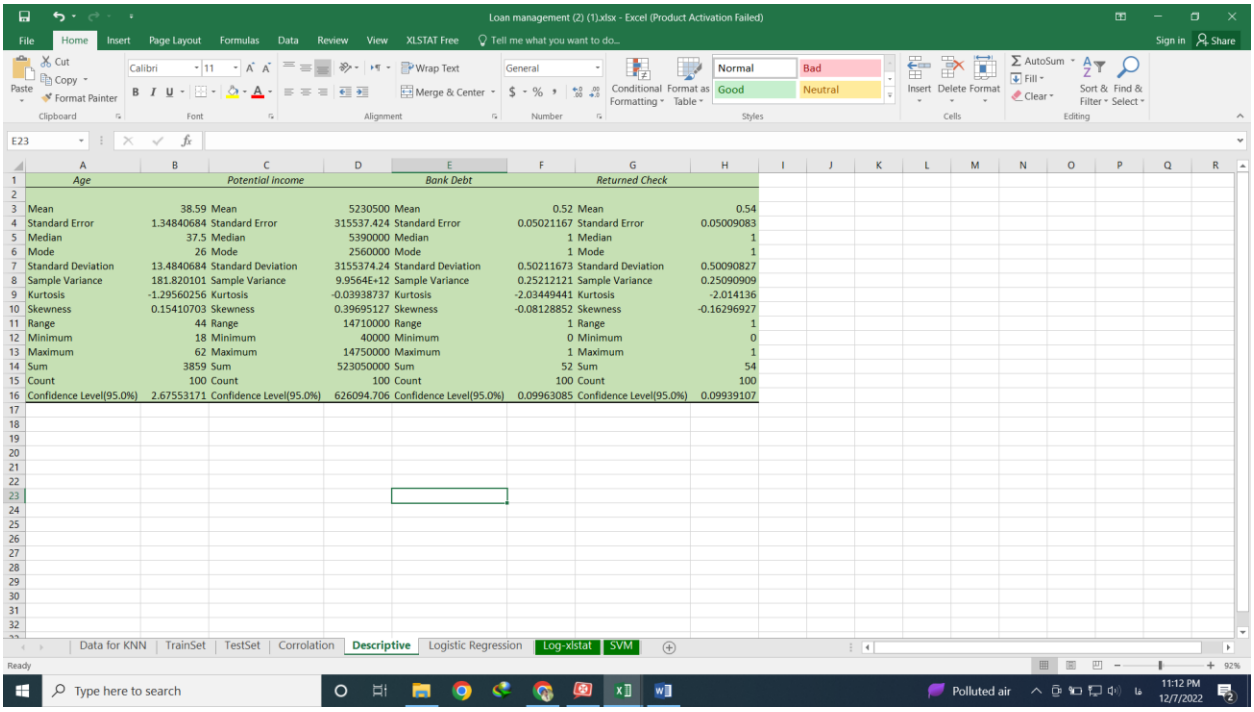
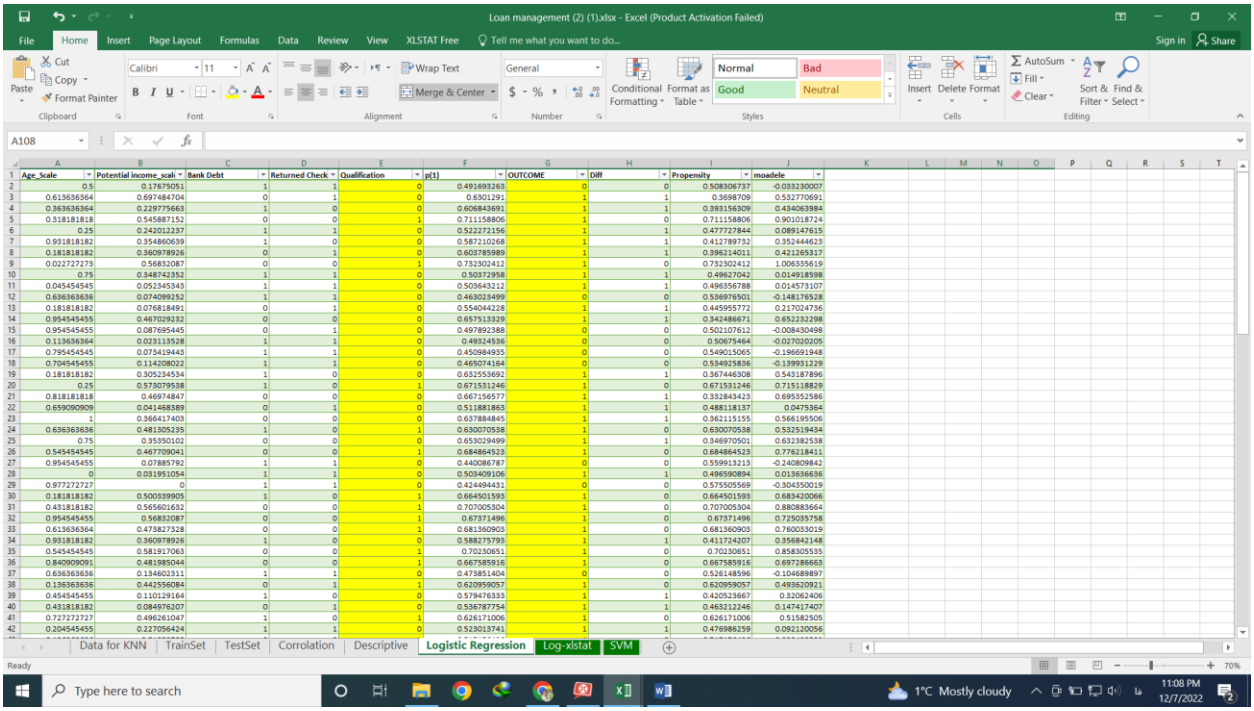
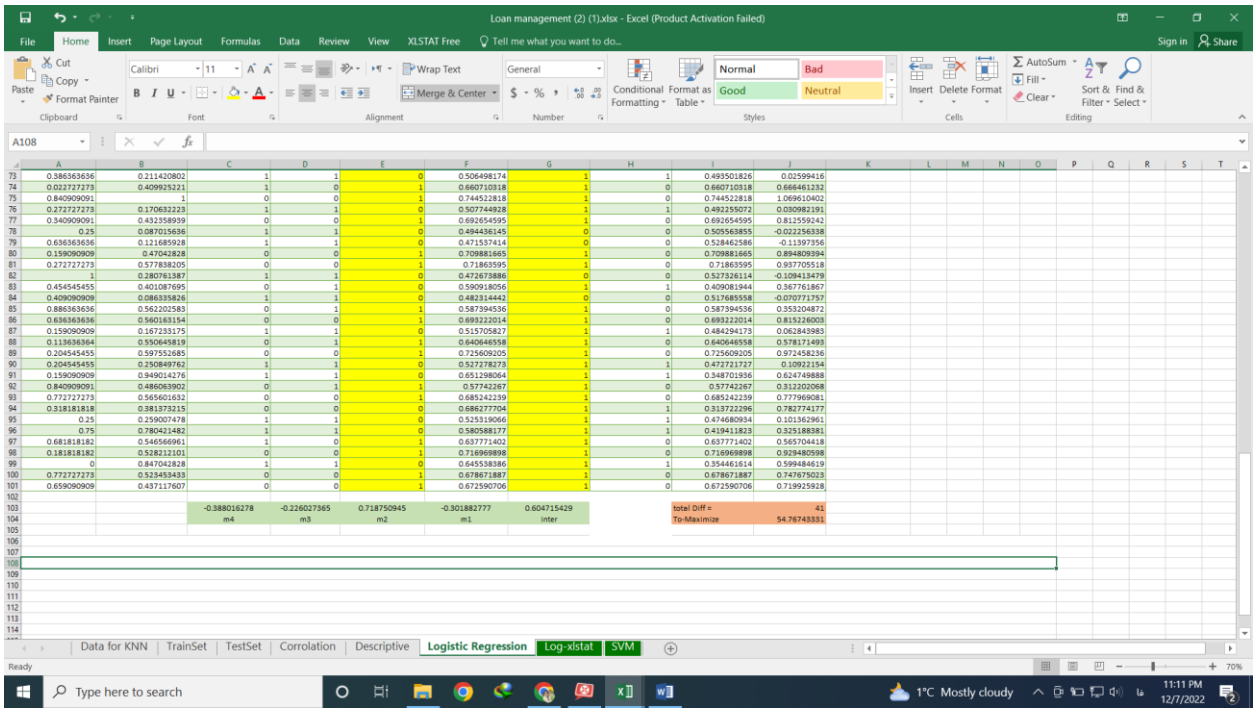


