

PM2.5

Features:17

Target:

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

1	AMB_TE	CH4	CO	NMHC	NO	NO2	NOx	O3	PM10	RAIN	FAIRH	SO2	THC	WD_HR	WIND_D	WIND_SF	WS_HR
	19.5	1.9	0.4175	0.089167	1.054167	10.78333	11.97917	42.95833	54.04167	0	71.29167	2.383333	1.9875	62.04167	62.16667	2.770833	2.320833
	18.70833	1.9625	0.60375	0.173333	5.420833	20.225	25.64583	26.17083	89.375	0.016667	83.66667	2.495833	2.120833	76.25	87.70833	1.529167	1.1125
	19.16667	2.079167	0.721667	0.27375	14.7375	29.58333	44.41667	4.904167	55.29167	0.816667	90.41667	2.408333	2.354167	94.91667	84.375	1.15	0.758333
	20.20833	2.154167	0.996667	0.360417	23.4	22.975	46.41667	9.4375	42.75	0.008333	88.54167	2.9125	2.504167	188.5417	186.7083	0.883333	0.491667
	20.375	2.033333	0.654583	0.324583	15.87391	24.69565	40.52174	7.291667	55.08696	0.008333	87.33333	3.433333	2.345833	119.9167	133.5417	1.3375	0.895833
	18.25	1.9	0.375417	0.16125	1.779167	17.14167	19.02083	24.61667	26.20833	1.825	86.45833	2.258333	2.075	62.04167	60.66667	2.633333	1.8375
	18	1.881818	0.41875	0.135455	2.426087	12.31304	14.68696	38.75	47.5	0	75.04167	2.583333	1.990909	61.33333	62.08333	3.675	2.920833
	16.875	1.891667	0.44875	0.12375	1.3125	14.025	15.44583	36.70833	51.20833	0.083333	77.375	2.420833	1.9875	65.04167	67.5	2.85	2.325
	17.33333	1.9125	0.557917	0.144583	3.441667	17.87083	21.26667	32.91667	49.16667	0	71.54167	3.554167	2.045833	63.66667	61.91667	2.8375	2.204167
	18.375	1.954167	0.542917	0.189583	5.879167	17.65	23.67083	21.25417	45.625	0	81.45833	2.654167	2.145833	94.125	85.79167	1.425	1.029167
	16.875	1.845833	0.415833	0.1325	3.841667	19.50417	23.35833	27.7375	19.625	2.3	89.41667	1.666667	1.991667	68	75.375	1.904167	1.329167
	16.29167	1.854545	0.382174	0.099545	2.104348	10.78261	12.90435	34	32.82609	0.066667	80.125	1.326087	1.945455	64.08333	64.875	3.629167	2.766667
	14.25	1.9	0.477917	0.117917	1.129167	11.39167	12.31667	33.79167	27.125	0	82.33333	1.1625	1.983333	64.54167	64.45833	3.525	2.641667
	14.66667	1.941667	0.64375	0.17375	3.016667	17.6125	20.625	31	61.33333	0	75.58333	3.420833	2.116667	66.875	66.79167	2.895833	2.266667
	15.625	1.958333	0.653333	0.213333	4.929167	26.125	31.08333	19.45	70.29167	0.083333	83.75	3.1625	2.183333	68.83333	69.25	1.725	1.408333
	18.66667	1.929167	0.515	0.146667	3.020833	17.99167	20.85	26.72083	33.25	0.016667	77	2.466667	2.075	74.91667	80.58333	2.116667	1.683333
	16.875	1.945833	0.51375	0.148333	7.004167	16.49167	23.54167	22.95417	26.45833	2.025	89.83333	1.795833	2.095833	114.8083	126.5833	1.5625	0.925
	14.83333	1.913043	0.542917	0.13913	2.816667	13.4625	16.33333	36.58333	51.3913	0.008696	70.33333	2.083333	2.03913	71.375	71.625	3.3625	2.741667
	14.58333	1.895833	0.5875	0.149167	3.3	15.51667	18.74583	36.58333	72.875	0.033333	74.58333	3.708333	2.05	71.20833	71.58333	3.7	3.0625
	15.5	1.883333	0.511667	0.18	4.416667	19.25	23.5875	26.81667	26.75	0	78	3.0625	2.075	82.58333	83.91667	2.7875	2.266667
	15.54167	1.858333	0.374583	0.15375	5.254167	16.1	21.29583	22.7875	18.45833	0.666667	88.95833	1.9125	2.033333	77.79167	74.41667	2.075	1.495833
	15.25	1.841667	0.329167	0.135417	4.941667	15.8375	20.725	26.81667	14.20833	1.341667	86	2.083333	1.9875	61.70833	61.625	2.9	2.254167
	10.0375	1.866667	0.420417	0.082917	2.216667	10.62083	12.74583	32.29167	34.95833	0.9	84.5	2.1375	1.9375	64.79167	67.20833	4.504167	3.716667
	4.970833	1.804167	0.3225	0.08375	1.604167	8.495833	9.991667	35.5	40.16667	0.258333	70.29167	2.583333	1.908333	77.125	76.79167	3.458333	2.920833
	7.833333	1.841667	0.335	0.098333	2.991667	11.48333	14.43333	33.33333	40.375	0	61.33333	2.270833	1.945833	71.75	69.95833	3.795833	3.045833
	13.87917	1.947826	0.59	0.266957	12.36696	23.82609	36.13043	22.23913	35.73913	0	64.25	3.117391	2.217391	111.9167	108.4583	1.445833	0.991667

#Date	PM2.5
2016/1/1	27.70833
2016/1/2	46.08333
2016/1/3	38.47826
2016/1/4	23.52174
2016/1/5	31.47826
2016/1/6	6.916667
2016/1/7	21.625
2016/1/8	25.5
2016/1/9	22.29167
2016/1/10	21.16667
2016/1/11	6.375
2016/1/12	11.13043
2016/1/13	11.75
2016/1/14	34.95833
2016/1/15	43.29167
2016/1/16	14.91667
2016/1/17	10.16667
2016/1/18	24.65217
2016/1/19	44
2016/1/20	10.58333
2016/1/21	5.666667
2016/1/22	2.958333
2016/1/23	17.625
2016/1/24	12.75

