

KDD CUP 2018

M10615073 李亞修

M10615077 方清鋒

指導教授：李漢銘 教授

Problem Description

You need to predict concentration levels of
PM2.5/PM10/O₃ , over the coming 48 hours for 35 stations in Beijing
PM2.5/PM10, over the coming 48 hours for 13 stations in London

Tools



Use Features

BJ(AQ+MEO)

Station_id

Time

PM25

PM10

NO2

CO

O3

SO2

Temperature

Pressure

Humidity

Wind_speed

Wind_direction

- LD(AQ)

- Station_id

- Time

- PM25

- PM10

- NO2

Missing Data

	A	B	C	D	E	F	G	H
1	station_id	time			station_id	time		
2	beijing_meo	2018/4/22 00:00			beijing_meo	2018/4/22 00:00		
3	chaoyang_meo	2018/4/22 01:00			beijing_meo	2018/4/22 01:00		
4	daxing_meo	2018/4/22 02:00			beijing_meo	2018/4/22 02:00		
5	fangshan_meo	2018/4/22 03:00			beijing_meo	2018/4/22 03:00		
6	fengtai_meo	2018/4/22 04:00			beijing_meo	2018/4/22 04:00		
7	hadian_meo	2018/4/22 05:00			beijing_meo	2018/4/22 05:00		
8	huairou_meo	2018/4/22 06:00			beijing_meo	2018/4/22 06:00		
9	mentougou_me	2018/4/22 07:00			beijing_meo	2018/4/22 07:00		
10	miyun_meo	2018/4/22 08:00			beijing_meo	2018/4/22 08:00		
11	pingchang_mec	2018/4/22 09:00			beijing_meo	2018/4/22 09:00		
12	pinggu_meo	2018/4/22 10:00			beijing_meo	2018/4/22 10:00		
13	shangdianzi_me	2018/4/22 11:00			beijing_meo	2018/4/22 11:00		
14	shijingshan_me	2018/4/22 12:00			beijing_meo	2018/4/22 12:00		
15	shunyi_meo	2018/4/22 13:00			beijing_meo	2018/4/22 13:00		
16	tongzhou_meo	2018/4/22 14:00			beijing_meo	2018/4/22 14:00		
17	xiayunling_mec	2018/4/22 15:00			beijing_meo	2018/4/22 15:00		
18	yanqing_meo	2018/4/22 16:00			beijing_meo	2018/4/22 16:00		
19	zhaitang_meo	2018/4/22 17:00			beijing_meo	2018/4/22 17:00		
20		2018/4/22 18:00			beijing_meo	2018/4/22 18:00		
21		2018/4/22 19:00			beijing_meo	2018/4/22 19:00		
22		2018/4/22 21:00			beijing_meo	2018/4/22 21:00		
23		2018/4/22 22:00			beijing_meo	2018/4/22 22:00		
24		2018/4/22 23:00			beijing_meo	2018/4/22 23:00		
25		2018/4/23 00:00			beijing_meo	2018/4/23 00:00		
26		2018/4/23 01:00			beijing_meo	2018/4/23 01:00		