Age		Potential income		Bank Debt		Returned Check	
Mean	38.59	Mean	5230500	Mean	0.52	Mean	0.54
Standard Error	1.34840684	Standard Error	315537.424	Standard Error	0.05021167	Standard Error	0.05009083
Median	37.5	Median	5390000	Median	1	Median	1
Mode	26	Mode	2560000	Mode	1	Mode	1
Standard Deviation	13.4840684	Standard Deviation	3155374.24	Standard Deviation	0.50211673	Standard Deviation	0.50090827
Sample Variance	181.820101	Sample Variance	9.9564E+12	Sample Variance	0.25212121	Sample Variance	0.25090909
Kurtosis	-1.29560256	Kurtosis	-0.03938737	Kurtosis	-2.03449441	Kurtosis	-2.014136
Skewness	0.15410703	Skewness	0.39695127	Skewness	-0.08128852	Skewness	-0.16296927
Range	44	Range	14710000	Range	1	Range	1
Minimum	18	Minimum	40000	Minimum	0	Minimum	0
Maximum	62	Maximum	14750000	Maximum	1	Maximum	1
Sum	3859	Sum	523050000	Sum	52	Sum	54
Count	100	Count	100	Count	100	Count	100
Confidence Level(95.0%)	2.67553171	Confidence Level(95.0%)	626094.706	Confidence Level(95.0%)	0.09963085	Confidence Level(95.0%)	0.09939107

