

Research Article

The Use of Machine Learning Approaches for the Diagnosis of Acute Appendicitis

Omer F. Akmese ¹, Gul Dogan,² Hakan Kor,³ Hasan Erbay ⁴, and Emre Demir⁵

¹Department of Computer Technologies, University of Hitit, University of Kırıkkale, Çorum 19500, Turkey

²Department of Surgical Medical Sciences, University of Hitit, Çorum 19040, Turkey

³Department of Computer Technologies, University of Hitit, Çorum 19300, Turkey

⁴Department of Computer Engineering, University of Turkish Aeronautical Association, Ankara 06790, Turkey

⁵Department of Biostatistics, University of Hitit, Çorum 19040, Turkey

Correspondence should be addressed to Omer F. Akmese; ofarukakmese@hitit.edu.tr

Received 2 December 2019; Accepted 2 March 2020; Published 25 April 2020

Academic Editor: Selim Suner

Copyright © 2020 Omer F. Akmese et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Acute appendicitis is one of the most common emergency diseases in general surgery clinics. It is more common, especially between the ages of 10 and 30 years. Additionally, approximately 7% of the entire population is diagnosed with acute appendicitis at some time in their lives and requires surgery. The study aims to develop an easy, fast, and accurate estimation method for early acute appendicitis diagnosis using machine learning algorithms. Retrospective clinical records were analyzed with predictive data mining models. The predictive success of the models obtained by various machine learning algorithms was compared. A total of 595 clinical records were used in the study, including 348 males (58.49%) and 247 females (41.51%). It was found that the gradient boosted trees algorithm achieves the best success with an accurate prediction success of 95.31%. In this study, an estimation method based on machine learning was developed to identify individuals with acute appendicitis. It is thought that this method will benefit patients with signs of appendicitis, especially in emergency departments in hospitals.

1. Introduction

Acute appendicitis occurs in a wide range of patients of all ages. It is common especially in young adults between the ages of 10 and 30. In addition, about 7% of the population suffers from appendicitis throughout their lives [1–3]. Appendectomy for acute appendicitis is one of the most common emergency surgical procedures performed by general surgeons. Since delayed surgery can increase the rate of perforated appendicitis, appendectomy must be performed a few hours after diagnosis [4–7]. Although appendicitis is the most common surgical pathology of the abdomen, its diagnosis, follow-up, and treatment vary according to the surgeon's experience and preference [8, 9]. Approximately one-third of patients present with atypical complaints [10]. This condition is often associated with complications and delays in diagnosis. Many patients

suspected of acute appendicitis develop a complication of 22% to 62%, even if they are diagnosed and operated on without any complications [11]. A definitive diagnosis is essential to prevent complications of appendicitis and to prevent negative appendectomy. The diagnosis of appendicitis is based primarily on biochemical tests that are decisive for inflammation. The use of this method alone revealed a negative appendectomy rate of 12.3–19%. Despite the remarkable improvements in modern radiography imaging and diagnostic laboratory examinations, the correct diagnosis of acute appendicitis remains a challenge [12–15]. In order to make the diagnosis without any complications, studies were performed with easily accessible methods such as CRP (C-reactive protein), leukocyte count, neutrophil ratio, bilirubin, multislice computed tomography, and ultrasonography imaging techniques [16, 17]. Radiological methods, especially ultrasonography and computed

tomography, are widely and successfully used in the diagnosis of acute appendicitis and its complications [18, 19]. However, since these methods require specialized equipment and experienced radiologists, data mining, a different method for diagnosing the disease, has been used. Data mining is a methodology to discover hidden patterns from large datasets using statistical approaches [20]. Unlike traditional statistical techniques, data mining is mainly concerned with transforming data into information and learning from that data [21]. Among the various applications of data mining, its use in the health field is no exception. It has proven to be very useful for analyzing medical datasets and extracting strong molds [22, 23]. In this article, we tried to estimate the necessity of surgery by using data from blood samples of patients. Thus, it is aimed to test the accuracy of the diagnosis related to the disease to minimize the resource consumption and to provide more accurate utilization of the limited health services.

In most patients suspected of acute appendicitis, diagnostic procedures are performed using blood values or images obtained. In the study, which investigated the relationship between basic statistical methods and acute appendicitis, the patients who underwent 302 appendectomies in 2008 were examined for gender, age, and leukocyte values [24].

There are few studies on predicting or treating acute appendicitis with data mining using computers. Demirhan et al. conducted a survey of artificial intelligence applications in medicine. A system has been developed to evaluate the role of artificial neural networks (ANNs) in the diagnosis of patients with acute right groin pain. In the developed system, patient data collected from a hospital were used in the training and testing of ANN. The performance of the system was compared with the Alvarado score and the evaluations of experienced doctors. The patient's symptoms and signs, hematological assessment, and demographic information were used as input data in the ANN [25].

