



Journal of Environmental Studies

Vol. 46, No. 1, Spring 2020

Journal Homepage: www.Jes.ut.ac.ir

Print ISSN: 1025-8620 Online ISSN 2345-6922

Optimization of Meteorological Variables to Predict Air Pollutant Concentrations for Use in Artificial Neural Network Model to Reduce the Cost and Time of Analysis

Document Type
Research Paper

Received
October 31, 2019

Accepted
March 11, 2020

Afsaneh Ghasemi¹, Jamil Amanollahi¹, Mohammad Darand²

1 Department of Environment, Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran

2 Department of Climatology, Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran

DOI: [10.22059/JES.2021.300440.1007998](https://doi.org/10.22059/JES.2021.300440.1007998)

Abstract

Kermanshah city is one of the polluted cities of the country due to traffic, industries and thunderstorms. In this study, five pollutants O₃, So₂, PM₁₀, No₂, Co were predicted using multilayer perceptron neural network for two time periods of today and tomorrow. Independent data include seven meteorological quantities of temperature, relative humidity, visibility, wind speed, dew point, pressure and precipitation. collinearity test and Forward Select technique were used to remove additional input variables and create subset of predictor variables. The optimal model was selected for each pollutant using RMSE, NMSE, R², IOA and FB indices. The results show that model 2 with 6 independent quantities is optimal for predicting CO and NO₂ pollutants and for predicting O₃ pollutants model 5 with 3 input quantities is a satisfactory model also for predicting SO₂ model 6 with two input variables and for prediction of PM₁₀, model 4 with 4 input variables were the most appropriate models. The results of this study show that using Forward Select technique to optimize the number of variables increases the accuracy and decreases prediction costs

Keywords: Air pollution, FS technique, Kermanshah, Multi-collinearity, Multi-layer perceptron network.

Extended abstract

Introduction

Today, air pollution is one of the main and most harmful problems in human societies, which has caused many environmental problems. Air quality is changing daily, even when the amount of pollutants entering the air is constant, factors that determine climate change, such as wind speed, wind direction, air mass thermal profile, amount of solar energy to perform photochemical reactions, wind duration or rainfall, alter air quality specifically. The air has a limited capacity and does not tolerate the discharge of various wastes and toxins that humans enter today.

Air pollution in cities is always a permanent and serious threat to the health and safety of the community and the environment. Recent studies show that the potential effects of air pollution on human health include increased mortality, increased hospitalization, and increased physiological changes in the body, especially respiratory and cardiovascular function. The exponential increase in population has led to rapid deforestation, rapid growth in industries and multiplicity of vehicles. Accelerating unmanaged urban development has led to a change in the chemical composition of the atmosphere which is associated with human activities. Industries, vehicles and other natural or human resources add a huge amount of air pollutants to the environment, which will lead to the destruction of air quality, and this has led to damage part of the environment. Predicting air pollution before increasing the level of these contaminations and prompt alarm can contribute to the health of the community. On the other hand, an increase in the number of independent variables would increase the cost and also increase the time required for predicting air pollution. Therefore, in this study, Forward Select (FS) technique was used to obtain the most suitable combination of independent variables with the most accurate prediction of specific contaminants in the artificial neural network (ANN) model.

Material and Methods

Kermanshah is an extensive city in the west of Iran country having industries and thunderstorms. The total population of Kermanshah city was 1952434 in the 2016 population census and its area is 250.45/4 km². This city is located in the middle of the western part of the country between a geographic orbit of 33°41' N to 35°17' N latitude of the equator and 45°24' E to 48° 0.6' E longitude of the Greenwich meridian. In this research, relative humidity, temperature, dew point, precipitation, pressure, wind speed, and visibility with the previous day's pollutant concentration were used as independent data. The Bureau of Meteorology (main synoptic station) of Kermanshah provided the meteorological data from 2014 to 2016. The five pollutant data consisting of Sulphur dioxide (SO₂), particulate matter PM₁₀, carbon monoxide (CO), nitrogen dioxide (NO₂), ozone (O₃) related to the air quality of Kermanshah were acquired from the Department of Environment (monitoring station) in Kermanshah from 2014 to 2016. The collinearity indicates that an independent variable is a function of the other variables. The Multi-collinearity test means that there is a linear relationship between two or more independent variables in the regression. If the collinearity of a regression equation is high, it means that there is a high dependence between the independent variables and it may not have a high validity due to the high coefficient of determining the model. In other words, although the model looks good but has no significant independent variables. Therefore, if the variables have a large linear relationship, the estimated statistics and regression coefficients may not be able to show the unique role and effect of each of the independent variables.

Multi-collinearity test was performed to remove additional input variables in SPSS software. The correlation between the independent and dependent variables is measured by two indicators of Variance Inflation Factor (VIF) and tolerance. The Variance Inflation Factor (VIF) indicates how much the variance of the estimated regression coefficients has increased since there are no correlated variables in the model. If the value of this index is close to one, there is no linearity. The relative tolerance factor is the relative scatter of a variable. If its value is close to one, it means that in an independent variable a small part of its scattering is justified by other independent variables, and if this value is close to zero, it means that one the variable is almost a linear combination of other independent variables. If the VIF values of the independent variables are more than 10 and tolerance is less than 0.2, then it can be said that the model is suffering from multi-collinearity.

To quantify the severity of the multi-collinearity, tolerance and Variance Inflation Factor (VIF) results were used. The Forward Selection (FS) method is based on regression and was used to select the best subsets of input variables. The Forward Selection (FS) technique has been used by many researchers to build powerful predictive models. This method is based on the degree of dependence of the independent variables with the dependent. After that, the variable that creates the most dependence with the dependent variable is considered as the first input and the variables with less dependence constitute a set of subsequent inputs. This step is repeated $n-1$ to evaluate the effect of each of the variables on the model output. A subset of the input variables is obtained to predict the outputs. The linear relation between variables created several models for each pollutant. Then the application of Multi-Layer Perceptron (MLP) network was used to predict pollutants in Matlab software.

To reduce errors and increase accuracy in forecasting, both independent and dependent data were normalized between zero and one. Neural networks are nonlinear models that are widely used to identify systems, predict time periods, and pattern. These networks can be tools for the flexibility of nonlinear regressions, which are generally composed of one or more layers with different neurons. The structure of the neural network typically consists of three layers, the input layer that distributes the data in the network, the hidden layer that processes the data, and the output layer that extracts the results for specific inputs. The Multi-Layer Perceptron (MLP) network has a hidden layer, and the flow of input to the network takes place in a forward path from layer to layer. There are two kind of signals in the Multi-Layered Perceptron (MLP) network, the function signal that travels in the path of departure, and the error signal that propagates in the return path of the network. One of the features of this network is its good computing features.

The Multi-Layer Perceptron (MLP) model is the most general network for predicting air pollution. It can solve nonlinear and very complex problems through the network topology. In this study, 20% of data were used in the training phase and 80% data in the test phase. The results obtained from the models were evaluated with indicators. The Root Mean Square Error (RMSE), Coefficient of determination (R^2), Normalized mean square error (NMSE), Fractional bias (FB), and the Index of agreement (IOA) indicators were used to determine the performance of the models. The Root Mean Square Error (RMSE), summarizes the difference between the observed and predicted concentration and shows the actual error model. Therefore, in an optimal model, the Root Mean Square Error (RMSE), should be close to zero.

The Coefficient of determination (R^2), indicates how much of the changes in the data observed by the model have been reproduced. The high Coefficient of determination (R^2) is a sign of desirability. A very small amount of the Normalized mean square error (NMSE), index indicates the implementation of the model in space and time, which is closer to zero, indicating the optimal model. Fractional bias (FB) indicates the low and high values of the predictions. If Fractional bias (FB), is equal to zero, it indicates that the forecast was very accurate. The Index of agreement (IOA) is a useful measure of model performance and has been proposed as an alternative to R and R^2 . This index is considered as the standard for measuring the mean square error.

Discussion and Results

In this study, to predict air pollutants in Kermanshah city and optimization effective variables in forecasting were used from the multi-layer perceptron model and Feature Selection (FS) technique. Results showed that relative humidity and temperature with VIF values of more than 10 and tolerance values below 0.2 exceeded the recommended value. After eliminating the both mentioned variables, the multi-collinearity test was repeated and results indicate that all variables were obtained within the recommended limit. The different subset of variables was developed using FS method by added one by one the variables constitute most to least correlation between input variables and dependent variable. For each pollutant, seven models were evaluated, but for O_3 pollutant, eight models were calculated due to the effect of pollutant NO_2 . The results of multilayer perceptron neural network analysis show that, MODEL 1 with $FB= 0.0170$, $IOA= 0.967$, $NMSE= 0.100$ and the highest $R^2= 0.7341$ was suitable for same-day predicting of CO. To predict one-day advance of CO, MODEL 2 and MODEL 5 have the highest R^2 values, but IOA statistics in MODEL 2 is more than MODEL 5,

and the values of FB and NMSE in MODEL 2 is lower than that of MODEL 5. So, MODEL 2 is more suitable for one-day advance of CO pollutants.

In predicting the PM₁₀ pollutant, the MODEL 4 has a maximum value of IOA= 0.960 and FB= 0.00151 with one of the lower value than other models in same day predicting of PM₁₀, and the MODEL 4 has the lowest amount of NMSE= 0.487 and RMSE= 0.0718 in one-day advance predicting of PM₁₀. So, MODEL 4 is selected for prediction of PM₁₀ pollutant as the optimal model. The prediction results of SO₂ pollutant indicate that the MODEL 3 has the lowest FB= -0.00302 and NMSE= 0.135 and the highest IOA= 0.943 and R²= 0.6118, respectively. Therefore, this MODEL is perfect for same-day prediction of SO₂ concentration.

Based on the result MODEL 6 with lowest values of NMSE= 0.105, FB= -0.0048, and the highest IOA= 0.972 is suitable MODEL for predicting one-day advance of SO₂.

In predicting NO₂ the MODEL 2 and MODEL 3 represents the highest performance compared to other models in same-day and one-day advance in prediction of NO₂ pollutant. Comparing the two models mentioned for NO₂ shows that both models have the same conditions in minimum and maximum values of the statistics, so considering the RMSE of the test phase, which is less in model 2 than model 6, indicate that model 2 it is a more appropriate model in predicting NO₂. The prediction results of the O₃ pollutant indicate that the MODEL 7 in the same day forecasting and the MODEL 5 in one-day forecasting in terms of the IOA index have the same value, and the indexes NMSE= 0.00120 and FB= 0.00137 in the MODEL 5 have the minimum values and in model 7, the value of R² = 0.711 is highest, so the input composition of MODEL 5 is considered as the optimal model.

Advantages of the Forward Selection (FS) technique include improving model accuracy, reducing computational time in model construction, facilitating data visualization and model understanding, and reducing excessive risk. The main idea of this technique is to evaluate the most useful subset of variables for a given learning algorithm, and the best subset of features is determined by the performance of the model. So, selecting the optimal model using FS technique has provided the possibility of saving time and reducing calculations cost, and by determining the final optimal model for each contaminant and selecting the most effective quantities in forecasting, it is possible to predict with more accuracy and less error.

Conclusion

Air pollution is a concern in many societies today, so choosing the best model for predicting pollutants in the atmosphere and using effective methods to determine the most important effective quantities is beneficial. Based on the results obtained, the final optimal model of model 2 with 6 independent variables is optimal for CO pollutant, and model 4 with 4 meteorological quantities is optimal for PM₁₀ contamination. For pollutants SO₂ model 6 with 2 independent variables and for pollutants NO₂ model 2 with 6 variables were selected as optimal and model 5 with 4 effective quantities is optimal for pollutant O₃. In general, the results of this study show that use of multi-collinearity test and forward select technique to eliminate linear relations and create a subset of effective variables is satisfactory. The results of this study showed that for the prediction of any pollutants, no need to use all seven variables from the output of the multi-collinearity test. The optimal number of independent variables for the prediction of each pollutant was obtained differently. Therefore, we can conclude that the selection of effective independent variables by FS method will reduce the analysis cost and time, as well as increase the accuracy of the pollutant predictions.