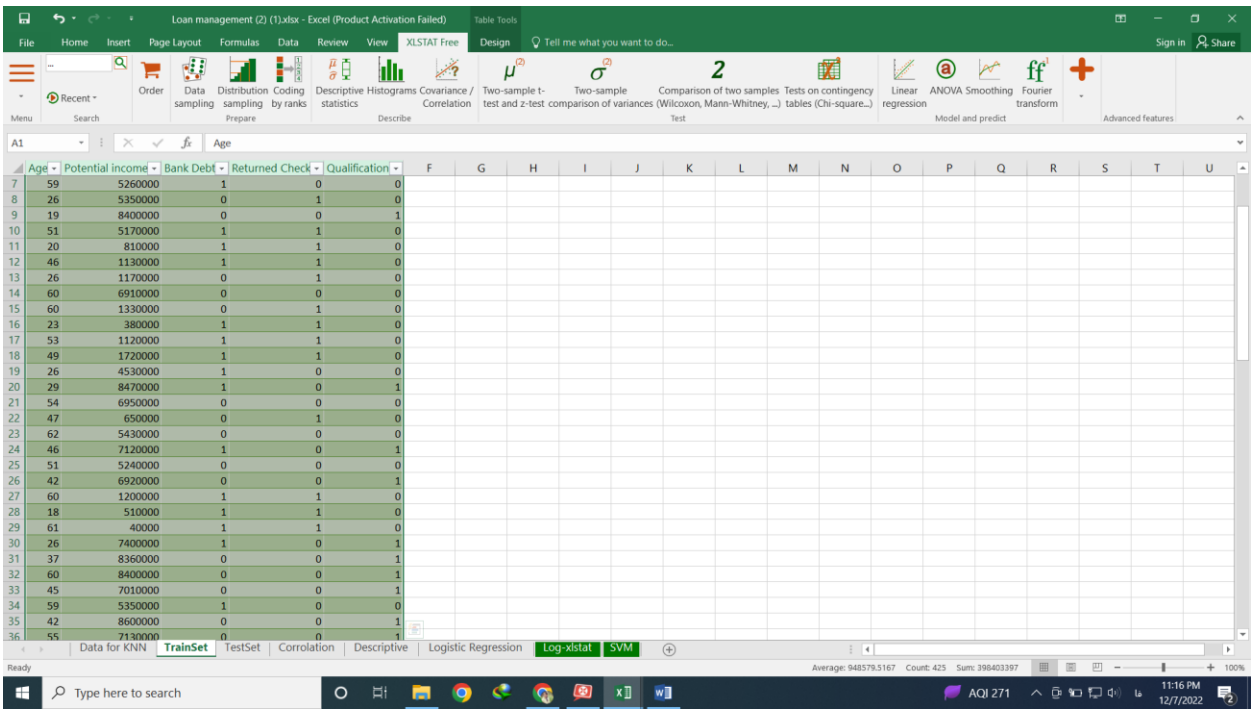


در این مسئله هم ابتدا گام های پیش پردازش داده انجام شده و داده ها هم مقیاس میشوند . سپس چون این مسیله از نوع طبقه بندی میباشد باید از روش های طبقه بندی مانند linear reg, svm و... استفاده نمود

