Mühendis ve Makina / Engineer and Machinery 65, 716, 432-461, 2024

Mühendis ve Makina / Engineer and Machinery
https://dergipark.org.tr/tr/pub/muhendismakina

THE IMPACT OF COVID-19 ON THE TECHNOLOGY SECTOR: THE CASE OF THE TURKISH CONSULTANCY COMPANY

Eda GÖZÜTOK[1], İlayda ÜLKÜ[2*]

[1] Industrial Engineering Department, Istanbul Kultur University, Istanbul, Turkey
ORCID No : http://orcid.org/0000-0001-9389-0564

[2] Industrial Engineering Department, Istanbul Kultur University, Istanbul, Turkey
ORCID No : http://orcid.org/0000-0003-0464-7007

Keywords

COVID-19, machine learning, regression models, KNIME, forecasting

Abstract

The COVID-19 pandemic has caused unprecedented changes in the global economy and society, with many studies attempting to understand the impact of the virus on different countries and industries. This study focuses on the effects of COVID-19 on a consulting company that specializes in technology services. By analyzing the company's sales data for the five-year period before the pandemic, and using machine learning techniques via the KNIME platform, the study aims to predict the sales data for the COVID-19 period. Three different regression models - linear, gradient boosting, and random forest - were used to make these predictions, and the models were compared based on their coefficient of determination (R2) to determine which model performed best. The chosen model was then used to interpret the impact of COVID-19 on the company. The findings of the study provide insights into how COVID-19 has affected the consulting company. The chosen model showed that the pandemic had a significant negative impact on the company's sales, with a sharp decline in the second quarter of 2020. However, the company was able to recover some of its losses by the fourth quarter of the year. The study also highlights the importance of using machine learning techniques to predict future sales data during unpredictable events such as the COVID-19 pandemic. Overall, this study sheds light on the impact of COVID-19 on a technology consulting company and demonstrates the importance of using data analysis and machine learning techniques to make predictions and interpret the effects of significant events on business operations.

* i.karabulut@iku.edu.tr
doi : 10.46399/muhendismakina.1362765

COVİD-19'UN TEKNOLOJİ SEKTÖRÜNE ETKİSİ: BİR DANIŞMANLIK ŞİRKETİ ÖRNEĞİ

Anahtar Kelimeler

COVID-19, makine öğrenimi, regresyon modelleri, KNIME, tahmin

Öz

COVID-19 salgını, dünya genelinde benzeri görülmemiş bir etki yaratmıştır; hem ekonomik hem de sosyal açıdan büyük değişimlere neden olmuştur. Bu salgının farklı ülkeler ve endüstriler üzerindeki etkisini anlamak için birçok çalışma yapılmıştır. Bu çalışma ise, teknoloji hizmetleri alanında uzmanlaşmış bir danışmanlık firmasının COVID-19'un etkilerini analiz etmektedir. Çalışma kapsamında, şirketin pandemi öncesine ait beş yıllık satış verileri incelenmiş ve bu veriler üzerinden KNIME platformu aracılığıyla makine öğrenimi teknikleri kullanılarak COVID-19 dönemine ilişkin satış verilerinin tahmini amaçlanmıştır. Bu tahminler için üç farklı regresyon modeli (doğrusal, gradyan artırma ve random forest) kullanılmış ve bu modellerin performansı, belirleme katsayılarına (R2) göre karşılaştırılmıştır. Seçilen model daha sonra, COVID-19'un şirket üzerindeki etkilerini yorumlamak için kullanılmıştır. Araştırmanın bulguları, COVID-19'un danışmanlık şirketini nasıl etkilediğine dair derinlemesine içgörü sağlamaktadır. Seçilen model, salgının şirketin satışları üzerinde önemli bir olumsuz etki yarattığını ve özellikle 2020 yılının ikinci çeyreğinde sert bir düşüş yaşandığını göstermiştir. Ancak, şirket yılın dördüncü çeyreğine doğru kayıplarının bir kısmını telafi etmeyi başarmıştır. Çalışma ayrıca, COVID-19 salgını gibi öngörülemeyen olaylar sırasında gelecekteki satış verilerini tahmin etmek için makine öğrenimi tekniklerinin önemini vurgulamaktadır. Genel olarak, bu çalışma, bir teknoloji danışmanlık şirketi üzerinde COVID-19'un etkilerine dair önemli bilgiler sunmakta ve iş operasyonları üzerindeki etkileri tahmin etmek ve yorumlamak için veri analizi ve makine öğrenimi tekniklerinin değerini ortaya koymaktadır.

Araştırma Makalesi			Research Article		
Başvuru Tarihi	:	19.09.2023	Submission Date	:	19.09.2023
Kabul Tarihi	:	29.04.2024	Accepted Date	:	29.04.2024

1. Introduction

The coronavirus pandemic, also known as the COVID-19 pandemic was first shown up in Wuhan, Hubei Province on 31st December 2019. COVID-19 was confirmed as a pandemic on 2020, March 11 by World Health Organization. Before the declaration as a pandemic, it spread to 18 countries from China (WHO 2022). The first case of Turkey's COVID-19 was discovered on March 11, 2020. More than 5 million people were infected after the first incidence, and nearly 100,000 people died until the beginning of May (Anon n.d.).

According to World Health Organization (WHO), more than 520 million people were affected by COVID-19 and also more than 6 million people passed away until May 2022 (WHO, 2022).

The COVID-19 pandemic affects the world socially, physiologically, and also economically. Several papers, studies, and, articles were published to explain those effects with different analyses from different angles. Out of 193 countries' economics, 167 economies saw a negative rate in 2020. Almost 79 percent of the world economy accounted for by these 167 economies. China was the worst affected country in the first quarter of 2020 with an economic decrease of 9,7 percent. Iceland and France followed China on the list. India, United Kingdom, and Spain were the most damaged countries during the second quarter of 2020 and the world economy shrunk 32,9 percentages. In the third quarter of 2020, 33,1 percent growth came amid a resurgence in consumer activity. In the last quarter of 2020, the economy maintained growing by 4 percent even with a slight rate compared to the third quarter. Total growth in 2020 was 3,6 percent and only two countries which are China and Turkey saw a positive economic rate at end of the 2020 (Organisation for Economic Co-operation and Development, 2021).

COVID-19 appeared suddenly and changed the expectations of companies. Forecasting is substantial to plan production, efficiently allocate resources and workforces, marketing strategies for companies. Machine learning (ML) is the study of computer algorithms that can learn and develop on their own through experience and data (Mitchell, 1997). Forecasting is one of the applications of machine learning.

This research concentrated to analyze the impacts of COVID-19 in the case of a Turkish consultancy company in technology sector with machine learning. With seven years of sales data collected made forecasting using different regression model from the Konstanz Information Miner (KNIME) platform. The regression models are gradient boosting regression (GBR), linear regression (LR) and random forest regression (RFR).

Section 1 contains general information about COVID-19 and its economic impacts. The remaining parts of this study are as follows. Section 2 discusses the review of previous studies published which provides a summary of similar studies about the economic impacts of COVID-19 in different sectors around the world. Machine learning, forecasting methodologies, and the KNIME tool which is used for modeling information are given in Section 3. Section 4 explains preparing sales data, and forecasting methods step-by-step. Numerical results and comments are given in Section 5. Finally, Section 6 highlighted and presented results and findings in this case.

2. Literature Review

After COVID-19 appeared, more than 50.000 papers were published on the Science Direct website. Pandemic affected various sectors like tourism, accommodation, education, aviation, production, automotive, and food. More than one-fifth of papers are related to the economic effect of COVID-19 from different angles (ScienceDirect, n.d.). According to Organization for Economic Co-operation and Development (OECD), countries' economies affected adversely in the beginning of the pandemic. Recovery started however lower-income economies still needs time (Organisation for Economic Co-operation and Development, 2021). Table 1 summarizes the main studies to understand the effect of COVID-19 worldwide except in the technology sector.

Table 1. Significant Studies about COVID-19 on Sectoral Basis

Author(s)	**Country**	**Sector**	**Impact**
Blazy et al., 2021	Caribbean	Agriculture	COVID-19 affected strongly the agricultural system such as a decrease in revenue, production decreasing.
Vall Castelló & Lopez Casasnovas, 2021	Spain	Supermarket	There was stockpiling especially in the first week of lockdown, with fear of pandemic was decreased sales rates.
Wen et al., 2021	China	Electric Vehicle Industry	There was a negative effect on sales in the short term.
Wieczorek-Kosmala, 2021	4 countries	Hospitality	There were financial slack holdings during COVID-19.

