Meetup

THE 7TH SESSION
A SYSTEMATIC VIEW OF
DATA SCIENCE

TITLE : LONG-TERM PREDICTION

ALIREZA SHADMAN

HEAD OF DEPARTMENT OF INDUSTRIAL
ENGINEERING FACULTY OF ENGINEERING

ICDS Innovation Center for Data Science

به نام نامی او

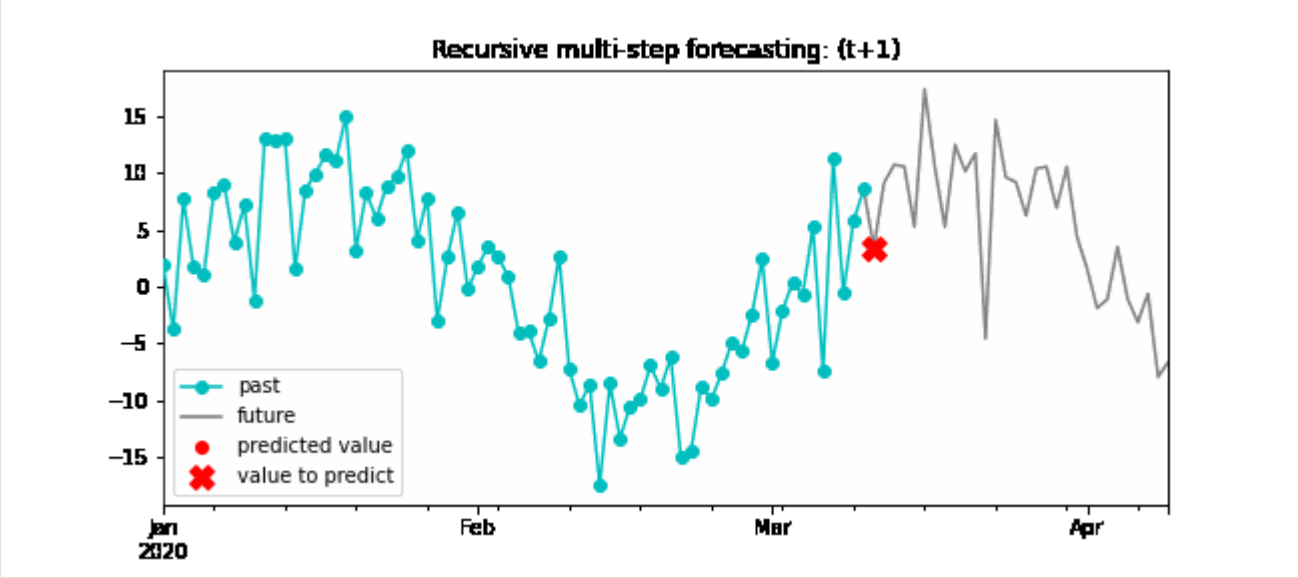
پیش بینی چند گام جلوتر

Multi-Step Ahead Prediction

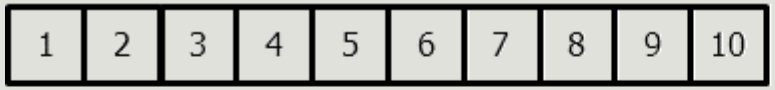
علیرضا شادمان

استادیار گروه مهندسی صنایع دانشگاه فردوسی مشهد

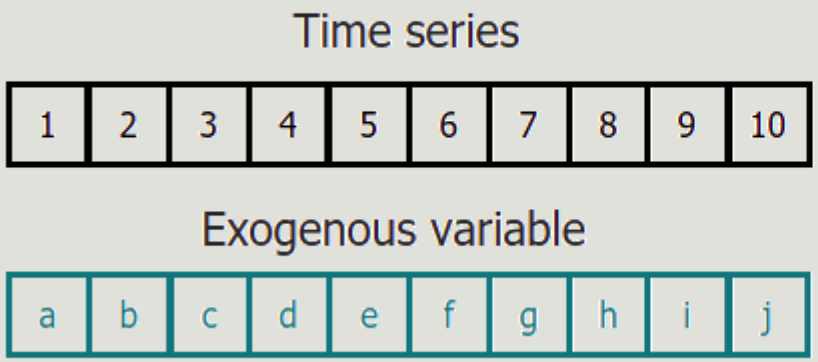
پیش بینی سری های زمانی



پیش بینی سری های زمانی (یک گام جلوتر)

