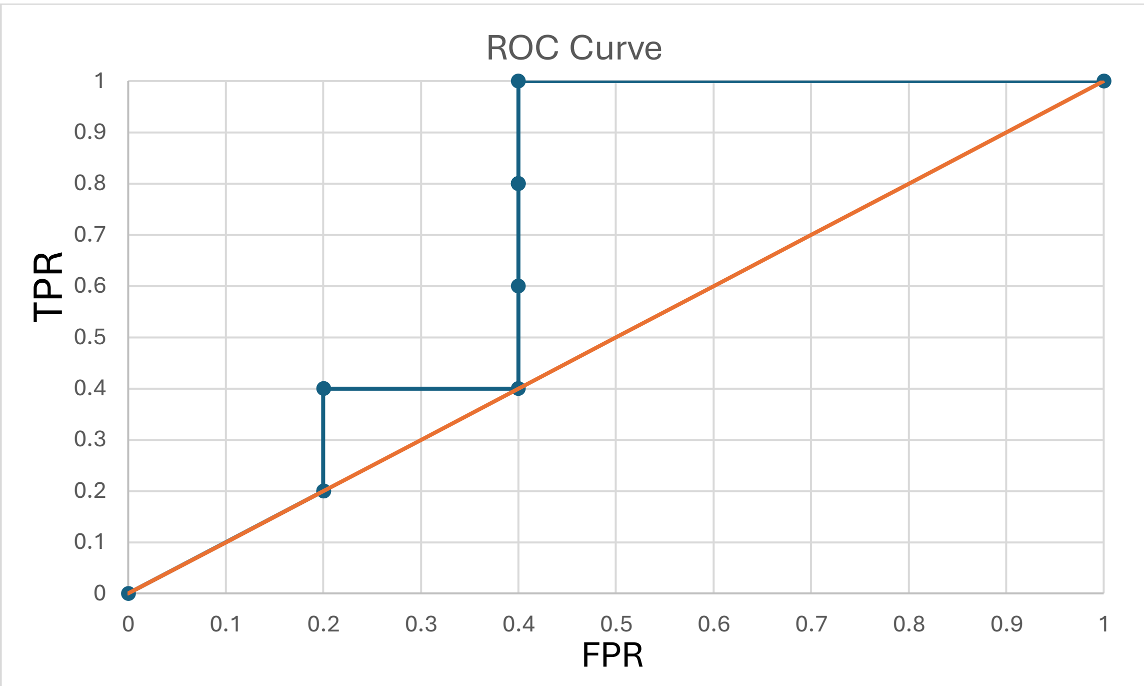
ML Assignment 3 – Theoretical

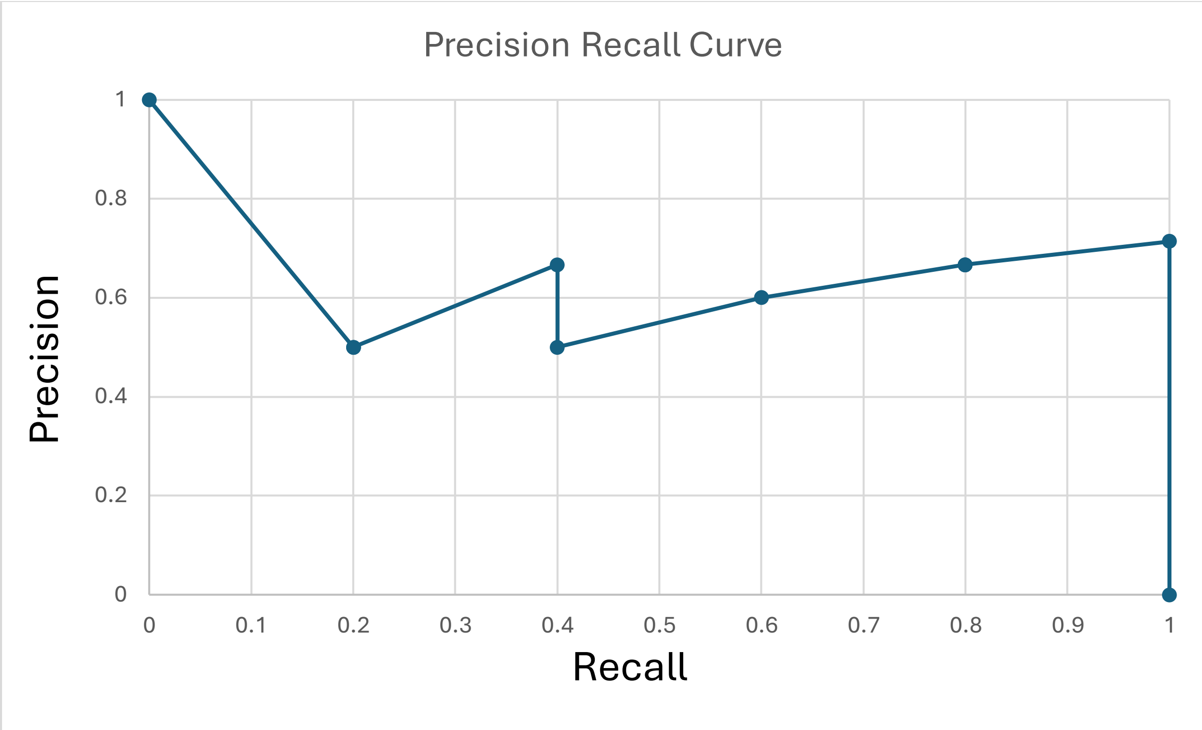
Names: Victor Gavrilenko 209406255 , *Ilay Cohen* : 206515744

