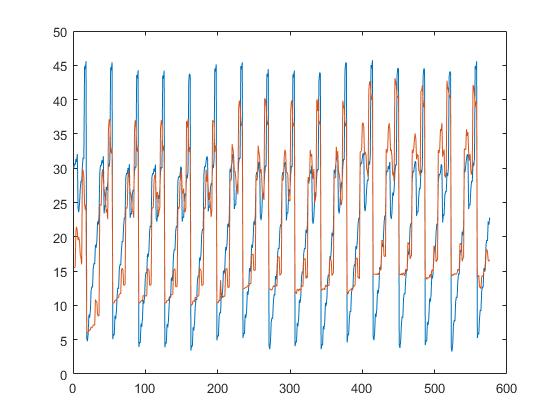
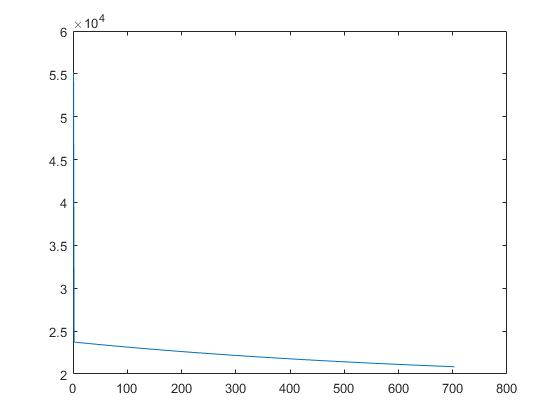
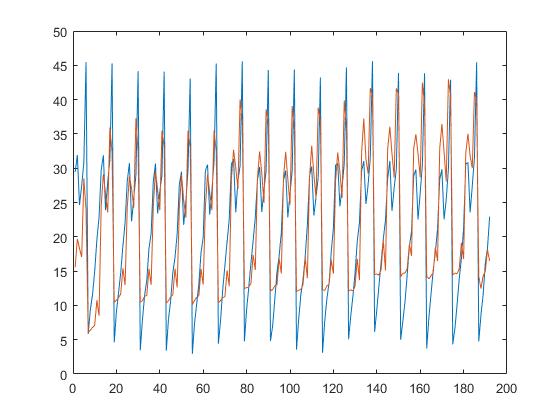
Deep Learning hw1

1. regression  
   i.   
   (1)ALL FEATURE

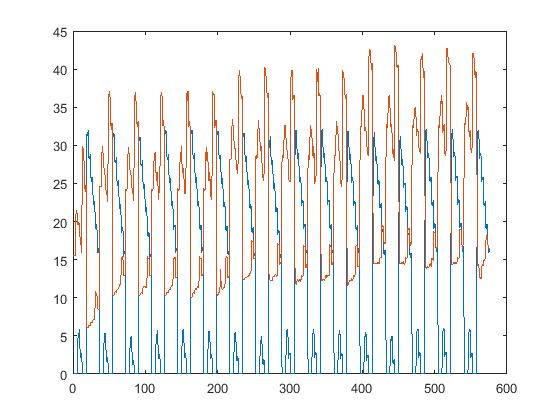
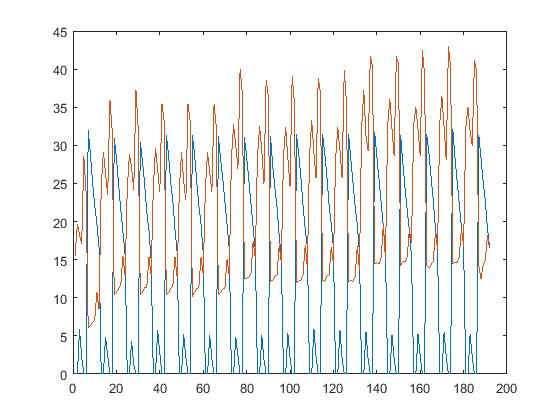
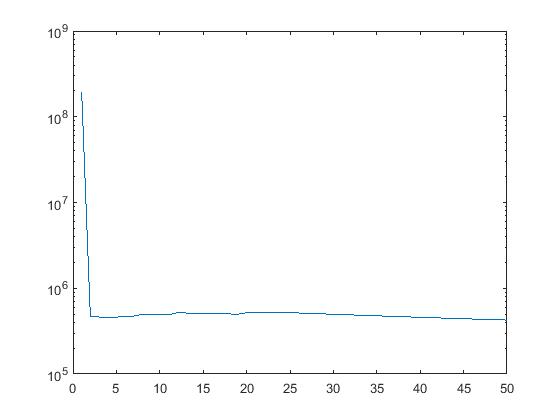
|  |  |
| --- | --- |
| Network Architecture | 16 – 5 – 4 – 1 |
| Selected features | All Features |
| Training ERMS | 5.9625 |
| Test ERMS | 6.0562 |



