NATURENGS
E-ISSN 2717-8013
https://dergipark.org.tr/en/pub/naturengs

An Effective Hybrid Model for Predicting Air Quality of Ankara

Anıl Utku [a], Ümit Can [a,*]

[a] Munzur University, Department of Computer Engineering, Tunceli, Turkey.

*Corresponding author

ARTICLE INFO

Received 29.12.2023
Accepted 17.03.2024

Doi: 10.46572/naturengs.1411983

ABSTRACT

Increasing industrialization, population growth, urbanization and increase in fossil fuel consumption lead to air pollution that affects human health by polluting the atmosphere. Particulate matter, known as PM10 and PM2.5, are air pollutants that can remain suspended in the air in solid, liquid or both states. Substances are described according to their aerodynamic diameter, known as particle size. Estimating particulate matter concentrations is very important for human health and the environment. In this study, a hybrid deep learning model was developed for air quality prediction using $PM_{2.5}$ and PM_{10} concentration data obtained from Bahçelievler, Demetevler, Sincan and Törekent air quality monitoring stations in Ankara. In the developed model, it was aimed to use the successful features of CNN and LSTM models. The developed CNN-LSTM model was compared with LR, RF, SVM, MLP, CNN and LSTM using MSE, RMSE, R^2, and MAE. Experimental results showed that the CNN-LSTM model outperformed the compared models and each station had an R^2 of approximately 0.9.

Keywords: CNN, LSTM, PM_{10}, $PM_{2.5}$.

1. Introduction

Particulate Matter (PM), which has negative effects on human health and the environment, refers to particulate matter present in the air in solid or liquid form [1]. Particulate matter with a diameter of 10 micrometers or less is called PM10, and particulate matter with a diameter of 2.5 micrometers or less is called PM2.5 [2]. Regular monitoring of PM2.5 and PM10 concentrations, which can cause cardiovascular and respiratory diseases, and controlling air pollution are important for the environment and human health [2, 3].

Particulate matter also damages ecosystems by causing air and water pollution [4, 5]. Predicting particulate matter concentrations using artificial intelligence methods is important for controlling and monitoring air quality and developing preventive strategies [6]. Artificial intelligence methods can extract patterns in complex and large amounts of data obtained through sensors. In this way, more successful results can be achieved for air quality monitoring systems.

Artificial intelligence-based air pollution detection systems allow air pollution to be monitored at specific time intervals [7]. A more successful forecast performance can be achieved by integrating data obtained from different external sources such as meteorology, industrial processes and traffic into forecasting processes with artificial intelligence methods [8].

Prediction models developed using artificial intelligence methods can adapt to changing seasonal and environmental conditions. Traditional air quality monitoring methods are often expensive and resource-intensive. Artificial intelligence-based forecasting models can monitor a wider geographic area with fewer resources and reduce maintenance costs.

For these reasons, $PM_{2.5}$ and PM_{10} predictions with artificial intelligence methods have great potential in air quality management and environmental health and significantly contribute to the literature in this field.

In this study, a hybrid deep learning model was developed due to the limitations of traditional methods used in weather forecasting. Traditional methods make predictions by making use of weather parameters such as humidity, temperature, and wind. Artificial intelligence-based methods enable the extraction of relationships and patterns in data based on historical data. In this way, complex relationships in sequential data can be modelled. This study presents a comparative analysis of traditional methods and deep learning techniques used in air pollution prediction. Using data from Bahçelievler, Demetevler, Sincan, and Törekent air quality stations in Ankara, applying these forecasts to an actual geographical region made this

* Corresponding author. e-mail address: ucan@munzur.edu.tr
ORCID : 0000-0002-8832-6317

study a valuable and practical application on a regional scale.

The originality of this study to the literature is as follows:

- A hybrid air pollution prediction model was developed using CNN and LSTM.
- The developed model was compared with traditional methods such as Convolutional Neural Network (CNN), Linear Regression (LR), Multilayer Perceptron (MLP), Random Forest (RF), Long Short Term Memory, and Support Vector Machine (SVM).
- This is the first study in the literature on air pollution prediction in Ankara using this dataset.
- CNN-LSTM outperformed the compared models.

1.1. Related Work

In this section, studies in the literature using machine learning and deep learning methods for air quality prediction were examined.

