



دانشگاه صنعتی خواجه نصیرالدین طوسی

به نام خدا

دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده برق

تشخیص و شناسایی خطا ارائهی پروژه پایانی

استاد درس: دکتر مهدی علیاری شوره‌دلی

به کوشش: جواد حسن‌پور سنگلجی – علیرضا انصاری

مقدمه

در راستای اعتلا بخشیدن به مفاهیم مطرح شده‌ی درس تشخیص و شناسایی خطا، تصمیم بر آن شد تا به عنوان پروژه‌ی پایانی درس، به تحلیل و بررسی داده‌های اخذ شده از توربین باد واقعی پرداخته شود. این داده‌ها حاوی اطلاعاتی هم‌چون دمای قسمت‌های مختلف توربین باد همانند پره‌ها و کابینت‌های کنترل و ...، سرعت باد و اطلاعاتی مربوط به توان الکتریکی است. داده‌های استفاده شده برای این پروژه، شامل ۶۲ بعد و ۵ نوع خطا می‌باشد. همچنین تعداد کل داده‌ها نیز قریب به ۵۰/۰۰۰ است.

هدف

همانطور که گفته شد، داده‌های اخذ شده دارای ۶۲ بعد می‌باشند. به منظور طراحی یک جداکننده، با استفاده از روش‌های مطرح شده در کلاس، ابتدا به کاهش بعد داده‌های اخذ شده می‌پردازیم. (روش استفاده شده جهت کاهش بعد برای این پروژه، روش LDA می‌باشد). سپس با استفاده از یک جداکننده خطی (Linear Classifier) و یک جداکننده غیرخطی (Support Vector Machine) اقدام به جداسازی کلاس‌ها از یکدیگر به منظور تشخیص نوع خطا، می‌کنیم.

ساختار داده‌ها

تمامی ۴۹۰۲۸ داده‌ی اخذشده از سیستم به همراه تاریخ و زمان داده‌برداری، در فایل `scada_data.csv` موجود می‌باشد. همچنین تعداد ۵۵۴ داده‌ی خطا نیز به همراه تاریخ و زمان وقوع، در فایل `scada_378650_748912_fault_data.csv` قرار دارد. با استفاده از کد `data_create.m`، داده‌های خطا را به کل داده‌ها ملحق کرده و بر اساس نوع خطای ایجاد شده، عددی بین ۱ تا ۶ (کلاس ۶ همان کلاس Normal) را بعنوان ستون ۶۳ ام و ستون `target` در نظر می‌گیریم.

Classification Methods

```
graph TD; CM[Classification Methods] --> LS[L.S]; CM --> NN[N.N]; CM --> KNN[K-NN]; CM --> SVM[S.V.M]; NN --> MLP[MLP]; NN --> RBF[RBF]; SVM --> Poly[Polynomial]; SVM --> Linear[Linear];
```

L.S

N.N

K-NN

S.V.M

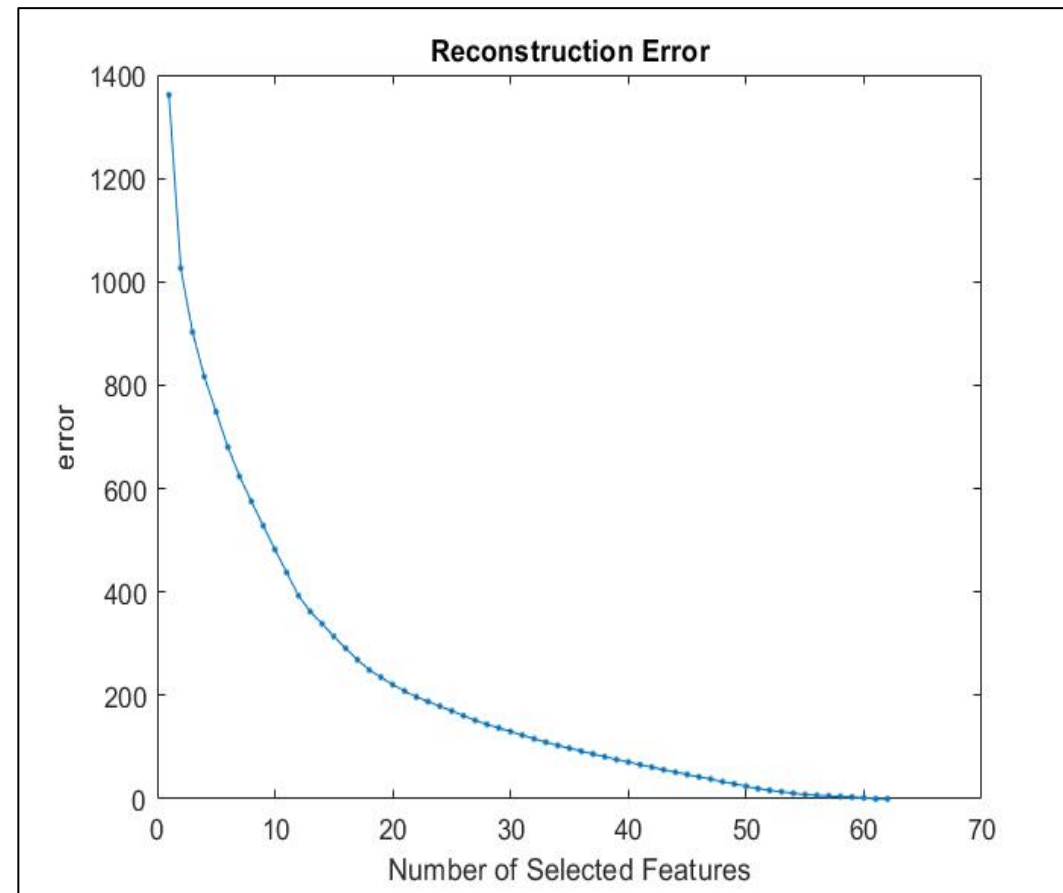
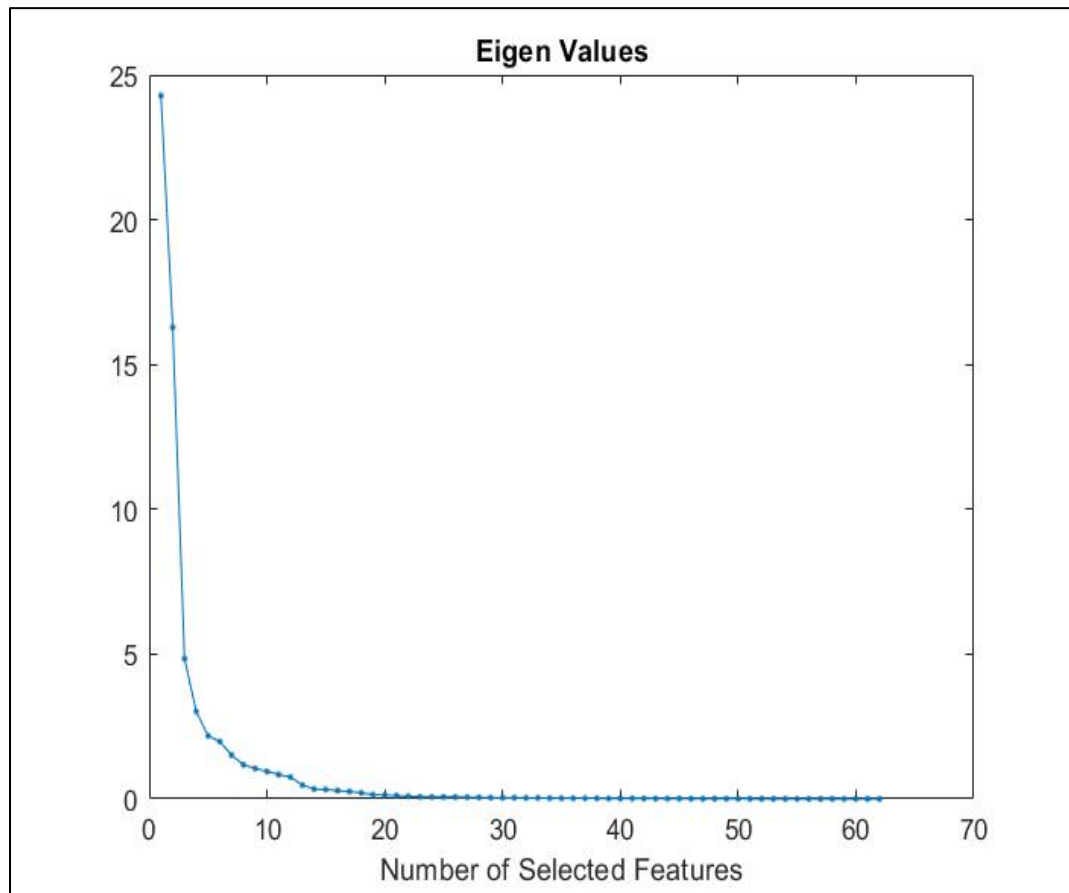
MLP