Hayakawa & Mukunoki, 2021	34 Different Countries	International Trade	There were negative effects on leather, footwear products, and transport kits especially in the fourth and fifth months of 2020 whereas there were effects positively in industries manufacturing medical products.
Maneenop & Kotcharin, 2020	9 Different Countries	Airline	Australia, Canada, the United Kingdom, and United State showed the worst reaction to proclamation of COVID-19 as a global pandemic.
Fu & Shen, 2020	China	Energy Industry	COVID-19 damaged the productivity of energy companies, and companies could not effort their fixed costs and expenses, as a result, an important damaging effect was found on the development of energy companies.
Menhat et al., 2021	Malaysia	Marine (Shipping, fisheries, marine tourism, energy)	Marine tourism was the most affected sector in Malaysia because the government considered as a non-essential sector. Shipping was affected lightly compared to other sectors.
He et al., 2020	China	Various sectors	Electric energy, heating, mining, and environmental industry have been impacted negatively during the COVID-19. Even so, some industries showed strong immunity for instance production, information technology, and healthcare in China.

Wang et al., 2020	China	Insurance	There were negative effects on property and personal insurance, and personal insurance was damaged more than that property insurance from the angle of insurance density and depth.

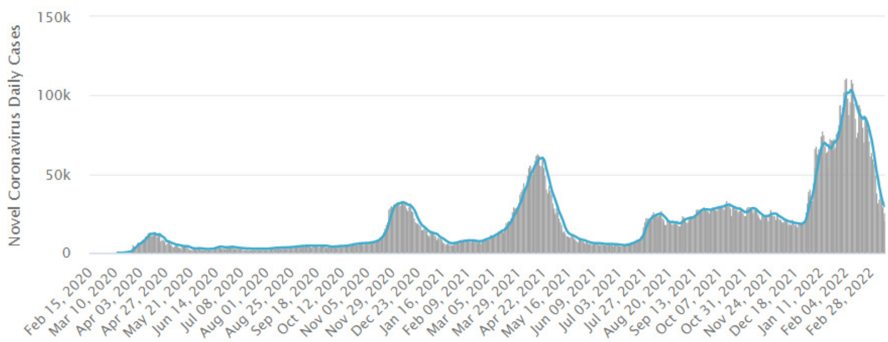

Figure 1. Daily Cases of Turkey

This study will be a research that will focus on analyzing the impacts of COVID-19 on the technology sector in a Turkish company. The first case of COVID-19 appeared on 11th March 2020 in Turkey. After the first case, more than 5 million people were infected and almost 100.000 people died until the beginning of the May (Anon n.d.). Figure 1 shows daily cases of Turkey (Worldometers n.d.).

Note. Daily COVID-19 cases in Turkey. From Worldometers, COVID Live - Coronavirus Statistics – by Worldometers, n.d. (Worldometers n.d.).

- Turkey Government took various precautions during this time. Important dates are;
- 1st February 2020- All the flights from China were stopped.
- 29th February 2020- All flights to South Korea, Italy, and Iraqi were stopped (The Economist, 2020).
- 11th March 2020- First COVID-19 case was announced.
- 12th March 2020- Schools was closed from 16th March (Kandemir, 2020).
- 16th March 2020- Indoor activities such as businesses and places of worship were halted (Cantekin, 2020).

- 21st March 2020- A total curfew was announced for people who are elder 65 and who have immune system diseases such as asthmatics, cardiovascular problem, high blood pressure (Kandemir, 2020).
- 10th April 2020- A curfew was announced to everyone during weekends (Ghosh, 2020).
- 4th May 2020- Return to normal life rules explained (Anon n.d.).
- 1st June 2020- Public spaces were opened and domestic flights were resumed (The Economist, 2020).
- 26th November 2020- A new curfew was announced for people elder than 65 and younger than 20. Also, indoor activities such as businesses and places of worship were halted again (Overseas Security Advisory Council, 2020).
- 14th April 2021- A two-week partial closure was declared (Anon n.d.).
- 26th April 2021- New full lockdown announced from 29th April to 17th May (The UN Refugee Agency, 2021).
- 17th May 2021- Partial closure was declared until 1st July (Anon n.d.).
- 1st July 2021- The curfew was completely lifted (Anon n.d.).

In compliance with data from the Presidency of Strategy and Budget, in the second quarter of 2020, most sectors were affected negatively by COVID-19 such as industry, construction, services, export, and import. The most affected sector in Turkey became exporting with a 15,4 percent decrease(T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı n.d.) Exporting became a general problem in the world because of restrictions on air flights and curfews. During the pandemic, electronic microchip demand was increasing around the world. In Turkey, the automotive sector was affected negatively as a consequence of an interruption in production in the second quarter (Sabah, 2020).

The consultancy sector was affected adversely by COVID-19 in the first quarter of 2020. The consultancy sector continues to grow over the last 12 years. In 2019, the value of the sector reached $160 billion. Source Global Research gathered the opinions of hundreds of consulting organizations from around the world and concluded that the consultant sector could lose $30 billion in value by 2020. The United States has a major percentage in the consultancy sector and according to the study, 1 percent decrease is expected. Study shows Europe will be the worst affected by COVID-19 (Consultancy n.d.). While business services, healthcare, energy and resources, manufacturing, and the public sector were in the higher risk group, financial services, pharma, retail and, technology and telecoms were in the lower-risk group (Consultancy n.d.).

Information technology (IT) consulting, often known as technology consulting, refers to services that assist clients in determining how to best employ information technology and digital to meet their business objectives. Remote working and curfews affected the demand for information and communication technologies (ICT) demands in a positive way (Taser, Aydin, Torgaloz and Rofcanin 2022). People need to do various things online such as shopping, working, studying and ICT makes them possible. That's why companies investing in digitalization for adopting this period.

A survey about people's grocery shopping choices shows there is a significant passing to online shopping during COVID-19. And it explains people who experienced online shopping are more likely to keep (Shen, Namdarpour, and Lin, 2022). Besides online shopping, the study by Mouratidis and Papagiannakis shows online learning, telework, telehealth, and teleconferencing also increased when compared to before COVID-19. While teleconferencing and e-learning increased by 34 percent, telework has a 31 percent increase (Mouratidis and Papagiannakis, 2021).

Microsoft Teams is a platform that combines teleconference, chat, and notes. Since it was released in November 2016, there was no significant growth like during the pandemic. From March 2020 to April 2020, the number of users increased by 4,5 times. According to a survey of Microsoft Teams users in the United States, 29.71 percent of businesses used Microsoft Teams for remote work during the COVID-19 pandemic in 2020 (Statista, 2022).

During a pandemic, ICT generates solutions to understand the spread of COVID-19. For instance, the Australian Government released an application called CovidSafe to monitor and keep down the spreading of COVID-19 and invest $5 million for consulting, development, and maintenance (Consultancy n.d.). Also, Turkey Government launched a similar application named "Hayat Eve Sığar" to minimize spreading. The application generates a unique code and people use this code when they went to indoors such as in shopping malls, theaters, cafés and restaurants, and schools.

The National Institute of Standards and Technology (NIST) defines cloud computing as a model that provides ubiquitous, on-demand network access to a common pool of configurable computing resources (e.g., servers, networks, applications, storage, and services) that can be provisioned and released quickly with minimal management effort or service provider interaction (Mell and Grance, 2017). According to a study, cloud computing is helping countries in the fight against COVID 19, economically and commercially in the education and health sectors (Alhomdy, Thabit, Abdulrazzak, Haldorai and Jagtap, 2021).

According to a survey of 263 senior UK banking executives, shows two-thirds are planning operational costs and want to invest in technology and automation. After COVID-19 they realized their potential agility and the necessity of digitalization (Appian, 2020).

Ntasis, Koronios, and Pappas (2021) investigate the impact of COVID-19 on the technology sector for TATA Consultancy Services (TCS). TATA Consultancy Services is a multinational information technology services and consulting company in India. They have 285 offices across 46 countries. They used the business and stock value performance of TCS for their study. According to a study, economic policy uncertainty has a negative effect on the stock price of TCS (Ntasis, Koronios, and Pappas, 2021).

Forecasting future demand is quite important to manage production, plan resources, avoid overstocking, and supply process. There are a number of methodologies that can be used at this point, and most of these methodologies are based on past demand or sales (Rohaan, Topan, and Groothuis-Oudshoorn, 2022). As examples in the energy sector, Ağbulut (2022), Ayvaz et al. (2017), and, Sun and Liu (2016) used ML to forecast energy demand and CO_2 emission in different models. In the health sector, Yang et al. (2022) studied the prediction of lung cancer recurrence and survivability, and Chowdhury et al. (2022) studied the diagnosis of COVID-19 from coughs sounds, Sabeti, and friends (2022) searched for detection of craniosynostosis in newborns with machine learnings. Madhurya et al. (2022), and Şahin et al. (2013) examined the detection of credit card fraud with different machine learnings techniques (Ağbulut, 2022; Ayvaz, Kusakci, and Temur, 2017; Chowdhury, Kabir, Rahman and Islam, 2022; Madhurya, Gururaj, Soundarya, Vidyashree and Rajendra, 2022; Sabeti, Boostani, Moradi and Shakoor, 2022; Sahin, Bulkan, and Duman, 2013; Sun and Liu, 2016; Yang, Xu, Sun, Zhang and Farid, 2022). While Ensafi et al. (2022), used time series forecasting techniques to predict item sales of a retail store, Chen and Ou (2011) used extreme learning machines to forecast sales in the retail industry (Chen and Ou, 2011; Ensafi, Amin, Zhang, Shah, 2022).

