature Reduction

4.3 Association Rule Analysis and Results

In this section, the Apriori algorithm was used to discover association relationships between the attributes in the dataset. Specifically, the virus infection status (cov_enf_ol) was categorized into two classes: "Yes" and "No," and rules were extracted for both classes. During the tests, constraints were imposed on confidence and support values to obtain stronger rules.

The algorithm discovered only 90 rules for the "Yes" class, while approximately 7,469,192 rules were obtained for the "No" class. After rule reduction, the number of rules for the "Yes" class was reduced to 21. Table 5 lists the top 3 rules for the "Yes" and “No” class based on their confidence values. These results help us better understand the interaction between attributes concerning the virus infection status.

The association rules displayed in Table 5 highlight the first three rules identified for the “Yes” and “No” COVID-19 infection classes. In the “Yes” class, the first rule shows that individuals who live alone, have 1-3 daily close contacts, and spend 4-6 hours in close contact outside the home have a high likelihood of being infected. Similarly, men with a bachelor's degree, who spend 4-6 hours in closed spaces on weekends and have 1-3 close contacts daily, are also likely to be infected.

Shown in Table 5, In the "No" class, the first rule reveals that high school graduates who often believe in false information and experience no changes in body pain are not infected with COVID-19.

Likewise, those who mostly work remotely, attend small social gatherings, and follow TV news are not infected either. The third rule indicates that individuals living in a household of 4, who spend 1-3 hours outside the home on weekdays and mostly follow internet news, are not infected.

Table 5. Association rules for “Yes” and “No” classes.

Class	Association Rule	Sup.	Conf.
Evet	{household_size=1, daily_close_contacts=1-3, time_spent_outside_close_contacts =4-6 hours}	0.02	1
Evet	{gender=Male, last_degree=Bachelor, weekend_indoor_time=4-6 hours, daily_close_contacts=1-3}	0.02	1
Evet	{gender=Male, household_size=1, num_of_children=0, chronic_illness=No, daily_close_contacts=1-3, keep_distance=Most of the time}	0.02	0.92
Hayır	{last_degree=High School, belief_in_false_information=Often, headache_bodypain=No Change}	0.07	1
Hayır	{work_remotely=Most of the time, small_social_gatherings=Sometimes , tv_news=Mostly}	0.07	1
Hayır	{household_size=4, weekday_time_outside_home=1-3, internet_news=Mostly}	0.08	0.98

6. Results and Discussion

Understanding whether people's habits and behavioral characteristics are effective in the spread of epidemics can improve our ability to cope with future outbreaks. In this study, the impact of human habits and behavioral characteristics on the spread of COVID-19 during the pandemic in Türkiye was examined using artificial intelligence data mining techniques. Data collected through surveys were first cleaned and organized. The performance of classification algorithms was tested on the prepared data, and it was observed that the Random Committee algorithm demonstrated the best performance. This test, performed with 43 features, was repeated after feature reduction. It was found that the number of features could be reduced to 14. With these findings, performance tests of all classifiers were repeated, and an increase in performance was observed for all algorithms except Random Committee. The RRF algorithm, which showed the highest performance increase and yielded better results after feature reduction, was identified as the most suitable algorithm for classification. Additionally, the fact that the highest performance was observed with tree-based ensemble learning algorithms suggests that these algorithms are more effective for classifying this type of categorical nominal data.

According to the results of feature reduction, the variables mask_in_indoor, avoid_travel, num_children, hand_wash_soap, and age_range have been identified as the top 5 behavioral traits with the most significant impact on the course of the pandemic.

The association rule analysis provides significant findings regarding factors influencing COVID-19 infection for both the "Yes" and "No" classes. For the "Yes" class, rules such as living alone, frequent close contact, and spending 4-6 hours in close contact outside the home show a strong confidence level and high lift value. Additionally, an important rule involves males with a bachelor's degree who spend 4-6 hours in closed spaces on weekends and have frequent close contact, reinforcing the association with COVID-19 infection. The survey's findings, particularly the strong correlation between mask-wearing in enclosed spaces and avoiding travel, highlight potential crucial preventative measures for future pandemics. The association between virus transmission and lack of children also suggests a connection to loneliness. Loneliness might increase individuals' need to leave home and socialize, thereby potentially increasing the risk of virus transmission and spread. This warrants consideration of measures to address loneliness in future quarantines, such as teleconferencing-based interventions specifically targeting individuals living alone.

On the other hand, for the "No" class, key rules indicate that frequent mask usage, avoidance of travel, and a lower number of children are strongly associated with not contracting COVID-19. Furthermore, behaviors such as adhering to proper handwashing practices and avoiding travel unless necessary are also significant for this class.

