Research Article
Submission Date: July 31, 2023 Acceptance Date: August 11, 2023

Cece O., Gençtürk M. (2023). Net Profit Margin Forecasting with Machine Learning Methods in Hospital Finance Management. Journal of Health Systems and Policies (JHESP), V, 103-119, DOI: 10.52675/jhesp.1335249

Net Profit Margin Forecasting with Machine Learning Methods in Hospital Finance Management

Oğuz CECE[1]
Mehmet GENÇTÜRK[2]

ABSTRACT

Hospital information management systems (HIMS) were managed using paper-based systems with individual efforts during the pre-computer era. Today, parallel to technological developments, tasks are carried out digitally in an electronic environment. HIMS software typically includes modules such as patient registration and appointment follow-up, clinical and medical records, radiology, laboratory, drug management, billing, reporting, and hospital management. Accounting records are processed in the finance management submodule within the hospital management module. Artificial intelligence models used in various sectors for financial estimation in hospital finance management have been found to be worth researching, given the benefits they offer to the hospital's financial management. Financial data from private hospitals traded on the stock exchange between 2009-2023 were used in the study. A total of 97 financial reports from 5 different private hospitals and 776 raw data obtained from these reports constitute the dataset for the study. "Net Profit Margin" has been estimated for the data set. The most reliable and closest-to-reality algorithm was determined by conducting five different algorithm trials in the PHYTON programming language.

1 Süleyman Demirel University, Institute of Social Sciences, Health Institutions Management, Isparta, Türkiye
2 Süleyman Demirel University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Isparta, Türkiye
* Corresponding author: Oğuz CECE, oguzcece@gmail.com

The most successful result was obtained with the Random Forest algorithm. The correlation coefficient between Random Forest prediction values and actual values was over 0.90 for both training and testing phases. This shows that there is a strong positive relationship between the predicted values and the actual values. It has been observed that hospitals can use Random Forest to make this estimation when they want to predict financial data for future periods.
Keywords: Financial Forecasting, Financial Forecasting Methods, Health Information Management System, Machine Learning

INTRODUCTION

Financial forecasting in the healthcare industry involves utilizing financial data and analysis techniques to anticipate future financial outcomes and trends, specifically for healthcare organizations such as hospitals, clinics, pharmaceutical companies, and health insurance providers. It aims to support decision-making, budgeting, resource allocation, and financial planning in the healthcare industry (Davenport and Kalakota, 2020).

Regulatory factors in the health sector encompass a range of complex dynamics, such as technological advancements, demographic changes, and evolving health policies. These factors may introduce additional uncertainties and challenges to financial forecasts in the health sector. However, financial forecasting can serve as a policy tool for strategic planning, resource management, and financial decision-making in healthcare organizations (Kaya et al., 2019).

Financial forecasting in healthcare involves a systematic approach that combines historical data, industry trends, and relevant factors to predict future financial outcomes.

Data Collection: Relevant financial data is collected from various sources within the healthcare organization. This includes historical financial statements, billing records, patient volumes, reimbursement rates, and other financial metrics. Additionally, it may include external data sources, such as industry reports, government data, and economic indicators (Koçyiğit et al., 2022).

Determining the Main Factors: The main factors that affect the financial performance of the health institution are determined. These factors may include patient volumes, service mix, reimbursement rates, pricing, staffing levels, regulatory changes, and market dynamics.

Define the Forecast Period: The time frame for the forecast is determined, whether it is short-term (e.g., monthly, or quarterly) or long-term (e.g., annual, or multi-year). Estimation time varies depending on the specific needs and goals of the healthcare institution (Özdemir and Bilgin, 2021).

Choosing Forecasting Methods: The appropriate estimation methods are selected based on the available data and the characteristics of the predicted financial variables. Standard forecasting techniques used in healthcare include time series analysis, regression analysis, trend analysis, and predictive modeling. In addition, healthcare-specific models, such as patient volume estimation or reimbursement models, can be used.