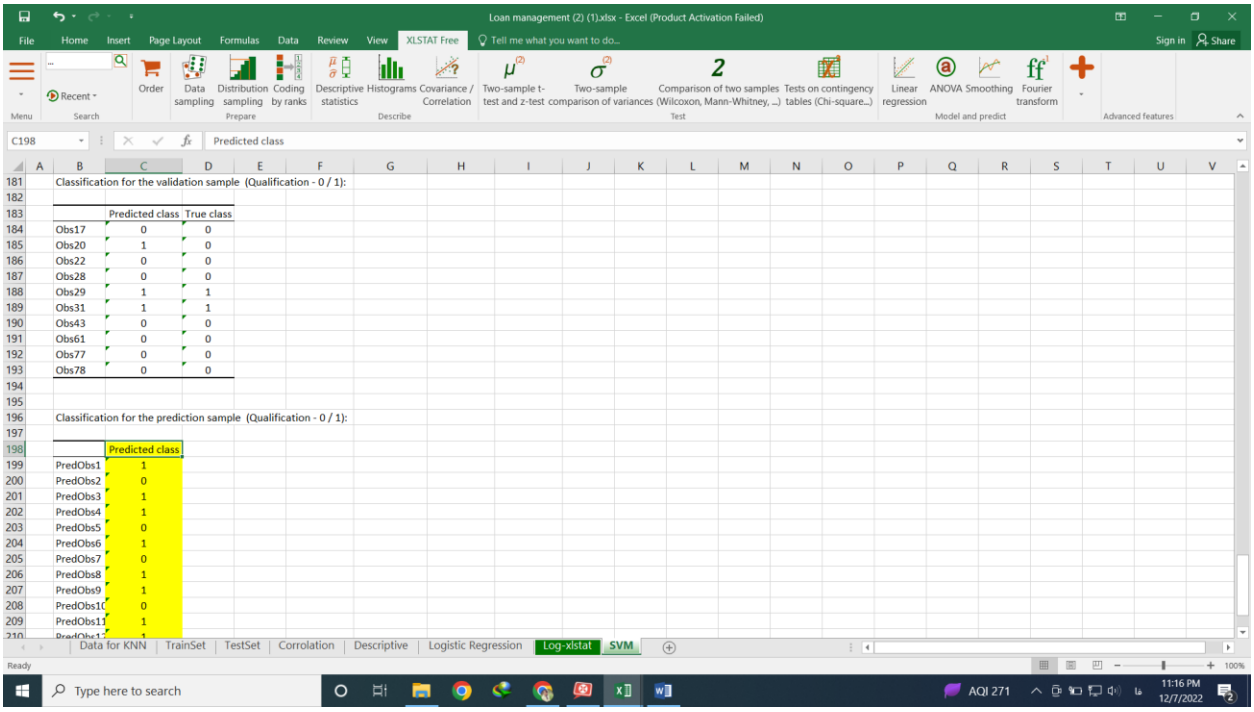
بعد محاسبه شیب خط ها از تابع linest و محاسبه ی معادله به جمع 41خطا ازمجموع ۱۰۱ داده میرسیم



سپس از طریق افزونه xlstat اقدام به دسته بندی و طبقه بندی دادگان مینماییم در این تمرین ما ۸۵ داده را بصورت رندوم بعوان مجموعه آموزش و ۱۶ داده را به عنوان مجموعه test در نظر گرفته ایم که بعد از اعمال روش های دسته بندی نتایج ذیل حاصل شد که حاکی از درصد درستی بالای مدل میباشد



طبقه بندی svm



Loan management (2) (1).xlsx - Excel (Product Activation Failed)

Classification for the validation sample (Qualification - 0 / 1):

