HTTPS://WWW.INFORMATIONANALYTICS.ORG/

E-ISSN: 3108-3927

JOURNAL OF INFORMATION ANALYTICS

VOLUME: 01
ISSUE: 01
YEAR 2025

E-ISSN: 3108-3927

JOURNAL OF INFORMATION ANALYTICS

Volume: 01 Issue: 01 2025

Correspondence

İ.Ü. Faculty of Transportation and Logistics Istanbul University Beyazit Campus 34452 Fatih, İstanbul, Turkiye
+90 537-623-00-26
journalofinformationanalytics@gmail.com

https://www.informationanalytics.org/

E-ISSN: 3108-3927

JOURNAL OF INFORMATION ANALYTICS

Volume: 01 Issue: 01 2025

Publication Ethics and Publication Malpractice Statement

Journal of Information Analytics adheres to the principles and guidelines outlined by the Committee on Publication Ethics (COPE), ensuring ethical and responsible behavior from all parties involved in the publication process, including authors, reviewers, editors, and the publisher. The journal also refers to the best practices described in Elsevier's Publishing Ethics Resource Kit (PERK) in case of ethical concerns or disputes

Author's Responsibility

- **Participation in Peer Review:** By submitting a manuscript, authors agree to engage in the journal's peer review process actively and constructively.
- **Reporting Standards:** Authors must present original, accurate, and comprehensive accounts of their research. Misleading or intentionally false statements are unethical and unacceptable.
- **Data Transparency and Retention:** Authors should maintain and, if requested, provide access to raw data for editorial review. Data must be retained for a reasonable period following publication.
- **Originality and Proper Attribution:** Submitted manuscripts must be the authors' original work. Any use of others' ideas or words must be properly cited. The journal employs plagiarism detection software such as iThenticate.
- **Simultaneous or Duplicate Submissions:** Submitting the same manuscript to multiple journals is considered unethical. Redundant or duplicate publication is strictly prohibited.
- **Acknowledgment of Sources:** All relevant sources of data and previously published work must be appropriately cited.
- **Authorship Criteria:** Only those who significantly contributed to the research should be listed as authors. All listed authors must approve the final version and agree to submission.
- **Disclosure of Conflicts of Interest:** Any potential conflicts of interest, financial or otherwise, must be disclosed at the time of submission.
- **Ethics Committee Approval:** Studies involving human or animal participants, or data collection via surveys or interviews, must include documentation of ethical committee approval in accordance with national or institutional guidelines.

Reviewer's Responsibility

- **Contribution to Editorial Decision:** Reviewers support editors in making publication decisions and help authors improve their manuscripts through constructive feedback.
- **Timeliness and Qualification:** Reviewers should only accept assignments for which they have the appropriate expertise and are able to provide reviews in a timely manner.
- **Confidentiality:** Review materials must be treated as confidential and not shared or discussed with unauthorized individuals.
- Objectivity and Constructive Feedback: Reviews should be impartial, respectful, and supported by clear arguments. Personal criticism is inappropriate.
- **Acknowledgment of Sources:** Reviewers should alert the editor to any relevant published work not cited by the authors, and report significant similarities to other manuscripts or publications.
- **Conflicts of Interest:** Reviewers must inform the editor of any conflicts and decline to review submissions where such conflicts exist.

Editorial Responsibility

- **Publication Decisions:** Editors are responsible for deciding which manuscripts are suitable for publication based on scientific merit, originality, clarity, and relevance to the journal's scope.
- **Fair and Unbiased Review:** Manuscripts are evaluated without regard to the authors' race, gender, institutional affiliation, or personal background.
- **Confidentiality:** Editors and editorial staff must not disclose any information about a manuscript under review to anyone other than those directly involved in the review process.
- **Disclosure and Conflict of Interest:** Editors must recuse themselves from handling manuscripts in which they have conflicts of interest and must not use unpublished materials for personal research without consent.

Publishing Ethics Issues

- **Allegations of Misconduct:** Any party may report suspected unethical behavior. Sufficient documentation must be provided for the editor to initiate an investigation.
- **Investigation Process:** The editorial team will confidentially review allegations in consultation with relevant parties. If necessary, institutional authorities may be informed.
- **Handling of Minor Breaches:** Minor issues may be resolved through dialogue with the author. Corrections or clarifications may be issued.
- **Serious Violations:** For serious breaches (e.g., data fabrication, plagiarism), the manuscript may be withdrawn, and published articles may be retracted. In extreme cases, the authors' affiliated institutions may be notified.

E-ISSN: 3108-3927

JOURNAL OF INFORMATION ANALYTICS

Volume: 01 Issue: 01 2025

Content

Journal of Information Analytics, 1(1), 1-11

Submit 02.11.2024
Revision 13.01.2025
Accept 19.02.2025
Publish 12.03.2025

Journal of Information Analytics

Open Access

Research Article

Machine Learning-Powered Forecasting for Social Media Campaign Optimization in E-Commerce

Mahmut Osman Biricik[a], Yunus Eroğlu[a], Suleyman Mete[a]

Industrial Engineering Department, Gaziantep University, Gaziantep, Türkiye

Abstract

In today's conditions, competition is increasing day by day, companies that want to survive in these difficult market conditions are looking for new solutions. This is where e-commerce and social media sales come into play. In this paper, artificial intelligence applications have been used to increase the efficiency of the social media campaigns of an e-commerce company, which sells through retail and e-commerce. In this way, it will be predicted whether the campaigns will be popular or not before campaign advertisements are published on social media accounts and all waste will be prevented. Stocking of raw materials, unnecessary advertising shooting costs, and costs spent to highlight videos will be exactly eliminated. This prediction is provided through the artificial intelligence application Orange Software. The necessary data for the study was obtained via the internet from companies operating in this sector. These data have been tested in artificial intelligence applications and increase of the efficiency has been analyzed.

Keywords E-commerce, artificial intelligence, forecast, machine learning

Citation: Biricik, M.O, Eroğlu, Y., & Mete, S. (2025). Machine learning-powered forecasting for social media campaign optimization in e-commerce. *Journal of Information Analytics*, 1(1), 1-11.

Corresponding Author: Suleyman Mete smete@gantep.edu.tr

1. Introduction

In today's trade, competition pushes companies more and more every day. The competitive rate is much higher in e-commerce and social media sales than in normal commerce, because customers can reach the price of the product with just a click. While the conditions are like this, customers who shop on the Internet search for various features in the company where they make purchases and compare the features found with other companies. For example, a customer who wants to make a purchase examines all the parameters such as paying at the door, delivery time of the offered product, return conditions, free shipping option and makes his purchase from the most advantageous place. During the day, sales, especially on Instagram, are organized by offering a discounted price under the title of a campaign and attracting the attention of customers. Instagram has risen to an important position among all social networks, and in terms of businesses, marketing has become a social media platform that must be included in terms of marketing, promotion activities, increasing sales and market share and thus strengthening competitive advantages (Deniz, 2020). In this paper, it is aimed to increase the success of a company that sells via Instagram and to reduce the difficulty it faces to a minimum.

Bringing together the company, which wants to get a share in online sales and prevent waste, with artificial intelligence applications is the most important pillar of targeted success. Nowadays, many discount offers and opportunities are offered on social media to increase sales, acquire customers, and increase market share. Some special time periods make campaigns more effective. These special times are periods when consumers' purchasing tendencies are high. Brands that want to turn these special days into opportunities organize various campaigns and discounts both in stores and on e-commerce websites (Akçadağ and Keklik, 2021). This trend also applies to the nuts industry. Currently, mixed discount packages are popular in online sales of nuts. For example, preparing half a kilo of pistachios, cashews, almonds, hazelnuts and peanuts in packages and offering them to customers with shipping included will increase sales significantly with the right price and a solid infrastructure, but this process is not as easy as it seems from the outside. Distance to raw materials, labor costs, fixed costs, time and cost spent on advertising shooting cause companies to suffer great losses.

In this paper, companies that sell through social media in the nuts sector have been contacted and it has been decided which parameters are needed to measure the consistency of a campaign video to be published on the journey to success. In addition to experienced companies in the sector, the papers in the literature were examined and additions were made to the missing parameters. The main purpose of the paper is to predict whether a promotional campaign on social media will be successful or not, in order to reach this consistency, social media campaign videos have been examined accompanied by various parameters and the number of likes, comments that came to the videos beforehand , the data of the number of shares and the number of orders, which are the objective function of the article, were obtained.

All the parameters were got from papers, companies and videos have been transferred to digital with the help of Orange Software and different results have been achieved here with the help of various algorithms. Data mining tools such as Orange Software also offer powerful features in terms of visualizing and analysing data (Aksu and Güzeller, 2024). An absolute solution has been reached with an artificial neural network algorithm that gives the optimal result in them, and campaign forecasting has been successfully performed. Among the artificial intelligence algorithms that we have scanned the literature and tried; the best results are obtained with artificial neural networks (Eroğlu, 2020).

2. Literature Review

Data mining has attracted a lot of attention in the community in recent years, being able to convert large amounts and large amounts of data into useful information and knowledge. The information and knowledge obtained can be used to apply such as market analysis, fraud detection, and customer retention, for production control and exploration science. Data mining is a process that uses statistical, mathematical, artificial intelligence, and machine learning techniques to extract and identify useful information and related knowledge from various large databases (Ishak et al., 2020). Orange Software is an open-source data mining software that provides users with capabilities such as data preparation, exploratory data analysis, and modelling (Tekerek, 2011). Today, to define a relationship between a parameter and its dependent parameters and to build a model to estimate or predict the chosen parameter, a variety of computational intelligence methods are used, and of course, they also provide favourable results. The purpose of this research is to use these types of algorithms in order to create an efficient model for predicting the deformation modulus on a database. In this regard, the performance of three models created by artificial neural network, K-nearest neighbor and random forest methods have been evaluated with the Orange Software (Fattahi and Jiryaee, 2023). K-Nearest Neighbor classification method has been studied for economic forecasting. Due to the effects of companies' financial distress on stakeholders, financial distress prediction models have been one of the most attractive areas in financial research (Imandoust and Bolandraftar, 2013). The machine learning algorithm, K-Nearest Neighbor (KNN) is introduced for human action recognition. A classifier is trained using KNN and the training set. It is aimed to recognize human actions when given acceleration signals (Wang et al., 2021). Predictive models for prognosis of small sample advanced schistosomiasis patients have not been well studied. We aimed to construct prognostic predictive models of small sample advanced schistosomiasis patients using two machine learning algorithms, k nearest neighbour (kNN) and support vector machine (SVM) utilizing routinely available data under the government medical assistance programme (Zhou et al., 2022). Given that the population is increasing and energy resources are decreasing, in this study we examine the amount of domestic energy consumption. The purpose of this study is to predict the factors affecting energy consumption in buildings. For this prediction, algorithms of decision tree, random forests and K-nearest neighbors have been used. These algorithms are available in Orange Software (Hosseini and Fard, 2021).

3. Methodology

In this paper, we used a software called Orange3, which is one of the machine learning applications of artificial intelligence. Orange3 is an open-source data visualization and machine learning software package used for data mining and data analysis. The data in this article were obtained from companies that trade nuts on social media. When deciding which data should be collected, first of all, articles and studies in the field of e-commerce were examined after that one-on-one interviews were conducted with companies that have been engaged in e-commerce for many years in the light of these sources. After reviewing the source and conducting interviews, it has been decided which data should be used in this article. As a common opinion of most businesses, it was determined that campaigns, especially for new companies, should offer payment at the door service for reliability. In addition, it was decided that the payment at the door service should be added to the campaign amount and that it should not be reflected as an extra fee. Our other determining parameters include how many kilograms the products can be sold more effectively, the positive effects of product diversity, price policies, effective use of the Instagram profile, credit card payment and having a sales store to overcome the trust problem. All data can be clearly observed in Table 1.

Direct transfer of data into a table when entering data into this application makes the work easier. It is possible to easily import an Excel table from the file section. In the Orange application, it is determined which features will be processed from the select columns section, where the data that should be ignored and used is distinguished, and the target value is also determined here. There are various algorithms for statistics and prediction in practice, and it is possible to choose the most appropriate one by connecting data to these algorithms. After testing and training in the algorithms are completed, the scores are transferred to another table. Orange 3 application, the working logic of which is explained above, plays an important role in preventing waste and high costs.

Table 1. Data of the company selling nuts online

Campaigns	Price	How many types	Weight /Kg	Background of Video	Accept Credit Card	Cash on Delivery	Extra Payment at door	Having Sales Outlet	Efficent Profile Managment	Tendency to Trend	Number of Likes	Number of Comment	Number of Share	Number of Order
1	985	6	3	NORMAL	YES	YES	YES	YES	NO	YES	3500	405	556	15
2	700	4	2	NORMAL	NO	YES	YES	NO	NO	NO	55	1	3	0
3	1900	5	4	NORMAL	YES	YES	YES	NO	NO	YES	1435	228	300	7
4	2072	5	4	NORMAL	YES	YES	NO	NO	NO	NO	350	68	50	2
5	1640	9	5	NORMAL	NO	YES	NO	NO	NO	NO	300	47	53	2
6	985	6	3	NORMAL	NO	YES	NO	YES	NO	YES	2610	241	320	8
7	1550	8	4	STRONG	NO	YES	NO	YES	YES	YES	2383	267	205	4
8	1500	8	4	STRONG	YES	NO	NO	YES	YES	YES	95.000	7500	108.000	200
9	1500	8	4	STRONG	YES	NO	NO	YES	YES	YES	3000	434	2700	15
10	1500	8	4	STRONG	YES	NO	NO	YES	YES	YES	9761	515	5225	30
11	1400	4	5	STRONG	YES	NO	NO	YES	YES	YES	17.700	730	12.300	50
12	2500	8	8	STRONG	YES	NO	NO	YES	YES	YES	4950	237	1700	40
13	1500	10	5	NORMAL	YES	YES	YES	YES	YES	NO	271	242	70	5
14	2000	14	7	NORMAL	YES	YES	YES	YES	YES	NO	386	28	89	1
15	1000	5	3	NORMAL	YES	YES	NO	YES	YES	NO	302	24	39	1
16	2000	10	5	STRONG	YES	NO	NO	YES	YES	YES	1500	55	806	6
17	1450	10	5	NORMAL	YES	NO	NO	YES	YES	YES	308	41	227	2
18	1400	10	5	NORMAL	YES	NO	NO	YES	YES	YES	3780	267	1904	25
19	1300	8	4	STRONG	NO	NO	NO	YES	NO	NO	82	18	18	1
20	1500	15	7	STRONG	NO	NO	NO	YES	NO	YES	695	182	194	10

3.1 Neural Network

Artificial Neural Networks are very powerful brain-inspired computational models, which have been employed in various areas such as computing, medicine, engineering, economics, and many others. An artificial neural network is based on the optimization theory. An Artificial Neural Network is a computational model inspired by the functioning of the human brain. It is composed of a set of artificial neurons (known as processing units) that are interconnected with other neurons, and these neurons depend on the weights of the neural network. As the word "network" in Neural Network refers to the interconnection between neurons present in various layers of a system, these weights represent the connections between the neurons, which determine the impact of one neuron on another (Zakaria et al., 2014). The structure of neural network example is given Figure 1.

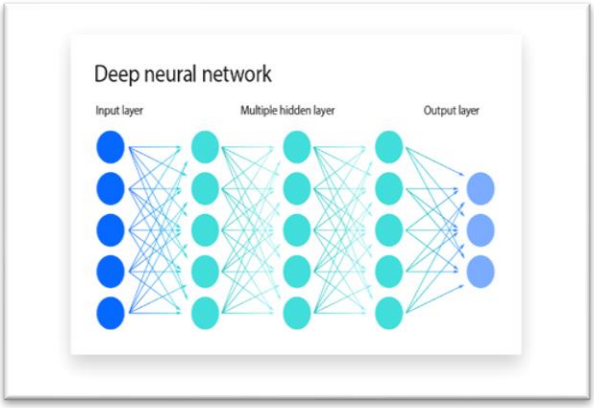

Figure 1. Representative view of deep neural network algorithm

3.1.2 AdaBoost

The AdaBoost algorithm, introduced by Freund and Schapire (1997), is a key method in the boosting approach to machine learning, which combines multiple weak learners to create a highly accurate prediction model. AdaBoost iteratively adjusts the weights of incorrectly classified data points, allowing the algorithm to focus more on difficult cases in subsequent iterations (Schapire, 2013). The representation of the Adaboost algorithm is shown in Figure 2.

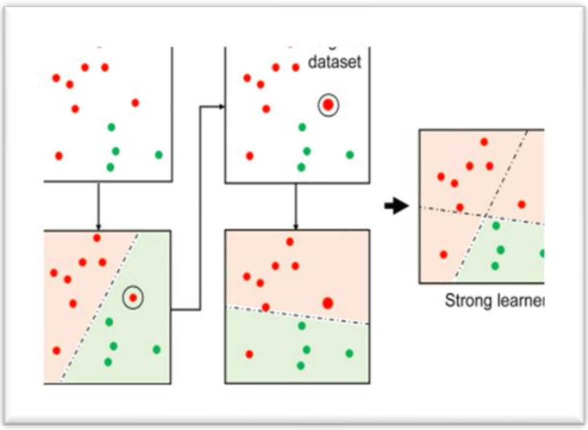

Figure 2. Representative view of adaboost structure

3.1.3 Random Forest

The random forest algorithm, proposed by Breiman (2001), has been highly successful as a general-purpose classification and regression method. This approach combines several randomized decision trees and aggregates their predictions by averaging. It has demonstrated excellent performance in situations where the number of variables is significantly larger than the number of observations. Furthermore, it is versatile enough to be applied to large-scale problems, easily adaptable to various ad hoc learning tasks, and provides measures of variable importance (Biau and Scornet, 2016). The representation of this method can be seen in Figure 3.

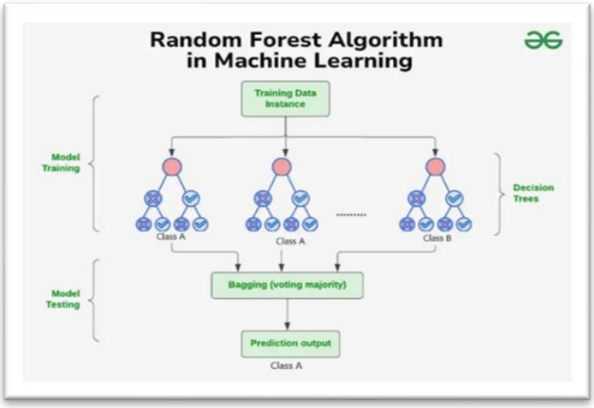

Figure 3. Representative view of random forest method

3.1.4 Tree Algorithms

Decision-tree learning is one of the most successful learning algorithms, due to its various attractive features: simplicity, comprehensibility, no parameters, and being able to handle mixed-type data. In decision-tree learning, a decision tree is induced from a set of labelled training instances represented by a tuple of attribute values and a class label. Because of the vast search space, decision-tree learning is typically a greedy, top-down and recursive process starting with the entire training data and an empty tree (Su and Zhang, 2006). Tree Algorithms showed by Figure 4.

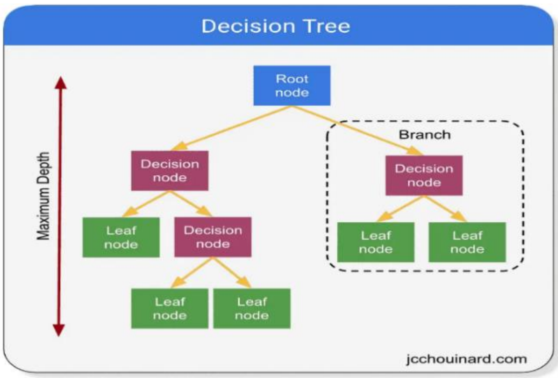

Figure 4. Representative view of decision tree method

3.1.5 KNN (Kernel – Nearest Neighbors)

The K-Nearest Neighbors (KNN) algorithm is a supervised learning method widely used for classification and regression problems, known for its simplicity. In this algorithm, the class of a given data point is determined by examining the "K" nearest neighbours in the training dataset. KNN typically uses distance metrics, such as Euclidean distance, to identify the nearest neighbours and then performs classification by majority voting. One of the main advantages of KNN is that it does not require learning any model parameters during training; in other words, it does not make any assumptions about the data. However, KNN can be computationally expensive for large datasets, as it must check the entire training set for each new instance. Additionally, the scaling of features plays a crucial role in the algorithm's accuracy because unequal feature scales can lead to misleading distance calculations (Altman, 1992). The schematic representation of KNN method shown in Figure 5.

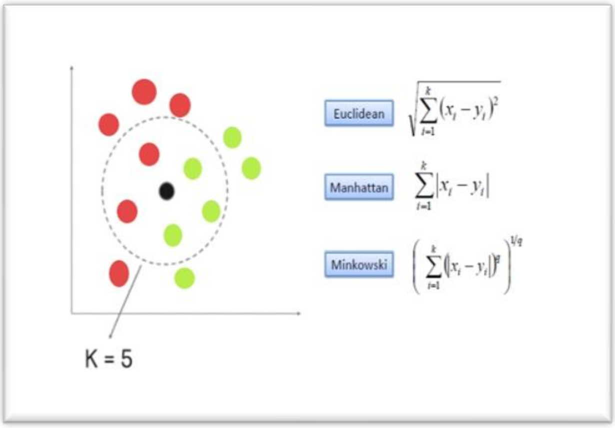

Figure 5. Representative view of KNN method

4. Result And Discussion

During this paper, real-world data was collected from companies active in this sector when gathering datasets. For forecasting, criteria were first determined with experts in the field. These criteria are essential for measuring the consistency of the campaign. In e-commerce, the most important factor is undoubtedly trust, and there are several possible solutions to establish this trust. In this article, Cash on Delivery, the most preferred solution, has been included in the criteria.

Afterward, the importance of several key criteria was identified: the campaign amount, the number of different products included in the campaign, the weight of the package, the efficient management of the advertising platform, and, finally, the absence of any additional fees for cash on delivery.

To observe the consistency of these features, previously published campaign videos on social media platforms was analysed. To measure the effectiveness of the campaigns, outputs based on likes, comments, shares, and the number of orders made in relation to these metrics was collected. We found that the features we defined were consistent at an optimal level. Therefore, this data was entered into the Orange 3 application and performed training sessions using machine learning statistical and forecasting algorithms such as Adaboost, Neural Network, and KNN. In our experiments, generally correct results in the test phase were obtained. However, in the actual application, the Neural Network model yielded almost flawless results with an R2 value of 0.99. Thus, significant progress in achieving optimal results and measuring the consistency of the campaigns were made. The Figure 6 below gives an overview of the Orange model.

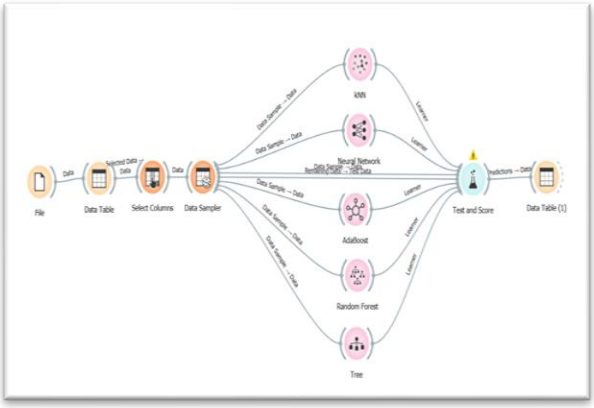

Figure 6. Overview of Orange software

After the data is collected and solved in various algorithms, each algorithm produces its own results. In such cases, the training data is not important because training data usually gives optimistic results. Below, the test on final solution phase of the paper is given in a Table 2.

Mean Squared Error (MSE) is an accepted measure of quality and control in various fields, particularly in statistical process control (Köksoy, 2006). MSE (Mean Squared Error) is also a metric used to evaluate the accuracy of a regression model. It calculates the average of the squared differences between the predicted values and the actual values. A low MSE value indicates that the model is making accurate predictions, while a high MSE value indicates poor performance.

Root Mean Squared Error (RMSE), as the name suggests, is the square root of MSE. Similar to MSE, it measures the average magnitude of the differences between the actual values and the predicted values. However, instead of using the mean squared differences, RMSE takes the square root of the average squared differences. The main difference between RMSE and MSE is that due to this mathematical operation, RMSE gives more weight to large errors, acting almost like a penalty for them. On the other hand, very small errors are, in a way, "rewarded" since their impact is lessened.

Mean Absolute Error (MAE) is an error metric that measures the average of the absolute differences between the predicted values and the actual values. This metric calculates the prediction errors directly based on absolute values and treats both large and small errors equally. In other words, the absolute difference is calculated for each error, and then the average of these differences is taken. The advantage of MAE is that it presents the magnitude of errors in a more understandable way, as it directly reflects the units of measurement.

The Mean Absolute Percentage Error (MAPE) is calculated using absolute errors in each period divided by the actual observation values for that period. Then, the average absolute percentage error. This approach is useful when the size or size of predictive variables is important in evaluating the accuracy of predictions. MAPE indicates how much error in forecasting is compared to the real value in the series (Prayudani et al., 2019).

R^2: The R^2 (R-squared) metric is used to evaluate the effectiveness of a regression model. It indicates the proportion of variance in the target variable explained by the model and essentially reflects how accurate the predicted values are compared to the actual values. A negative R^2 does not indicate that the solution is wrong, it indicates that the model performs poorly relative to the real data; This may be due to over-tuning of training data or poor performance on testing data.

Table 2. Performance Results of the Models

MODEL	MSE	RMSE	MAE	MAPE	R^2
Neural Network	11.044	3.323	2.183	0.123	0.998
KNN	7804.644	88.344	51.236	0.852	-0.189
Random Forest	7495.025	86.574	50.120	0.815	-0.142
Adaboost	7334.250	85.640	49.250	0.938	-0.117
Tree	7171.599	84.685	47.688	0.620	-0.092

5. Conclusion

Social media has now become a platform that people constantly prefer for shopping. Companies that can observe these purchasing behaviours in the best way certainly achieve success. Companies that can find out what their customers want to get the best results in this context. In this study, the preliminary estimation for sales to be realized within social media and websites was successfully realized. Before the social media discount campaign was launched in the nuts sector, how successful the campaign would be was estimated based on various parameters and previously available data, and 98% success was achieved in test results. As a result of the paper, artificial neural networks provide the best results among algorithms. With the results was obtained, the company using our application will be able to activate the transactions that other companies plan to do very quickly. Companies pay a lot of money to agencies for advertising shoots, it becomes a big waste to shoot these videos and not get trending and interaction, the studies using artificial neural networks prevented this.