Building Models and Analyzing Data: Selected forecasting methods are applied to the collected data. This includes analyzing historical patterns, identifying trends, and building mathematical models that capture relationships between critical factors and financial results. Statistical software and data analysis tools can be used in this process (AbdelSalam et al., 2022).

Validating and Refining Models: Validate forecasting models by comparing predicted results with historical data. Evaluate the accuracy of the predictions and make adjustments or improvements to the models if necessary. This iterative process helps to increase the accuracy and reliability of the estimates.

Sensitivity Analysis: Sensitivity analysis is performed to assess the impact of changes in fundamental assumptions or variables on financial projections.

Interpreting and Reporting Results: The financial results are predicted, analyzed and interpreted for the health institution. Results are communicated to relevant stakeholders, such as management, finance teams, and decision-makers, to provide information on strategic planning, budgeting, and resource allocation (Abdullah, 2021; Ural et al., 2015).

Monitoring and Updating: The financial performance is regularly monitored according to the predicted results. Estimates are revised as new data becomes available or as significant changes occur in healthcare (Abdullah, 2021).

Financial forecasting in healthcare is subject to uncertainties and external factors, such as regulatory changes, patient demographics, and changes in healthcare policies. Therefore, ongoing monitoring and flexibility in the forecasting process are essential for adapting to changing conditions and increasing the accuracy of forecasts.

Financial Forecasting Methods in the Literature

Financial forecasting is a well-studied field in finance and economics, and a substantial body of literature explores various aspects of economic forecasting. Some of the commonly studied topics and techniques in financial forecasting research include:

Time Series Analysis: Time series analysis is a fundamental technique used in financial forecasting. It involves analyzing historical data to identify patterns, trends, and seasonality of economic variables. Techniques such as Autoregressive Integrated Moving Average (ARIMA) models, exponential smoothing, and state space models are widely used in time series analysis (Agirbas et al., 2018).

Machine Learning and Artificial Intelligence: With advancements in machine learning and artificial intelligence, these techniques have gained popularity in financial forecasting research. Methods such as neural networks, support vector machines (SVM), random forests, and ensemble models are applied to predict stock prices, exchange rates, and other financial variables (Anal et al., 2010).

Financial Statement Analysis: Financial statement analysis techniques are used to predict financial performance indicators, such as a company's sales revenues, earnings, and cash flows. These techniques often include ratio analysis, joint dimensional analysis, and regression analysis using financial statement data.

Case Studies: Case studies examine the impact of specific events on financial variables, such as mergers and acquisitions, initial public offerings (IPOs), or regulatory changes. These studies assess the short- and long-term effects of events on stock prices, trading volumes, and other financial outcomes (Akkaya et al., 2009).

Volatility Estimation: Volatility estimation involves estimating the volatility and risk of financial assets. Techniques such as Generalized Autoregressive Conditional Variance Variance (GARCH) and stochastic volatility models are widely used to predict volatility in financial markets (Ural et al., 2015).

Predictive Analytics in Healthcare Finance: In the healthcare industry, predictive analytics is increasingly applied to financial forecasting. This includes analyzing healthcare usage data, reimbursement rates, patient demo-

graphics, and other factors to estimate healthcare costs, revenue streams, and budgetary requirements.

Forecasting in Risk Management: Financial forecasting is an integral part of risk management in multiple industries such as, including insurance, banking, and investment management. Studies often focus on estimating credit risk, default probabilities, market risks, and other financial variables related to risks (Akkaya et al., 2009).

Prediction Combination and Model Selection: Forecast combination techniques involve combining multiple individual predictions to increase accuracy and reduce prediction errors. Model selection methods aim to determine the most appropriate forecasting model or style for a particular financial variable or time series data.

Forecast Evaluation and Accuracy Metrics: Evaluating forecast accuracy is crucial in financial forecasting research. Various metrics such as mean absolute error (MAE), square mean error (MSE), square mean error (RMSE), and estimation-encompassing tests are used to evaluate the performance of forecasting models and compare different forecasting techniques (Fletcher, 2012).