RBF

Polynomial

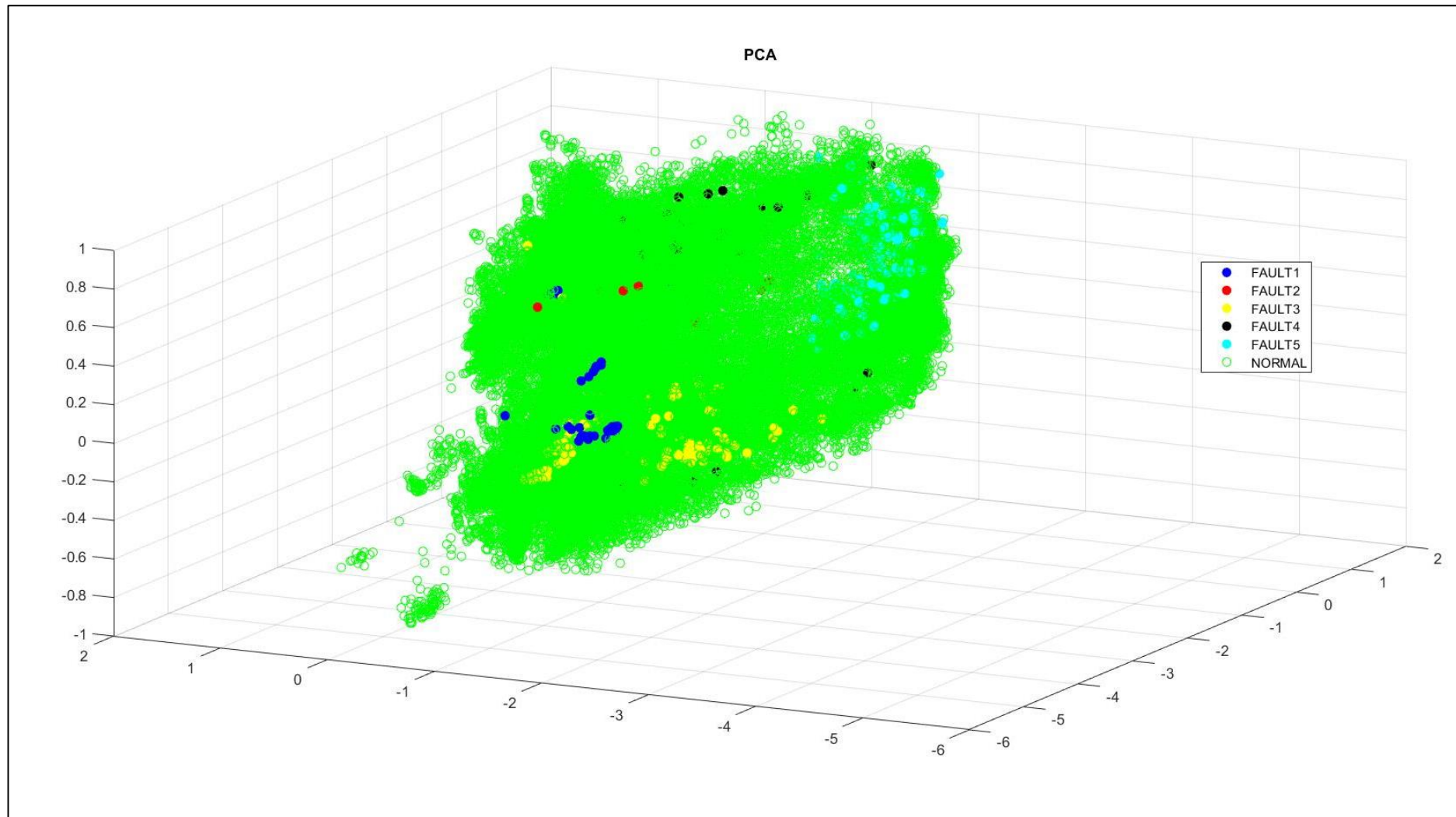
Linear

کاهش بعد PCA

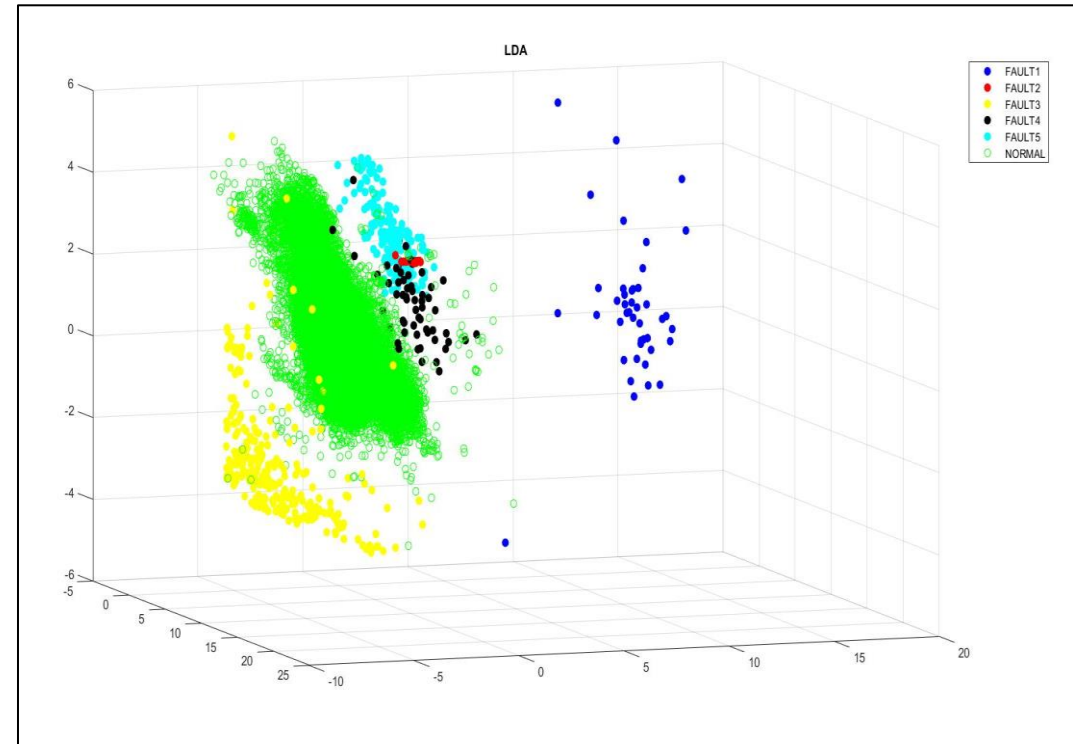
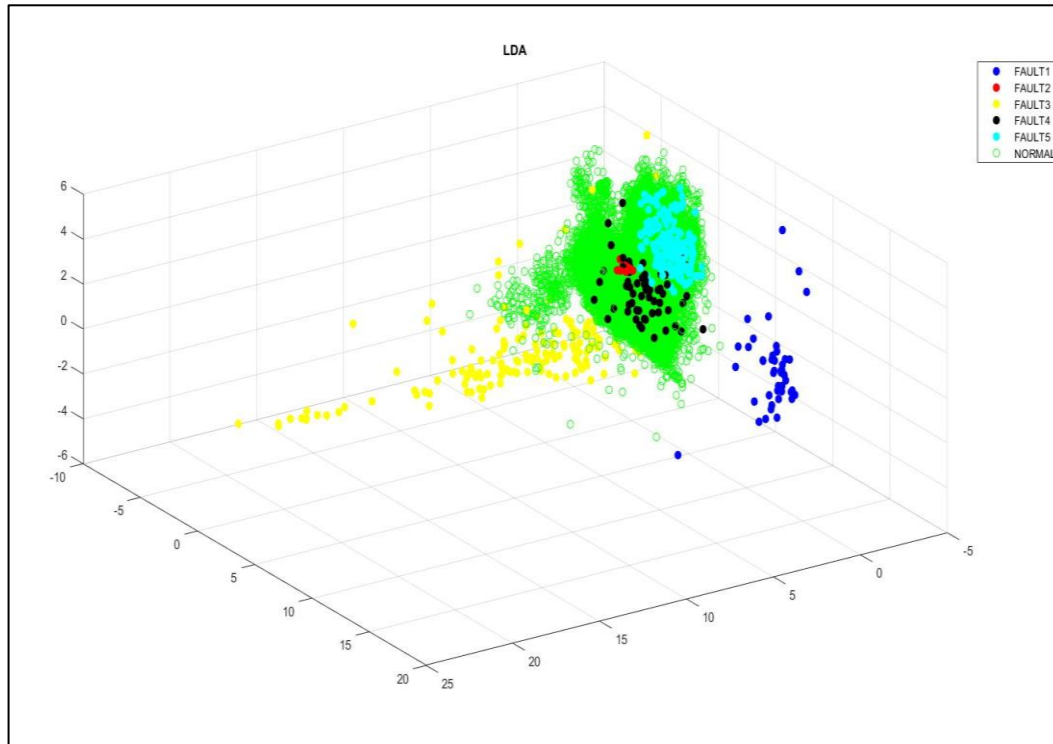


با توجه به نمودار مقادیر ویژه، تعداد Feature انتخاب شده، ۲۰ عدد (۲۰ مقدار ویژه بزرگتر) است.