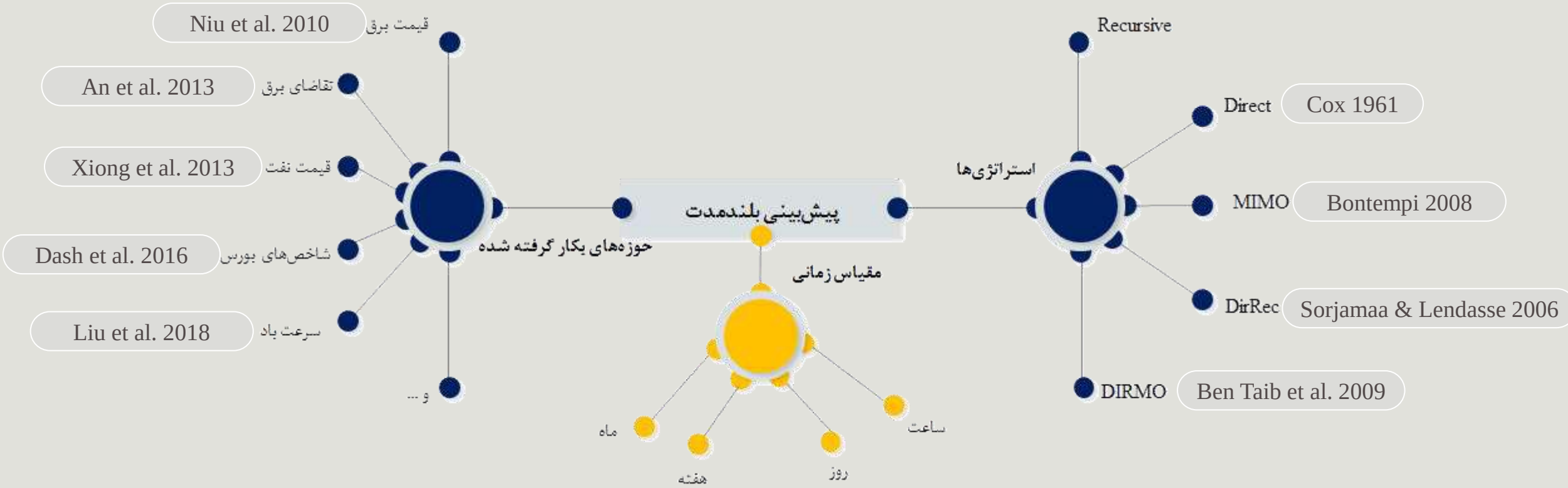


پیش بینی سری های زمانی (یک گام جلوتر)



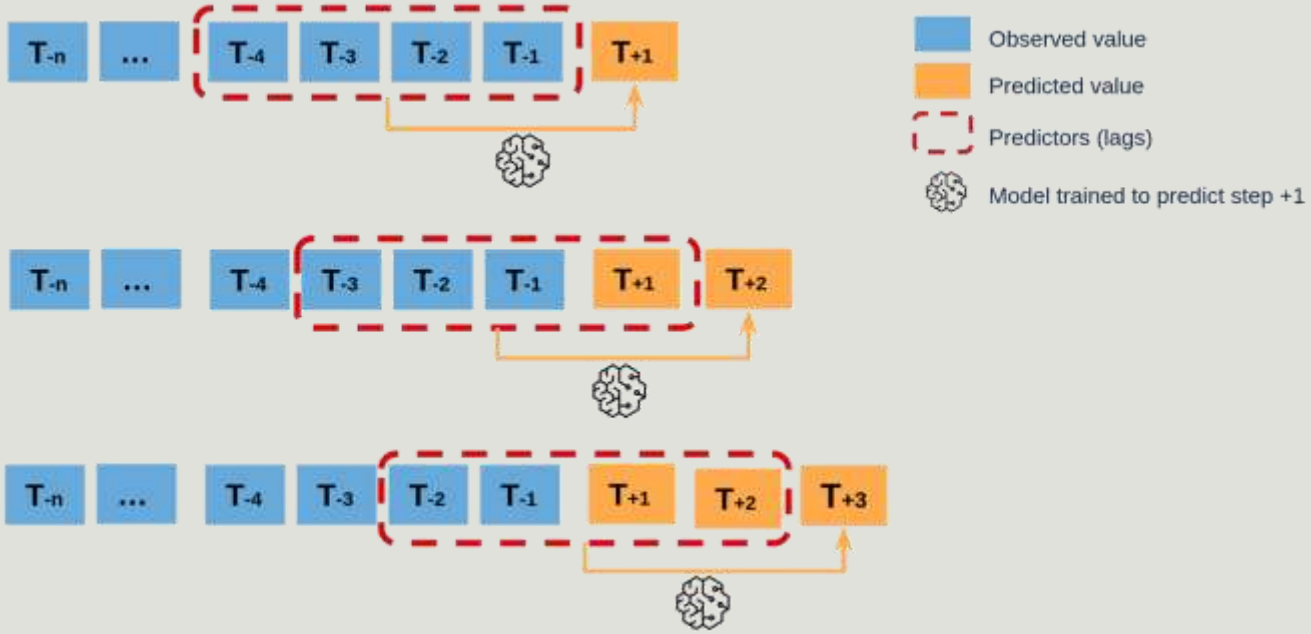
X						y
1	2	3	4	5	f	6
2	3	4	5	6	g	7
3	4	5	6	7	h	8
4	5	6	7	8	i	9
5	6	7	8	9	j	10

پیش بینی سری های زمانی (چند گام جلوتر)



پیش بینی سری های زمانی (چند گام جلوتر)