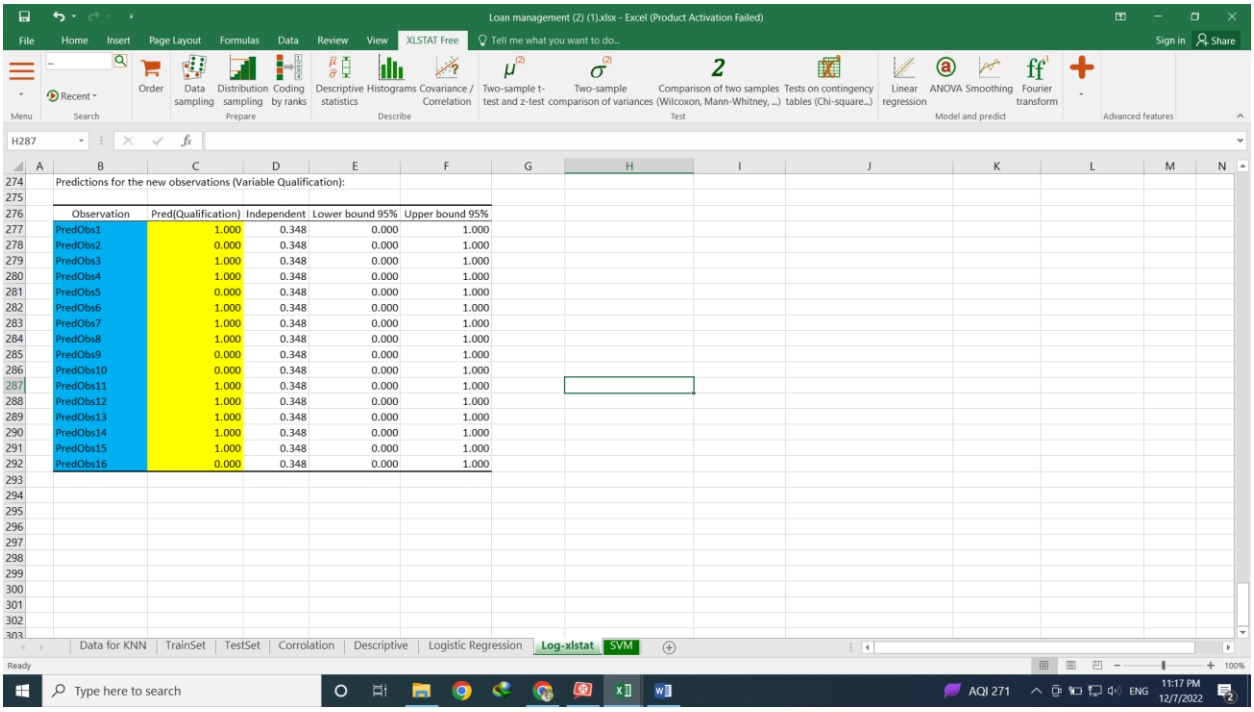
	Predicted class	True class
Obs17	0	0
Obs20	1	0
Obs22	0	0
Obs28	0	0
Obs29	1	1
Obs31	1	1
Obs43	0	0
Obs61	0	0
Obs77	0	0
Obs78	0	0

Classification for the prediction sample (Qualification - 0 / 1):

	Predicted class
PredObs1	1
PredObs2	0
PredObs3	1
PredObs4	1
PredObs5	0
PredObs6	1
PredObs7	0
PredObs8	1
PredObs9	1
PredObs10	0
PredObs11	1

Data for KNN | TrainSet | TestSet | Corrolation | Descriptive | Logistic Regression | Log-xlstat | SVM