بهینه‌سازی متغیرهای هواشناسی به منظور پیش‌بینی غلظت آلاینده‌های هوا به منظور کاهش هزینه و زمان محاسبات در مدل شبکه عصبی مصنوعی

افسانه قاسمی^۱، جمیل امان‌اللهی^{۱*}، محمد دارند^۲

۱ گروه محیط‌زیست، دانشکده منابع طبیعی، دانشگاه کردستان، سنندج، ایران

۲ گروه آب و هواشناسی، دانشکده منابع طبیعی، دانشگاه کردستان، سنندج، ایران

تاریخ وصول مقاله: ۱۳۹۸/۸/۹ تاریخ پذیرش مقاله: ۱۳۹۸/۱۲/۲۱

چکیده

شهر کرمانشاه به دلیل وجود صنایع، ترافیک و توفان‌های گرد و غبار یکی از شهرهای آلوده کشور است. در این پژوهش پنج آلاینده PM_{10} ، CO ، O_3 ، NO_2 ، SO_2 با استفاده از شبکه عصبی پرسپترون چند لایه برای دو بازه زمانی امروز و فردا پیش‌بینی شدند. داده‌های مستقل شامل هفت کمیت هواشناسی دما، رطوبت نسبی، میزان دید، سرعت باد، نقطه شبنم، فشار، و بارش است. آزمون هم‌خطی و تکنیک انتخاب پیشرو برای حذف متغیرهای ورودی اضافی و ایجاد زیرمجموعه‌ای از متغیرهای اثرگذار در پیش‌بینی استفاده شد. مدل بهینه با استفاده از شاخص‌های $RMSE$ ، $NMSE$ ، R^2 ، IOA و FB برای هر آلاینده انتخاب شد. نتایج نشان می‌دهد که مدل ۲ با تعداد ۶ کمیت مستقل برای پیش‌بینی غلظت آلاینده مونوکسید کربن و دی‌اکسید نیتروژن مدلی بهینه است و برای پیش‌بینی آلاینده ازن مدل ۵ با تعداد ۳ کمیت ورودی مدل مطلوبی است همچنین برای پیش‌بینی دی‌اکسید گوگرد مدل ۶ با دو متغیر ورودی و برای پیش‌بینی ذرات معلق (PM_{10}) مدل ۴ با ۴ متغیر ورودی مناسب‌ترین مدل بوده‌اند. نتایج این پژوهش نشان می‌دهد که استفاده از تکنیک انتخاب پیشرو برای بهینه‌سازی تعداد متغیرها سبب افزایش دقت و کاهش هزینه‌های پیش‌بینی خواهد شد.

کلیدواژه‌ها: آزمون هم‌خطی، پیش‌بینی، تکنیک انتخاب پیشرو، شبکه عصبی پرسپترون چندلایه، کرمانشاه

۱. سرآغاز

توان محاسباتی زیادی هستند زیرا باید بسیاری از تعاملات بین پارامترهای مختلف مانند انتشار، هواشناسی و پوشش زمین را پردازش کنند در مقابل مدل‌های تصادفی عملکرد ساده‌تری دارند؛ زیرا سعی می‌کنند براساس تشخیص برخی از الگوها، رابطه ریاضی بین متغیرهای ورودی و خروجی را شکل دهند (Coman et al., 2008). بعضی از محققان برای مطالعه آلودگی هوا از بررسی الگوهای همدیدی

پیش‌بینی غلظت آلاینده‌های هوا در کاهش تأثیر آنها گام اساسی محسوب می‌شود. مدل‌سازی برای پیش‌بینی غلظت آلاینده‌های هوا به دو صورت قطعی و تصادفی انجام می‌شود. مدل‌های قطعی از چندین معادله استفاده می‌کنند؛ بنابراین غلظت آلاینده‌ها را در دامنه محدودی پیش‌بینی می‌کنند. این مدل‌ها علاوه بر پیچیدگی نیازمند

استفاده شده است. مطالعات متعددی درباره بهبود عملکرد مدل شبکه عصبی مصنوعی با استفاده از تکنیک‌های مختلف طبقه‌بندی برای داده‌های ورودی انجام شده است (Hrust et al., 2009; Kurt & Oktay, 2010; Cheng et al., 2012; Perez, 2012) در ادامه به چند مورد اشاره می‌شود. Cabaneros و همکاران (۲۰۱۷) از مدل پرسپترون چند لایه برای پیش‌بینی غلظت دی‌اکسید نیتروژن در کنار جاده استفاده کردند در این پژوهش داده‌های هواشناسی و آلاینده‌های جو به عنوان متغیر استفاده شد، مدل پرسپترون چند لایه با چهار تکنیک انتخاب پیشرو، حذف پسرو، آنالیز مؤلفه‌های اصلی و درخت رگرسیون ترکیب شد براساس نتایج تأثیر اکثر آلاینده‌ها مانند اکسیدهای نیتروژن در پیش‌بینی دی‌اکسید نیتروژن قابل توجه است و ترکیب مدل پرسپترون چند لایه با چهار تکنیک ذکر شده نشانگر برتری این مدل در پیش‌بینی است. Unnikrishnan و Madhu (۲۰۱۹) به منظور پیش‌بینی غلظت دی‌اکسید گوگرد و ارتباط متغیرهای هواشناسی با آلاینده‌های هوا از مدل پرسپترون چند لایه به همراه تکنیک انتخاب پیشرو و حذف پسرو برای بهینه‌سازی متغیرها استفاده کردند براساس نتایج استفاده از تکنیک بهینه‌سازی پارامترهای ورودی مدل به کاهش خطا در پیش‌بینی و عملکرد بهتر مدل شبکه عصبی می‌انجامد. جدی و همکاران (۱۳۹۶) از مدل‌های شبکه عصبی مصنوعی برای پیش‌بینی غلظت مونوکسیدکربن در تهران استفاده کردند در این پژوهش از دو روش آماری برای حذف داده‌های نویز استفاده شد نتایج حاصله علاوه بر نشان دادن قابلیت خوب شبکه عصبی در پیش‌بینی مونوکسیدکربن، بیانگر اهمیت آماده‌سازی داده‌ها برای ورود به شبکه عصبی است زیرا با توجه به آمارهای استفاده شده در آموزش شبکه و رابطه مستقیم آن با توان عصبی شبکه باید در ورود داده‌های مناسب به شبکه دقت کرد تا شبکه تعمیم‌پذیری خوبی با داده‌های جدید داشته باشد. رفیع پورگنابی و همکاران (۱۳۹۵) به منظور پیش‌بینی غلظت ازن در شهر تهران از

(سلطانی گرد فرامرز و همکاران، ۱۳۹۴)، تحلیل ماهواره (رایگانی و خیراندیش، ۱۳۹۶)، بررسی تغییرات ارتفاع و ضخامت لایه مرزی (نصیری و همکاران، ۱۳۹۵) و مدل‌سازی مکانی-زمانی (بخشی‌زاده و همکاران، ۱۳۹۴) استفاده کرده‌اند. در این میان کاربرد شبکه عصبی مصنوعی^۱ برای پیش‌بینی غلظت آلاینده‌های هوا به طور موفقیت‌آمیزی مورد استفاده قرار گرفته است. یکی از پرکاربردترین شبکه‌ها در پیش‌بینی غلظت آلاینده‌های هوا شبکه پرسپترون چند لایه^۲ است. از مزایای این شبکه دقت بالا و توانایی برقراری روابط عملکردی بین متغیرهای پیش‌بینی کننده است (Alexandrov et al., 2005). استفاده از شبکه عصبی پرسپترون چند لایه در سال‌های اخیر نتایج سودمندی از پیش‌بینی‌ها را نشان داده است. Stamenkovic و همکاران (۲۰۱۶) برای کشورهای اروپایی، ایالات متحده آمریکا، چین، ژاپن، روسیه و هند در سال‌های ۲۰۰۱ تا ۲۰۰۸ انتشار گاز اکسیدهای نیتروژن را با مدل شبکه عصبی مصنوعی بررسی کردند براساس نتایج مدل شبکه عصبی مصنوعی می‌تواند در انتشار اکسیدهای نیتروژن عملکرد خوبی داشته باشد. Alves و همکاران (۲۰۱۹) از مدل پرسپترون چند لایه برای پیش‌بینی غلظت ازن در برزیل استفاده کردند در این پژوهش غلظت آلاینده‌های هوا و کمیت‌های هواشناسی به عنوان ورودی مدل لحاظ شد و مدل عملکرد مطلوبی را نشان داد. با توجه به موارد ذکر شده شبکه پرسپترون چند لایه به عنوان مدلی غیرخطی و آماری در پیش‌بینی غلظت آلاینده‌های هوا دارای کاربردی رضایت‌بخش است؛ اما این شبکه برخلاف مدل‌های خطی در تعداد متغیرهای ورودی محدودیت دارد. روش‌های متعددی برای انتخاب بهترین متغیرهای ورودی به مدل وجود دارد که مشهورترین آن‌ها انتخاب پیشرو^۳، حذف پسرو^۴ و آنالیز مؤلفه‌های اصلی^۵ است. در این پژوهش از تکنیک انتخاب پیشرو به منظور به دست آوردن بهترین و مؤثرترین متغیرها در پیش‌بینی هر آلاینده و افزایش صحت پیش‌بینی در مدل پرسپترون چند لایه

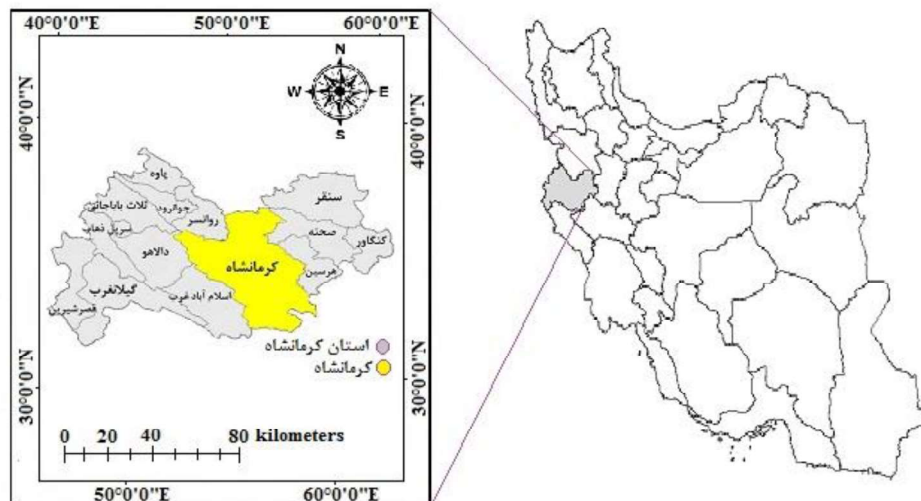
هواشناسی در پیش‌بینی غلظت آلاینده‌های هوا و ایجاد بهترین مدل از متغیرهای هواشناسی با استفاده از تکنیک انتخاب پیشرو توجه کمتری شده است. هدف اصلی در این پژوهش تعیین تأثیرگذارترین متغیرهای هواشناسی با استفاده از تکنیک انتخاب پیشرو بر غلظت آلاینده‌های هوا است به گونه‌ای که بعد از پیش‌بینی غلظت هر آلاینده با مدل پرسپترون چند لایه، برای هر آلاینده مدلی بهینه با حداکثر تأثیرگذاری متغیرهای هواشناسی و کمترین خطا و حداکثر ضریب همبستگی تهیه می‌شود که می‌تواند در کنترل غلظت آلاینده‌های هوا و مدیریت بهتر آلودگی هوا توسط مدیران شهری استفاده شود.

۲. مواد و روش‌ها

۲-۱. منطقه مورد مطالعه

استان کرمانشاه یکی از شهرهای پرجمعیت و بزرگ کشور با جمعیت ۱۹۵۲۴۳۴ نفر در آبان ماه ۹۵ است. کرمانشاه در میانه ضلع غربی کشور بین مدار جغرافیایی ۳۳ درجه و ۴۱ دقیقه تا ۳۵ درجه و ۱۷ دقیقه عرض شمالی از خط استوا و ۴۵ درجه و ۲۴ دقیقه تا ۴۸ درجه و ۶/۰ دقیقه طول شرقی از نصف النهار گرینویچ قرا گرفته است و دارای ۲۵۰۴۵/۴ کیلومترمربع گستردگی و ارتفاع ۱۲۰۰ متر از سطح دریاهاى آزاد است (سالنامه آماری استان کرمانشاه، ۱۳۹۷).

مدل پرسپترون چند لایه استفاده کردند و به‌منظور استخراج متغیرهای تأثیرگذار در پیش‌بینی از تکنیک تحلیل مؤلفه‌های اصلی بهره بردند. براساس نتایج پژوهش مدل طراحی شده میزان غلظت ازن را با دقت ۶۷ تا ۹۷ درصد در ۲۴ ساعت آینده مشخص کرد و روش تحلیل مؤلفه‌های اصلی در انتخاب و بهینه‌سازی متغیرهای ورودی دارای کارایی مناسبی است. در پژوهش‌های خارجی امکان‌سنجی در به دست آوردن مدلی مناسب و به نوعی تعیین ارتباط بین متغیرهای هواشناسی و آلاینده‌های جو با استفاده از مدل پرسپترون چند لایه و تکنیک انتخاب پیشرو در بهینه‌سازی متغیرها بسیار انجام می‌شود و بیانگر کاهش خطا و صحت بیشتر در پیش‌بینی است اما در پژوهش‌های داخلی به‌ویژه درباره شهر کرمانشاه (کارایی خوب مدل پرسپترون چند لایه در پیش‌بینی شاخص کیفیت هوای شهر کرمانشاه (رستمی فصیح و همکاران، ۱۳۹۴)، بررسی تغییرات فصلی گرد و غبار (PM_{10}) در کرمانشاه در دوره چهار ساله ۱۳۸۷-۱۳۹۰ (الماسی و همکاران، ۱۳۹۳)، مطلوبیت مدل پرسپترون چند لایه در پیش‌بینی وقوع طوفان‌های گرد و غبار در کرمانشاه (Zinatizadeh et al., 2014)، بررسی کیفیت گرد و غبار (PM_{10}) با تجزیه و تحلیل آماری در شهر کرمانشاه طی چهار سال ۲۰۰۸-۲۰۱۱ (Sharifi et al., 2015) به ترتیب اهمیت متغیرهای



شکل ۱. محدوده منطقه مورد مطالعه (سازمان مدیریت و برنامه‌ریزی استان کرمانشاه، معاونت آمار و اطلاعات)

۲-۲. داده‌ها

مشخص می‌کند واریانس ضرایب رگرسیونی برآورد شده تا چه حد بیشتر از زمانی افزایش یافته است که متغیرهای همبسته در مدل وجود ندارند. اگر مقدار این شاخص نزدیک به یک باشد هم‌خطی وجود ندارد و فاکتور تحمل نسبتی از پراکندگی نسبی یک متغیر است. اگر میزان تورم واریانس یک متغیر مستقل بیشتر از ۱۰ و میزان فاکتور تحمل آن کمتر از ۰/۲ باشد می‌توان گفت که این متغیر برای استفاده در مدل مناسب نیست (Wang et al., 2006; Dirk & Bart, 2004).