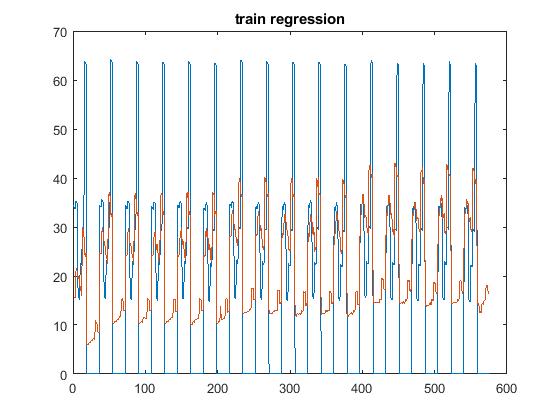
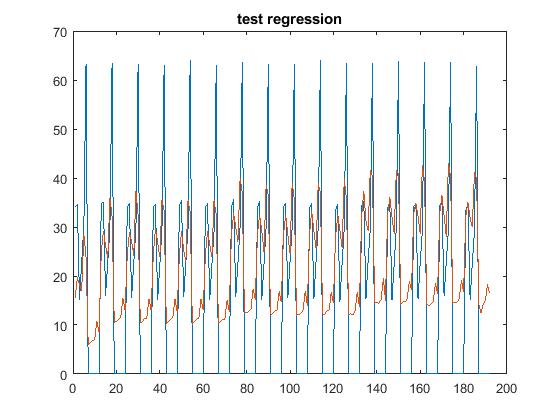
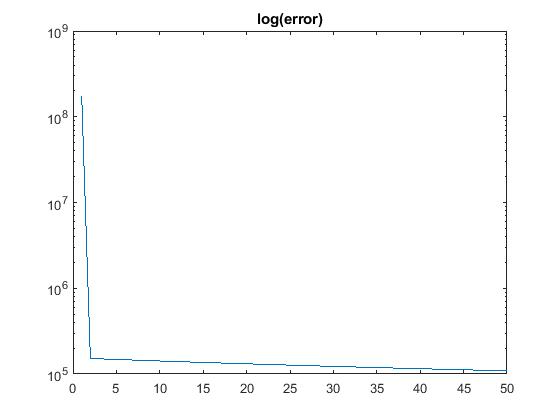
learning curve training regression result