These findings highlight the impact of specific behavioral traits on virus spread, showing that behaviors like mask-wearing and travel restrictions significantly affect infection rates. The performance improvements observed in classification algorithms when feature reduction was applied emphasize the importance of these factors in predicting COVID-19 outcomes.

7. Conclusions

In conclusion, it has been observed that the behaviors and habits exhibited by individuals during the pandemic in Türkiye had a significant impact on the progression of the outbreak. This impact can be utilized through artificial intelligence techniques for both classifier model development and association rule mining. The findings indicate that tree-based algorithms are more suitable for classification success and that feature reduction positively affects classification performance. During the pandemic, it was evident that patterns in the behaviors and habits of individuals infected and not infected with the virus could be revealed using artificial intelligence techniques. In the future, researchers can use or develop the methods employed in this study to conduct research on populations in different regions of the world and compare their findings with those from Türkiye.

8. Limitations and Recommendations

It should be noted that this study is subject to a number of acknowledged limitations. Primarily, the dataset comprises 515 responses, which, while sufficient for preliminary data mining analyses, may not fully represent the population of Turkey. Furthermore, the majority of participants are from Istanbul and the Marmara region, which may limit the generalizability of the findings to other regions with different cultural, social, or economic characteristics.

Due to the fact that the survey was conducted within a particular time frame, it is likely that the social and economic conditions that were in place at the time had an impact on the responses that the participants gave. Future research can aim to obtain a larger and more diverse data set with the objective of enhancing the findings' generalizability and robustness. Furthermore, the validity and generalizability of subsequent findings may be enhanced through the utilization of a combination of methods, approaches and/or longitudinal studies.

Declaration of Ethical Standards

Ethical approval for this study was obtained from the Scientific Research and Publication Ethics Committee of Tekirdağ Namık Kemal University, with the decision dated 18/06/2021, numbered T2021-648, and decision number 8.

This study is derived from a master's thesis (thesis number: 10529711) titled *"Analysis of the Effect of Life Habits in the Spreading Process of COVID-19 with Artificial Intelligence,"* defended on January 30, 2023, under the supervision of Assist. Prof. Erkan Özhan by İrem Sena Tekin in the Department of Computer Engineering at Tekirdağ Namık Kemal University, Institute of Natural and Applied Sciences.

Acknowledgments

This study is part of the Master of Science thesis by İrem Sena Tekin, conducted in the Department of Computer Engineering within the Institute of Natural and Applied Sciences at Tekirdağ Namık Kemal University, under the supervision of thesis advisor Erkan Özhan. The authors would like to thank the Institute for its support and all survey participants for their valuable contributions.

Credit Authorship Contribution Statement

Author-1 : Research, Analysis, Writing, Data curation, Review.
Author-2 : Research, Analysis, Data curation, Writing, Visualization, Editing and Supervision.

Declaration of Competing Interest

The authors have no conflicts of interest to declare regarding the content of this article.

Data Availability Statement

All data generated or analyzed during this study are included in this published article.

The data used in this study are retained for use in future research and are not currently available for public access. For further information about the data, please contact the authors.

9. References

Adamo, J.-M. (2001). *Data Mining for Association Rules and Sequential Patterns*. Springer New York. https://doi.org/10.1007/978-1-4613-0085-4

Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, 487–499.

Alakus, T. B., & Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals*, *140*, 110120. https://doi.org/10.1016/j.chaos.2020.110120

Älgå, A., Eriksson, O., & Nordberg, M. (2020). Analysis of Scientific Publications During the Early Phase of the COVID-19 Pandemic: Topic Modeling Study. *Journal of Medical Internet Research*, *22*(11), e21559. https://doi.org/10.2196/21559

Asia, S. (2022). *Americas Covid-19 South-East Asia*. 1–5.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brinati, D., Campagner, A., Ferrari, D., Locatelli, M., Banfi, G., & Cabitza, F. (2020). Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. *Journal of Medical Systems*, *44*(8), 135. https://doi.org/10.1007/s10916-020-01597-4

De Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives*, *6*, 100150. https://doi.org/10.1016/j.trip.2020.100150

Duval, K. L., & Collaborator, S. B. (2020). *iCARE Collaborator Documents*. https://doi.org/10.17605/OSF.IO/NSWCM

Epsi, N. J., Chenoweth, J. G., Blair, P. W., Lindholm, D. A., Ganesan, A., Lalani, T., Smith, A., Mody, R. M., Jones, M. U., Colombo, R. E., Colombo, C. J., Schofield, C., Ewers, E. C., Larson, D. T., Berjohn, C. M., Maves, R. C., Fries, A. C., Chang, D., Wyatt, A., … Richard, S. A. (2024). Precision Symptom Phenotyping Identifies Early Clinical and Proteomic Predictors of Distinct COVID-19 Sequelae. *The Journal of Infectious Diseases*, jiae318. https://doi.org/10.1093/infdis/jiae318