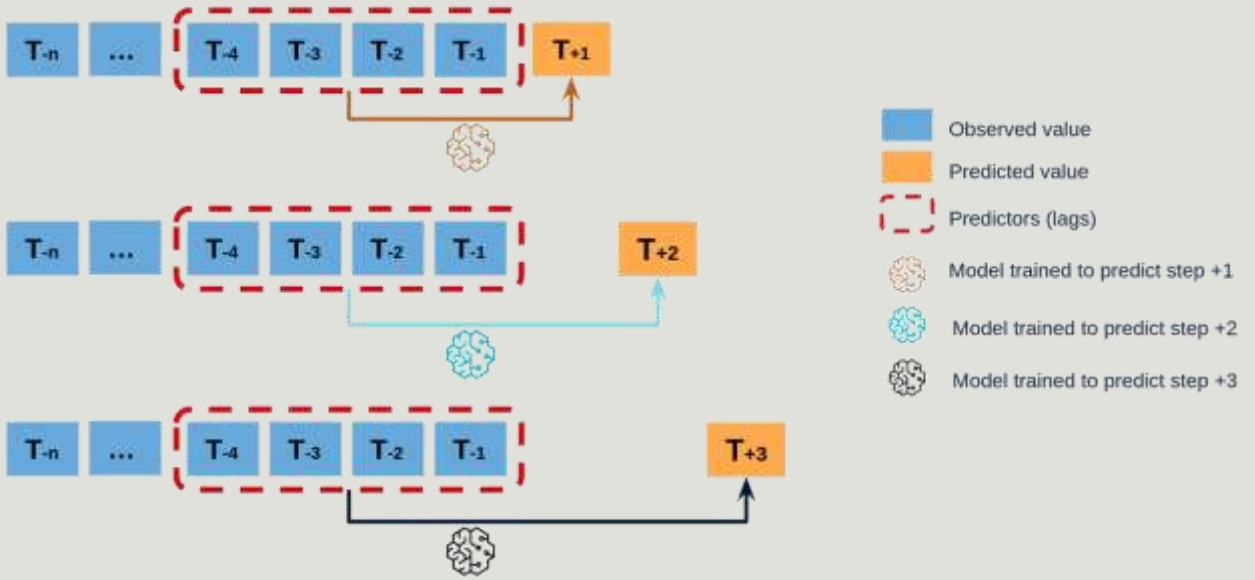
استراتژی تکراری (Recursive):



در استراتژی Rec ، مدل پیش‌بینی برای یک دوره بعد ایجاد می‌شود. سپس به منظور پیش‌بینی مقدار متغیر در H گام جلوتر، این مدل H بار مورد استفاده قرار می‌گیرد و در هر بار مقدار پیش‌بینی شده در دوره قبلی به عنوان ورودی مدل لحاظ می‌گردد.

پیش بینی سری های زمانی (چند گام جلوتر)

استراتژی مستقیم (Direct):



در استراتژی Dir ، برای هر گام زمانی مدلی مجزا ایجاد می‌گردد و به این ترتیب در این استراتژی H مدل پیش‌بینی بایستی تعلیم یابد.

پیش بینی سری های زمانی (چند گام جلوتر)

استراتژی چند ورودی – چند خروجی (MIMO): در استراتژی MIMO تنها یک مدل به منظور پیش‌بینی متغیر در H گام جلوتر تعلیم داده می‌شود و خروجی این مدل به جای یک مقدار، به صورت برداری با H مقدار خواهد بود.

استراتژی های ترکیبی:

- ترکیب استراتژی های تکراری و مستقیم
- ترکیب استراتژی های MIMO و مستقیم

Atmospheric Pollution Research
Volume 10, Issue 3, May 2019, Pages 689-700

Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas

Ida Kalate Ahani, Majid Salari, Alireza Shadman

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.apr.2018.11.006

Get rights and content

Abstract

In recent years, the atmospheric pollution in most metropolitan cities has become a crisis and the necessity of air quality forecasting has increased. Among different air pollutants, $PM_{2.5}$ is considered as the major air pollutant in urbanized regions, especially because of serious harmful health effects on human being. So, there is an urgent need to develop air quality forecast programs capable of providing accurate predictions over a long future horizon. Predicting $PM_{2.5}$ concentrations for several steps ahead of time is of great interest, especially in decision-making related to control policies and emergency measures such as traffic limitations, school closures, or temporarily shutting down major polluting industrial units. In this paper, commonly used multi-step ahead prediction strategies, including Recursive (Rec), Direct (Dir), Direct-Recursive (DirRec), Multi-Input Multi-Output (MIMO) and Direct-MIMO (DIRMO) along with Autoregressive integrated moving average with exogenous variables (ARIMAX) and Multi-Layer Perceptron (MLP) modelling techniques are examined. Also, the independent variables are considered as time series variables and are forecasted using ARIMA/MLP model in order to be used for prediction of the dependent variables in multi-steps ahead of time. The experimental study is performed using $PM_{2.5}$ data in Mashhad, Iran. Daily $PM_{2.5}$ forecasts for this city is provided for the next 10 days. Four different feature selection methods are also implemented and compared. The results indicate that recursive strategy with LASSO feature selection in ARIMAX model overcomes in most of time steps.