Utku et al. [9] proposed GA-LSTM model consisting of LSTM and genetic algorithm to detect $PM_{2.5}$ concentrations in Shanghai, London, Beijing, and Singapore. As a result of comparisons with popular methods, their proposed model got the best result.

Kristiani et al [10] propose a model combining LSTM and statistical methods for $PM_{2.5}$ air pollution prediction. Correlation Analysis, Ekstreme gradient boosting (XGBoost), and Chemical Processing methods were used to select key features. For SO2, PM10, and NO2, the chemically processed model (model B) had the highest accuracy compared to the other models. This model obtained approximately 1 percentage point lower RMSE values than the others. In addition, according to the RMSE values obtained, it was revealed that training with all station datasets has 3 points higher RMSE value than training with each station dataset.

In another study, one-year meteorological parameters and $PM_{2.5}$ values from Hunan Province, China were taken. XGBoost MLP models were used. The results of this study provide an important contribution to the development of accurate and effective models for $PM_{2.5}$ pollution prediction [11].

Zhang et al. [12] propose a hybrid model for utilizing data from air quality monitoring stations. The model is tested for $PM_{2.5}$ concentration prediction at three monitoring stations in three different regions in Lanzhou City, China. Experimental results showed that the proposed model has good temporal stability and generalization ability.

Czernecki et al. [13] conducted a study for short-term PM_{10} and $PM_{2.5}$ predictions. This study also investigated the influence of important meteorological variables. The dataset used in the study includes hourly PM_{10} and $PM_{2.5}$ concentrations measured during the winter season for 10 years at 11 urban air quality monitoring stations located in four large Polish regions. These regions have high population densities, low plains, and high plateaus. Four different machine learning methods were used in the study. As a result of the experimental studies, XGBoost showed the best performance.

Menares et al. [14] used ten years of air pollution and meteorological measurement data from monitoring stations in Santiago, Chile. Missing data in the dataset were reconstructed using a method based on discrete cosine transforms and photo-chemical estimators selected by unsupervised clustering. Different models for predicting $PM_{2.5}$ maximum concentrations were proposed in the study. LSTM and Deep Feedforward Neural Network are some of these models. As a result of the experimental studies, the LSTM model gave better results than the deterministic models used for the same region.

Determining the concentration of particulate matter in atmospheric air is of great importance for human health. Harishkumar et al. [15] conducted air quality forecasting using Taiwan Air Quality Monitoring data from 2012-2017. Their study proposed various machine learning models that were compared using metrics such as RMSE, MAE, MSE, and R^2. Their proposed models gave more successful results than other models.

2. Material and Method

This study was carried out to compare the hybrid deep learning model developed with traditional machine learning and deep learning methods in predicting PM2.5 and PM10 particulate matter pollution. The primary material of the study consists of rich data sources belonging to Bahçelievler, Demetevler, Sincan, and Törekent air quality stations located in the city of Ankara. The time series obtained from these stations were used to monitor the change in air quality parameters.

2.1. Dataset

In this study, data from Bahçelievler, Demetevler, Sincan and Törekent air quality monitoring stations in Ankara were used [16]. The data used includes daily recorded PM2.5 and PM10 values. The first 10 lines of the dataset are shown in Figure 1.

The dataset consists of 366 lines of observation data obtained between 01/01/2020 and 31/12/2020.

Datetime	BAHÇELİEVLER STATION		DEMETEVLER STATION		SİNCAN STATION		TÖREKENT STATION	
	PM10 (μg/m3)	PM 2.5 (μg/m3)	PM10 (μg/m3)	PM 2.5 (μg/m3)	PM10 (μg/m3)	PM 2.5 (μg/m3)	PM10 (μg/m3)	PM 2.5 (μg/m3)
01.01.2020 00:00:56	24.11	7.81	34.32	16.57	16.52	18.07	28.19	18.07
02.01.2020 00:00:56	9.55	1.68	15.15	5.93	13.11	34.01	44.48	34.01
03.01.2020 00:00:56	9.61	2.65	17.86	13.67	46.92	12.95	16.92	12.95
04.01.2020 00:00:56	26.12	18.34	42.14	28.36	39.02	7.89	24.59	7.89
05.01.2020 00:00:56	33.85	20.89	39.04	25.4	42.6	26.26	42.43	26.26
06.01.2020 00:00:56	29.55	8.14	19.06	19.05	50.32	18.05	29.31	18.05
07.01.2020 00:00:56	18.54	2.77	14.3	22.68	76.43	21.48	32.05	21.48
08.01.2020 00:00:56	22.09	7.96	22.09	16.92	37.51	27.2	36.41	27.2
09.01.2020 00:00:56	9.6	1.96	16.62	7.19	21.34	10.58	16.73	10.58
10.01.2020 00:00:56	36.19	11.85	61.66	29.49	59.7	43.08	64.1	43.08