Fernández, R. R., Martín De Diego, I., Aceña, V., Fernández-Isabel, A., & Moguerza, J. M. (2020). Random forest explainability using counterfactual sets. *Information Fusion*, *63*, 196–207. https://doi.org/10.1016/j.inffus.2020.07.001

Grekousis, G., Feng, Z., Marakakis, I., Lu, Y., & Wang, R. (2022). Ranking the importance of demographic, socioeconomic, and underlying health factors on US COVID-19 deaths: A geographical random forest approach. *Health & Place*, *74*, 102744. https://doi.org/10.1016/j.healthplace.2022.102744

Işik, K., & Kapan Ulusoy, S. (2021). Metal Sektöründe üretim sürelerine etki eden faktörlerin veri madenciliği yöntemleriyle tespit edilmesi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, *36*(4), 1949–1962. https://doi.org/10.17341/gazimmfd.736659

Israfil, S. M. H., Sarker, Md. M. R., Rashid, P. T., Talukder, A. A., Kawsar, K. A., Khan, F., Akhter, S., Poh, C. L., Mohamed, I. N., & Ming, L. C. (2021). Clinical Characteristics and Diagnostic Challenges of COVID−19: An Update From the Global Perspective. *Frontiers in Public Health*, *8*, 567395. https://doi.org/10.3389/fpubh.2020.567395

Jana, R. K., Ghosh, I., Das, D., & Dutta, A. (2021). Determinants of electronic waste generation in Bitcoin network: Evidence from the machine learning approach. *Technological Forecasting and Social Change*, *173*, 121101. https://doi.org/10.1016/j.techfore.2021.121101

Kartsonaki, C., Baillie, J. K., Barrio, N. G., Baruch, J., Beane, A., Blumberg, L., Bozza, F., Broadley, T., Burrell, A., Carson, G., Citarella, B. W., Dagens, A., Dankwa, E. A., Donnelly, C. A., Dunning, J., Elotmani, L., Escher, M., Farshait, N., Goffard, J.-C., … Zucman, D. (2023). Characteristics and outcomes of an international cohort of 600 000 hospitalized patients with COVID-19. *International Journal of Epidemiology*, *52*(2), 355–376. https://doi.org/10.1093/ije/dyad012

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. https://doi.org/10.1038/nrg3920

Liemohn, M. W., Shane, A. D., Azari, A. R., Petersen, A. K., Swiger, B. M., & Mukhopadhyay, A. (2021). RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. *Journal of Atmospheric and Solar-Terrestrial Physics*, *218*, 105624. https://doi.org/10.1016/j.jastp.2021.105624

Linden, T., Hanses, F., Domingo-Fernández, D., DeLong, L. N., Kodamullil, A. T., Schneider, J., Vehreschild, M. J. G. T., Lanznaster, J., Ruethrich, M. M., Borgmann, S., Hower, M., Wille, K., Feldt, T., Rieg, S., Hertenstein,

B., Wyen, C., Roemmele, C., Vehreschild, J. J., Jakob, C. E. M., … Fröhlich, H. (2021). Machine Learning Based Prediction of COVID-19 Mortality Suggests Repositioning of Anticancer Drug for Treating Severe Cases. *Artificial Intelligence in the Life Sciences*, *1*, 100020. https://doi.org/10.1016/j.ailsci.2021.100020

Mallah, S. I., Ghorab, O. K., Al-Salmi, S., Abdellatif, O. S., Tharmaratnam, T., Iskandar, M. A., Sefen, J. A. N., Sidhu, P., Atallah, B., El-Lababidi, R., & Al-Qahtani, M. (2021). COVID-19: Breaking down a global health crisis. *Annals of Clinical Microbiology and Antimicrobials*, *20*(1), 35. https://doi.org/10.1186/s12941-021-00438-7

Moore, K. A., Lipsitch, M., Barry, J. M., & Osterholm, M. T. (2020). COVID-19: The CIDRAP Viewpoint. *Center of Infectious Diseases Research and Policy*, 1–9.

Moraglio, A., Di Chio, C., & Poli, R. (2007). Geometric Particle Swarm Optimisation. In M. Ebner, M. O'Neill, A. Ekárt, L. Vanneschi, & A. I. Esparcia-Alcázar (Eds.), *Genetic Programming* (Vol. 4445, pp. 125–136). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-71605-1_12

Muralidharan, N., Gupta, S., Prusty, M. R., & Tripathy, R. K. (2022). Detection of COVID19 from X-ray images using multiscale Deep Convolutional Neural Network. *Applied Soft Computing*, *119*, 108610. https://doi.org/10.1016/j.asoc.2022.108610