For this reason, it becomes clear how important the determining parameters, which are written in an ordered manner, are. Before the campaign starts, it has been determined which route should be followed and which way this tracking should be done. In addition, the entry of sponsored ads into a campaign video is a widely used method on Instagram, and if the desired interaction is not received, the investments made in this video again lead to waste. The preparation of raw materials and packaging materials for the campaign, which is one of the most important costs, also entails inventory costs, and if this is not prepared well, customer evasion costs arise in campaigns, all these costs disappear after the consistency of the campaign is predicted in advance and provide a great profit to the company.

References

Tekerek, A. (2011). Data mining processes and open source data mining tools. Academic Informatics, 11, 2-4.

Ishak, A., Siregar, K., Ginting, R., & Afif, M. (2020). Orange software usage in data mining classification method on the dataset lenses. In *IOP Conference Series: Materials Science and Engineering*, 1003 (1), 012113. IOP Publishing.

Eroğlu, Y. (2020). Forecasting models for Covid-19 cases of turkey using artificial neural networks and deep learning. *Endüstri Mühendisliği*, *31*(3), 353-372.

Aydın, N. (2017). Where is e-commerce going in the world. *Реформа*, (1), 52-60.

Akçadağ, M. & Keklik, B. (2021). E-commerce applications of public employees on special days: Sivas province PTT example V. International Kaoru Ishikawa Business Administration And Economy Congress, 21-36, May 2012, Ankara, Turkiye.

Altman, N. S. (1992). An introduction to kernel and nearest-neighbour nonparametric regression. *The American Statistician*, *46*(3), 175-185.

Deniz, R. B. (2020). A review of marketing strategies and applications on Instagram. European Journal of Managerial Research (Eujmr), 4(7), 150-167.

Schapire, R. E. (2013). Explaining adaboost. In empirical inference: festschrift in honour of Vladimir *N.* Vapnik (pp. 37-52). Berlin, Heidelberg: Springer Berlin Heidelberg.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*, 197-227.

Köksoy, O. (2006). Multiresponse robust design: Mean square error (MSE) criterion. *Applied Mathematics and Computation*, *175*(2), 1716-1729.

Prayudani, S., Hizriadi, A., Lase, Y. Y. & Fatmi, Y. (2019). Analysis accuracy of forecasting measurement technique on random K-nearest neighbor (RKNN) using MAPE and MSE. In *Journal of Physics: Conference Series*, 1361(1), 012089. IOP Publishing.

Aksu, G., & Güzeller, C. (2019). Büyük veri: Sosyal bilimler ile eğitim bilimlerinde kullanımı ve uygulama alanları. *Akdeniz İnsani Bilimler Dergisi*, 9(1), 13-26.

Imandoust, S. B. & Bolandraftar, M. (2013). Application of k-nearest neighbor (KNN) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications*, *3*(5), 605-610.

Wang, P., Zhang, Y., & Jiang, W. (2021, June). Application of K-Nearest neighbor (KNN) algorithm for human action recognition. In *2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (Vol. 4, pp. 492-496). IEEE.

Zhou, X., Wang, H., Xu, C., Peng, L., Xu, F., Lian, L. & Li, G. (2022). Application of kNN and SVM to predict the prognosis of advanced schistosomiasis. *Parasitology Research*, *121*(8), 2457-2460.

Hosseini, S., & Fard, R. H. (2021). Machine learning algorithms for predicting electricity consumption of buildings. *Wireless Personal Communications*, *121*, 3329-3341.

Journal of Information Analytics, 1(1), 12-22

Journal of Information Analytics

Submit 05.12.2024
Revision 16.01.2025
Accept 18.02.2025
Publish 12.03.2025

Open Access

Research Article

Machine Learning-Based Forecasting of Minimum Wage in Turkey: A Case Study (2005-2024)

Nesrin Altınyüzük Gezer[1], Yunus Eroğlu[2], Suleyman Mete[2]

[1]Graduate School of Natural & Applied Sciences, Gaziantep University, Gaziantep, Türkiye

[2]Department of Industrial Engineering, Gaziantep University, 27100 Gaziantep, Türkiye

altinyuzuknesrinn@hotmail.com, eroglu@gantep.edu.tr, smete@gantep.edu.tr

Abstract

The minimum wage represents the lowest legally permissible level of compensation that employers are obligated to pay their employees. Determining the minimum wage involves a complex interplay of economic, social, and political factors, making it a crucial indicator for labor market policies. This study examines Turkey's minimum wage trends over the period 2005-2024, leveraging historical data to uncover the key determinants influencing wage adjustments. The aim is to predict the 2025 minimum wage by identifying the parameters that have significant effects on the determination of the minimum wage. Historical data was analyzed using machine learning and a forecasting analysis was performed for the 2025 minimum wage. In this context, studies were carried out using machine learning algorithms such as AdaBoost, Neural network, Linear regression. Orange 3 program was used with the data for machine learning and prediction processes. This research contributes to the literature by demonstrating the applicability of machine learning in economic forecasting, providing valuable insights for policymakers, employers, and labor market stakeholders.

Keywords Minimum wage, machine learning, forecasting, prediction models

Citation: Altınyüzük Gezer, N., Eroğlu, Y., & Mete, S. (2025). Machine learning-based forecasting of minimum wage in Turkey: A case study (2005-2024). *Journal of Information Analytics*, 1(1), 12-22.

Corresponding Author: Suleyman Mete smete@gantep.edu.tr

1. Introduction

Wage can be defined as the monetary level that employees earn because of physical or mental work to sustain their lives and meet their needs and the needs of their dependents (Çınar and Öz, 2018). A minimum wage is the lowest remuneration that employers can legally pay their employees. The purpose of minimum wages is to protect workers against unduly low pay and to help ensure a fair and equitable share of fruits of progress to all, and a minimum wage to all who are employed and in need of such protection (ILO, 2015).

This study aims to estimate the minimum wage in Turkey for the year 2025 by analyzing critical economic indicators that influence its determination. Various parameters play an active role in shaping the minimum wage, reflecting the intricate dynamics of the country's economy and labor markets. Specifically, the parameters identified in this study include the Wholesale Price Index (WPI), the Producer Price Index (PPI), Turkey's unemployment rate, and exchange rates for major currencies such as the USD and EUR. These indicators have been selected based on their historical correlation with wage adjustments and their capacity to reflect broader economic trends.

To ensure a robust analysis, the study incorporates a comprehensive dataset covering the period from 2005 to 2024. This dataset includes the historical values of WPI, PPI, unemployment rates, exchange rates for the USD and EUR, and the corresponding minimum wage figures during this timeframe. By analyzing the interplay between these variables, the study seeks to uncover patterns and relationships that can inform the projection of the minimum wage for 2025. This approach provides a holistic perspective, taking into account both domestic economic pressures, such as unemployment, and external factors, like currency fluctuations, which significantly impact Turkey's economic environment.

The findings of this study offer valuable insights for a wide range of stakeholders. Policymakers and labor organizations can leverage this data to make informed decisions during wage-setting negotiations, ensuring that the minimum wage reflects both the needs of workers and the realities of the economic environment. Moreover, these insights serve as a crucial resource for businesses and organizations, enabling them to prepare accurate and strategic budget forecasts for the coming year. By aligning financial planning with anticipated changes in labor costs, companies can better navigate economic uncertainties and maintain operational efficiency. This comprehensive analysis underscores the importance of integrating economic trends into decision-making processes, benefiting all parties involved in Turkey's labor market dynamics.

2. Literature Review

There is limited academic research that directly focuses on minimum wage estimation using machine learning. However, there are some studies that apply machine learning methods to salary and wage estimation.

Ghei and Lee (2020) discuss a study using machine learning methods to predict annual wages. The study compares the performance of basic Mincer regression on Current Population Survey (CPS) data with advanced machine learning algorithms such as XGBoost, LightGBM and deep learning. The best results are obtained with the Stacking method, which is a combination of different models. For forecasts limited to a small set of variables, the machine learning methods provided only a modest improvement compared to Mincer regression, but greatly outperformed forecasts with an expanded set of variables. In particular, Stacking provided significant improvements in the RMSE-log, RMSE-level and absolute deviation measures, combining the strengths of different algorithms to achieve higher prediction accuracy.

Çınar and Öz (2018) applying the human capital model to wage estimation for the service sector in Turkey, it analyzes the effect of individuals' characteristics such as education and experience on wage levels. In the study, a survey was applied to 2000 people working in Bursa. The study examined the contribution of years of education, work experience and other demographic factors to wage differences based on Jacob Mincer's wage equation. The data were collected from individuals working in the service sector and analyzed with econometric methods. The findings showed that education level and work experience have a significant and positive effect on wages, but these effects offer a diminishing return after a certain level. The study emphasizes that service sector wage policies can be improved with strategies aimed at developing human capital.

Cazcarra (2024) conducted a study analyzing the impact of increasing the minimum wage on income inequality between 2001-2021 in Spain. Using national census data provided by the Spanish Tax Administration, he showed that increasing the minimum wage reduced income inequality. It was found that increasing the minimum wage did not lead to inflation or unemployment, but rather increased net employment, kept prices under control, and increased company profit margins. Analysis using Multivariate Linear Regression, Random Forest Regressor, and Time Series Regression Model machine learning models revealed that increasing the minimum wage increased the country's wealth, increased employment and company profits, and was an effective method for wealth redistribution. For example, the multivariate linear regression model explained 99.3% of the Gini index, using p-values and coefficients to determine the effects of independent variables on the Gini index. The random forest regression model optimized the mean square error (MSE) to 0.0039 and determined the significance of the variables. These findings show that the minimum wage increase is effective in reducing income inequality and increasing economic welfare

3. Material and Method

In the application part of the study, the last 20 years of dataset was used. In this context, WPI (Wholesale Price Index), PPI (Producer Price Index), Turkey's unemployment rate and exchange rates data between 2005-2024 were collected from TÜİK. Table 1 shows the last 20 years of data on the parameter values affecting the minimum wage.

Table 1. Last 20 years wage dataset

Year	Minimum Wage (TRY)	Year+1	WPI (Wholesale Price Index)	PPI (Producer Price Index)	Unemployment Rate (%)	USD/TRY Average Exchange Rate	EUR/TRY Average Exchange Rate
2005	350,15	380,46	7,72	2,66	10,6	1,3405	1,67
2006	380,46	419,15	9,65	11,58	10,2	1,4297	1,798
2007	419,15	503,26	8,39	5,94	10,3	1,3003	1,7773
2008	503,26	546,48	10,06	8,11	11	1,2976	1,8969
2009	546,48	599,12	6,53	5,93	14	1,5457	2,1508
2010	599,12	658,95	6,4	8,87	11,9	1,499	1,9886
2011	658,95	739,8	10,45	13,33	9,8	1,6708	2,3244
2012	739,8	803,68	6,16	2,45	9,2	1,7922	2,3041
2013	803,68	891,03	7,4	6,97	9,7	1,9033	2,529
2014	891,03	1000,54	8,17	6,36	9,9	2,1865	2,9042
2015	1000,54	1.300,99	8,81	5,71	10,3	2,7191	3,0187
2016	1.300,99	1.404,06	8,53	9,94	10,9	3,0181	3,3375
2017	1.404,06	1.603,12	11,92	15,47	10,9	3,6445	4,1159
2018	1.603,12	2.020,90	20,3	33,64	11	4,8301	5,6789
2019	2.020,90	2.324,71	11,84	7,36	13,7	5,6712	6,3481
2020	2.324,71	2.825,90	14,6	25,15	13,2	7,0034	8,014
2021	2.825,90	5.500,35	36,08	79,89	13,4	8,8557	10,4408
2022	5.500,35	11.402,32	64,27	97,72	10,4	16,5512	17,3642
2023	11.402,32	17.002,12	64,77	44,22	9,4	23,7482	25,6852
2024	17.002,12	?	44,38	28,52	9	32,7825	35,4779

The aim is to make a minimum wage estimate for 2025 by teaching the last 20-year values with machine learning. Machine learning techniques were employed for the analysis process. Initially, the dataset was loaded, preprocessed, and subjected to exploratory data analysis to understand its structure and key features. The data was then split into training and testing subsets, ensuring proper validation. Various machine learning algorithms, specific models such as Linear Regression, Gradient Boosting, Neural Network, Adaboost, Stochastic Gradient Descent were trained on the training subset using cross-validation to optimize hyperparameters and minimize overfitting.

After training, the performance of each model was evaluated using appropriate metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R-squared (R^2) etc. Based on these metrics, the model yielding the best results was selected. Finally, the chosen model was used to generate predictions on unseen data, and the results were analyzed to interpret the model's behavior and assess its real-world applicability.

Estimating the minimum wage with data from previous years is a time series analysis. In time series analysis, it is a common approach to predict future values using past data. In order to predict future values of economic variables, such as minimum wage forecasts, historical data needs to be analyzed. In this context, machine learning models are usually trained and tested using the 1-year shift method. While this method allows the model to learn based on historical data, it also allows to evaluate how the model performs not only with training data but also with new data that it will actually encounter in the testing phase. In other words, testing by shifting by one year provides a more realistic measure of the model's ability to predict the future and reduces the risk of overfitting. This strategy helps the model learn the changing dynamics over time, so that it can make more accurate predictions when faced with future data.

In this study, the minimum wage data in the dataset has been shifted by 1 period and a column named "Year+1" has been added to calculate the 2025 minimum wage as target.

After the data was exported to Orange with the "File" function, it was displayed as a table with the "Data table" function and the data was checked. Then, with the "Select Columns" function, unnecessary data was ignored, and the target data was determined as follows. In Figure 1, it is seen that the data to be estimated is Year+1.

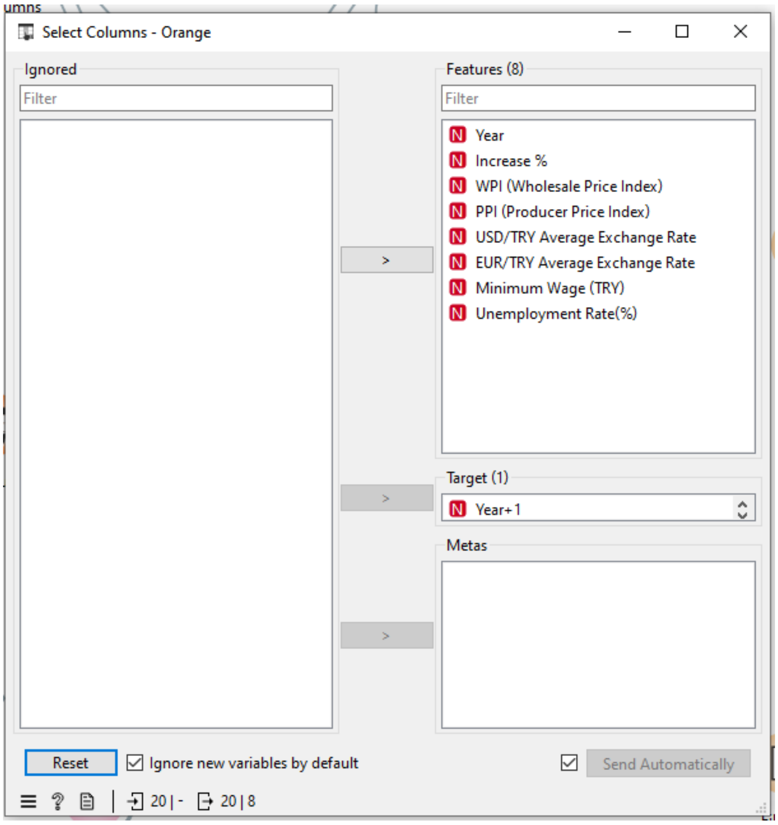

Figure 1. Determining target data

Five machine learning techniques were used in the analysis with Orange. These are Neural Network, Adaboost, Linear Regression, Stochastic Gradient Descent, Gradient Boosting learning methods.

3.1 Machine Learning Algorithms

a. Neural Network

Artificial Neural Network (ANN) has been a hot topic in artificial intelligence since the 1980s. It abstracts the human brain neural network from the perspective of information processing, establishes a simple model and composes different networks according to different connections (Dong and Hu, 1997). Trying to simulate brain neural network processing, memory information in the way of information processing.

In engineering and academia are often directly referred to as neural network or neural network. Neural network is a computing model, by a large number of nodes (or neurons) connected to each other (Jenkins and Tanguay, 1995). Each node represents a specific output function, called the activation function. The connection between every two nodes represents a weight for the signal passing through the connection, which is called the weight, which is equivalent to the memory of the artificial neural network (Bnlsabi, 1993). The output of the network will vary depending on how the network is connected, the weight value, and the incentive function. However, the network itself is usually an approximation to some kind of algorithm or function in nature, or it may be an expression of a logic strategy (Luo et al., 1998).

b. Adaboost

The AdaBoost algorithm corrects the misclassifications made by weak classifiers, and it is less susceptible to overfitting than most learning algorithms. Recognition performances of the AdaBoost-based classifiers are generally encouraging (Freund and Schapire, 1997).

c. Linear Regression

Linear regression is the building block for many modern modeling tools. In particular, when the sample size is small or the signal is relatively weak, linear regression often provides a satisfactory approximation to the underlying regression function (James et al., 2023).

d. Stochastic Gradient Descent (SGD)

Used to provide optimized learning processes on large datasets, this method provides a stochastic approach to minimize the cost function.

e. Gradient Boosting

Based on the principle of boosting tree-based models, this method aims to gradually reduce errors by succeeding in adding weak learners. With these 5 machine learning techniques in Orange, the dataset was first trained and then predicted. Figure 2 shows the process flow in Orange.

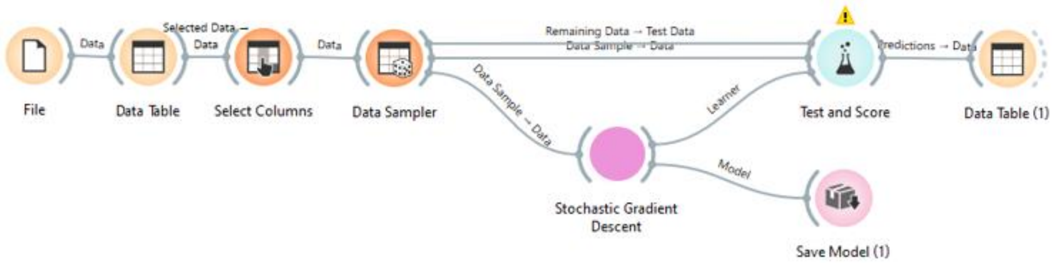

Figure 2. Process flow in Orange

The data analysis process began with loading data into Orange. In the first step, the general structure of the data was examined by providing data representation. Then, the target value (the output to be predicted) and the inputs (the variables used by the model for prediction) were determined. In order to make a more accurate assessment of the data, the training and test data were separated into separate groups by applying the sample separation (train-test split) method. In this step, the 'Test & Score' function used for testing and model validation was activated. Along with the sample separation process, the testing process was carried out by matching the model with the training data. The results were observed, and the performance of the model was evaluated. Figure 3 shows the flow of loading the model into Orange to make predictions.

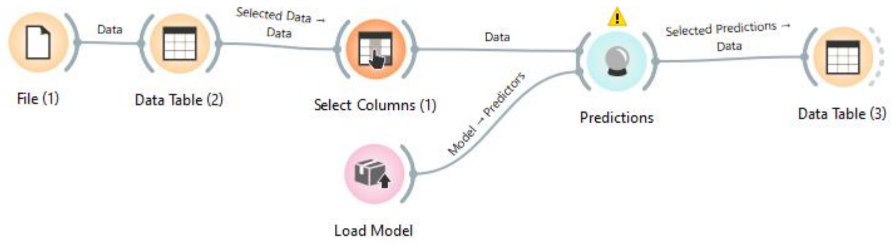

Figure 3. Loading the model with Orange

The mean squared error (MSE) or mean squared deviation (MSD) of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the true value.

If vector n of predictions is generated from a sample of n data points on all variables, and Y is the vector of observed values of the variable being predicted, with $\hat{Y}$ being the predicted (e.g. as from a least-squares fit), then the within-sample MSE of the predictor is computed as (Wikipedia, 2024a):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

RMSE (Root Mean Squared Error): The root mean square deviation (RMSD) or root mean square error (RMSE) is either one of two closely related and frequently used measures of the differences between true or predicted values on the one hand and observed values or an estimator on the other.

The RMSD of an estimator $\theta^{\wedge}$ with respect to an estimated parameter θ is defined as the square root of the mean squared error (Wikipedia, 2024b):

$$\mathrm{RMSD}(\hat{\theta}) = \sqrt{\mathrm{MSE}(\hat{\theta})} = \sqrt{\mathrm{E}((\hat{\theta}-\theta)^2)}.$$

MAE (Mean Absolute Error): Measures the average absolute difference between predicted and actual values. It is more robust to outliers than MSE (Eroğlu, 2024).

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n}.$$

It is thus $|e_i|=|y_i-x_i|$ and arithmetic average of the absolute errors , where y_i is the prediction and x_i the true value (Wikipedia, 2024c).

MAPE (Mean Absolute Percentage Error): Shows the percentage difference between predicted and actual values, useful for understanding relative errors (Eroğlu, 2024).

$$\mathrm{MAPE} = 100\frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

Where A_t is the actual value and F_t is the forecast value. Their difference is divided by the actual value A_t. The absolute value of this ratio is summed for every forecasted point in time and divided by the number of fitted points n (Wikipedia, 2024d).

R^2 (R-Square): Represents the proportion of variance in the target variable explained by the model. An R^2 value closer to 1 indicates a better fit (Eroğlu, 2024).

R^2 formulation is as below where SS_{res} is residual sum of squares and SS_{tot} total sum of squares; (Wikipedia, 2024e).

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

4. Results and Discussion

The results obtained in the study were compared using five different metrics (MSE, RMSE, MAE, MAPE and R^2) to evaluate the performance of each model. These metrics comprehensively measure the accuracy, error rate and overall performance of each model. Table 2 shows the performance results of the machine learning models.

Table 2. Results of machine learning models

Model	MSE	RMSE	MAE	MAPE	R^2
Neural Network	173.595.602	416.648	261.018	0.111	0.992
Stochastic Gradient Descent	66.205.537	257.304	204.619	0.222	0.997
Linear Regression	473.863.523	217.684	164.023	0.178	0.998
AdaBoost	4.265.772	65.313	37.237	0.076	1
Gradient Boosting	0.020	0.142	0.106	0.000	1

MSE (Mean Squared Error) and RMSE (Root Mean Squared Error) values indicate the distance of the model's predictions from the actual data. Lower values in these metrics indicate that the model makes more accurate predictions.

MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error) measure error rates as absolute values. These metrics reflect how accurate the predictions are and how many errors the model makes.

R^2 (R-squared) is a coefficient that shows how well the model fits the data and the explanatory power of the independent variables on the dependent variable. A high R^2 value indicates that the model explains the data set well.

Gradient Boosting and AdaBoost Models

When the results are examined, Gradient Boosting and AdaBoost models stand out with their low error rates and high accuracy values. Both models have low MSE, RMSE, MAE and MAPE values and exhibit high R^2 values.

Gradient Boosting has modeled the data set flawlessly by showing excellent performance in all metrics. The fact that the MSE and RMSE values are almost zero reveals that the model's predictions are extremely close to the real data. The fact that the MAE and MAPE values are also close to zero shows that the model minimizes prediction errors. The R^2 value is also close to 1, indicating that this model explains the data by 100%. These results indicate that the Gradient Boosting model is the model that produces the most accurate and reliable predictions for this data set.

AdaBoost is also a model that exhibits high performance in a similar way. The MSE and RMSE values are quite low, and the MAE and MAPE values are at a minimum level. This shows that AdaBoost has a high capacity to make accurate predictions, and that the model's performance is quite reliable. In addition, the R^2 value is close to 1, explaining the data with very high accuracy. Although AdaBoost shows slightly higher error rates compared to Gradient Boosting, it still stands out as a strong alternative.

Neural Network and Linear Regression Models

Neural Network and Linear Regression models stand out with higher error rates and lower accuracy levels.

The Neural Network model demonstrates a high performance in terms of R^2 value, explaining 99.2% of the data set. However, since the MSE, RMSE, MAE and MAPE values are high, it is seen that there are large errors in the model's predictions. This shows that although the Neural Network model learns the data quite well, its overall accuracy is not sufficient, and that improvement is needed. In particular, the MAPE value of 11.1% indicates that the model's predictions are significantly incorrect.

Linear Regression performed very poorly on this dataset. The high MSE, RMSE, MAE and MAPE values indicate that the linear regression model's estimates are far from the real data. In particular, the MSE value (473,863,523) is quite high, and the model's estimates contain a lot of errors. Although the R^2 value is high at 0.998, this model was generally inadequate due to its high error rates.

Stochastic Gradient Descent

The Stochastic Gradient Descent model provides much lower values in error metrics such as MSE and RMSE than the Linear Regression and Neural Network models, indicating that the accuracy of the predictions is high. This shows that SGD manages to approach the lowest error more quickly and efficiently when updating the model parameters.

5. Conclusion

As a result, Gradient Boosting achieved the most successful results among the models tested within the scope of this study. Gradient Boosting modeled the dataset very well with low error rates and high accuracy values and provided excellent prediction accuracy. AdaBoost also stood out as a very effective model and exhibited high performance in terms of prediction accuracy and error rates.

However, Neural Network and Linear Regression models attracted attention with their higher error rates and low accuracy levels and produced less effective results for this dataset. Although Neural Network has learned the dataset well with its high R^2 value, its error rates in predictions are quite high. Linear Regression can be considered as the least successful model with both low accuracy and high error rates.

The Stochastic Gradient Descent (SGD) model, when compared to other models, still has some error rates compared to advanced techniques such as Gradient Boosting and AdaBoost, but generally achieves better results compared to Linear Regression and Neural Network models. SGD's MSE, RMSE, MAE and

MAPE values indicate lower error rates than the performance of these models. Since the R^2 value shows that the model explains the data by 99.7%, it can be said that it has a very successful prediction capacity.

In this context, it can be said that the Stochastic Gradient Descent model shows slightly lower accuracy than Gradient Boosting and AdaBoost but still offers a strong alternative. This model can be an effective option for those looking for a simpler and faster solution.

As a result, it can be said that the Gradient Boosting model is the most successful model on this dataset, offering excellent prediction accuracy with minimal errors. AdaBoost also stands out as an important alternative, demonstrating high performance and accuracy, although it shows slightly higher error rates compared to Gradient Boosting. Additionally, Stochastic Gradient Descent (SGD) performs well, delivering lower error rates than models like Linear Regression and Neural Networks, though it does not achieve the same level of accuracy as Gradient Boosting and AdaBoost. Therefore, Gradient Boosting and AdaBoost are the models to prefer for this dataset due to their superior performance, while Stochastic Gradient Descent provides a reliable alternative with a good balance between performance and computation efficiency.

When the estimates made with the model are evaluated, the Stochastic Gradient Descent (SGD) model estimated the minimum wage as 22170 TRY for 2025. The minimum wage announced for 2025 in Turkey was determined as 22104 TRY. Compared to the estimates of other models, the SGD model presented a result that is quite close to the actual value.