PM2.5(M=1)

Hypothesis: $h_{\theta}(x) = \theta_0 x_0 + \theta_1 x_1 + \dots + \theta_{17} x_{17}$

Cost function: $J(\theta_0, \theta_1, \dots, \theta_{16}, \theta_{17}) = \frac{1}{2N} \sum_{i=1}^N (h_{\theta}(x^i) - T^i)^2$

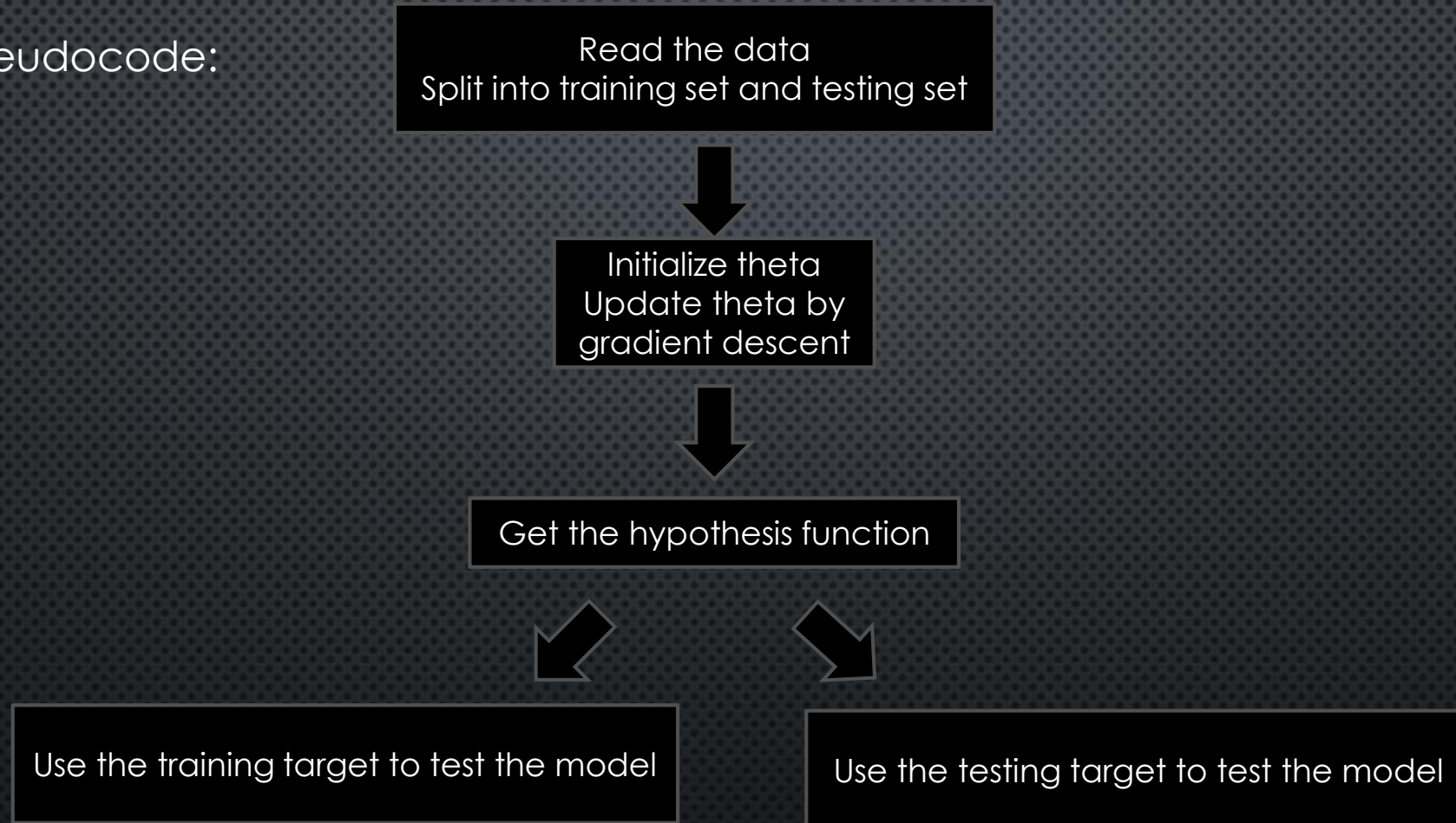
Gradient descent: *repeat before convergence*{

$$\theta_n = \theta_n - \alpha \frac{\partial J(\theta_0, \theta_1, \dots, \theta_{17})}{\partial \theta_n}$$

}

PM2.5(M=1)

Pseudocode:



PM2.5(M=1)hypothesis function

x_i^j : the i th feature in the j th data

$$h_{\theta}(x) = [\theta_0 \quad \theta_1 \quad \dots \quad \theta_{16} \quad \theta_{17}]_{1 \times 18} \cdot \begin{bmatrix} x_0^1 & x_0^2 & \dots & x_0^{N-1} & x_0^N \\ x_1^1 & x_1^2 & \dots & x_1^{N-1} & x_1^N \\ \dots & \dots & \dots & \dots & \dots \\ x_{16}^1 & x_{16}^2 & \dots & x_{16}^{N-1} & x_{16}^N \\ x_{17}^1 & x_{17}^2 & \dots & x_{17}^{N-1} & x_{17}^N \end{bmatrix}_{18 \times N} = [y^1 \quad y^2 \quad \dots \quad y^{N-1} \quad y^N]_{1 \times N}$$

$x_0^i = 1$
Just to put the constant of the hypothesis function into the matrix

Transpose from the original data

```
#hypothesis function(M=1):  
def hypothesis(theta,X):  
    return np.matmul(theta,np.transpose(X))
```



```
#gradient descent
def gradient_descent(theta,X,T,learning_rate,iteration):
    N=len(X)
    cost_function=[]
    for i in range(iteration):
        cost_function.append(np.sum((hypothesis(theta,X)-T)**2)/len(X)/2)
        theta_grad=(1/N)*np.matmul((hypothesis(theta,X)-T),(X))
        theta-=learning_rate*theta_grad
        if i %1000000==0:
            print("it is the %d time of iterations, rmse is %.10lf and cost function is %.10lf" %(i,rmse(hypothesis(theta,X),T),np.sum((hypothesis(theta,X)-T)**2)/len(X)/2))
    return theta,cost_function
```