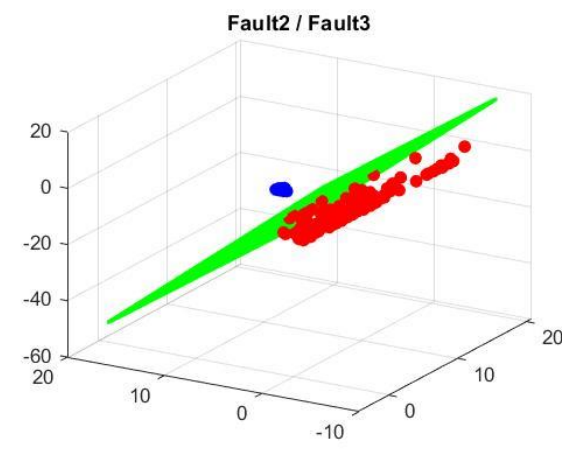
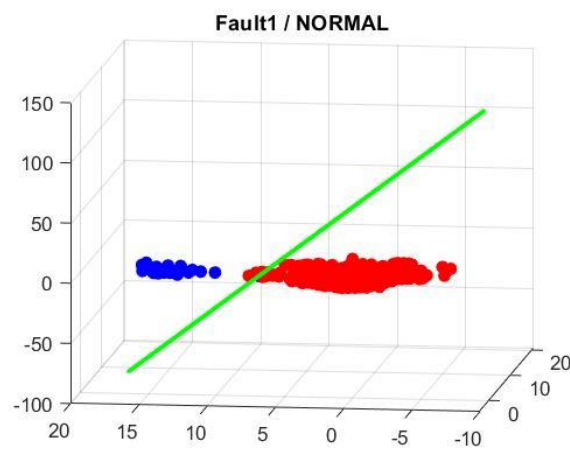
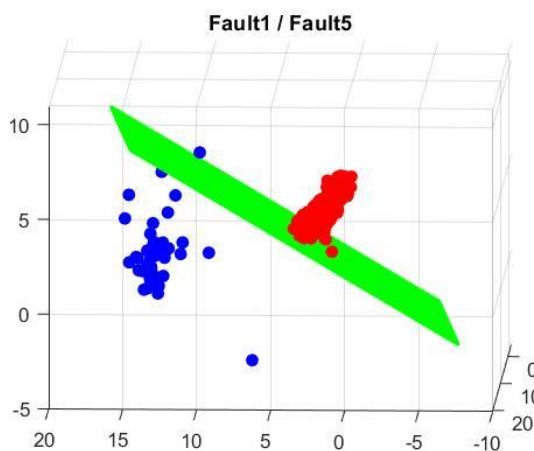
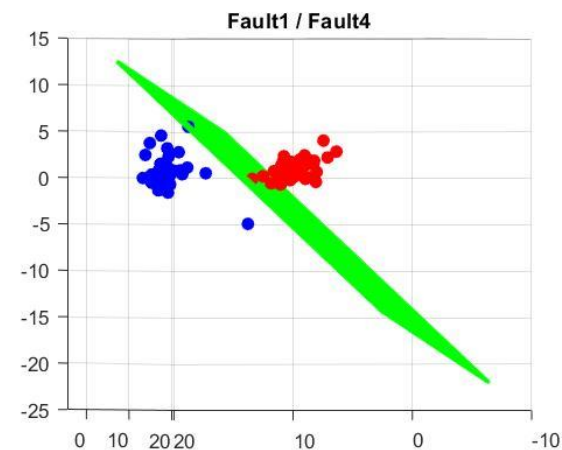
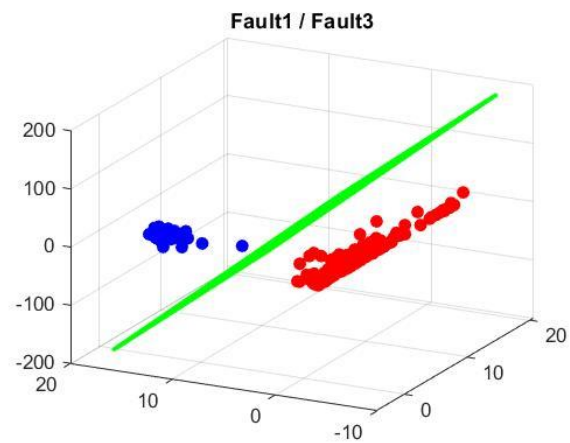
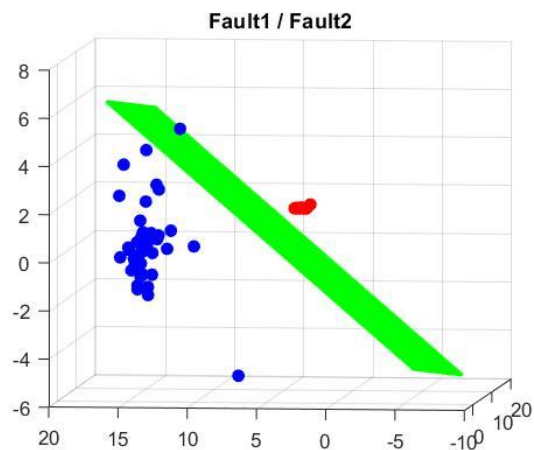
کاهش بعد PCA



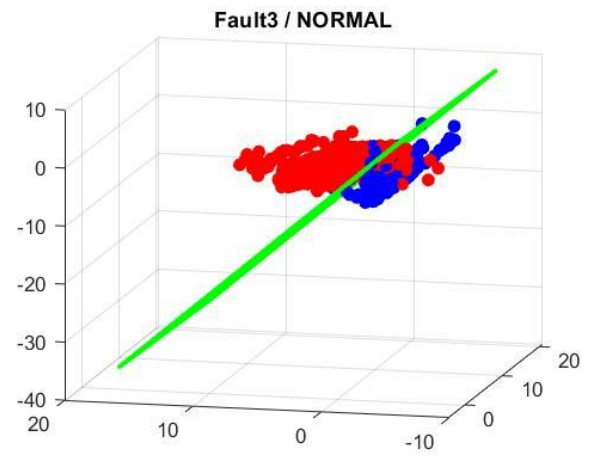
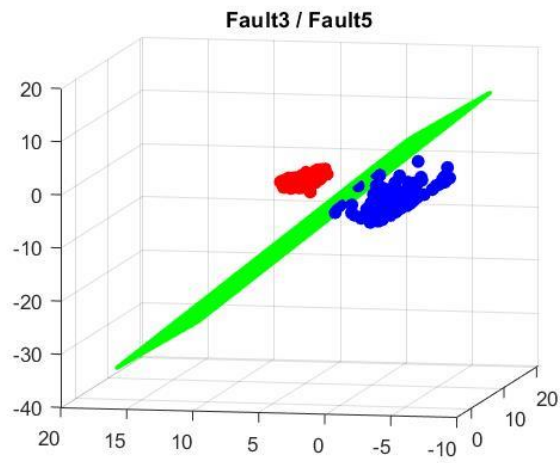
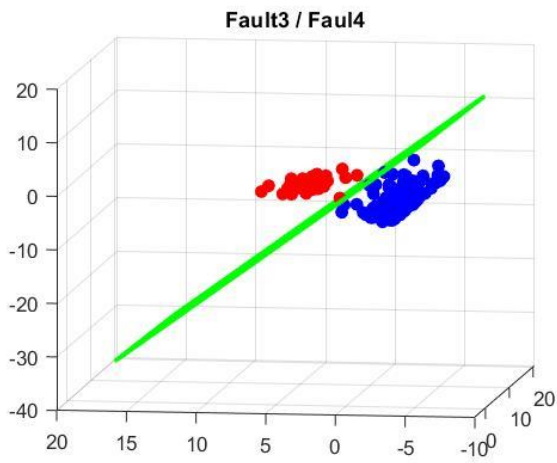
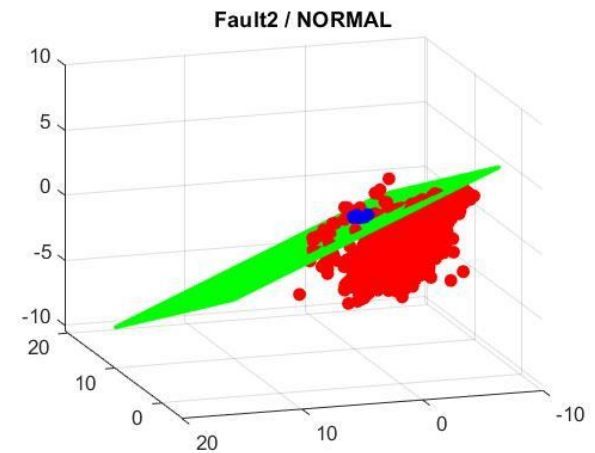
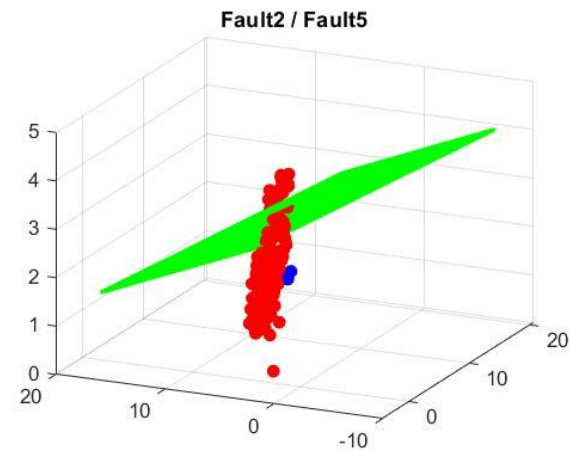
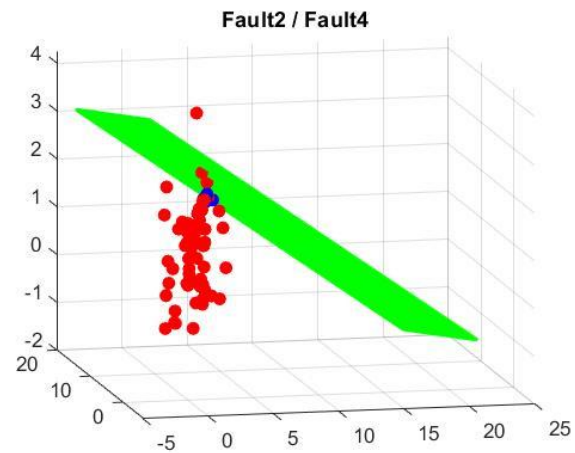
کاهش بعد LDA