3. Methodology

This section explains general information about machine learning and implementation areas, three regression models that we used for prediction, and explains which data were collected and preprocessing of collected data. The study adhered to research and publication ethics standards.

3.1 Machine Learning

Machine learning (ML) is the study of computer algorithms that can learn and develop on their own through experience and data (Mitchell 1997). It is becoming one of the hottest topics for research. Researchers are used ML in various areas such as health, energy, biology, social science, banking, etc. ML helps to detect some diseases earlier, predict energy consumption, forecast sales in various sectors, identification of fraud or spam e-mail, and recommend products or videos.

There are various platforms to implement machine learning. In this study, Konstanz Information Miner (KNIME) was used for implementation. KNIME is an open-source and free platform based on Java that used data analysis, integration, and reporting. Figure 2 is a summary of input and outputs values used in three regression models. Different models were implemented to sales data and compared with each other. The best prediction was used to explain COVID-19 effects on sales.

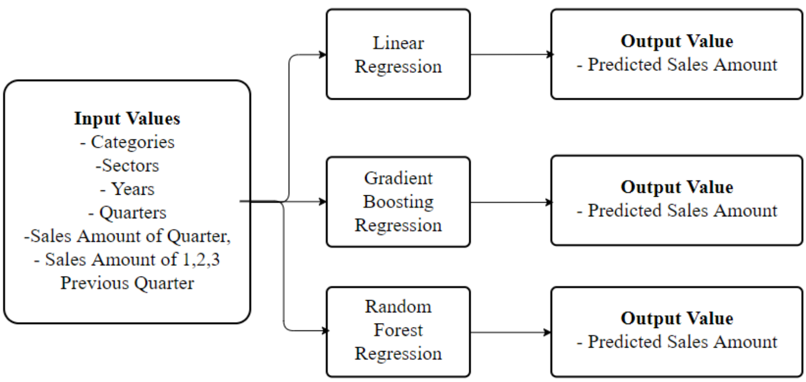

Figure 2. Diagram of Input and Output Value

3.2 Data Preparation

In this study, sales data of a technology company was collected between 2015 and 2022 years. At the first, sales data have 3.072 rows that include 8 service categories and 12 sectors. The 3 most known services categories were selected according to sales amount from the 8 services. Since 5 sectors have minor sales amount like 4 percent as total, they were grouped in others. Sales amount was calculated quarterly every year and 1,2, and 3 previous quarter amount was added as new columns to understand sales trends. Every row that has "0" as sales amount was deleted to make more successful forecasting. After preprocessing, sales data have 531 rows. Finally, historical information of COVID-19 was taken from the Republic of Turkey Ministry of Health website (Anon n.d.).

Sales amount normalized with min-max normalization. This method increases consistency of models. The purpose of min-max normalization method is to normalize the smallest value to 0 and the largest value to 1, and spread all other data to this 0-1 range. Its formula is as below. This approach serves to obscure the company's data while also standardizing all parameters onto a consistent positive scale.

$$X' = \frac{X - X_{min}}{X_{max-}X_{min}} \quad (1)$$

X = The value that will normalized

X_{min} = The smallest value of variable

X_{max} = The biggest value of variable

After normalization, descriptive analysis implemented to sales amount. Normal distribution was controlled and sales amount variables was found non-parametric. Figure 3 shows the distribution of values. The outliers that were seen in figure was checked in sales data. They did not remove from data because any mistake was not found.

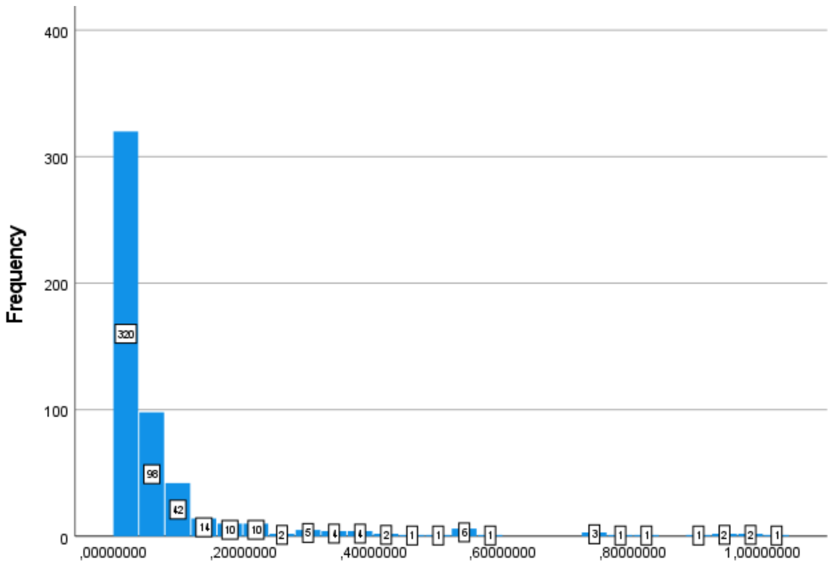

Figure 3. Frequency of Sales Amount

Kruskal Wallis test implemented for category and sector. This method used to test whether the mean of two or more samples shows a significant difference from each other. Table 2 shows the results of test. Significance level is lower 0,001, which means category and sector has significance on sales amount.

Table 2. Results of Kruskal Wallis test for Category and Sector

Variables	Test	Significance
Category	Kruskal-Wallis Test	<0,001
Sector	Kruskal-Wallis Test	<0,001

Figure 4 shows the sales amount of three service categories every quarter. 3rd service category has a major sales percentage every quarter. The revenue line has a significant peak in every last quarter of a year until the 2021 year.

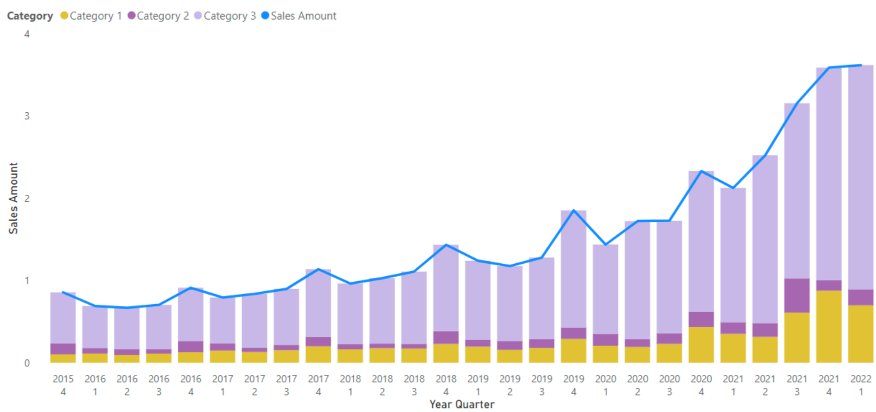

Figure 4. Categorical Sales Amount Quarterly

3.3 Machine Learning Models

A study by Chicco and friends shows that R2 has more advantages over other statistical rates such as MAE, MSE, RMSE, and MAPE in different analyses. According to the study, only MAPE or MSE, or RMSE is not enough to explain a regression performance quality (Chicco, Warrens, and Jurman, 2021).

3.1.1 Linear Regression

Linear regression is a classical statistical algorithm and a machine learning algorithm. The first type of regression analysis was developed by Adrien Marie Legendre in 1806 (Legendre 1806). The relationship between variables is modeled using linear predictor functions. Simple linear regression is used when there is just one variable. Multiple linear regression is used when there is more than one variable (Rencher and Christensen, 2012). In linear regression, the main goal is finding the best equation which is nearest to the values.

3.3.2 Gradient Boosting Trees Regression

The gradient boosting technique used in classification and regression was suggested by Breiman et al. in 1984 (Breiman, Friedman, Olshen and Stone, 1984). The gradient-boosting algorithm improved weak learners to strong learners (Nie, Roccotelli, Fanti, Ming and Li, 2021). The main goal is to minimize the sum of squared errors. Figure 6 shows the relationship between iterations and error. With every iteration, the error is decreasing and the model becomes more successful.