First Classifier:

label	classification	0	>0.1	>0.2	>0.3	>0.4	>0.5	>0.6	>0.7	>0.8	>0.9	1
1	0.41	1	1	1	1	1	0	0	0	0	0	0
0	0.00	1	0	0	0	0	0	0	0	0	0	0
0	0.00	1	0	0	0	0	0	0	0	0	0	0
1	0.73	1	1	1	1	1	1	1	1	0	0	0
0	0.62	1	1	1	1	1	1	1	0	0	0	0
1	1.00	1	1	1	1	1	1	1	1	1	1	0
0	1.00	1	1	1	1	1	1	1	1	1	1	0
1	0.14	1	1	0	0	0	0	0	0	0	0	0
0	0.00	1	0	0	0	0	0	0	0	0	0	0
1	0.55	1	1	1	1	1	1	0	0	0	0	0

Threshold	0	>0.1	>0.2	>0.3	>0.4	>0.5	>0.6	>0.7	>0.8	>0.9	1.0
TPR	1	1	0.8	0.8	0.8	0.6	0.4	0.4	0.2	0.2	0
FPR	1	0.4	0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0
Precision	0	0.714286	0.666667	0.666667	0.666667	0.6	0.5	0.666667	0.5	0.5	1
Recall	1	1	0.8	0.8	0.8	0.6	0.4	0.4	0.2	0.2	0