Financial forecasting is an active field of study that constantly explores new techniques, improves existing methods, and addresses the challenges and complexities of estimating financial variables in various contexts (Abdullah, 2021).

The 2013 study by Soyiri et al. provides a comprehensive review of financial forecasting methods used in healthcare institutions. Outlines the approaches and techniques used to forecast financial performance and their strengths and limitations (Soyiri and Reidpath, 2013).

The study by Lee and Miller in 2022 focuses on estimating health expenditures and examines various models and methods used to evaluate health costs. Discusses challenges and opportunities in healthcare expenditure estimation and suggests potential areas for future research (Lee and Miller, 2002).

The 2012 study by Astolfi et al. reviews the literature on financial forecasting in hospitals, highlighting the different approaches and models used to predict economic outcomes (Astolfi et al., 2012).

Regression analysis of methods used to make financial forecasts, time estimates, etc., appears to be. However, it is seen that the methods used to make financial estimations in the field of health are more limited.

METHODOLOGY

Purpose and Importance of the Research

It has been observed that there is mostly no financial forecasting module among the Hospital Information Management Systems (HIMS) modules used by public and private hospitals. It has been evaluated that integrating such a module is into the HIMS, the financial situation of the health business will become predictable, and it will be protected from potential financial losses in various industries such as the literature, banking, manufacturing, construction, defense, etc. Although there are similar studies for different sectors, there are few scientific studies in hospital management.

Scope of the Study and Data Set

Private hospitals that are actively or passively involved in the healthcare sector and are members of the Public Disclosure Platform (KAP) are included in the scope of the study. The financial data of the hospitals in question since 2009 constitute dataset for the study. In the study's dataset, the relevant items in the financial reports of the hospitals in Table 1 were used as raw data. These data were first processed by calculating financial ratios and then evaluated as learning data in the financial estimation process using a machine learning method.

Table 1: Data Used in the Study

SHARE CODE	COMPANY TITLE	FINANCIAL REPORT STATEMENT	
		Date Range	Number
ACIBD	ACIBADEM SAĞLIK HİZMETLERİ VE TİCARET A.Ş.	2009-2012	13
LKMNH	LOKMAN HEKİM ENGÜRÜSAĞ SAĞLIK TURİZM EĞİTİM HİZMETLERİ VE İNŞAAT TAAHHÜT A.Ş.	2011-2023	50
MPARK	MLP SAĞLIK HİZMETLERİ A.Ş.	2017-2023	24
EGEPO	NASMED ÖZEL SAĞLIK HİZMETLERİ TİCARET A.Ş.	2021-2023	7
TNZTP	TAPDİ OKSİJEN ÖZEL SAĞLIK VE EĞİTİM HİZMETLERİ SANAYİ TİCARET A.Ş.	2023	2
DNYGZ	DÜNYA GÖZ HASTANESİ SANAYİ VE TİCARET A.Ş.	2015	1

Limitations of the Study

The study utilized financial data shared by private hospitals that are publicly traded in Borsa Istanbul and operate in the health sector through the Public Disclosure Platform.

Variables of the Study

The dependent variables of the research are the "gross profit margin, net profit margin, current ratio, debt-capital ratio, and acid-test ratio," which are the most frequently used financial ratios to evaluate the financial condition of an enterprise. The independent variable of the research is "time," as it allows for the estimation of past values.

Parameters Used in the Research

Table 2: Parameters Used in the Research

Financial Ratios	Income Statement and Balance Sheet Items
	Sales Revenue
Gross Margin	Cost of sales
Net Profit Margin	Net profit
Current rate	Current Assets
Debt-Capital Ratio	Stocks
Acid-Test Ratio	Short Term Debts
	Long Term Debts
	Capital

Methods Used in the Research

Python is widely used in artificial intelligence because of its advantages and features. It supports high-level programming and offers an interpreted language, which makes writing and understanding code easy. Python has pre-built libraries like Numpy and Scipy for scientific calculations and scikit-learn for machine learning tasks. These libraries provide a solid foundation for AI development in Python (Polimis et al., 2017).

