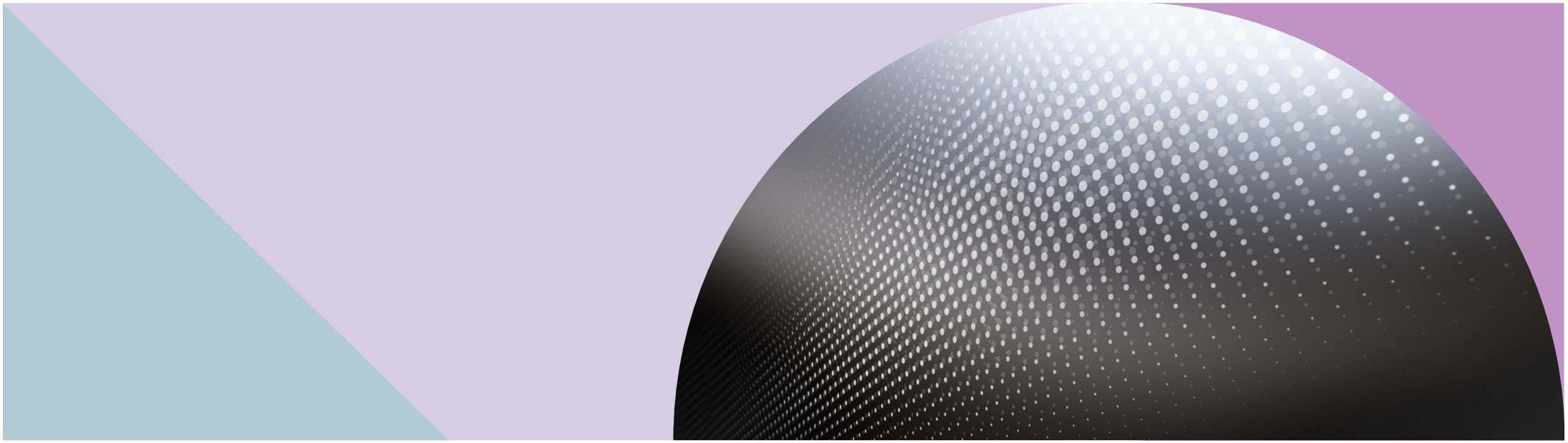
Chemical Analysis Machine Learning Model

Sameer Mohammed Al Harbi



بسم الله الرحمن الرحيم

اللهم لا سهلا الا ما جعلته سهلا وانت تجعل الحزن اذا شئت سهلا

Introduction

- The geologists work so hard in order to collect a valuable geological samples from the field.
- The Samples are brought to laboratories for chemical analysis.

The Problem

Some time the result is missing by chemecal analysis device

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
0.5	0.02	5.5	234.2		33.89	134.11	111.77	2.02	11.14	38.43	37		228.64	54.85	6.5
0.5	0.02	5.5	185.41		24.58	108.11	76.72	3.07	12.14	18.87	9		216.29	54.23	11.18
0.5	0.02	5.5	135.94		26.54	117.66	76.57	3.22	12.01	19.63	10		224.2	61.2	10.59
0.5	0.02	5.5	149.31		26.62	123.62	82.46	3.02	11.86	20.81	7		231.28	58.28	10.65
0.5	0.02	5.5	169.88		36.72	128.06	149.43	3.5	14.18	34.05	37		270.93	60.29	5.94
0.5	0.02	5.5	121.28		24.78	112.2	58.31	5.05	13.66	20.42	16		231.65	50.89	10.38
0.5	0.02	5.84	118.29		32.19	68.84	135.82	5.99	15.2	34.34	41		275.15	58.03	6.35
0.5	0.02	7.52	229.66		28.1	76.96	78.82	5.02	14.91	21.8	16		245.72	63	11.69
0.5	0.02	5.5	221.46		26.06	70.66	92.07	4.77	13.77	20.62	13		231.37	61.67	11.63
0.5	0.02	5.5	226.19		28.21	63.53	69.04	4.78	14	20.22	35		222.43	59.37	12.36
0.5	0.02	6.75	245.16		26.88	78.91	78.2	4.87	14.21	20.86	14		241.13	64.82	12.54
0.5	0.02	5.51	250.84		27.54	86.28	91.67	4.92	14.53	21.47	13		246.74	67.07	13.18
0.5	0.02	5.5	164.08		25.5	107.22	63.35	4.53	13.88	19.97	26		227.2	60.17	12.11
0.5	0.02	5.5	183.97		25.73	116.93	79.16	4.81	13.99	19.69	13		241.51	61.13	13.26
0.5	0.02	7.45	169.73		29.47	127.2	118.98	4.84	14.5	22.6	14		253.19	58.67	12.39
0.5	0.02	7.1	149		30.47	117.87	100.87	4.71	14.17	21.77	19		251.2	55.43	12.21
0.5	0.02	5.5	105.52		26.32	90.78	94.4	4.72	14.49	20.41	17		238.13	50.27	13.08
0.5	0.02	6.56	111.64		44.47	109.54	170.52	3.95	21.82	29	16		347.27	62.67	6.18
0.5	0.02	5.5	129.83		27.91	79.12	116.12	4.88	13.81	20.63	18		239.98	66.41	12.92
0.5	0.02	5.5	171.93		27.31	109.47	85.55	4.46	14.01	24.09	16		244.88	69.91	11.27
0.5	0.02	5.64	178.08		28.41	99.16	80.11	5.16	14.49	21.01	74		246.11	61.62	12.24
0.5	0.02	5.5	128.08		28.06	119.02	105.13	4.92	14.44	21.39	16		246.8	59.08	12.95
0.5	0.02	5.91	134.25		34.46	123.28	128.3	5.19	15.7	22.8	14		252.74	58.55	14.13
0.5	0.02	7.06	51.77		27.65	111.18	89.4	4.58	13.61	20.62	46		228.43	51.49	13.22
0.5	0.02	6.8	385.86		29.07	113.8	113.24	4.73	14.21	20.22	14		231.24	51.07	15.24
0.5	0.02	7.55	88.41		25.32	105.2	79.33	4.75	13.46	19.66	15		235.48	53.33	15.24