Figure 1. The first 10 rows of the dataset

Figure 2 shows the change graphs of $PM_{2.5}$ and PM_{10} concentrations over time for each station.

2.2. Prediction Models

LR assumes that this relationship is linear; a line can represent the relationship between the dependent and independent variables [17]. LR tries to find the regression line that best fits the distribution of the dataset. This line determines the effect of the independent variable on the dependent variable and the magnitude of this effect [18].

RF is an ensemble learning technique combining multiple decision trees [19]. A more robust and more stable prediction model is obtained by learning each tree separately and then combining the results of these trees. RF is based on the principle of randomness. Randomness means that each tree is trained on a different subset of data samples and a random subset of variables. This allows each tree to learn differently and reduces the problem of overfitting [20]. RF brings together many decision trees and combines the results of these trees. In classification problems, the class with the most votes is estimated using the voting method. In regression problems, the predictions of these trees are averaged.

SVM is a machine learning technique used in regression and classification problems. In regression problems of SVM, the dependent variable is a continuous numerical value [21]. SVM uses a hyperplane to describe the relationship between data. This hyperplane tries to maximize a space around the data called the margin. This margin helps improve the accuracy of the regression estimates. The data points closest to the hyperplane are support vectors. These data points determine the position and margin of the hyperplane [22]. SVM can use kernel functions to process linearly separable data. These functions transform data in high-dimensional space, thus expressing nonlinear relationships [23].

MLP is an artificial neural network consisting of multiple layers, and there are complete connections between these layers. Each layer contains many artificial nerve cells or neurons [24]. MLP has a multi-layer structure that includes input layer, hidden layers, and output layer. Each layer comprises neurons and is fully interconnected, meaning each neuron is connected to every neuron in the previous layer [25]. Each neuron passes its inputs through an activation function. MLP's hidden layers increase the complexity of the model and help it learn nonlinear relationships. The output layer of the MLP produces the predictions or results of the model. The back propagation algorithm is used while MLP is trained on training data. This algorithm calculates the error (loss) by comparing the model's predictions with the actual results and makes the weight updates backward layer by layer to reduce this error [26].

CNN is generally used in image processing but can also be used indirectly for regression problems [27]. CNN is designed to process image data and uses convolution and pooling layers to learn feature maps. These feature maps represent different data features and allow these features to be learned hierarchically. CNNs can successfully perform classification tasks using these feature maps. CNN-based approaches can be used, especially regarding data with complex structures or image-based regression problems [28]. For example, specially adapted CNN models can be used in regression problems based on visual data, such as estimating the size of an object from an image or estimating a particular feature value from an image.

LSTM has been successfully applied in sequential data processing problems such as time series analysis [29]. LSTM includes memory cells in addition to traditional recurrent neural network models. Through these cells, information is forgotten and remembered. LSTM includes input gate, forget gate, and output gate that control the information flow [30].

2.3. Developed Model

The CNN-LSTM model uses the prominent features of CNN and LSTM to increase prediction success. CNN is successful for feature extraction from time series data. LSTM, on the other hand, can extract dependencies and relationships over time. The hybrid CNN-LSTM model is an effective model for processing large and complex time series data and producing future predictions.

Data preprocessing was performed before the models were applied to the dataset. Missing fields in the data were checked. Missing rows were filled with the average values of the columns. The dataset was normalized using MinMaxScaler. 80% of the dataset was split for training and 20% for testing. 10% of the training dataset was used for optimization of model parameters.