References

Bnlsabi, A. (1993). Some analytical solutions to the general approximation problem for feed forward neural networks. Neural Networks, 6, 991–996.

Cazcarra, M. L., (2024). Machine learning analysis of the impact of increasing the minimum wage on income inequality in Spain from 2001 to 2021," arXiv preprint arXiv:2402.02402, 2024. https://arxiv.org/abs/2402.02402

Çınar, M., & Öz, R., (2018). Hizmet sektörü için ücret tahmini: insan sermayesi modeli, International Conference on Economics and Administrative Sciences (IZCEAS), Bursa, Turkiye, Aralık.

Dong, J., & Hu, S. (1997). The progress and prospects of neural network research. Information and Control, 26(5), 360–368.

Eroğlu, Y. (2024). IMDb score estimation using movie dialogues: A text mining and machine learning hybrid approach, International Data Science and Statistics Congress, Ankara, Türkiye.

Freund, Y., & Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences*, 55 (1), 119–139. https://ieeexplore.ieee.org/document/4454220.

Ghei, D., & Lee, S.M., (2020). Annual wage prediction: machine learning competition, https://static1.squarespace.com/static/62014b6e32679565e6c69891/t/621e43aaace5152a7b76ff1a/1646150570815/2020_Winners_sangminlee-dhanajay-report-updated.pdf

International Labor Organization, "How to define a minimum wage?," https://www.ilo.org/global/topics/wages/minimum-wages/definition/WCMS_439072/lang--en/index.htm.

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor J. (2023). Linear regression, in *An Introduction to Statistical Learning: With Applications in Python*, Cham: Springer International Publishing, 69–134.

Jenkins, B.K., & Tanguay, A.R. (1995). Handbook of neural computing and neural networks. Boston: MIT Press.

Luo, Z.H., Xie, Y., & Zhu, C. (1997). The study of convergence of CMAC learning process. Acta Automatic Sinica, 23(4), 455–461.

Wikipedia "Mean absolute percentage error," Wikipedia, Oct. 4, 2024 https://en.wikipedia.org/wiki/Mean_absolute_percentage_error

Wikipedia, "Coefficient of determination," *Wikipedia*, Dec. 17, 2024. https://en.wikipedia.org/wiki/Coefficient_of_determination.

Wikipedia, "Mean absolute error," Wikipedia, Dec. 21, 2024. https://en.wikipedia.org/wiki/Mean_absolute_error.

Wikipedia, "Mean squared error," Wikipedia, Oct. 22, 2024. https://en.wikipedia.org/wiki/Mean_squared_error.

Wikipedia, "Root means square deviation," Wikipedia, Oct. 15, 2024. https://en.wikipedia.org/wiki/Root_mean_square_deviation.

Journal of Information Analytics, 1(1), 23-37

Journal of Information Analytics

Submit	12.12.2024
Revision	26.02.2025
Accept	13.03.2025
Publish	18.03.2025

Open Access

Research Article

Analyzing the Emergency Assembly Points Criteria Using the Best-Worst Method under Interval Type-2 Fuzzy Sets

Taner Yigit[1], Erkan Celik[2]

[1] Directorate of Planning and Urban Planning, Mus Municipality, Muş, Türkiye

[2]Department of Transportation and Logistics, İstanbul University, İstanbul, Türkiye

Abstract

More than 400 high-impact natural disasters affect the global population each year. There have been many disasters in our country throughout its history. Earthquakes are the most common and most damaging type of disaster in our country. Within the scope of this paper, especially in the first 72 hours after the earthquake, which is called the golden times, the priority of the emergency assembly areas and the prioritization of the emergency assembly points, which should be planned within the scope of the disaster management system, were examined in order for the citizens who were not yet affected by the disaster or were affected with less damage to continue their shelter and vital activities. .

First, 5 main criteria and 13 sub-criteria were determined within the scope of emergency assembly points site selection criteria. The main criteria were determined as the preferability of the land, electrical infrastructure, plumbing system, safety and security and proximity. The weights of the criteria for location selection of emergency assembly points were calculated with 20 different decision makers who are experts in their fields. At this stage, the weights of each main criterion and sub-criteria were calculated using the Best-Worst approach in the literature. According to these results, the preferability of the land was determined as the most important main criterion, while landslides, flooding, etc. was determined as an important sub-criterion.

Keywords Disaster, disaster management system, earthquake, emergency assembly areas,

intermittent type-2 fuzzy sets, best-worst method

Citation: Yigit, T., & Celik, E. (2025). Analyzing the emergency assembly points criteria using the best-worst method under interval type-2 fuzzy sets. *Journal of Information Analytics*, 1(1), 23-37.

Corresponding Author: Erkan Celik erkancelik@istanbul.edu.tr

1. Introduction

Disaster management systems aim to prevent, mitigate, or eliminate the impacts of any disaster, regardless of its type. These systems involve coordinated planning that encompasses pre-disaster phases (risk reduction and preparedness) and post-disaster phases (response and recovery), ensuring the efficient and purposeful use of resources for the benefit of society and all living beings. Earthquakes, the most common natural disaster in our country, are a primary focus for disaster risk management efforts. As part of these efforts, emergency assembly areas were designated in collaboration between AFAD and municipalities in 2018.

In the literature, emergency assembly areas also referred to as emergency gathering areas or post-disaster assembly areas are predefined locations identified by AFAD and relevant municipalities. These areas are intended to prevent panic during the critical first 72 hours following a disaster, facilitate efficient information exchange, and provide safe locations where the public can gather away from hazardous zones. These areas are free of physical risks and are crucial for post-disaster management. The location of emergency assembly areas within urban settings is a critical aspect of urban planning and disaster management. These areas must meet specific criteria, including appropriate distribution, adequate size, visibility, accessibility, and suitable infrastructure features.

This study focuses on prioritizing 16 emergency assembly points in the Muş province. To achieve this, the criteria for emergency assembly area selection were first examined in the literature. Five main criteria and 13 sub-criteria were determined, including land suitability, electrical infrastructure, sanitation systems, safety and security, and proximity. Since multiple criteria are involved in prioritizing emergency assembly points, this problem is considered a multi-criteria decision-making (MCDM) problem. The weights of the criteria were calculated with the participation of 20 experts in the field. The Best-Worst Method (BWM), developed by Rezaei (2015), was employed to determine the weights of each main and sub-criterion.

The BWM is a multi-criteria decision-making method used in various fields such as disaster management, logistics and supply chain management, engineering, and agriculture. One notable feature of the BWM is its reduced need for pairwise comparison data, making it more consistent than the Analytical Hierarchy Process (AHP). According to the results of the BWM, land suitability was identified as the most important main criterion, while sub-criteria such as landslides and flooding were found to be significant considerations.

2. Proposed Method

This section presents the fundamental steps of the Best-Worst Method (BWM) under interval type-2 fuzz sets, which was used in this study to determine the importance weights of 5 main criteria and their 13 sub-criteria for prioritizing emergency assembly points. The BWM was developed using pairwise comparison for alternatives and criteria by Rezaei (2015). The best and worst criterion are used as two vectors as that needs fewer data here, and it leads to more reliability (Rezaei, 2016; Rezaei et al. 2016). Some fuzzy versions of the BWM is presented for different application areas. The triangular fuzzy BWM is applied by Hafezalkotob and Hafezalkotob (2017), Guo and Zhao (2017), and Moslem et al. (2020). Tian et al. (2018a) proposed F-BWM for calculating the risk factors of FMEA. Mou et al. (2016) proposed the intuitionistic fuzzy multiplicative BWM for healthcare. The best power plant alternative is selected by Omrani et al. (2018) using fuzzy BWM. Tian et al. (2018b) proposed the intuitionistic F-BWM for the green supplier selection problem. Wu et al. (2019) integrated the IT2Fs and BWM using centroids for green supplier selection problems. Altay et al. (2023) applied BWM under IT2FSs for for location selection of e-scooter sharing stations. A detailed survey about BWM was presented by Mi et al. (2019). This detailed review can be analyzed by interested researchers for taking inspire the later research related to the BWM. In this section, we will present the steps of the IT2F-BWM using center of area method.

Step 1. A set of decision criteria n is determined. The criteria $(c_1, c_1, \dots, c_n)$ is used to calculate the importance weights.

$$\tilde{\tilde{E}} = \begin{array}{c} \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} \\ \begin{pmatrix} \tilde{\tilde{e}}_{11} & \tilde{\tilde{e}}_{12} & \cdots & \tilde{\tilde{e}}_{1n} \\ \tilde{\tilde{e}}_{21} & \tilde{\tilde{e}}_{22} & \cdots & \tilde{\tilde{e}}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\tilde{e}}_{n1} & \tilde{\tilde{e}}_{n1} & \cdots & \tilde{\tilde{e}}_{nn} \end{pmatrix} \end{array} \tag{1}$$

where $\tilde{\tilde{e}}_{ij}$ shows the IT2F preference degree of criterion i over criterion j. IT2F pairwise comparisons on these n criteria can be applied based on the linguistic terms of decision-makers. All linguistic variables and IT2FSs are presented in Table 1. In this evaluation matrix, the diagonal elements are considered as Equally important (EI) and $\tilde{\tilde{e}}_{11}, \tilde{\tilde{e}}_{22}, \dots, \tilde{\tilde{e}}_{nn} = ((1;1;1;1;1;1),(1;1;1;1;0.9;0.9))$.

Table 1. Linguistic terms for importance weights (Celik et al. 2015)

Linguistic term	IT2FSs
EMI	((8;9;9;10;1;1), (8.5;9;9;9.5;0.9;0.9))
IV8	((7;8;8;9;1;1), (7.5;8;8;8.5;0.9;0.9))
VSMI	((6;7;7;8;1;1), (6.5;7;7;7.5;0.9;0.9))
IV6	((5;6;6;7;1;1), (5.5;6;6;6.5;0.9;0.9))
SMI	((4;5;5;6;1;1), (4.5;5;5;5.5;0.9;0.9))
IV4	((3;4;4;5;1;1), (3.5;4;4;4.5;0.9;0.9))
MMI	((2;3;3;4;1;1), (2.5;3;3;3.5;0.9;0.9))
IV2	((1;2;2;3;1;1), (1.5;2;2;2.5;0.9;0.9))
EI	((1;1;1;1;1;1), (1;1;1;1;0.9;0.9))

Step 2. The best and the worst criterion is decided using decision-maker preference.

Step 3. the preference of the best criterion and the worst criterion over all the other criteria is determined using IT2FSs. The resulting Best-to-Others (BtO) vector would be: $\tilde{\tilde{E}}_B = (\tilde{\tilde{e}}_{B1}, \tilde{\tilde{e}}_{B2}, \dots, \tilde{\tilde{e}}_{Bn})$,

where $\tilde{\tilde{e}}_{Bj}$ indicates the preference of the best criterion B over criterion j. It is clear that $\tilde{\tilde{e}}_{BB} = ((1;1;1;1;1;1),(1;1;1;1;0.9;0.9))$

The resulting Others-to-Worst (OtW) vector would be $\tilde{\tilde{E}}_B = (\tilde{\tilde{e}}_{1W}, \tilde{\tilde{e}}_{2W}, \dots, \tilde{\tilde{e}}_{nW})^T$

where a_{jW} indicates the preference of the criterion j over the worst criterion W. It is clear that $\tilde{\tilde{e}}_{WW} = ((1;1;1;1;1;1),(1;1;1;1;0.9;0.9))$.

Step 4. The optimal weights $(w_1^*, w_2^*, \dots, w_n^*)$ is calculated. In this process, the center of area is utilized. The constrained optimization model is built as Wu et al. (2019). The optimal weight for the criteria is the one where, for each pair of $\tilde{\tilde{w}}_B / \tilde{\tilde{w}}_j$ and $\tilde{\tilde{w}}_j / \tilde{\tilde{w}}_W$. The maximum absolute differences $\left|\tilde{\tilde{w}}_B / \tilde{\tilde{w}}_j - \tilde{\tilde{E}}_{Bj}\right|$ and

$\left|\tilde{\tilde{w}}_j / \tilde{\tilde{w}}_W - \tilde{\tilde{E}}_{jW}\right|$ should be minimized for finding optimal solutions. The consistency index as shown in Table 2 is used to calculate consistency ratio. It is created using the same process of Rezaei (2015) and Wu et al. (2019). The bigger the δ^*, the higher the consistency ratio.

Table 2. Consistency Index table

Linguistic Term	EI	IV	MMI	IV	SMI	IV	VSMI	IV	EMI
Defuzzified	0.975	1.95	2.925	3.9	4.875	5.85	6.825	7.8	8.775
CI	2.9582	4.4872	5.8948	7.2373	8.5373	9.8069	11.0533	12.2812	13.494

$$\min\max\left\{\left|\tilde{\tilde{w}}_B \,\tilde{\tilde{/}}\, w_j - \tilde{\tilde{E}}_{Bj}\right|, \left|\tilde{\tilde{w}}_j \,\tilde{\tilde{/}}\, w_W - \tilde{\tilde{E}}_{jW}\right|\right\}$$

$$s.t.\begin{cases} \sum_{j=1}^{n} COA\left(\tilde{\tilde{w}}_j\right) = 1 \\ w_{j1}^U \le w_{j1}^L, w_{j4}^L \le w_{j4}^U, \\ w_{j1}^L \le w_{j2}^L \le w_{j3}^L \le w_{j4}^L \\ w_{j1}^U \le w_{j2}^U \le w_{j3}^U \le w_{j4}^U, \\ w_{j1}^U \ge 0,\ j = 1, 2, \ldots, N \end{cases} \tag{2}$$

where $\tilde{\tilde{w}}_B = \left(\tilde{w}_B^U, \tilde{w}_B^L\right) = \left(\left(w_{B1}^U, w_{B2}^U, w_{B3}^U, w_{B4}^U; H_1\left(\tilde{w}_B^U\right), H_2\left(\tilde{w}_B^U\right)\right), \left(a_{B1}^L, a_{B2}^L, a_{B3}^L, a_{B4}^L; H_1\left(\tilde{w}_B^L\right), H_2\left(\tilde{w}_B^L\right)\right)\right)$

$$\tilde{\tilde{w}}_W = \left(\tilde{w}_W^U, \tilde{w}_W^L\right) = \left(\left(w_{W1}^U, w_{W2}^U, w_{W3}^U, w_{W4}^U; H_1\left(\tilde{w}_W^U\right), H_2\left(\tilde{w}_W^U\right)\right), \left(a_{W1}^L, a_{W2}^L, a_{W3}^L, a_{W4}^L; H_1\left(\tilde{w}_W^L\right), H_2\left(\tilde{w}_W^L\right)\right)\right)$$

$$\tilde{\tilde{w}}_j = \left(\tilde{w}_j^U, \tilde{w}_j^L\right) = \left(\left(w_{j1}^U, w_{j2}^U, w_{j3}^U, w_{j4}^U; H_1\left(\tilde{w}_j^U\right), H_2\left(\tilde{w}_j^U\right)\right), \left(a_{j1}^L, a_{j2}^L, a_{j3}^L, a_{j4}^L; H_1\left(\tilde{w}_j^L\right), H_2\left(\tilde{w}_j^L\right)\right)\right)$$

$$\tilde{\tilde{w}}_{B,j} = \left(\tilde{w}_{B,j}^U, \tilde{w}_{B,j}^L\right) = \left(\left(w_{B,j1}^U, w_{B,j2}^U, w_{B,j3}^U, w_{B,j4}^U; H_1\left(\tilde{w}_{B,j}^U\right), H_2\left(\tilde{w}_{B,j}^U\right)\right), \left(a_{B,j1}^L, a_{B,j2}^L, a_{B,j3}^L, a_{j4}^L; H_1\left(\tilde{w}_{B,j}^L\right), H_2\left(\tilde{w}_{B,j}^L\right)\right)\right)$$

$$\tilde{\tilde{w}}_{j,W} = \left(\tilde{w}_{j,W}^U, \tilde{w}_{j,W}^L\right) = \left(\left(w_{j1,W}^U, w_{j2,W}^U, w_{j3,W}^U, w_{j4,W}^U; H_1\left(\tilde{w}_{j,W}^U\right), H_2\left(\tilde{w}_{j,W}^U\right)\right), \left(a_{j1,W}^L, a_{j2,W}^L, a_{j3,W}^L, a_{j4,W}^L; H_1\left(\tilde{w}_{j,W}^L\right), H_2\left(\tilde{w}_{j,W}^L\right)\right)\right)$$

The maximum absolute gaps between $\left|\tilde{\tilde{w}}_B \,\tilde{\tilde{/}}\, w_j - \tilde{\tilde{E}}_{Bj}\right|$ and $\left|\tilde{\tilde{w}}_j \,\tilde{\tilde{/}}\, w_W - \tilde{\tilde{E}}_{jW}\right|$ are aimed to minimize to eliminate for $\left|\tilde{\tilde{w}}_B \,\tilde{\tilde{/}}\, w_j - \tilde{\tilde{E}}_{Bj}\right|$ and $\left|\tilde{\tilde{w}}_j \,\tilde{\tilde{/}}\, w_W - \tilde{\tilde{E}}_{jW}\right|$. The model is transformed to nonlinear programming for minimizing the absolute gap as $\delta^* = ((\delta^*; \delta^*; \delta^*; \delta^*; 1; 1), (\delta^*; \delta^*; \delta^*; \delta^*; 0.9; 0.9))$.

$$
\min \delta^*
$$

$$
s.t.\begin{cases}
\left|\tilde{\tilde{w}}_{B1}^{U}-\tilde{\tilde{w}}_{j1}^{U}\tilde{\tilde{w}}_{Bj,1}^{U}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{B2}^{U}-\tilde{\tilde{w}}_{j2}^{U}\tilde{\tilde{w}}_{Bj,2}^{U}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{B3}^{U}-\tilde{\tilde{w}}_{j3}^{U}\tilde{\tilde{w}}_{Bj,3}^{U}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{B4}^{U}-\tilde{\tilde{w}}_{j4}^{U}\tilde{\tilde{w}}_{Bj,4}^{U}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{B1}^{L}-\tilde{\tilde{w}}_{j1}^{L}\tilde{\tilde{w}}_{Bj,1}^{L}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{B2}^{L}-\tilde{\tilde{w}}_{j2}^{L}\tilde{\tilde{w}}_{Bj,2}^{L}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{B3}^{L}-\tilde{\tilde{w}}_{j3}^{L}\tilde{\tilde{w}}_{Bj,3}^{L}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{B4}^{L}-\tilde{\tilde{w}}_{j4}^{L}\tilde{\tilde{w}}_{Bj,4}^{L}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{j1}^{U}-\tilde{\tilde{w}}_{W1}^{U}\tilde{\tilde{w}}_{jW,1}^{U}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{j2}^{U}-\tilde{\tilde{w}}_{W2}^{U}\tilde{\tilde{w}}_{jW,2}^{U}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{j3}^{U}-\tilde{\tilde{w}}_{W3}^{U}\tilde{\tilde{w}}_{jW,3}^{U}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{j4}^{U}-\tilde{\tilde{w}}_{W4}^{U}\tilde{\tilde{w}}_{jW,4}^{U}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{j1}^{L}-\tilde{\tilde{w}}_{W1}^{L}\tilde{\tilde{w}}_{jW,1}^{L}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{j2}^{L}-\tilde{\tilde{w}}_{W2}^{L}\tilde{\tilde{w}}_{jW,2}^{L}\right|\le\delta^*,\\
\left|\tilde{\tilde{w}}_{J3}^{L}-\tilde{\tilde{w}}_{W3}^{L}\tilde{\tilde{w}}_{JW,3}^{L}\right|\le\delta^*,\left|\tilde{\tilde{w}}_{J4}^{L}-\tilde{\tilde{w}}_{W4}^{L}\tilde{\tilde{w}}_{jW,4}^{L}\right|\le\delta^*,\\
\sum_{j=1}^{n} COA\left(\tilde{\tilde{w}}_j\right)=1\\
w_{j1}^{U}\le w_{j1}^{L},\, w_{j4}^{L}\le w_{j4}^{U},\, w_{j1}^{L}\le w_{j2}^{L}\le w_{j3}^{L}\le w_{j4}^{L}\\
w_{j1}^{U}\le w_{j2}^{U}\le w_{j3}^{U}\le w_{j4}^{U},\, w_{j1}^{U}\ge 0,\ j=1,2,\ldots,N
\end{cases}
\tag{3}
$$

3. Application

3.1. Criteria for Selecting Assembly Areas

In its 2018 report, AFAD listed the criteria for determining assembly areas as follows: "Population density, accessibility, and ease of evacuation, suitability for access by the elderly and disabled, flat and unobstructed terrain, and areas that are not affected by secondary hazards such as fire, flood, or tsunami. The areas should also not be near seas, rivers, or locations prone to liquefaction, and they must be distant from fault lines. They should be close to residential areas but not affected by structural or non-structural hazards, and they should be public spaces that can provide basic needs such as electricity, water, and toilets" (AFAD, 2018). Studies in the literature on emergency assembly areas consider different criteria.

Çınar et al. (2018) examined the factors influencing assembly area planning in Karşıyaka District, İzmir Province. They highlighted that assembly areas were initially determined in 2006 under İzmir's provincial emergency assistance plan and later revised by AFAD in 2018 to include district-based assembly areas. Their study revealed that the 30 assembly areas designated for 15 of the 27 neighborhoods in Karşıyaka were insufficient in both quantity and size and did not meet international standards. Özel (2019) conducted a thesis study on assembly areas in Kastamonu Province. The purpose was to designate green spaces in zoning plans as assembly areas and propose additional areas where needed. His study found that the number of designated areas was inadequate and some were deemed risky. He emphasized that assembly areas should be open or green spaces large enough for the population, accessible, and evenly distributed. Taylan (2018) investigated whether the assembly areas in Çankırı Province were effectively planned to serve citizens post-disaster. He found that public awareness of these areas was low, and there were no regulations or laws setting national standards for assembly areas. Palazca (2020) analyzed the qualitative and quantitative characteristics of the 93 assembly areas in Denizli's 64 neighborhoods. Criteria included land use, slope, area size, and neighborhood population. He stated that assembly areas must be accessible, safe, and suitable for use during emergencies, with at least 2.5 m² of space per person. Dursun (2021) examined the standards and requirements for determining assembly areas in

Esenler District, Istanbul. His findings revealed that the assembly areas were insufficient, not easily accessible, and had low usability. Studies from various articles, theses, and reports emphasize the following main criteria for assembly areas: Accessibility, connectivity to road axes, usability, multi-functionality, ownership, and area size (Ju et al., 2012; Nappi & Souza, 2014; Kılcı et al., 2014; Trivedi & Singh, 2016; Yu & Wen, 2016; Çelik, 2017; Trivedi, 2018). Based on a literature review, this study identifies five main criteria and 13 sub-criteria, described below:

Soil Hardness, Slope, Topography, and Tree Presence: The hardness of the soil, slope, topographical features, and the presence of trees are crucial for shelter areas. The soil should be resistant to the effects of rainfall. The slope of the area should not be excessive, and rugged terrains should be avoided as much as possible. The presence of trees is particularly important for providing shade during hot summer months. Wetlands are unsuitable for shelter areas (Trivedi, 2018). Soil hardness is especially significant in terms of rainfall. Hard soil is less affected by precipitation and is more suitable for daily living (Kılcı et al., 2014).

Topography: Green areas are preferred for shelter areas due to their ability to provide shade and abundant oxygen, particularly during the summer months (Kılcı et al., 2014). Savannas are considered the most suitable type of terrain for these areas. Plains are more favorable than hilly terrains for this purpose (Kılcı et al., 2014).

Slope: According to data from the Turkish Red Crescent (Kızılay), the slope of the area should not exceed 7%, with the optimal slope ranging between 2% and 4% (Kılcı et al., 2014).

Ownership Status: Land controlled by central or local governments can be transformed into shelter areas more easily compared to privately or corporately owned lands. Publicly owned lands are more accessible (Trivedi, 2018). Compared to privately or corporately owned lands, public ownership facilitates the acquisition of permissions (Kılcı et al., 2014). In its 2002 report, JICA recommended that gathering areas should generally include green spaces, playgrounds, neighborhood or district parks, as well as the yards of schools, mosques, and hospitals. Each designated area should not be smaller than 500 m^2 (JICA and IBB, 2002).

Population Density (Per Capita Area): In its 2002 report, JICA recommended that areas designated as "Preliminary Evacuation Zones" should provide 1.5 m^2 per person. In their study, Tarabanis and Tsionas (1999) suggested an area of approximately 2 m^2 per person. However, in IZAMP and TAMP-Izmir, a significantly higher value of 4 m^2 per person, exceeding international standards, was specified (Çınar et al., 2018). Gathering areas, which are crucial to prevent chaos, should be selected based on the population of the settlement. Multiple areas, rather than a single large one, should be identified, and their locations should be communicated to the public (Taylan, 2018).

Accessibility for the Elderly and Disabled: Post-disaster and emergency gathering areas must be accessible to elderly and disabled individuals. The pathways leading to these areas and the routes continuing to safe zones should feature wheelchair-friendly paving and appropriate gradients.

Electrical Infrastructure: Electricity is a critical resource for maintaining daily life. Heaters, medical equipment when necessary, and many devices used today rely on electricity. Furthermore, shelters must be equipped with infrastructure to support communication, telecommunications, and alert systems (Trivedi, 2018).

Operational Electricity: Electricity is essential for heating and powering devices used in daily life. Areas must have optimal levels of electrical infrastructure (Kılcı et al., 2014).

Lighting Electricity: Since gathering areas are expected to be used during the first 72 hours following a disaster or emergency, it is crucial to provide sufficient lighting for citizens sheltered in these areas at night.

Telecommunication Facilities: The distance of designated gathering areas from telecommunication facilities is critical. These areas should support the network infrastructure of all GSM providers.

Sanitation System: Shelter areas must be equipped with infrastructure such as drainage and clean water systems. In particular, sewage infrastructure is vital to prevent the spread of epidemics and infectious diseases (Trivedi, 2018).

Drinking Water: Water, one of the basic human necessities, is essential not only for survival but also for cooking and personal hygiene (Kılcı et al., 2014).

Toilet Facilities: Sewage infrastructure should be available in the designated areas for citizens who will be sheltered for approximately 72 hours following a disaster or emergency (Kılcı et al., 2014).

Safety and Security: When planning the locations of shelter areas, safety and security are of utmost importance. Since shelter areas tend to be crowded, they should be situated on terrains that are not vulnerable to secondary natural disasters such as landslides, rockfalls, or floods (Trivedi, 2018).

Landslides, Floods, etc.: It is crucial to select gathering areas that are free from additional disaster risks and will not be affected by secondary disasters (Taylan, 2018).