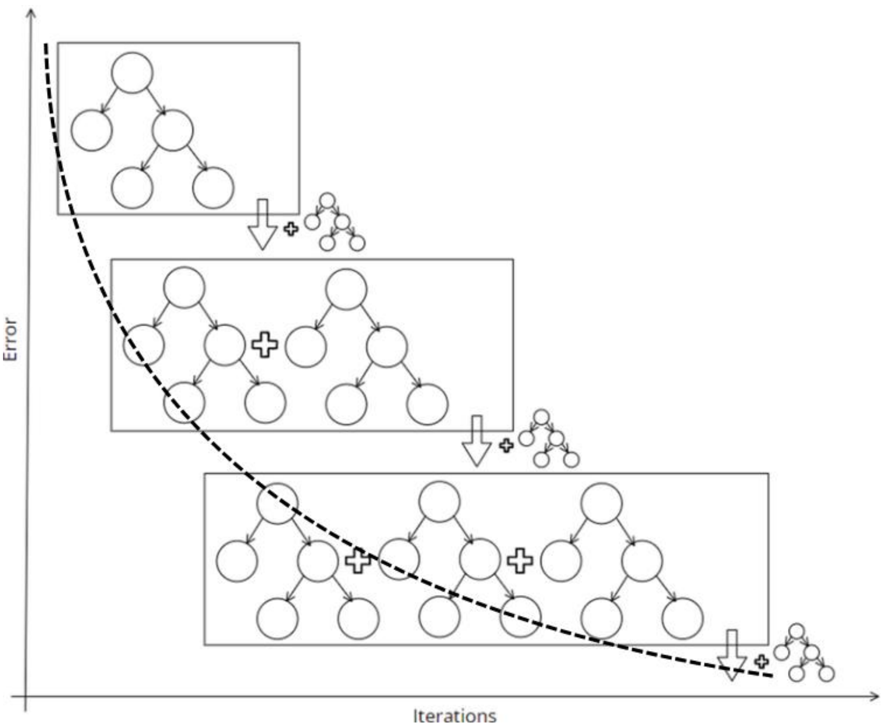

Figure 5. Schematic Representation of Gradient Boosting Regression

3.3.3 Random Forest Regression

Like gradient boosting, the random forest technique is used for classification and regression. Tin Kam Ho created the first algorithm in 1995 (Ho, 1995). The random forest algorithm generates reasonable forecasting with little configuration. In the regression model, the average forecasting of each tree are returned and overfitting of the training set is prevented (Ho, 1998; Minasny, 2009). Figure 7 shows the diagram of random forest regression. This method is very popular to analyze large data sets (Borup, Christensen, Milbach and Nielsen, 2022).

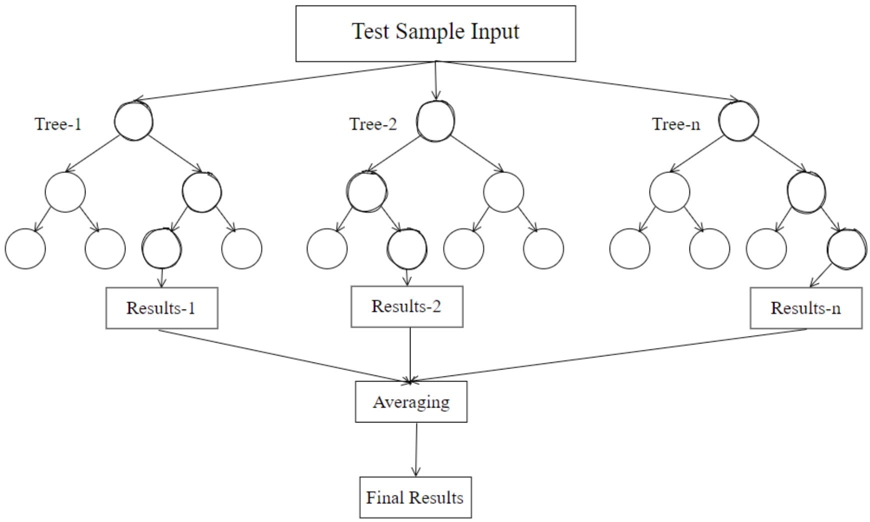

Figure 6. Diagram of Random Forest Regression

4. Implementation

The monthly total sales figures have been calculated on a quarterly basis starting from the first month of the year. The amounts of the 1, 2, and 3 previous quarters were added as new columns to understand sales trends. Three service categories,7 sectors, years and quarterly knowledge, quarterly revenue amounts, and sales amount of 1,2, and 3 previous quarters are used as inputs for machine learning. After splitting the sales data into training and test data, three regression models were implemented and the sales figures for the next year were predicted with different year combinations. Models are compared to each other according to the R^2 values.

Three regression models were constructed using KNIME. The inclusion of parameters in the regression model was determined based on their p-values. Parameters with a p-value less than 0.05 were included in the model. Figure 7 shows the construction of the models. Excel reader module is used to import collected sales data. KNIME understands data type automatically but also allows to change ma-

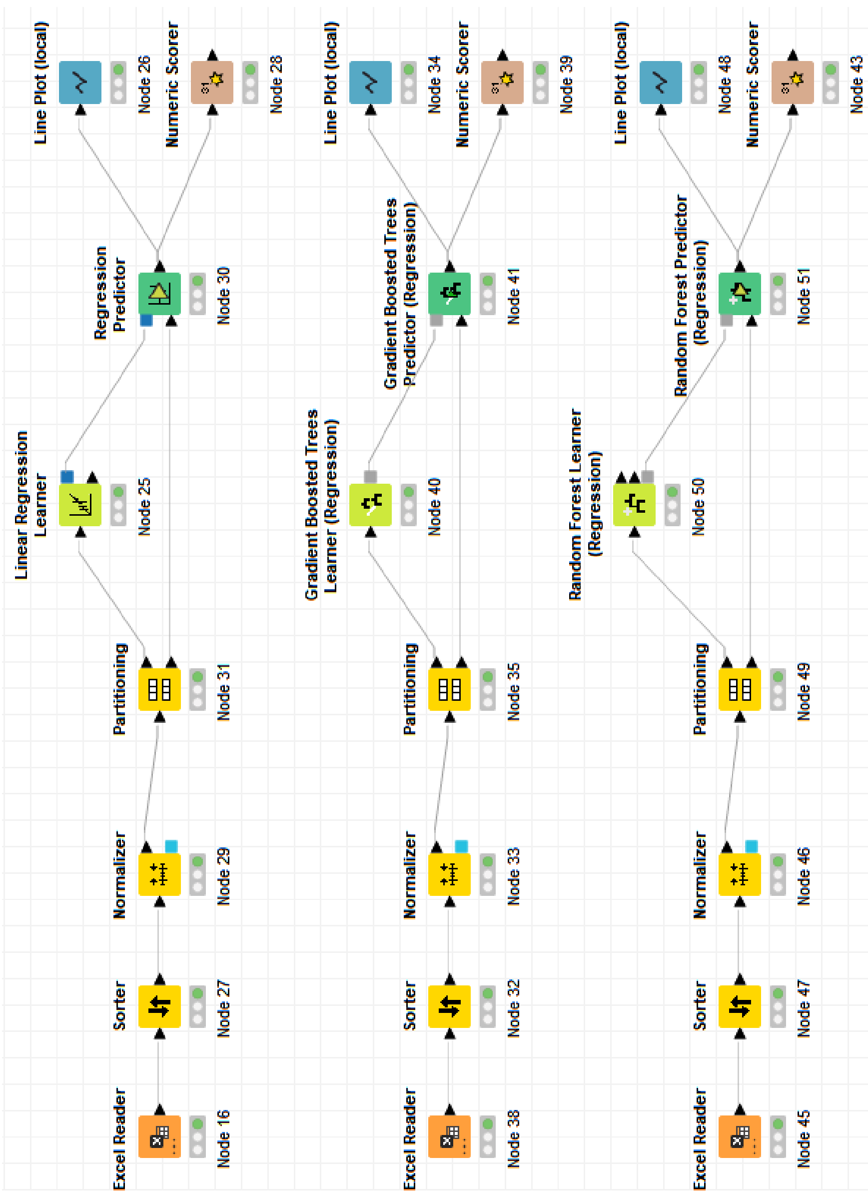

Figure 7. Regression models on KNIME

nually. The imported data were sorted by year and quarter. With the normalizer module, data consistency increased for sales amount. The partitioning module helps to split data as train and test. Partitioning rate change according to the years which belongs to the train or test. Three regression model is added with two modules that are learner and predictor. The learner module teaches machine learning which data are input and which are predicted. The predictor module is configured to predict sales amounts. The numeric scorer calculates the success of models and the line plot shows custom graphics.

5. Numerical Results

In this section, three various forecastings were calculated with KNIME. One of them belongs to pre COVID-19 era and the others prediction of the COVID-19 era.

5.1 Forecast for 2019 with Data Between 2015 and 2018

From 2015 to 2018 years belong to the pre-COVID-19 era. Data from 2015-2018 were used as training data and data from 2019 were used as test data. The sales data has 270 rows for training and 82 rows are predicted with three models. Figure 9 shows the predicted sales amount with linear regression and the actual sales amount for all sectors in 2019. Even if the quality of prediction looks good, it changed from sector to sector.