Python libraries are pre-written collections of code that provide developers with various functions and tools. These libraries contain modules, functions,

and classes that can be imported into Python programs to perform specific tasks without having to write code from scratch. It covers various fields such as data processing, scientific computing, web development, machine learning, natural language processing, and more (Lakshmi and Scholar, 2016).

Python libraries for financial analysis and forecasting include Pandas, NumPy, and Matplotlib. These libraries offer functions for data manipulation, numerical calculations, and visualization. Additionally, libraries such as Scikit-learn and Stats models provide powerful statistical modeling and machine-learning capabilities that can be applied to financial data (Wasserbacher and Spindler, 2022). In this study, Pandas, NumPy, and Scikit-learn libraries were used.

Stages of Research

1. Data Preprocessing

Hospital data was preprocessed to prepare it for data analysis. PYTHON version 3.11.4 was used for research, and data processing was performed using nodes (Number Filter).

Scanning and Marking (Number Filter): In this step, the text is scanned to remove extra spaces and identify numbers that will not be used.

After this process, the missing five values in the data were determined and filled by taking the average of the relevant columns.

2. Calling Libraries

Pandas (pd), Numpy (Np), and Sklearn libraries are used in Python for financial data analysis and forecasting. Thus, the values are fixed within a range of minimum-maximum, and many useful functions are activated.

3. Processing of Data for Prediction

The data to be estimated is entered as input. The target variable was created by delaying the period in the selected column value. The last period data of each hospital was deleted due to the delay in the target variable. The data has been selected for the feature selection process. The data is partitioned as training and test data. The accuracy and usability of the dataset have been tested. The categorical variables of Hospital and Period information were converted

into numerical data. It was checked whether there were any empty values in the data set, and then the existiopenpty values were filled with the average value of the relevant columns.

4. Feature Selection

Feature selection methods have been applied to improve the algorithm results. The mutual_info_regression algorithm is used in the method. "Mutual_information" is a non-parametric measure revealing two random variables' dependency structure. This value is 0 if and only if these random variables are independent. In other cases, the mutual information value increases with the increase in dependency (Polimis et al., 2017). In the trials conducted with this algorithm, it was observed that eight features yielded the best results. These data are;

- Hospital
- Debt-Capital Ratio
- Net Profit
- Current Assets
- Stocks
- DVB
- UVB

It is equity.

The selected features divide the data into two training and test data. Here, 0.30 is used as the training data ratio.

5. Machine Learning Algorithms Used in Research

Within the scope of the research, Random Forest, Linear Regression, Multilayer Perception, Nearest Neighbors Regression, and Decision Tree Regression methods were used. Hyperparameter tuning was not performed using k-fold cross-validation because the number of data for each way was small.

It aims to determine the Net Profit Margin of the next period using the data from a specific hospital in the model. For this reason, a new variable with a lag according to the Net Profit Margin column was created in the data structure, and the idle observations were deleted. The number of data points examined is 97 pieces. With the implementation of the delay in the data, the total number of data decreased to 91.

Random Forest Algorithm: Random forest is a method for estimating many community learning tasks. Prediction variability can indicate how effective the training set is in generating the observed random forest and provide additional information about prediction accuracy. Forest-confidence-interval is a Python module for calculating variance and adding confidence intervals to scikit-learn Random forest regression or classification objects (Polimis et al., 2017).

Linear Regression Algorithm: Linear regression is a popular machine learning algorithm that predicts continuous values based on input variables. It is a statistical technique to model the relationship between a dependent variable and one or more independent variables. It aims to find the most appropriate line that represents the linear relationship between the variables. Its primary purpose is to minimize the difference between the estimated and actual values. It is widely used for forecasting and prediction tasks. It is helpful to understand the relationship between variables, identify trends, make predictions, and evaluate the effect of changes in independent variables on the dependent variable (Lakshmi and Scholar, 2016).