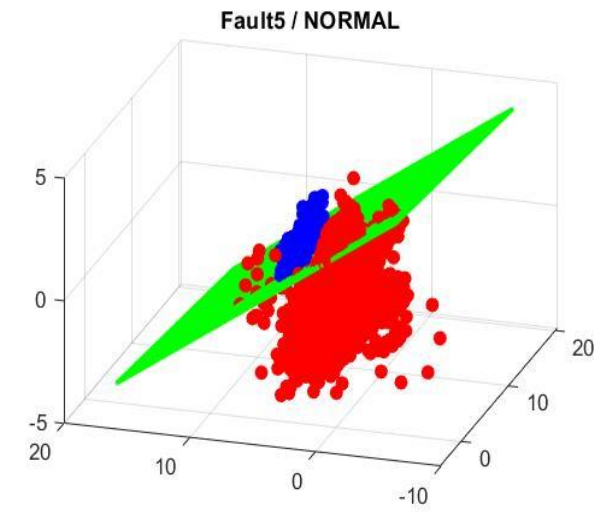
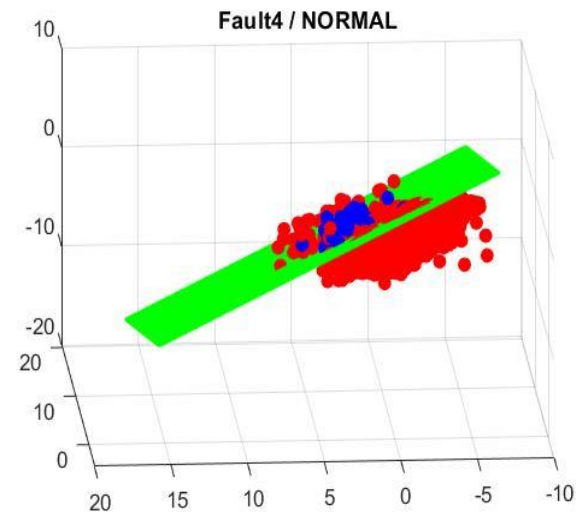
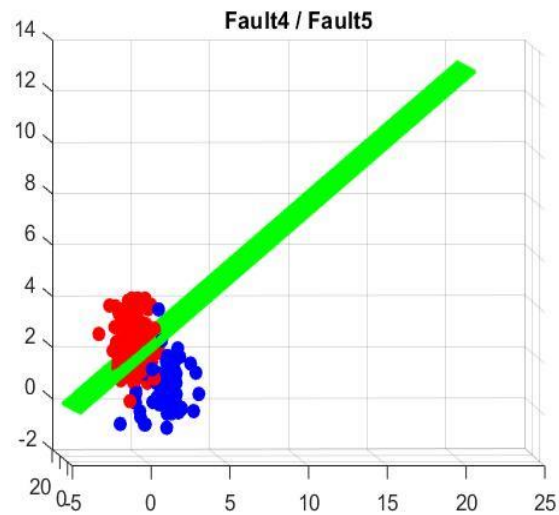
Least Square



Least Square



Least Square



Least Square

پارامترهای خطای برای Train

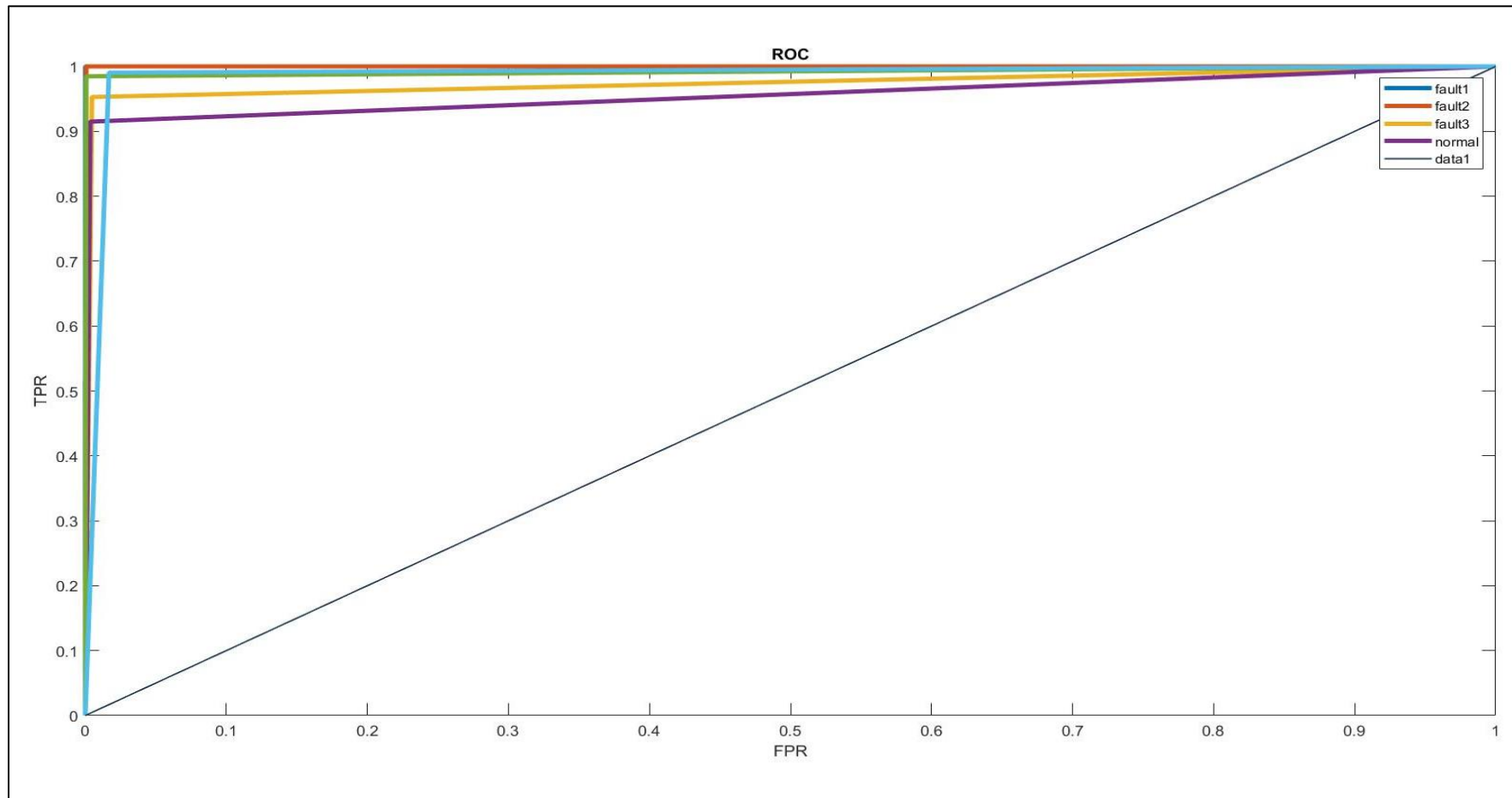
<u>error_unsupervised</u>	<u>error_supervised</u>	<u>performance</u>
0.0098773	0.026179	0.99012