Journal of Cleaner Production
Volume 263, 1 August 2020, 120983

An ensemble multi-step-ahead forecasting system for fine particulate matter in urban areas

Ida Kalate Ahani, Majid Salari, Alireza Shadman

Show more

https://doi.org/10.1016/j.jclepro.2020.120983

Abstract

In recent years, growing air pollution has become a significant issue due to its detrimental effects on the environment and different living organisms. Providing accurate and reliable forecasts of air quality over a long future horizon is an effective way to mitigate health risks. In this paper, the problem of urban PM2.5 forecasts for several days ahead is considered. An ensemble multi-step-ahead forecasting system is introduced for this problem, which combines different multi-step-ahead strategies (including single-output and multi-output approaches). The proposed hybrid framework consists of three parts. In the first part, the Ensemble Empirical Mode Decomposition (EEMD) technique is combined with a prediction tool and multi-step-ahead strategies. Boosting idea is considered in the second part of the algorithm. Finally, the stacked ensemble of boosted hybrid structures is developed to provide the final multi-step-ahead forecasts. Least Square Support Vector Regression (LSSVR), and Long Short-Term Memory neural network (LSTM) are employed as the prediction tools in the proposed hybrid framework. Through real PM2.5 data examples from Mashhad, Iran, the proposed ensemble model is investigated for 1-day-ahead to 10-days-ahead. The results reveal the effectiveness of the ensemble model in comparison with the multi-step-ahead strategies in all time-steps. The proposed model with LSSVR prediction tool shows the smallest mean RMSE, MAE, and MAPE values of 7.810, 5.562, and 18.104% over all time-steps, and RMSE improvement rates of more than 35% compared to simply combining different multi-step-ahead strategies with LSSVR approach.

Atmospheric Pollution Research
Volume 11, Issue 9, September 2020, Pages 1645-1656

Multi-step-ahead prediction of fine particulate matter considering real-time decomposition techniques and uncertainty of input variables