source

+

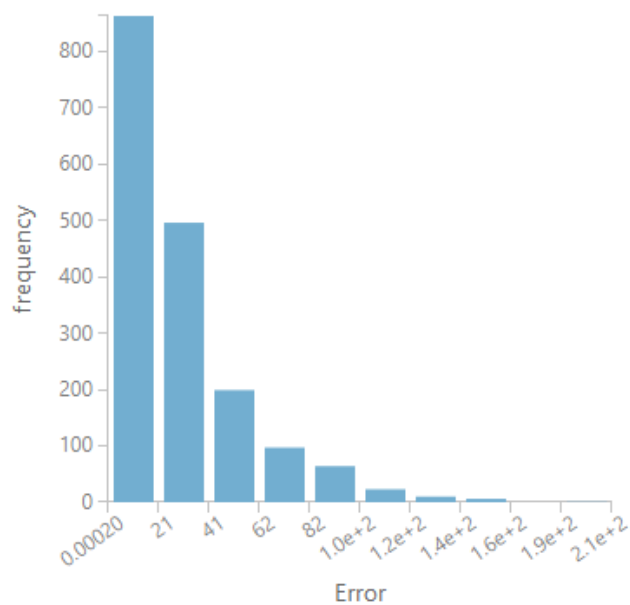
就绪

E	F	G
station_id	time	Scored Labels
beijing_meo	2018/4/22 00:00	120.9276516
beijing_meo	2018/4/22 01:00	51.12714487
beijing_meo	2018/4/22 02:00	22.4392003
beijing_meo	2018/4/22 03:00	39.2027768
beijing_meo	2018/4/22 04:00	77.89494911
beijing_meo	2018/4/22 05:00	
beijing_meo	2018/4/22 06:00	122.6932144
beijing_meo	2018/4/22 07:00	123.5269423
beijing_meo	2018/4/22 08:00	123.3502077
beijing_meo	2018/4/22 09:00	83.32207809
beijing_meo	2018/4/22 10:00	75.8996726
beijing_meo	2018/4/22 11:00	
beijing_meo	2018/4/22 12:00	116.3912609
beijing_meo	2018/4/22 13:00	115.8156483
beijing_meo	2018/4/22 14:00	133.9757379
beijing_meo	2018/4/22 15:00	
beijing_meo	2018/4/22 16:00	144.5904509
beijing_meo	2018/4/22 17:00	146.0321727
beijing_meo	2018/4/22 18:00	100.5775532
beijing_meo	2018/4/22 19:00	
beijing_meo	2018/4/22 21:00	107.0996208
beijing_meo	2018/4/22 22:00	112.5802181
beijing_meo	2018/4/22 23:00	95.05385088
beijing_meo	2018/4/23 00:00	84.12323843
beijing_meo	2018/4/23 01:00	79.7339324