Theta_grad:
$$= \begin{bmatrix} \frac{\partial J}{\partial \theta_0} & \frac{\partial J}{\partial \theta_1} & \cdots & \frac{\partial J}{\partial \theta_{16}} & \frac{\partial J}{\partial \theta_{17}} \end{bmatrix}_{1 \times 18}$$

$$= \frac{1}{N} \begin{bmatrix} \sum_{i=1}^N (h_{\theta}(x^i) - T^i) \cdot x_0^i & \sum_{i=1}^N (h_{\theta}(x^i) - T^i) \cdot x_1^i & \cdots & \sum_{i=1}^N (h_{\theta}(x^i) - T^i) \cdot x_{16}^i & \sum_{i=1}^N (h_{\theta}(x^i) - T^i) \cdot x_{17}^i \end{bmatrix}_{1 \times 18}$$

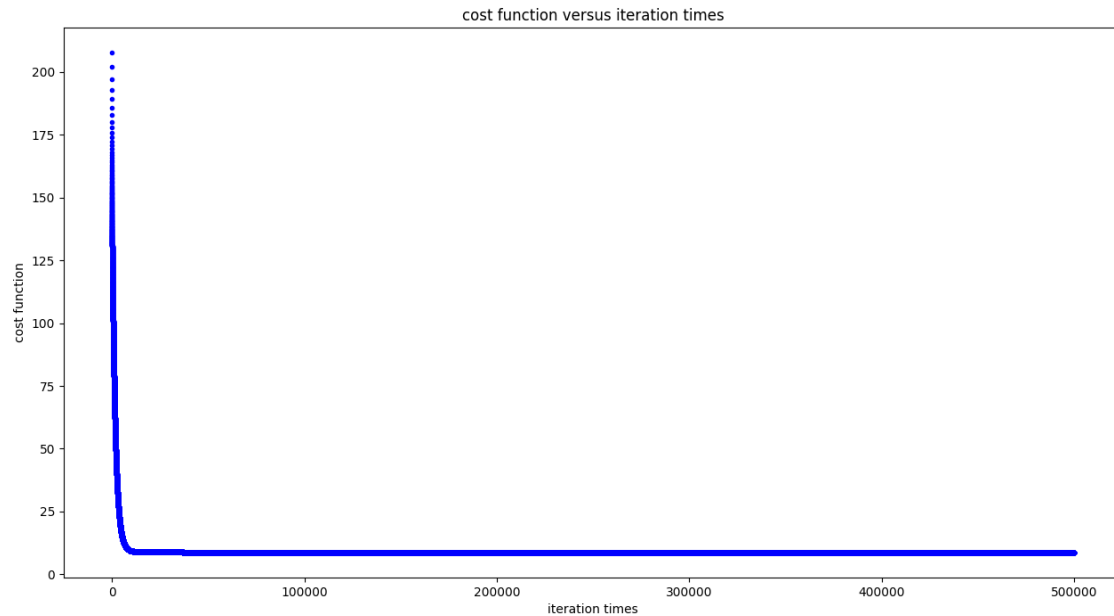
$$= \frac{1}{N} ([y^1 \quad y^2 \quad \cdots \quad y^{N-1} \quad y^N]_{1 \times N} - [T^1 \quad T^2 \quad \cdots \quad T^{N-1} \quad T^N]_{1 \times N}) \cdot \begin{bmatrix} x_0^1 & x_0^2 & \cdots & x_0^{N-1} & x_0^N \\ x_1^1 & x_1^2 & \cdots & x_1^{N-1} & x_1^N \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{16}^1 & x_{16}^2 & \cdots & x_{16}^{N-1} & x_{16}^N \\ x_{17}^1 & x_{17}^2 & \cdots & x_{17}^{N-1} & x_{17}^N \end{bmatrix}_{18 \times N}$$

Hypothesis: $h_{\theta}(x) = \theta_0 x_0 + \theta_1 x_1 + \cdots + \theta_{17} x_{17}$

Cost function: $J(\theta_0, \theta_1, \dots, \theta_{16}, \theta_{17}) = \frac{1}{2N} \sum_{i=1}^N (h_{\theta}(x^i) - T^i)^2$

Gradient descent: $\theta_n = \theta_n - \alpha \frac{\partial J(\theta_0, \theta_1, \dots, \theta_{17})}{\partial \theta_n}$

PM2.5(M=1) Outcome



Initial state:

```
theta= [[0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]]
```

Running for the method of gradient descent

before iteration, rmse is 20.38007928 and cost function is 207.67381581

it is the 50000 time of iterations, rmse is 4.19507851 and cost function is 8.79934186

it is the 100000 time of iterations, rmse is 4.18962794 and cost function is 8.77649112

it is the 200000 time of iterations, rmse is 4.18349911 and cost function is 8.75083239

it is the 250000 time of iterations, rmse is 4.18081514 and cost function is 8.73960763

it is the 300000 time of iterations, rmse is 4.17823396 and cost function is 8.72881952

it is the 350000 time of iterations, rmse is 4.17573201 and cost function is 8.71836891

it is the 400000 time of iterations, rmse is 4.17329904 and cost function is 8.70821242

it is the 450000 time of iterations, rmse is 4.17092925 and cost function is 8.69832540

it is the 500000 time of iterations, rmse is 4.16861854 and cost function is 8.68869028

Final state:

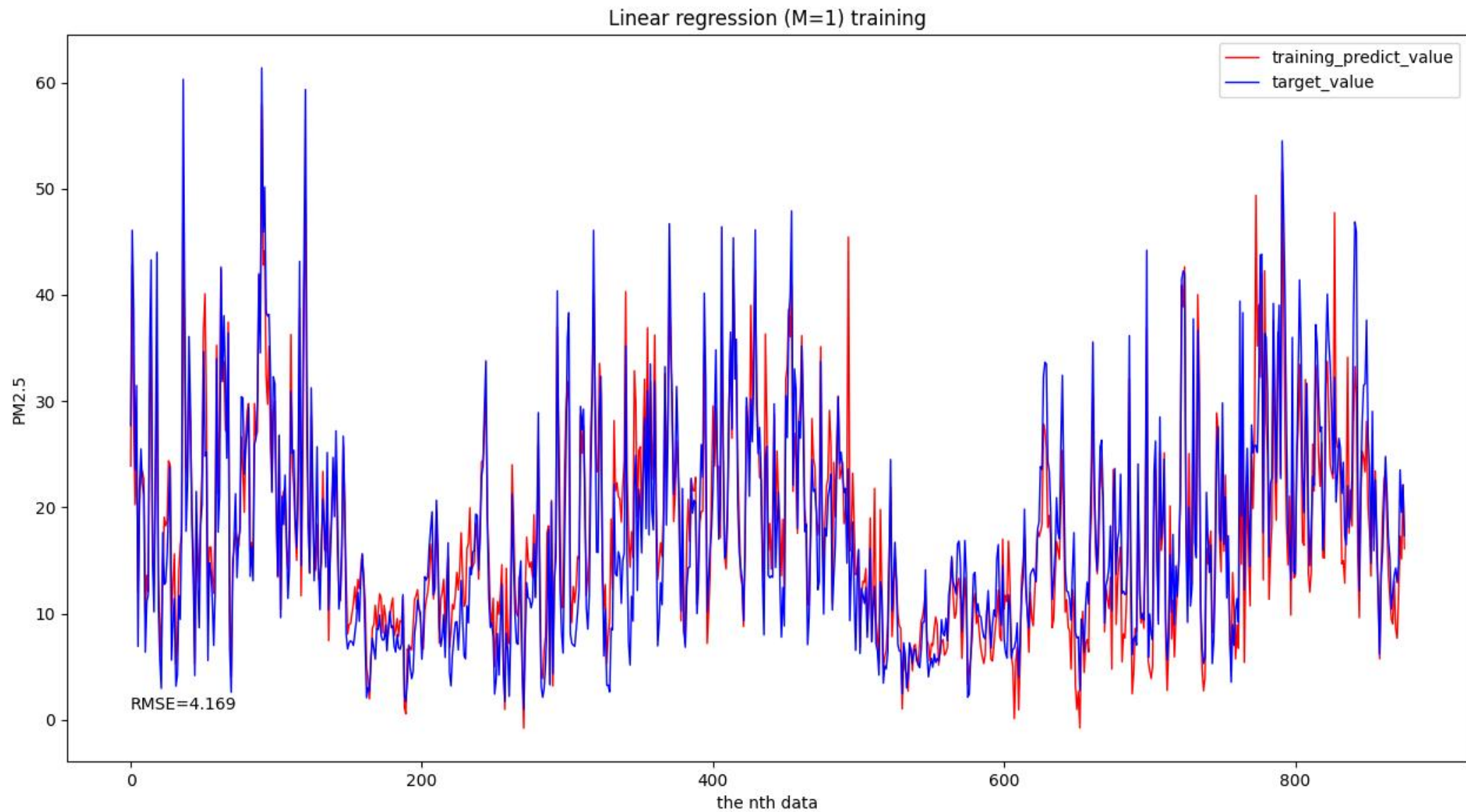
```
theta= [[-0.62542292 -0.22271807 -0.4065718   1.11799464 -0.17069211 -0.24752272
```

0.20173709 0.03665635 0.04426917 0.46077071 -0.43757445 0.04076274

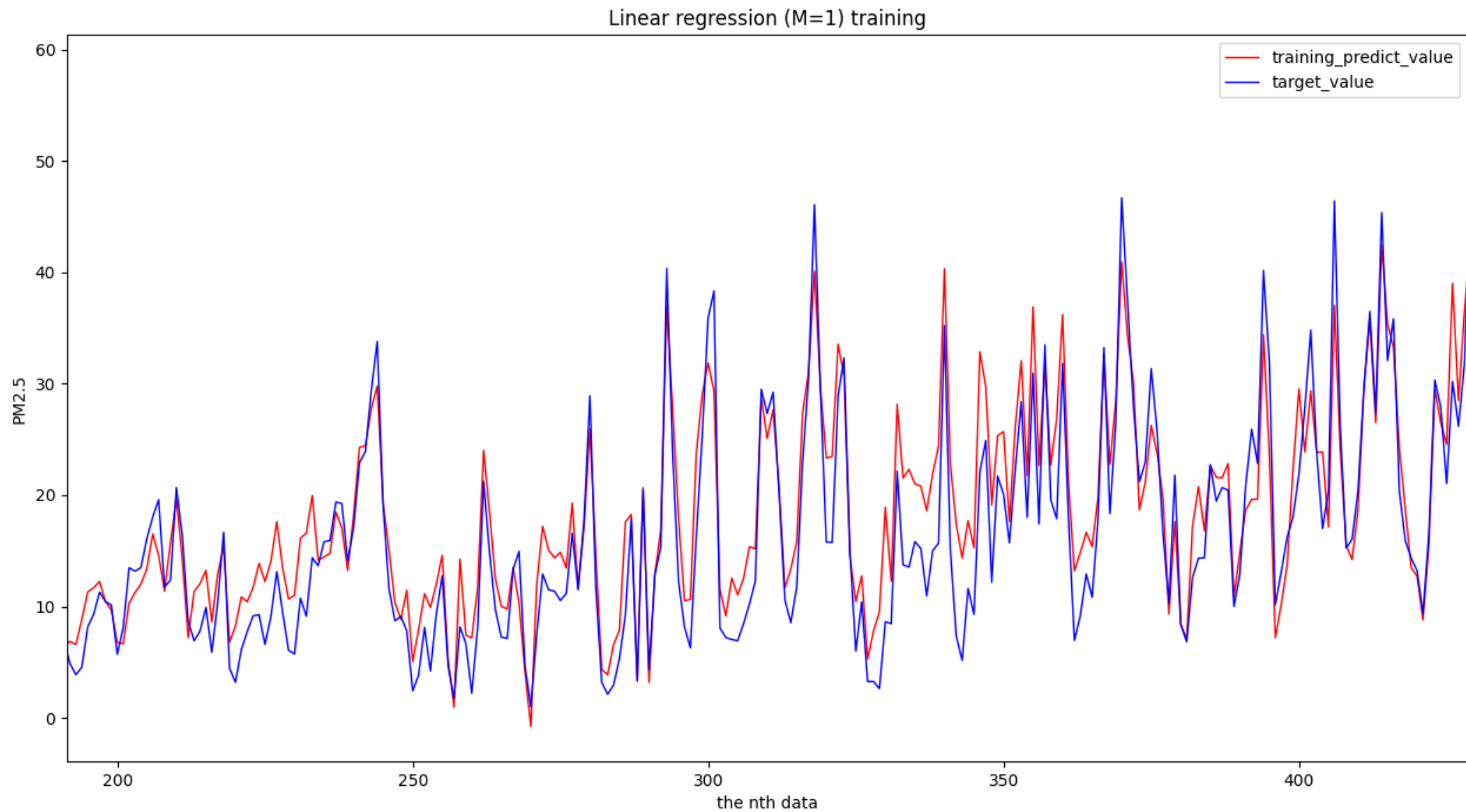
```
0.1783273 -0.61653299 0.04520647 -0.0308441 -0.55604122 -0.61305675]]
```

```
the runtime of this program:94.919
```