Ida Kalateh Ahani, Majid Salari, Alireza Shadman

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.apr.2020.06.028

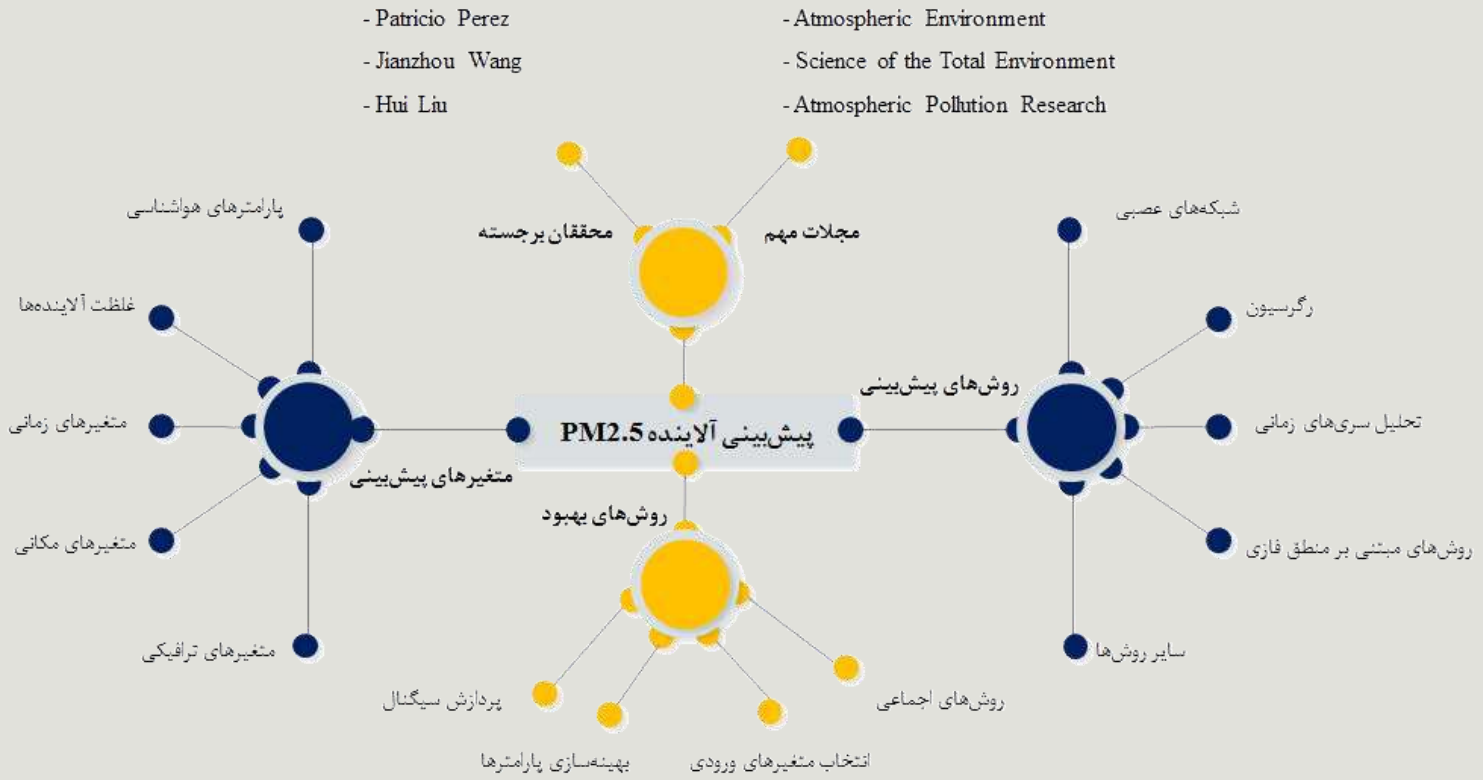
Get rights and content

Abstract

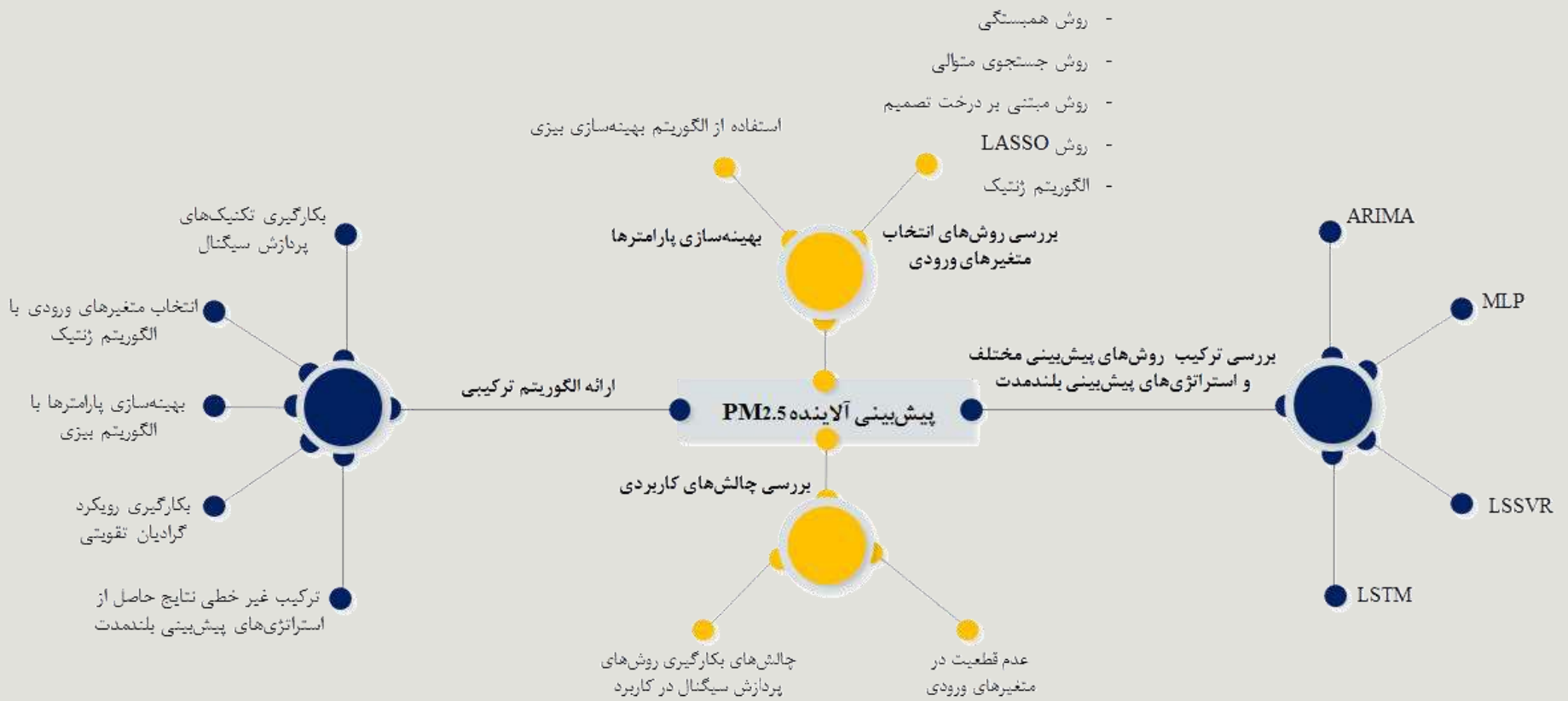
Fine particulate matter is one of the major air pollutants in urban areas, which adversely affects people's health and is considered as a serious threat to the society. Effective prediction of this pollutant can provide information for sensitive people to avoid or reduce outdoor activities and on the other hand can help regulators for efficient decision-making related to precautionary measures. This paper develops a novel hybrid algorithm for multi-step-ahead prediction of PM2.5 in which two challenges of forecasting in real applications has been taken into account. First, the challenge of employing decomposition techniques in practice is addressed. Different real-time approaches have been explored and compared in decomposition-based and noise-removal-based frameworks. Also, a real-time approach which combines feature selection and noise-removal-based technique is implemented in the framework of the proposed algorithm. The second challenge is the lack of access to real values of meteorological parameters, as influential predictors, in practical cases and use of forecasted values instead. To address the uncertainty in model inputs, Monte Carlo simulation is employed in the framework of the proposed algorithm. Probabilistic forecasts of the proposed algorithm output a distribution of the predicted values for a given time-point which can be used to calculate prediction intervals and probability of exceeding PM2.5 warning threshold. These results provide valuable information for decision-makers and regulators to take precautions and put control measures in place. According to the results, if a public alert is issued on days with at least 50% probability of exceeding PM2.5 warning level, it is observed that in 76.88% of days the algorithm indicates a correct warning for 1-day-ahead prediction which decreases to 55.00%, 41.25% and 37.50% for 2-days-ahead to 4-days-ahead forecasts.