Multilayer Perception Algorithm: An artificial neural network commonly used for classification tasks. It consists of many interconnected layers of artificial neurons, where each neuron is a computational unit that processes and transmits information to the next layer. MLP uses a feed-forward mechanism, in which information flows from the input layer to the output layer through hidden layers. It can learn complex patterns and make predictions based on known patterns. MLP is trained using a process called backpropagation, in which the network adjusts its weights and biases to minimize the difference between the predicted and actual outputs (Car et al., 2020).

Nearest Neighbors Regression Algorithm: A machine learning algorithm for classification and regression tasks. It works by finding the closest data points to a test data point in the training dataset. It can be used for both supervised and unsupervised learning. Label the majority the nearest neighbors in the controlled environment to the test data point, in order to find the most similar data points for clustering or anomaly detection purposes in the uncontrolled environment (Samruddhi and Ashok Kumar, 2020).

Decision Tree Regression Algorithm: It is a machine learning algorithm for regression tasks. It works by recursively partitioning the training data based on feature values to create a tree-like model. Each inner node of the tree represents a decision based on a feature, while each leaf node represents the predicted output value. The algorithm selects the best feature and split point that minimizes overall mean square error (MSE) during training. It provides a useful way to capture nonlinear relationships between features and target variables (Pilnenskiy and Smetannikov, 2020).

RESULTS

Correlation coefficient, mean absolute error (MAE), mean square error (RMSE), relative mean errors (RAE), and relative error squares (RRSE), Mean Absolute Percentage Error (MAPE), R^2 values were used to compare the performances of the methods. The formulas used are given below.

- $MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$
- $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
- $RAE = \sum_{i=1}^{n}\frac{|y_i-\hat{y}_i|}{|y_i-\bar{y}|}$
- $RRSE = \sqrt{\sum_{i=1}^{n}\frac{(y_i-\hat{y}_i)^2}{(y_i-\bar{y})^2}}$
- $MAPE = \frac{100/\%}{n}\sum_{i=1}^{n}\frac{|y_i-\hat{y}_i|}{|y_i|}$
- $R^2 = \sum_{i=1}^{n}\frac{(y_i-\hat{y}_i)^2}{(y_i-\bar{y})^2}$

Here, is the observation value (the target value in the test or training data), does the model predict the target value, and is the mean of the observation values. In machine learning algorithms, MAE, RMSE, RAE, RRSE, MAPE, R^2 values are expected to be as small as possible.

Machine learning algorithm results are given in Table 3 and Table 4.

Table 3: Prediction Data Test Results

Algorithm	Prediction						
	Correlation Coefficient	Mean Absolute Error (MAE)	Mean Error Squares (RMSE)	Relative Mean Errors (RAE)	Relative Error Squares (RRSE)	Mean Absolute Percentage Error (MAPE)	R^2
Random Forest	0.954	22.744.521	75.374.108	0.214	0.299	0.456	0.887
Linear Regression	0.894	82.867.890	112.600.301	0.779	0.447	1.357	0.750
Multilayer Perception	0.998	9.576.807	14.854.148	0.090	0.059	0.827	0.997
Nearest Neighbors	0.843	48.421.537	148.780.191	0.455	0.591	0.838	0.053
Decision Tree Regression	1	0	0	0	0	0	1

Table 4: Test Data Test Results

Algorithm	Test						
	Correlation Coefficient	Mean Absolute Error (MAE)	Mean Error Squares (RMSE)	Relative Mean Errors (RAE)	Relative Error Squares (RRSE)	Mean Absolute Percentage Error (MAPE)	R^2
Random Forest	0.935	29.130.646	46.332.685	0.445	0.370	1.050	0.710
Linear Regression	0.894	59.510.417	73.111.997	0.910	0.583	1.945	0.784
Multilayer Perception	0.886	36.499.620	63.937.853	0.558	0.510	2.037	0.606
Nearest Neighbors	0.935	25.482.721	47.434.492	0.390	0.378	1.011	0.861
Decision Tree Regression	0.635	57.374.939	102.752.293	0.877	0.820	1.024	0.935

As can be seen in Table 3 and Table 4, the Random Forest method gave the best results.