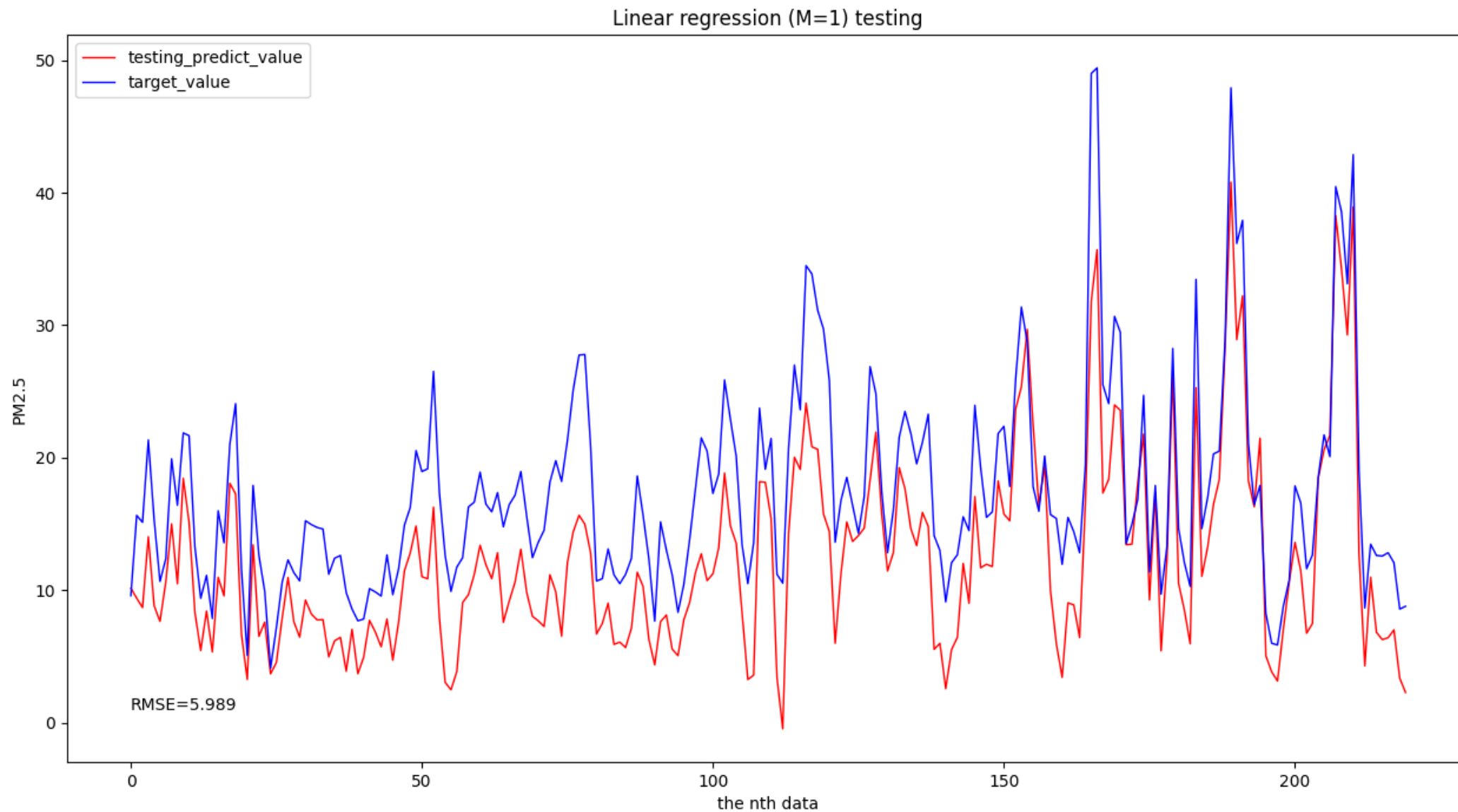
PM2.5(M=1)Outcome



PM2.5(M=1)Outcome



PM2.5(M=1)Outcome



PM2.5(M=2)