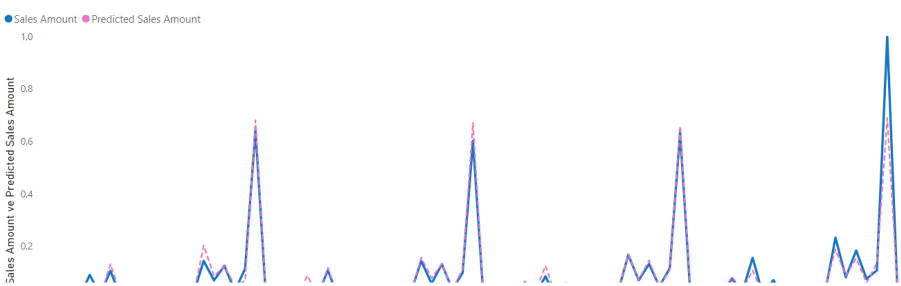

Figure 8. Actual and Predicted Sales Amount of 2019 Year with LR for All Sectors

Figure 9 and 10 details the actual and projected sales amounts for the technology and holding sector. There is a noticeable deviation at the end of the year. There may be delays in the collection of revenues.

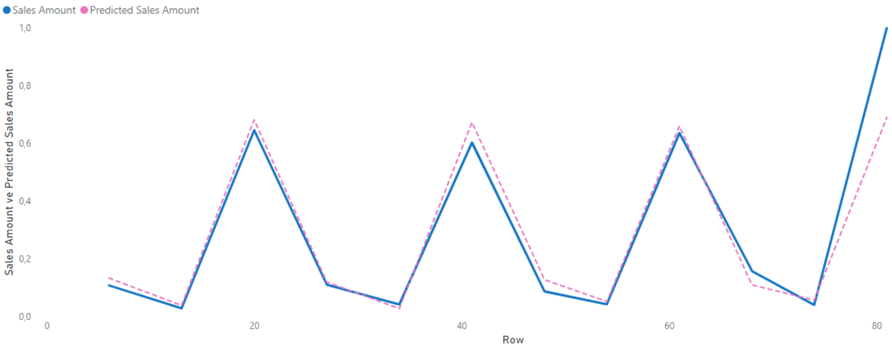

Figure 9. Actual and Predicted Sales Amount of 2019 Year with LR for Technology Sector

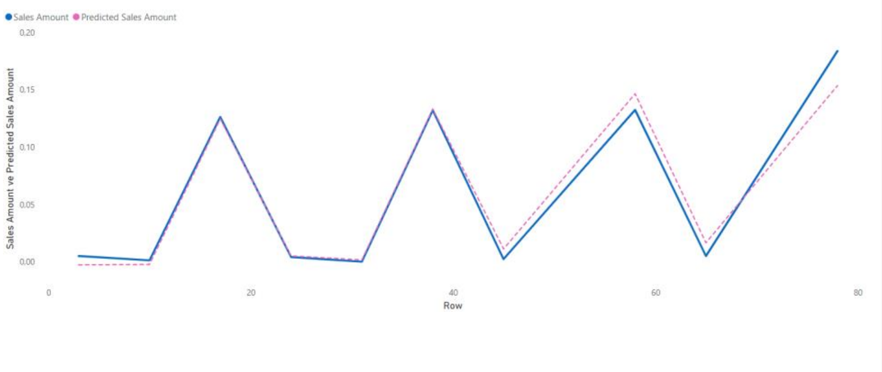

Figure 10. Actual and Predicted Sales Amount of 2019 Year with LR for Holding

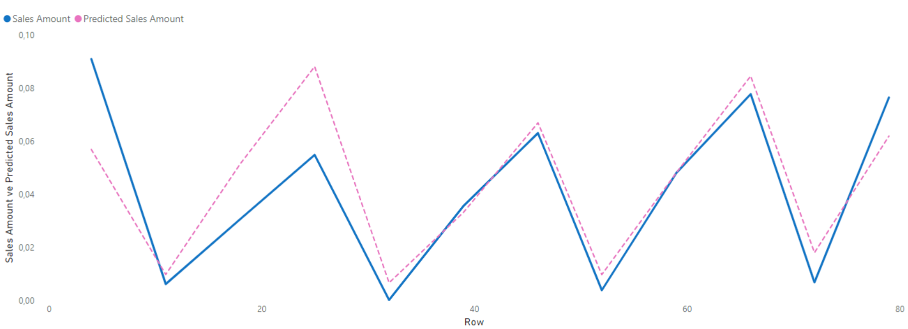

Figure 11. Actual and Predicted Sales Amount of 2019 Year with LR for Automotive Sector

The finance sector had a difference between actual and predicted sales amount almost all year is shown in Figure 12. While predicted sales amount was more than actual in the first quarter, there was not a difference in the second quarter. This means the collection revenue is not associated with this difference.

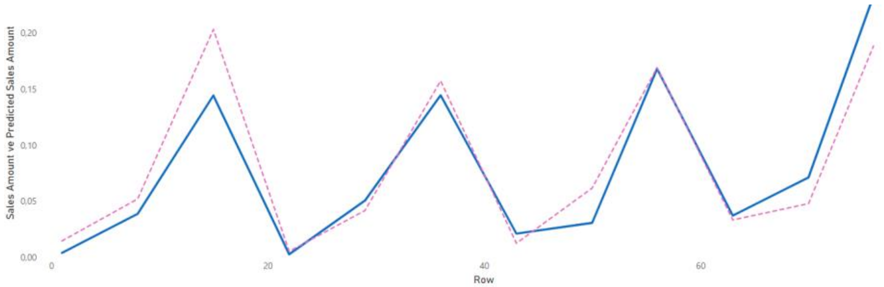

Figure 12. Actual and Predicted Sales Amount of 2019 Year with LR for Finance Sector

Figure 13 explains the values of the retail sector. There is a significant difference in the first quarter of the year like the finance sector. However, in this case, this difference is a positive way.

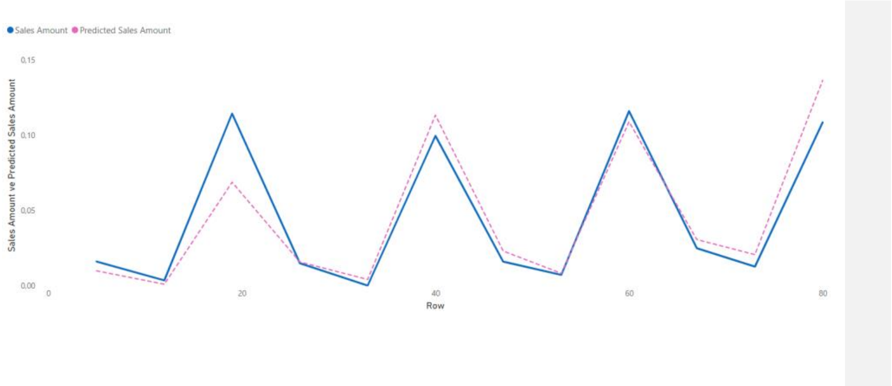

Figure 13. Actual and Predicted Sales Amount of 2019 Year with LR for Retail Sector

Service and others were shown in Figures 14 and 15. Both of them had significant deviations all year. It is hard to explain this difference with a collection of revenue. Because of the irregularity of sales amount, different models can be improved for those sectors.

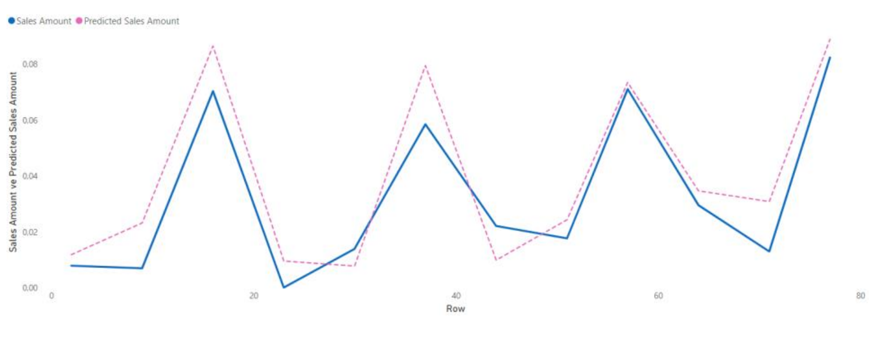

Figure 14. Actual and Predicted Sales Amount of 2019 Year with LR for Service Sector

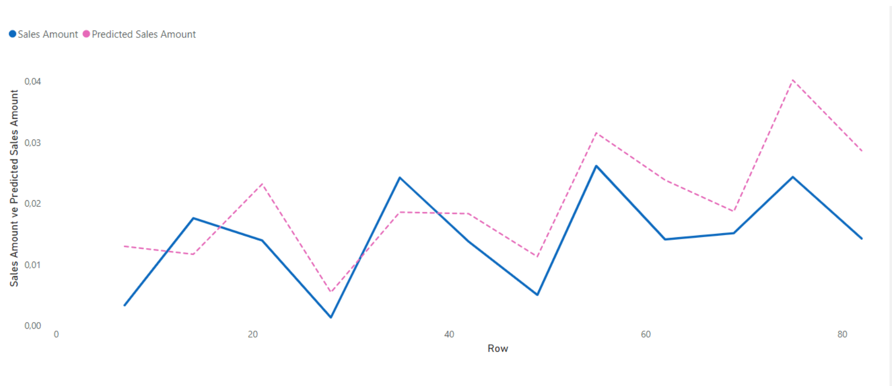

Figure 15. Actual and Predicted Sales Amount of 2019 Year with LR for Others

Table 4 is a summary of the results of three regression models. Results shows, that while linear regression has the biggest R^2 value which means has better goodness of fit, random forest regression is the most ineffective regression model.