testing regression result  
(2)without the first feature



|  |  |
| --- | --- |
| Training ERMS | 13.9625 |
| Test ERMS | 14.0562 |

(3)without the second feature  


|  |  |
| --- | --- |
| Training ERMS | 13.7667 |
| Test ERMS | 13.7948 |

(3) without the third feature

|  |  |
| --- | --- |
| Training ERMS | 33.7473 |
| Test ERMS | 33.3946 |

(4) without the fourth feature

|  |  |
| --- | --- |
| Training ERMS | 31.7629 |
| Test ERMS | 32.7845 |

(5) without the fifth feature

|  |  |
| --- | --- |
| Training ERMS | 11.7667 |
| Test ERMS | 11.7948 |

(6) without the sixth feature

|  |  |
| --- | --- |
| Training ERMS | 15.7424 |
| Test ERMS | 15.7648 |

(7) without the seventh feature

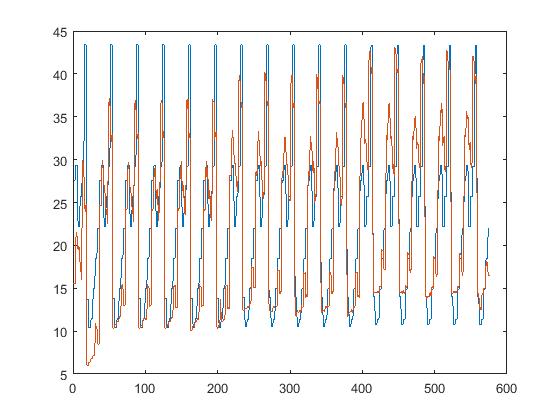
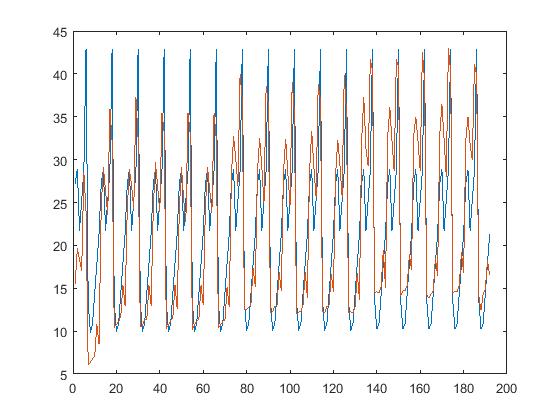
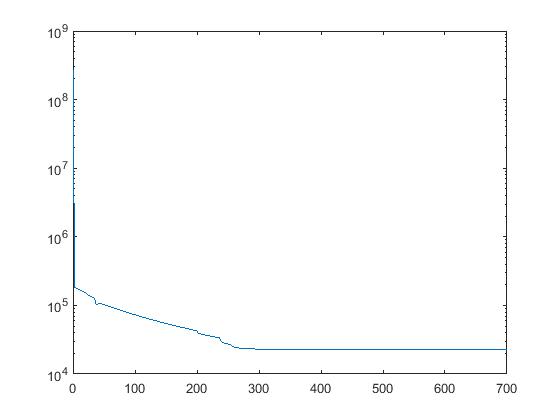
|  |  |
| --- | --- |
| Training ERMS | 43.2667 |
| Test ERMS | 43.2948 |

(8) without the eighth feature

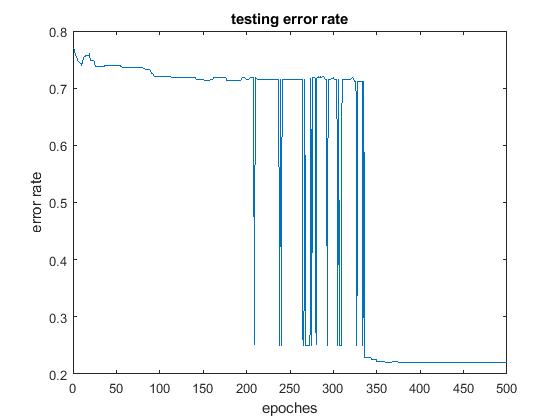
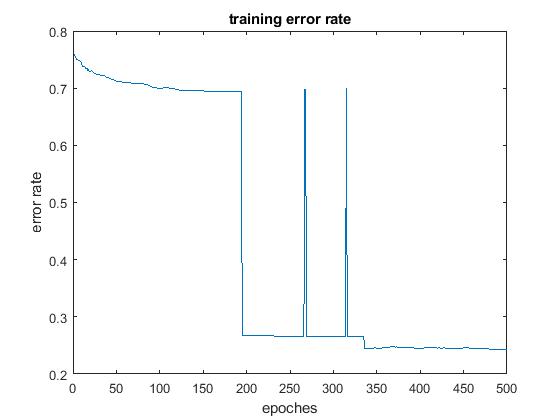
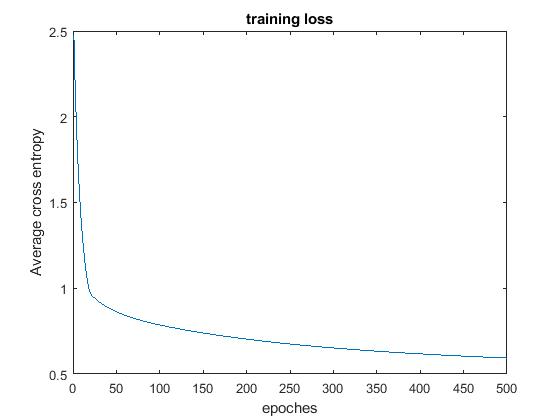
|  |  |
| --- | --- |
| Training ERMS | 12.1357 |
| Test ERMS | 12.1948 |