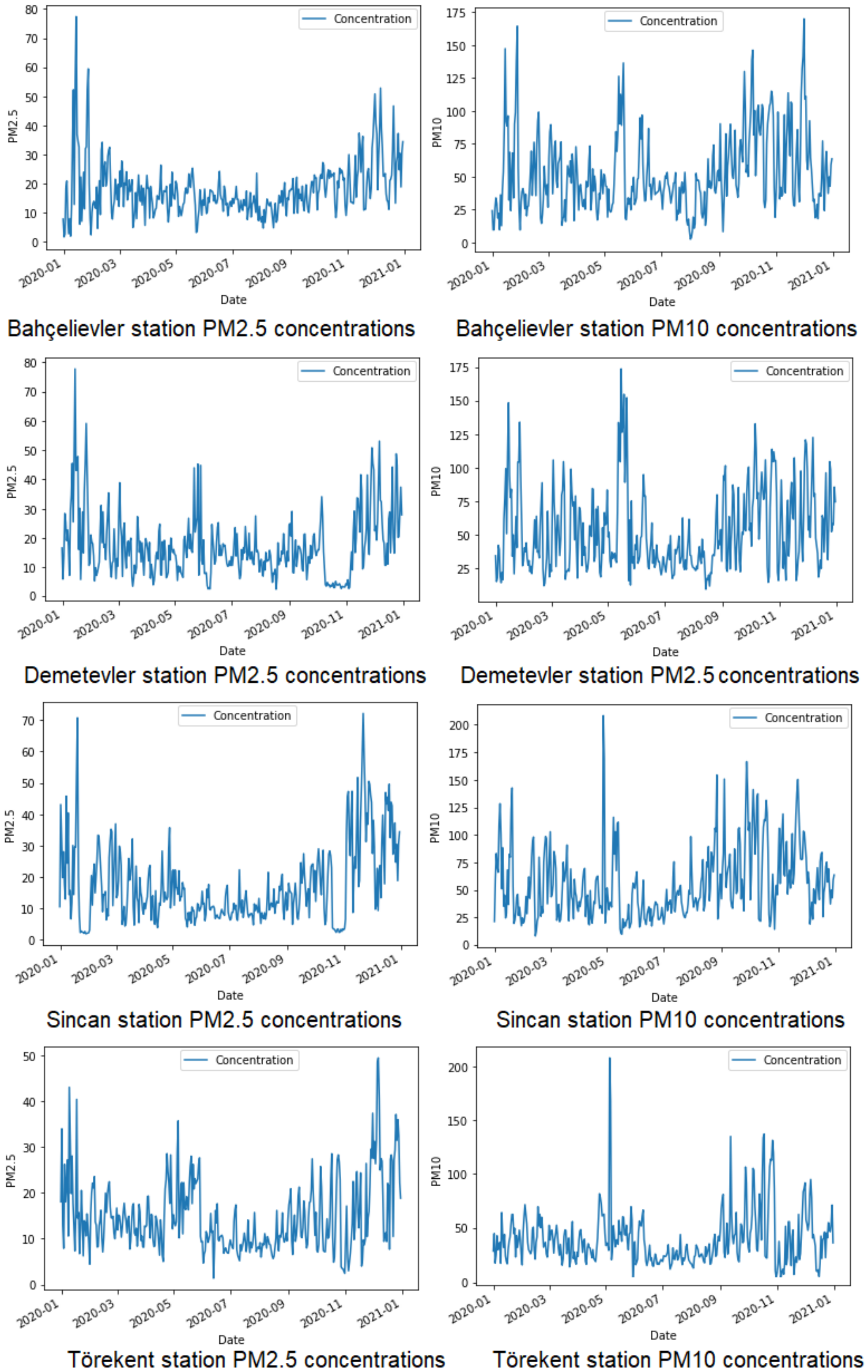

Figure 2. The graphs of change of PM2.5 and PM10 concentrations over time for each station

GridSearch was used to optimize the hyperparameters of the compared models. The hyper-parameters of the CNN-LSTM hybrid model and their values obtained as a result of GridSearchCV are shown in Table 1.