Table 4. Result of Three Regression Models for 2019 Forecasting

Regression Model	**Train Data**	**Test Data**	**R^2**	**MAE**	**MSE**	**RMSE**	**MSD**	**MAPE**
LR		2020	0,896	0,028	0,003	0,056	0,001	2,108
GBR	2015-2019	0,928	0,023	0,002	0,047	-0,009	1,093	
RFR		0,894	0,026	0,003	0,057	-0,013	1,757	

5.2 Forecast for 2020 with Data Between 2015 and 2019

The First COVID-19 case has announced at the beginning of 2020 in Turkey. We try to forecast the 2020 year that has COVID-19 with data between 2015 and 2019. The construction of regression models is the same as 2019 forecasting which is explained in the implementation section. 352 rows of sales data were used for training and 80 rows were predicted. The actual and predicted sales amount was shown in Figure 16.

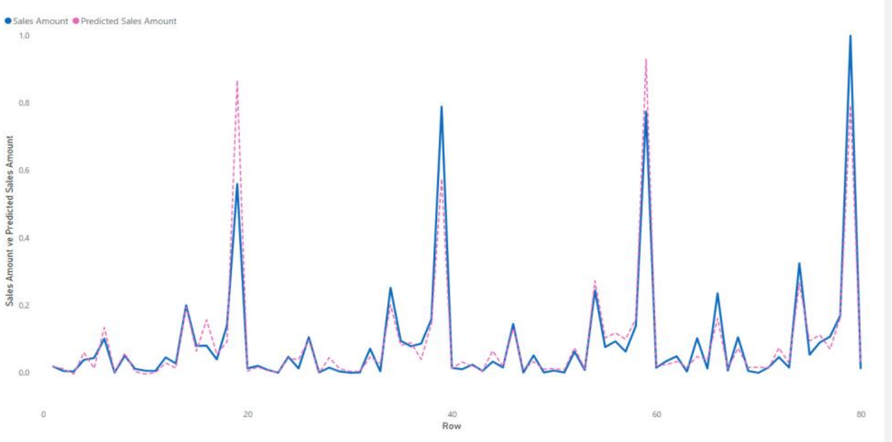

Figure 16. Actual and Predicted Sales Amount of 2020 Year with Linear Regression

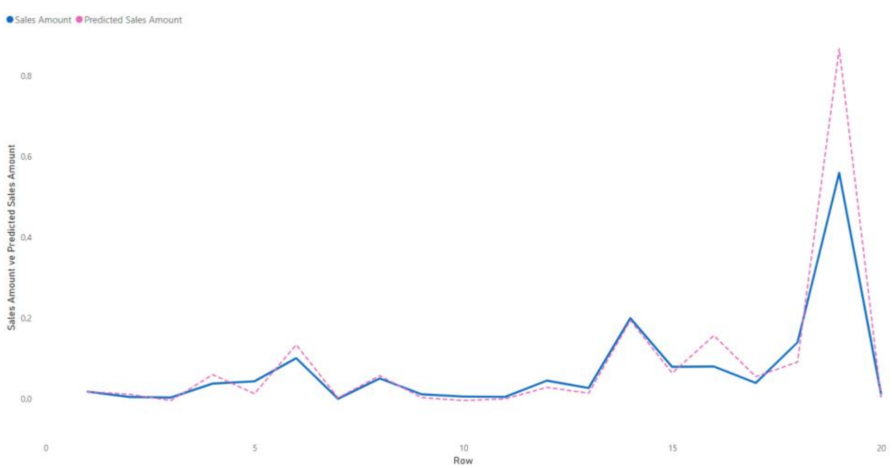

Figure 17. Actual and Predicted Sales Amount of Q1 2020 Year with Linear Regression

It seems that the actual sales amount is lower than the predicted sales amount in the first quarter of 2020. Figure 17 shows the detail of the first quarter. As we understand, there is a notable variation in category 3 that has the biggest sales value. The first quarter revenue is lower than the predicted sales amount.

When on closer inspection of the second quarter in Figure 18, we can see total sales amount is bigger than predicted. The deficit in the first quarter may have been offset here.

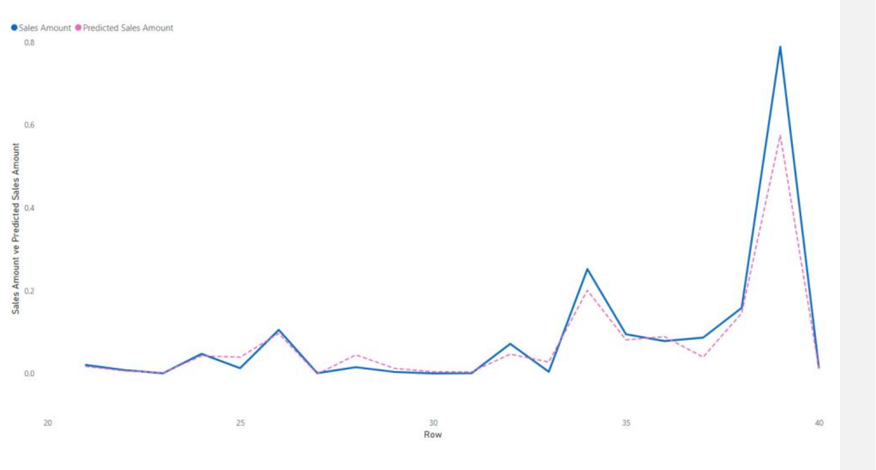

Figure 18. Actual and Predicted Sales Amount of Q2 2020 Year with Linear Regression

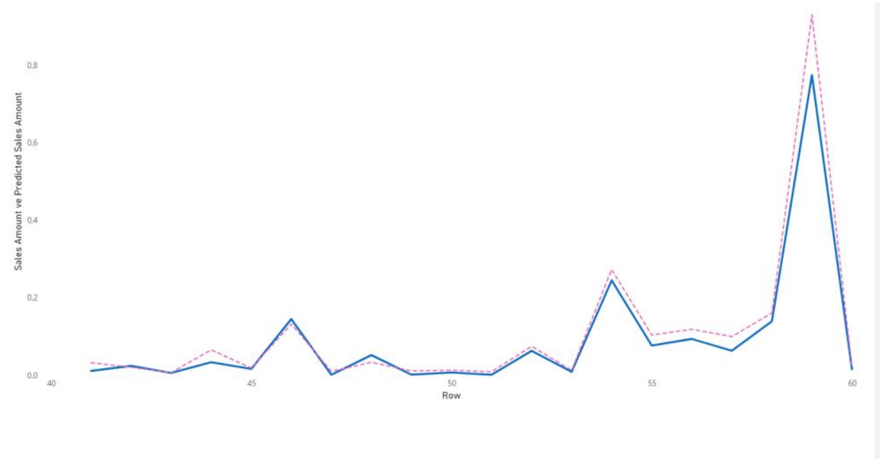

Figure 19. Actual and Predicted Sales Amount of Q3 2020 Year with Linear Regression

The third quarter of the year has better accuracy as we can see in Figure 19 when compared first and second quarters. This adaptation impact is kept in the fourth quarter which is shown in Figure 20.

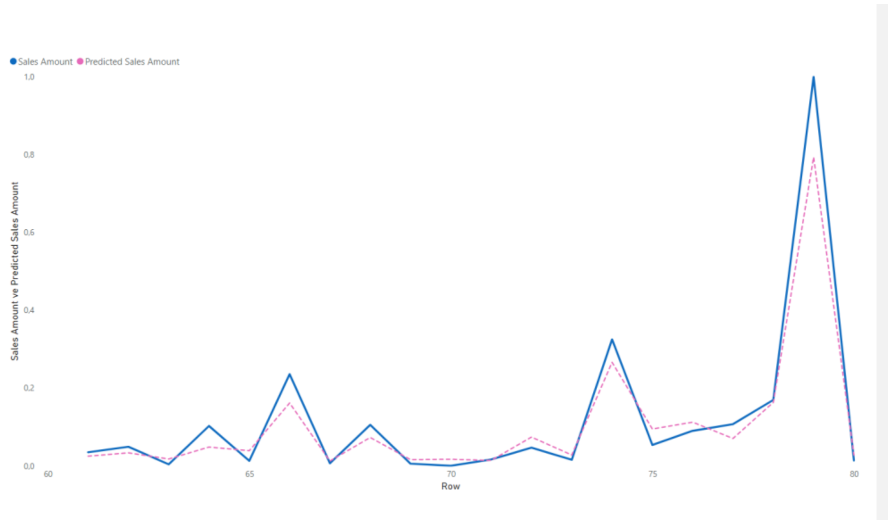

Figure 20. Actual and Predicted Sales Amount of Q4 2020 Year with Linear Regression

Table 5 shows that R^2 has decreased compared to the 2019 prediction. COVID -19 may cause this decrease.