	<u>FAR</u>	<u>MAR</u>	<u>Sensitivity</u>	<u>Specificity</u>	<u>Accuracy</u>
fault 1	8.1477e-05	0	1	0.99992	0.99992
fault 2	0.00070581	0	1	0.99929	0.99929
fault 3	0.0047462	0.04712	0.95288	0.99525	0.99503
fault 4	0.003505	0.085106	0.91489	0.9965	0.99639
fault 5	0.00059911	0.015267	0.98473	0.9994	0.99935

	<u>Fault1</u>	<u>Fault2</u>	<u>Fault3</u>	<u>Fault4</u>	<u>Fault5</u>	<u>Normal</u>	<u>MissClass</u>
Fault1	32	0	0	0	0	0	0
Fault2	0	15	0	0	0	0	0
Fault3	0	0	182	0	0	6	3
Fault4	0	0	0	43	3	1	0
Fault5	0	0	0	2	129	0	0
Normal	3	26	174	127	19	36087	0

Least Square

نمودار ROC برای Train



Least Square

پارامترهای خطای Test

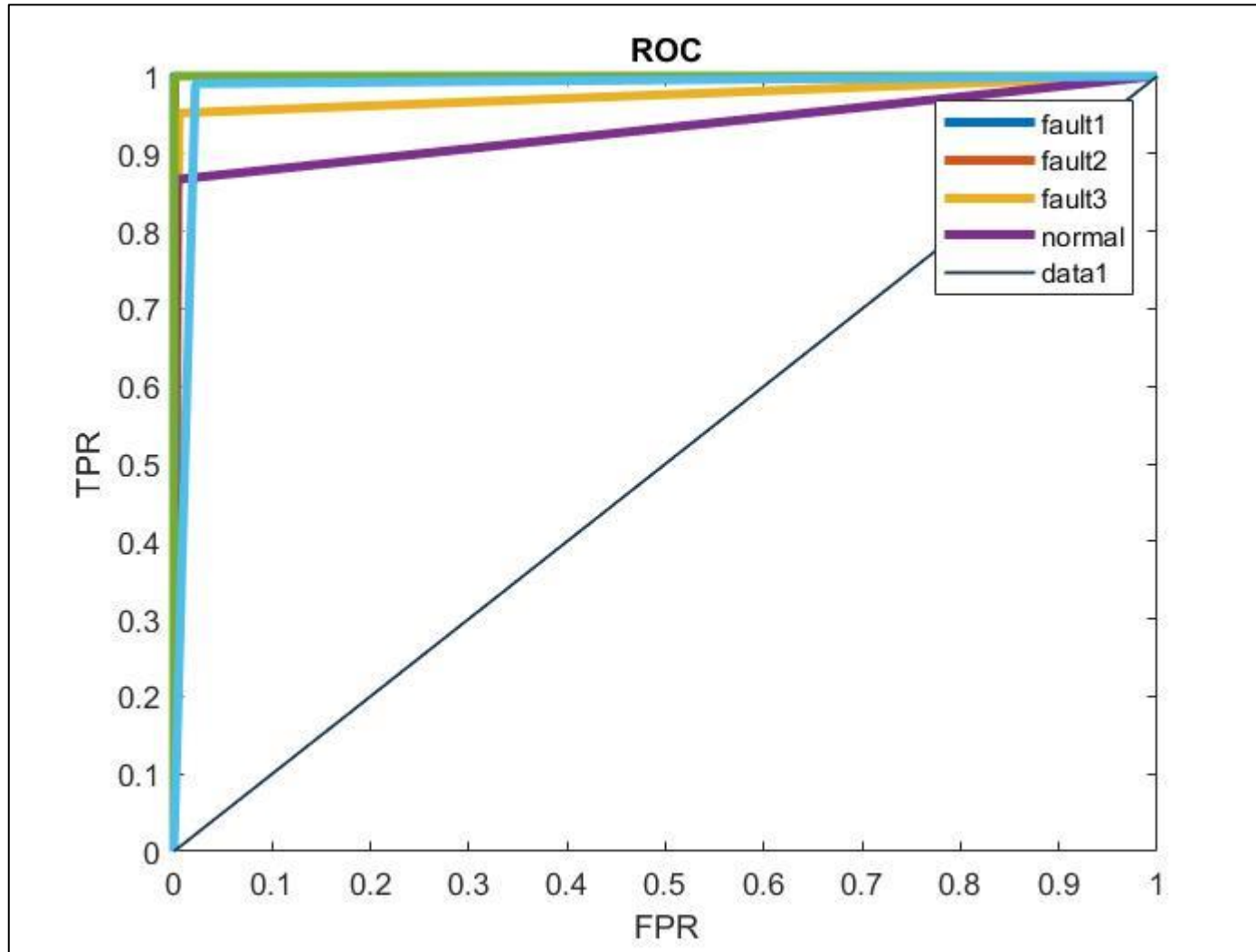
error_unsupervised	error_supervised	performance
0.010015	0.031778	0.98999

	FAR	MAR	Sensitivity	Specificity	Accuracy
fault 1	0	0	1	1	1
fault 2	0.00097744	0	1	0.99902	0.99902
fault 3	0.0053196	0.047619	0.95238	0.99468	0.99446
fault 4	0.0031793	0.13333	0.86667	0.99682	0.99666
fault 5	0.00024512	0	1	0.99975	0.99976

	Fault1	Fault2	Fault3	Fault4	Fault5	Normal	MissClass
Fault1	11	0	0	0	0	0	0
Fault2	0	5	0	0	0	0	0
Fault3	0	0	60	0	0	2	1
Fault4	0	0	0	13	1	1	0
Fault5	0	0	0	0	43	0	0
Normal	0	12	65	39	2	12027	0