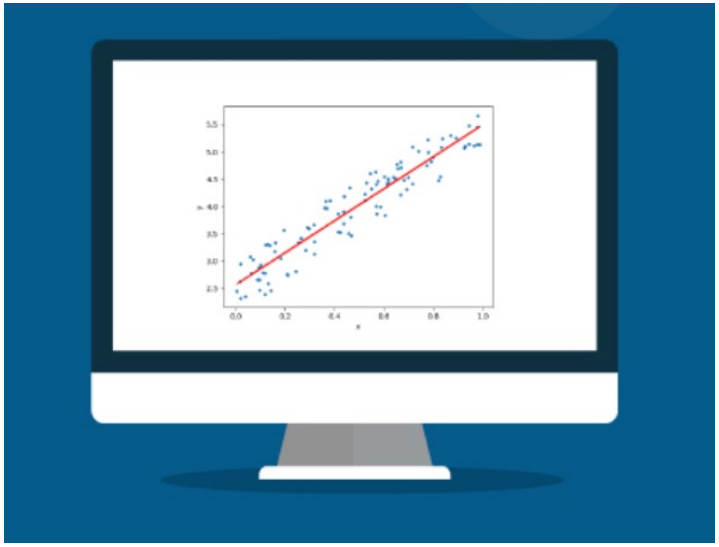
What else

Sometimes the analysis devices throw negative values

0.5	0.02	5.5	163.72		49.06	384.32	73.08	4.26	10	123.1	17		328.31	61.96	9.68
0.5	0.02	5.5	120.93		55.27	799.64	6.65	3.11	10	268.3	15		291.39	67.86	4.13
0.5	0.02	5.5	138.46		30.31	114.69	15.64	3.52	10	42.3	19		293.84	48.41	9.45
0.5	0.02	5.5	86.78		45.43	98.74	73.97	5	10	35.56	16		344.97	59.43	15.43
0.5	0.02	5.5	88.95		45.22	210.82	140.11	4.43	10	82.5	15		336.05	77.62	11.01
0.5	0.02	5.5	195.86		65.13	636.26	192.46	3.67	10	238.89	13		342.91	76.67	5.03
0.5	0.02	5.5	193.88		59.7	646.39	123.58	3.53	10	227.6	14		330.58	74.57	5.55
0.5	0.02	5.5	176.62		48.35	457.88	142.1	4.4	10	143.55	15		350.74	72.15	8.37
0.5	0.02	5.5	197.26		50.63	591.75	144.38	3.93	10	176.94	16		330.15	91.22	8.11
0.5	0.02	5.5	151.45		56.61	756.08	163.46	3.63	10	231.16	15		318.44	78.55	5.97
0.5	0.02	5.5	130.19		44.95	424.34	171.29	4.23	10	116.55	17		344.42	68.7	9.26
0.5	0.02	5.5	74.32		33.96	142.95	129.73	5.36	10	32.97	15		302.71	62.77	14.38
0.5	0.02	5.5	95.13		35.03	149.21	78.68	5.66	10	33.64	16		305.72	68.55	15.34
1	0.04	-50	283	23	25	34	1076	-20	-20	16	-10	204	239	104	-250
-1	0.05	51	370	47	-5	-10	35	22	-20	-10	-10	151	91	35	-250
-1	0.05	-50	497	47	10	-10	37	22	-20	-10	-10	202	99	66	-250
1.1	0.12	-50	754	49	22	11	819	23	-20	29	34	155	123	588	-250
1	0.22	82	419	27	26	33	563	-20	-20	19	86	166	269	631	-250
1.2	1	-50	238	28	15	20	3998	-20	-20	16	192	214	123	1650	-250
1	2.01	51	360	27	9	-10	3038	-20	-20	16	632	200	121	1556	-250
-1	0.57	61	330	22	9	-10	4119	-20	-20	14	40	173	56	656	-250
-1	0.43	128	168	34	9	-10	1815	-20	-20	11	20	183	57	218	-250
-1	0.18	191	214	36	9	-10	134	-20	-20	10	-10	113	40	145	-250
-1	0.38	-50	183	35	6	14	300	-20	-20	16	-10	106	56	95	-250
-1	0.1	55	198	37	8	-10	609	-20	-20	16	10	147	47	169	-250
-1	0.08	-50	220	32	8	-10	82	-20	-20	13	-10	167	46	178	-250

The Solution

- Using machine learning AI to predict missing values and fill the gabs



First Step Clean The Data

First step is cleaning the data and isolate the good and complete data rows from the the rows with null or nigative values.

Also, we need to keep the the rows with missing data in order to be tested with our model.

Next Creating The AI Model

Our problem is numercal problem so I prefere to use Liner regression model in order to solve the issue.