In another study, the performance of fuzzy inference systems and logistic regression classifiers was compared using appendicitis data [26]. In the study conducted on samples selected from the KEEL (Knowledge Extraction based on Evolutionary Learning) database with the R program, the accuracy values of the results obtained from fuzzy inference systems and logistic regression were high. In another study where data were processed by the clustering method, gender, abdominal pain, age parameters and leukocyte, platelet, lymphocyte, neutrophil, and C-reactive protein (CRP) values were used. The number of ideal clusters in the study was determined as four. This clustering study was intended to design an ideal decision support system, including physician views [27].

2. Materials and Methods

Machine learning is the term automatic data analysis for analytical and statistical pattern recognition and modeling. It allows recursively learning from the new data, updating the models that have been created, and thus finding hidden information or patterns in large datasets. Data mining

focuses on the exploration of previously unknown features of an extensive existing dataset [28]. Data mining methods are used by many researchers for prediction purposes. The classification and evaluation under data mining techniques help in the creation of training data, the classification of the estimation model, and the testing of classification efficiency [29]. It can be ensured that the models can be adapted to future datasets.

In this study, the CRISP-DM methodology was applied for data mining. The CRISP-DM methodology of data mining proposes step-by-step procedures so that the process can be reliable and standard. This methodology follows a cyclical process that includes business understanding, data understanding, data preparation, modeling, evaluation, and deployment [30].

In data mining, models can be classified under two main headings: predictive and descriptive. In predictive models, a model is developed by acting on the results of known data. It is intended to estimate the results for unknown data, that is, make inferences that will predict the future. However, in descriptive models, it is provided to define the patterns in the available data that can be used to guide decision making [31].

This study was carried out with the data of clinical records collected at the Hitit University Training and Research Hospital. These data show records of gender, laboratory markers, and surgical conditions belonging to patients applied to the hospital with suspicion of appendicitis. In the design of the study, blood samples of 595 patients admitted to Hitit University Training and Research Hospital were taken. Those diagnosed with acute appendicitis and who underwent surgery were identified as one group; others were identified as a separate group. In addition to descriptive statistical analysis, machine learning methods have been used to predict groups. Before starting the research, priori power analysis was performed using *t*-test in independent groups. In order to reach 95% power with $\alpha=0.05$ error, the power analysis found using Cohen $d=0.35$ effect size calculated using the literature information, and it was decided to include a total of 428 patients, 214 in each group.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (the Ethics Committee of Noninterventional Research in Hitit University (reference number: 2019-45)) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individuals included in this study.

2.1. Data and Statistical Analysis. Independent sample *t*-test was used for parametric test assumptions in numerical measurement comparisons in groups with and without surgery. Pearson correlation analysis and chi-square tests were used for weighting. Machine learning algorithms were used for disease classification. Classification success was evaluated with accuracy, precision, and recall results. Both qualitative and quantitative data were analyzed by machine learning algorithms. In the study, data analyses were carried

TABLE 1: Dataset description.

| Name | Type | Description | Role |
|--------|-------------|--------------------------------|--------|
| Group | Categorical | No surgery/having surgery | Target |
| Gender | Categorical | Female/male | Input |
| HGB | Numeric | Hemoglobin | Input |
| NEU | Numeric | Neutrophil | Input |
| LYM | Numeric | Lymphocytes | Input |
| MCV | Numeric | Mean corpuscular volume | Input |
| MPV | Numeric | Mean platelet volume | Input |
| HTC | Numeric | Hematocrit | Input |
| PLT | Numeric | Thrombosis | Input |
| CRP | Numeric | C-reactive protein | Input |
| WBC | Numeric | White blood cells (leukocytes) | Input |

Categorical data represent types of data which may be divided into groups. Numerical data are data expressed in numbers, unlike letters or words. Target: dependent variable; input: independent variable.

TABLE 2: Blood test results in acute appendicitis dataset.

| Name | Value range | Having surgery | No surgery | <i>p</i> value |
|--------|-------------|----------------|----------------|------------------|
| Group | 0 or 1 | | | |
| Gender | 1 or 2 | 77/137 | 104/110 | 0.008 |
| HGB | 1.4–130 | 13.63 ± 8.18 | 13.02 ± 1.43 | 0.348 |
| NEU | 0.8–29 | 11.82 ± 5.31 | 9.11 ± 5.77 | <0.001 |
| LYM | 0.2–94 | 3.09 ± 7.09 | 2.44 ± 1.17 | 0.016 |
| MCV | 4–97.3 | 80.98 ± 8.76 | 81.20 ± 5.79 | 0.434 |
| MPV | 6–99 | 14.03 ± 20.02 | 9.31 ± 7.93 | 0.708 |
| HTC | 3–99 | 40.32 ± 6.24 | 39.69 ± 3.49 | 0.143 |
| PLT | 111–593 | 272.49 ± 73.58 | 274.37 ± 82.73 | 0.954 |
| CRP | 0–302 | 39.61 ± 51.62 | 22.86 ± 40.64 | <0.001 |
| WBC | 6–31590 | 15280 ± 5302 | 12541 ± 6259 | <0.001 |