۲-۴. تکنیک انتخاب پیشرو

روش انتخاب پیشرو به طور موفقیت‌آمیزی توسط بسیاری از محققان به منظور ساخت مدل‌های پیش‌بینی قدرتمند استفاده شده است (Chen et al., 2004)؛ Khan et al., 2005؛ Eksioglu et al., 2006؛ Wang et al., 2007) این روش عمدتاً مبتنی بر رگرسیون است و براساس میزان وابستگی متغیرهای مستقل با وابسته است. بنابراین به منظور اجرای این تکنیک نخست همبستگی پیرسون بین پنج آلاینده و هفت کمیت هواشناسی و تأثیر هر یک از پنج آلاینده‌ها در روز گذشته به دلیل مؤثر بودن آلاینده‌های روز قبل در تشدید آلودگی روز بعد و تأثیر آلاینده دی‌اکسید نیتروژن فقط بر آلاینده ازن به علت نقش این آلاینده در تشکیل ازن به عنوان متغیر مستقل در دوره زمانی سه سال ۹۳، ۹۴ و ۹۵ به صورت میانگین ۲۴ ساعته از ۴ ایستگاه سنجش آلودگی فعال در شهر کرمانشاه ارزیابی شد و پس از تعیین بیشترین همبستگی هر آلاینده با متغیرهای مستقل مدل‌هایی از بیشترین به کمترین همبستگی بین متغیرها برای هر آلاینده در دو بازه امروز (مقادیر غلظت آلاینده‌های امروز) و فردا (قرار دادن مقادیر غلظت آلاینده روز پیشین) تنظیم شد. برای همه آلاینده‌ها هفت مدل و برای آلاینده ازن به دلیل بررسی آلاینده دی‌اکسید نیتروژن به عنوان متغیر مستقل، هشت مدل تنظیم شد. در نخستین

در این مطالعه از داده‌های میانگین روزانه مربوط به غلظت پنج آلاینده ذرات معلق (PM_{10})، مونوکسید کربن (CO)، ازن (O_3)، دی‌اکسید نیتروژن (NO_2)، دی‌اکسید گوگرد (SO_2) به عنوان متغیرهای وابسته و داده‌های میانگین روزانه هفت کمیت هواشناسی شامل دما، رطوبت نسبی، نقطه شبنم، سرعت باد، میزان دید، فشار و بارش به عنوان متغیرهای مستقل استفاده شد. داده‌های مربوط به غلظت آلاینده‌ها از سازمان حفاظت محیط‌زیست ایستگاه پایش آلودگی و داده‌های هواشناسی از سازمان کل هواشناسی ایستگاه سینوپتیک اصلی موجود در استان کرمانشاه اخذ شد. بازه زمانی داده‌های مورد استفاده به صورت میانگین ۲۴ ساعته در سال‌های ۹۳، ۹۴ و ۹۵ است.

۲-۳. آزمون هم‌خطی

بعد از اخذ داده‌های مذکور روزهایی که داده‌ای برای آن‌ها ثبت نشده بود از روند آنالیز حذف شد سپس داده‌های مربوط به هر آلاینده به عنوان متغیر وابسته با تأثیر هفت کمیت هواشناسی به عنوان متغیر مستقل مرتب شدند. برای آلاینده ازن علاوه بر تأثیر کمیت‌های هواشناسی تأثیر آلاینده دی‌اکسید نیتروژن و آلاینده ازن در روز گذشته نیز لحاظ شد و این امر به دلیل تأثیر بسیار این دو کمیت در تشکیل ازن تروپوسفر است. سپس آزمون هم‌خطی اجرا شد. اگر هم‌خطی در یک معادله رگرسیون بالا باشد؛ بدین معنی است که بین متغیرهای مستقل وابستگی بالایی وجود دارد و ممکن است آمار برآورد شده و ضرایب رگرسیون توانایی نشان دادن نقش منحصر به فرد و اثر هر یک از متغیرهای مستقل را نداشته باشند (Cogliani, 2001)؛ (Chen, 2008) بنابراین، آزمون هم‌خطی به منظور حذف متغیرهای ورودی اضافی، در نرم‌افزار SPSS ver16 انجام شد. رابطه خطی بین متغیرها با دو شاخص تورم واریانس و فاکتور تحمل محاسبه می‌شود. شاخص تورم واریانس

۷-۲. تعیین مدل بهینه

در پایان به‌منظور تعیین عملکرد مدل‌ها شاخص‌های ریشه میانگین مربعات خطا (RMSE)، ضریب تعیین (R^2)، میانگین نرمال مربعات خطا (NMSE)، FB و شاخص توافق (IOA) استفاده شدند. روابط (۲-۶)

$$RMSE = \sqrt{1/N \sum (P_i - O_i)^2} \quad (2)$$

$$NMSE = 1/N \sum (P_i - O_i)^2 / OP \quad (3)$$

$$R^2 = 1 - \sum (P_i - O_i)^2 / \sum (O_i - O^-)^2 \quad (4)$$

$$FB = 2(O^- - P^-) / (O^- + P^-) \quad (5)$$

$$IOA = 1 - \sum (P_i - O_i)^2 / \sum (P_i - O^- + O_i + O^-)^2 \quad (6)$$

در رابطه‌های فوق O_i داده‌های مشاهده شده؛ P_i داده‌های پیش‌بینی شده؛ O^- میانگین داده‌های مشاهده شده و P^- میانگین داده‌های پیش‌بینی شده است.

۳. نتایج

۳-۱. تحلیل آماری غلظت آلاینده‌ها و کمیت‌های

هواشناسی طی سه سال

نتایج آماری حاکی از آن است که طی سال‌های ۹۳، ۹۴ و ۹۵ متوسط غلظت ازن، گرد و غبار و مونوکسید کربن دارای روند غیریکنواخت بوده و میانگین غلظت دی‌اکسید نیتروژن در طی سه سال کاهش یافته و دی‌اکسید گوگرد نیز افزایش یافته است و متوسط دما، میزان دید، سرعت باد و بارش دارای مقادیر غیریکنواخت است نقطه شبنم دارای روند افزایشی است و رطوبت نسبی و فشار کاهش یافته است. (جدول ۱)

تصاویر نشانگر متغیر بودن غلظت آلاینده‌ها طی سال‌های ۹۳، ۹۴ و ۹۵ است. حداکثر غلظت مونوکسید کربن (۲۶/۶۴ میکروگرم بر متر مکعب) و دی‌اکسید نیتروژن (۸۹/۵۵ میکروگرم بر متر مکعب) در سال ۹۴ قابل مشاهده است. سال ۹۵ دارای بیشترین مقدار از غلظت دی‌اکسید گوگرد (۴۱/۲۵ میکروگرم بر متر مکعب) و گرد و غبار (PM_{10}) (۷۷۹/۰۷ میکروگرم بر متر مکعب) است و بیشترین غلظت از ازن (۱۸۹/۶۴ میکروگرم بر متر مکعب) در سال ۹۳ مشاهده شده است. (شکل ۲)

مدل متغیرها از حداکثر به حداقل همبستگی مرتب شدند و برای مدل‌های بعدی متغیرها با حداقل همبستگی یکی یکی برای هر مدل حذف شد تا در مدل نهایی یک متغیر با حداکثر همبستگی باقی ماند. استفاده از این تکنیک برای تعیین زیرمجموعه‌ای از متغیرهای ورودی مؤثر و حذف متغیرهای اضافی مطلوب است.

۲-۵. نرمال‌سازی داده‌ها

بعد از اجرای تکنیک انتخاب پیشرو از آنجایی که ورود داده‌های خام به مدل به کاهش سرعت و دقت در مدل منجر خواهد شد (Zhu et al., 2007) بنابراین نرمال‌سازی داده‌ها قبل از ورود داده‌ها به مدل نهایی انجام شد تا داده‌های مربوط به غلظت آلاینده‌ها و کمیت‌های هواشناسی در دامنه بین صفر و یک استاندارد شوند و میزان خطای ناشی از پیش‌بینی کاهش یابد. برای نرمال‌سازی از رابطه (۱) استفاده شد.

$$X_n = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (1)$$

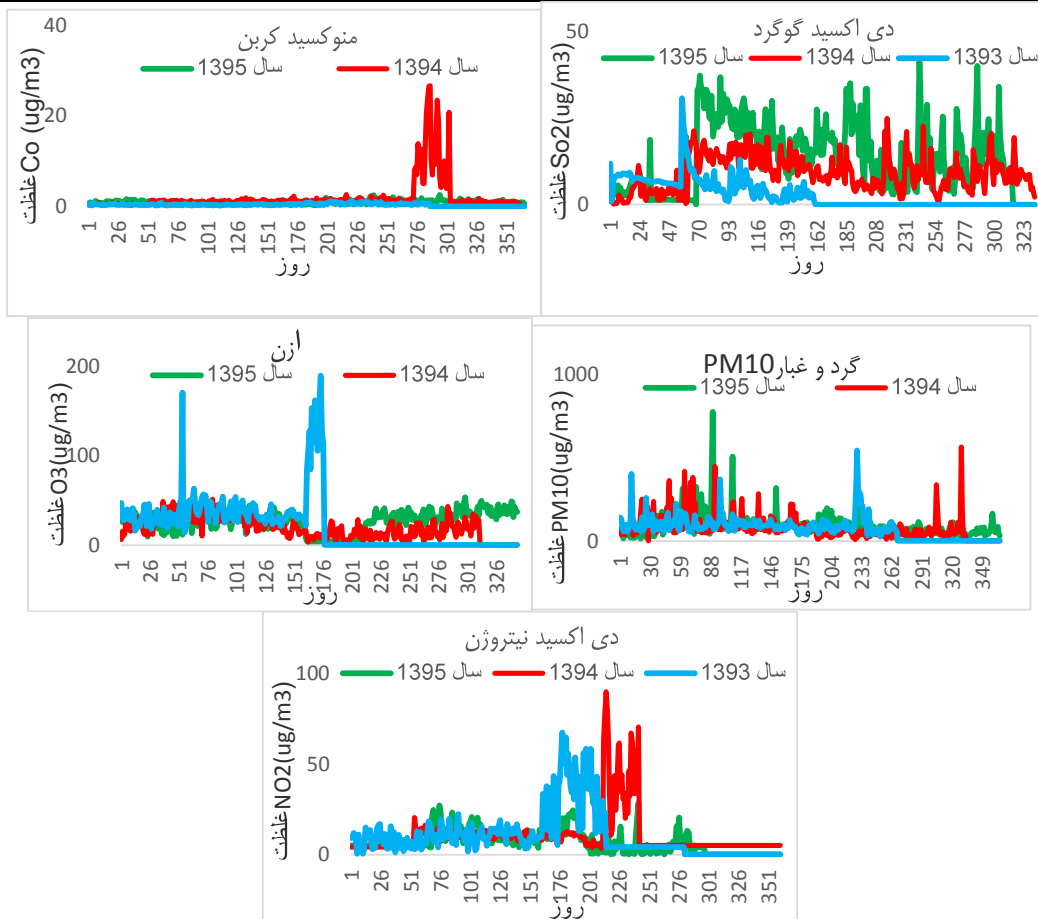
X_n داده نرمال شده؛ X داده خام؛ X_{\min} و X_{\max} حداقل و حداکثر مقدار داده است.