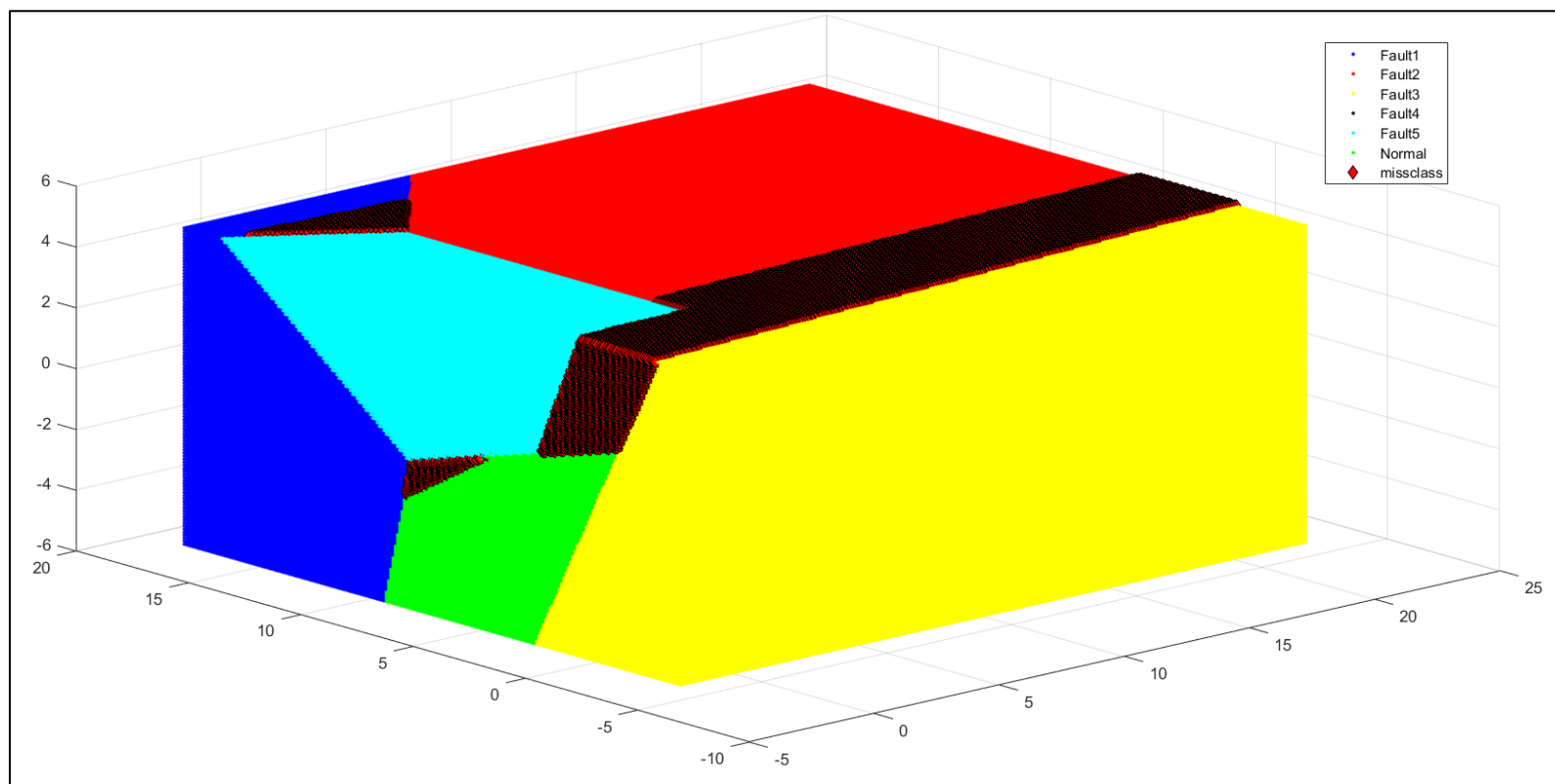
Least Square

نمودار ROC برای Test



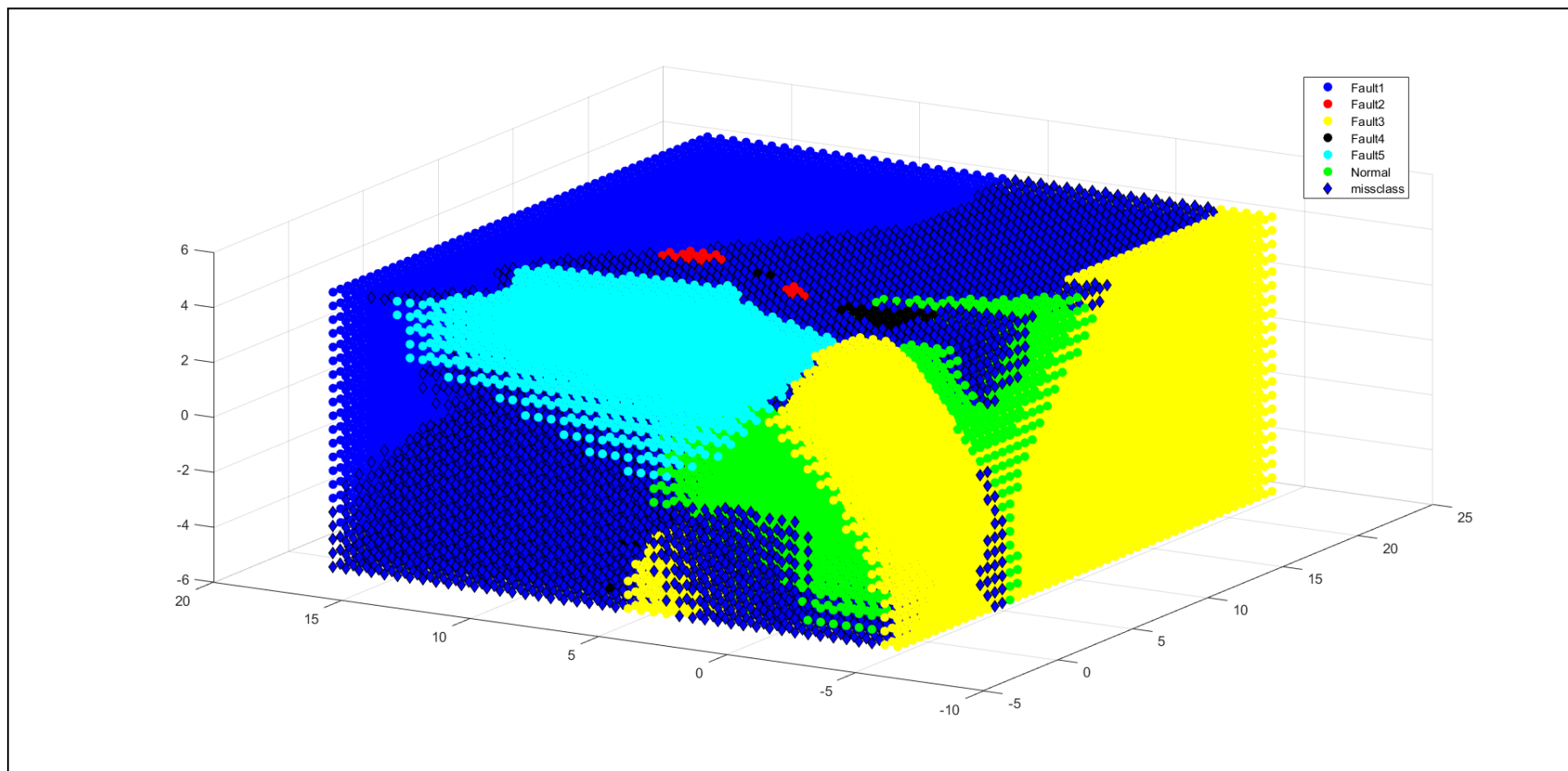
جداسازی با استفاده از ماشین بردار پشتیبان (Support Vector Machine)

ماشین بردار پشتیبان خطی (Linear S.V.M)

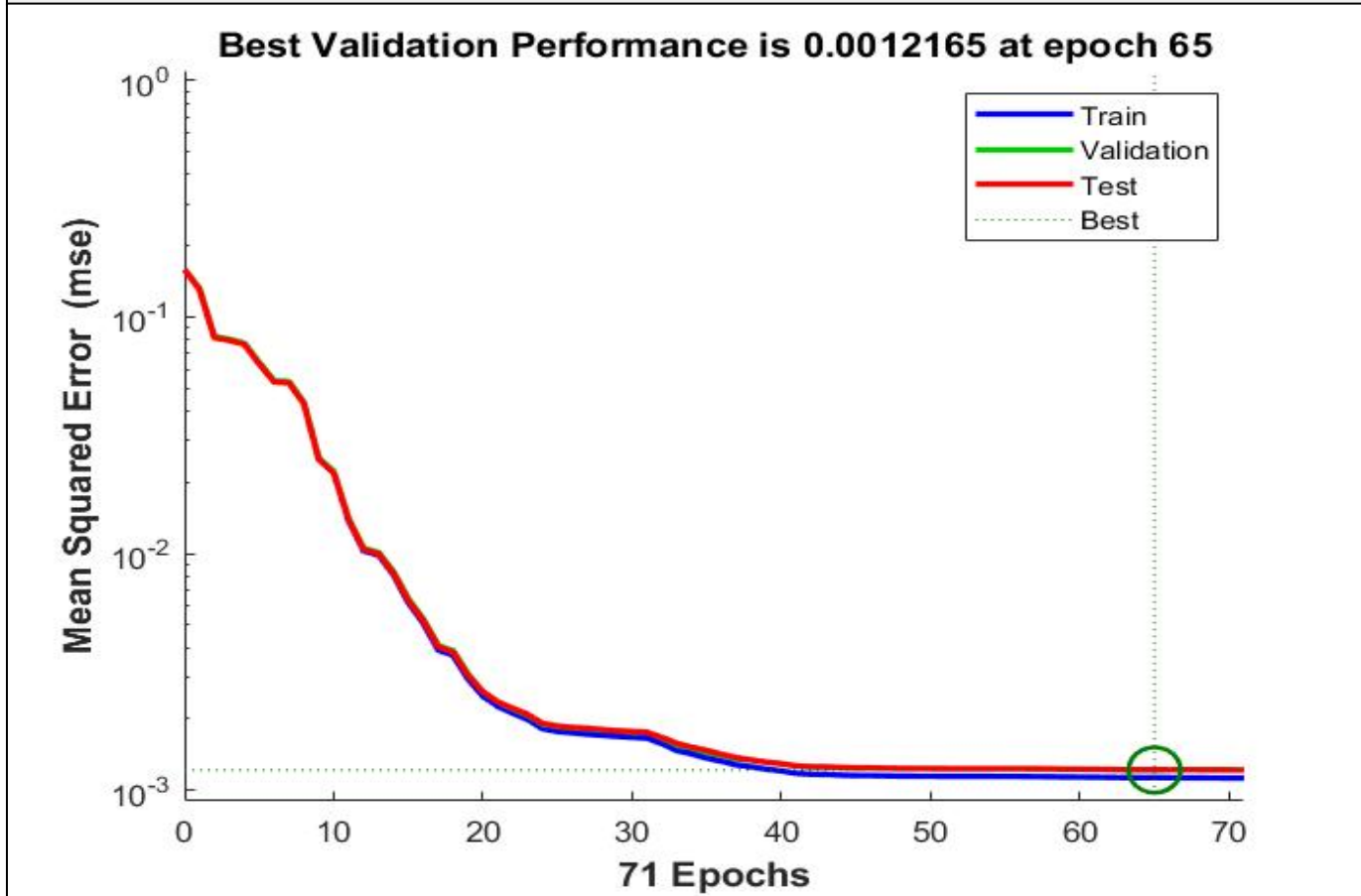
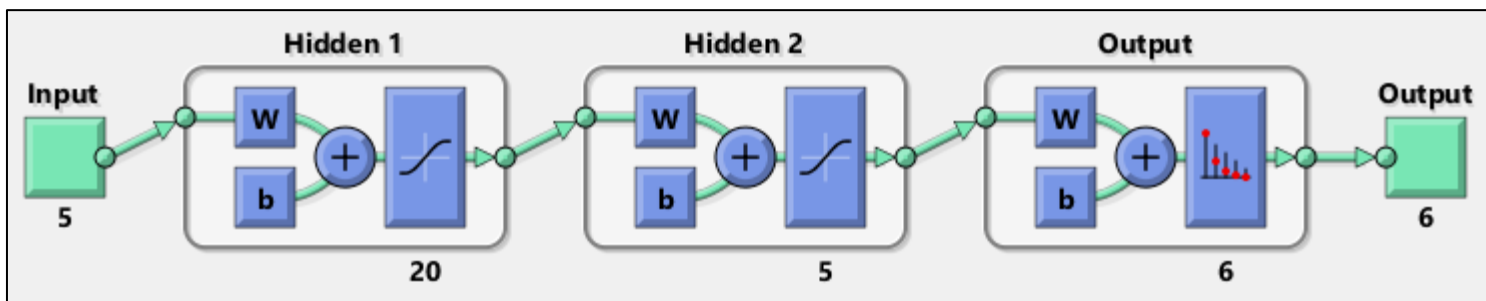