Table 5. Results of Three Regression Models for 2020 Forecasting

Regression Model	Train Data	Test Data	R^2	MAE	MSE	RMSE	MSD	MAPE
LR		2020	0,896	0,028	0,003	0,056	0,001	2,108
GBR	2015-2019	0,928	0,023	0,002	0,047	-0,009	1,093	
RFR		0,894	0,026	0,003	0,057	-0,013	1,757	

The linear regression model was the best model when we predict the non-COVID-19 era. The linear regression was implemented quarterly for the 2020 year. Results are written in table 6. As we understand, prediction quality gets better towards the end of the year. Especially, the R2 value of the 1st quarter has a disruptive impact on the year. When R2 is calculated with the 2nd, 3rd, and

4th quarters, the best value has been obtained. The first case of COVID-19 was appeared end of 2019 and announced as a pandemic in March 2020. In Turkey, the full lockdown was from March to June. This situation explains the lower R2 value in the first and second quarters. Until 21st November, Turkey did not have a lockdown.

Table 6. Results of Linear Regression Quarterly in the 2020 Year

Regression Model	Train Data	Test Data	R^2
LR	2015-2019	1st Quarter	0,42
LR		1st,2nd Quarter	0,778
LR		1st,2nd,3rd Quarter	0,841
LR		2nd,3rd,4th Quarter	0,931

Forecast for 2020-2021 with Data Between 2015 and 2019The sales data between 2015 and 2019 were used to predict the 2020 and 2021 years which has pandemic. 352 rows of sales data were used for training and 158 rows were predicted. The line graph of actual and predicted sales data is given in Figure 21.

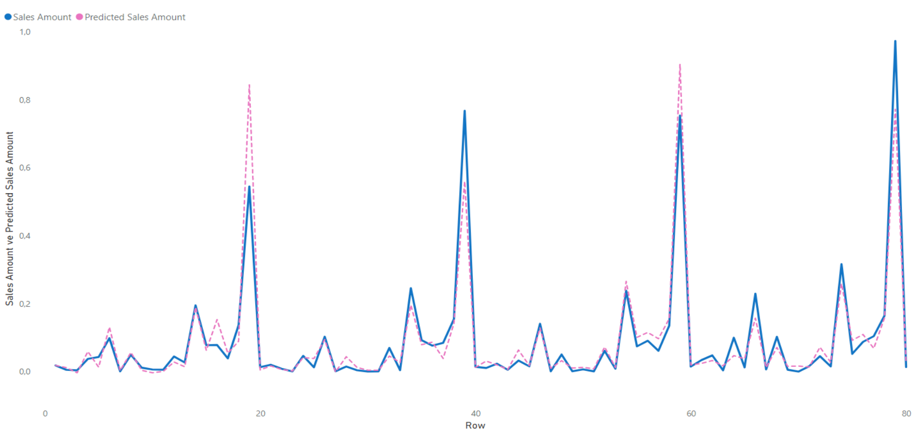

Figure 21. Actual and Predicted Sales Amount of 2021 Year with Linear Regression

Table 7 shows the statistical measures of each regression model. According to the results, linear regression has the best value and is better when compared to forecasting 2020. This improvement can be proof that after the first year of COVID-19 normalization process started.

Table 7. Results of Three Regression Models for 2020-2021 Forecasting

Regression Model	Train Data	Test Data	R^2	MAE	MSE	RMSE	MSD	MAPE
LR	2015		0,921	0,033	0,003	0,058	-0,003	1,459
GBR	-	2020-2021	0,832	0,039	0,007	0,084	-0,021	0,874
RFR	2019		0,839	0,039	0,007	0,083	-0,028	1,074

When the actual and predicted total sales amount in 2020 was compared predicted value is lower than actual. It can be a negative effect of COVID-19 at the beginning of the 2020 year. When collection data examined, the first quarter of 2020 has missing collection. There may have been a problem in collection due to the uncertainty brought by the COVID-19 period. However, in 2020 there is a positive effect on the total sales amount. Table 8 is a summary of the difference between actual and predicted sales amounts.

Table 8. The Difference Between Actual and Predicted Total Sales Amounts

Year	Quarter	Actual Total Sales Amount	Predicted Total Sales Amount	Difference
2020	Q1	1,44	1,73	-20,48%
	Q2	1,72	1,45	15,74%
	Q3	1,73	2,09	-20,69%
	Q4	2,33	2,03	12,87%
2021	Q1	2,13	2,32	-8,99%
	Q2	2,52	2,18	13,64%
	Q3	3,16	2,90	8,21%
	Q4	3,59	3,47	3,45%

5.4 Forecast for 2021 with Data Between 2015 and 2020

By using the values between 2015-2020, 2021 values were tried to be predicted. 432 lines belong to train data and 78 lines are predicted as sectoral. Table 9 summarized metrics of three regression models. Linear regression has the best R2 value.

Table 9. Results of Three Regression Models for 2021 Forecasting

Regression Model	Train Data	Test Data	R^2	MAE	MSE	RMSE	MSD	MAPE
LR	2015-2020	2021	0,928	0,039	0,004	0,063	-0,012	0,752
GBR			0,912	0,041	0,005	0,069	-0,023	0,596
RFR			0,889	0,043	0,006	0,078	-0,033	0,518

2020 year was the most affected from COVID-19. 2021 values predicted with two different train data for understanding that 2020 data has disruptive effect on machine learning whether or. Table 10 shows the R2 value of two different models. The coefficient of determination has better value without 2020 sales data. To ensure that the 2020 sales data will be ignored, the model can be run again by continuing to collect data after the COVID-19.

Table 10. Comparison of 2021 R^2 Values with Different Train Data

Regression Model	Train Data	Test Data	R^2
LR	2015-2019	2021	0,933
LR	2015-2020	2021	0,928

6. Conclusion

After COVID-19 first appeared in China, various sectors were affected tourism, accommodation, education, aviation, production, automotive, energy, technology, and food. Every country took various cautions like using masks, remote working, and curfews to control this pandemic.

Remote working and curfews affected the demand for information and communication technologies (ICT). People need to do various things online such as shopping, working, studying and ICT makes them possible. That's why companies investing in digitalization for adopting this period.

This study aims to understand the impact of COVID-19 and explore the popularity of information technologies in the technology sector through various machine learning models, utilizing sales data from pre-COVID-19 periods of a Turkish consulting firm.

Results show COVID-19 has adverse effects, especially in the first quarter of the 2020 year. After the adaptation period, the sales amount normalizes. Even, in 2021 a positive effect was seen on sales amount. Further, it can be investigated whether the machine learning model tested afterward is suitable for predicting future years or whether it is more appropriate to make predictions by ignoring the COVID-19 period. In addition, the inflation impact can be studied during the COVID-19 era, and sales revenue convert as dollars because of changes in currency. With the increasing demand for the technology sector, being a developer is more interesting day by day, and companies seeking good developers in Turkey and also around the world. It can be examined as a new study, impacts of increasing demand for information technologies on developers, and their salary.

Conflict of Interest

No conflict of interest is declared by the authors.

References

Ağbulut, Ümit. 2022. "Forecasting of Transportation-Related Energy Demand and CO_2 Emissions in Turkey with Different Machine Learning Algorithms." *Sustainable Production and Consumption* 29:141–57. doi: https://doi.org/10.1016/J.SPC.2021.10.001.

Alhomdy, Sharaf, Fursan Thabit, Fua'ad Hasan Abdulrazzak, Anandakumar Haldorai, and Sudhir Jagtap. 2021. "The Role of Cloud Computing Technology: A Savior to Fight the Lockdown in COVID 19 Crisis, the Benefits, Characteristics and Applications." *International Journal of Intelligent Networks* 2:166–74. doi: https://doi.org/10.1016/J.IJIN.2021.08.001.

Anon. n.d. "COVID-19 Pandemic in Turkey - Wikipedia." Retrieved May 19, 2022a (https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Turkey).

Anon. n.d. "Ministry of Health." Retrieved May 19, 2022b (https://www.saglik.gov.tr/).

Appian. 2020. "Innovating in a Crisis: How Covid-19 Changed the Way UK Banks Deliver Core Operations." Retrieved May 19, 2022 (https://appian.com/resources/resource-center/analyst-reports/how-covid-19-changed-the-way-uk-banks-deliver-core-operations.html).

Ayvaz, Berk, Ali Osman Kusakci, and Gül T. Temur. 2017. "Energy-Related CO2 Emission Forecast for Turkey and Europe and Eurasia: A Discrete Grey Model Approach." *Grey Systems 7*(3):436–52. doi: https://doi.org/10.1108/GS-08-2017-0031.