Warning Systems (Audio Systems): Following a disaster or emergency, citizens require areas where they can obtain accurate information and gather safely. Gathering areas are especially critical during the first 12 to 24 hours post-disaster to ensure that victims have access to reliable information (Taylan, 2018).

Proximity: Potential areas should be located near healthcare facilities, supply depots, and transportation routes (Trivedi, 2018).

Distance to Settlements (Accessibility): Proximity to healthcare facilities is vital for early intervention for those in need. Similarly, being close to supermarkets or supply depots is important for the provision of goods to these areas (Kılcı et al., 2014). Each block within the zoning plan should be no more than a 15-minute walk or 500 meters from a gathering area. In cases where access to a designated gathering area is disrupted (e.g., due to damaged or blocked transportation networks), continuity must be ensured with alternative gathering areas (Çınar et al., 2018).

Distance from Potential Disaster Debris: Gathering areas, where citizens may stay for up to 72 hours post-disaster, should be located at a safe distance from potential disaster debris zones to facilitate effective crisis management.

3.2. Best-Worst Method Results

In this section, the importance weights of the 5 main criteria and the 13 sub-criteria for determining emergency gathering points are determined using the Best-Worst Method. In this phase, evaluations from 20 experts with competence in the field were considered. These 20 experts work in various units in Muş province, specializing in disaster and emergency management. Three of the experts are from the Muş Provincial Disaster and Emergency Directorate, thirteen work at Muş Municipality, Muş Provincial Governorship Environmental and Urbanism Directorate, and two at Muş Provincial Special Administration. One expert holds a master's degree, one holds an associate degree, and the rest have bachelor's degrees. Among the experts, 2 are Electrical Engineers, 1 is a Geomatics Engineer, 1 is a Geomatics Technician, 7 are Civil Engineers, 1 is a Geological Engineer, 4 are Architects, 3 are Urban Planners, and 1 is a branch manager. 11 of the experts have 5-10 years of work experience, 4 have 10-15 years of work experience, and 5 have over 15 years of experience.

For example, Expert 1 works at the Muş Provincial Disaster and Emergency Directorate. According to the evaluation of the main criteria by Expert 1, the best criterion is safety and security (C4), while the worst criterion is electrical infrastructure (C2). The evaluations made are shown in Table 3 below.

Table 3. Evaluation of best and worst criteria for main criteria using linguistic terms

Criteria	C1	C2	C3	C4	C5
Best Criterion (C4)	MMI	VSMI	VSMI	EI	SMI
Worst Criterion (C2)	VSMI	1 (Equal)	MMI	VSMI	SMI

This table reflects the evaluation made by Expert 1, where "C4" (Safety and Security) is rated as the best criterion, and "C2" (Electrical Infrastructure) is rated as the worst.

$$\begin{aligned}
&\min \ \xi \\
&\textit{subject to} \\
&\left|\frac{w_4}{w_1} - \tilde{\tilde{3}}\right| \le \xi, \left|\frac{w_4}{w_2} - \tilde{\tilde{7}}\right| \le \xi, \\
&\left|\frac{w_4}{w_3} - \tilde{\tilde{7}}\right| \le \xi, \left|\frac{w_4}{w_5} - \tilde{\tilde{5}}\right| \le \xi, \\
&\left|\frac{w_2}{w_1} - \tilde{\tilde{7}}\right| \le \xi, \left|\frac{w_2}{w_3} - \tilde{\tilde{3}}\right| \le \xi, \\
&\left|\frac{w_2}{w_5} - \tilde{\tilde{5}}\right| \le \xi, \\
&w_1 + w_2 + w_3 + w_4 + w_5 = 1 \\
&w_1, w_2, w_3, w_4, w_5 \ge 0
\end{aligned}$$

When the mathematical model obtained from the evaluations made by Expert 1 was solved, the importance of the criteria were as follows:

- Land Preference (C1) = 0.217
- Electrical Infrastructure (C2) = 0.052
- Sanitation System (C3) = 0.093
- Safety and Security (C4) = 0.507
- Proximity (C5) = 0.130

Similarly, based on the evaluation of the sub-criteria for Land Preference (C1), the best sub-criterion is Population Density (Area per Person) (C13), while the worst criterion is Property Status (C12). The evaluations are shown in Table 4 below.

Table 4. Evaluation of best and worst criteria for sub-criteria using linguistic terms

Criteria	C11	C12	C13	C14
Population Density (Area per Person) (C13)	MMI	VSMI	EI	SMI
Property Status (C12)	VSMI	EI	VSMI	SMI

Table 4 reflects the evaluation by Expert 1, where C13 (Population Density) is the best sub-criterion, and C12 (Property Status) is the worst sub-criterion.

$$
\begin{aligned}
&\min \ \xi \\
&ş.k.a \\
&\left|\frac{w_{13}}{w_{11}} - \tilde{\tilde{3}}\right| \le \xi, \left|\frac{w_{13}}{w_{12}} - \tilde{\tilde{7}}\right| \le \xi, \\
&\left|\frac{w_{13}}{w_{14}} - \tilde{\tilde{5}}\right| \le \xi, \left|\frac{w_{12}}{w_{11}} - \tilde{\tilde{7}}\right| \le \xi, \\
&\left|\frac{w_{12}}{w_{14}} - \tilde{\tilde{5}}\right| \le \xi, \\
&w_{11} + w_{12} + w_{13} + w_{14} = 1 \\
&w_{11}, w_{12}, w_{13}, w_{14} \ge 0
\end{aligned}
$$

Evaluation of the best and worst criteria for the electric infrastructure (C2) sub-criteria is as follows: Telecommunication Facility (C23) is determined as the best criterion, while Lighting Electricity (C22) is identified as the worst criterion. The evaluations are shown in Table 5.

Table 5. Evaluation of the best and worst criteria for the main criteria with linguistic terms

Criteria	C21	C22	C23
Telecommunication Facility (C23)	MMI	SMI	EI
Lighting Electricity (C22)	MMI	EI	SMI

$$
\begin{aligned}
&\min \ \xi \\
&ş.k.a \\
&\left|\frac{w_{23}}{w_{21}} - \tilde{\tilde{3}}\right| \le \xi, \left|\frac{w_{23}}{w_{22}} - \tilde{\tilde{5}}\right| \le \xi, \\
&\left|\frac{w_{22}}{w_{21}} - \tilde{\tilde{3}}\right| \le \xi, \\
&w_{21} + w_{22} + w_{23} = 1 \\
&w_{21}, w_{22}, w_{23} \ge 0
\end{aligned}
$$

When the mathematical model obtained from the evaluations made by Expert 1 is solved, the importance of the criteria is as follows:

- Usable electricity (C11) = 0.245
- Lighting electricity (C12) = 0.111
- Telecommunication facility (C13) = 0.644

For the sub-criteria of the Sanitary Plumbing System (STS) (C3), the weight of the Drinking water (C31) criterion is 0.756, while the weight of the Toilet condition (C32) criterion is 0.244. For the Safety and Security (GE) sub-criteria, the weight of the Landslides, floods, etc. (C41) criterion is 0.836, showing a very prominent importance, while the Warning systems (Sound systems) (C42) criterion has a weight of 0.164, indicating a less significant importance. Finally, in the evaluation of the sub-criteria under the Proximity (Yk) (C5) main criterion, the Distance from the settlement (Accessibility) (C51) sub-criterion is 0.756, and the Distance from potential disaster debris areas (C52) sub-criterion is 0.244. Each evaluation made by the expert was calculated using the best-worst method. The results for each main and sub-criterion are shown in Table 6.

Table 6. The importance of criteria for experts

Criteria	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Expert 8	Expert 9	Expert 10	Expert 11	Expert 12	Expert 13	Expert 14	Expert 15	Expert 16	Expert 17	Expert 18	Expert 19	Expert 20
(C1)	0.217	0.489	0.224	0.36	0.507	0.497	0.206	0.51	0.497	0.132	0.124	0.085	0.19	0.37	0.53	0.504	0.253	0.139	0.124	0.184
(C2)	0.052	0.116	0.053	0.062	0.13	0.071	0.054	0.053	0.091	0.054	0.088	0.198	0.056	0.109	0.101	0.119	0.05	0.1	0.061	0.054
(C3)	0.093	0.193	0.09	0.082	0.093	0.091	0.088	0.092	0.071	0.073	0.061	0.04	0.082	0.064	0.051	0.119	0.101	0.051	0.088	0.111
(C4)	0.507	0.145	0.409	0.36	0.052	0.128	0.498	0.129	0.128	0.22	0.521	0.198	0.482	0.182	0.177	0.059	0.427	0.477	0.521	0.184
(C5)	0.13	0.057	0.224	0.136	0.217	0.213	0.154	0.216	0.213	0.52	0.206	0.478	0.19	0.274	0.141	0.199	0.169	0.232	0.206	0.466
(C11)	0.24	0.226	0.547	0.095	0.569	0.528	0.475	0.493	0.572	0.136	0.126	0.437	0.144	0.185	0.528	0.518	0.572	0.547	0.256	0.226
(C12)	0.057	0.066	0.078	0.429	0.069	0.075	0.055	0.29	0.066	0.066	0.07	0.197	0.057	0.074	0.226	0.196	0.066	0.078	0.083	0.066
(C13)	0.559	0.572	0.234	0.286	0.121	0.226	0.188	0.101	0.136	0.226	0.315	0.296	0.559	0.278	0.075	0.196	0.226	0.234	0.154	0.572
(C14)	0.144	0.136	0.141	0.19	0.241	0.17	0.282	0.116	0.226	0.572	0.49	0.07	0.24	0.463	0.17	0.089	0.136	0.141	0.508	0.136
(C21)	0.245	0.257	0.571	0.429	0.429	0.644	0.292	0.536	0.257	0.244	0.542	0.244	0.244	0.244	0.244	0.6	0.244	0.257	0.321	0.244
(C22)	0.111	0.6	0.143	0.143	0.143	0.111	0.542	0.143	0.143	0.644	0.167	0.111	0.111	0.111	0.111	0.257	0.111	0.6	0.143	0.111
(C23)	0.644	0.143	0.286	0.429	0.429	0.244	0.167	0.321	0.6	0.111	0.292	0.644	0.644	0.644	0.644	0.143	0.644	0.143	0.536	0.644
(C31)	0.756	0.75	0.333	0.75	0.667	0.75	0.667	0.667	0.25	0.75	0.25	0.75	0.75	0.75	0.75	0.25	0.75	0.25	0.75	0.75
(C32)	0.244	0.25	0.667	0.25	0.333	0.25	0.333	0.333	0.75	0.25	0.75	0.25	0.25	0.25	0.25	0.75	0.25	0.75	0.25	0.25
(C41)	0.836	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.833	0.75	0.25	0.833	0.75	0.75	0.75	0.75	0.833	0.75	0.833
(C42)	0.164	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.167	0.25	0.75	0.167	0.25	0.25	0.25	0.25	0.167	0.25	0.167
(C51)	0.756	0.25	0.667	0.75	0.333	0.75	0.333	0.75	0.25	0.75	0.25	0.75	0.75	0.75	0.167	0.75	0.75	0.75	0.667	0.75
(C52)	0.244	0.75	0.333	0.25	0.667	0.25	0.667	0.25	0.75	0.25	0.75	0.25	0.25	0.25	0.833	0.25	0.25	0.25	0.333	0.25

The cumulative local and global weights of the evaluations made by 20 experts were calculated using the simple average approach, and the results are shown in Table 7. When the main criteria are examined, it is observed that the most important main criterion is Land Preference (ATE) (C1), with an importance weight of 0.307. The second most important main criterion is Safety and Security (GE) (C4), with a weight of 0.290. The third most important main criterion is also Safety and Security (GE) (C4), with the same weight of 0.290. Considering the five main criteria, the relatively less important main criteria are Sanitary Plumbing System (STS) (C3) and Electric Infrastructure (EA) (C2), with weights of 0.087 and 0.084, respectively.

Table 7. The local and global weights

Criteria	Local weights	Global weights
(C1)	0.307	
(C2)	0.084	
(C3)	0.087	
(C4)	0.290	
(C5)	0.232	
(C11)	0.371	0.114
(C12)	0.118	0.036
(C13)	0.278	0.085
(C14)	0.233	0.072
(C21)	0.349	0.029
(C22)	0.228	0.019
(C23)	0.418	0.035
(C31)	0.617	0.054
(C32)	0.383	0.033
(C41)	0.746	0.216
(C42)	0.254	0.074
(C51)	0.596	0.138
(C52)	0.404	0.094

When the cumulative global weight is examined, the most important sub-criterion is Landslides, floods, etc. (C41). The importance weight for this sub-criterion is determined to be 0.216. The other two most important sub-criteria are Distance from the settlement (Accessibility) (C51), with an importance weight of 0.138, and Topography and slope (C11), with an importance weight of 0.114. Landslides, Floods, etc. (C41) hold significant global importance because natural disasters (such as landslides, floods, and storms) can lead to substantial losses worldwide and are a critical factor when considering environmental risks. These types of disasters are not only significant at the local level but also draw attention due to their impact on the global economy and environment. Global disaster management focuses on preparedness and risk mitigation strategies for such events. On a global scale, Distance from the settlement (Accessibility) (C51) is an important criterion, particularly during disasters. Accessibility is a fundamental factor for quickly delivering aid and providing logistical support to disaster-stricken areas. Additionally, the operational efficiency of international aid organizations and disaster response teams on a global scale depends on the accessibility of settlements. Therefore, access and reach are of critical importance on a global level. Topography is particularly important under the influence of natural disasters. The slope of mountainous regions, for example, can affect the frequency of landslides and floods, making it a key factor in shaping global disaster strategies. The management of such terrains is critical for construction and settlement planning, thus holding significant importance at the global level.

According to the cumulative global weights, the three least important criteria are Lighting electricity (C22), Ownership status (C12), and Usable electricity (C21). Lighting electricity (0.019) holds very low importance on a global scale. Most developed countries and regions have largely completed their lighting infrastructure, so it is less emphasized as a global priority. However, this issue may be more significant in developing regions, but on a global scale, the focus tends to be more on basic infrastructure elements (e.g., drinking water, energy infrastructure). Ownership status (0.036) may be important at the local level, but it is a less prioritized factor on a global scale. Globally, ownership status is not directly related to managing environmental risks, especially in the context of disaster management and environmental planning. Instead, broader-scale issues such as the conservation of natural resources and the environment take precedence. The global importance of electric infrastructure is not as pronounced as it is at the local level. On a global scale, energy access is more closely related to energy policies and sustainable energy production. Electric infrastructure (0.029) is generally sufficient in developed countries and is considered an element that needs to be restored after disaster situations or energy crises. On a global level, elements such as water supply or disaster preparedness are prioritized over electricity.

4. Conclusion

Disasters are naturally occurring events that cause loss of life and property. For an event to be classified as a disaster, it must be large enough to significantly impact people or the environment where people live. From this perspective, a disaster is not just an event but the consequence of an event. Disaster management, due to the nature of disasters, follows a cyclical structure consisting of four consecutive phases. As a result of this structure, the risk and damage reduction phase does not begin and end in a specific place within the cycle. On the contrary, after a disaster, risk and damage reduction efforts, which start with the recovery phase, continue until the next disaster. The anticipation of a disaster is the phase where efforts are made to reduce the potential impact of the disaster threat. Efforts are made through structural and non-structural measures to prevent the negative impacts of environmental and technological hazards caused by natural dangers. During the damage reduction phase, the negative impacts of hazards and risks cannot be entirely prevented, but the area or severity of these threats can be significantly reduced through various strategies and actions. To successfully manage disaster risk, it is essential to strengthen the damage reduction phase and ensure investments in risk and hazard-reducing measures to build resilience against disasters.

The identification of the locations of emergency gathering points used after a disaster is one of the most important aspects of urban planning and disaster management. The distribution, size, visibility, accessibility, and infrastructure requirements of these areas must be ensured to make them suitable for their purpose.

In this study, 16 different emergency gathering points have been considered for prioritization in Muş province. To prioritize these gathering points, the criteria for emergency gathering locations were first examined in the literature. In this stage, five main criteria and 13 sub-criteria were determined within the context of emergency gathering points' location selection. The main criteria are land preference, electric infrastructure, sanitary plumbing system, safety and security, and proximity.

This paper presents a model for prioritizing emergency gathering points in Muş province based on the established criteria, which is applicable to other provinces in Turkey. The model proposed for other provinces in our country can be implemented. In addition to the five main criteria and 13 sub-criteria, other criteria can also be considered. The Bayesian Best-Worst Method can be used as an alternative approach for determining the criteria's weights. Furthermore, the interval type-2 fuzzy set-based approach can be developed with heuristic fuzzy sets, value-range fuzzy sets, and other different fuzzy sets.

References

AFAD (2018). URL-7, 2020. Afet ve Acil Durum Başkanlığı Resmî Web Sitesi, https://www.afad.gov.tr/ 06.12.2020

Altay, B. C., Celik, E., Okumus, A., Balin, A., & Gul, M. (2023). An integrated interval type-2 fuzzy BWM-MARCOS model for location selection of e-scooter sharing stations: The case of a university campus. Engineering Applications of Artificial Intelligence, 122, 106095.

Celik, E. (2024). Analyzing the Shelter Site Selection Criteria for Disaster Preparedness Using Best–Worst Method under Interval Type-2 Fuzzy Sets. Sustainability, 16(5), 2127.

Çelik, E., 2017. A cause and effect relationship model for location of temporary shelters in disaster operations management. International Journal of Disaster Risk Reduction, 22:257-268.

Celik, E., Akyuz, E., 2018. An interval type-2 fuzzy AHP and TOPSIS methods for decision-making problems in maritime transportation engineering: the case of ship loader. Ocean Engineering, 155, 371-381.

Celik, E., Gul, M., Gumus, A.T., Guneri, A.F., 2012. A fuzzy TOPSIS approach based on trapezoidal numbers to material selection problem, Journal of Information Technology Applications & Management, 19(3):19-30

Ciga, B., 2020. Peyzaj planlama açısından acil durum toplanma alanlarının yeterliliklerinin Istanbul ili kartal ilçesi örneğinde incelenmesi. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, İstanbul. 164s.

Çınar, A.K., Akgün, Y., Maral, H., 2018. Afet sonrası acil toplanma ve geçici barınma alanlarının planlamasındaki faktörlerin incelenmesi: İzmir-Karşıyaka örneği. Planlama Dergisi, 28(2):179-200.

Dursun, B., 2021. Afet toplanma alanlarının türlerine ve yer seçim kriterlerine göre uygunluğunun değerlendirilmesi: İstanbul, Esenler örneği. Yüksek Lisans Tezi, Aydın Üniversitesi, İstanbul, 92s.

Guo, S., & Zhao, H. (2017). Fuzzy best-worst multi-criteria decision-making method and its applications. Knowledge-based systems, 121, 23-31.

Hafezalkotob, A., & Hafezalkotob, A. (2017). A novel approach for combination of individual and group decisions based on fuzzy best-worst method. Applied Soft Computing, 59, 316-325.

JICA and IBB (2002). The study on a disaster prevention / mitigation basic plan ın Istanbul ıncluding microzonation ın the republic of Turkey. Final Report, Volume V. (In Turkish)

Ju, Y., Wang, A., Liu, X., 2012. Evaluating emergency response capacity by fuzzy AHP and 2-tuple fuzzy linguistic approach, Expert Systems with Applications, 39(8), 6972-6981.

Kılcı, F., Kara, B.Y., Bozkaya, B., 2014. Locating temporary shelter areas after an earthquake: A case for Turkey, European Journal of Operational Research, 243(1), 323-332.

Mi, X., Tang, M., Liao, H., Shen, W., & Lev, B. (2019). The state-of-the-art survey on integrations and applications of the best worst method in decision making: Why, what, what for and what's next?. Omega, 87, 205-225.

Moslem, S., Gul, M., Farooq, D., Celik, E., Ghorbanzadeh, O., & Blaschke, T. (2020). An integrated approach of best-worst method (BWM) and triangular fuzzy sets for evaluating driver behavior factors related to road safety. Mathematics, 8(3), 414.

Nappi, M.M.L., Souza, J.C., 2014. Disaster management: hierarchical structuring criteria for selection and location of temporary shelters, Natural Hazards, 75(3): 2421-2436.

Özel, S., 2019. Afet sonrası toplanma alanlarının kentsel açık ve yeşil alan sistemlerindeki yeri – Kastamonu kenti örneği. Yüksek Lisans Tezi, Kastamonu Üniversitesi, Kastamonu, 173s.

Palazca, A., 2020. Afet sonrası toplanma alanlarının analizi: Denizli örneği. Yüksek Lisans Tezi, Pamukkale Üniversitesi, Denizli. 218s.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. Omega, 53, 49-57.

Tarabanis, K., & Tsionis, I. (1999). Using network analysis for emergency planning in case of an earthquake. Transactions in GIS, 3(2), 187-197.

Taylan, S., 2018. Afet sonrası acil toplanma ve geçici barınma alanları standartlarının değerlendirilmesi (Çankırı ili örneği). Yüksek Lisans Tezi, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, 134s.

Trivedi, A., 2018 A multi-criteria decision approach based on DEMATEL to assess determinants of shelter site selection in disaster response, International Journal of Disaster Risk Reduction, 31: 722-728.

Trivedi, A., Singh, A., 2016. Prioritizing emergency shelter areas using hybrid multi-criteria decision

Yu, J., Wen, J., 2016. Multi-criteria satisfaction assessment of the spatial distribution of urban emergency shelters based on high-precision population estimation. International Journal of Disaster Risk Science, 7(4):413-429.

Journal of Information Analytics, 1(1), 38-58

Journal of Information Analytics

Submit 22.04.2025
Revision 29.04.2025
Accept 14.05.2025
Publish 15.05.2025

Open Access

Research Article

A DEMATEL-based causal model for understanding the key determinants of physician migration

Umran Tepe[1], Gokhan Agac[2], Ece Colkesen Tefiroglu[3]

[1]*Patient Admissions Department, Acıbadem Healthcare Group, İstanbul, Türkiye. E-mail:* umran.tepe20000@gmail.com

[2]*Department of Healthcare Management, Faculty of Health Sciences, Sakarya University of Applied Sciences, Türkiye. Visiting Researcher, Institute of Health Administration, Robinson College of Business, Georgia State University, USA. ORCID: 0000-0002-4753-4689 E-mail:gokhanagac@subu.edu.tr*

[3]*Department of Healthcare Management, Institute of Health Sciences, Ankara University, Türkiye and Department of Healthcare Management, Faculty of Health Sciences Sakarya University of Applied Sciences, Türkiye. ORCID: 0000-0002-9818-6362 E-mail: ececolkesen@subu.edu.tr*

Abstract

This study aims to determine the reasons for the migration of physicians and to reveal their impact levels. Migration has critical effects on both receiving and sending countries, leading to a decrease in the quality of health care, loss of qualified employment, decreased productivity and health inequality. The push factors leading to physician migration are low salaries, poor working conditions and political insecurity, while the pull factors are higher living standards, career opportunities and personal development opportunities. In the study, critical success factors were determined by using the DEMATEL method. Based on the opinions of specialized physicians, 6 main criteria and 50 sub-criteria were determined. The main criteria are economic, political-social, educational and career-related, social-cultural, health-related, labor and working conditions-related criteria. The criteria and the relationships between them are visualized with diagrams. According to the analysis, the most influential main criterion on physician migration was found to be work and working conditions, while the least influential criterion was found to be social and cultural factors. According to the degree of influence, economic reasons and education-career opportunities are the most influential criteria. In the sub-criteria dimension, “low salaries”, “violence”, “lack of career development” and “political insecurity” are the most prominent criteria. The findings reveal that the migration decision is shaped in a multidimensional structure and the impact of different criteria. In this study, physician migration is analyzed with the MCDM. With the results of the research, strategic recommendations are developed for health policy makers. It is emphasized that to ensure the return of physicians, the problem should be viewed from a health inequality perspective. To prevent migration, it is emphasized that working conditions should be improved, career and educational opportunities should be increased, economic incentives should be provided, and social security should be ensured.

Keywords Brain drain, DEMATEL, multi-criteria decision making, international migration, physician migration, reasons for migration

Citation: Tepe, U., Agac, G., & Colkesen Tefiroglu, E. (2025). A DEMATEL-based causal model for understanding the key determinants of physician migration. *Journal of Information Analytics*, 1(1), 38-58.

Corresponding Author: Gokhan Agac gokhanagac@subu.edu.tr

1. Introduction

Physicians leaving their home countries to practice medicine in other countries is a global health problem. While international migration of physicians causes significant losses for the sending country, it is considered as a gain of qualified labor for the receiving countries. In this context, physician migration affects the capacity in health care delivery and causes differences between countries (Buchan, 2008; Saluja et al., 2020). Medical education, clinical internships, specialty training and continuing professional development programs have been provided for many years to train physicians. Although there is no return in terms of public resources, investments and expenses, it creates additional costs for new physicians to be trained. At the end of the process, the migration of physicians to other countries negatively affects the health system from the perspective of the physician training countries. Therefore, it is important to understand the reasons for migration for the sending country to reduce the loss of physicians (Davda et al., 2021; Morley et al., 2017; Murataj et al., 2022).

Physician migration leads to disruptions in health services in origin countries. With the migration of physicians, the gap in specialized areas reduces the quality of health care. The lack of physicians creates an excessive workload on other health professionals in the destination country and significantly affects the health system (Byrne et al., 2021). In addition, with the decrease in the number of physicians, patients' access to health care is restricted and inequality in health emerges. In the long run, it can have an impact on the planning of health policies and cause instability in the health system, leading to a system that is dependent on external resources. These situations also undermine social trust and can become a political pressure factor (Dohlman et al., 2019; Humphries et al., 2019).

Factors related to why physicians migrate from their home countries are divided into two categories: attractive and repulsive factors. Repulsive factors, as an important reason for brain drain in undeveloped and developing countries, include health professionals' discomfort with unfavorable conditions in their home countries. Attractive factors, on the other hand, are defined as advantages such as better working conditions, higher living standards and higher salaries in the countries they intend to migrate to (Adovor et al., 2021). Economic factors, working conditions, lack of professional satisfaction, political and social insecurity stand out among the reasons why physicians migrate (Ebeye & Lee, 2023; Obinna et al., 2022; Vujicic et al., 2004). In addition to excessive workload, the increasing number of physical and emotional violence against doctors has a significant impact on the decision to migrate (Kadaifci et al., 2024). Depending on the family, the decision to migrate can change in terms of economic, career, education, etc., and reasons such as personal development opportunities, including self-realization of physicians, affect the decision (Becker & Teney, 2020). In addition, after those who leave with physician migration, there is a feeling of motivational burnout for the physicians who stay (Sweileh, 2024).