۲-۶. شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی می‌توانند ابزارهای انعطاف‌پذیری از رگرسیون‌های غیرخطی باشند که به‌طور کلی از یک یا چند لایه با نورون‌های متفاوت تشکیل شده‌اند (Pastor Barsenas et al., 2005). عمومی‌ترین شبکه‌ای که برای پیش‌بینی آلودگی‌های هوا کاربرد دارد، شبکه پرسپترون چند لایه است (Pastor-Barsenas et al., 2001؛ Kolehmainen et al., 2001؛ Gardner and Dorling, 1999). شبکه پرسپترون چندلایه توانایی مدل‌سازی مشکلات بسیار پیچیده و غیرخطی را از طریق توپولوژی شبکه دارا است. مدل‌های حاصل از تکنیک انتخاب پیشرو پس از نرمال‌سازی داده‌ها در شبکه پرسپترون چندلایه در نرم‌افزار Matlab اجرا شدند. ۸۰ درصد داده‌ها در مرحله آموزش و ۲۰ درصد داده‌ها در مرحله تست استفاده شد و داده‌های حاصل از پیش‌بینی غلظت پنج آلاینده و تأثیر متفاوت متغیرها در هر مدل ارزیابی شد.

جدول ۱. نتایج آماری مربوط به تغییرات غلظت آلاینده‌ها و متغیرهای هواشناسی طی سه سال

بارش	سرعت باد	میزان دید	نقطه شبنم	دما	رطوبت نسبی	فشار	O ₃	CO	PM ₁₀	NO ₂	SO ₂	
۰	۰	۰	-۳۶/۶۰	-۱۳/۶۰	۱	۹۹۳/۸۰	۱۴/۲۸	۰/۱۹	۲۵/۹۹	۰/۵۷	۰/۵۵	حداقل
۲۱	۱۲	۲۰۰۰۰	۱۸/۲۰	۴۱/۶۰	۹۹	۱۰۳۲/۷۰	۱۸۹/۶۴	۱/۲۲	۵۴۳/۹۸	۶۷/۲۹	۳۰/۷۱	حداکثر
۰/۲۴	۲/۸۱	۹۷۲۲	-۰/۷۸	۱۵/۸۲	۴۳/۴۳	۱۰۱۳/۰۷	۴۲/۸۹	۰/۴۹	۹۹/۳۵	۱۳/۱۸	۵/۹۵	میانگین
۱/۲۵	۲/۱۳	۱۷۶۱	۵/۴۹	۱۱/۳۴	۲۷/۷۷	۹/۳۸	۳/۰۷	۰/۱۹	۵/۶۵	۱/۳۳	۳/۹۶	انحراف معیار
۰	۰	۰	-۳۶/۷۰	-۷/۶۰	۱	۹۹۹/۱۰	۳/۱۸	۰/۲۵	۸/۶۴	۳/۵۸	۰/۲۰	حداقل
۳۵	۱۳	۴۰۰۰۰	۱۴/۳۰	۴۱/۶۰	۹۷	۱۰۳۵	۵۱/۶۵	۲۶/۶۴	۵۶۶/۴۵	۸۹/۵۵	۲۴/۹۲	حداکثر
۰/۲۴	۳/۰۴	۹۶۳۴	-۳/۰۶	۱۶/۷۰	۳۷/۱۴	۱۰۱۲/۴۶	۲۱/۷۱	۱/۹۷	۸۳/۹۶	۱۰/۵۷	۸/۹۸	میانگین
۱/۷۸	۲/۳۱	۱۸۵۱	۶/۳۵	۱۱/۳۵	۲۶/۸۹	۹/۰۹	۱/۱۹	۳/۶۵	۶/۹۴	۱/۲۳	۵/۰۶	انحراف معیار
۰	۰	۰	-۵۸/۱۵	-۸/۳۰	۱	۹۹۲/۸۰	۳/۹۷	۰/۲۲	۱۴/۶۳	۰/۰۴	۰/۰۷	حداقل
۲۴	۵۰	۲۰۰۰۰	۱۳/۶۰	۴۵/۷۰	۱۰۰	۱۰۳۲/۴۰	۵۳/۷۱	۲/۵۹	۷۷۹/۰۷	۲۸/۲۲	۴۱/۲۵	حداکثر
۱/۸۷	۲/۸۹	۹۷۷۵	-۳/۱۴	۱۶/۰۲	۳۷/۰۳	۱۰۱۲/۴۴	۲۶/۹۲	۰/۸۴	۹۰/۹۱	۸/۳۶	۱۴/۱۵	میانگین
۰/۲۲	۲/۳۸	۱۴۱۷	۷/۰۳	۱۰/۹۰	۲۵/۵۵	۸/۹۰	۱/۲۱	۰/۳۹	۶/۸۷	۶/۰۳	۹/۶۰	انحراف معیار



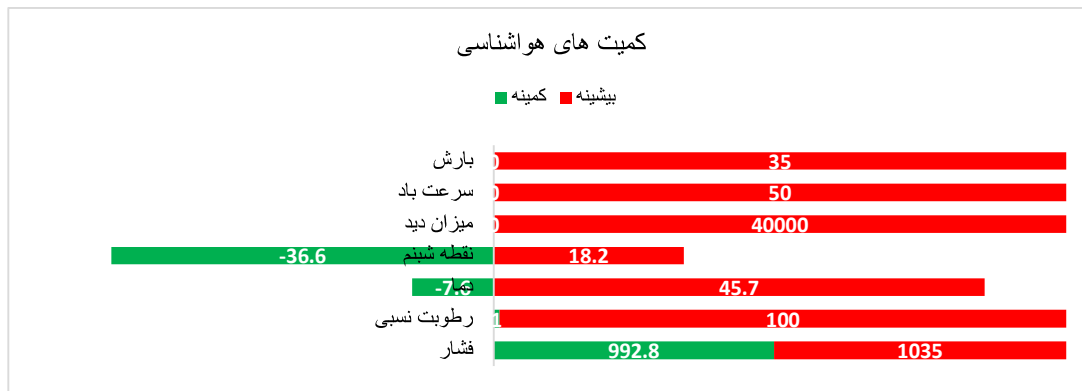
شکل ۲. تغییرات غلظت آلاینده‌های مونوکسید کربن، ازن، دی‌اکسید نیتروژن، دی‌اکسید گوگرد و غبار طی سه سال

رابطه خطی هستند. این مقادیر حد استاندارد در نظر گرفته شده است. نتایج حاصل از این آزمون در جدول (۱) نشان داده شده است. بر طبق نتایج حاصل از جدول مقادیر حاصل شده از فاکتور تورم واریانس (VIF) و تحمل (Tolerance) برای سه کمیت دما، نقطه شبنم و رطوبت نسبی نسبت به دیگر کمیت‌ها خارج از حد استاندارد است. کمیت رطوبت نسبی نسبت به دو کمیت دیگر دارای حداکثر رابطه خطی است. میزان تورم واریانس بیشتر از ۱۰ و فاکتور تحمل کمتر از ۰/۲ این موضوع را به وضوح نشان می‌دهد. با حذف پارامتر رطوبت نسبی و اجرای مجدد آزمون هم‌خطی تورم واریانس و فاکتور تحمل تمامی پارامترها در دامنه استاندارد قرار گرفت.

مقادیر متفاوتی از کمیت‌های هواشناسی طی سه سال قابل مشاهده است. از میان مقادیر متغیر کمیت‌های هواشناسی مربوط به سه سال ۹۳، ۹۴ و ۹۵ حداقل مقدار برای رطوبت نسبی ۱ برای دما ۷/۶۰- برای فشار ۹۹۲/۸۰ برای نقطه شبنم ۳۶/۶۰- و برای پارامترهای سرعت باد، میزان دید و بارش صفر است. (شکل ۳)

۳-۲. نتایج آزمون هم‌خطی

آزمون هم‌خطی در نرم‌افزار SPSS Ver16 به‌منظور تعیین رابطه خطی بین آلاینده‌ها و کمیت‌های هواشناسی طی سال‌های ۹۳، ۹۴ و ۹۵ انجام شد. در اجرای این آزمون متغیرهایی که دارای فاکتور تورم واریانس کمتر از ۱۰ و تحمل بیشتر از ۰/۲ باشند نشان دهنده مقادیر مطلوبی از



شکل ۳. مقادیر کمینه و بیشینه کمیت‌های هواشناسی طی سه سال

جدول ۲. مقادیر تورم واریانس و آماره تحمل برای کمیت‌های مستقل در پیش‌بینی آلاینده‌های هوای شهر کرمانشاه طی سه سال

آزمون هم‌خطی			آزمون هم‌خطی بعد از حذف متغیر همبسته		
آلاینده‌ها	VIF	Tolerance	VIF	Tolerance	
دما	۱۰/۵۷۶	۰/۰۹۵	۰/۰۹۶	۰/۹۱۳	دما
نقطه شبنم	۴/۸۳۳	۰/۲۰۷	۱/۱۲۷	۰/۸۸۸	نقطه شبنم
بارش	۱/۲۵۳	۰/۷۹۸	۱/۱۶۴	۰/۸۵۹	بارش
فشار	۱/۰۰۲	۰/۹۹۸	۱/۰۰۲	۰/۹۹۸	فشار
رطوبت نسبی	۱۵/۱۳۶	۰/۰۶۶	۱/۱۴۳	۰/۸۷۵	رطوبت نسبی
میزان دید	۱/۱۸۰	۰/۸۴۸	۱/۱۳۶	۰/۸۸۰	میزان دید
سرعت باد	۱/۱۶۷	۰/۸۵۷			سرعت باد

ادامه جدول ۲. مقادیر تورم واریناس و آماره تحمل برای کمیت‌های مستقل در پیش‌بینی آلاینده‌های هوای شهر کرمانشاه طی سه سال

آزمون هم‌خطی			آزمون هم‌خطی بعد از حذف متغیر همبسته		
آلاینده‌ها	VIF	Tolerance	VIF	Tolerance	
ذرات معلق (PM ₁₀)	دما	۱۰/۹۴۴	۰/۰۹۱	دما	۱/۲۷۴
	نقطه شبنم	۵/۳۳۳	۰/۱۸۸	نقطه شبنم	۱/۱۲۶
	بارش	۱/۲۶۲	۰/۷۹۲	بارش	۱/۱۶۲
	فشار	۱/۳۰۱	۰/۷۶۹	فشار	۱/۲۰۰
	رطوبت نسبی	۱۶/۷۲۱	۰/۰۶۰	میزان دید	۱/۱۶۹
	میزان دید	۱/۲۳۴	۰/۸۱۰	سرعت باد	۱/۱۲۹
دی اکسید گوگرد	سرعت باد	۱/۱۶۴	۰/۸۵۹	سرعت باد	۱/۱۲۹
	دما	۱۰/۹۵۳	۰/۰۹۱	دما	۱/۰۹۹
	نقطه شبنم	۴/۹۱۶	۰/۲۰۳	نقطه شبنم	۱/۱۴۶
	بارش	۱/۲۲۷	۰/۸۱۵	بارش	۱/۱۵۸
	فشار	۱/۰۰۲	۰/۹۹۸	فشار	۱/۰۰۲
	رطوبت نسبی	۱۵/۶۶۸	۰/۰۶۴	میزان دید	۱/۱۳۷
دی اکسید نیتروژن	میزان دید	۱/۱۸۲	۰/۸۴۶	سرعت باد	۱/۱۰۹
	سرعت باد	۱/۱۴۰	۰/۸۷۷	سرعت باد	۱/۱۰۹
	دما	۱۰/۵۰۹	۰/۰۹۵	دما	۱/۱۰۴
	نقطه شبنم	۴/۶۶۷	۰/۲۱۴	نقطه شبنم	۱/۱۱۸
	بارش	۱/۲۱۴	۰/۸۲۴	بارش	۱/۱۳۶
	فشار	۱/۰۰۲	۰/۹۹۸	فشار	۱/۰۰۲
ازن	رطوبت نسبی	۱۴/۸۱۲	۰/۰۶۸	میزان دید	۱/۱۳۰
	میزان دید	۱/۱۶۳	۰/۸۵۹	سرعت باد	۱/۱۲۴
	سرعت باد	۱/۱۵۵	۰/۸۶۶	سرعت باد	۱/۱۲۴
	دما	۱۱/۴۳۵	۰/۰۸۷	دما	۱/۲۰۷
	نقطه شبنم	۵/۱۳۲	۰/۱۹۵	نقطه شبنم	۱/۱۳۵
	بارش	۱/۲۴۳	۰/۸۰۵	بارش	۱/۱۵۳
ازن روز گذشته	فشار	۱/۰۰۲	۰/۹۹۸	فشار	۱/۰۰۲
	رطوبت نسبی	۱۵/۶۸۹	۰/۰۶۴	میزان دید	۱/۱۴۶
	میزان دید	۱/۱۷۲	۰/۸۵۴	سرعت باد	۱/۱۰۹
	سرعت باد	۱/۱۲۶	۰/۸۸۸	سرعت باد	۱/۵۶۹
	دی اکسید نیتروژن	۱/۵۶۹	۰/۶۳۷	دی اکسید نیتروژن	۱/۵۶۹
	ازن روز گذشته	۱/۵۹۳	۰/۶۲۸	ازن روز گذشته	۱/۵۳۲

۳-۳. نتایج همبستگی بین متغیرهای مستقل با آلاینده‌های هوای شهر کرمانشاه

به منظور تعیین همبستگی بین متغیرها همبستگی پیرسون در دو بازه امروز و فردا طی سه سال ۹۳، ۹۴ و ۹۵ به منظور تعیین تأثیرگذاری غلظت آلاینده روز گذشته در میزان تشدید غلظت آلاینده فردا و وابستگی آلاینده‌ها به کمیت‌های هواشناسی اجرا شد که نشان دهنده تأثیر مستقیم

یا عکس یک کمیت هواشناسی در کاهش یا ازدیاد آلاینده است. به دلیل نقش آلاینده دی اکسید نیتروژن در تشکیل ازن، همبستگی ازن با این آلاینده نیز بررسی شد. شایان ذکر است علاوه بر موارد فوق به منظور اجرای تکنیک انتخاب پیشرو میزان همبستگی بین متغیرها محاسبه شده است. نتایج همبستگی بین متغیرهای مستقل در پیش‌بینی امروز و فردای آلاینده‌های هوای کرمانشاه در جدول ۲

مدل و برای آلاینده‌ ازن به دلیل بررسی پارامتر دی‌اکسید نیتروژن به عنوان متغیر مستقل، هشت مدل تنظیم شد مدل‌ها از بیشترین به کمترین همبستگی بین متغیرهای مستقل با وابسته و حذف حداقل همبستگی در هر مدل مرتب شدند جدول (۳). سر انجام هر کدام از مدل‌های تنظیم شده به عنوان ورودی وارد مدل پرسپترون چند لایه شد و نتایج مدل‌ها بر اساس آماره‌های توضیح داده شده ارزیابی می‌شود.