Borup, Daniel, Bent Jesper Christensen, Nicolaj Nørgaard Mühlbach, and Mikkel

Slot Nielsen. 2022. “Targeting Predictors in Random Forest Regression.” *International Journal of Forecasting.* doi: https://doi.org/10.1016/J.IJFORECAST.2022.02.010.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees.*

Cantekin, Kayahan. 2020. “Turkey: Provincial Bar Associations Sue Government for Prohibiting Large Public Meetings and Elections Because COVID-19 Fears.” *Library of Congress.*

Chen, F. L., and T. Y. Ou. 2011. “Sales Forecasting System Based on Gray Extreme Learning Machine with Taguchi Method in Retail Industry.” *Expert Systems with Applications* 38(3):1336–45. doi: https://doi.org/10.1016/J.ESWA.2010.07.014.

Chicco, Davide, Matthijs J. Warrens, and Giuseppe Jurman. 2021. “The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation.” doi: https://doi.org/10.7717/peerj-cs.623.

Nihad Karim Chowdhury, Muhammad Ashad Kabir, Md Muhtadir Rahman, and Sheikh Mohammed Shariful Islam. 2022. “Machine Learning for Detecting COVID-19 from Cough Sounds: An Ensemble-Based MCDM Method.” *Computers in Biology and Medicine* 145:105405. doi: https://doi.org/10.1016/J.COMPBIOMED.2022.105405.

Consultancy. n.d. “The Impact of the Coronavirus on the Global Consulting Industry.” 2020. Retrieved May 19, 2022 (https://www.consultancy.org/news/162/the-impact-of-the-coronavirus-on-the-global-consulting-industry).

Ensafi, Yasaman, Saman Hassanzadeh Amin, Guoqing Zhang, and Bharat Shah. 2022. “Time-Series Forecasting of Seasonal Items Sales Using Machine Learning – A Comparative Analysis.” *International Journal of Information Management Data Insights* 2(1):100058. doi: https://doi.org/10.1016/J.JJIMEI.2022.100058.

Ghosh, Bobby. 2020. “Turkey’s Late Response to Coronavirus Overshadows Ramadan.” *BloombergQuint.*

Ho, Tin Kam. 1995. “Random Decision Forests.” Pp. 278–82 vol.1 in *Proceedings of 3rd International Conference on Document Analysis and Recognition.* Vol. 1.

Ho, Tin Kam. 1998. “The Random Subspace Method for Constructing Decision Forests.” *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*(8):832–44. doi: https://doi.org/10.1109/34.709601.

Kandemir, Aslı. 2020. "Turkey Imposing Curfew for People Over Age 65." *Bloomberg.*

Legendre, Adrien Marie. 1806. *Nouvelles Méthodes Pour La Détermination Des Orbites Des Comète.*

Madhurya, M. J., H. L. Gururaj, B. C. Soundarya, K. P. Vidyashree, and A. B. Rajendra. 2022. "Exploratory Analysis of Credit Card Fraud Detection Using Machine Learning Techniques." *Global Transitions Proceedings.* doi: 10.1016/J.GLTP.2022.04.006.

Mell, Peter, and Tim Grance. 2017. "The NIST Definition of Cloud Computing." *Cloud Computing and Government: Background, Benefits, Risks* 267–69. doi: https://doi.org/10.1016/B978-0-12-804018-8.15003-X.

Minasny, Budiman. 2009. "The Elements of Statistical Learning, Trevor Hastie, Robert Tishirani, Jerome Friedman, Second Edition (2009), Springer Series in Statistics, ISBN 0172-7397, 745 Pp." Geoderma 153(1–2):291. doi: 10.1016/J.GEODERMA.2009.08.001.

Mitchell, Tom. 1997. *Machine Learning.*

Mouratidis, Kostas, and Apostolos Papagiannakis. 2021. "COVID-19, Internet, and Mobility: The Rise of Telework, Telehealth, e-Learning, and e-Shopping." Sustainable Cities and Society 74:103182. doi: https://doi.org/10.1016/J.SCS.2021.103182.

Nie, Peng, Michele Roccotelli, Maria Pia Fanti, Zhengfeng Ming, and Zhiwu Li. 2021. "Prediction of Home Energy Consumption Based on Gradient Boosting Regression Tree." *Energy Reports* 7:1246–55. doi: https://doi.org/10.1016/J.EGYR.2021.02.006.

Ntasis, Lazaros, Konstantinos Koronios, and Theodoros Pappas. 2021. "The Impact of COVID-19 on the Technology Sector: The Case of TATA Consultancy Services." *Strategic Change* 30(2):137–44. doi: https://doi.org/10.1002/jsc.2397.

Organisation for Economic Co-operation and Development. 2021. "OECD | Building a Resilient Recovery: How We Can Emerge Stronger from the COVID-19 Pandemic." 2022. Retrieved May 19, *2022* (https://www.oecd.org/coronavirus/en/).

Overseas Security Advisory Council. 2020. "Health Alert: Turkey, New COVID-Related Measures In Effect Starting November 20." *Retrieved May* 19, 2022 (https://www.osac.gov/Country/Turkey/Content/Detail/Report/9bc0839f-2e83-446a-860f-1a2d5567cf7a).

Rencher, A. C., and W. F. Christensen. 2012. *Methods of Multivariate Analysis, Wiley Series in Probability and Statistics.* 19th ed.

Rohaan, D., E. Topan, and C. G. M. Groothuis-Oudshoorn. 2022. "Using Supervised Machine Learning for B2B Sales Forecasting: A Case Study of Spare Parts Sales Forecasting at an after-Sales Service Provider." *Expert Systems with Applications* 188:115925. doi: https://doi.org/10.1016/J.ESWA.2021.115925.

Sabah. 2020. "Otomotiv Devi Corona Virüs Yüzüne Üretimi Durdurdu! COVID-19 Nedeniyle Üretimleri Durduran Markalar." *2020*. Retrieved May 19, 2022 (https://www.sabah.com.tr/galeri/otomobil/otomotiv-devi-corona-virus--yuzune-uretimi-durdurdu-covid-19-nedeniyle-uretimleri-durduran-markalar).

Sabeti, Malihe, Reza Boostani, Ehsan Moradi, and Mohammad Hossein Shakoor. 2022. "Machine Learning-Based Identification of Craniosynostosis in Newborns." *Machine Learning with Applications* 8:100292. doi: 10.1016/J.MLWA.2022.100292.

Sahin, Yusuf, Serol Bulkan, and Ekrem Duman. 2013. "A Cost-Sensitive Decision Tree Approach for Fraud Detection." *Expert Systems with Applications* 40(15):5916–23. doi: https://doi.org/10.1016/J.ESWA.2013.05.021.

ScienceDirect. n.d. "ScienceDirect.Com | Science, Health and Medical Journals, Full Text Articles and Books." Retrieved May 21, 2022 (https://www.sciencedirect.com/).

Shen, Hui, Farnoosh Namdarpour, and Jane Lin. 2022. "Investigation of Online Grocery Shopping and Delivery Preference before, during, and after COVID-19." *Transportation Research Interdisciplinary Perspectives* 14:100580. doi: https://doi.org/10.1016/J.TRIP.2022.100580.

Statista. 2022. "What Software or Tools Does Your Firm Use for Remote Work?" Retrieved May 19, 2022 (https://www.statista.com/statistics/892994/staffing-industry-types-of-text-messaging-software-used-in-the-united-states/).

Sun, Wei, and Mohan Liu. 2016. "Prediction and Analysis of the Three Major Industries and Residential Consumption CO2 Emissions Based on Least Squares Support Vector Machine in China." *Journal of Cleaner Production* 122:144–53. doi: https://doi.org/10.1016/J.JCLEPRO.2016.02.053.

T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı. n.d. "Presidential Strategy and Budget Department of Turkish Republic." Retrieved May 21, 2022 (https://www.sbb.gov.tr/).

Taser, Didem, Esra Aydin, Alev Ozer Torgaloz, and Yasin Rofcanin. 2022. "An Examination of Remote E-Working and Flow Experience: The Role of Technostress and Loneliness." *Computers in Human Behavior* 127:107020. doi: https://doi.org/10.1016/J.CHB.2021.107020.

The Economist. 2020. "What Turkey Got Right about the Pandemic." *The Economist.*

The UN Refugee Agency. 2021. "Announcements - UNHCR Turkey." Retrieved May 19, 2022 (https://help.unhcr.org/turkey/coronavirus/announcements/).

WHO. 2022. "WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data." 2022. Retrieved May 19, 2022 (https://covid19.who.int/).

Worldometers. n.d. "Turkey COVID - Coronavirus Statistics - Worldometer." Retrieved May 19, 2022 (https://www.worldometers.info/coronavirus/country/turkey).

Yang, Yang, Li Xu, Liangdong Sun, Peng Zhang, and Suzanne S. Farid. 2022. "Machine Learning Application in Personalised Lung Cancer Recurrence and Survivability Prediction." *Computational and Structural Biotechnology Journal.* doi: https://doi.org/10.1016/J.CSBJ.2022.03.035.