The number of physicians per capita in the world varies among countries. The average number of physicians per capita in OECD countries is 3.7% (OECD, 2023). According to the World Health Statistics report, the number of physicians per ten thousand people between 2014 and 2021 is 36.6 in the European region, 24.5 in the USA, 20.9 in the Western Pacific, 11.2 in the Eastern Mediterranean, 7.7 in South East Asia and 2.9 in Africa (WHO, 2023). This inequality creates great differences in access to health services. Looking at the literature, physician migration is common in countries such as India, Nigeria, Lebanon, Romania, and South Africa (Botezat & Moraru, 2020; Onah et al., 2022; Tankwanchi et al., 2021).

According to OECD (2021) statistics, 19% of physicians in OECD countries received their first medical education in other countries. According to the Turkish Medical Association (2021) study report, the number of physicians applying to migrate in Türkiye increased 15-fold from 2012 to 2020 (Karatuzla, 2024).

Most studies on physician migration examine the effects of economic, social and professional factors on migration (Apostu et al., 2022; Dubas-Jakóbczyk et al., 2020; Hadley, 2024; Okeke et al., 2014). Goštautaitė et al. (2023) examined the impact of physician migration on health systems and made different recommendations on how to prevent it. Apostu et al. (2022) studied the factors affecting physicians in Romania. Teney (2019) analyzed the reasons for the migration of highly qualified professionals from European Union countries to Germany. Domagala et al. (2022) examined the reasons for migration of health professionals in Poland and found that the push factors for migration are inadequate salaries, lack of favorable working conditions and lack of support for personal development and education. Attractive factors were identified as rising standard of living, lack of salary inequality and easy access to emerging technologies. These factors lead to a decrease in the number of qualified personnel in the health system and a decrease in the quality of patient care.

The aim of this study is to examine the reasons affecting physician migration and to determine the effects of these reasons. The movement of physician migration was analyzed using the DEMATEL (Decision-making Trial and Evaluation Laboratory) method and critical success factors were identified. The following research questions were formulated in line with this purpose:

RQ1. What are the opinions of physicians towards the phenomenon of international migration?

RQ2. What are the most dominant reasons that push physicians towards international migration?

RQ3. What are the levels of importance and relationship between the reasons that push physicians to the phenomenon of migration?

In line with the purpose of the study, when the relevant literature is examined, Multi-Criteria Decision Making (MCDM) techniques, which are frequently preferred and examine criteria using both qualitative and quantitative data at the same time, come to the fore. MCDM methods are widely used to evaluate different criteria and factors in complex decision-making processes (Gokler & Boran, 2024; Komasi et al., 2023). DEMATEL, one of the MCDM techniques, is an effective method that analyzes the relationships between system factors and visualizes this structure through cause-effect relationship maps. It has been widely used in various fields such as health, economy, tourism, management, engineering (Agarwal & Kapoor, 2022; Aka & Yavuz, 2024; Braga, 2021; Che et al., 2022; Gedam et al., 2021; Gokler & Boran, 2024; Parmar & Desai, 2020). The method not only transforms the interdependencies of factors into cause-and-effect relationships, but also identifies the critical components of a system with the help of influence relationship diagrams. In this direction, the DEMATEL technique was preferred in the study to determine the critical factors related to physician migration.

The contribution of this study to literature can be listed as follows:

- Providing a framework that addresses the factors that cause physician migration in a broad and detailed manner.
- Identifying which factors are the dominant factors among the factors that cause physician migration.

- To reveal the importance and relationship levels of the factors that cause physician migration.
- Provide broad recommendations to health decision makers or policy makers in light of the findings
- Contributing to the lack of such an approach in the literature by addressing the issue using the DEMATEL technique, one of the MCDM techniques.

The rest of the paper is organized as follows: In the next section, the authors present the methodology of this study; the third section discusses the implementation of the study and the findings. The fourth section presents a broad discussion in line with the findings of the study. Finally, the study concludes with a conclusion.

2. Research Methodology

Multi-criteria decision-making tools are frequently preferred for solving problems with multiple criteria. Among these tools, DEMATEL, which is used in the simplified analysis of complex problems, attracts attention (Parmar & Desai, 2020). It was developed by Gabus and Fontela in 1972 and is used to identify and visualize the relationship between criteria in a system (Agarwal & Kapoor, 2022; Farooque et al., 2020). Thanks to this method, an improvement in a criterion that is important and affects other criteria can lead to a similar improvement in other criteria (Celikbilek & Ozdemir, 2020).

DEMATEL analyzes the cause and effect relationship between the criteria and produces a diagram of the relationship between the criteria. In the diagram, the magnitude of the relationship between the criteria is expressed in numerical values. In this way, it enables decision makers to better interpret the criteria (Braga, 2021). The stages of the DEMATEL method are as follows (Gedam et al., 2021; Shieh et al., 2010):

Step 1: Creating a Direct Relationship Matrix

The first step of the DEMATEL method is the creation of a direct relationship matrix. The values in this matrix indicate the direct relationship of variable (i) with variable (j). The direct relationship matrix is shown in equation (1):

$$\mathrm{D} = [\mathrm{d_{ij}}]_{\mathrm{nxn}} \qquad (1)$$

The variables studied are mutually dependent. Each is assigned a score to indicate the degree of influence. Experts made their ratings according to the 0-4 scale. In the scale used; '0' means zero effect, '1' means moderately low effect, '2' means moderately high effect, '3' means high effect and '4' means very high effect.

Step 2: Creating the Normalization Matrix

Each row and column value in the direct relationship matrix is summed for normalization. All values in the matrix are divided by the largest value obtained from the sum. Eqs. (2) and (3) are used for the normalization process.

$$\mathrm{X} = \mathrm{s.D.} \qquad (2)$$

$$\mathrm{s} = \min\left[\frac{1}{\left(\max_i \sum_{j=1}^{n} |\mathrm{d_{ij}}|\right)}, \frac{1}{\left(\max_{\mathrm{j}} \sum_{j=1}^{n} |\mathrm{d_{ij}}|\right)}\right] \qquad (3)$$

Step 3: Creating the Total Impact Matrix

The normalized direct relationship matrix is transformed into the total influence matrix shown in Eq. (5) using Eq. (4). The normalized direct relationship matrix is subtracted from the unit matrix and inverted. The resulting matrix is then multiplied by itself to find the total influence matrix.

$$\mathrm{T} = \mathrm{X} + \mathrm{X}^2 + \mathrm{X}^3 + \cdots + \mathrm{X}^h = \mathrm{X}(\mathrm{I} - \mathrm{X})^{-1} \quad (4)$$

$$\mathrm{T} = \left[t_{ij}\right]_{\mathrm{nxn}} \quad (5)$$

Step 4: Identification of Affecting and Affected Variables

After calculating the sum of rows (D_i) and columns (R_j) of the total influence matrix with Eq. (6) and Eq. (7), $D_i + R_j$ and $D_i - R_j$ values are found.

$$\mathrm{D} = (r_i)_{\mathrm{nx1}} = \left[\sum\nolimits_{j=1}^{n} t_{ij}\right]_{\mathrm{nx1}} \quad (6)$$

$$\mathrm{R} = (c_j)_{\mathrm{1xn}} = \left[\sum\nolimits_{i=1}^{n} t_{ij}\right]_{\mathrm{1xn}} \quad (7)$$

Step 5: Calculation of Criteria Weights and Drawing the Influence Diagram

Weight values for the criterion are calculated as the average of the squares of D_i+R_j and D_i-R_j. With this step, the criteria with priority importance are determined.

$$w_i = \sqrt{(D_i + R_j)^2 + (D_i - R_j)^2} \quad (8)$$

$$W_i = \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (9)$$

The influence diagram is drawn with $D_i + R_j$ values on the horizontal axis and $D_i - R_j$ values on the vertical axis. This diagram shows the relationship between the criteria and is therefore important.

3. Application of the Study

In this section, the main and sub-criteria affecting physician migration, data collection and the findings obtained by applying the DEMATEL approach are presented. The application of the study is based on Türkiye. In addition, in this section, the reasons for migration of physicians are evaluated under the main dimensions.

Determination of Criteria

In this study, a literature review was conducted to identify the critical main and sub-factors affecting the brain drain of physicians. After the research, six main factors and their sub-criteria were identified: economic factors, political and social factors, factors related to education and career, social and cultural factors, factors related to health, factors related to work and environmental conditions (Supplementary file 1). The list of these factors is presented in Table 1.

Table 1. Sub-criteria affecting the migration of physicians

Economic Factors (C1)	Political and Social Factors (C2)	Education and Career Related Factors (C3)	Social and Cultural Factors (C4)	Health-Related Factors (C5)	Work and Working Conditions Related Factors (C6)
High Unemployment Rate **(C11)**	Political Rights and Civil Freedoms **(C21)**	Inadequate Research and Development Studies (**C31**)	Unfavourable Conditions for Family Building **(C41)**	Resource Deficiencies in the Health System **(C51)**	Feelings of Inadequacy due to Job Dissatisfaction **(C61)**
Wage Differentials **(C12)**	Bad Political Climate **(C22)**	Existence, Adequacy, Superiority of the School **(C32)**	Cultural Mentality **(C42)**	Governance and Management Deficiencies in Health Services **(C52)**	Unsatisfactory Working Conditions **(C62)**
Low Salaries **(C13)**	Management Dissatisfaction **(C23)**	Opportunity to develop knowledge and skills **(C33)**	Personal Safety and Security Needs **C43)**	Access to Health Services **(C53)**	Violence Against Physicians **(C63)**
High Costs of Living **(C14)**	Racism **(C24)**	Education Curriculums **(C34)**	Social Intolerance **C(44)**	Fear of Infectious Disease (HIV, COVID-19) **(C54)**	Job Dissatisfaction **(C64)**
Economic Collapse **(C15)**	Ethnic, Religious and Political Tensions **(C25)**	Similarity of Professional Qualifications **(C35)**	Desire to Recognise New Culture **(C45)**	Access to Clean Water Sanitation **(C55)**	High Stress Level at Work **(C65)**
	Human Rights Violations (**C26)**	Similarity and Mutual Recognition of the Language of Education **(C36)**	Sexual Preferences **C(46)**	Stress Levels **(C56)**	Heavy Work Load **(C66)**
	Human Favouritism **(C27)**	Limited Funding for Medical Research **(C37)**	Religious and Political Beliefs **(C47)**		Inadequacy of Technological Infrastructure **(C67)**
	Need for Freedom of Expression **(C28)**	Lack of Opportunities for Career Development **(C38)**	Facilities such as housing, car, pension provisions **(C48)**		
	Crime Rates and Corruption **(C29)**	Desire to Gain Experience **(C39)**	Cultural Affinity **(C49)**		
			Size of Social Networks **(C410)**		
			Housing Problems **(C411)**		
			Proximity (Distance between two countries) **(C412)**		
			Transport **(C413)**		
			Family Existence **(C414)**		

Identification of Experts and Collection of Data

The data of the study were collected from 12 physicians who are experts in their fields and actively working in public hospitals in Türkiye. Face-to-face interviews of 15-30 minutes were conducted with specialists in the fields of Obstetrics and Gynecology, Medical Genetics, General Surgery, Neurology, Cardiology, Psychiatry, Skin and Venereal Diseases and Internal Medicine.

Application of the DEMATEL Approach

In this section, the data obtained from the experts are presented in the findings by applying the steps of the DEMATEL technique.

Step 1: Creating the Direct Relationship Matrix

A Direct Relationship Matrix was created by taking the arithmetic mean of the scores of the 12 physicians participating in the study in the main criteria section of the DEMATEL form (Table 2). After the direct relationship matrix, the highest value was determined by calculating the row and column sums in the matrix (Table 2).

Table 2. Direct relationship matrix

MAIN CRITERIA	C1	C2	C3	C4	C5	C6	TOTAL
C1	0	2.583	2.916	2.083	2.666	3.5	13.75
C2	1.75	0	2.583	2.333	2.833	3.333	12.833
C3	2.75	2.166	0	2.916	3.25	3.333	14.416
C4	2.166	2.833	2.166	0	3	2.666	12.833
C5	2.166	3	2.833	1.916	0	2.166	12.083
C6	3.5	2.916	2.75	2.583	2.416	0	14.166
TOTAL	12.333	13.5	13.25	11.833	14.166	**15**	

Step 2: Creating the Normalization Matrix

After the Direct Relationship Matrix, a normalized Direct Relationship Matrix was created (Table 3). In Table 2, the numbers in the direct relationship matrix are divided by the value "15", which is the maximum of the sum of the numbers in the rows and columns, to obtain the value "z=0.0666". The "z" value obtained was multiplied by the direct relationship matrix values to form a normalized direct relationship matrix.

Table 3. Normalization matrix

MAIN CRITERIA	C1	C2	C3	C4	C5	C6
C1	0	0.172	0.194	0.138	0.127	0.233
C2	0.116	0	0.172	0.155	0.188	0.222
C3	0.183	0.144	0	0.194	0.216	0.222
C4	0.144	0.188	0.144	0	0.2	0.177
C5	0.144	0.2	0.188	0.127	0	0.144
C6	0.233	0.194	0.183	0.172	0.161	0

Step 3: Creating the Total Impact Matrix and Identification of Influencing Variables and Affected Variables

The normalized direct relationship matrix is subtracted from the unit matrix and first inverted. Then, the resulting matrix is multiplied by itself and the total influence matrix is calculated. The unit matrix (I) required for the formation of the total influence matrix is given in Table 4.

Table 4. Unit Matrix of main criteria

MAIN CRITERIA	C1	C2	C3	C4	C5	C6
C1	1	0	0	0	0	0
C2	0	1	0	0	0	0
C3	0	0	1	0	0	0
C4	0	0	0	1	0	0
C5	0	0	0	0	1	0
C6	0	0	0	0	0	1

To construct the total impact matrix, the normalized matrix (N) was subtracted from the unit matrix (Table 5).

Table 5. Extraction of identity matrix of main criteria from normalized relationship matrix (I-N)

MAIN CRITERIA	C1	C2	C3	C4	C5	C6
C1	1	-0.172	-0.194	-0.138	-0.127	-0.233
C2	-0.116	1	-0.172	-0.155	-0.188	-0.222
C3	-0.183	-0.144	1	-0.194	-0.216	-0.222
C4	-0.144	-0.188	-0.144	1	-0.2	-0.177
C5	-0.144	-0.2	-0.188	-0.127	1	-0.144
C6	-0.233	-0.194	-0.183	-0.172	-0.161	1

The inverse of the resulting matrix was then taken as $(I\text{-}N)^{-1}$ (Table 6).

Table 6. $(I\text{-}N)^{-1}$ matrix

MAIN CRITERIA	C1	C2	C3	C4	C5	C6
C1	2.190	1.423	1.421	1.262	1.473	1.580
C2	1.222	2.198	1.327	1.204	1.400	1.485
C3	1.387	1.450	2.303	1.343	1.549	1.620
C4	1.234	1.352	1.301	2.063	1.402	1.447
C5	1.183	1.304	1.280	1.130	2.180	1.366
C6	1.406	1.468	1.442	1.312	1.493	2.424

The matrix was then multiplied by the normalized direct relationship matrix to form the total relationship matrix $T = N(I - N)^{-1}$ (Table 7). The sum of rows and columns of the total relationship matrix was calculated.

Table 7. $T = N(I - N)^{-1}$, the total effect matrix

MAIN CRITERIA	C1	C2	C3	C4	C5	C6	TOTAL
C1	1.190	1.423	1.421	1.262	1.473	1.580	8.351
C2	1.222	1.198	1.327	1.204	1.400	1.485	7.839
C3	1.387	1.450	1.303	1.343	1.549	1.620	8.654
C4	1.234	1.352	1.301	1.063	1.402	1.447	7.801
C5	1.183	1.304	1.280	1.130	1.180	1.366	7.445
C6	1.406	1.468	1.442	1.312	1.493	1.424	8.547
TOTAL	7.626	8.197	8.076	7.316	8.499	8.924	

Step 4: Calculating Criteria Weights

The sum of D_i+R_j is used to determine the importance of the criteria, while the D_i-R_j value is used to determine their influence status. If the D_i-R_j value is negative, it is influenced by other criteria, i.e. it is in the position of receiver, and if the D_i-R_j value is positive, it is in the position of cause, i.e. it has an impact on other criteria. As a result, relationships and effects between criteria were determined (Table 8).

Table 8. Weighting of Relationship direction and importance of main factors

D_i+R_j	D_i-R_j	Group	w_i	W_i
15.977	0.725	cause	15.993	0.164
16.036	-0.357	**effect**	16.040	0.164
16.730	0.577	cause	16.740	0.171
15.117	0.484	cause	15.125	0.155
15.945	-1.053	effect	15.980	0.164
17.471	-0.376	effect	17.475	0.179
	Total		97.355	

As a result of the calculations, the total relationship matrix of the main criteria (Table 9) was created.

Table 9. Total relationship matrix of main criteria $T = N(I - N)^{-1}$

MAIN CRITERIA	C1	C2	C3	C4	C5	C6
C1	1.190	1.423	1.421	1.262	1.473	**1.580**
C2	1.222	1.198	1.327	1.204	1.400	**1.485**
C3	1.387	1.450	1.303	1.343	**1.549**	**1.620**
C4	1.234	1.352	1.301	1.063	1.402	1.447
C5	1.183	1.304	1.280	1.130	1.180	1.366
C6	1.406	1.468	1.442	1.312	**1.493**	1.424

In order to determine the interaction, a threshold value (1.48) was calculated by summing the mean and standard deviation. The cells above this value are highlighted in bold in Table 9. According to the research findings, among the main factors affecting the reasons for physicians to migrate, the most important criterion is C6 with D_i+R_j = 17.47 and the least important criterion is C4 with D_i+R_j = 15.11. When the D_i-R_j values are analyzed, the factors that are affected by other criteria are C2, C5 and C6, and the most affected factor is determined to be criterion C6. When the factors affecting the other criteria more than the other criteria are examined, C1 and C3 criteria are identified, and C1 is the factor affecting more than the other with a value of D_i-R_j = 0.7251.

Step 5: Drawing the Impact Diagram

The interaction between the criteria is better understood by drawing the influence diagram. The influence directional graph diagram is constructed with the points D_i+R_j and D_i-R_j positioned D_i+R_j on the horizontal axis and D_i-R_j on the vertical axis. The horizontal axis represents the degree of importance of the criteria and the vertical axis represents the degree of influence of the criteria. The influence diagram is important in terms of showing the variables that affect and are affected by each other. The influence directional graph diagram is shown in Figure 1.

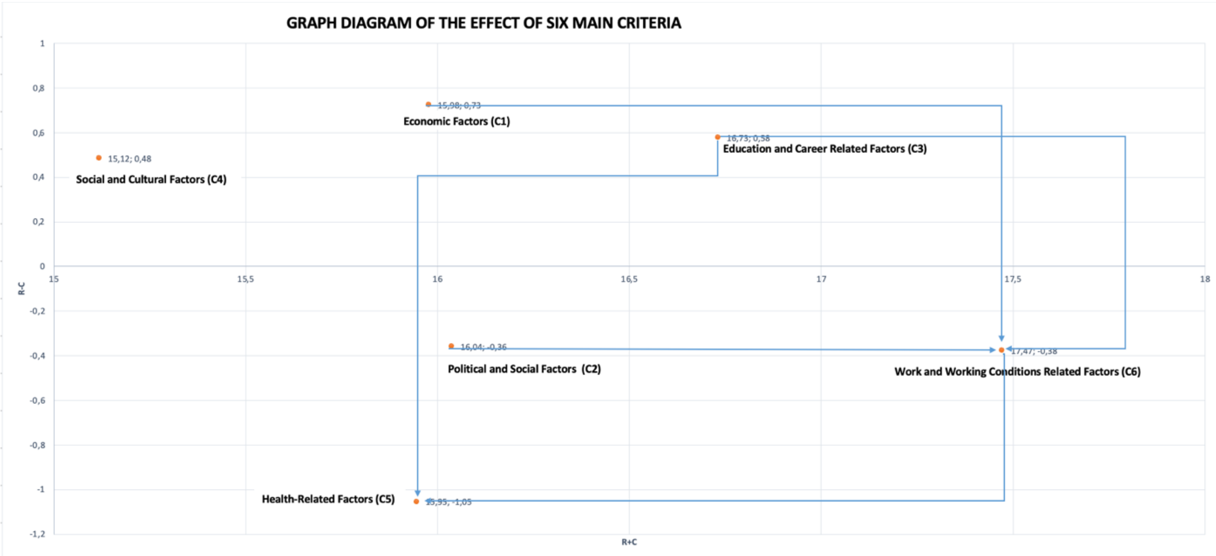

Figure 1. Impact-direction graph diagram of main criteria

Accordingly, criteria C1 and C2 are effective on criterion C6. Criterion C3 is effective on criteria C5 and C6. Criteria C4 and C5 criteria are not effective. Criteria C1, C2, C3, and C4 are not influenced by any criteria. Criterion C5 is influenced by criteria C3 and C6; C6 is influenced by criteria C1, C2 and C3.

Analysis of Economic Factors Dimension with DEMATEL

The sub-criteria for the economic factors dimension are given in Table 9. Analyses related to these criteria are given in the Supplementary File. Impact-Way Graph Diaphragm of Economic Criteria is given in Figure 2.

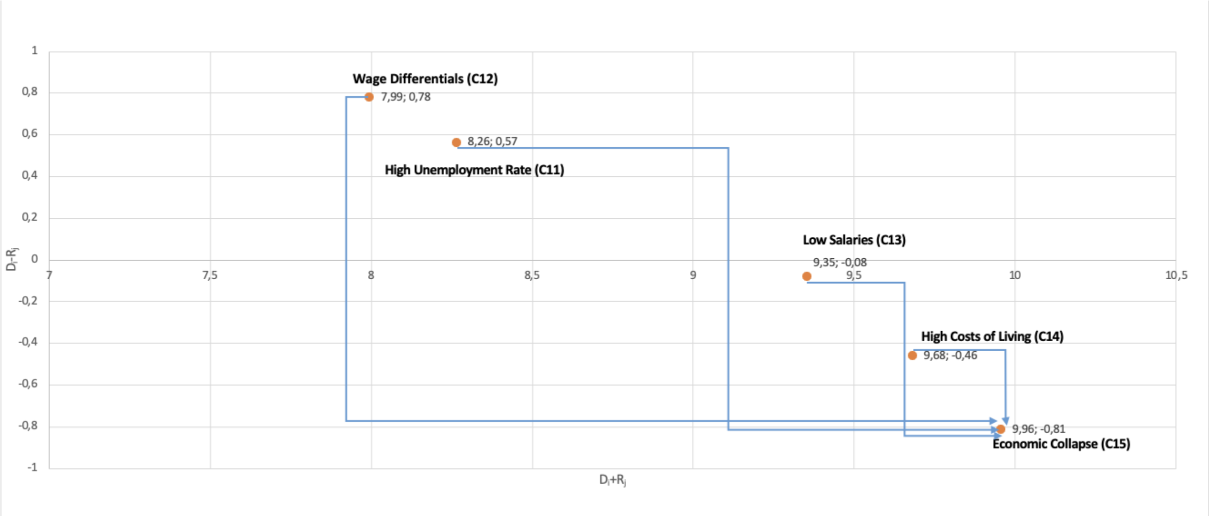

Figure 2. Impact-way graph diaphragm of economic criteria

According to the findings related to C1, it was determined that C15 was the most determinant factor among the factors affecting physicians' reasons for migration, while C12 was the least effective. When the D_i-R_j values are analyzed, it is determined that C13, C14 and C15 are the factors that are influenced by the other criteria, and C15 is the one that is most influenced by the other criteria. The criteria that affect the other factors more are C11 and C12, and C12 is the most determining factor. According to Figure 3; C11, C12, C13 and C14 criteria are found to be effective on C15. C15 is not effective on other criteria.

Analysis of Political and Social Factors Dimension with DEMATEL

The sub-criteria for C2 are given in Table 9. Analyses related to these criteria are given in the Supplementary File. The Influence-Directional Graph Diagram of Criterion C2 is given in Figure 3.

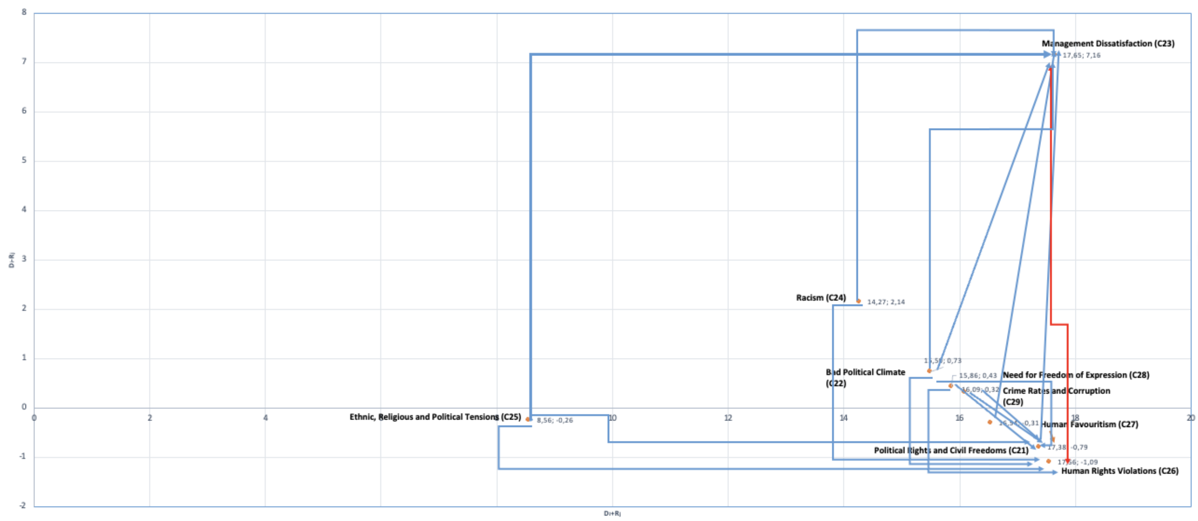

Figure 3. Influence-way graph diaphragm of C2

According to the findings related to C2, among the factors affecting the reasons for physicians to migrate, C23 was found to be the most determinant and C25 was the least important. When the D_i-R_j value is analyzed, it is determined that the factors affected by other criteria are C21, C22, C25, C26 and the most affected factor is C26. The factors that affect other factors more are C23, C24 and the most affected factor is C23.

Figure 3 shows that C21 affects C23 and C26 criteria; C22 affects C23 and C26 criteria; C24 affects C26 criteria; C25 affects C21 and C26 criteria; C27 affects C23 and C26 criteria; C24 affects C23 criteria; C28 affects C21 criteria; C29 affects C23 and C26 criteria. C23 and C26 are criteria that are affected by each other.

Analysis of Education and Career Related Factors Dimension with DEMATEL

The sub-criteria for C3 are given in Table 9. Analyses related to these criteria are given in the Supplementary File. The Influence-Directional Graph Diagram of Criterion C3 is given in Figure 4.