Group: 0 refers to those without surgery and 1 refers to those who were operated on. Gender variable: 1 refers to females and 2 refers to males. HGB: hemoglobin, NEU: neutrophil, LYM: lymphocytes, MCV: mean corpuscular volume, MPV: mean platelet volume, HTC: hematocrit, PLT: thrombosis, CRP: C-reactive protein, and WBC: white blood cells (leukocytes). *p* values < 0.05 are statistically significant. Bold denotes a significant *p* value less than 0.05.

out using SPSS (22.0), Rapidminer (9.5) packet programs, and Python (3.7) programming language.

When Table 1 is examined, the data were summarized according to the gender of the patients, their operating status, and blood test results. While gender and blood test results are used as input data, the group variable is determined as the target (dependent) variable.

Table 2 shows the blood test results from the dataset. In the variable Group, 0 refers to those without surgery and 1 refers to those who were operated on. For the gender variable, 1 refers to females and 2 refers to males. Other data are the values in the blood samples. Values are presented as mean ± standard deviation. Bold denotes a significant *p* value less than 0.05. *p* values are given in Table 2. Only values less than one-thousand were shown as $p < 0.001$ in accordance with APA reporting.

In this study, the attributes and doctor’s recommendations, which are stated as important for diagnosis in the literature, were taken into consideration, and the weights of the qualities were calculated. The relationship between two or more variables has not been investigated. The weights of

TABLE 3: Attributes and weighting.

| No | Attributes | Weighting (chi-square) |
|----|------------|------------------------|
| 1 | NEU | 1 |
| 2 | WBC | 0.994 |
| 3 | CRP | 0.436 |
| 4 | MPV | 0.201 |
| 5 | LYM | 0.120 |
| 6 | HTC | 0.119 |
| 7 | Gender | 0.079 |
| 8 | MCV | 0.075 |
| 9 | MPV | 0.049 |
| 10 | HB | 0 |

HGB: hemoglobin, NEU: neutrophil, LYM: lymphocytes, MCV: mean corpuscular volume, MPV: mean platelet volume, HTC: hematocrit, PLT: thrombosis, CRP: C-reactive protein, and WBC: white blood cells (leukocytes).

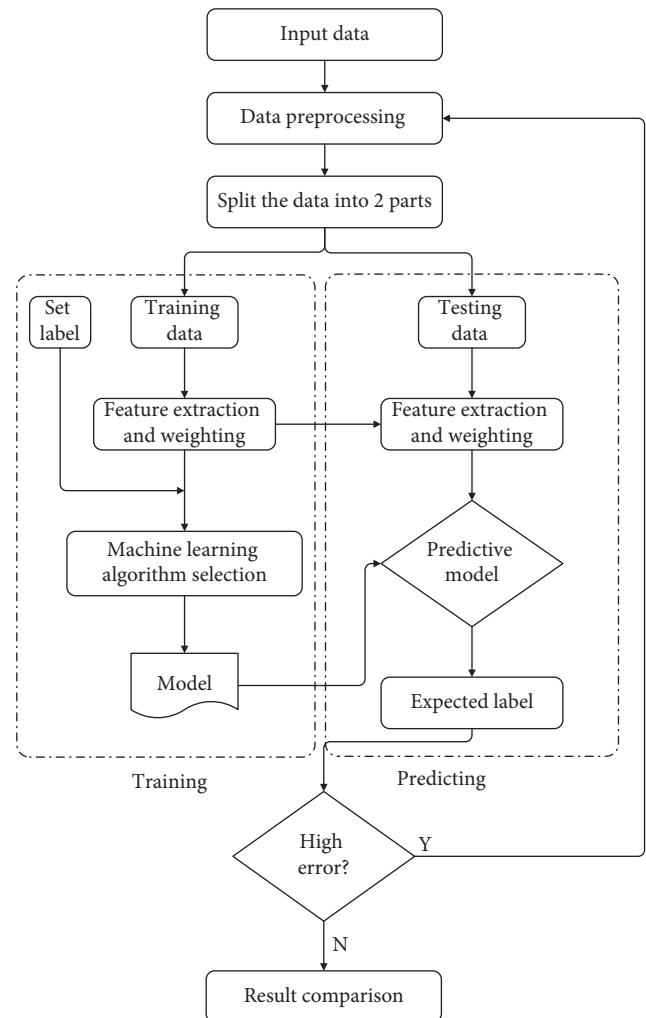


FIGURE 1: The architecture of the proposed system.

numerical and categorical variables in predicting a categorical outcome variable were determined. The chi-square weighting method was used to determine these weights. The higher the weight of an attribute, the more relevant it is