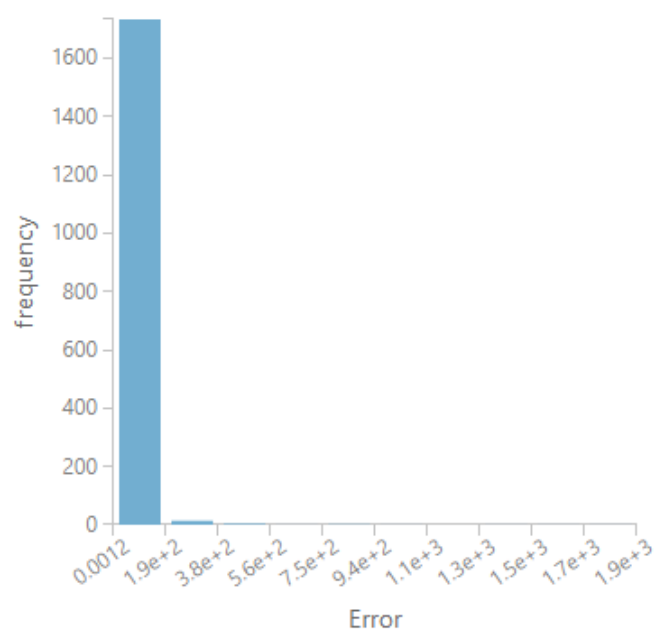
Linear Regression

BJ –PM25,PM10,O3

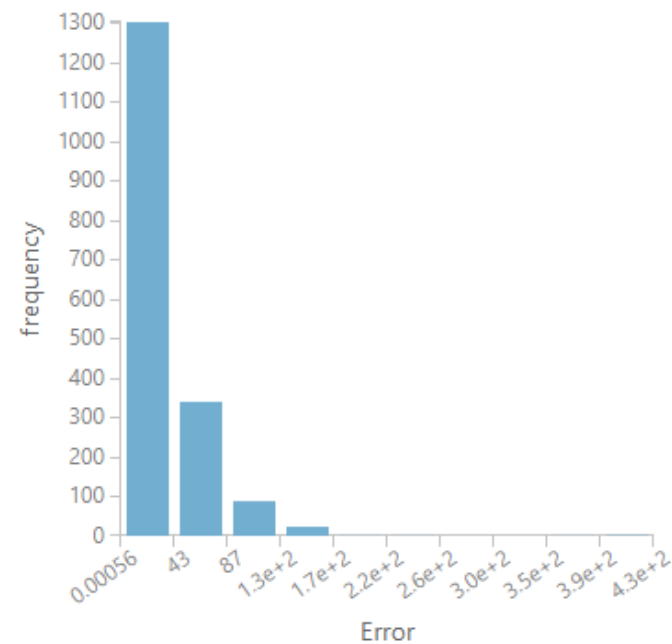
▲ Error Histogram



▲ Error Histogram



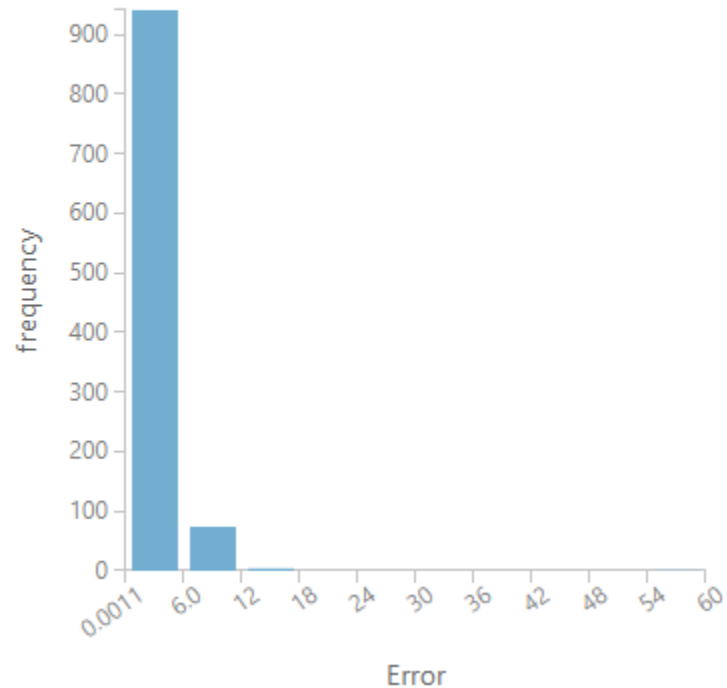
▲ Error Histogram



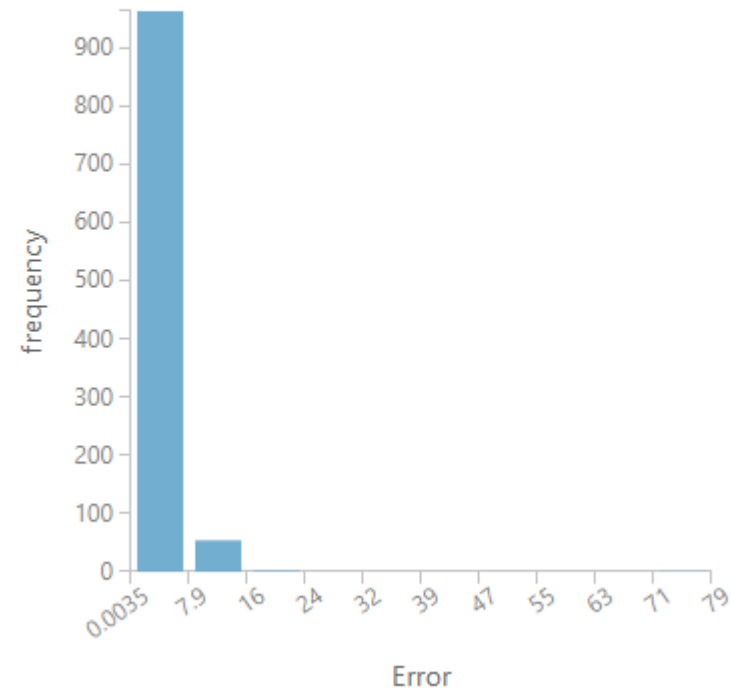
Linear Regression

LD –PM2.5,PM10

▲ Error Histogram



▲ Error Histogram



Future work

Noise reduction

Add more features

Missing data processing

Thanks!