پیش‌بینی بلندمدت غلظت آلاینده ذرات معلق کوچکتر از ۲.۵ میکرومتر با استفاده از روش‌های آماری

خلاصه مطالعات:

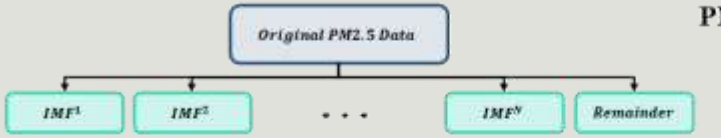


پیش‌بینی بلندمدت غلظت آلاینده ذرات معلق کوچکتر از ۲.۵ میکرومتر با استفاده از روش‌های آماری



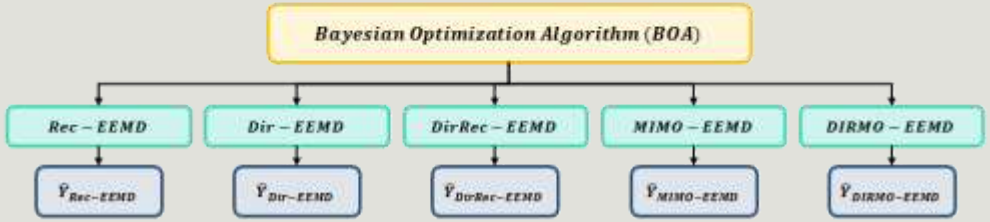
Part I: Combination of multi-step-ahead strategies with EEMD

A: تجزیه سری زمانی آلاینده PM2.5



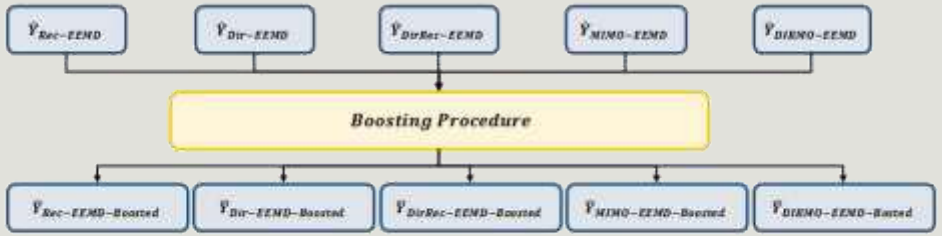
B: تعیین پارامترهای LSSVR/LSTM توسط الگوریتم BOA برای هر یک از ساختارهای پیش‌بینی چندگام جلوتر در ترکیب با الگوریتم EEMD