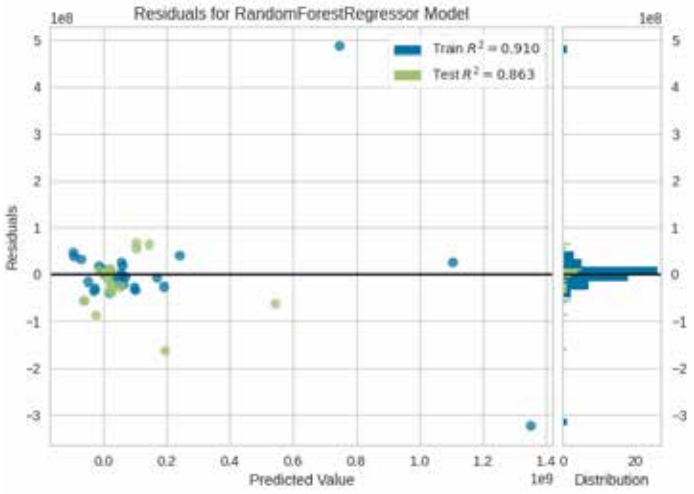

Figure 1. Distribution of residuals of values estimated by Random Forest Regression

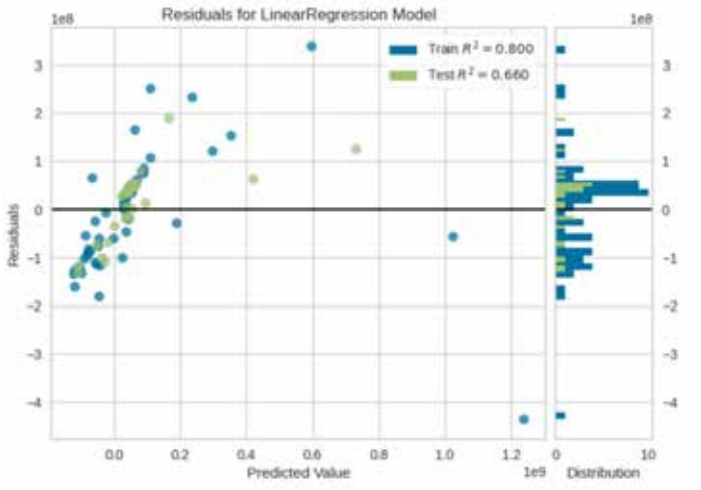

Figure 2. Distribution of residuals of values estimated by Linear Regression

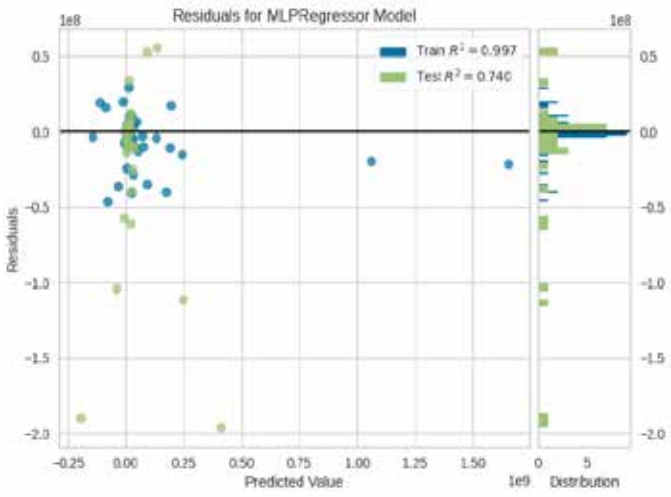

Figure 3. Distribution of residuals of values estimated by Multilayer Perception Regression