Hypothesis: $h_{\theta}(x) = \theta_0 x_0 + \theta_1 x_1 + \dots + \theta_{17} x_{17} +$
 $\theta_{18} x_1 x_1 +$
 $\theta_{19} x_2 x_1 + \theta_{20} x_2 x_2 +$
 $\theta_{21} x_3 x_1 + \theta_{22} x_3 x_2 + \theta_{23} x_3 x_3 +$
 \dots
 $\theta_{154} x_{17} x_1 + \theta_{155} x_{17} x_2 + \dots + \theta_{169} x_{17} x_{16} + \theta_{170} x_{17} x_{17}$

$$\Rightarrow h_{\theta}(x) = \sum_{i=0}^{17} \theta_i x_i + \sum_{i=1}^{17} \sum_{j=1}^i \theta_n x_i x_j, n \text{ in range}(18, 171)$$

$$\text{Cost function: } J(\theta_0, \theta_1, \dots, \theta_{169}, \theta_{170}) = \frac{1}{2N} \sum_{i=1}^N (h_{\theta}(x^i) - T^i)^2$$

Gradient descent: *repeat before convergence*{

$$\theta_n = \theta_n - \alpha \frac{\partial J(\theta_0, \theta_1, \dots, \theta_{170})}{\partial \theta_n}$$

}

PM2.5(M=2)hypothesis function-1

Append the data (or features) by multiplied two features :

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22.....169 170

```
#append the dataX to match the theta (171 features)
k=18
for i in range(1,18):
    for j in range(1,i+1):
        if k in range(18,171):
            dataX=np.insert(dataX,k,values=dataX[:,i]*dataX[:,j],axis=1)
            k+=1
```

1	AMB_TE	CH4	CO	NMHC	NO	NO2	NOx	O3	PM10	RAINF	HR	SO2	THC	WD_HR	WIND_D	WIND_SF	WS_HR	1	2	2	...	16	17
	19.5	1.9	0.4175	0.089167	1.054167	10.78333	11.97917	42.95833	54.04167	0	71.29167	2.383333	1.9875	62.04167	62.16667	2.770833	2.320833						
	18.70833	1.9625	0.60375	0.173333	5.420833	20.225	25.64583	26.17083	89.375	0.016667	83.66667	2.495833	2.120833	76.25	87.70833	1.529167	1.1125						
	19.16667	2.079167	0.721667	0.27375	14.7375	29.58333	44.41667	4.904167	55.29167	0.816667	90.41667	2.408333	2.354167	94.91667	84.375	1.15	0.758333						
	20.20833	2.154167	0.996667	0.360417	23.4	22.975	46.41667	9.4375	42.75	0.008333	88.54167	2.9125	2.504167	188.5417	186.7083	0.883333	0.491667						
	20.375	2.033333	0.654583	0.324583	15.87391	24.69565	40.52174	7.291667	55.08696	0.008333	87.33333	3.433333	2.345833	119.9167	133.5417	1.3375	0.895833						
	18.25	1.9	0.375417	0.16125	17.779167	17.14167	19.02083	24.61667	26.20833	1.825	86.45833	2.258333	2.075	62.04167	60.66667	2.633333	1.8375						
	18	1.881818	0.41875	0.135455	2.426087	12.31304	14.68696	38.75	47.5	0	75.04167	2.583333	1.990909	61.33333	62.08333	3.675	2.920833						
	16.875	1.891667	0.44875	0.12375	1.3125	14.025	15.44583	36.70833	51.20833	0.083333	77.375	2.420833	1.9875	65.04167	67.5	2.85	2.325						
	17.33333	1.9125	0.557917	0.144583	3.441667	17.87083	21.26667	32.91667	49.16667	0	71.54167	3.554167	2.045833	63.66667	61.91667	2.8375	2.204167						
	18.375	1.954167	0.542917	0.189583	5.879167	17.65	23.67083	21.25417	45.625	0	81.45833	2.654167	2.145833	94.125	85.79167	1.425	1.029167						
	16.875	1.845833	0.415833	0.1325	3.841667	19.50417	23.35833	27.7375	19.625	2.3	89.41667	1.666667	1.991667	68	75.375	1.904167	1.329167						
	16.29167	1.854545	0.382174	0.099545	2.104348	10.78261	12.90435	34	32.82609	0.066667	80.125	1.326087	1.945455	64.08333	64.875	3.629167	2.766667						
	14.25	1.9	0.477917	0.117917	1.129167	11.39167	12.31667	33.79167	27.125	0	82.33333	1.1625	1.983333	64.54167	64.45833	3.525	2.641667						
	14.66667	1.941667	0.64375	0.17375	3.016667	17.6125	20.625	31	61.33333	0	75.58333	3.420833	2.116667	66.875	66.79167	2.895833	2.266667						
	15.625	1.958333	0.653333	0.213333	4.929167	26.125	31.08333	19.45	70.29167	0.083333	83.75	3.1625	2.183333	68.83333	69.25	1.725	1.408333						
	18.66667	1.929167	0.515	0.146667	3.020833	17.99167	20.85	26.72083	33.25	0.016667	77	2.466667	2.075	74.91667	80.58333	2.116667	1.683333						
	16.875	1.945833	0.51375	0.148333	7.004167	16.49167	23.54167	22.95417	26.45833	2.025	89.83333	1.795833	2.095833	114.8083	126.5833	1.5625	0.925						
	14.83333	1.913043	0.542917	0.13913	2.816667	13.4625	16.33333	36.58333	51.3913	0.008696	70.33333	2.083333	2.03913	71.375	71.625	3.3625	2.741667						
	14.58333	1.895833	0.5875	0.149167	3.3	15.51667	18.74583	36.58333	72.875	0.033333	74.58333	3.708333	2.05	71.20833	71.58333	3.7	3.0625						
	15.5	1.883333	0.511667	0.18	4.416667	19.25	23.5875	26.81667	26.75	0	78	3.0625	2.075	82.58333	83.91667	2.7875	2.266667						
	15.54167	1.858333	0.374583	0.15375	5.254167	16.1	21.29583	22.7875	18.45833	0.666667	88.95833	1.9125	2.033333	77.79167	74.41667	2.075	1.495833						
	15.25	1.841667	0.329167	0.135417	4.941667	15.8375	20.725	26.81667	14.20833	1.341667	86	2.083333	1.9875	61.70833	61.625	2.9	2.254167						
	10.0375	1.866667	0.420417	0.082917	2.216667	10.62083	12.74583	32.29167	34.95833	0.9	84.5	2.1375	1.9375	64.79167	67.20833	4.504167	3.716667						
	4.970833	1.804167	0.3225	0.08375	1.604167	8.495833	9.991667	35.5	40.16667	0.258333	70.29167	2.583333	1.908333	77.125	76.79167	3.458333	2.920833						
	7.833333	1.841667	0.335	0.098333	2.991667	11.48333	14.43333	33.33333	40.375	0	61.33333	2.270833	1.945833	71.75	69.95833	3.795833	3.045833						
	13.87917	1.947826	0.59	0.266957	12.38696	23.82609	36.13043	22.23913	35.73913	0	64.25	3.117391	2.217391	111.9167	108.4583	1.445833	0.991667						