C: پیش‌بینی H گام جلوتر با استفاده از ساختارهای توسعه یافته



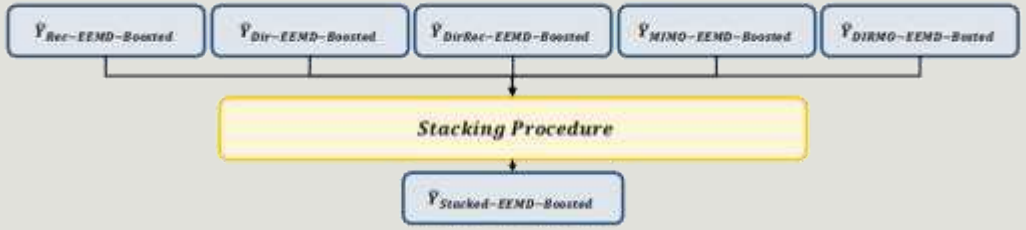
Part II: Boosting prodecure

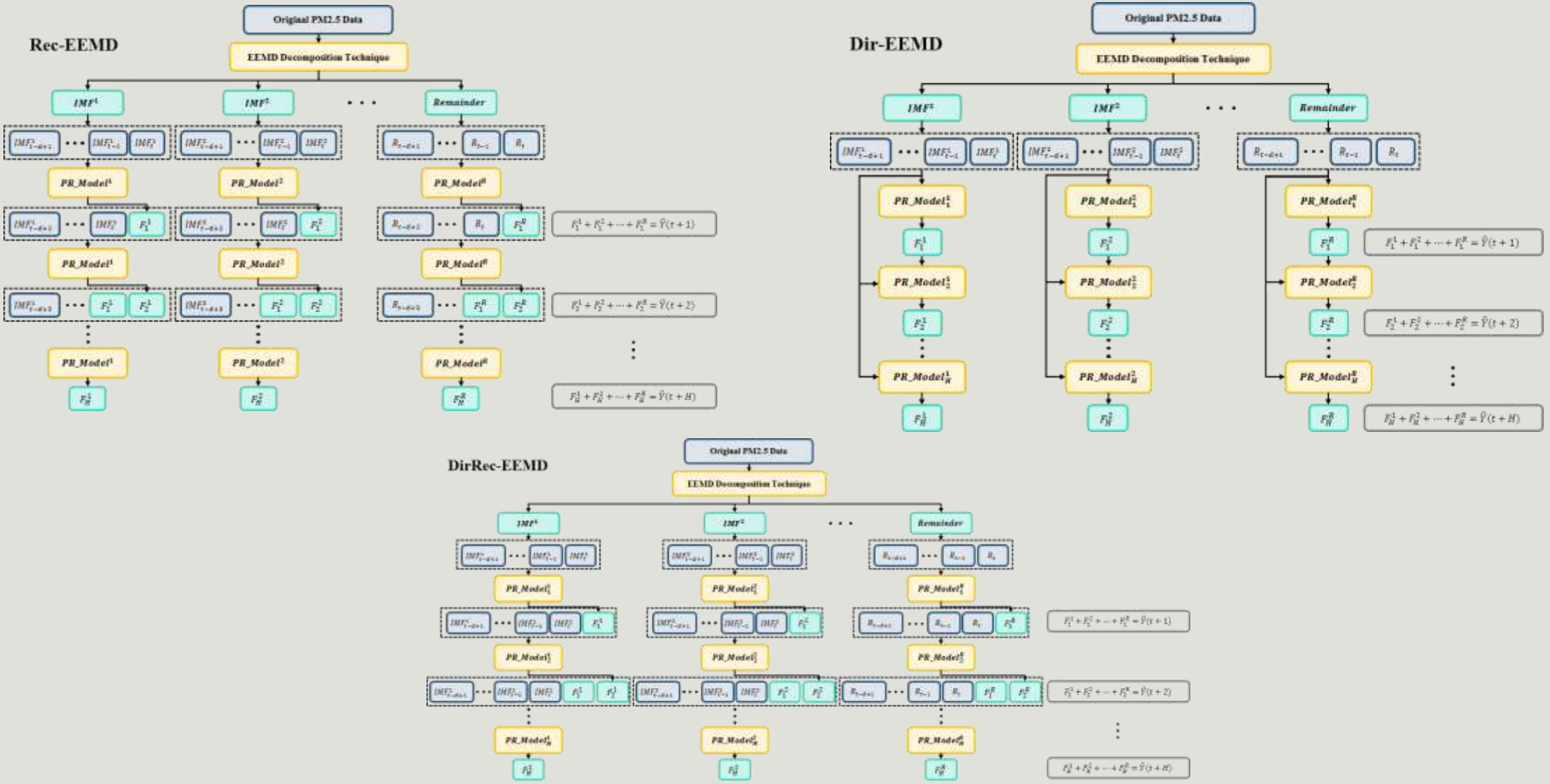
A: پیاده‌سازی رویکرد گرادیان تقویتی برای هر یک از ساختارهای پیش‌بینی چند گام جلوتر (۵-۷)