طبقه بندی linear regression



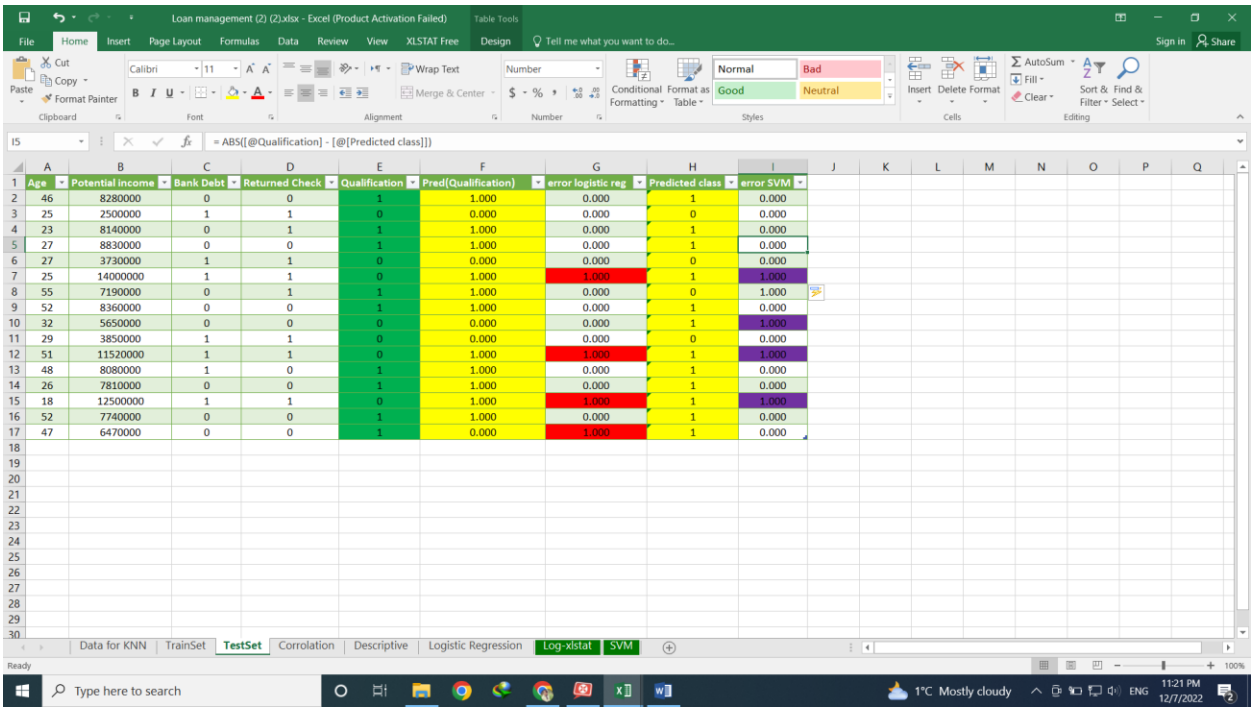
Loan management (2) (1).xlsx - Excel (Product Activation Failed)

Predictions for the new observations (Variable Qualification):

Observation	Pred(Qualification)	Independent	Lower bound 95%	Upper bound 95%
PredObs1	1.000	0.348	0.000	1.000
PredObs2	0.000	0.348	0.000	1.000
PredObs3	1.000	0.348	0.000	1.000
PredObs4	1.000	0.348	0.000	1.000
PredObs5	0.000	0.348	0.000	1.000
PredObs6	1.000	0.348	0.000	1.000
PredObs7	1.000	0.348	0.000	1.000
PredObs8	1.000	0.348	0.000	1.000
PredObs9	0.000	0.348	0.000	1.000
PredObs10	0.000	0.348	0.000	1.000
PredObs11	1.000	0.348	0.000	1.000
PredObs12	1.000	0.348	0.000	1.000
PredObs13	1.000	0.348	0.000	1.000
PredObs14	1.000	0.348	0.000	1.000
PredObs15	1.000	0.348	0.000	1.000
PredObs16	0.000	0.348	0.000	1.000

Data for KNN | TrainSet | TestSet | Corrolation | Descriptive | Logistic Regression | Log-xlstat | SVM

نمای کلی از نتایج اکتسابی



Age	Potential income	Bank Debt	Returned Check	Qualification	Pred(Qualification)	error logistic reg	Predicted class	error SVM
46	8280000	0	0	1	1.000	0.000	1	0.000
25	2500000	1	1	0	0.000	0.000	0	0.000
23	8140000	0	1	1	1.000	0.000	1	0.000
27	8830000	0	0	1	1.000	0.000	1	0.000
27	3730000	1	1	0	0.000	0.000	0	0.000
25	14000000	1	1	0	1.000	1.000	1	1.000
55	7190000	0	1	1	1.000	0.000	0	1.000
52	8360000	0	0	1	1.000	0.000	1	0.000
32	5650000	0	0	0	0.000	0.000	1	1.000
29	3850000	1	1	0	0.000	0.000	0	0.000
51	11520000	1	1	0	1.000	1.000	1	1.000
48	8080000	1	0	1	1.000	0.000	1	0.000
26	7810000	0	0	1	1.000	0.000	1	0.000
18	12500000	1	1	0	1.000	1.000	1	1.000
52	7740000	0	0	1	1.000	0.000	1	0.000
47	6470000	0	0	1	0.000	1.000	1	0.000

خطاهای مدل برحسب خطای MSE ارزیابی شده است.