نشان داده شده است. براساس نتایج پنج آلاینده مذکور با متغیر آلاینده روز گذشته دارای همبستگی بالایی هستند ولی حداقل همبستگی هر آلاینده با کمیت‌های هواشناسی مختلف، متفاوت است.

۳-۴. نتایج حاصل از تکنیک انتخاب پیشرو در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه
با استفاده از همبستگی پیرسون برای هر آلاینده هفت

جدول ۳. نتایج حاصل از همبستگی متغیرهای مستقل با آلاینده‌های هوای شهر کرمانشاه در بازه‌ امروز و فردا طی سه سال

دی‌اکسید نیتروژن	آلاینده‌ روز گذشته	سرعت باد	میزان دید	فشار	بارش	نقطه شبنم	دما		
	۰/۸۳۱	-۰/۰۱۷	۰/۰۱۲	-۰/۰۰۱	۰/۰۵۰	۰/۰۰۵	۰/۰۸۷	مونوکسیدکربن	امروز
	۰/۵۰۴	۰/۰۴۶	-۰/۰۳۲	-۰/۰۵۷	۰/۰۴۱	۰/۱۲۱	-۰/۰۴۳	ذرات معلق (PM ₁₀)	
	۰/۸۱۴	-۰/۰۰۹	-۰/۰۹۵	۰/۰۰۰	-۰/۰۱۲	۰/۰۳۲	-۰/۰۵۴	دی‌اکسیدگوگرد	
	۰/۸۲۲	۰/۰۹۵	۰/۰۳۸	-۰/۰۰۳	-۰/۰۴۱	۰/۰۱۶	۰/۳۵۸	دی‌اکسید نیتروژن	
۰/۵۲۴	۰/۸۳۰	-۰/۰۱۷	۰/۰۷۲	-۰/۰۱۰	-۰/۰۴۹	-۰/۰۰۱	-۰/۱۳۶	ازن	
	۰/۸۳۱	-۰/۰۲۰	۰/۰۰۹	۰/۰۰۸	۰/۰۸۴	۰/۰۰۰	۰/۰۹۱	مونوکسیدکربن	فردا ۱
	۰/۵۰۴	۰/۰۵۶	-۰/۰۵۸	-۰/۰۳۴	۰/۰۴۷	۰/۱۳۷	-۰/۰۴۵	ذرات معلق (PM ₁₀)	
	۰/۸۱۴	-۰/۰۰۳	-۰/۰۸۵	۰/۰۰۱	۰/۰۰۰	۰/۰۲۳	-۰/۰۵۳	دی‌اکسیدگوگرد	
	۰/۸۲۲	۰/۱۰۱	۰/۰۴۲	-۰/۰۰۲	-۰/۰۵۷	۰/۰۱۵	۰/۳۶۲	دی‌اکسید نیتروژن	
۰/۵۳۸	۰/۸۱۷	-۰/۰۱۴	۰/۰۵۰	۰/۰۱۳	-۰/۰۰۸	۰/۰۰۸	-۰/۱۴۲	ازن	

جدول ۴. نتایج حاصل از تکنیک انتخاب پیشرو به منظور انتخاب متغیرهای مؤثر در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه

آلاینده‌ها	مدل	کمیت‌های ورودی
	مدل ۱	DP, P, Visi, WS, Preci, T, P-CO
	مدل ۲	P, Visi, WS, Preci, T, P-CO
	مدل ۳	Visi, WS, Preci, T, P-CO
مونوکسید کربن	مدل ۴	WS, Preci, T, P-CO
	مدل ۵	Preci, T, P-CO
	مدل ۶	T, P-CO
	مدل ۷	P-CO

ادامه جدول ۴. نتایج حاصل از تکنیک انتخاب پیشرو به منظور انتخاب متغیرهای مؤثر در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه

آلاینده‌ها	مدل	کمیت‌های ورودی
ذرات معلق (PM10)	مدل ۱	P, T, Preci, WS, Visi, DP, P-PM10
	مدل ۲	T, Preci, WS, Visi, DP, P-PM10
	مدل ۳	Preci, WS, Visi, DP, P-PM10
	مدل ۴	WS, Visi, DP, P-PM10
	مدل ۵	Visi, DP, P-PM10
	مدل ۶	DP, P-PM10
	مدل ۷	P-PM10
دی‌اکسید گوگرد	مدل ۱	Preci, P, WS, DP, T, Visi, P-SO2
	مدل ۲	P, WS, DP, T, Visi, P-SO2
	مدل ۳	WS, DP, T, Visi, P-SO2
	مدل ۴	DP, T, Visi, P-SO2
	مدل ۵	T, Visi, P-SO2
	مدل ۶	Visi, P-SO2
	مدل ۷	P-SO2
دی‌اکسید نیتروژن	مدل ۱	P, DP, Visi, Preci, WS, T, P-NO2
	مدل ۲	DP, Visi, Preci, WS, T, P-NO2
	مدل ۳	Visi, Preci, WS, T, P-NO2
	مدل ۴	Preci, WS, T, P-NO2
	مدل ۵	WS, T, P-NO2
	مدل ۶	T, P-NO2
	مدل ۷	P-NO2
ازن	مدل ۱	Preci, DP, P, WS, Visi, T, NO2, P-O3
	مدل ۲	DP, P, WS, Visi, T, NO2, P-O3
	مدل ۳	P, WS, Visi, T, NO2, P-O3
	مدل ۴	WS, Visi, T, NO2, P-O3
	مدل ۵	Visi, T, NO2, P-O3
	مدل ۶	T, NO2, P-O3
	مدل ۷	NO2, P-O3
	مدل ۸	P-O3

پیش‌بینی امروز پنج آلاینده با مدل پرسپترون چند لایه طی سه سال ۹۳، ۹۴ و ۹۵ در جدول (۴) نشان داده شده است. حداقل و حداکثر شاخص‌ها برای هر آلاینده در مدل‌های

۳-۵. نتایج آماره‌های RMSE, NMSE, FB, IOA, R2 در پیش‌بینی امروز آلاینده‌های هوای شهر کرمانشاه نتایج آماره‌های RMSE, NMSE, FB, IOA, R2 در

(PM₁₀) بیشترین مقدار شاخص IOA مربوط به مدل ۴ و ۵ با IOA= ۰/۹۶۰ و کمترین مقدار این شاخص مربوط به مدل‌های ۶ و ۷ با IOA= ۰/۹۳۱ است.

مختلف، متفاوت است مثلاً برای آلاینده مونوکسید کربن مدل ۲ دارای کمترین میزان شاخص FB با مقدار ۰/۰۰۲- FB= و مدل ۶ دارای بیشترین میزان این شاخص با مقدار ۰/۱۰۳- FB= است. برای آلاینده ذرات معلق

جدول ۵. نتایج پنج آماره RMSE, FB, IOA, R2, RMSE در پیش‌بینی امروز آلاینده‌های هوای شهر کرمانشاه طی سه سال

آلاینده	مدل	RMSE آموزش	RMSE تست	NMSE	FB	IOA	R ²
مونوکسید کربن	مدل ۱	-۰/۰۶۹	-۰/۰۱۱	-۰/۱۰۰	-۰/۰۱۷	-۰/۹۶۷	-۰/۷۳۴
	مدل ۲	-۰/۰۷۸	-۰/۰۱۴	-۰/۲۰۰	-۰/۰۰۲	-۰/۹۵۰	-۰/۵۴۴
	مدل ۳	-۰/۰۸۳	-۰/۰۱۳	-۰/۲۱۴	-۰/۰۵۷	-۰/۹۴۲	-۰/۵۲۰
	مدل ۴	-۰/۰۷۷	-۰/۰۱۱	-۰/۱۰۸	-۰/۰۴۶	-۰/۹۷۳	-۰/۶۷۲
	مدل ۵	-۰/۰۷۲	-۰/۰۱۷	-۰/۲۷۱	-۰/۰۱۲	-۰/۹۴۲	-۰/۵۶۵
	مدل ۶	-۰/۰۸۴	-۰/۰۱۵	-۰/۲۵۵	-۰/۱۰۳	-۰/۹۳۵	-۰/۲۳۴
	مدل ۷	-۰/۰۸۶	-۰/۰۹۹	-۰/۸۹۹	-۰/۰۴۰	-۰/۸۵۱	-۰/۰۲۲
ذرات معلق (PM ₁₀)	مدل ۱	-۰/۰۸۹	-۰/۰۴۳	-۰/۱۸۱	-۰/۰۷۸	-۰/۹۴۹	-۰/۲۲۳
	مدل ۲	-۰/۰۸۹	-۰/۰۴۸	-۰/۲۴۸	-۰/۰۸۶	-۰/۹۳۷	-۰/۴۴۰
	مدل ۳	-۰/۰۸۰	-۰/۰۴۰	-۰/۱۴۴	-۰/۰۶۱	-۰/۹۵۱	-۰/۳۰۹
	مدل ۴	-۰/۰۷۷	-۰/۰۴۰	-۰/۱۳۴	-۰/۰۰۱	-۰/۹۶۰	-۰/۴۷۲
	مدل ۵	-۰/۰۷۹	-۰/۰۴۷	-۰/۲۰۷	-۰/۰۰۶	-۰/۹۶۰	-۰/۳۵۷
	مدل ۶	-۰/۰۸۶	-۰/۰۵۱	-۰/۲۴۶	-۰/۰۱۶	-۰/۹۳۱	-۰/۲۸۱
	مدل ۷	-۰/۰۸۴	-۰/۰۵۴	-۰/۳۰۳	-۰/۰۴۲	-۰/۹۳۱	-۰/۴۵۸
دی‌اکسید گوگرد	مدل ۱	-۰/۰۹۱	-۰/۰۱۱	-۰/۲۲۶	-۰/۰۱۳	-۰/۹۲۸	-۰/۰۷۶
	مدل ۲	-۰/۰۸۴	-۰/۰۱۳	-۰/۲۱۸	-۰/۰۱۰	-۰/۹۰۷	-۰/۰۷۷
	مدل ۳	-۰/۰۷۸	-۰/۰۱۱	-۰/۱۳۵	-۰/۰۰۳	-۰/۹۴۳	-۰/۶۱۲
	مدل ۴	-۰/۰۹۳	-۰/۰۱۲	-۰/۲۰۵	-۰/۰۳۸	-۰/۹۴۵	-۰/۱۱۴
	مدل ۵	-۰/۰۷۵	-۰/۰۱۴	-۰/۱۸۰	-۰/۰۸۶	-۰/۹۱۸	-۰/۱۴۹
	مدل ۶	-۰/۰۸۵	-۰/۰۱۱	-۰/۱۳۸	-۰/۰۳۸	-۰/۹۴۸	-۰/۱۰۸
	مدل ۷	-۰/۰۸۵	-۰/۰۱۳	-۰/۲۶۰	-۰/۰۳۸	-۰/۹۰۶	-۰/۴۱۲
دی‌اکسید نیتروژن	مدل ۱	-۰/۱۱۰	-۰/۰۹۹	-۰/۷۷۰	-۰/۰۴۵	-۰/۸۴۹	-۰/۵۳۰
	مدل ۲	-۰/۱۰۵	-۰/۰۶۲	-۰/۳۰۰	-۰/۰۲۷	-۰/۹۳۰	-۰/۵۷۳
	مدل ۳	-۰/۱۲۲	-۰/۱۰۲	-۰/۸۹۱	-۰/۰۴۴	-۰/۸۴۸	-۰/۵۷۷
	مدل ۴	-۰/۱۱۵	-۰/۰۸۴	-۰/۷۱۸	-۰/۱۱۹	-۰/۸۸۳	-۰/۴۸۹
	مدل ۵	-۰/۱۱۶	-۰/۰۸۵	-۰/۵۴۶	-۰/۰۵۳	-۰/۸۹۶	-۰/۴۰۰
	مدل ۶	-۰/۱۰۳	-۰/۰۹۹	-۰/۷۶۱	-۰/۰۵۰	-۰/۸۷۵	-۰/۳۶۳
	مدل ۷	-۰/۱۲۱	-۰/۰۹۳	-۰/۷۷۷	-۰/۲۴۶	-۰/۸۶۴	-۰/۳۷۴