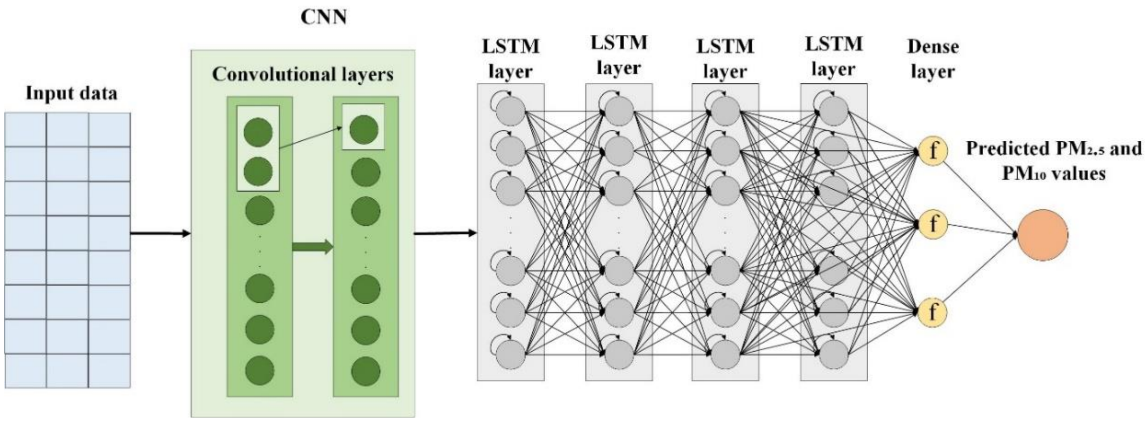

Figure 3. The architecture of the developed hybrid model

Table 1. The hyper-parameters of the CNN-LSTM hybrid model and their values

Hyper-parameter	Value
Epoch	80
Learning Rate	0.1
Batch size	64
Conv1D filters	32
Pool size	2
LSTM neurons	8
LSTM layers	4
Activation function	ReLU
Optimizer	Adam

The architecture of the developed hybrid model is shown in Figure 3.

As seen in Figure 3, the CNN-LSTM hybrid model consists of the input, convolutional, LSTM, fully connected, and output layers. The time series data is first fed to the input layer. The input layer converts to the appropriate dimensions to transmit the data to the CNN and LSTM layers.

The convolution layer is used to extract the features of the time series data. These layers use filters to recognize patterns and structures of data over time. Each convolution layer contains various filters for generating feature maps. After each layer, ReLU activation functions are used, and feature maps are generated. Pooling layers are used to reduce size and highlight important features.

Feature maps from CNN are transmitted to one or more LSTM layers. LSTM is used to capture dependencies and patterns over time. Each LSTM layer contains cells and gates (input, output, and forget gates). This mechanism is used to analyse the data's long-term dependencies and intra-time patterns. LSTM layers learn how the model combines historical information and current data. The outputs of the LSTM layers are transmitted to the fully connected layers. These layers allow the model to make predictions and produce results. An output layer generates predictive values, which are the model's final outputs. This layer predicts time series data target variables, such as $PM_{2.5}$ and PM_{10} values.

3. Experimental Results

This study aimed to develop a hybrid model for air quality prediction using PM2.5 and PM10 data obtained from Bahçelievler, Demetevler, Sincan and Törekent air monitoring stations in Ankara and to compare this model with other traditional methods. The developed hybrid model combines CNN and LSTM models and aims to make more precise predictions by integrating these two powerful methods. Experimental results evaluate the performance of the developed model by comparing it with traditional methods and reveal the contribution of this study to the field of air quality prediction. In this section, the performance of the model and the comparative results will be presented in detail. Table 2 shows the experimental results for Bahçelievler.

Table 2. The experimental results for Bahçelievler

	Model	MSE	RMSE	MAE	R^2
$PM_{2.5}$	LR	35.276	5.939	4.443	0.683
	RF	34.926	5.910	4.368	0.689
	SVM	33.724	5.807	4.314	0.706
	MLP	33.173	5.760	4.304	0.714
	CNN	33.589	5.796	4.347	0.708
	LSTM	23.184	4.815	3.502	0.819
	CNN-LSTM	**17.498**	**4.183**	**2.916**	**0.904**
PM_{10}	LR	500.016	22.361	18.114	0.636
	RF	570.618	23.887	18.697	0.585
	SVM	480.507	21.920	18.179	0.650
	MLP	465.643	21.579	17.731	0.661
	CNN	471.073	21.704	17.882	0.657
	LSTM	359.193	18.952	13.329	0.781
	CNN-LSTM	**234.288**	**15.306**	**11.496**	**0.906**

As seen in Table 2, CNN-LSTM outperformed the compared models. Following CNN-LSTM, LSTM, MLP, CNN, SVM, RF, and LR have been successful.