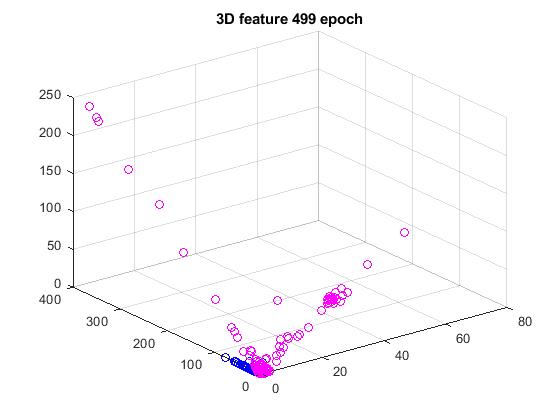
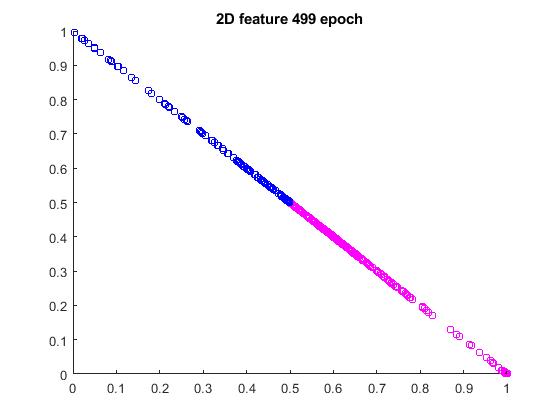
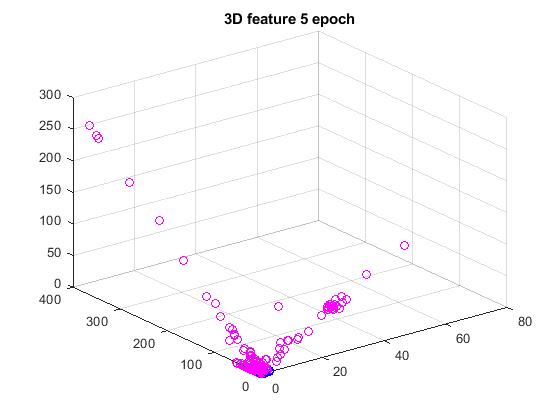
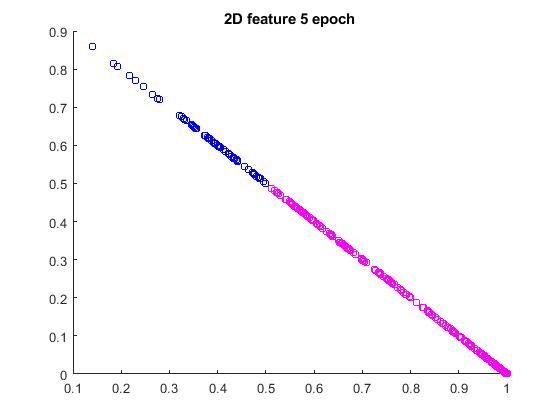
ii. 3 significant features  
We found that there are three key features control the dataset. The figures and chart are as follows.

|  |  |
| --- | --- |
| Network Architecture | 16 – 5 – 4 – 1 |
| Selected features | Wall area, roof area, glazing area |
| Training ERMS | 5.6188 |
| Test ERMS | 5.6248 |



2. classification  
 i.



ii.   


iii.

We found that as the epoch increase, the separation of the data becomes better. The error rate decreases because of the separation of the data in the latent space.