ادامه جدول ۵. نتایج پنج آماره RMSE, FB, IOA, R2 در پیش‌بینی امروز آلاینده‌های هوای شهر کرمانشاه طی سه سال

آلاینده	مدل	RMSE آموزش	RMSE تست	NMSE	FB	IOA	R ²
ازن	مدل ۱	۰/۰۴۸	۰/۰۴۱	۰/۱۴۰	۰/۰۰۵	۰/۹۷۳	۰/۶۹۲
	مدل ۲	۰/۰۴۶	۰/۰۴۰	۰/۱۱۰	۰/۰۶۰	۰/۹۷۷	۰/۶۰۱
	مدل ۳	۰/۰۴۸	۰/۰۴۳	۰/۱۴۰	۰/۰۵۸	۰/۹۷۱	۰/۵۹۶
	مدل ۴	۰/۰۴۶	۰/۰۴۲	۰/۱۰۶	۰/۰۲۲	۰/۹۷۷	۰/۵۸۰
	مدل ۵	۰/۰۴۹	۰/۰۴۰	۰/۰۹۶	۰/۰۱۹	۰/۹۷۹	۰/۶۱۰
	مدل ۶	۰/۰۵۰	۰/۰۵۰	۰/۱۷۴	۰/۰۰۴	۰/۹۶۶	۰/۷۱۰
	مدل ۷	۰/۰۴۸	۰/۰۳۵	۰/۰۹۱	۰/۰۱۸	۰/۹۸۱	۰/۷۱۱
	مدل ۸	۰/۰۵۰	۰/۰۳۷	۰/۰۹۹	۰/۰۰۸	۰/۹۷۹	۰/۷۱۰

برای آلاینده ازن حداکثر مقدار RMSE در مرحله آزمون در مدل ۸ با $RMSE=0/038$ و حداقل مقدار این شاخص در مدل ۷ با $RMSE=0/034$ مشاهده شد. درباره آلاینده دی‌اکسید گوگرد مدل ۶ با $NMSE=0/105$ کمترین مقدار و مدل ۴ با $NMSE=0/561$ بیشترین مقدار این شاخص را به خود اختصاص می‌دهند.

۳-۶. نتایج آماره‌های RMSE, NMSE, FB, IOA, R2 در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه

نتایج حاصل از آماره‌های RMSE, NMSE, FB, IOA, R2 در پیش‌بینی فردا پنج آلاینده با مدل پرسپترون چند لایه در سه سال ۹۳، ۹۴ و ۹۵ در جدول (۵) نشان داده شده است. برای آلاینده دی‌اکسید نیتروژن حداکثر $R^2=0/573$

جدول ۶. نتایج پنج آماره RMSE, FB, IOA, R2 در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه طی سه سال

آلاینده	مدل	RMSE آموزش	RMSE تست	NMSE	FB	IOA	R2
مونوکسید کربن	مدل ۱	۰/۰۷۸	۰/۰۱۱	۰/۱۳۵	۰/۰۰۶	۰/۹۵۰	۰/۵۲۶
	مدل ۲	۰/۰۱۰	۰/۰۱۱	۰/۱۵۱	۰/۰۰۲	۰/۹۸۰	۰/۸۲۸
	مدل ۳	۰/۰۷۰	۰/۰۱۴	۰/۱۶۷	۰/۰۱۸	۰/۹۳۱	۰/۵۲۱
	مدل ۴	۰/۰۸۰	۰/۰۱۳	۰/۱۶۱	۰/۰۱۳	۰/۹۵۵	۰/۵۸۳
	مدل ۵	۰/۰۷۶	۰/۰۱۳	۰/۱۵۷	۰/۰۲۳	۰/۹۶۸	۰/۷۲۵
	مدل ۶	۰/۰۷۴	۰/۰۱۳	۰/۱۶۴	۰/۰۲۳	۰/۹۱۰	۰/۴۹۰
	مدل ۷	۰/۰۸۶	۰/۰۹۹	۰/۱۸۱	۰/۰۴۰	۰/۸۵۱	۰/۰۲۲
ذرات معلق (PM10)	مدل ۱	۰/۰۸۷	۰/۰۸۵	۰/۷۲۴	۰/۱۴۴	۰/۸۵۳	۰/۳۰۰
	مدل ۲	۰/۰۸۹	۰/۰۷۷	۰/۶۳۱	۰/۰۳۱	۰/۸۶۲	۰/۰۹۲
	مدل ۳	۰/۰۷۸	۰/۰۹۰	۰/۶۸۴	۰/۰۷۵	۰/۸۵۰	۰/۲۵۳
	مدل ۴	۰/۰۸۸	۰/۰۷۲	۰/۴۸۷	۰/۰۸۲	۰/۸۹۳	۰/۴۶۵
	مدل ۵	۰/۰۸۸	۰/۰۸۵	۰/۷۲۰	۰/۰۲۹	۰/۸۵۶	۰/۲۷۲
	مدل ۶	۰/۰۷۸	۰/۰۸۶	۰/۶۵۹	۰/۱۳۹	۰/۸۵۳	۰/۳۵۰
	مدل ۷	۰/۰۸۶	۰/۰۸۰	۰/۶۰۴	۰/۱۶۶	۰/۸۷۳	۰/۴۵۳
ذرات معلق (PM10)	مدل ۱	۰/۰۸۷	۰/۰۸۵	۰/۷۲۴	۰/۱۴۴	۰/۸۵۳	۰/۳۰۰
	مدل ۲	۰/۰۸۹	۰/۰۷۷	۰/۶۳۱	۰/۰۳۱	۰/۸۶۲	۰/۰۹۲
	مدل ۳	۰/۰۷۸	۰/۰۹۰	۰/۶۸۴	۰/۰۷۵	۰/۸۵۰	۰/۲۵۳
	مدل ۴	۰/۰۸۸	۰/۰۷۲	۰/۴۸۷	۰/۰۸۲	۰/۸۹۳	۰/۴۶۵
	مدل ۵	۰/۰۸۸	۰/۰۸۵	۰/۷۲۰	۰/۰۲۹	۰/۸۵۶	۰/۲۷۲

ادامه جدول ۶. نتایج پنج آماره RMSE, FB, IOA, R2 در پیش‌بینی فردا آلاینده‌های هوای شهر کرمانشاه طی سه سال

آلاینده	مدل	RMSE آموزش	RMSE تست	NMSE	FB	IOA	R2
ذرات معلق (PM10)	مدل ۶	۰/۰۷۸	۰/۰۸۶	۰/۶۵۹	۰/۱۳۹	۰/۸۵۳	۰/۳۵۰
	مدل ۷	۰/۰۸۶	۰/۰۸۰	۰/۶۰۴	۰/۱۶۶	۰/۸۷۳	۰/۴۵۳
دی‌اکسید گوگرد	مدل ۱	۰/۰۸۷	۰/۰۱۰	۰/۱۳۵	۰/۰۵۰	۰/۹۵۹	۰/۱۹۲
	مدل ۲	۰/۰۹۰	۰/۰۱۰	۰/۱۶۳	-۰/۰۳۰	۰/۹۶۰	۰/۰۴۹
	مدل ۳	۰/۰۹۲	۰/۰۱۰	۰/۲۰۵	-۰/۰۱۸	۰/۹۵۱	۰/۶۲۲
	مدل ۴	۰/۰۹۳	۰/۰۱۷	۰/۵۶۱	-۰/۰۸۷	۰/۸۷۴	۰/۳۲۴
	مدل ۵	۰/۰۹۳	۰/۰۰۹	۰/۱۲۱	-۰/۰۲۰	۰/۹۶۸	۰/۵۵۵
	مدل ۶	۰/۰۹۳	۰/۰۰۸	۰/۱۰۵	-۰/۰۰۵	۰/۹۷۲	۰/۶۴۵
	مدل ۷	۰/۰۹۶	۰/۰۱۳	۰/۳۴۲	-۰/۰۵۸	۰/۹۳۶	۰/۳۲۰
دی‌اکسید نیتروژن	مدل ۱	۰/۱۰۸	۰/۰۸۵	۰/۵۴۶	۰/۰۴۰	۰/۹۱۳	۰/۵۵۵
	مدل ۲	۰/۱۰۴	۰/۰۹۳	۰/۷۰۲	۰/۰۷۵	۰/۸۷۵	۰/۴۸۲
	مدل ۳	۰/۱۱۷	۰/۰۶۴	۰/۳۲۶	-۰/۰۱۰۲	۰/۹۶۸	۰/۵۷۳
	مدل ۴	۰/۱۱۵	۰/۱۰۷	۰/۹۹۳	۰/۱۲۹	۰/۸۴۷	۰/۳۸۰
	مدل ۵	۰/۱۰۷	۰/۰۹۷	۰/۷۳۲	۰/۰۹۷	۰/۸۷۳	۰/۳۷۷
	مدل ۶	۰/۱۱۴	۰/۰۹۶	۰/۷۵۱	۰/۱۲۳	۰/۸۷۰	۰/۲۶۳
	مدل ۷	۰/۱۲۱	۰/۰۶۳	۰/۳۵۵	۰/۱۵۸	۰/۹۲۷	۰/۳۷۴
ازن	مدل ۱	۰/۰۴۷	۰/۰۳۸	۰/۱۲۱	-۰/۰۳۱	۰/۹۷۶	۰/۶۵۰
	مدل ۲	۰/۰۴۷	۰/۰۳۷	۰/۰۸۸	-۰/۰۲۰	۰/۹۸۲	۰/۴۶۴
	مدل ۳	۰/۰۴۹	۰/۰۳۷	۰/۱۰۱	-۰/۰۲۸	۰/۹۷۹	۰/۶۶۲
	مدل ۴	۰/۰۴۶	۰/۰۳۵	۰/۰۸۲	-۰/۰۴۰	۰/۹۸۲	۰/۶۷۹
	مدل ۵	۰/۰۴۶	۰/۰۳۵	۰/۰۰۱	۰/۰۰۱	۰/۹۸۴	۰/۶۹۴
	مدل ۶	۰/۰۵۰	۰/۰۳۵	۰/۰۸۱	-۰/۰۶۶	۰/۹۸۳	۰/۴۸۸
	مدل ۷	۰/۰۴۹	۰/۰۳۴	۰/۰۸۴	-۰/۰۱۸	۰/۹۸۲	۰/۶۶۷
	مدل ۸	۰/۰۴۸	۰/۰۳۸	۰/۱۰۵	۰/۰۰۹	۰/۹۷۸	۰/۶۸۷

۴. بحث و نتیجه‌گیری

در این پژوهش به‌منظور پیش‌بینی آلاینده‌های هوای شهر کرمانشاه و بهینه‌سازی متغیرهای مؤثر در پیش‌بینی از مدل شبکه عصبی پرسپترون چند لایه و تکنیک انتخاب ویژگی انتخاب پیشرو استفاده شد. نتایج حاصل از خروجی مدل پرسپترون چند لایه و مقایسه آماره‌های مذکور در جدول ۵ و ۶ به شرح زیر است مدل ۱ آلاینده مونوکسید کربن با حداکثر میزان IOA=۰/۹۶۷ و R²=۰/۷۳۴ و کمترین میزان خطا RMSE=۰/۰۶۹ و NMSE=۰/۱۰۰ نسبت به شش مدل دیگر در پیش‌بینی امروز آلاینده بهینه است برای پیش‌بینی فردا آلاینده مونوکسید کربن مدل ۲ و مدل ۵ دارای بیشترین میزان R²=۰/۸۲۸ برای مدل ۲ و

R²=۰/۷۲۵ برای مدل ۵ هستند ولی میزان آماره ۰/۹۸۰ IOA= در مدل ۲ بیشتر از مدل ۵، IOA=۰/۹۶۸ و همچنین میزان آماره‌های FB=-۰/۰۰۲، NMSE=۰/۱۵۱ و IOA=۰/۰۱۱ RMSE= در مدل ۲ کمتر از مدل ۵ است. لذا مدل ۲ برای پیش‌بینی فردا آلاینده مونوکسید کربن مناسب‌تر است. برای انتخاب مدل بهینه از میان مدل ۱ و ۲ در پیش‌بینی امروز و فردا آلاینده مونوکسید کربن می‌توان مدل ۲ را که دارای مقادیر کمتر FB=-۰/۰۰۲، NMSE=۰/۰۱۱ و IOA=۰/۱۵۱ و NMSE= نسبت به مدل ۵ با مقادیر، FB=-۰/۰۲۳ و IOA=۰/۹۸۰ و R²=۰/۸۲۸ است، را به عنوان مدل بهینه انتخاب کرد با توجه به نتایج حاصل شده برای پیش‌بینی