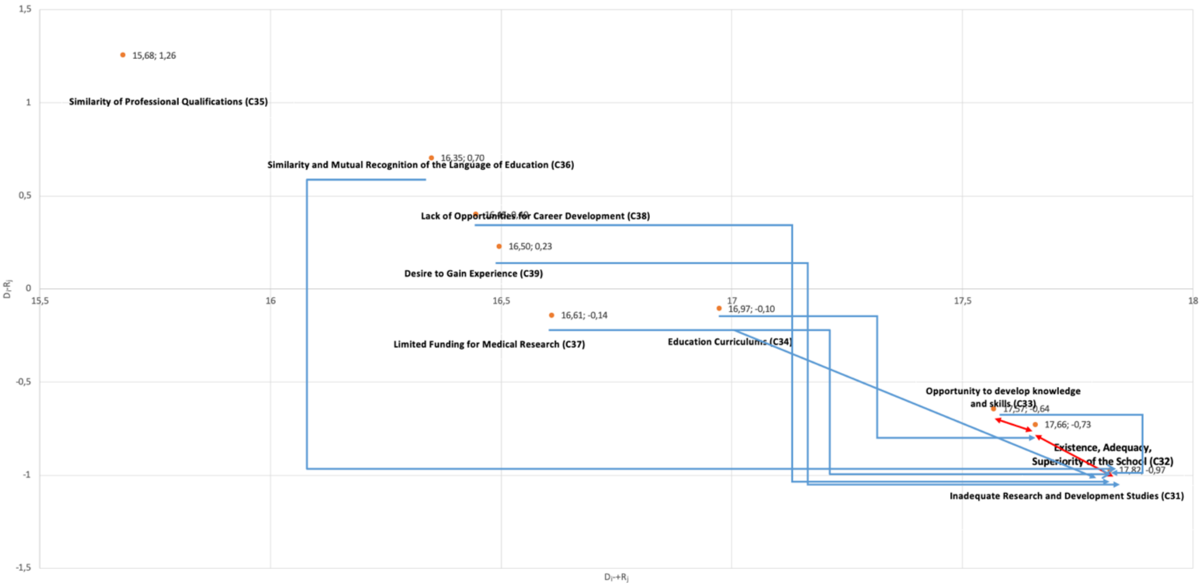

Figure 4. Effect-way graph diaphragm of C3

According to the findings related to C3, among the factors affecting the reasons for physicians to migrate, C31 is the most determinant factor, while C35 is the least important factor. According to the D_i-R_j value, among the factors that are influenced by other criteria; C32, C33, C34, and C37 criteria; it is noteworthy that C37 criterion is the factor that is influenced more than others. The factors with a higher power to influence other factors are C35 and C36, and the criterion with the highest power to influence others is C35.

According to Figure 4, C31 and C32 and C32 and C33 have mutual influence on each other. Criterion C31 influenced criterion C33; criterion C33 influenced criterion C31; criterion C34 influenced criteria C31 and C32; criterion C35 influenced criteria C31 and C33; criterion C36 influenced criteria C31 and C32; criterion C37 influenced criterion C31; criterion C38 influenced criteria C31 and C33; criterion C39 influenced criteria C31 and C32.

Analysis of Social and Cultural Factors Dimension with DEMATEL

The sub-criteria for C4 are given in Table 9. Analyses related to these criteria are given in the Supplementary File. The Influence-Directional Graph Diagram of Criterion C4 is given in Figure 5.

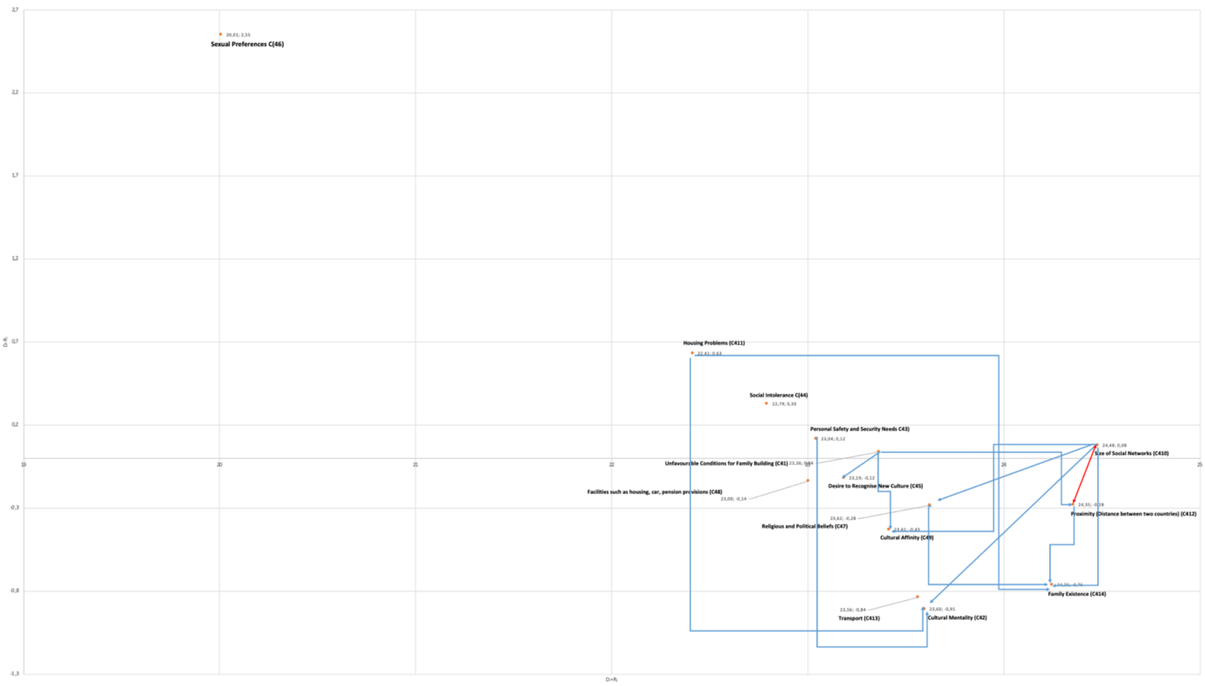

Figure 5. Influence-directional graph diaphragm of C4

According to the findings related to C4, it was determined that the most determinant factor among the factors affecting the reasons for physicians to migrate is C410, while the least important factor is C46. When analyzed according to the D_i-R_j value, among the factors influenced by other criteria; there are factors such as C42, C45, C47, C48, C49, C412, C413, C414. Among these factors, C42 is the most influential factor. The factors with a higher power to influence other factors are C41, C43, C44, C45, C410, C411. Among these, the factor with the highest influencing power on others is C46 since its D_i-R_j value is 2.55. According to Figure 5, criterion C41 influenced criteria C42 and C412, and criterion C410 influenced criteria C47 and C49. C42, C43, C44, C45 and C46 did not affect any criteria.

Analysis of Health Related Factors Dimension with DEMATEL

The sub-criteria for C5 are given in Table 9. Analyses related to these criteria are given in the Supplementary File. The Influence-Directional Graph Diagram of C5 Criteria is given in Figure 6.

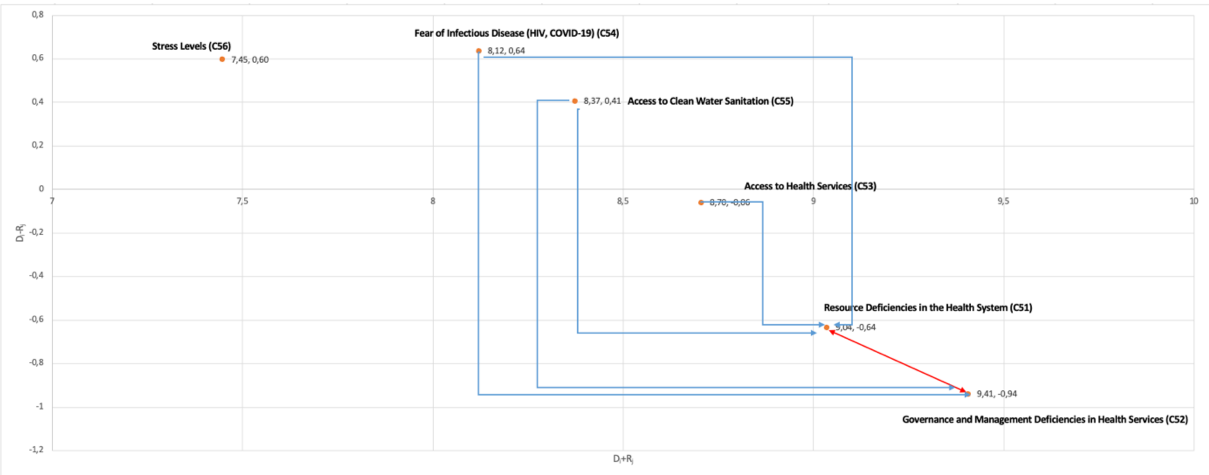

Figure 6. Influence-way graph diaphragm of C5

According to the findings related to C5, the most important factor affecting physicians' reasons for migration is C51, while the least important factor is C56. When analyzed according to the Di-Rj value, among the factors influenced by other criteria; there are factors such as C52 and C53. Among these factors, the most influential factor is C51. The factors with a higher power to influence other factors are C54 and C56. Among these two factors, the factor with the highest influence power on others is C54 since its D_i-R_j value is 0.63. According to Figure 7; C51 and C52 mutually influenced each other. Criterion C53 influenced criterion C51; C54 influenced criteria C51 and C52; C55 influenced criteria C51 and C52. Criterion C56 did not affect any criterion.

Analysis of the Factors Related to Work and Working Conditions Dimension with DEMATEL

The sub-criteria for C6 are given in Table 9. Analyses related to these criteria are given in the Supplementary File. The Influence-Directional Graph Diagram of Criterion C6 is given in Figure 7.

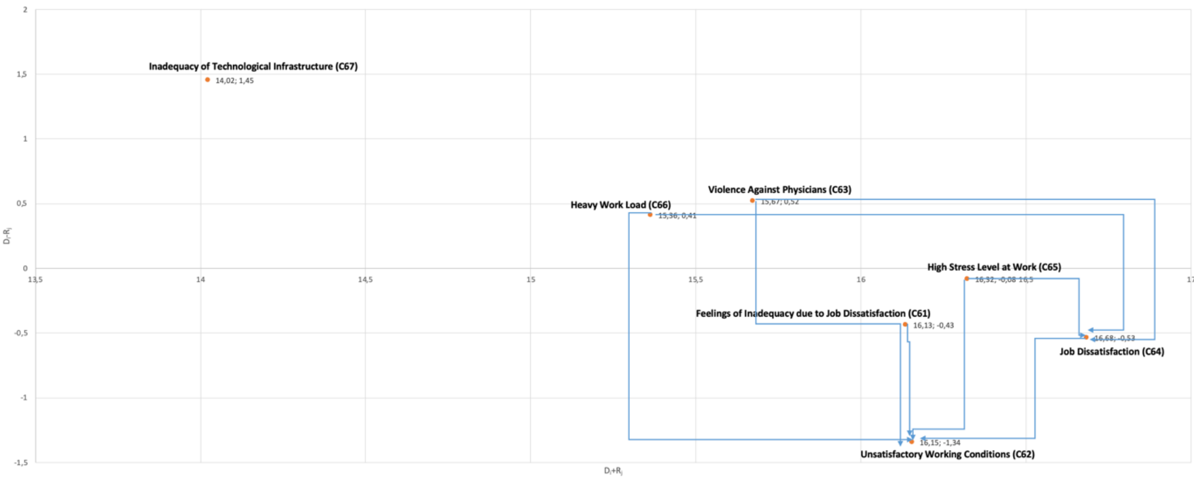

Figure 7. Influence-way graph diagram of C6

According to the findings related to C6, the most important factor affecting physicians' reasons for migration is C64, while the least important factor is C67. When analyzed according to the D_i-R_j value, the factors influenced by other criteria include C61, C62, C64 and C65. The most influential factor among these is C62. The factors with a higher power to influence other factors are C63 and C67. Among these two factors, C67 has the highest influence on others. According to Figure 8; C61 criterion influenced C62 criterion; C63 criterion influenced C62 and C64 criteria; C64 criterion influenced C62 criterion; C65 criterion influenced C62 and C64 criteria; C66 criterion influenced C64 criterion. C67 and C62 did not affect any criteria.

4. Discussion

The structuring and analysis of the research is based on a solid theoretical foundation of the DEMATEL methodology. This methodology requires less effort and cost than structural analysis models in terms of data acquisition, less involvement and computational simplicity (Agarwal & Kapoor, 2022). DEMATEL is a method that does not require the large number of participants required in structural models. Based on 12 experts' opinions, the analysis yielded similar results to previous studies. Therefore, it is seen that the method used provides advantages in terms of time and convenience.

This study was conducted to determine the causes of the physician brain drain and to examine the prioritization of these causes using the DEMATEL method. The importance levels of the main criteria that cause the migration of physicians were determined as: factors related to work and working conditions, factors related to education and career, political and social factors, economic factors, factors related to health, social and cultural factors. When considering the main factors that cause the brain drain phenomenon of physicians, significant and meaningful results were obtained in relation to the literature in terms of importance and priority.

Push-attract factors affecting health worker migration encompass a range of economic, professional and environmental considerations. Economic incentives, poor working conditions and lack of career progression emerge as important push factors (Sweileh, 2024). Studies reveal that better working conditions are an important factor in keeping physicians in the country of origin (Bidwell et al., 2014). It has been found that physicians who do not migrate experience burnout as non-migrating physicians take on the workload of migrating physicians. Burnout and workload overload among physicians also negatively affect health system delivery and reduce the quality of care (Ebeye & Lee, 2023). Limited opportunities for career development, promotion and professional advancement in their home countries lead health professionals to seek better prospects abroad (Castro-Palaganas et al., 2017). In a study of 1129 medical students in 2021, it was found that 52% of the students planned to go abroad after graduation (Uzun, 2021). A study on migrants by Adeniyi et al. (2022) found that early-career physicians migrate for better postgraduate education and salaries. Although some estimates vary from country to country, it is generally reported that approximately 100,000 Euros are spent annually to train a physician (Saluja et al., 2020).

In general, health professionals in developing countries migrate to countries with strong health systems and more opportunities for progress (Glinos et al., 2014; WHO, 2022). The migration of health professionals has increased with the migration of highly skilled specialists and the changing structure of the European Union (OECD, 2019). Labonte et al. (2015) note that in South Africa, this process has been addressed politically, with various initiatives contributing to reducing the shortage of health professionals. The countries' reputation for corruption at the government level, lack of accountability, limited health systems and inadequate legal systems are cited as driving factors for migration. It has been determined that physicians migrate because they do not want to take part in sociopolitical unrest that affects their personal and family security (Karan et al., 2016). Humphries et al. (2017) stated that Australia is an attractive migration location with both language suitability and recruitment policies. Both sending and receiving countries need to make political interventions by implementing appropriate retention strategies, improving working conditions, and encouraging international cooperation (Sweileh, 2024).

It has been noted that professionals often tend to reduce their dissatisfaction with their salary or working conditions by engaging in a dual practice in both the private and public sectors, and when they fail to do so, it becomes a reason for migration (Russo et al., 2018). Apostu et al. (2022) found that salary supplements without other measures had a low impact on the retention of migrant physicians in

Romania. It is estimated that low- and middle-income countries lose USD 15.86 billion annually due to physician migration (Saluja et al., 2020). It has been found that the remuneration policies of physicians are not sufficient to provide both personal and family livelihoods, and in this case, the social lives of physicians are negatively affected. It has been stated that they migrate with better prospects for their families and themselves (Karan et al., 2016; Ossa et al., 2020).

Latukha et al. (2021) emphasize the importance of talent management in their study on how to reverse brain drain and make gains. Goštautaitė et al. (2023) suggest focusing on human resource practices and ensuring equal opportunities to reduce physician migration. It is thought that increasing the number of other health professionals working in health institutions and organizing their duties will have a significant impact on alleviating the workload of doctors and increasing the number of doctors. In addition, it is suggested that incentives for research grants, personal development training, congresses, courses and other activities of physicians who migrate for educational purposes should be developed.

5. Conclusions

It is recommended that countries and policymakers develop a system to be aware of emerging migration and labor force trends and develop policies accordingly. Emphasis should be placed on why physicians migrate and what incentives can be developed to encourage them to return. These can be incentives such as improving overtime wages, providing guidance on housing and transportation, improving salaries, personal development opportunities, and educational support. In addition, in order to prevent health workers from preferring to go to rural areas, they can be directed to regions where health services are limited with appropriate incentives. It is thought that increasing the number of other health professionals working in health institutions and organizing their duties will have a significant impact on alleviating the workload of doctors and increasing the number of doctors. In addition, it is suggested that incentives for research grants, personal development training, congresses, courses and other activities of physicians who migrate for educational purposes should be developed.

Brain drain of physicians is a global health problem. Therefore, both high-income countries and low- and middle-income countries, in other words, both sending and receiving countries, need to address this problem mutually. This problem should be viewed from a health inequality perspective and national and regional strategies should be developed. It is thought that it would be valuable to examine each of the effective factors from a separate political perspective after determining the order of importance and to develop solutions accordingly.

In the study, the DEMATEL method was used to determine the factors related to physician migration. In the DEMATEL method, the opinions of experts are accepted with the same weight. Depending on the experience of the experts, a different method can be preferred to determine their weights. To deepen this issue, it is thought that conducting research by utilizing different MCDM techniques will contribute to the literature.

Acknowledgements
We would like to thank TÜBİTAK for their support of this study, which was carried out under the scope of the TÜBİTAK 2209-A Research Project Support Programme for Undergraduate Students.

Conflict of Interests

The authors have no potential conflicts of interest to declare in respect of the research or the writing and publication of the paper.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Ethical Aspects of the Research

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Council of Sakarya University of Applied Sciences (11.11.2023).

Supplementary

References

Adeniyi, M. A., Efuntoye, O., Popoola, G., Adebayo, O., Ekundayo, O., Ibiyo, M., Igbokwe, M. C., Ogunsuji, O., Fagbule, F., Egwu, O., et al. (2022). Profile and determinants of intention to migrate by early career doctors in Nigeria: A report from CHARTING study. International Journal of Health Planning and Management, 37, 1512–1525.

Agarwal, A. & Kapoor, K. (2022), Analysis of barriers in green human resource implementation for enhancing organizational sustainability: a DEMATEL approach, Transnational Corporations Review, Vol. 15 No. 2, pp. 1-15, in Press, doi: 10.1080/19186444.2021.1982575.

Apostu, S. A., Vasile, V., Marin, E., & Bunduchi, E. (2022). Factors influencing physicians' migration:A case study from Romania. Mathematics, 10, 505.

Apostu, S.A., Vasile, V., Marin, E., & Bunduchi, E. (2022). Factors influencing physicians migration-a case study from Romania. Mathematics, 10, 505.

Becker, R., & Teney, C. (2020). Understanding high-skilled intra-European migration patterns: the case of European physicians in Germany. Journal of Ethnic and Migration Studies, 46(9), 1737- 1755.

Bidwell, P., Laxmikanth, P., Blacklock, C., Hayward, G., Willcox, M., Peersman, W., Moosa, S., & Mant, D. (2014). Security and skills: The two key issues in health worker migration. Global Health Action, 7, 24194.

Botezat, A., & Moraru, A. (2020). Brain drain from Romania: What do we know so far about the Romanian medical diaspora? Eastern Journal of European Studies, 11, 309–334.

Braga, I. F., Ferreira, F. A., Ferreira, J. J., Correia, R. J., Pereira, L. F., & Falcão, P. F. (2021). A DEMATEL analysis of smart city determinants. Technology in Society, 66, 101687.

Buchan, J. (2008). How can the migration of health service professionals be managed so as to reduce any negative effects on supply? World Health Organization. Regional Office for Europe.

Byrne, J.-P., Conway, E., McDermott, A. M., Matthews, A., Prihodova, L., Costello, R. W., & Humphries, N. (2021). How the organisation of medical work shapes the everyday work experiences underpinning doctor migration trends: The case of Irish-trained emigrant doctors in Australia. Health Policy, 125(4), 467–473 https://doi.org/10.1016/j.healthpol.2021.01.002

Castro-Palaganas E., Spitzer D. L., Kabamalan M. M., Sanchez M. C., Caricativo R., Runnels V., Labonté R., Murphy G. T., & Bourgeault I. L. (2017). An examination of the causes, consequences, and policy responses to the migration of highly trained health personnel from the Philippines: The high cost of living/leaving—A mixed method study. Human Resources for Health, 15(1), 25. https://doi.org/10.1186/s12960-017-0198-z

Çelikbilek, Y. & Özdemir, M. (2020). Çok kriterli karar verme yöntemleri. Ankara: Nobel Akademik Yayıncılık.

Che, Y., Deng, Y., & Yuan, Y. (2022). Maximum-entropy-based decision-making trial and evaluation laboratory and its application in emergency management. Journal of Organizational and End User Computing (JOEUC), 34(7), 1-16. https://doi.org/10.4018/JOEUC.302891

Davda, L. S., Radford, D.R. & Gallagher, J. E. (2021). Migration, Retention and return migration of health professionals; Comment on doctor retention: A cross-sectional study of how Ireland has been losing the battle. International Journal of Health Policy and Management, 10(10), 667-669. doi: 10.34172/ijhpm.2020.225

Dohlman, L., DiMeglio, M., Hajj, J., & Laudanski, K. (2019). Global brain drain: How can the Maslow theory of motivation improve our understanding of physician migration? International journal of environmental research and public health, 16(7), 1182. https://doi.org/10.3390/ijerph16071182

Domagała, A., Kautsch, M., Kulbat, A., & Parzonka, K. (2022). Exploration of estimated emigration trends of Polish health professionals. International Journal of Environmental Research and Public Health, 19(2), 940. https://doi.org/10.3390/ijerph19020940

Dubas-Jakóbczyk, K., Domagala, A., Kiedik, D., Peña-Sánchez, J. N. (2020). Exploring satisfaction and migration intentions of physicians in three university hospitals in Poland. Int. J. Environ. Res. Public Health, 17, 43.

Ebeye, T., & Lee, H. (2023). Down the brain drain: A rapid review exploring physician emigration from West Africa. Global Health Research and Policy, 8(1), 23. https://doi.org/10.1186/s41256-023-00307-0

Farooque, M., Jain, V., Zhang, A., & Li, Z. (2020). Fuzzy DEMATEL analysis of barriers to blockchain-based life cycle assessment in China. Computers & Industrial Engineering, 147, 106684.

Gedam, V., Raut, R., Inamdar, Z., Narkhede, B., Dharaskar, S., & Narvane, V. (2022). COVID-19 critical success factors in Indian healthcare industry—A DEMATEL approach. Journal of Multi-Criteria Decision Analysis, 29(1-2), 135–149. https://doi.org/10.1002/mcda.1763

Glinos, I., & Buchan, J. (2014). Health professional mobility in a changing Europe: Lessons and findings. In J. Buchan, M. Wismar, I. A. Glinos, & J. Brenner (Eds.), Health professional mobility in a changing Europe: New dynamics, mobile individuals and diverse responses (pp. 17–33). World Health Organization.

Gökler, S. H., & Boran, S. (2024). A hybrid multi-criteria decision-making approach for hospital sustainability performance assessment. Business Process Management Journal. https://doi.org/10.1108/BPMJ-03-2024-0156

Goštautaitė, B., Mayrhofer, W., Bučiūnienė, I., & Jankauskienė, D. (2023). Mitigating medical brain drain: the role of developmental HRM and the focus on opportunities in reducing the self-initiated expatriation of young professionals. The International Journal of Human Resource Management, 35(3), 367–393. https://doi.org/10.1080/09585192.2023.2241815

Hadley, A. (2024). Migration and professional mobility: Rural attraction and retention of South African educated physicians. Social Science & Medicine, 350, 116884. https://doi.org/10.1016/j.socscimed.2024.116884

Humphries N., Connell J., Negin J., & Buchan J. (2019). Tracking the leavers: Towards a better understanding of doctor migration from Ireland to Australia 2008–2018. Human Resources for Health, 17(1), 36. https://doi.org/10.1186/s12960-019-0365-5

Kadaifci, C., Topcu, Y. I., & Mu, E. (2024). An analytic hierarchy process contingency analysis of factors affecting the emigration decision of medical doctors in Turkey. European Journal of Investigation in Health, Psychology and Education, 14(3), 623–647. https://doi.org/10.3390/ejihpe14030041

Karan, A., DeUgarte, D., & Barry, M. (2016). Medical "brain drain" and health care worker shortages: How should international training programs respond? AMA Journal of Ethics, 18(7), 665–675.

Karatuzla, M. "ağlık çalişanlarinda güncel konu incelemesi: beyin göçü, Genel Sağlık Bilimleri Dergisi, İstanbul Üniversitesi-Cerrahpaşa, Lisansüstü Eğitim Enstitüsü, 6 Sayı/Issue:1 Yıl:2024

Labonté R., Sanders D., Mathole T., Crush J., Chikanda A., Dambisya Y., Runnels V., Packer C., MacKenzie A., Murphy G.T., &Bourgeault I.L. (2015). Health worker migration from South Africa: Causes, consequences and policy responses. Human Resources for Health, 13, 92. https://doi.org/10.1186/s12960-015-0093-4

Latukha, M., Shagalkina, M., Mitskevich, E., & Strogetskaya, E. (2021). From brain drain to brain gain: the agenda for talent management in overcoming talent migration from emerging markets. The International Journal of Human Resource Management, 33(11), 2226–2255. https://doi.org/10.1080/09585192.2021.1949374

Morley, C. P., Wang, D., Mader, E. M., & Wallace, S. (2017). Analysis of the association between Millennium Development Goals 4 & 5 and the physician workforce across international economic strata. BMC International Health and Human Rights, 17(1), 18. https://doi.org/10.1186/s12914-017-0126-2

Murataj, N., Syla, B., Krasniqi, Y., Bahtiri, S., Bekaj, D., Beqiri, P., & Hoxha, I. S. (2022). Migration intent of health care workers during the COVID-19 pandemic in Kosovo. International Journal of Environmental Research and Public Health, 19(17), 11122. https://doi.org/10.3390/ijerph191711122

OECD (2019), Recent trends in International migration of doctors, Nurses and Medical Students, OECD Publishing, Paris, https://doi.org/10.1787/5571ef48-en.

OECD (2023), Ready for the next crisis? Investing in health system resilience, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/1e53cf80-en.

Okeke, E. N. (2014). Do higher salaries lower physician migration? Health Policy and Planning, 29, 603-614.