جداسازی با استفاده از ماشین بردار پشتیبان (Support Vector Machine)

ماشین بردار پشتیبان چند جمله‌ای (Polynomial S.V.M)



جداسازی با استفاده از شبکه‌ی عصبی (MLP) Multi Layer Perceptron



جداسازی با استفاده از شبکه‌ی عصبی (MLP) Multi Layer Perceptron

Training Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	Accuracy
1	30132	0	0	0	0	12	100.0%
2	0	14020	0	0	0	7	100.0%
3	0	0	17811	0	0	14	100.0%
4	0	0	0	41900	0	36	99.9%
5	0	0	0	1377	2201	1	98.9%
6	0	0	0	0	0	33859	100%
Macro Avg	100%	100%	100%	96.8%	100%	99.8%	99.7%
Weighted Avg	0.0%	0.0%	0.0%	3.2%	0.0%	0.2%	0.3%

Validation Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	Accuracy
1	6539	0	0	0	0	3	100.0%
2	0	3033	0	0	0	1	100.0%
3	0	0	37783	0	0	1	100.0%
4	0	0	0	9186	0	20	99.8%
5	0	0	0	311	26106	0	98.8%
6	0	0	0	0	0	7337	100%
Macro Avg	100%	100%	100%	96.7%	100%	99.7%	99.6%
Weighted Avg	0.0%	0.0%	0.0%	3.3%	0.0%	0.3%	0.4%

Test Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	Accuracy
1	6372	0	0	0	0	4	99.9%
2	0	2967	0	0	0	1	100.0%
3	0	0	38352	0	0	4	100.0%
4	0	0	0	8974	0	13	99.9%
5	0	0	0	314	26051	0	98.8%
6	0	0	0	0	0	7268	100%
Macro Avg	100%	100%	100%	96.6%	100%	99.7%	99.6%
Weighted Avg	0.0%	0.0%	0.0%	3.4%	0.0%	0.3%	0.4%

All Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	Accuracy
1	43043	0	0	0	0	19	100.0%
2	0	20020	0	0	0	9	100.0%
3	0	0	25425	0	0	19	100.0%
4	0	0	0	50060	0	69	99.9%
5	0	0	0	2002	7417	1	98.9%
6	0	0	0	0	0	48464	100%
Macro Avg	100%	100%	100%	96.8%	100%	99.8%	99.6%
Weighted Avg	0.0%	0.0%	0.0%	3.2%	0.0%	0.2%	0.4%

جداسازی با استفاده از شبکه‌ی عصبی (MLP) Multi Layer Perceptron

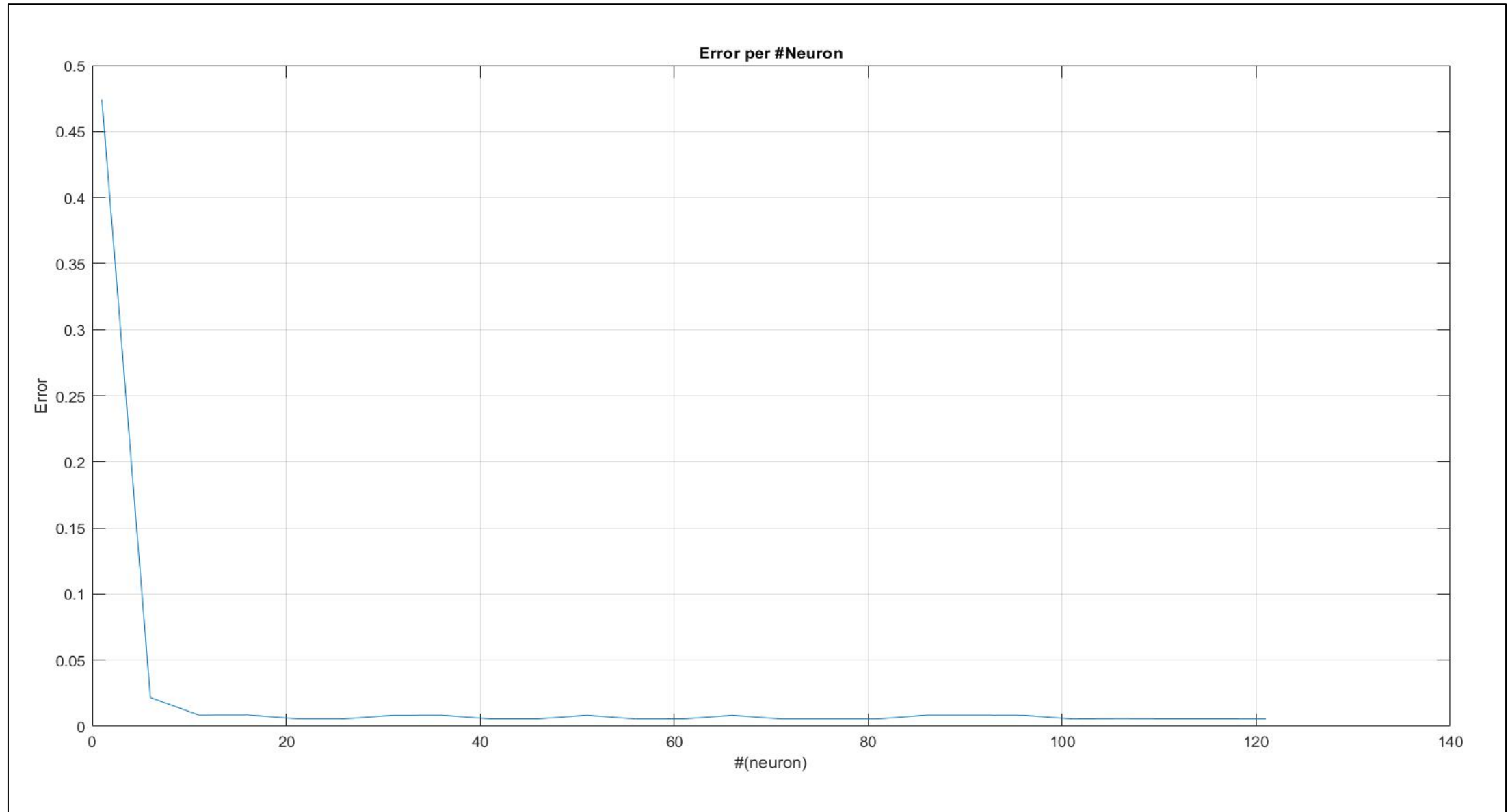
پارامترهای خطا

	error_unsupervised	error_supervised	performance
	0.0035192	0.0057777	0.99648

	FAR	MAR	Sensitivity	Specificity	Accuracy
fault 1	3.3984e-05	0	1	0.99997	0.99997
fault 2	1.5461e-05	0	1	0.99998	0.99999
fault 3	5.4617e-05	0	1	0.99995	0.99997
fault 4	0.00012776	0.032258	0.96774	0.99987	0.99656
fault 5	0.0046803	0	1	0.99532	0.99667

	Fault1	Fault2	Fault3	Fault4	Fault5	Normal
Fault1	43043	0	0	0	0	0
Fault2	0	20020	0	0	0	0
Fault3	0	0	2.5425e+05	0	0	0
Fault4	0	0	0	60060	2002	0
Fault5	0	0	0	0	1.7417e+05	0
Normal	19	9	19	69	1	48464