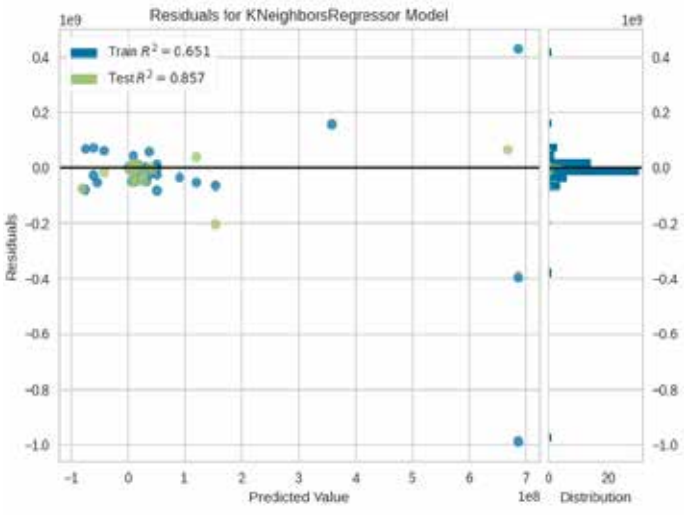

Figure 4. Distribution of residuals of values estimated by K-nearest Regression

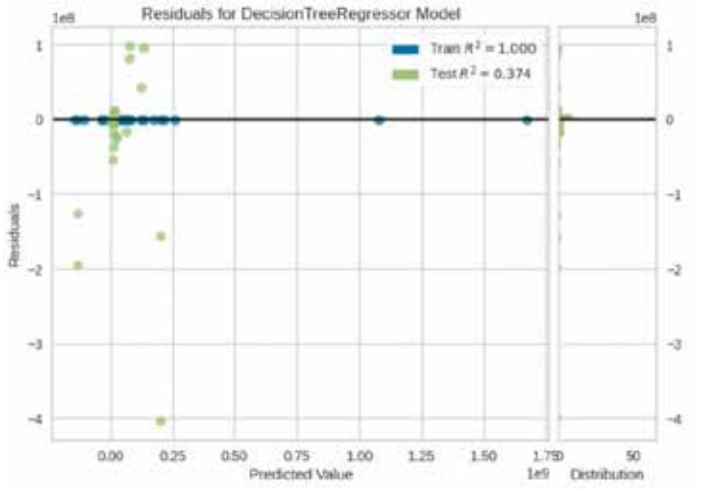

Figure 5. Distribution of residuals of values estimated by Decision Tree Regression

The residuals for all methods exhibit a near-symmetric scatter around zero during both the training and testing phases. This indicates that the machine learning methods used predict the actual value unbiasedly. In cases where the estimated value is too large, it is observed that the residual value is too large for at most two points, and all the residual values are close to each other and remain small.

In this study, the net profit margin of hospitals was determined using machine learning techniques over the data set created from hospital data. Regarding the net profit margin, the estimation of the financial data and ratios in the dataset for the next period is made using some algorithms. The results obtained from five different machine learning algorithms (Random Forest, Linear Regression, Multilayer Perception, Nearest Neighbors, Decision Tree Regression) are shown in Figure 1-5. As can be seen, the most successful results were obtained with the Random Forest algorithm. Since hospitals aim to predict financial data for future periods, this estimation can be made using Random Forest.

DISCUSSIONS AND CONCLUSIONS

A literature review and several studies have been conducted on the utilization of machine learning methods in the field of artificial intelligence. Generally, studies on banking, stocks, the stock market, and the economy have been widely encountered. However, the number of studies in the health field is quite limited.

No analysis has been found that directly estimates profit margins in healthcare institutions. Although health institutions seem like a business, they provide services in a different manner than a typical company and are traded using different accounting items. For this reason, directly comparing an enterprise's data with data from the health institutions will not yield an accurate result.

In this study, an estimation was made only on "net profit," but there are plans to make estimations on other items in future studies. In addition, the study is believed to contribute to literature as it is a rare study in its field.

Ethical Approval: Ethical approval was not required as the study was "reviewed."

Authors' Contributions: Oğuz Cece (%80), Mehmet Gençtürk (%20)

Funding and Acknowledgement: There is no financial support for the study.