Oleribe, O. O., Udofia, D., Giwa, A., Foster, M., & Taylor-Robinson, S. D. (2022). Migration of highly-skilled workers: Personal perspectives. Pan African Medical Journal, 41, 292. https://doi.org/10.11604/pamj.2022.41.292.34644

Onah, C. K., Azuogu, B. N., Ochie, C. N., Akpa, C. O., Okeke, K. C., Okpunwa, A. O., Bello, H. M., & Ugwu, G. O. (2022). Physician emigration from Nigeria and the associated factors: The implications to safeguarding the Nigerian health system. Human Resources for Health, 20(1), 85. https://doi.org/10.1186/s12960-022-00770-1

Ossai, E. N., Una, A. F., Onyenakazi, R. C., & Nwonwu, E. U. (2020). Emigration plans after graduation of clinical medical students of Ebonyi State University Abakaliki, Nigeria: Implications for policy. Nigerian Journal of Clinical Practice, 23(7), 995–1003.

Parmar, P. S., & Desai, T. N. (2020). Evaluating sustainable lean six sigma enablers using fuzzy DEMATEL: A case of an Indian manufacturing organization. Journal of Cleaner Production, 265, 121802.

Russo, G., Fronteira, I., Jesus, T. S., & Buchan, J. (2018). Understanding nurses' dual practice: A scoping review of what we know and what we still need to ask on nurses holding multiple jobs. Human Resources for Health, 16, 14.

Saluja, S., Rudolfson, N., Massenburg, B. B., Meara, J. G., & Shrime, M. G. (2020). The impact of physician migration on mortality in low and middle-income countries: an economic modelling study. BMJ global health, 5(1), e001535. https://doi.org/10.1136/bmjgh-2019-001535

Shieh, J. I., Wu, H. H., & Huang, K. K. (2010). A DEMATEL method in identifying key success factors of hospital service quality. Knowledge-Based Systems, 23(3), 277–282. https://doi.org/10.1016/j.knosys.2010. 01.013

Sweileh, W. (2024). Research trends and patterns on international migration of health workers (1950–2022). SAGE Open, 14(4). https://doi.org/10.1177/21582440241293190

Tankwanchi, A. S., Hagopian, A., & Vermund, S. H. (2021). African physician migration to high-income nations: Diverse motives to emigrate ("We are not Florence Nightingale") or stay in Africa ("There is no place like home"): Comment on "Doctor retention: A cross-sectional study of how Ireland has been losing the battle." International Journal of Health Policy and Management, 10(10). https://doi.org/10.34172/ijh

Teney, C. (2019). Immigration of highly skilled European professionals to Germany: intra-EU brain gain or brain circulation? Innovation: The European Journal of Social Science Research, 34(1), 69–92. https://doi.org/10.1080/13511610.2019.1578197

Türk Tabipler Birliği (TTB) (2021). Türk tabipleri birliği merkez konseyi çalışma raporu 1 Eylül 2020 – 31 Mayıs 2021. https://l24.im/lTtsUov

Uzun, S. U. (2021). Pamukkale Üniversitesi Tıp Fakültesi Öğrencilerinin Beyin Göçüne Yönelik Tutumları ve İlişkili Etmenler. 5. Uluslararası 23. Ulusal Halk Sağlığı Kongresi, 13-18 Aralık 2021.

Vujicic, M., Zurn, P., Diallo, K., Adams, O., & Dal Poz, M. R. (2004). The role of wages in the migration of health care professionals from developing countries. Human Resources for Health, 2(1), 3. https://doi.org/10.1186/1478-4491-2-3

WHO (2023), World health statistics 2023, Available online:https://www.who.int/publications/i/item/9789240074323 (accessed on 18 March 2025)

WHO. Global Strategy on human resources for health: Workforce 2030; WHO: Geneva, Switzerland, 2016; Available online: https://apps.who.int/iris/bitstream/handle/10665/250368/9789241511131-eng.pdf (accessed on 11 March 2022).

Journal of Information Analytics, 1(1), 59-72

Submit 13.06.2025
Revision 01.07.2025
Accept 01.07.2025
Publish 02.07.2025

Journal of Information Analytics

Research Article

Open Access

Forecasting Lip Landmark Movements Using Time Series Models

Gözde Nergiz[1], Faruk Serin[2],

[1] Department of Computer Tecnology, Adana Alparslan Turkes Science and Technology University, Adana, Türkiye

[2]Department of Computer Engineering, Mersin University, Mersin, Türkiye

Abstract

This study examines the predictability of time series data derived from lip movements during speech using traditional statistical methods. The dataset was generated from a publicly available video, where x and y coordinates of 40 lip landmarks were extracted for each frame using Google's MediaPipe Face Mesh technology. Comprising a total of 3,242 frames and 80 time series, the dataset was analyzed by applying ARIMA and SARIMA models with various parameter combinations. The lowest Mean Absolute Percentage Error (MAPE) achieved was 0.0994 for the ARIMA model and 0.1331 for the SARIMA model. The most successful parameter combinations for the ARIMA model were typically p=5, d=0, q=1, while for the SARIMA model, the parameters p=1, d=0, q=3, P=0, D=0, Q=1, s=25 demonstrated the best performance.

Keywords Lip movement analysis, MediaPipe face mesh, time series forecasting, ARIMA, SARIMA

Citation: Nergiz, G., & Serin, F. (2025). Forecasting Lip Landmark Movements Using Time Series Models. *Journal of Information Analytics*, 1(1), 59-72.

Corresponding Author: Gözde Nergiz gnergiz@atu.edu.tr

1. Introduction

Humans rely on words to convey emotions and thoughts; however, facial expressions form the core of communication. These expressions, reflecting an individual's thought processes, emotions, and experiences, reveal their unique identity (Ekman, 2003). As social beings, humans employ not only verbal communication but also nonverbal elements such as gestures, facial expressions, and body language (Mast, 2007). The human face is the primary domain where these nonverbal communication elements are most prominently observed.

The face is an integrated structure composed of various subcomponents (eyes, eyebrows, nose, cheeks, and mouth). Among these, the mouth and its surrounding area serve as a central region for both speech production and the vivid expression of emotions. In particular, the lips are key elements that visually enhance the intensity and sincerity of expressions. This rich expressive capacity has laid the foundation for the development of modern facial recognition and tracking systems. With technological advancements, applications based on facial and lip analysis are increasingly integrated into daily life. Examples include facial recognition systems in smartphones, automatic tagging features on social media platforms, driver fatigue detection systems, and security applications. In healthcare, these technologies are utilized to enhance eye contact abilities in individuals with autism spectrum disorder, measure consumer reactions in marketing research, and support augmented reality applications. Given the significance of the human face, many new technologies and studies are actively being pursued. One prominent framework for researchers is the MediaPipe Face Mesh application (Lugaresi et al., 2019). The Face Mesh application is frequently preferred across various domains for its ability to automatically detect and track faces in images or videos (Adhikari et al., 2025; Aripin & Setiawan, 2024; Balaji & Sujatha, 2025; Jakhete & Kulkarni, 2024). Comparative studies have demonstrated its superiority over other methods. For instance, Jakhete and Kulkarni (2024) conducted a comprehensive study on emotion recognition through facial expressions, comparing various methods and datasets. They found that the MediaPipe Face Mesh model, capable of detecting 468 3D facial landmarks in real time, outperformed models like DLIB and OpenPose in terms of accuracy and speed. These technologies are employed in converting speech to text for individuals with hearing impairments, ensuring dubbing synchronization in film and video content, and detecting speech content remotely in forensic science. However, a key challenge in facial analysis is the temporary occlusion of specific facial regions, particularly the lips. Occlusion by hands, hair, or other objects can lead to missing or corrupted movement data (W. Zhang et al., 2023). The human face is characterized by relatively stable positions and sizes of facial components, largely unaffected by environmental factors, which impose strong structural constraints. This is a critical aspect often overlooked by current methods when addressing occluded landmarks (Li et al., 2024). Time series analysis methods have shown promising results in addressing such challenges. Time series analysis is a statistical approach used to examine patterns, trends, and variations in temporally observed data (Shumway & Stoffer, 2025). It is widely applied in fields such as finance (Lu & Xu, 2024; Sui et al., 2024), healthcare (Kong et al., 2024), energy (Gulay et al., 2024), and meteorology (Mishra et al., 2024; Ansari & Alam, 2024).

In this study, the human face in videos is detected using Face Mesh technology, with a focus on the lip region. The x and y coordinates of 40 distinct lip landmarks (resulting in 80-dimensional data) identified through Face Mesh are treated as time series across the video duration. The primary objective of the research is to apply ARIMA and SARIMA methods with different parameter combinations to these coordinate movements, compare the resulting MAPE values, and determine the most optimal forecasting model. The study's findings will contribute to predicting missing or corrupted lip coordinates in cases of temporary occlusion, ensuring the continuity of facial movement analysis. The second section of this study presents a literature review. The third section details the proposed methodology. The fourth section evaluates the findings under the results and discussion section. The final section provides the conclusions.

2. Literature Review

Time series forecasting models can be categorized as univariate and multivariate. Univariate models include AR (Ding et al., 2010), MA (Tsay, 2005), ARMA (Brockwell & Davis, 2016), and ARIMA (Box et al., 2015), while SARIMA (Rosychuk et al., 2016) is prominent for seasonal data, and SARIMAX (Elamin & Fukushige, 2018) is used when exogenous variables are incorporated. For multivariate data, models such as VAR (Zivot & Wang, 2006), VARMA (Tsay, 2013), and VARMAX (Casals et al., 2012) are preferred. For short-term forecasting, methods like SES (X. Zhang et al., 2020) and Holt-Winters (Jiang et al., 2020) offer effective solutions. Various time series models have been employed for forecasting. ARIMA models, characterized by three parameters (p, d, q), are widely applied across domains such as the furniture industry (Yucesan et al., 2018), healthcare (Kadri et al., 2014; Wei et al., 2016; Xu et al., 2016), finance (Zhang et al., 2016), energy (Yuan et al., 2016; Cadenas et al., 2016), food industry (Tripathi et al., 2014), transportation (Mete et al., 2022; Serin et al., 2021), aquaculture (Siddique et al., 2025), climate (Wahyudi & Febriani, 2024). Variants such as vector-ARIMA (Mai et al., 2015), ARMA (Aboagye-Sarfo et al., 2015), SARIMA (Butler et al., 2016; Rosychuk et al., 2016), and MSARIMA (Aroua & Abdul-Nour, 2015) are also frequently utilized by researchers.

Siddique et al. (2025) analyzed air temperature and precipitation data from 2011–2022 in Mymensingh, Bangladesh, using ARIMA models to forecast trends for 2023–2030. Their aim was to predict the impact of climate variables on aquaculture and provide data-driven insights for planning. Data sourced from NASA was validated using Bangladesh Meteorological Department records. The optimal models were ARIMA (2,1,2) for temperature and ARIMA (3,0,2) for precipitation, selected based on statistical metrics such as BIC, RMSE, and MAPE, supported by ACF and PACF graphs. The forecasts indicate a significant temperature increase and precipitation decrease in Mymensingh in the coming years.

Kong et al. (2024) aimed to predict missing values in healthcare data in a time-aware manner. They used Truncated SVD to compress data, reducing redundancy and noise, followed by ARIMA for missing value prediction. Their approach improved accuracy by considering temporal dimensions and capturing essential data patterns, with experiments on the WISDM dataset demonstrating its effectiveness and efficiency.

Wahyudi and Febriani (2024) compared SARIMA and SARIMAX models to predict particulate organic carbon (POC) levels in Indonesia's Sunda Shelf waters using MODIS data from 2002–2022. The models were SARIMA (3,1,3) x (2,0,0,60) and SARIMAX (3,1,3) x (2,0,0,60), with SARIMAX incorporating exogenous variables like sea surface temperature, chlorophyll-a, and salinity. Although SARIMAX had a lower AIC, validation metrics (MAPE, RMSE, correlation coefficient) showed SARIMA's superior performance. Forecasts suggest POC levels will fluctuate seasonally between 108.3–135.9 mg/m^3 from 2022–2030, peaking during the northwest monsoon season.

Kumar et al. (2024) compared Holt-Winters Exponential Smoothing (HWES) and ARIMA models to enhance demand forecasting and dynamic pricing strategies. Tested on real-world data, the models were evaluated for reducing lost sales and optimizing revenue under uncertain market conditions. Their dynamic pricing model, designed for limited sales seasons, also analyzed lost sales patterns. The findings indicate ARIMA's superior performance over HWES in volatile market conditions.

3. Material and Method

Dataset

The dataset used in this study was created from a publicly available YouTube video, with its characteristics detailed in Table 1, utilizing MediaPipe Face Mesh technology (Lugaresi et al., 2019).

Table 1. Basic Information About the Video Used for the Dataset

	Video File Info
Duration	129.68 saniye (~2 dakika 10 saniye)
Frame Rate	25 fps
Total Frame Count	3242

MediaPipe operates in real time, detecting 468 three-dimensional (3D) landmarks on a face. To precisely track lip movements during speech, specific lip landmarks defined by MediaPipe were selected. These landmarks were divided into two groups: the outer lip contour and the inner lip contour, each comprising 20 points. The x and y coordinates of each lip landmark were extracted, forming a dataset with 3,242 rows (corresponding to the total number of video frames) and 83 columns (including id, time, frame, and x and y coordinate values for the 40 landmarks).

ARIMA and SARIMA Models

The AR(p) and MA(q) models applied to forecasting are represented as in Equations (1) and (2), respectively (Yule, 1926; Wold, 1938).

$$Y_t = \sum_{i=1}^{p} a_i\, Y_{t-i} + \varepsilon_t \tag{1}$$

$$Y_t = \varepsilon_t + \sum_{j=1}^{q} b_j\, \varepsilon_{t-j} \tag{2}$$

where a_i are non-seasonal AR parameters, ε_t is zero mean Gaussian noise and b_j are non-seasonal MA parameters.

The ARMA (p, q) model combines p autoregressive terms and q moving average terms, as shown in Equation (3):

$$Y_t = c + a_1 Y_{t-1} + \cdots + a_p Y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q} \tag{3}$$

In cases of non-stationary data, differencing is required to achieve stationarity, as in the ARIMA model (Box et al., 2015).

Performance Measurements

The forecasting results were evaluated using the Mean Absolute Percentage Error (MAPE), as defined in Equation (4).

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n} \frac{|Y_t - \hat{Y}_t|}{Y_t}\right) * 100 \tag{4}$$

4. Experimental Results

In this study, the ARIMA method was applied with different combinations to each of the 80 time series (40 x-coordinates and 40 y-coordinates) derived from the coordinates of 40 distinct lip landmarks. The combinations were determined by setting the hyperparameter ranges for p, d, and q as 0–9, 0–1, and 0–3, respectively, resulting in 80 different combinations per series, totaling 6,400 evaluations.

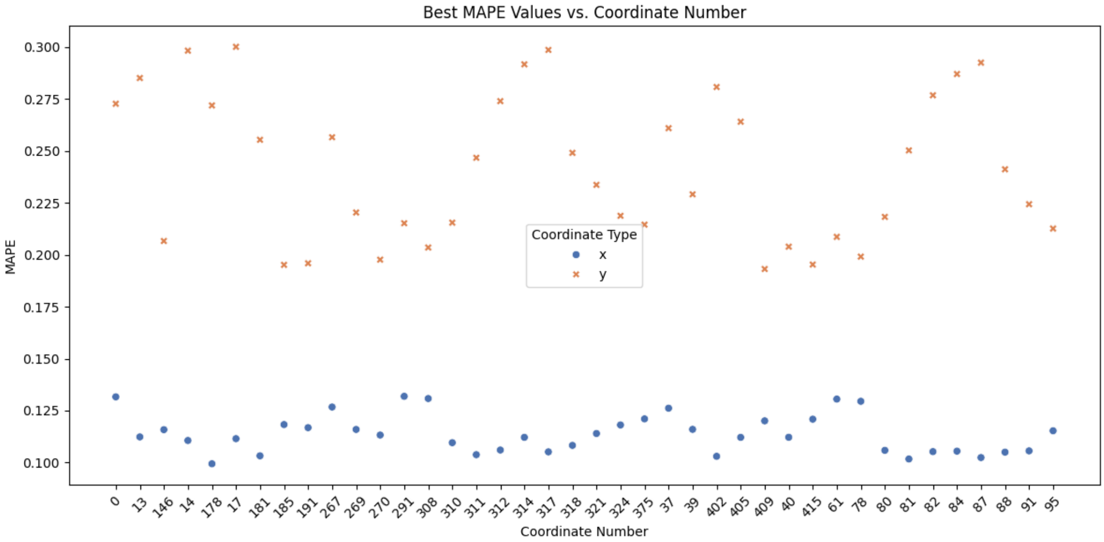

Figure 1. Best MAPE Value for Each Series

Figure 1 displays the best MAPE values for each time series. In the figure, "x" markers (orange) represent the values obtained from the time series of y-coordinates of lip landmarks, while dot markers (blue) represent those from x-coordinates. The x-coordinates (blue) are generally concentrated in the 0.10–0.15 MAPE range, while y-coordinates (orange) are scattered in the 0.20–0.30 range. This suggests that the coordinate type (x or y) is a determining factor in the obtained values.

Table 2. MAPE Statistics by Coordinate Type

Coord. Type	Min	Max	Mean	Median
x	0.10	0.13	0.11	0.11
y	0.19	0.30	0.24	0.24

Table 2 presents the basic statistics obtained by considering the best MAPE values for x and y coordinates. The MAPE average for x-coordinates is 0.113807, with a median of 0.112231, while for y-coordinates, these values are 0.24 and 0.24, respectively. The minimum and maximum values also show that y-coordinates are distributed over a wider range (0.19–0.30) compared to x-coordinates, suggesting greater variation in y-series and thus greater difficulty in prediction.

Table 3. Top 10% of Coordinates with the Lowest MAPE Values and ARIMA Parameter Values

Coordinate	p value	d value	q value	MAPE Score
178_x	5	0	1	0.0994
81_x	5	0	1	0.1017
87_x	5	0	1	0.1024
402_x	5	0	3	0.1030
181_x	9	0	1	0.1032
311_x	3	0	2	0.1038
88_x	5	0	1	0.1050
317_x	6	0	3	0.1051

* MAPE values were rounded to four decimal places to enhance numerical clarity across coordinates.

Table 3 lists the p, d, q values used in the ARIMA method and the corresponding MAPE values for the top 10% of coordinate series with the lowest MAPE values. The first row indicates that for the time series of the x-coordinate of lip landmark 178, applying the ARIMA method with p=5, d=0, and q=1 resulted in a MAPE value of 0.0994.

Table 4. Top 10% of Coordinates with the Highest MAPE Values and ARIMA Parameter Values

Coordinate	p value	d value	q value	MAPE Score
17_y	7	0	2	0.3002
317_y	1	0	3	0.2987
14_y	1	0	3	0.2983
87_y	1	0	3	0.2925
314_y	2	0	3	0.2917
84_y	7	0	2	0.2871
13_y	9	0	1	0.2851
402_y	2	0	3	0.2808

* MAPE values were rounded to four decimal places to enhance numerical clarity across coordinates.

Table 4 shows the highest MAPE values (the top 10% of the 8-coordinate series), ranging from 0.3002 for 17_y to 0.2808 for 402_y. All series in the table consist of y-coordinates of the landmarks. The ARIMA parameters are generally observed to be p=1 or p=2, q=3. This indicates that the modeling challenges for y-coordinates are exacerbated with certain parameter combinations.

Table 5. Top 5 Most Successful ARIMA Parameter Combinations

Count	p value	d value	q value
8	5	0	1
7	7	0	2
6	6	0	1
6	8	0	3
6	4	0	2

Table 5 presents the top five ARIMA parameter combinations based on the frequency of their use in achieving the best MAPE values for the 80 different series. The table shows that the combination p=5, d=0, q=1, used 8 times, ranks first. Additionally, the fact that the d value is 0 in the combinations listed in the table suggests that differencing is generally unnecessary.

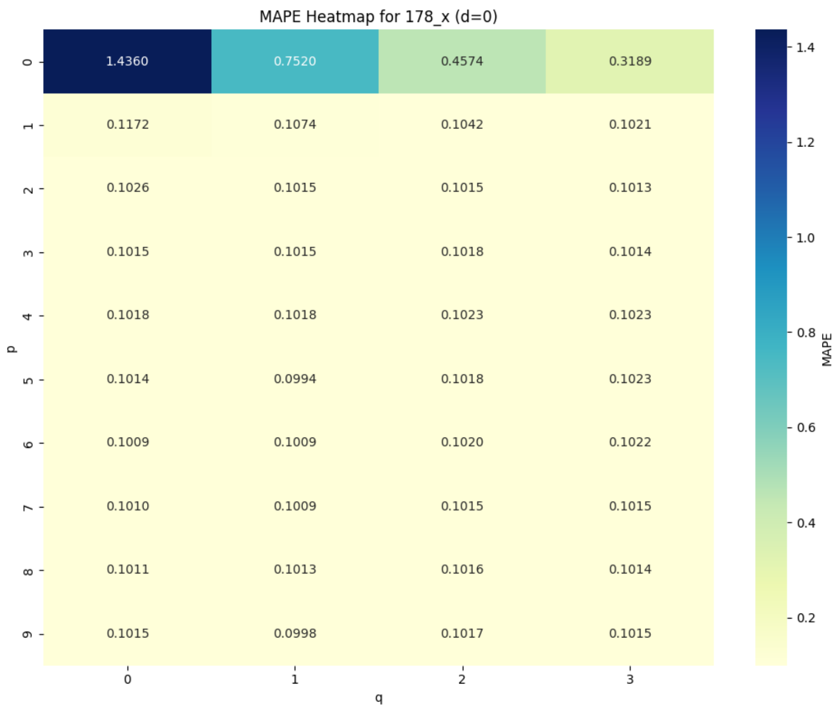

Figure 2. MAPE Heatmap for the x-Coordinate of Landmark 178 (d=0)

Figure 2 displays the MAPE heatmap for the x-coordinate of landmark 178, which achieved the best result among the MAPE values obtained with the ARIMA model. With d=0, the lowest MAPE (approximately 0.0994) was obtained with the combination p=5, q=1, while the highest MAPE (1.4360) was obtained with p=0, q=0.

For the application of the SARIMA method to the 80 different time series, hyperparameter ranges were set as p (0–9), d (0–1), q (0–3), P (0–1), D (0–1), Q (0–1), and s (25), resulting in 640 combinations per series and a total of 51,200 evaluations.

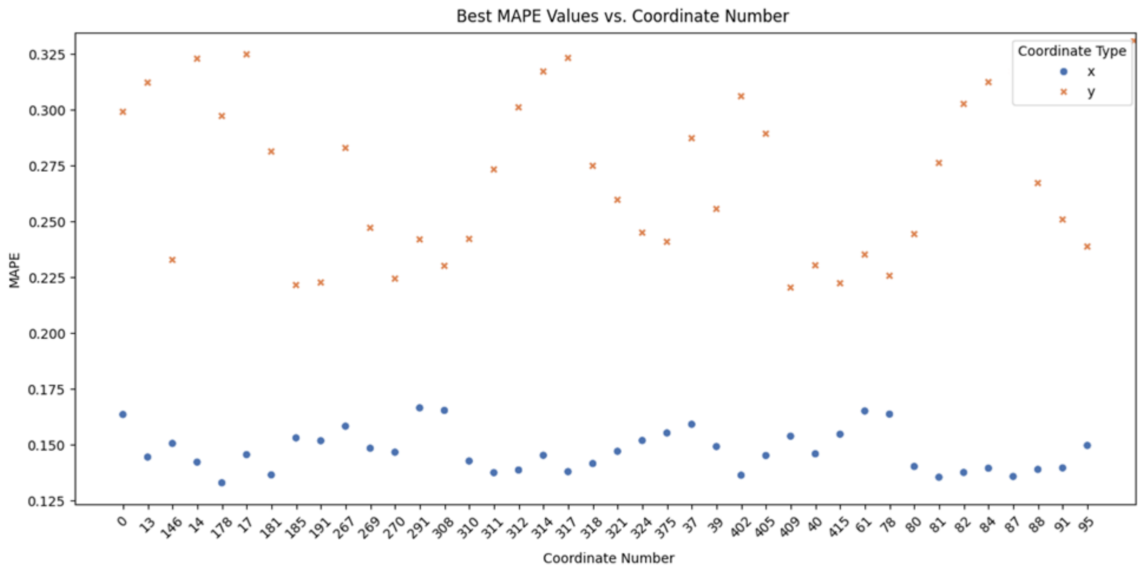

Figure 3. Best MAPE Value for Each Series

Figure 3 shows the best MAPE values for each series according to coordinate numbers. The x-coordinates (blue dots) are generally concentrated in the 0.133–0.166 MAPE range, while y-coordinates (orange crosses) are distributed in the 0.220–

0.325 range. This indicates that, similar to the ARIMA results, the coordinate type (x or y) is a determining factor in MAPE values. The lower error rates for x-coordinates suggest that these series are easier to predict compared to y-coordinates.

Table 6. MAPE Statistics by Coordinate Type

Coord. Type	Min	Max	Mean	Median
x	0.13	0.17	0.15	0.15
y	0.22	0.32	0.27	0.26

Table 6 shows that for x-coordinates, MAPE statistics are min 0.13, max 0.17, mean 0.15, and median 0.15; for y-coordinates, they are min 0.22, max 0.32, mean 0.27, and median 0.26. These values indicate that y-coordinates have approximately 80% higher average error rates and are distributed over a wider range. Overall, x-coordinates exhibit more consistent and lower error rates, while y-coordinates are more variable and challenging to predict.

Table 7. Top 10% of Series with the Lowest MAPE Values

Coordinate	p value	d value	q value	P value	D value	Q value	S value	MAPE Score
178_x	1	0	3	0	0	1	25	0.1331
81_x	1	0	3	1	0	0	25	0.1356
87_x	1	0	3	0	0	0	25	0.1359
402_x	1	0	3	0	0	0	25	0.1365
181_x	1	0	3	0	0	1	25	0.1366
311_x	1	0	3	0	0	0	25	0.1376
82_x	1	0	3	0	0	0	25	0.1377
317_x	1	0	3	0	0	0	25	0.1381

* MAPE values were rounded to four decimal places to enhance numerical clarity across coordinates.

Table 7 presents the combination values for the series in the top 10% with the lowest MAPE values. It is observed that all series in the top 10% percentile consist of X series.

Table 8. Top 10% of Series with the Highest MAPE Values

Coordinate	p value	d value	q value	P value	D value	Q value	S value	MAPE Score
17_y	0	1	3	0	0	0	25	0.3248
317_y	0	1	3	0	0	0	25	0.3232
14_y	0	1	3	0	0	0	25	0.3229
87_y	0	1	3	0	0	0	25	0.3172
314_y	0	1	3	0	0	0	25	0.3172
84_y	0	1	3	0	0	0	25	0.3124
13_y	0	1	3	0	0	0	25	0.3122
402_y	0	1	3	0	0	0	25	0.3061

* MAPE values were rounded to four decimal places to enhance numerical clarity across coordinates.