پیش‌بینی انتخاب کرد. در پیش‌بینی امروز آلاینده دی‌اکسید نیتروژن مدل ۲ و در پیش‌بینی فردا مدل ۳ با کمترین مقادیر آماره‌های $FB=0/029$ و $NMSE=30/0$ برای مدل ۲ و $FB=-0/102$ و $NMSE=0/326$ برای مدل ۳ و بیشترین مقادیر $IOA=0/930$ و $R^2=0/573$ برای مدل ۲ و $IOA=0/968$ و $R^2=0/573$ برای مدل ۳ در مقایسه با سایر مدل‌ها، مدل‌های مطلوبی هستند. مقایسه دو مدل ذکر شده برای دی‌اکسید نیتروژن نشان می‌دهد هر دو مدل دارای شرایط یکسانی از نظر حداکثر و حداقل مقادیر آماره‌های ذکر شده است. لذا با توجه به در نظر گرفتن $RMSE=0/105$ مرحله آموزش که در مدل ۲ کمتر از مدل ۳ است می‌توان گفت مدل ۲ با شش کمیت بهینه شده (آلاینده در روز گذشته، دما، بارش، میزان دید، سرعت باد و فشار) در پیش‌بینی مقدار دی‌اکسید نیتروژن مدل مناسب‌تری است. نتایج حاصل از پیش‌بینی آلاینده ازن حاکی از این است که مدل ۷ در پیش‌بینی امروز و مدل ۵ در پیش‌بینی فردا آلاینده از نظر شاخص $IOA=0/980$ دارای مقدار یکسانی هستند و شاخص‌های $RMSE=0/035$ ، $FB=0/001$ و $NMSE=0/001$ در مدل ۵ در مقایسه با مدل ۷ با مقادیر $RMSE=0/035$ و $IOA=0/980$ $NMSE=0/035$ برای مدل ۷ در مقایسه با مدل ۵ با داشتن امتیازات مثبت بیشتری در مقادیر آماره‌های ذکر شده به عنوان مدل بهینه در نظر گرفته می‌شود. در واقع انتخاب پیشرو بر مبنای رگرسیون خطی است و با حذف متغیرهایی که همبستگی کمتری با آلاینده‌های هوا دارند امکان استفاده از مؤثرترین متغیرها با حداکثر همبستگی و کمترین میزان خطا را فراهم ساخته است این تکنیک تا حد امکان تعداد ورودی به مدل را کاهش می‌دهد تا به جای استفاده از تمام متغیرها در پیش‌بینی تنها از متغیرهای مؤثرتر در پیش‌بینی هر آلاینده استفاده شود و این امر به پیش‌بینی بهتر و خطای کمتر مدل پرسپترون چند لایه منجر خواهد شد. با مشاهده خطاهای ناشی از پیش‌بینی آلاینده‌ها در جدول ۵ و ۶ که در

مونوکسید کربن به جای استفاده از هفت کمیت هواشناسی می‌توان از مدل ۲ با شش کمیت هواشناسی (آلاینده در روز گذشته، دما، بارش، میزان دید، سرعت باد و فشار) بهره برد در واقع انتخاب پیشرو با حذف کمیتی که کمترین همبستگی با آلاینده مذکور را دارد امکان پیش‌بینی دقیق‌تر همراه با خطای کمتر را فراهم کرده است. در پیش‌بینی امروز آلاینده ذرات معلق (PM_{10}) مدل ۴ دارای حداکثر مقدار شاخص $R^2=0/472$ است و آماره $FB=0/001$ و $NMSE=0/134$ و $RMSE=0/077$ دارای مقدار کمتری نسبت به سایر مدل‌ها است و مدل ۴ در پیش‌بینی فردا آلاینده دارای حداقل مقدار $NMSE=0/487$ است، بنابراین با مقایسه دو مدل امروز و فردا، به این نتیجه می‌رسیم که ترکیب ورودی مربوط به مدل ۴ با چهار کمیت هواشناسی (آلاینده روز گذشته، نقطه شبنم، میزان دید و سرعت باد) در پیش‌بینی امروز آلاینده ذرات معلق (PM_{10}) با کمترین میزان خطا $RMSE=0/077$ ، $NMSE=0/134$ و $FB=0/001$ و بیشترین مقدار $IOA=0/960$ در مقایسه با مدل ۴ فردا با مقادیر $IOA=0/472$ و $R^2=0/472$ ، $NMSE=0/088$ ، $RMSE=0/088$ ، $FB=0/082$ و $IOA=0/465$ و $R^2=0/465$ مدل مناسب‌تری برای پیش‌بینی است. نتایج پیش‌بینی آلاینده دی‌اکسید گوگرد نشان می‌دهد که مدل ۳ در پیش‌بینی امروز و مدل ۶ در پیش‌بینی فردا آلاینده دی‌اکسید گوگرد دارای حداقل مقدار آماره‌های $FB=-0/003$ و $FB=-0/005$ برای مدل ۳ و $NMSE=0/135$ برای مدل ۶، و کمترین مقدار $NMSE=0/105$ و همچنین حداکثر میزان $IOA=0/943$ و $R^2=0/645$ برای مدل ۳ و $IOA=0/972$ برای مدل ۶ است و مدل‌های بهینه محسوب می‌شوند. برای انتخاب یک مدل از میان دو مدل ذکر شده می‌توان مدل ۶ را با دو کمیت هواشناسی مؤثر در پیش‌بینی (آلاینده روز گذشته و میزان دید) با مقدار کمتر $NMSE=0/105$ و $RMSE=0/008$ و مقادیر بیشتر $IOA=0/972$ و $R^2=0/645$ در مقایسه با مدل ۳ برای

استفاده از روش انتخاب پیشرو در مدل‌های بهینه نسبت به سایر مدل‌ها به میزان قابل توجهی کاهش یافته است که نشان‌دهنده کاربرد خوب روش انتخاب پیشرو در بهینه‌سازی کمیت‌ها و کاهش خطای مدل پرسپترون چند لایه در پیش‌بینی و افزایش صحت مدل است. (جدول ۷)

مدل‌های بهینه منتخب نسبت به سایر مدل‌ها دارای مقادیر نزدیک به صفر و قابل قبولی هستند و از مقایسه خطای دو مدل بهینه انتخابی در جدول ۷ نسبت به یکدیگر که میزان خطا در مدل بهینه نهایی کمتر است می‌توانیم به این نتیجه برسیم میزان خطای پیش‌بینی در مدل پرسپترون چند لایه با

جدول ۷. مقایسه میزان خطا در مدل‌های بهینه مربوط به هر آلاینده در مدل پرسپترون چند لایه

FB	NMSE	RMSE تست	RMSE آموزش	مدل‌های بهینه	آلاینده‌ها
۰/۰۱۷	۰/۱۰۰	۰/۰۱۱	۰/۰۶۹	مدل ۱	مونوکسیدکربن امروز
-۰/۰۰۲	۰/۱۵۱	۰/۰۱۱	۰/۰۱۰	مدل ۲	مونوکسیدکربن فردا (مدل بهینه نهایی)
۰/۰۰۱	۰/۱۳۴	۰/۰۴۱	۰/۰۷۸	مدل ۴	گرد و غبار امروز (مدل بهینه نهایی)
۰/۰۸۲	۰/۴۸۷	۰/۰۷۲	۰/۰۸۸	مدل ۴	گرد و غبار فردا
-۰/۰۰۳	۰/۱۳۵	۰/۰۱۱	۰/۰۷۸	مدل ۳	دی‌اکسید گوگرد امروز
-۰/۰۰۵	۰/۱۰۵	۰/۰۰۸	۰/۰۹۳	مدل ۶	دی‌اکسید گوگرد فردا (مدل بهینه نهایی)
۰/۰۲۷	۰/۳۰۰	۰/۰۶۲	۰/۱۰۵	مدل ۲	دی‌اکسید نیتروژن امروز (مدل بهینه نهایی)
-۰/۱۰۲	۰/۳۲۶	۰/۰۶۴	۰/۱۱۷	مدل ۳	دی‌اکسید نیتروژن فردا
-۰/۰۱۸	۰/۰۹۱	۰/۰۳۵	۰/۰۴۸	مدل ۷	ازن امروز
۰/۰۰۱	۰/۰۰۱	۰/۰۳۵	۰/۰۴۶	مدل ۵	ازن فردا (مدل بهینه نهایی)

آلاینده‌های جو با کمترین خطای خروجی و کاهش هزینه محاسباتی و زمان و ایجاد مدل‌های متنوعی برای پیش‌بینی هر آلاینده و Balram و همکاران (۲۰۱۹) در تعیین کارآمدی سیستم هشدار کیفیت هوا برای برآورد $PM_{2.5}$ با استفاده از مدل شبکه عصبی مبتنی بر تکنیک انتخاب پیشرو با کمترین میزان خطا و حداکثر میزان R^2 مؤید این پژوهش هستند. Guajardo و همکاران (۲۰۰۶) برای ایجاد یک روش پیش‌بینی ترکیبی برای انتخاب مناسب‌ترین ویژگی‌ها و ایجاد مدل بهینه از مدل‌های ترکیبی مبتنی بر انتخاب پیشرو استفاده کردند نتایج نشان داد که این تکنیک در انتخاب و بهینه‌سازی متغیرها کارآمد است. همچنین Zhao (۲۰۱۶) به‌منظور پیش‌بینی کیفیت هوا با استفاده از شبکه عصبی از تکنیک انتخاب پیشرو به‌منظور انتخاب و بهینه‌سازی متغیرها استفاده کرد براساس نتایج حاصله مدل مبتنی بر تکنیک انتخاب پیشرو می‌تواند پیش‌بینی‌های دقیقی از کیفیت هوا ارائه دهد. در این راستا Azid و همکاران (۲۰۱۳) از شبکه عصبی بر مبنای تکنیک انتخاب پیشرو و مدل آنالیز اجزای اصلی به‌منظور پیش‌بینی

از مزایای تکنیک انتخاب پیشرو می‌توان به بهبود دقت مدل، کاهش زمان محاسبات در ساخت مدل، تسهیل تجسم داده‌ها و درک مدل و کاهش ریسک بیش از حد اشاره کرد. در واقع ایده اصلی این تکنیک ارزیابی سودمندترین زیر مجموعه از متغیرها برای یک الگوریتم یادگیری معین است و بهترین زیر مجموعه از ویژگی‌ها با توجه به عملکرد مدل تعیین می‌شود (Guyon & Elisseeff, 2003; Rakotomamonjy, 2002; Famili et al., 1997). بنابراین انتخاب مدل بهینه با استفاده از تکنیک انتخاب پیشرو و مدل پرسپترون چند لایه امکان صرفه‌جویی در زمان اجرای مدل به دلیل استفاده از حداقل متغیرهای تأثیرگذار در پیش‌بینی و کاهش هزینه محاسبات به دلیل جایگزینی این روش با روش‌هایی که نیازمند محاسبات پیچیده‌تری هستند را میسر ساخته است و با در نظر گرفتن مدل بهینه نهایی برای هر آلاینده و انتخاب مؤثرترین کمیت‌ها در پیش‌بینی امکان پیش‌بینی با دقت بیشتر و خطای کمتر فراهم می‌شود. نتایج حاصل از پژوهش Prasad و همکاران (۲۰۱۶) در پیش‌بینی

کمتر است که نمایانگر عملکرد رضایت‌بخش مدل پرسپترون چند لایه و روش انتخاب پیشرو در بهینه‌سازی متغیرهای هواشناسی است و چون استفاده از این مدل و تکنیک، پیچیدگی محاسباتی کمتری دارد بنابراین میزان هزینه محاسباتی آن نیز کم است و می‌توان از آن به عنوان روشی مناسب استفاده کرد. به‌طور کلی نتایج حاصل از این پژوهش نشان می‌دهد که استفاده از آزمون هم‌خطی و تکنیک انتخاب پیشرو به‌منظور حذف رابطه خطی و ایجاد زیرمجموعه‌ای از متغیرهای مؤثر در پیش‌بینی رضایت‌بخش بوده زیرا برای پیش‌بینی هیچ کدام از آلاینده‌ها نیاز به همه ۷ متغیر حاصل از خروجی آزمون هم‌خطی نبود و تنها متغیرهایی که وابستگی بیشتری با آلاینده‌ها داشتند، برای پیش‌بینی لحاظ شدند و از آنجایی که تعداد بهینه متغیر مستقل برای پیش‌بینی هر آلاینده متفاوت است پس می‌توان گفت که انتخاب متغیر مستقل اثرگذار با استفاده از تکنیک انتخاب پیشرو سبب کاهش میزان هزینه و زمان در انجام محاسبات و همچنین سبب افزایش دقت در پیش‌بینی مورد نظر خواهد شد.