Table 3 shows the experimental results for Demetevler station.

Table 3. The experimental results for Demetevler

	Model	MSE	RMSE	MAE	R^2
$PM_{2.5}$	LR	80.038	8.946	6.858	0.521
	RF	78.300	8.848	6.447	0.534
	SVM	74.611	8.637	6.785	0.553
	MLP	60.155	7.756	5.888	0.640
	CNN	62.553	7.909	6.248	0.625
	LSTM	49.542	7.038	5.025	0.761
	CNN-LSTM	**37.158**	**6.095**	**4.459**	**0.901**
PM_{10}	LR	418.800	20.464	16.383	0.542
	RF	411.212	20.278	16.308	0.552
	SVM	394.634	19.865	15.662	0.573
	MLP	389.167	19.727	15.593	0.580
	CNN	392.812	19.819	15.658	0.575
	LSTM	234.676	15.319	12.117	0.847
	CNN-LSTM	**215.767**	**14.689**	**11.644**	**0.885**

As seen in Table 3, CNN-LSTM outperformed the compared models. Following CNN-LSTM, LSTM, MLP, CNN, SVM, RF and LR have been successful.

Table 4 shows the experimental results for Sincan.

Table 4. The experimental results for Sincan

	Model	MSE	RMSE	MAE	R^2
$PM_{2.5}$	LR	93.776	9.683	6.853	0.589
	RF	79.235	8.901	6.473	0.665
	SVM	72.382	8.507	6.077	0.701
	MLP	57.110	7.557	5.412	0.781
	CNN	58.734	7.664	5.427	0.772
	LSTM	48.676	6.976	4.844	0.807
	CNN-LSTM	**44.626**	**6.680**	**4.772**	**0.893**
PM_{10}	LR	452.035	21.261	16.800	0.641
	RF	437.628	20.919	16.440	0.655
	SVM	390.185	19.753	15.379	0.701
	MLP	383.574	19.585	15.316	0.707
	CNN	387.216	19.678	15.467	0.703
	LSTM	361.337	19.008	14.879	0.858
	CNN-LSTM	**317.813**	**17.827**	**13.902**	**0.883**

As seen in Table 4, CNN-LSTM outperformed the compared models. Following CNN-LSTM, LSTM, MLP, CNN, SVM, RF and LR have been successful.

Table 5 shows the experimental results for Törekent.

Table 5. The experimental results for Törekent

	Model	MSE	RMSE	MAE	R^2
$PM_{2.5}$	LR	39.009	6.245	4.859	0.609
	RF	37.543	6.127	4.652	0.627
	SVM	36.790	6.065	4.627	0.636
	MLP	32.131	5.668	4.255	0.695
	CNN	33.332	5.773	4.358	0.680
	LSTM	23.511	4.848	3.623	0.787
	CNN-LSTM	**21.225**	**4.607**	**3.385**	**0.855**
PM_{10}	LR	451.449	21.247	17.336	0.545
	RF	439.060	20.953	17.283	0.558
	SVM	411.131	20.276	16.442	0.586
	MLP	341.888	18.490	14.677	0.655
	CNN	355.548	18.856	14.764	0.642
	LSTM	282.969	16.821	12.987	0.907
	CNN-LSTM	**239.320**	**15.469**	**11.789**	**0.924**

As seen in Table 5, CNN-LSTM outperformed the compared models. Following CNN-LSTM, LSTM, MLP, CNN, SVM, RF, and LR have been successful.

Experimental results showed that the CNN-LSTM model was more successful than the compared models for each station. CNN-LSTM had an R^2 value above 0.85 for all stations and for both $PM_{2.5}$ and PM_{10}. For Bahçelievler station, it had an R^2 value over 0.9 for both $PM_{2.5}$ and PM_{10}.

The CNN-LSTM model outperforms the compared models in predicting both $PM_{2.5}$ and PM_{10} for each station. The fact that RF is more successful than LR can be interpreted as the relationships in the dataset are complex and nonlinear. In such datasets, models based on decision trees, such as RF, may perform better. RF can automatically assess the importance of features and give more weight to essential features. LR does not directly evaluate the importance of features and treats all features similarly.