PM2.5(M=2)hypothesis function-2

$$h_{\theta}(x) = \sum_{i=0}^{17} \theta_i x_i + \sum_{i=1}^{17} \sum_{j=1}^i \theta_n x_i x_j, n \text{ in range}(18,171)$$

$$\Rightarrow h_{\theta}(x) = \sum_{i=0}^{17} \theta_i x_i + \sum_{i=18}^{170} \theta_i x_i = \sum_{i=0}^{170} \theta_i x_i = [\theta_0 \quad \theta_1 \quad \dots \quad \theta_{169} \quad \theta_{170}]_{1 \times 171} \cdot \begin{bmatrix} x_0^1 & x_0^2 & \dots & x_0^{N-1} & x_0^N \\ x_1^1 & x_1^2 & \dots & x_1^{N-1} & x_1^N \\ \dots & \dots & \dots & \dots & \dots \\ x_{169}^1 & x_{169}^2 & \dots & x_{169}^{N-1} & x_{169}^N \\ x_{170}^1 & x_{170}^2 & \dots & x_{170}^{N-1} & x_{170}^N \end{bmatrix}_{171 \times N}$$

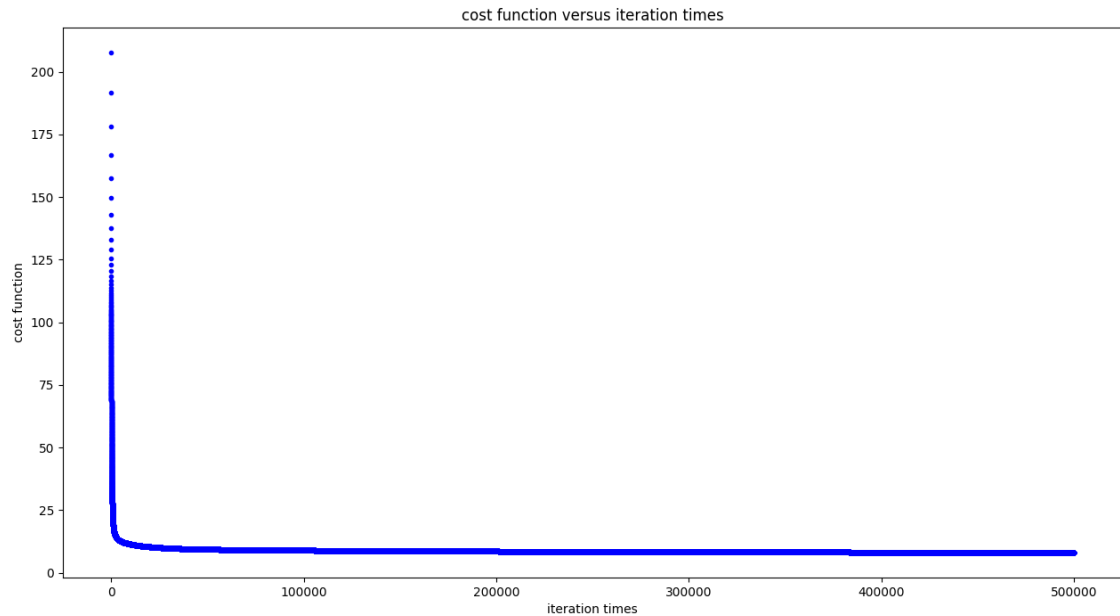
$$= [y^1 \quad y^2 \quad \dots \quad y^{N-1} \quad y^N]_{1 \times N}$$