Conflict of Interest Statement: There is no conflict of interest.

REFERENCES

AbdelSalam, K., Hany, S., El Hady, D., Essam, M., Mahmoud, O., Mohammed, M., Samir, A., Magdy, A. (2022). Robo-nurse healthcare complete system using artificial intelligence. In International Conference on Advanced Machine Learning Technologies and Applications. Springer International Publishing.

Abdullah, M. (2021). The implication of machine learning for financial solvency prediction: An empirical analysis on public listed companies of Bangladesh. Journal of Asian Business and Economic Studies, 28, 303-320.

Agirbas, İ., Turgut, M., Işıkçelik, F. (2018). Özel hastanelerde finansal analiz. III. Uluslararası Al-Farabi Sosyal Bilimler Kongresi, November.

Akkaya, G. C., Demireli, E., Yakut, Ü. H. (2009). İşletmelerde finansal başarısızlık tahminlemesi: Yapay sinir ağları modeli ile İMKB üzerine bir uygulama. Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 10, 187-216.

Sarıkale, H. İlter, B. (2018). Finansal analiz sürecindeki temel mali tablo düzeltmelerinin oran analizine etkisi. Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi, 20, 165-184.

Astolfi, R., Lorenzoni, L., Oderkirk, J. (2012). A comparative analysis of health forecasting methods. OECD Health Working Papers, 59.

Car, Z., Baressi Šegota, S., Anđelić, N., Lorencin, I., Mrzljak, V. (2020). Modeling the spread of COVID-19 infection using a multilayer perceptron. Computational and Mathematical Methods in Medicine. 2020.

Davenport, T., Kalakota, R. (2020). The potential for artificial intelligence in healthcare. SSRN Electronic Journal, 6, 94-98.

Fletcher, T. S. (2012). Machine learning for financial market prediction. UCL (University College London), Computer Science, London.

Kaya, U., Yılmaz, A., Dikmen, Y. (2019). Sağlık alanında kullanılan derin öğrenme yöntemleri. European Journal of Science and Technology, 16, 792-808.

Koçyiğit, S., Bıyık, E., Ertaş, Ş. (2022). Özel bir sağlık işletmesinin finansal performansının trend analizi ile değerlendirilmesi. Abant Sosyal Bilimler Dergisi, 22, 165-180.

Lakshmi, J. V. N. (2016). Stochastic gradient descent using linear regression with python. International Journal on Advanced Engineering Research and Applications, 2, 519-524.

Lee, R., Miller, T. (2002). An approach to forecasting health expenditures, with application to the U.S. Medicare system. Health Services Research, 37, 1365-1386.

Özdemir, L., Bilgin, A. (2021). The use of artificial intelligence in health and ethical problems. Sağlık ve Hemşirelik Yönetimi Dergisi, 8, 439-445.

Pilnenskiy, N., Smetannikov, I. (2020). Feature selection algorithms as one of the Python data analytical tools. Future Internet, 12, 54.

Polimis, K., Rokem, A., Hazelton, B. (2017). Confidence intervals for random forests in Python. The Journal of Open Source Software, 2, 124.

Samruddhi, K., Ashok Kumar, R. (2020). Used car price prediction using K-Nearest Neighbor Based Model. International Journal of Innovative Research in Applied Sciences and Engineering, 4, 629-632.

Soyiri, I. N., Reidpath, D. D. (2013). An overview of health forecasting. Environmental Health and Preventive Medicine, 18, 1-9.

Ural, K., Gürarda, Ş., Önemli, M. B. (2015). Lojistik regresyon modeli ile finansal başarısızlık tahminlemesi: Borsa İstanbul'da faaliyet gösteren gıda, içki ve tütün şirketlerinde uygulama. Muhasebe ve Finansman Dergisi, 67, 85-100.

Wasserbacher, H., Spindler, M. (2022). Machine learning for financial forecasting, planning and analysis: Recent developments and pitfalls. Digital Finance, 4, 63-88.