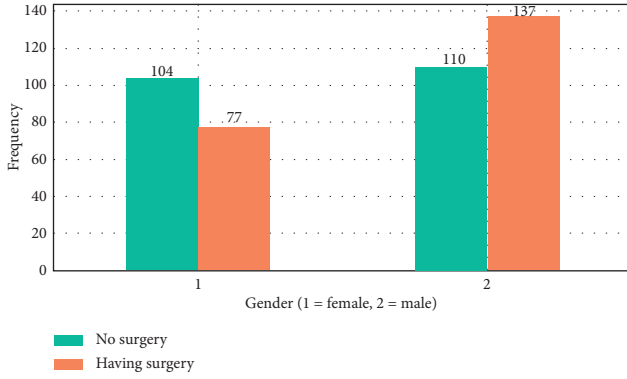


FIGURE 2: Graph of those who underwent surgery according to gender after data preprocessing.

considered. Calculated weights were normalized from 0 to 1. Table 3 shows the attributes and weightings.

When the analysis of the established model is performed, the performance of the measurements obtained is evaluated. In the calculation of these values, true positive, false positive, true negative, and false negative values are used.

TP (correct decision): in the actual case, it means the accurate estimate of the patients who were operated on.

FP (type I error): patients who were not operated on in the real situation were defined as operated.

TN (correct decision): it means the accurate estimation of patients who do not undergo surgery in real condition.

FN (type II error): patients who underwent surgery in real condition were defined as not operated.

Accuracy: the ratio of correctly estimated samples to the number of all samples. That is, the test is the rate of total correct diagnoses.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}}. \quad (1)$$

Precision (positive predictive value): it is the ratio of correctly predicted positive samples to the number of samples estimated in the positive class.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{TP}}{\text{P}'}. \quad (2)$$

Recall (sensitivity, true positive rate): it is the ratio of correctly predicted positive samples to the number of samples in the true positive class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{\text{P}}. \quad (3)$$

Predictive data mining was performed by machine learning algorithms, and the accuracy of these methods was compared. In the models, some of the data were used as training data, and some were used to test the accuracy of the model. The proposed architecture is shown in Figure 1.

The data collected according to the architecture seen in Figure 1 were subjected to preprocessing. The quality of the data dramatically influences the outcome of the estimation. This means that preprocessing plays an important role in the model [32]. After determining the dependent variable, the data were divided into two groups as training and test data.

Seventy percent of the data was used as training data, and thirty percent was used for testing the model. Finally, the performance of the model was evaluated. The missing values in the records have been deleted to maintain the accuracy of the attributes. In the preprocessing stage, the missing data have been cleared; as a result, the number of clinical records decreased from 595 records to 428.

According to Figure 2, there were 77 (36%) females and 137 (64%) males in the having surgery group and 104 (48.6%) females and 110 (51.4%) males in the no surgery group. In addition, half of the 428 records were female, and half were male.

It was found that the “gradient boosted trees” algorithm achieves the best success with an accurate prediction success rate of 95.31%. As a general principle, boosting algorithms try to achieve a strong learner by combining the weak learner obtained in each iteration under specific rules. The basic idea is to minimize the error and determine the target outputs for the next model. This technique is based on the advancement of subsequent estimates by learning from previous prediction errors.

It is a machine learning technique for gradient boosting, regression, and classification problems. Each tree is grown using information from previously created trees. The basic algorithm can be generalized to the following where x represents features and y represents response:

- (1) Gradient boosting creates a f_1 function that generates predictions in the first iteration.

$$f_1(x) = y. \quad (4)$$

- (2) It calculates the difference between the estimates and the target value and creates the function h_1 for these differences.

$$h_1(x) = y - f_1(x). \quad (5)$$

- (3) It creates a new tree.

$$f_2(x) = f_1(x) + h_1(x). \quad (6)$$

- (4) It recalculates the difference between estimates and goals.

$$h_2(x) = y - f_2(x). \quad (7)$$

- (5) In this way, it continually tries to increase the success of the “ F ” function and to reduce the difference between predictions and targets to zero.

Gradient descent is a very general optimization algorithm that can find optimal solutions to a wide range of problems. Estimates are updated so that the loss function (MSE) is minimum, and the estimated values are close to the