Transpose from the original data

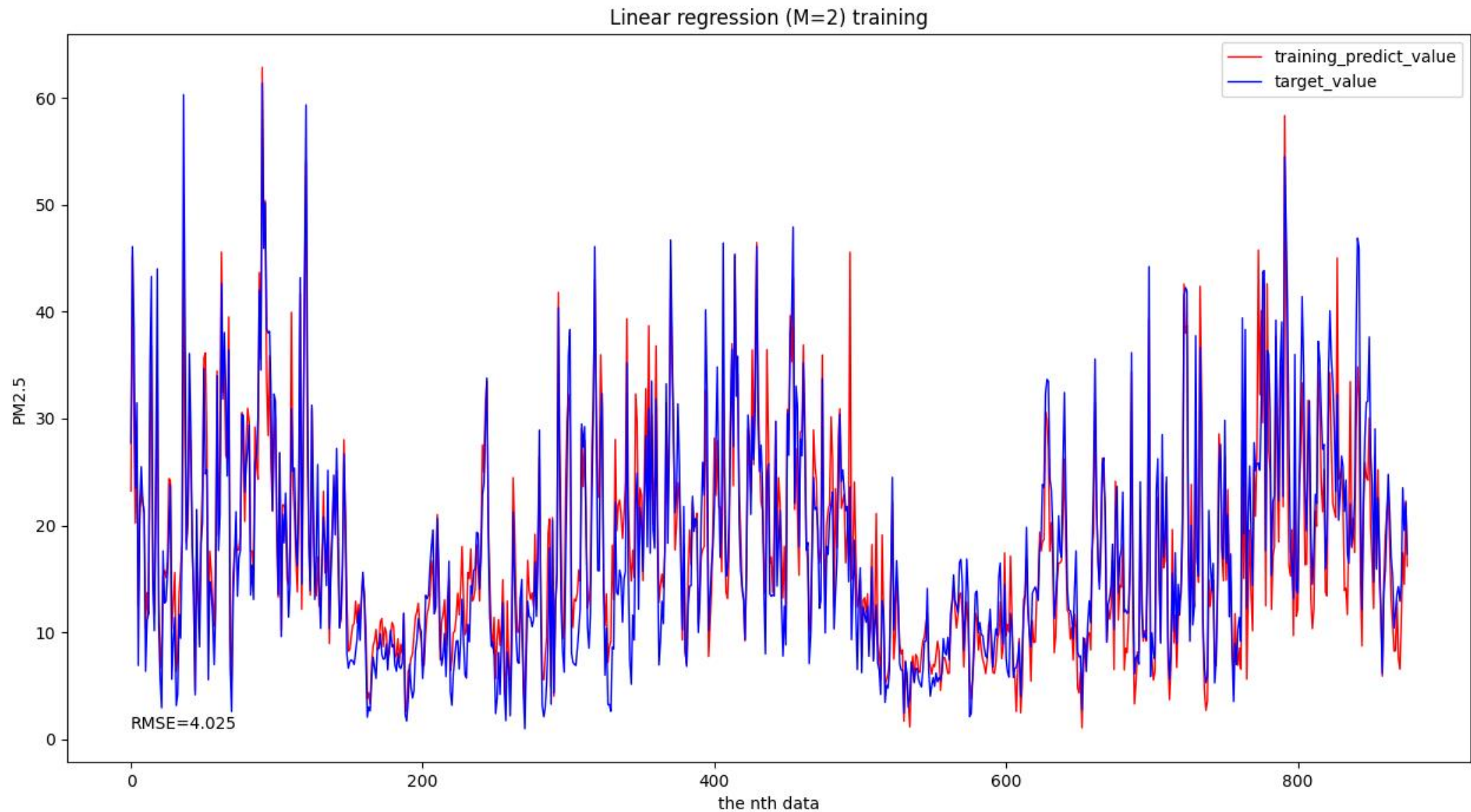
```
#hypothesis function
def hypothesis(theta,X):
    return np.matmul(theta,np.transpose(X))
```

The only different from M=1

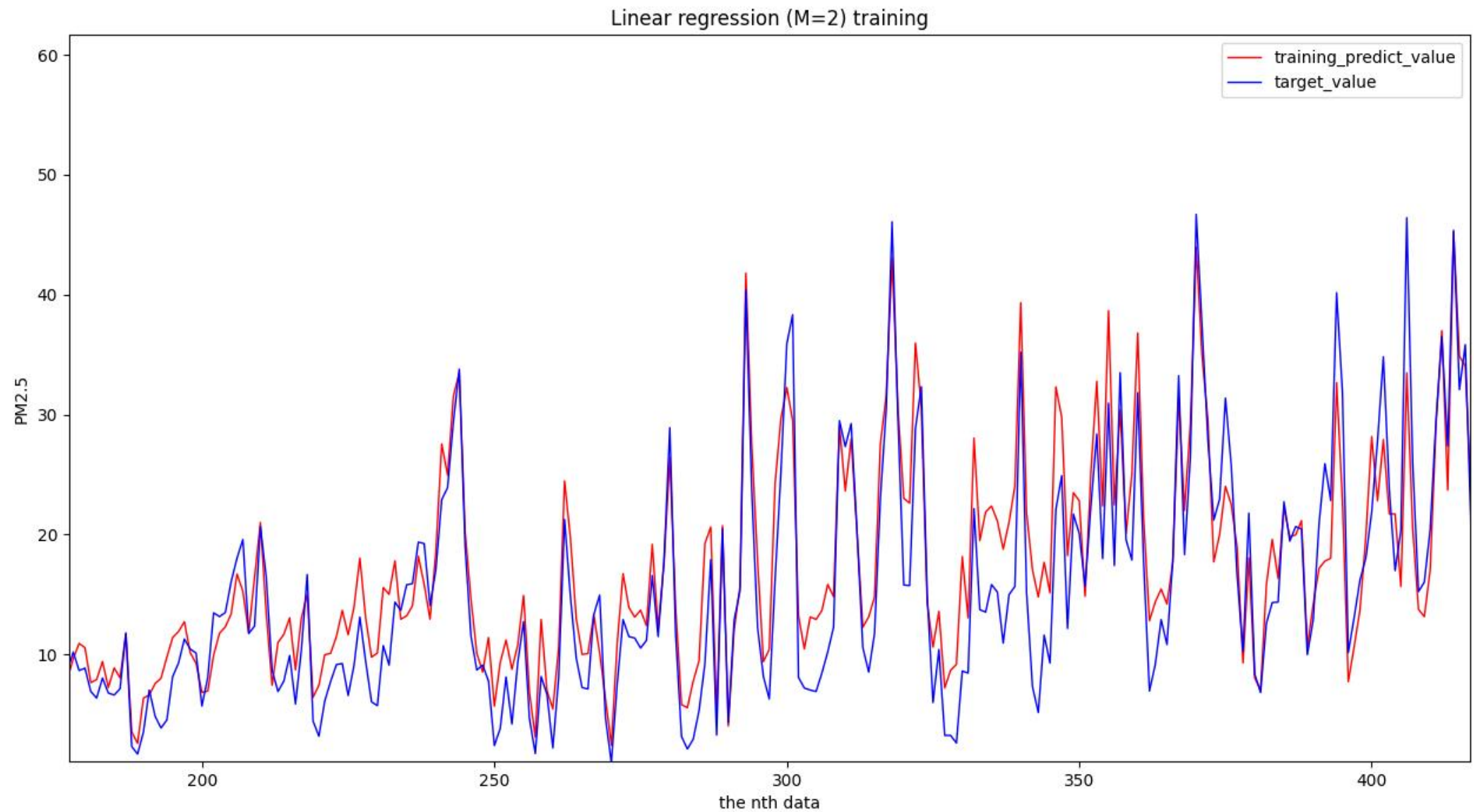
PM2.5(M=2) Outcome

[illegible]

PM2.5(M=2)Outcome



PM2.5(M=2)Outcome



PM2.5(M=2)Outcome