The fact that SVM is more successful than RF can be interpreted by its ability to capture nonlinear relationships. SVM can better model complex and nonlinear relationships in time series data.

The fact that MLP is more successful than SVM can be interpreted with the multi-layered structure of MLP. Thanks to the activation functions in its structure, MLP better captures complex and non-linear relationships and automatically learns features. CNN, on the other hand, is designed especially for multidimensional image data. Therefore, CNN's ability to learn time series data is limited.

LSTM is a deep learning model specifically designed for sequential data. LSTM has special units that allow capturing long-term dependencies in time series. With the CNN-LSTM model, it is aimed to use the successful features of CNN and LSTM effectively. CNN is good at feature extraction, while LSTM is good at modelling and predicting time dependencies.

4. Conclusions

Air pollutants such as $PM_{2.5}$ and PM_{10} negatively affect human health and the environment by causing respiratory diseases and cardiovascular diseases. Therefore, accurately predicting the concentration of particulate matter is important for monitoring air quality and protecting public health and the environment. Predicting air pollution according to different time periods of the day can help protect children, the elderly and chronically ill patients in the risk group during periods when pollution is high. Additionally, it can contribute to strategies to be developed regarding air pollution, environment and climate. Due to the limitations of traditional approaches used in air quality prediction, the use of artificial intelligence methods in air quality prediction comes to the fore. Artificial intelligence methods enable the successful extraction of patterns and relationships in the data, thus increasing prediction accuracy. For this purpose, a hybrid model was developed using the prominent features of deep learning methods. In this study, it was aimed to contribute to taking measures to protect public health and the environment by accurately predicting $PM_{2.5}$ and PM_{10}

concentrations with the developed hybrid CNN-LSTM model. Experimental results show that the prediction results of the developed model are quite successful and promising. Integrating artificial intelligence-based methods into air quality models will increase both long-term and short-term forecast accuracy. In this way, the planning and strategies to be developed by governments and municipalities will be more effective.

References

[1] **Harishkumar, K. S., Yogesh, K. M., & Gad, I.** (2020). Forecasting air pollution particulate matter (PM2. 5) using machine learning regression models. *Procedia Computer Science*, 171, 2057-2066.

[2] **Kalia, P., & Ansari, M. A. (2020).** IOT based air quality and particulate matter concentration monitoring system. Materials Today: Proceedings, 32, 468-475.

[3] **Pio, C., Rienda, I. C., Nunes, T., Gonçalves, C., Tchepel, O., Pina, N. K., ... & Alves, C. A. (2022).** Impact of biomass burning and non-exhaust vehicle emissions on PM10 levels in a mid-size non-industrial western Iberian city. *Atmospheric Environment*, 289, 119293.

[4] **Zoran, M. A., Savastru, R. S., Savastru, D. M., & Tautan, M. N. (2020).** Assessing the relationship between surface levels of PM2. 5 and PM10 particulate matter impact on COVID-19 in Milan, Italy. Science of the total environment, 738, 139825.

[5] **Meo, S. A., Almutairi, F. J., Abukhalaf, A. A., & Usmani, A. M. (2021).** Effect of green space environment on air pollutants PM2. 5, PM10, CO, O3, and incidence and mortality of SARS-CoV-2 in highly green and less-green countries. International Journal of Environmental Research and Public Health, 18(24), 13151.

[6] **Olabi, A. G., Obaideen, K., Elsaid, K., Wilberforce, T., Sayed, E. T., Maghrabie, H. M., & Abdelkareem, M. A. (2022).** Assessment of the pre-combustion carbon capture contribution into sustainable development goals SDGs using novel indicators. Renewable and Sustainable Energy Reviews, 153, 111710.

[7] **Lutz, É., & Coradi, P. C. (2022).** Applications of new technologies for monitoring and predicting grains quality stored: Sensors, internet of things, and artificial intelligence. Measurement, 188, 110609.

[8] **Liu, X., Lu, D., Zhang, A., Liu, Q., & Jiang, G. (2022).** Data-driven machine learning in environmental pollution: gains and problems. Environmental science & technology, 56(4), 2124-2133.

[9] **Utku, A., Can, Ü., Kamal, M., Das, N., Cifuentes-Faura, J., & Barut, A. (2023).** A long short-term memory-based hybrid model optimized using a genetic algorithm for particulate matter 2.5 prediction. *Atmospheric Pollution Research*, 14(8), 101836.