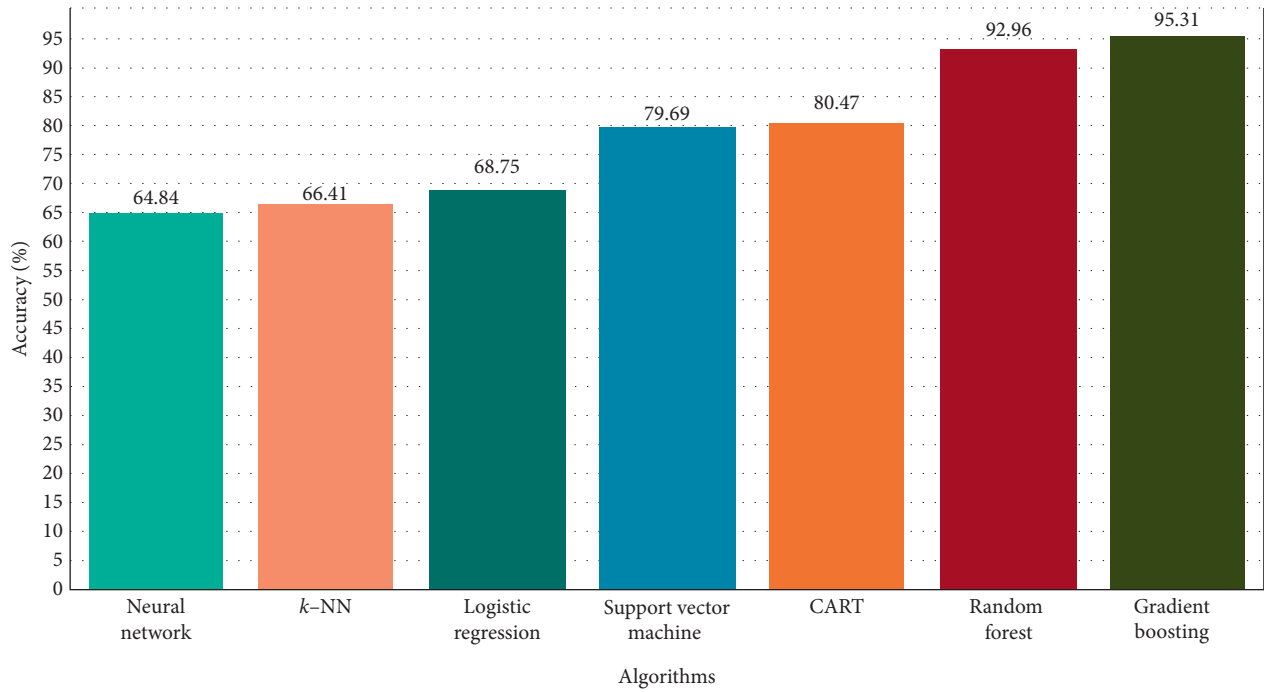


FIGURE 3: Accuracy percentages of algorithms.

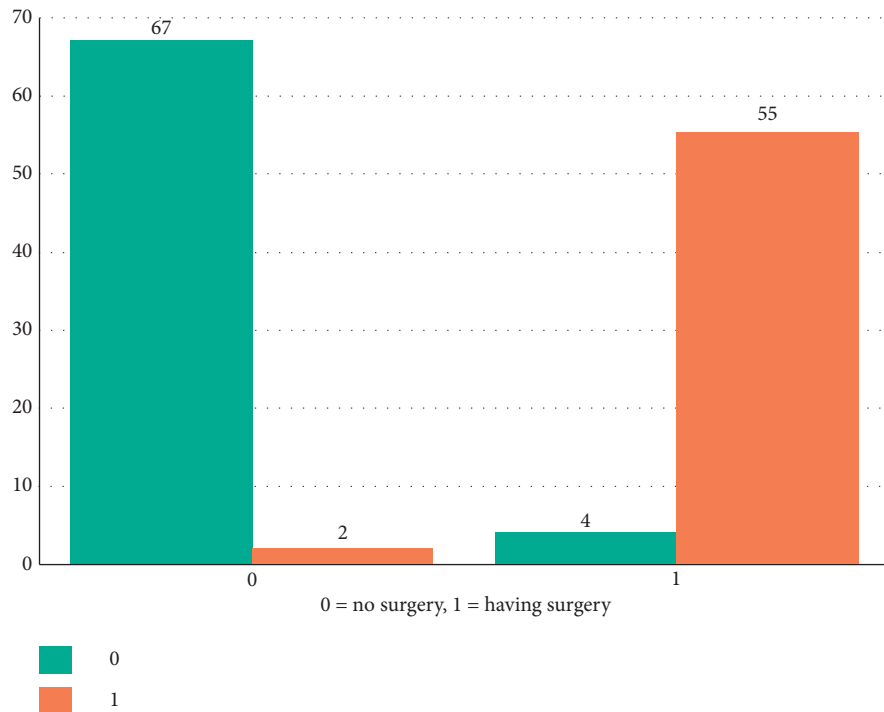


FIGURE 4: Gradient boosting tree prediction.

actual values. The general idea of the gradient descent is to adjust the parameters repeatedly to minimize a cost function.

$$J(\theta) = \frac{1}{n} \sum_{i=1}^n (h_{\theta}(x^{(i)}) - y^{(i)})^2, \quad (8)$$

where n is the number of training examples, $x^{(i)}$ is the input vector for the i^{th} training examples, $y^{(i)}$ is the class label for the i^{th} training examples, θ is the chosen parameter values or weights, and $h_{\theta}(x^{(i)})$ is the algorithm's prediction for the i^{th} training examples using the parameters θ .

