P1: Precision, Recall and F1 measure

The values of the Precision, Recall and F1 measure calculated for each threshold value for this system are as follows.

	Threshold	Precision	Recall	F1_measure	TP_R	FN_R	FP_R	TN_R
0	1	0.909091	0.20	0.327869	0.20	0.80	0.02	0.98
1	5	0.909091	0.50	0.645161	0.50	0.50	0.05	0.95
2	10	0.857143	0.60	0.705882	0.60	0.40	0.10	0.90
3	15	0.800000	0.80	0.800000	0.80	0.20	0.20	0.80
4	20	0.745763	0.88	0.807339	0.88	0.12	0.30	0.70
5	25	0.692308	0.90	0.782609	0.90	0.10	0.40	0.60
6	30	0.655172	0.95	0.775510	0.95	0.05	0.50	0.50
7	35	0.615385	0.96	0.750000	0.96	0.04	0.60	0.40
8	40	0.580838	0.97	0.726592	0.97	0.03	0.70	0.30
9	50	0.550562	0.98	0.705036	0.98	0.02	0.80	0.20

We can see from the plot that as the threshold values increase, the Precision values decrease until a minimum of 0.550562. However, the Recall values are increasing and reaches 0.98.

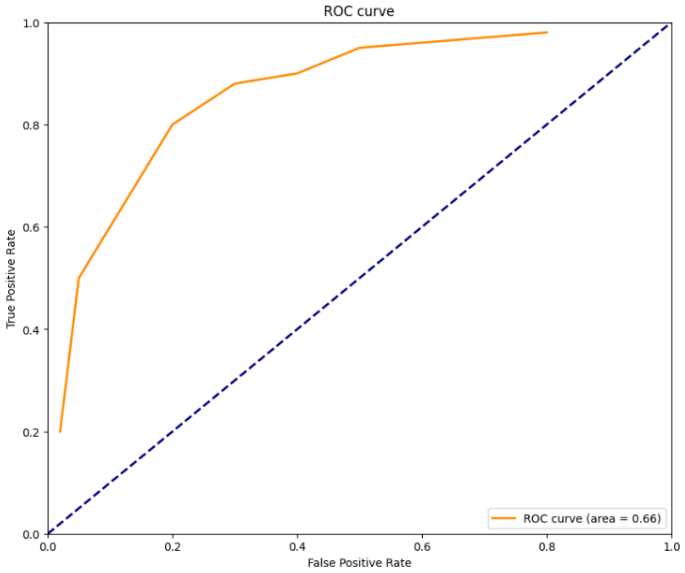
Recall reflects the system's ability to identify positive samples, and higher Recall indicates the system's ability to identify positive samples. In contrast, Precision reflects the system's ability to distinguish negative samples, and higher Precision indicates the system's ability to distinguish negative samples. When calculating Recall, as threshold values increase, the percentage of correctly classified positives also increases. The denominator is getting smaller, so the result obtained is getting larger. Since false positives increase with increasing threshold values, the false positives keep increasing in the calculation of Precision, resulting in progressively larger denominators and smaller and smaller results. However, F1_measure, is the summed average of the two, and the higher the F1_measure, the more robust the classification system is.

We can see that the maximum F1_measure is 0.807339 and the corresponding threshold value is 20. Therefore, we can get the best threshold value of 20 in this system.

P2: ROC

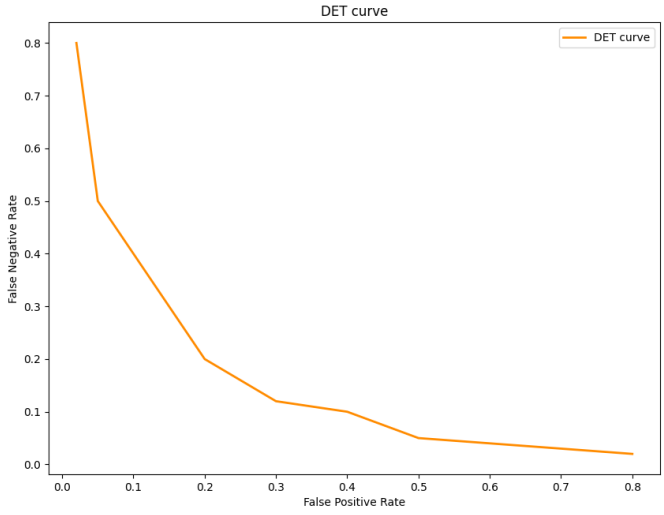
The plot for the ROC curve as follows.

The ROC curve is located in the upper left, and the classifier with a larger AUC value has a higher correct rate. The AUC is about 0.66 and $0.5 < AUC < 1$, which is better than random guesses and shows that we can distinguish between different categories. It also means that this system can have predictive value if the threshold is appropriately set. We can predict that the farthest perpendicular distance from the line $y=x$ occurs roughly at some point between the point (0.2,0.8) and the point (0.3,0.88), which may be close to the maximum F1_measure obtained when the threshold value is 20.



P3: DET

The plot for the ROC curve as follows.



We can see that the false positive rate increase and the false negative rate decrease as the threshold values increase. The DET curve can tell us at which point the two reach equilibrium. Considering the four aspects of the threshold value, F1_measure, false positive rate, and false negative rate, the F1_measure can achieve the maximum value when the threshold value is between 20 and 25. Moreover, this interval's false positive and negative rates can also reach an equilibrium consistent with the results obtained and the plots shown.