Table 8 shows the series in the worst 10% of the 80 different series based on the best MAPE values obtained using the SARIMA model, along with their parameter values and MAPE scores. The dominance of y-coordinates in the highest MAPE values confirms the modeling challenges for these series. The combinations p=0 and q=3 is frequently observed, but d=1 appears to increase the error rate in some cases.

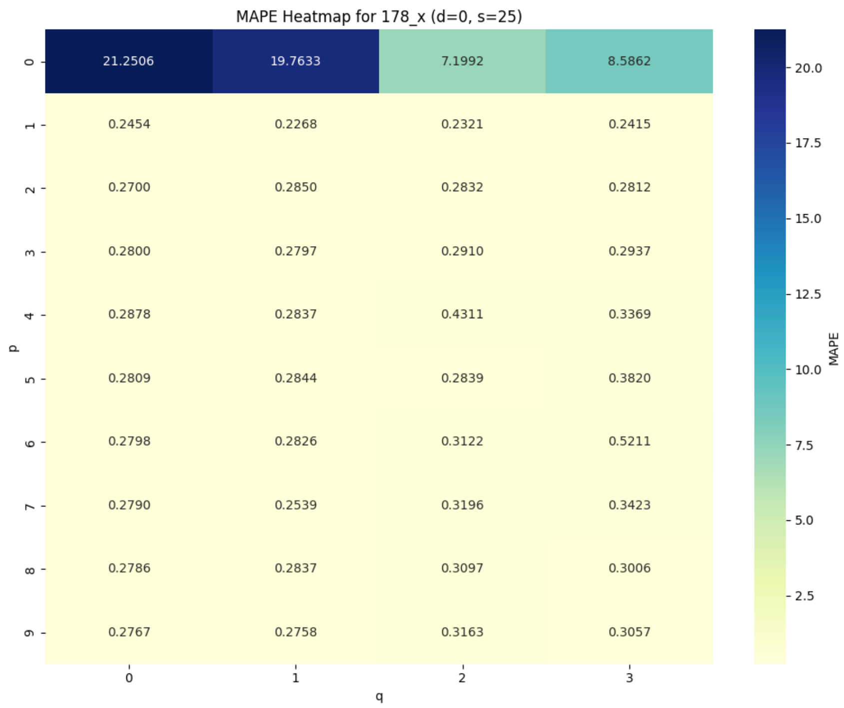

Figure 4. MAPE Heatmap for the x-Coordinate of Landmark 178 (d=0, s=25)

Figure 4 presents the MAPE heatmap for the x-coordinate of landmark 178, which achieved the best MAPE value among the 80-coordinate series when evaluated with the SARIMA method. With d=0 and s=25, the lowest MAPE (approximately 0.0994) was obtained with p=5, q=1, while the highest MAPE (1.4360) was obtained with p=0, q=0. With d=0 fixed, the effect of p and q values on MAPE is clearly demonstrated.

The most successful SARIMA parameter combinations and their frequency values are shown in Table 9. The prevalence of d=0 and q=3 combinations suggests that differencing is generally unnecessary, and a high moving average component is effective. The concentration of seasonal parameters (P, Q) at low values may indicate a limited seasonality effect.

Table 9. Top 5 Most Successful SARIMA Parameter Combinations

p value	d value	q value	P value	D value	Q value	S value	Freq.
0	1	3	0	0	0	25	37
1	0	3	0	0	0	25	27
0	1	3	0	0	1	25	5
0	1	3	1	0	0	25	4
1	0	3	0	0	1	25	3

Table 1-A (see appendix) lists the best and worst MAPE values obtained with both ARIMA and SARIMA models for each coordinate, along with the parameter combinations used to achieve these values. Generally, the best ARIMA combinations, typically 5, 0, 1 or similar parameter settings, yield MAPE values in the 0.09–0.13 range, while the best SARIMA combinations, mostly 1, 0, 3, 0, 0, 0, 25 or 1, 0, 3, 0, 0, 1, 25 produce MAPE values in the 0.13–0.32 range. Additionally, higher MAPE values are generally observed for y-coordinates. Among the worst-performing models, certain SARIMA combinations (e.g., 6, 1, 3, 0, 0, 1, 25) exhibit extremely high error rates (MAPE values exceeding 3000).

5. Discussion and Future Works

In this study, time series data derived from x and y coordinates of 40 lip landmarks, extracted using MediaPipe Face Mesh technology, were modeled using ARIMA and SARIMA methods, with forecasting performance evaluated via MAPE values. The results indicate that x-coordinates generally exhibit lower error rates compared to y-coordinates. This suggests that horizontal (x-axis) movements are temporally more regular and predictable, while vertical (y-axis) movements display greater variability. The ARIMA model, with its simpler structure and fewer parameter requirements, produced successful results for many coordinates. Notably, parameter combinations such as p=5, d=0, q=1, with minimal differencing (d=0), yielded low MAPE values. The SARIMA model, incorporating seasonal components, resulted in a broader error range for some series, indicating limited seasonality in lip coordinate data.

For future work, incorporating multivariate time series approaches (e.g., VAR, VARMA, VARMAX) could more effectively capture dependencies and simultaneous movements among lip landmarks, potentially yielding better results. Additionally, comparing machine learning and deep learning-based forecasting models (e.g., LSTM, GRU) with traditional methods is recommended as a meaningful direction for future research.

References

Aboagye-Sarfo, P., Mai, Q., Sanfilippo, F. M., Preen, D. B., Stewart, L. M., & Fatovich, D. M. (2015). A comparison of multivariate and univariate time series approaches to modelling and forecasting emergency department demand in Western Australia. Journal of Biomedical Informatics, 57, 62-73.

Adhikari, M., Joshi, P., Shrestha, S., & Shaik, S. (2025). Vision-based driver drowsiness and distraction detection through behavioral indicators of fatigue. SoutheastCon 2025, 261–266. https://doi.org/10.1109/SoutheastCon56624.2025.10971515

Ansari, M., & Alam, M. (2024). An intelligent IoT-cloud-based air pollution forecasting model using univariate time-series analysis. Arabian Journal for Science and Engineering, 49(3), 3135–3162. https://doi.org/10.1007/s13369-023-07876-9

Aripin, & Setiawan, A. (2024). Indonesian lip-reading detection and recognition based on lip shape using face mesh and long-term recurrent convolutional network. Applied Computational Intelligence and Soft Computing, 2024(1), 6479124. https://doi.org/10.1155/2024/6479124

Aroua, A., & Abdul-Nour, G. (2015). Forecast emergency room visits–a major diagnostic categories-based approach. International Journal of Metrology and Quality Engineering, 6(2), 204.

Balaji, S., & Sujatha, D. P. (2025). Hybrid eye-tracking system for cursor control: a kalman filter and exponential moving average-based approach for robust face tracking. Research Square. https://doi.org/10.21203/rs.3.rs-6541901/v1

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control. John Wiley & Sons.

Brockwell, P. J., & Davis, R. A. (2016). ARMA Models. In P. J. Brockwell & R. A. Davis, Introduction to time series and forecasting (pp. 73–96). Springer International Publishing. https://doi.org/10.1007/978-3-319-29854-2_3

Butler, M. B., Gu, H., Kenney, T., & Campbell, S. G. (2016). P017: Does a busy day predict another busy day? A time-series analysis of multi-centre emergency department volumes. CJEM, 18(S1), S83-S84.

Cadenas, E., Rivera, W., Campos-Amezcua, R., & Heard, C. (2016). Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. Energies, 9(2), 109.

Casals, J., García-Hiernaux, A., & Jerez, M. (2012). From general state-space to VARMAX models. Mathematics and Computers in Simulation, 82(5), 924–936. https://doi.org/10.1016/j.matcom.2012.01.001

Ding, J., Han, L., & Chen, X. (2010). Time series AR modeling with missing observations based on the polynomial transformation. Mathematical and Computer Modelling, 51(5–6), 527–536. https://doi.org/10.1016/j.mcm.2009.11.016

Ekman, P. (2003). Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.(2003). Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Xvii, 267.

Elamin, N., & Fukushige, M. (2018). Modeling and forecasting hourly electricity demand by SARIMAX with interactions. Energy, 165, 257–268. https://doi.org/10.1016/j.energy.2018.09.157

Gulay, E., Sen, M., & Akgun, O. B. (2024). Forecasting electricity production from various energy sources in Türkiye: A predictive analysis of time series, deep learning, and hybrid models. Energy, 286, 129566. https://doi.org/10.1016/j.energy.2023.129566

Jakhete, S. A., & Kulkarni, N. (2024). A comprehensive survey and evaluation of Mediapipe face mesh for human emotion recognition. 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), 1–8. https://doi.org/10.1109/ICCUBEA61740.2024.10775188

Jiang, W., Wu, X., Gong, Y., Yu, W., & Zhong, X. (2020). Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. Energy, 193, 116779. https://doi.org/10.1016/j.energy.2019.116779

Kadri, F., Harrou, F., Chaabane, S., & Tahon, C. (2014). Time series modelling and forecasting of emergency department overcrowding. Journal of Medical Systems, 38(9), 107-127.

Kong, L., Li, G., Rafique, W., Shen, S., He, Q., Khosravi, M. R., Wang, R., & Qi, L. (2024). Time-aware missing healthcare data prediction based on ARIMA model. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 21(4), 1042–1050. https://doi.org/10.1109/TCBB.2022.3205064

Kumar, L., Khedlekar, S., & Khedlekar, U. K. (2024). A comparative assessment of holt winter exponential smoothing and autoregressive integrated moving average for inventory optimization in supply chains. Supply Chain Analytics, 8, 100084. https://doi.org/10.1016/j.sca.2024.100084

Li, X., Wu, K., & Zhang, S. (2024). Landmark-in-facial-component: Towards occlusion-robust facial landmark localization. Image and Vision Computing, 151, 105289. https://doi.org/10.1016/j.imavis.2024.105289

Lu, M., & Xu, X. (2024). TRNN: An efficient time-series recurrent neural network for stock price prediction. Information Sciences, 657, 119951. https://doi.org/10.1016/j.ins.2023.119951

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: a framework for building perception pipelines (No. arXiv:1906.08172). arXiv. https://doi.org/10.48550/arXiv.1906.08172

Mai, Q., Aboagye-Sarfo, P., Sanfilippo, F. M., Preen, D. B., & Fatovich, D. M. (2015). Predicting the number of emergency department presentations in Western Australia: A population-based time series analysis. Emergency Medicine Australasia, 27(1), 16-21.

Mast, M. S. (2007). On the importance of nonverbal communication in the physician–patient interaction. Patient Education and Counseling, 67(3), 315–318.

Mete, S., Çelik, E., & Gül, M. (2022). Predicting the time of bus arrival for public transportation by time series models. Journal of Transportation and Logistics, 7(2), 541–555.

Mishra, P., Al Khatib, A. M. G., Yadav, S., Ray, S., Lama, A., Kumari, B., Sharma, D., & Yadav, R. (2024). Modeling and forecasting rainfall patterns in India: A time series analysis with XGBoost algorithm. Environmental Earth Sciences, 83(6), 163. https://doi.org/10.1007/s12665-024-11481-w

Rosychuk, R. J., Youngson, E., & Rowe, B. H. (2016). Presentations to emergency departments for COPD: A time series analysis. Canadian Respiratory Journal, 2016, 1–9. https://doi.org/10.1155/2016/1382434

Serin, F., Alisan, Y., & Kece, A. (2021). Hybrid time series forecasting methods for travel time prediction. Physica A: Statistical Mechanics and Its Applications, 579, 126134. https://doi.org/10.1016/j.physa.2021.126134

Shumway, R. H., & Stoffer, D. S. (2025). Time series analysis and its applications: With R examples. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70584-7

Siddique, M. A. B., Mahalder, B., Haque, M. M., & Ahammad, A. K. S. (2025). Forecasting air temperature and rainfall in Mymensingh, Bangladesh with ARIMA: Implications for aquaculture management. Egyptian Journal of Aquatic Research, S1687428525000251. https://doi.org/10.1016/j.ejar.2025.02.009

Sui, M., Zhang, C., Zhou, L., Liao, S., & Wei, C. (2024). An ensemble approach to stock price prediction using deep learning and time series models. 2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS), 793–797. https://doi.org/10.1109/ICPICS62053.2024.10796661

Tripathi, R., Nayak, A. K., Raja, R., Shahid, M., Kumar, A., Mohanty, S., ... & Gautam, P. (2014). Forecasting rice productivity and production of Odisha, India, using autoregressive integrated moving average models. Advances in Agriculture, 2014.

Tsay, R. S. (2005). Analysis of Financial Time Series. John Wiley & Sons.

Tsay, R. S. (2013). Multivariate time series analysis: With R and financial applications. John Wiley & Sons.

Wahyudi, A. J., & Febriani, F. (2024). Time-series forecasting of particulate organic carbon on the Sunda Shelf: Comparative performance of the SARIMA and SARIMAX models. Regional Studies in Marine Science, 80, 103863. https://doi.org/10.1016/j.rsma.2024.103863

Wei, W., Jiang, J., Liang, H., Gao, L., Liang, B., Huang, J., ... & Qin, F. (2016). Application of a combined model with autoregressive integrated moving average (ARIMA) and generalized regression neural network (GRNN) in forecasting hepatitis incidence in Heng county, China. PloS one, 11(6), e0156768.

Wold, H. (1938). A study in the analysis of stationary time series [PhD Thesis, Almqvist & Wiksell]. https://www.diva-portal.org/smash/record.jsf?pid=diva2:492042

Xu, Q., Tsui, K. L., Jiang, W., & Guo, H. (2016). A hybrid approach for forecasting patient visits in emergency department. Quality and Reliability Engineering International, 32(8), 2751-2759.

Yucesan, M., Gul, M., & Celik, E. (2018). Performance comparison between ARIMAX, ANN and ARIMAX-ANN hybridization in sales forecasting for furniture industry. Drvna industrija: Znanstveni časopis za pitanja drvne tehnologije, 69(4), 357-370.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?–A study in sampling and the nature of time-series. Journal of the Royal Statistical Society, 89(1), 1–63.

Yuan, C., Liu, S., & Fang, Z. (2016). Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. Energy, 100, 384-390.

Zhang, G., Zhang, X., & Feng, H. (2016). Forecasting financial time series using a methodology based on autoregressive integrated moving average and Taylor expansion. Expert Systems, 33(5), 501-516.

Zhang, W., Zhang, X., & Tang, Y. (2023). Facial expression recognition based on improved residual network. IET Image Processing, 17(7), 2005–2014. Scopus. https://doi.org/10.1049/ipr2.12743

Zhang, X., Yu, Y., Xiong, F., & Luo, L. (2020). Prediction of daily blood sampling room visits based on ARIMA and SES model. Computational and Mathematical Methods in Medicine, 2020(1), 1720134. https://doi.org/10.1155/2020/1720134

Zivot, E., & Wang, J. (Eds.). (2006). Vector autoregressive models for multivariate time series. In Modeling Financial Time Series with S-PLUS® (pp. 385–429). Springer. https://doi.org/10.1007/978-0-387-32348-0_11

APPENDIX

Table 1-A. Best and worst MAPE values obtained with ARIMA and SARIMA methods

Coord.	ARIMA			SARIMA			
	BEST		WORST	BEST		WORST	
	(p, d, q)	MAPE	MAPE	(p, d, q, P, D, Q, s)	MAPE	(p, d, q, P, D, Q, s)	MAPE
178_x	5,0,1	0.09937608	1.435977992	1,0,3,0,0,1,25	0.133050513	6,1,3,0,0,1,25	3133.920292
81_x	5,0,1	0.10168164	1.447421532	1,0,3,1,0,0,25	0.135553991	0,0,0,0,0,0,25	100
87_x	5,0,1	0.10237386	1.266574391	1,0,3,0,0,0,25	0.135892189	0,0,0,0,0,0,25	100
402_x	5,0,3	0.10296468	0.828397457	1,0,3,0,0,0,25	0.136452465	6,1,3,0,1,0,25	161.8177657
181_x	9,0,1	0.10321259	1.528204297	1,0,3,0,0,1,25	0.136598705	0,0,0,0,0,0,25	100
311_x	3,0,2	0.10375558	0.829421795	1,0,3,0,0,0,25	0.137593851	0,0,0,0,0,0,25	100
82_x	8,0,3	0.10522601	1.270162108	1,0,3,0,0,0,25	0.137692494	6,1,3,0,1,0,25	331.1700564
317_x	6,0,3	0.10511561	0.923345096	1,0,3,0,0,0,25	0.138056544	6,1,3,0,1,0,25	138.279693
312_x	3,0,2	0.10599842	0.921313308	1,0,3,0,0,0,25	0.138771163	0,0,0,0,0,0,25	100
88_x	5,0,1	0.10497560	1.592677386	0,1,3,1,0,0,25	0.139040264	6,0,3,0,1,0,25	359.5971969
84_x	5,0,1	0.10545934	1.315978034	1,0,3,0,0,0,25	0.139583	0,0,0,0,0,0,25	100
91_x	5,0,3	0.10558966	1.711664151	1,0,3,0,0,0,25	0.139723399	4,0,2,1,1,0,25	18412.73995
80_x	6,0,1	0.10584419	1.605907296	1,0,3,1,0,0,25	0.140368225	7,1,3,0,1,0,25	1007.48025
318_x	9,0,2	0.10823836	0.787790034	1,0,3,0,0,0,25	0.141668595	0,0,0,0,0,0,25	100
14_x	8,0,3	0.11058498	1.087601349	1,0,3,0,0,0,25	0.142316379	6,1,3,0,1,0,25	194.285881
310_x	9,0,2	0.10953666	0.792216458	1,0,3,0,0,0,25	0.142806091	0,0,0,0,0,0,25	100
13_x	8,0,3	0.11231879	1.089366615	1,0,3,0,0,0,25	0.144568805	4,1,2,1,1,1,25	3136.260776
405_x	3,0,2	0.11209947	0.781406271	0,1,3,0,0,1,25	0.145260372	6,1,3,0,1,0,25	193.120676
314_x	6,0,3	0.11210699	0.896873503	1,0,3,0,0,1,25	0.145364045	6,1,3,0,1,0,25	863.2600201
17_x	5,0,1	0.11146452	1.092059035	1,0,3,0,0,0,25	0.145661558	6,1,3,0,1,0,25	669.6962154
40_x	5,0,1	0.11214378	1.738231615	0,1,3,1,0,0,25	0.146025867	0,0,0,0,0,0,25	100
270_x	8,0,2	0.11321370	0.747115433	1,0,3,0,0,0,25	0.146750405	0,0,0,0,0,0,25	100
321_x	7,0,2	0.11395582	0.739733634	0,1,3,0,0,1,25	0.14715464	0,0,0,0,0,0,25	100
269_x	3,0,2	0.11591097	0.785626852	1,0,3,0,0,0,25	0.148525467	0,0,0,0,0,0,25	100
39_x	6,0,1	0.11595688	1.572304405	1,0,3,0,0,0,25	0.149286127	0,0,0,0,0,0,25	100
95_x	7,0,1	0.11527360	1.720214371	0,1,3,0,0,0,25	0.149770269	0,0,0,0,0,0,25	100
146_x	4,0,2	0.11580781	1.852531197	1,0,3,0,0,0,25	0.150677859	0,0,0,0,0,0,25	100
191_x	3,0,3	0.11677169	1.735104706	0,1,3,0,0,0,25	0.151879802	0,0,0,0,0,0,25	100
324_x	3,0,3	0.11807483	0.781246805	1,0,3,0,0,0,25	0.151995233	0,0,0,0,0,0,25	100
185_x	3,0,2	0.11825533	1.866160322	0,1,3,1,0,0,25	0.153164012	6,0,3,1,0,1,25	19807.02369
409_x	3,0,3	0.12004112	0.757987122	1,0,3,0,0,0,25	0.15392803	4,0,2,1,1,1,25	170.2169865
415_x	6,0,3	0.12085731	0.78717809	1,0,3,0,0,0,25	0.154813162	4,0,2,0,1,0,25	8528.834526
375_x	7,0,3	0.12099382	0.759801628	0,1,3,0,0,0,25	0.155368895	4,0,2,0,1,0,25	321.3556024
267_x	3,0,2	0.12671326	0.913918311	1,0,3,0,0,0,25	0.158385372	0,0,0,0,0,0,25	100
37_x	6,0,1	0.12614272	1.342226212	1,0,3,0,0,0,25	0.159250053	0,0,0,0,0,0,25	100
0_x	6,0,1	0.13160650	1.113006584	1,0,3,0,0,0,25	0.163629547	0,0,0,0,0,0,25	100
78_x	8,0,2	0.12947171	1.887034148	1,0,3,0,0,0,25	0.163796981	0,0,0,0,0,0,25	100
61_x	4,0,2	0.13052547	1.978311226	1,0,3,0,0,0,25	0.165201677	6,0,3,1,0,1,25	166.0745253
308_x	6,0,3	0.13079112	0.800457802	1,0,3,0,0,0,25	0.165458509	0,0,0,0,0,0,25	100
291_x	3,0,3	0.13185803	0.805768958	1,0,3,0,0,0,25	0.166597358	0,0,0,0,0,0,25	100

409_y	4,0,3	0.19317847	10.79446741	0,1,3,0,0,0,25	0.220444474	4,0,2,0,1,0,25	1334.940218
185_y	7,0,3	0.19516749	10.82031819	0,1,3,0,0,0,25	0.221582356	0,0,0,0,0,0,25	100
415_y	4,0,3	0.19530656	10.84401865	0,1,3,0,0,0,25	0.222381734	0,0,0,0,0,0,25	100
191_y	9,0,2	0.19591937	10.86357107	0,1,3,0,0,0,25	0.222718994	6,1,3,1,0,1,25	185.307472
270_y	6,0,2	0.19766462	10.89146523	0,1,3,0,0,0,25	0.224439292	4,1,2,0,1,0,25	137.2688519
78_y	6,0,2	0.19911389	10.77444918	0,1,3,0,0,0,25	0.225715871	6,1,3,0,1,0,25	179.4914227
308_y	4,0,3	0.20350221	10.74694029	0,1,3,0,0,0,25	0.230164229	4,0,2,0,1,0,25	386.3229999
40_y	4,0,2	0.20391242	10.90726028	0,1,3,0,0,0,25	0.230396475	0,0,0,0,0,0,25	100
146_y	6,0,1	0.20665702	10.79970753	0,1,3,0,0,0,25	0.232837765	6,1,3,0,1,0,25	189.7121921
61_y	9,0,2	0.20864385	10.7357768	0,1,3,0,0,0,25	0.235220248	0,0,0,0,0,0,25	100
95_y	5,0,1	0.21264610	10.82086875	0,1,3,0,0,1,25	0.23879431	6,1,3,0,1,0,25	288.5158727
375_y	8,0,2	0.21452793	10.79135543	0,1,3,0,0,0,25	0.240945359	6,0,3,0,0,1,25	2740891.654
291_y	7,0,0	0.21520494	10.70296935	0,1,3,0,0,0,25	0.242001198	4,0,2,0,1,1,25	9308.670746
310_y	9,0,3	0.21547824	10.93042232	0,1,3,0,0,0,25	0.242220283	0,0,0,0,0,0,25	100
80_y	4,0,2	0.21826290	10.94321673	0,1,3,0,0,0,25	0.244380664	4,1,2,0,0,1,25	1238.126899
324_y	4,0,3	0.21878916	10.81026512	0,1,3,0,0,0,25	0.245031666	0,0,0,0,0,0,25	100
269_y	7,0,2	0.22037003	10.98728604	0,1,3,0,0,0,25	0.247175417	0,0,0,0,0,0,25	100
91_y	6,0,1	0.22436393	10.87953525	0,1,3,0,0,0,25	0.250905597	0,0,0,0,0,0,25	100
39_y	7,0,2	0.22912624	10.99522356	0,1,3,0,0,0,25	0.255641984	6,1,3,0,0,0,25	124.8762493
321_y	5,0,0	0.23368296	10.8963746	0,1,3,0,0,0,25	0.259719968	0,0,0,0,0,0,25	100
88_y	8,0,3	0.24118974	10.86410158	0,1,3,0,0,0,25	0.267244044	0,0,0,0,0,0,25	100
311_y	9,0,3	0.24671998	11.01742902	0,1,3,0,0,0,25	0.273314605	0,0,0,0,0,0,25	100
318_y	6,0,2	0.24908672	10.86429984	0,1,3,0,0,0,25	0.274936472	0,0,0,0,0,0,25	100
81_y	4,0,2	0.25027510	11.02458964	0,1,3,0,0,0,25	0.276241699	0,0,0,0,0,0,25	100
181_y	7,0,2	0.25539508	10.97539704	0,1,3,0,0,0,25	0.281352583	6,0,3,1,0,1,25	5099.08556
267_y	7,0,2	0.25660685	11.06260426	0,1,3,0,0,1,25	0.282948285	0,0,0,0,0,0,25	100
37_y	8,0,3	0.26097292	11.06809729	0,1,3,0,0,1,25	0.287303265	0,0,0,0,0,0,25	100
405_y	9,0,1	0.26409151	11.00209626	0,1,3,0,0,0,25	0.289306693	0,0,0,0,0,0,25	100
178_y	8,0,3	0.27193200	10.91805419	0,1,3,0,0,0,25	0.297232966	0,0,0,0,0,0,25	100
0_y	8,0,1	0.27273381	11.1160408	0,1,3,1,0,0,25	0.299110378	0,0,0,0,0,0,25	100
312_y	9,0,1	0.27399102	11.08319996	1,1,3,0,0,1,25	0.301104343	0,0,0,0,0,0,25	100
82_y	4,0,2	0.27683224	11.08541415	1,1,3,0,0,1,25	0.302656093	0,0,0,0,0,0,25	100
402_y	2,0,3	0.28080079	10.92425628	0,1,3,0,0,0,25	0.30609919	0,0,0,0,0,0,25	100
13_y	9,0,1	0.28511938	11.11326614	0,1,3,0,0,0,25	0.31219073	0,0,0,0,0,0,25	100
84_y	7,0,2	0.28706051	11.04852261	0,1,3,0,0,0,25	0.312403903	6,0,3,0,0,1,25	121349.6688
314_y	2,0,3	0.29171824	11.06421396	0,1,3,0,0,0,25	0.317171768	0,0,0,0,0,0,25	100
87_y	1,0,3	0.29252835	10.96507303	0,1,3,0,0,0,25	0.317228345	6,1,3,1,0,1,25	140.527565
14_y	1,0,3	0.29830442	10.99049135	0,1,3,0,0,0,25	0.322901189	0,0,0,0,0,0,25	100
317_y	1,0,3	0.29866457	10.9702455	0,1,3,0,0,0,25	0.323231939	0,0,0,0,0,0,25	100
17_y	7,0,2	0.30015594	11.08009806	0,1,3,0,0,0,25	0.324837695	6,0,3,0,0,1,25	951.9562228