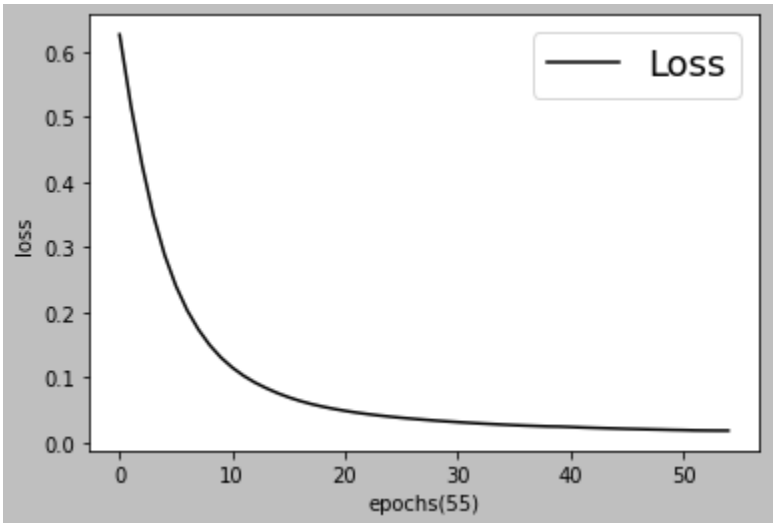
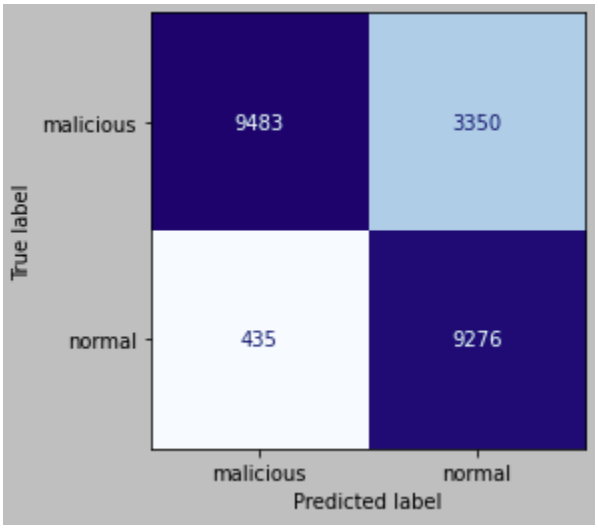
Task:
Generate data and train DL models

Output:



F1:	83.055%
AUC:	82.179%
Accuracy:	83.211%
Precision:	95.61%
Recall:	73.90%



Conclusion:

1. Used a shallow MLP for the detection task.
2. Due to the inherent bias in the original data, the models were inclined towards the larger number of instances.
3. Added a few thousand R2L attacks to the original dataset and used a shallow MLP for the detection task.
4. This improved the detection rate. Previously the accuracy was around 79% where a large number of attacks were missed by the model.
5. Now the miss rate has decreased due to the inclusion of a large number of R2L attacks in the training data.
6. Creating a stratified dataset did not improve the accuracy by much.

Next:

1. Use different autoencoders to reduce dimensions and to learn hidden patterns among malicious and normal traffic.
2. Use 5-class to check which attack has the highest miss rate.
3. Reduce dimensions using Undercomplete AE.
4. Build a denoising AE to make the autoencoders invariant to minor changes in the data. It will make the model robust and the model may be able to catch a variety of patterns in the test data.