۵. پیشنهاد

با توجه به کاربرد تکنیک انتخاب پیشرو در بهینه‌سازی متغیرهای پیش‌بینی‌کننده کیفیت هوا و مدیریت بهتر در زمینه‌های مرتبط پیشنهاد می‌شود از این تکنیک همراه با مدل‌های دیگر استفاده شود و دقت و صحت نتایج حاصل از آن با مدل پرسپترون چند لایه مقایسه شود.

یادداشت‌ها

1. Artificial Natural Network (ANN)
2. Multi-Layer Perceptron (MLP)
3. Forward Selection (FS)
4. Backward Elimination
5. Principal Component Analysis (PCA)
6. Fractional Bias (FB)

شاخص آلودگی هوا استفاده کردند نتایج حاصله نشان داد که مدل شبکه عصبی همراه با روش انتخاب پیشرو کاربرد موفقیت‌آمیزی در تصمیم‌گیری و حل مشکلات هواشناسی برای مدیریت بهتر دارد. همچنین در پیش‌بینی مونوکسیدکربن در هوای شهر تهران با استفاده از تکنیک انتخاب پیشرو و مدل‌های شبکه عصبی و ماشین بردار پشتیبان، مشخص شد روش انتخاب پیشرو در افزایش عملکرد مدل‌ها مؤثر است و به‌عنوان روشی سودمند توصیه می‌شود (Akbarzadeh et al., 2020). پژوهش حاضر نیز گویای پیش‌بینی مطلوب آلاینده‌ها و مطلوبیت تکنیک انتخاب پیشرو و مدل پرسپترون چند لایه در انتخاب مؤثرترین متغیرها و افزایش صحت در پیش‌بینی است. از آنجایی که استفاده از مدل پرسپترون چند لایه و تکنیک انتخاب پیشرو را می‌توان جایگزینی نسبت به مدل‌سازی‌های دیگر توصیه کرد که نیازمند قدرت محاسبات بسیاری هستند، می‌توان گفت استفاده از روش مذکور به کاهش هزینه محاسبات منجر شده است، زیرا در زمان خرابی تجهیزات و فقدان داده‌ها با اطلاع از مؤثرترین متغیرها در پیش‌بینی آلاینده‌ها و در واقع استفاده از ترکیب مدل شبکه عصبی و انتخاب پیشرو می‌توان در زمان لازم برای پیش‌بینی هر آلاینده به دلیل استفاده از حداقل متغیرهای تأثیرگذار به عنوان ورودی مدل، صرفه‌جویی کرد و به دلیل تعیین بهترین متغیرها در پیش‌بینی هر آلاینده با صحت و خطای کمتر پیش‌بینی را انجام داد. براساس نتایج حاصل شده برای آلاینده مونوکسید کربن مدل بهینه نهایی مدل ۲ با ۶ متغیر مستقل است و برای آلاینده ذرات معلق (PM₁₀) مدل ۴ با ۴ کمیت هواشناسی بهینه است. برای آلاینده دی‌اکسید گوگرد مدل ۶ با ۲ متغیر مستقل و برای آلاینده دی‌اکسید نیتروژن مدل ۲ با ۶ متغیر، بهینه انتخاب شدند و مدل ۵ با ۴ کمیت مؤثر برای آلاینده ازن بهینه است. میزان خطا در مدل‌های بهینه نسبت به سایر مدل‌ها

منابع

الماسی، ع، مرادی، م، شرفی، ک، و عباسی، ش. ۱۳۹۳. تغییرات فصلی کیفیت هوای کرمانشاه از نظر غلظت آلاینده PM₁₀ در

دوره ۴ ساله (۱۳۸۷-۹۰)، سلامت و بهداشت، ۵(۲): ۱۴۹-۱۵۸.

بخشی‌زاده، ف.، رضائیان، ه. و اکبری، م. ۱۳۹۴. مدل‌سازی مکانی-زمانی سه بعدی پراکنش آلاینده اکسیدهای ازت هوا ناشی از ترافیک در تقاطع خیابان ولی‌عصر-فاطمی شهر تهران، تحلیل فضایی مخاطرات محیطی، ۲(۱): ۴۳-۶۲.

جدی، ح.، عباسپور، ر.ع.، خالصیان، م. و علوی‌پناه، ک. ۱۳۹۶. پیش‌بینی غلظت آلاینده مونوکسیدکربن در کلانشهر تهران با استفاده از شبکه‌های عصبی مصنوعی، علوم و تکنولوژی محیط‌زیست، ۱۹(۵): ۱۵-۲۵.

رستمی فصیح، ز.، مصداقی‌نیا، ع.، ندافی، ک.، نبی‌زاده نودهی، ر.، محوی، ا.ح. و هادی، م. ۱۳۹۴. پیش‌بینی شاخص کیفیت هوا بر مبنای متغیرهای هواشناسی و مؤلفه‌های خودهمبسته با استفاده از شبکه عصبی مصنوعی، علوم پزشکی رازی، ۲۲(۱۳۷): ۳۱-۴۳.

رایگانی، ب. و خیراندیش، ز. ۱۳۹۶. بهره‌گیری از سری زمانی داده‌های ماهواره‌ای به‌منظور اعتبارسنجی کانون‌های شناسایی شده تولید گرد و غبار استان البرز، تحلیل فضایی مخاطرات محیطی، ۴(۴): ۱-۱۸.

رفیع‌پور گنتابی، م.، آل‌شیخ، ع.ا.، علیمحمدی، ع. و صادقی نیارکی، ا. ۱۳۹۵. توسعه مدل پیش‌بینی غلظت ازن در هوا با استفاده از شبکه عصبی مصنوعی، محیط‌زیست طبیعی، منابع طبیعی ایران، ۶۹(۱): ۴۷-۶۰.

سلطانی‌گردفرامرز، ط.، مفیدی، ع. و گندمکار، ا. ۱۳۹۴. بررسی همدیدی روزهای بسیار آلوده در شهر مشهد مورد مطالعه ۱۳ و ۱۴ نوامبر، تحلیل فضایی مخاطرات محیطی، ۲(۴): ۹۵-۱۱۲.

نصیری، ب.، زارعی چقابلکی، ز.، حلیمی، م. و رستمی فتح‌آبادی، م. ۱۳۹۵. بررسی تغییرات ارتفاع و ضخامت لایه مرزی در شرایط گردوغباری شهر اهواز، تحلیل فضایی مخاطرات محیطی، ۳(۲): ۵۲-۶۴.

Akbarzadeh, A., Vesali Naseh, M.R. and NodeFarahani, M. 2020. Carbon monoxide prediction in the atmosphere of Tehran using developed support vector machine. *Pollution*, 6(1): 43-57.

Alexandrov, V.D., Velikov, S.K., Donev, E.H. and Ivanov, D.M. 2005. Quantifying nonlinearities in ground level ozone behavior at mountain-valley station at ovnarsko, bulgaria by using neural networks. *Bulgarian Geophysical*, 31: 1-4.

Alves, L., Sperandio Nascimento, E.G. and Moreira, D.M. 2019. Hourly tropospheric ozone concentration forecasting using deep learning. *WIT Transactions on Ecology and the Environment*, 236: 129-138.

Azid, A., Juahir, H., Latif, M.T., Zain, S.M. and Osman, M.R. 2013. Feed-forward artificial neural network model for air pollutant index prediction in the southern region of peninsular malaysia. *J. Environmental Protection* 4(12):1-10.

Balram, D., Lian, K.Y. and Sebastian, N. 2019. Air quality warning system based on a localized PM_{2.5} soft sensor using a novel approach of Bayesian regularized neural network via forward feature selection. *Ecotoxicology and Environmental Safety*, 182(30): 1-9.

Cabaneros, S.M., Hughes, B.R. and Calautit, J.K. 2017. Hybrid artificial neural network models for effective prediction and mitigation of urban roadside NO₂ pollution. *Energy Procedia*, 142: 3524-3530.

Chen, G. 2008. *Encyclopedia of statistics in quality and reliability*. John Wiley and Sons Ltd pp. 1800.

Chen, S.X., Hong, X., Harris, C.J. and Sharkey, P.M. 2004. Sparse modeling using orthogonal forward regression with PRESS statistic and regularization. *IEEE Transactions on Systems Man and Cybernetics Part B*, 34(2): 898-911.

Cheng, S.Y., Li, L., Chen, D.S. and Li, J.B. 2012. A neural network based ensemble approach for improving the accuracy of meteorological fields used for regional air quality modeling. *Environmental Management*, 112: 404-414.

Cogliani, E. 2001. Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. *Atmospheric Environment*, 35(16): 2871-2877.

- Coman, A., Ionescu, A. and Candau, Y. 2008. Hourly ozone prediction for a 24-h horizon using neural networks. *Environmental Modelling and Software*, 23(12): 1407–1421.
- Dirk, V.P. and Bart. L. 2004. Customer attribution analysis for financial services using proportional hard models. *Operational Research*, 157(1): 196 -277.
- Eksioglu, B., Demirer, R. and Capar, I. 2005. Subset selection in multiple linear regression: a new mathematical programming approach. *Computers and Industrial Engineering*, 49(1): 155 -167.
- Famili, A., Shen, W.M., Weber, R. and Simoudis, E. 1997. Data preprocessing and intelligent data analysis. *Intelligent Data Analysis*, 1(1-4): 3–23.
- Gardner, M.W. and Dorling, S.R. 1999. Neural network modeling and prediction of hourly NOx and NO2 concentrations in urban air in London. *Atmospheric Environment*, 33(5): 709–719.
- Guajardo, J., Weber, R. and Miranda, J. 2006. A forecasting methodology using support vector regression and dynamic feature selection. *Information & Knowledge Management*, 5(4): 329-335.
- Guyon, I. and Elisseeff, A. 2003. An introduction to variable and feature selection. *Machine Learning Research*, 3: 1157–1182.
- Hrust, L., Klaic, Z.B., Krizan, J., Antonic, O. and Hercog, P. 2009. Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. *Atmospheric Environment*, 43(35): 5588–5596.
- Khan, J.A., Aelst, S.V. and Zamar. R.H. 2007. Building a robust linear model with forward selection and stepwise procedures. *Computational Statistics and Data Analysis*, 52(1): 239-248.
- Kolehmainen, M., Martikainen, H. and Ruuskanen. J. 2001. Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment*, 35(5): 815–825.
- Kurt, A. and Oktay, A.B. 2010. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. *Expert Systems with Applications*, 37(12): 7986–7992.
- Pastor Barsenas, B., Soria Ivas, E. and Martín-Guerrero, J.D. 2005. Unbiased sensitivity analysis and pruning techniques in neural networks for surface ozone modeling. *Ecological Modelling*, 182(2): 149–158.
- Perez, P. 2012. Combined model for PM₁₀ forecasting in a large city. *Atmospheric Environment*, 60: 271–276.
- Prasad, K., Gorai, A.k. and Goyal, P. 2016. Developmen to ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time. *Atmospheric environment*, 128: 246-262.
- Rakotomamonjy, A. 2002. Variable selection using SVM based criteria. *Machine Learning Research*, 3: 1357–1370.
- Sharifi, K., Khosravi, T., Moradi, M. and Pirsahab, M. 2015. Air quality and variations in PM₁₀ pollutant concentration in western Iran during a four-year period (2008-2011), Kermanshah- a case study. *Engineering Science and Technology*, 10(1): 47-56.
- Stamenkovic, L.J., Antanasijevic, D.Z., Ristic, M.D., Peric Grujic, A.A. and Pocajt, V.V. 2016. Prediction of nitrogen oxides emissions at the national level based on optimized artificial neural network model. *Air Quality Atmosphere & Health*, 10:15-23.
- Unnikrishnan, R. and Madhu, G. 2019. Comparative study on the effects of meteorological and pollutant parameters on ANN modelling for prediction of SO₂. *SN Applied Sciences*, 1: 1-12.
- Wang, X.X., Chen, S., Lowe, D. and Harris, C.J. 2006. Sparse support vector regression based on orthogonal forward selection for the generalized kernel model. *Neurocomputing*, 70(1-3): 462 -474.
- Zinatizadeh, A.A., Zinadini, S., Pirsahab, M., Atafar, Z., Kurdian, A.R., Dezfoulinejad, A. and Yavari, F. 2014. Dust level forecasting and its interaction with gaseous pollutants using artificial neural network: A case study for kermanshah, Iran. *Energy and Environment*, 5(1): 51-58.
- Zhao, C. 2016. *Air quality forecasting using neural networks*, master's thesis, Supervisor: Prof. Juha Karhunen, Department of Computer Science, Aalto University.
- Zhu, Y.M., Lu, X.X. and Zhou, Y. 2007. Suspended sediment flux modeling with artificial neural network: An example of the long Chuan Jiang River in the Upper Yangtze Catchment China. *Geomorphology*, 84(1): 111-125.