TABLE 4: Results of gradient boosting tree analysis.

| Accuracy: 95.31% | True 1 | True 0 | Total | Class precision (%) |
|------------------|-----------------------------|-----------------------------|-------------|---------------------|
| Pred. 1 | 55 (TP) Correct Decision | 2 (FP) Type I error | 57 (P') | 96.49 |
| Pred. 0 | 4 (FN) Type II error | 67 (TN) Correct Decision | 71 (N') | 94.36 |
| Total | 59 (P) | 69 (N) | 128 (P + N) | |
| Class recall | 93.22% | 97.10% | | |

TP: true positives, TN: true negatives, FN: false negatives, and FP: false positives. Precision: it is the ratio of correctly predicted positive samples to the number of samples estimated in the positive class. Recall: it is the ratio of correctly predicted positive samples the ratio to the number of samples in the true positive class.

At each iteration, the residue of the loss function is calculated using the gradient descent method and becomes the target value for the next iteration [33]. In summary, gradient boosting uses the gradient descent method to minimize the derivable loss function value of each weak classifier.

3. Results and Discussion

During the research, 7 machine learning algorithms were tried. The most successful algorithm was gradient boosted trees. The accuracy rates of the algorithms applied in the research process are shown in Figure 3.

After the data preprocessing step, 70% of the clinical records falling to 428 were used as training data. The model was tested with 30% of the data. Of the 128 clinical records, 122 were correctly estimated. Thus, the predictive accuracy of the model was 95.31%.

Figure 4 shows the estimation results according to the gradient boosted trees algorithm. Accordingly, 67 out of 69 patients without surgery were correctly estimated. Of the 59 patients who underwent surgery, 55 were correctly estimated.

The ratio of correctly estimated samples to the number of all samples according to Table 4 is 95.31%. This ratio represents the total accuracy. It is represented as TP (true positives), TN (true negatives), FN (false negatives), and FP (false positives).

Examination and simple laboratory tests can often diagnose acute appendicitis, but when signs and symptoms are atypical, diagnosis is difficult. It has been observed that researchers who carry out their studies using whole blood counts have more emphasis on neutrophil/lymphocyte rates in the diagnosis of acute appendicitis. In diagnostic studies, mostly computerized imaging techniques and statistical analysis of the values obtained from blood counts come to the forefront. With the use of various algorithms, the contribution of computer estimation skills is seen in a small number of studies.

In this retrospective study, we proposed a model that predicts the necessity of appendicitis surgery using laboratory data of patients aged 0–18 years. The proposed model has been developed as a diagnostic tool to improve surgical decision making in patients with suspected acute appendicitis. In this way, unnecessary appendectomies and possible postoperative complications can be prevented. Increased use of preoperative computed tomography (CT) for the diagnosis of acute appendicitis has helped to reduce negative

appendectomy rates [34–37]. However, the use of CT has raised concerns about unnecessary exposure to radiation [38]. In addition to radiation exposure, these methods require specialized equipment and experienced radiologists. A different aspect of this study compared to others is the use of machine learning methods based on laboratory data for the diagnosis of appendectomy. The most important features that add value to the research globally can be listed as follows. The proposed model may assist in the diagnosis in the absence of specialized equipment and personnel in making decisions for surgical operation. In addition, the accuracy of disease-related diagnoses can be tested, resource consumption can be reduced, and limited health care can be used more accurately. In addition, an easy, fast, and accurate estimation method has been developed for the diagnosis of appendectomy which is the main problem of the article. Thus, accurate diagnosis can reduce hospital stay and cost. In future studies, new patient data should be added to the continuing education of the model. It is thought that the accuracy of estimation will increase as the number of data increases.

4. Conclusions

In this article, different machine learning methods were investigated in order to predict the necessity of appendicitis surgery by examining only the blood sample data of patients presenting with abdominal pain complaints. An easy, fast, and accurate estimation method has been developed for the diagnosis of appendectomy, which is the main problem of the article. Machine learning algorithms were used for patient data with suspicion of appendicitis, and the accuracy of these algorithms was compared. It was found that the gradient boosting tree algorithm achieves the best success with an accurate prediction success of 95.31%. The closest value to this result was found to be 92.96% with a random forest algorithm. In the study, the accuracy rate was considered as the most important factor.

Data Availability

Full collected data can be obtained through email from the corresponding author.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of

the institutional and national research committee (the Ethics Committee of Noninterventional Research in Hitit University (reference number: 2019-45)) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent

Informed consent was obtained from all individuals included in this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

OFA analyzed the data and wrote the article. GD provided the data and commented on the article. HK made an important contribution to the writing of the article. HE has substantively revised the work. ED contributed to the writing and revision of this paper. All authors have read and approved the article.