[10] **Kristiani, E., Kuo, T. Y., Yang, C. T., Pai, K. C., Huang, C. Y., & Nguyen, K. L. P. (2021).** PM2. 5 forecasting model using a combination of deep learning and statistical feature selection. *IEEE Access*, 9, 68573-68582.Çalışma4

[11] **Peng, J., Han, H., Yi, Y., Huang, H., & Xie, L. (2022).** Machine learning and deep learning modeling and simulation for predicting PM2. 5 concentrations. *Chemosphere*, 308, 136353.Çalışma6

[12] **Zhang, Q., Wu, S., Wang, X., Sun, B., & Liu, H. (2020).** A PM2. 5 concentration prediction model based on multi-task deep learning for intensive air quality monitoring stations. *Journal of cleaner production*, 275, 122722.

[13] **Czernecki, B., Marosz, M., & Jędruszkiewicz, J. (2021).** Assessment of machine learning algorithms in short-term forecasting of pm10 and pm2. 5 concentrations in selected polish agglomerations. *Aerosol and Air Quality Research*, 21(7), 200586.

[14] **Menares, C., Perez, P., Parraguez, S., & Fleming, Z. L. (2021).** Forecasting PM2. 5 levels in Santiago de Chile using deep learning neural networks. *Urban Climate*, 38, 100906.

[15] **Harishkumar, K. S., Yogesh, K. M., & Gad, I. (2020).** Forecasting air pollution particulate matter (PM2. 5) using machine learning regression models. *Procedia Computer Science*, 171, 2057-2066.

[16] **https://ulasav.csb.gov.tr/dataset/06-hava-kalitesi-verileri**

[17] **Wang, Q., & Wang, L. (2020).** Renewable energy consumption and economic growth in OECD countries: A nonlinear panel data analysis. *Energy*, 207, 118200.

[18] **Wang, Q., Yang, T., & Li, R. (2023).** Does income inequality reshape the environmental Kuznets curve (EKC) hypothesis? A nonlinear panel data analysis. *Environmental Research*, 216, 114575.

[19] **Band, S. S., Janizadeh, S., Chandra Pal, S., Saha, A., Chakrabortty, R., Melesse, A. M., & Mosavi, A. (2020).** Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sensing*, 12(21), 3568.

[20] **Su, Y., Weng, K., Lin, C., & Zheng, Z. (2021).** An improved random forest model for the prediction of dam displacement. *IEEE Access*, 9, 9142-9153.

[21] **Ray, S. (2019, February).** A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE.

[22] **Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020).** A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189-215.

[23] **Taşkın, G., & Camps-Valls, G. (2021).** Graph embedding via high dimensional model representation for hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-11.

[24] **Zhang, J., Li, C., Yin, Y., Zhang, J., & Grzegorzek, M. (2023).** Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. *Artificial Intelligence Review*, 56(2), 1013-1070.

[25] **Xiao, X., Liu, J., Liu, D., Tang, Y., Dai, J., & Zhang, F. (2021).** SSAE-MLP: Stacked sparse autoencoders-based multi-layer perceptron for main bearing temperature prediction of large-scale wind turbines. *Concurrency and Computation: Practice and Experience*, 33(17), e6315.

[26] **Wang, J., Wu, H., Zhang, X., & Yao, Y. (2020).** Watermarking in deep neural networks via error back-propagation. *Electronic Imaging*, 2020(4), 22-1.

[27] **Mishra, M. (2021).** Machine learning techniques for structural health monitoring of heritage buildings: A state-of-the-art review and case studies. *Journal of Cultural Heritage*, 47, 227-245.

[28] **Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., ... & Mellit, A. (2023).** A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. *Sustainability*, 15(7), 5930.

[29] **Weerakody, P. B., Wong, K. W., Wang, G., & Ela, W. (2021).** A review of irregular time series data handling with gated recurrent neural networks. *Neurocomputing*, 441, 161-178.

[30] Harrou, F., Kadri, F., & Sun, Y. (2020). Forecasting of photovoltaic solar power production using LSTM approach. *Advanced Statistical Modeling, Forecasting, and Fault Detection in Renewable Energy Systems*, 3.