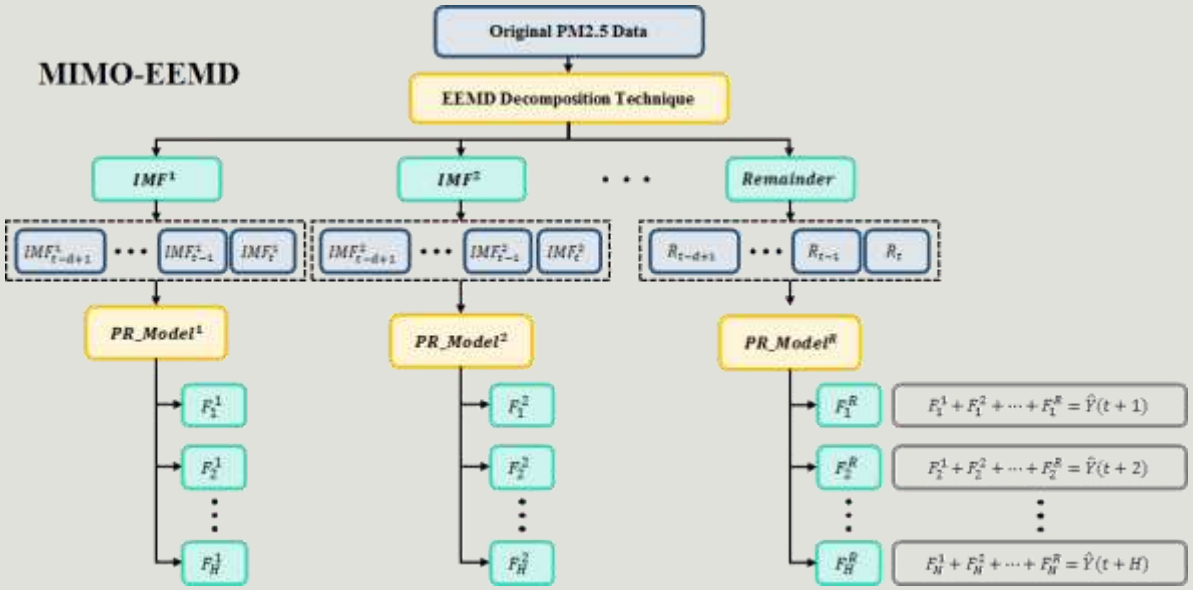


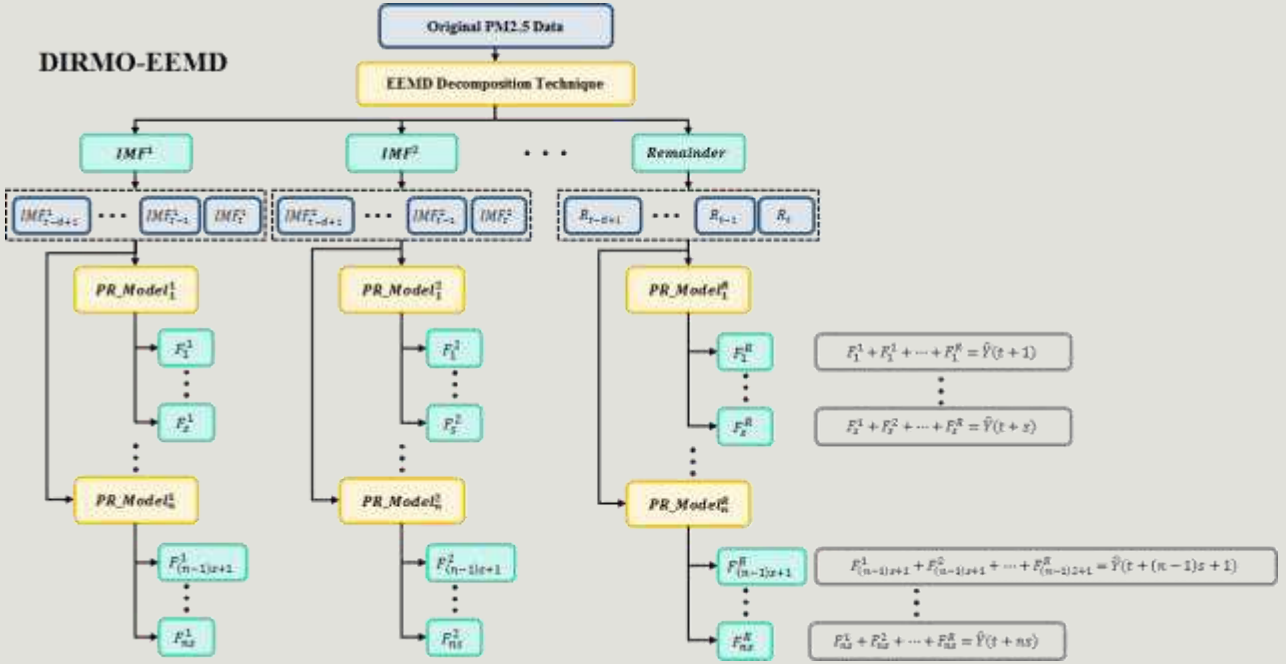
Part III: Stacking prodecure

A: ترکیب نتایج حاصل از مرحله قبل با استفاده از روش LSSVR/LSTM جهت بدست آوردن پیش‌بینی نهایی




Rec-EEMD
Original PM2.5 Data
EEMD Decomposition Technique
IMF¹
IMF²
Remainder
PR_Model¹
PR_Model²
PR_Model^R
Dir-EEMD
Original PM2.5 Data
EEMD Decomposition Technique
IMF¹
IMF²
Remainder
DirRec-EEMD
Original PM2.5 Data
EEMD Decomposition Technique
IMF¹
IMF²
Remainder


MIMO-EEMD
Original PM2.5 Data
EEMD Decomposition Technique
IMF^1
IMF^2
Remainder
PR_Model^1
PR_Model^2
PR_Model^R
$F_1^1 + F_1^2 + \cdots + F_1^R = \hat{Y}(t+1)$
$F_2^1 + F_2^2 + \cdots + F_2^R = \hat{Y}(t+2)$
$F_H^1 + F_H^2 + \cdots + F_H^R = \hat{Y}(t+H)$


DIRMO-EEMD
Original PM2.5 Data
EEMD Decomposition Technique
IMF^1
IMF^2
Remainder
$PR\_Model_1^1$
$PR\_Model_1^2$
$PR\_Model_1^R$
$PR\_Model_n^1$
$PR\_Model_n^2$
$PR\_Model_n^R$
$F_1^1 + F_1^2 + \cdots + F_1^R = \hat{Y}(t+1)$
$F_s^1 + F_s^2 + \cdots + F_s^R = \hat{Y}(t+s)$
$F_{(n-1)s+1}^1 + F_{(n-1)s+1}^2 + \cdots + F_{(n-1)s+1}^R = \hat{Y}(t+(n-1)s+1)$
$F_{ns}^1 + F_{ns}^2 + \cdots + F_{ns}^R = \hat{Y}(t+ns)$

با تشکر از توجه شما