Niranjan, A., Prakash, A., Veena, N., Geetha, M., Shenoy, P. D., & Venugopal, K. R. (2018). EBJRV: An Ensemble of Bagging, J48 and Random Committee by Voting for Efficient Classification of Intrusions. *WIECON-ECE 2017 - IEEE International WIE Conference on Electrical and Computer Engineering 2017*, 51–54. https://doi.org/10.1109/WIECON-ECE.2017.8468876

Papi, R., Attarchi, S., Darvishi Boloorani, A., & Neysani Samany, N. (2022). Knowledge discovery of Middle East dust sources using Apriori spatial data mining algorithm. *Ecological Informatics*, *72*, 101867. https://doi.org/10.1016/j.ecoinf.2022.101867

Rahimi, I., Chen, F., & Gandomi, A. H. (2023). A review on COVID-19 forecasting models. *Neural Computing and Applications*, *35*(33), 23671–23681. https://doi.org/10.1007/s00521-020-05626-8

Sabherwal, A. K., Sood, A., & Shah, M. A. (2024). Evaluating mathematical models for predicting the transmission of COVID-19 and its variants towards sustainable health and well-being. *Discover Sustainability*, *5*(1), 38. https://doi.org/10.1007/s43621-024-00213-6

Sahu, A. K., Mathew, R., Aggarwal, P., Nayer, J., Bhoi, S., Satapathy, S., & Ekka, M. (2021). Clinical Determinants of Severe COVID-19 Disease – A Systematic Review and Meta-Analysis. *Journal of Global Infectious Diseases*, *13*(1), 13–19. https://doi.org/10.4103/jgid.jgid_136_20

Tadie, M. B., Yimer, Y. S., & Taye, G. (2024). Determinants of COVID-19 severity in Ethiopia: A multicentre case–control study. *BMJ Open*, *14*(5), e083076. https://doi.org/10.1136/bmjopen-2023-083076

Tandan, M., Acharya, Y., Pokharel, S., & Timilsina, M. (2021). Discovering symptom patterns of COVID-19 patients using association rule mining. *Computers in Biology and Medicine*, *131*, 104249. https://doi.org/10.1016/j.compbiomed.2021.104249

Tarca, A. L., Carey, V. J., Chen, X., Romero, R., & Drăghici, S. (2007). Machine Learning and Its Applications to Biology. *PLoS Computational Biology*, *3*(6), e116. https://doi.org/10.1371/journal.pcbi.0030116

The RECOVERY Collaborative Group. (2021). Dexamethasone in Hospitalized Patients with Covid-19. *New England Journal of Medicine*, *384*(8), 693–704. https://doi.org/10.1056/NEJMoa2021436

Van Lissa, C. J., Stroebe, W., vanDellen, M. R., Leander, N. P., Agostini, M., Draws, T., Grygoryshyn, A., Gützgow, B., Kreienkamp, J., Vetter, C. S., Abakoumkin, G., Abdul Khaiyom, J. H., Ahmedi, V., Akkas, H., Almenara, C. A., Atta, M., Bagci, S. C., Basel, S., Kida, E. B., … Bélanger, J. J. (2022). Using machine learning to identify important predictors of COVID-19 infection prevention behaviors during the early phase of the pandemic. *Patterns*, *3*(4), 100482. https://doi.org/10.1016/j.patter.2022.100482

Walker, K. W. (2021). Exploring adaptive boosting (AdaBoost) as a platform for the predictive modeling of tangible collection usage. *The Journal of Academic Librarianship*, *47*(6), 102450. https://doi.org/10.1016/j.acalib.2021.102450

Xie, N., Zhang, W., Chen, J., Tian, F., & Song, J. (2023). Clinical Characteristics, Diagnosis, and Therapeutics of COVID-19: A Review. *Current Medical Science*, *43*(6), 1066–1074. https://doi.org/10.1007/s11596-023-2797-3

Zhang, N., Jia, W., Lei, H., Wang, P., Zhao, P., Guo, Y., Dung, C.-H., Bu, Z., Xue, P., Xie, J., Zhang, Y., Cheng, R., & Li, Y. (2021). Effects of Human Behavior Changes During the Coronavirus Disease 2019 (COVID-19) Pandemic on Influenza Spread in Hong Kong. *Clinical Infectious Diseases*, *73*(5), e1142–e1150. https://doi.org/10.1093/cid/ciaa1818

Zheng, N., Du, S., Wang, J., Zhang, H., Cui, W., Kang, Z., Yang, T., Lou, B., Chi, Y., Long, H., Ma, M., Yuan, Q., Zhang, S., Zhang, D., Ye, F., & Xin, J. (2020). Predicting COVID-19 in China Using Hybrid AI Model. *IEEE Transactions on Cybernetics*, *50*(7), 2891–2904.

https://doi.org/10.1109/TCYB.2020.2990162

Internet References

1- Zimmer, C., Corum, J., Wee, S.-L., & Kristoffersen, M., Coronavirus Vaccine Tracker, https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html, (10.06.2020).