References

- [1] D. G. Addiss, N. Shaffer, B. S. Fowler, and R. V. Tauxe, "The epidemiology of appendicitis and appendectomy in the United States," *American Journal of Epidemiology*, vol. 132, no. 5, pp. 910–925, 1990.
- [2] M. Feldman, L. S. Friedman, and L. J. Brandt, *Sleisenger and Fordtran's Gastrointestinal and Liver Disease: Pathophysiology, Diagnosis, Management*, Elsevier Health Sciences, Amsterdam, Netherlands, 2015.
- [3] H. Köksal, B. Uysal, and R. Sarıbağcı, "Bir devlet hastanesinin akut apandisit tecrübesi/acute appendicitis experience of a state hospital," *Journal of Academic Emergency Medicine*, vol. 9, pp. 41–44, 2010.
- [4] N. A. Bickell, A. H. Aufses Jr., M. Rojas, and C. Bodian, "How time affects the risk of rupture in appendicitis," *Journal of the American College of Surgeons*, vol. 202, no. 3, pp. 401–406, 2006.
- [5] M. Busch, F. S. Gutzwiller, S. Aellig, R. Kuettel, U. Metzger, and U. Zingg, "In-hospital delay increases the risk of perforation in adults with appendicitis," *World Journal of Surgery*, vol. 35, no. 7, pp. 1626–1633, 2011.
- [6] B. G. Jeon, H. J. Kim, and S. C. Heo, "CT scan findings can predict the safety of delayed appendectomy for acute appendicitis," *Journal of Gastrointestinal Surgery*, vol. 23, no. 9, pp. 1856–1866, 2019.
- [7] C. L. Temple, S. A. Huchcroft, and W. J. Temple, "The natural history of appendicitis in adults A prospective study," *Annals of Surgery*, vol. 221, no. 3, pp. 278–281, 1995.
- [8] M. Abeş and H. Ö. Apaydın, "Türk çocuk cerrahlarının akut apandisit yaklaşımı," *Çocuk Cerrahisi Dergisi*, vol. 29, no. 3, pp. 88–92, 2015.
- [9] K. Tsao, S. D. St. Peter, P. A. Valusek et al., "Management of pediatric acute appendicitis in the computed tomographic era," *Journal of Surgical Research*, vol. 147, no. 2, pp. 221–224, 2008.
- [10] F. R. Lewis, J. W. Holcroft, J. Boey, and J. E. Dunphy, "Appendicitis," *Archives of Surgery*, vol. 110, no. 5, pp. 677–684, 1975.
- [11] S. Xharra, L. Gashi-Luci, K. Xharra et al., "Correlation of serum C-reactive protein, white blood count and neutrophil percentage with histopathology findings in acute appendicitis," *World Journal of Emergency Surgery*, vol. 7, no. 27, pp. 2–26, 2012.
- [12] P. A. Boonstra, R. N. van Veen, and H. B. A. C. Stockmann, "Less negative appendectomies due to imaging in patients with suspected appendicitis," *Surgical Endoscopy*, vol. 29, no. 8, pp. 2365–2370, 2015.
- [13] N. Kansakar, P. N. Agarwal, R. Singh et al., "Evaluation of combined use of modified Alvarado score and Ultrasound in predicting acute appendicitis: a prospective study," *International Surgery Journal*, vol. 5, no. 11, pp. 3594–3597, 2018.
- [14] P. van Amstel, R. R. Gorter, J. H. van der Lee, H. A. Cense, R. Bakx, and H. A. Heij, "Ruling out appendicitis in children: can we use clinical prediction rules?" *Journal of Gastrointestinal Surgery*, vol. 23, no. 10, pp. 2027–2048, 2019.
- [15] B. W. Warner, R. M. Kulick, M. M. Stoops, S. Mehta, M. Stephan, and U. R. Kotagal, "An evidenced-based clinical pathway for acute appendicitis decreases hospital duration and cost," *Journal of Pediatric Surgery*, vol. 33, no. 9, pp. 1371–1375, 1998.
- [16] B. Behzatoglu, E. Hatipoğlu, S. Bayramoğlu et al., "Akut apandisit tanısında ultrasonografi ve bilgisayarlı tomografi bulgularının karşılaştırılması," *Bakırköy Tıp Dergisi*, vol. 2, pp. 22–24, 2006.
- [17] D. R. McGowan, H. M. Sims, K. Zia, M. Uheba, and I. A. Shaikh, "The value of biochemical markers in predicting a perforation in acute appendicitis," *ANZ Journal of Surgery*, vol. 83, no. 1-2, pp. 79–83, 2013.
- [18] J. J. Atema, C. C. van Rossem, M. M. Leeuwenburgh, J. Stoker, and M. A. Boermeester, "Scoring system to distinguish uncomplicated from complicated acute appendicitis," *British Journal of Surgery*, vol. 102, no. 8, pp. 979–990, 2015.
- [19] M. M. Sevinç, E. Kınacı, E. Çakar et al., "Diagnostic value of basic laboratory parameters for simple and perforated acute appendicitis: an analysis of 3392 cases," *Ulus Travma Acil Cerrahi Derg*, vol. 22, no. 2, pp. 155–162, 2016.
- [20] C. Colak, E. Karaman, and M. G. Turtay, "Application of knowledge discovery process on the prediction of stroke," *Computer Methods and Programs in Biomedicine*, vol. 119, no. 3, pp. 181–185, 2015.
- [21] N. Mdzingwa, "Data Mining with Oracle 10 g Using Clustering and Classification Algorithms," *Thesis*, Rhodes University, Grahamstown, South Africa, 2005, <https://research.ict.ru.ac.za/g05m5125/CSHnsThesis.pdf>.
- [22] S. E. Brossette and P. A. Hymel Jr., "Data mining and infection control," *Clinics in Laboratory Medicine*, vol. 28, no. 1, pp. 119–126, 2008.
- [23] T. K. Dao, F. Zabaneh, J. Holmes, L. Disrude, M. Price, and L. Gentry, "A practical data mining method to link hospital microbiology and an infection control database," *American Journal of Infection Control*, vol. 36, no. 3, pp. S18–S20, 2008.
- [24] A. Aren, A. H. Gökçe, F. S. Gökçe et al., "Akut apandisit yaş, cinsiyet, lökosit değerleri ile ilişkisi," *Istanbul Tıp Dergisi*, vol. 10, no. 3, pp. 126–129, 2009.
- [25] A. Demirhan, Y. A. Kılıç, and G. İnan, "Tıpta yapay zeka uygulamaları," *Yoğun Bakım Dergisi*, vol. 9, no. 1, pp. 31–41, 2010.
- [26] İ. Kar, B. Bakırarar, and S. K. Köse, "Bulanık çıkarsama sistemleri ile veri madenciliği yöntemlerinin sınıflama performansının benzetim çalışması ile karşılaştırılması," in *Proceedings of XIX National and II International Biostatistics Congress*, p. 33, Antalya, Turkey, 2017.