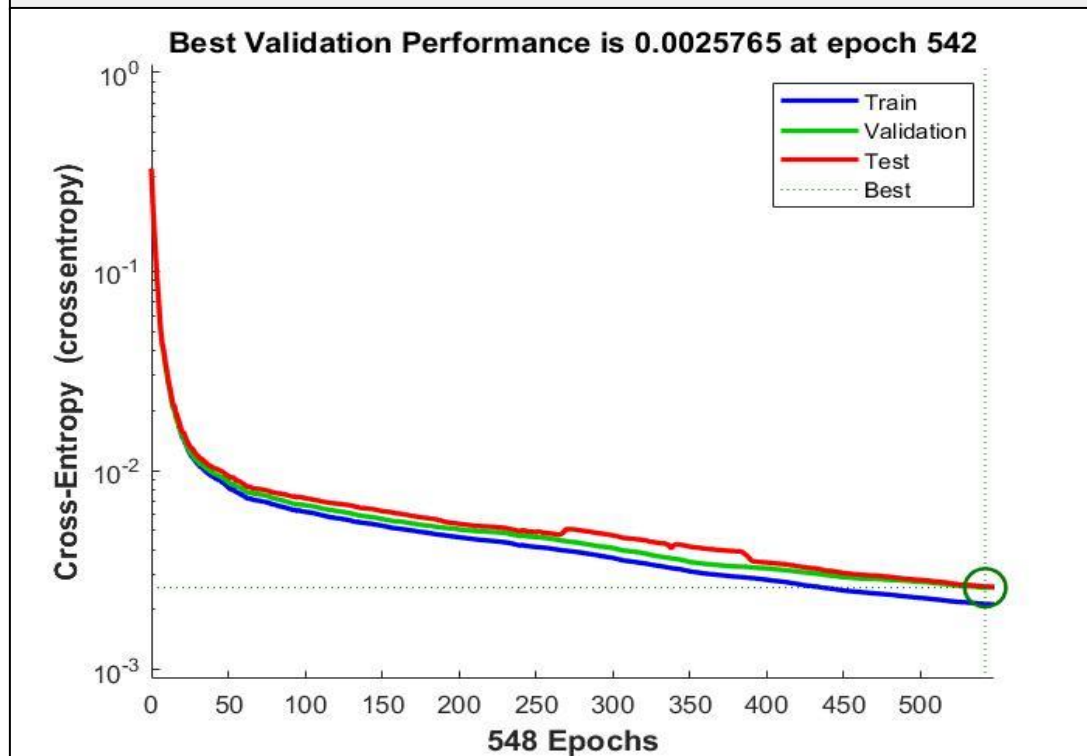
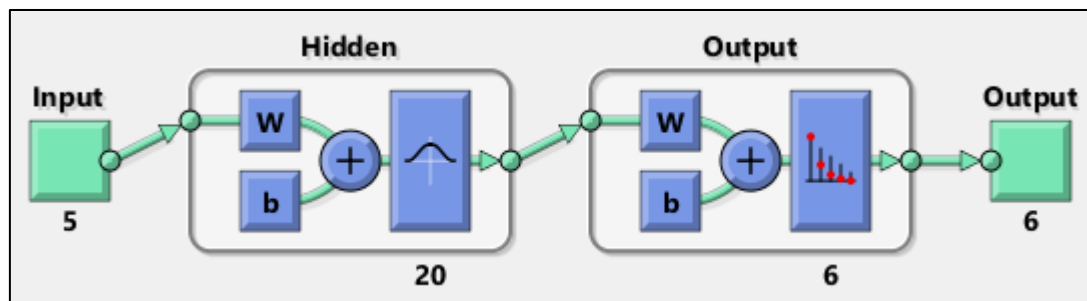
جداسازی با استفاده از شبکه‌ی عصبی Radial Base Function(RBF)



جداسازی با استفاده از شبکه‌ی عصبی (RBF)

با توجه به نمودار خطا برحسب تعداد نرون که در صفحه‌ی قبل ارائه شد، مقدار بهینه‌ی ۲۰ را برای تعداد نرون‌های شبکه RBF انتخاب

کرده و نتایج را ارائه می‌دهیم.



جداسازی با استفاده از شبکه‌ی عصبی (RBF) Radial Base Function

Training Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	3073	0	0	0	0	0	100%
	4.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	0	1396	0	0	0	0	100%
	0.0%	1.9%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0	0	17994	0	0	16	99.9%
	0.0%	0.0%	24.6%	0.0%	0.0%	0.0%	0.1%
4	0	0	0	4253	0	25	99.4%
	0.0%	0.0%	0.0%	5.8%	0.0%	0.0%	0.6%
5	0	0	0	130	12320	0	99.0%
	0.0%	0.0%	0.0%	0.2%	16.9%	0.0%	1.0%
6	0	0	0	0	0	33897	100%
	0.0%	0.0%	0.0%	0.0%	0.0%	46.4%	0.0%
	100%	100%	100%	97.0%	100%	99.9%	99.8%
	0.0%	0.0%	0.0%	3.0%	0.0%	0.1%	0.2%

Validation Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	661	0	0	0	0	0	100%
	4.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	0	296	0	0	0	1	99.7%
	0.0%	1.9%	0.0%	0.0%	0.0%	0.0%	0.3%
3	0	0	3920	0	0	4	99.9%
	0.0%	0.0%	25.0%	0.0%	0.0%	0.0%	0.1%
4	0	0	0	901	0	5	99.4%
	0.0%	0.0%	0.0%	5.8%	0.0%	0.0%	0.6%
5	0	0	0	34	2622	0	98.7%
	0.0%	0.0%	0.0%	0.2%	16.7%	0.0%	1.3%
6	0	0	0	0	0	7221	100%
	0.0%	0.0%	0.0%	0.0%	0.0%	46.1%	0.0%
	100%	100%	100%	96.4%	100%	99.9%	99.7%
	0.0%	0.0%	0.0%	3.6%	0.0%	0.1%	0.3%

Test Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	609	0	0	0	0	0	100%
	3.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	0	328	0	0	0	0	100%
	0.0%	2.1%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0	0	3740	0	0	4	99.9%
	0.0%	0.0%	23.9%	0.0%	0.0%	0.0%	0.1%
4	0	0	0	906	0	4	99.6%
	0.0%	0.0%	0.0%	5.8%	0.0%	0.0%	0.4%
5	0	0	0	38	2632	0	98.6%
	0.0%	0.0%	0.0%	0.2%	16.8%	0.0%	1.4%
6	0	0	0	0	0	7404	100%
	0.0%	0.0%	0.0%	0.0%	0.0%	47.3%	0.0%
	100%	100%	100%	96.0%	100%	99.9%	99.7%
	0.0%	0.0%	0.0%	4.0%	0.0%	0.1%	0.3%

All Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	4343	0	0	0	0	0	100%
	4.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	0	2020	0	0	0	1	100.0%
	0.0%	1.9%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0	0	25654	0	0	24	99.9%
	0.0%	0.0%	24.6%	0.0%	0.0%	0.0%	0.1%
4	0	0	0	6060	0	34	99.4%
	0.0%	0.0%	0.0%	5.8%	0.0%	0.0%	0.6%
5	0	0	0	202	17574	0	98.9%
	0.0%	0.0%	0.0%	0.2%	16.8%	0.0%	1.1%
6	0	0	0	0	0	48522	100%
	0.0%	0.0%	0.0%	0.0%	0.0%	46.5%	0.0%
	100%	100%	100%	96.8%	100%	99.9%	99.8%
	0.0%	0.0%	0.0%	3.2%	0.0%	0.1%	0.2%

جداسازی با استفاده از شبکه‌ی عصبی Radial Base Function(RBF)

پارامترهای خطا

error_unsupervised	error_supervised	performance
0.0024992	0.0055788	0.9975

	FAR	MAR	Sensitivity	Specificity	Accuracy
fault 1	0	0	1	1	1
fault 2	9.7643e-06	0	1	0.99999	0.99999
fault 3	0.00030465	0	1	0.9997	0.99977
fault 4	0.00034633	0.032258	0.96774	0.99965	0.99774
fault 5	0.0023256	0	1	0.99767	0.99807

	Fault1	Fault2	Fault3	Fault4	Fault5	Normal
Fault1	4343	0	0	0	0	0
Fault2	0	2020	0	0	0	0
Fault3	0	0	25654	0	0	0
Fault4	0	0	0	6060	202	0
Fault5	0	0	0	0	17574	0
Normal	0	1	24	34	0	48522

Method	Performance	Error	Process Time (sec)
Least Square (LS)	99%	4.3%	0.2
SVM Linear	97.3%	3.3%	0.1
SVM Polynomial	99.8%	1.8%	1.2
k-NN	99.9%	2.05%	44.5
MLP	99.6%	0.57%	0.2
RBF	99.9%	0.019%	0.4