- [27] K. Turhan, T. Kurt, B. Kurt, and İ. Buçan Kırkbir, "Akut apandisit nedeniyle opere edilen hastaların kümeleme yöntemi ile profillerinin belirlenmesi," in *Proceedings of Tıp Bilişiminde Yenilikler Sempozyumu*, Ankara, Turkey, 2017.
- [28] C. Petit, R. Bezemer, and L. Atallah, "A review of recent advances in data analytics for post-operative patient deterioration detection," *Journal of Clinical Monitoring and Computing*, vol. 32, no. 3, pp. 391–402, 2018.
- [29] P. Sinha and P. Sinha, "Comparative study of chronic kidney disease prediction using KNN and SVM," *International Journal of Engineering Research and Technology*, vol. 4, no. 12, pp. 608–612, 2015.
- [30] P. Chapman, J. Clinton, R. Kerber et al., "CRISP-DM 1.0 Step-by-step Data Mining Guide," in *SPSS*, pp. 9–13, 2000, <https://www.the-modeling-agency.com/crisp-dm.pdf>.
- [31] F. X. Bea and M. Schweitzer, *Allgemeine Betriebswirtschaftslehre 2: Bd. 2: Führung (Vol. 2)*, UTB, Munich, Germany, 2011.
- [32] K. M. Orabi, Y. M. Kamal, and T. M. Rabah, "Early predictive system for diabetes mellitus disease," in *Proceedings of Industrial Conference on Data Mining*, pp. 420–427, New York, NY, USA, July 2016.
- [33] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [34] B. A. Anderson, L. Salem, and D. R. Flum, "A systematic review of whether oral contrast is necessary for the computed tomography diagnosis of appendicitis in adults," *The American Journal of Surgery*, vol. 190, no. 3, pp. 474–478, 2005.
- [35] S. Krajewski, J. Brown, P. T. Phang, M. Raval, and C. J. Brown, "Impact of computed tomography of the abdomen on clinical outcomes in patients with acute right lower quadrant pain: a meta-analysis," *Canadian Journal of Surgery*, vol. 54, no. 1, p. 43, 2011.
- [36] P. M. Rao, J. T. Rhea, R. A. Novelline, A. A. Mostafavi, and C. J. McCabe, "Effect of computed tomography of the appendix on treatment of patients and use of hospital resources," *New England Journal of Medicine*, vol. 338, no. 3, pp. 141–146, 1998.
- [37] E. Sporn, G. F. Petroski, G. J. Mancini, J. A. Astudillo, B. W. Miedema, and K. Thaler, "Laparoscopic appendectomy is it worth the cost? trend analysis in the US from 2000 to 2005," *Journal of the American College of Surgeons*, vol. 208, no. 2, pp. 179–185, 2009.
- [38] D. J. Shah, R. K. Sachs, and D. J. Wilson, "Radiation-induced cancer: a modern view," *The British Journal of Radiology*, vol. 85, no. 1020, pp. e1166–e1173, 2012.