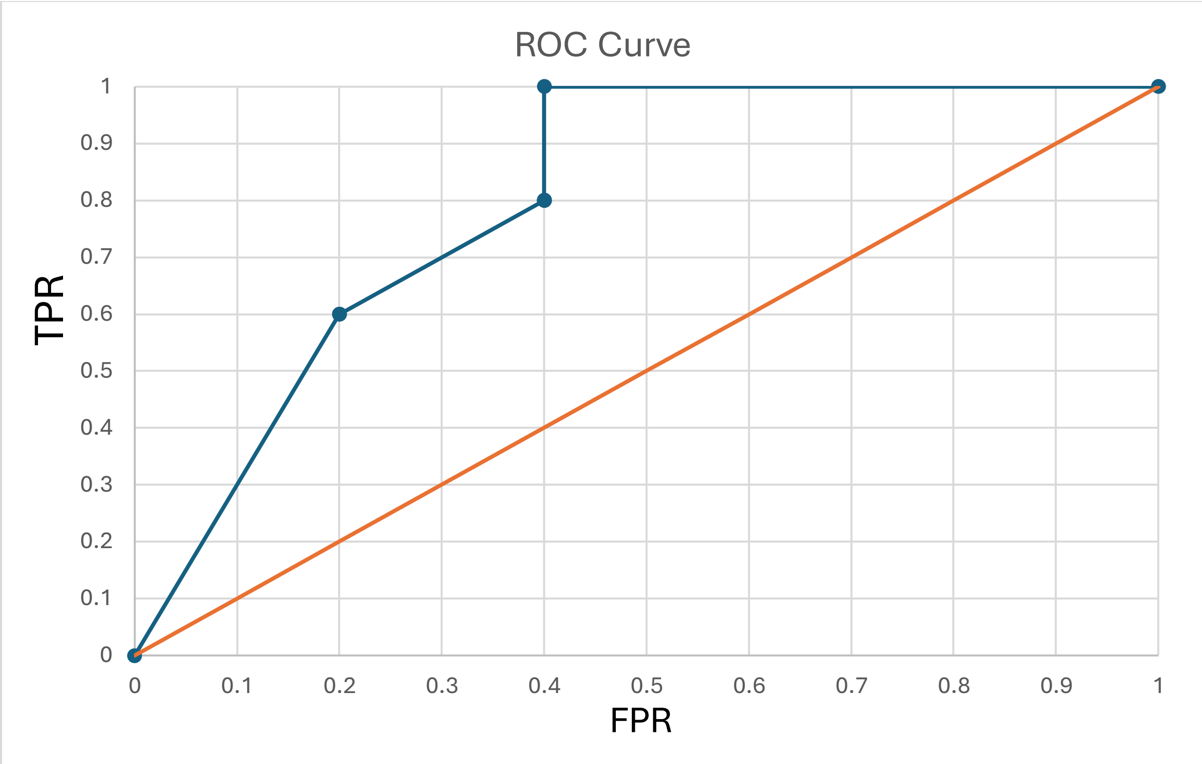


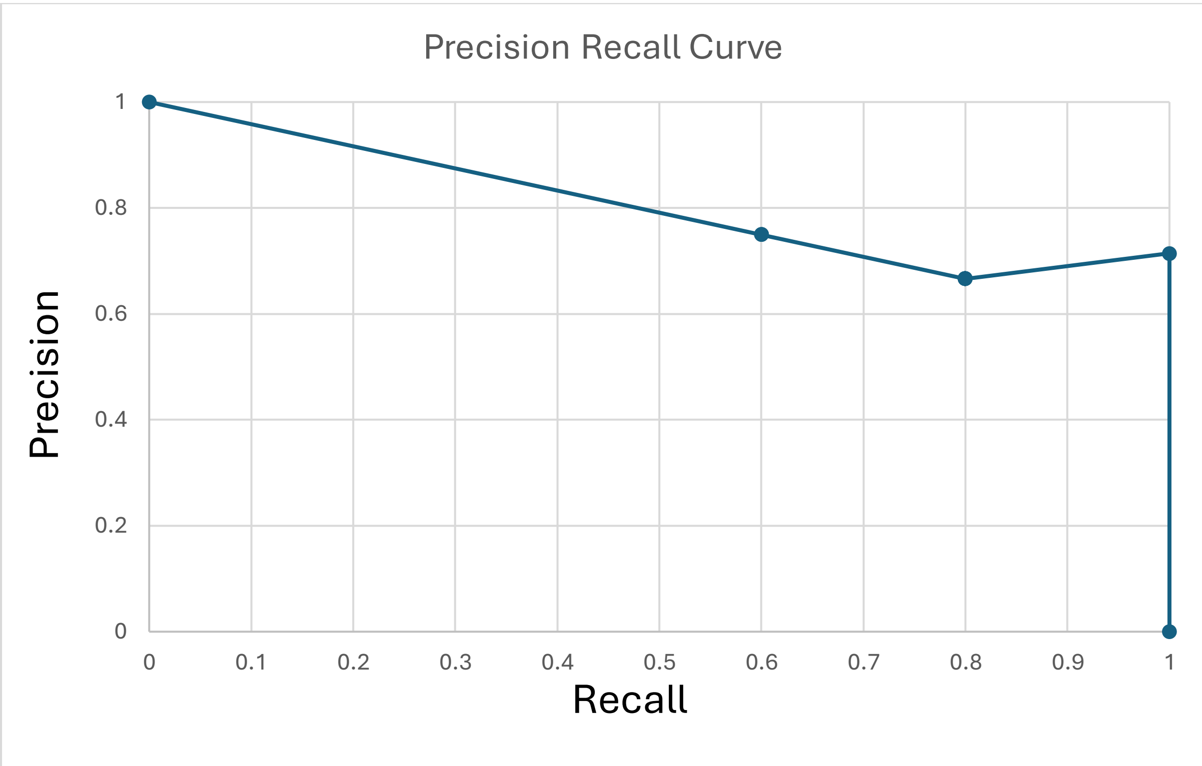


Second Classifier

Label	classification	0	>0.1	>0.2	>0.3	>0.4	>0.5	>0.6	>0.7	>0.8	>0.9	1.0
1	1.00	1	1	1	1	1	1	1	1	1	1	0
0	0.72	1	1	1	1	1	1	1	1	0	0	0
0	0.99	1	1	1	1	1	1	1	1	1	1	0
1	0.14	1	1	0	0	0	0	0	0	0	0	0
0	0.00	1	0	0	0	0	0	0	0	0	0	0
1	0.94	1	1	1	1	1	1	1	1	1	1	0
0	0.10	1	0	0	0	0	0	0	0	0	0	0
1	0.77	1	1	1	1	1	1	1	1	0	0	0
0	0.02	1	0	0	0	0	0	0	0	0	0	0
1	1.00	1	1	1	1	1	1	1	1	1	1	0

Threshold	0	>0.1	>0.2	>0.3	>0.4	>0.5	>0.6	>0.7	>0.8	>0.9	1.0
TPR	1	1	0.8	0.8	0.8	0.8	0.8	0.8	0.6	0.6	0
FPR	1	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.2	0.2	0
Precision	0	0.714286	0.666667	0.666667	0.666667	0.666667	0.666667	0.666667	0.75	0.75	1
Recall	1	1	0.8	0.8	0.8	0.8	0.8	0.8	0.6	0.6	0





***Answer the following questions*:**

1. *Which model is better based on those graphs?*

2. *If those models predict heart attack, choose a 'one threshold' metric (such as acc, precision, etc..) and explain why it's a suitable metric for this case.*

1.In the ROC curves, Classifier 2's curve is higher up and looks better than Classifier 1's. It means Classifier 2 is doing a better job overall at separating the positive and negative cases.

So, it seems like Classifier 2 is doing a better job than Classifier 1.

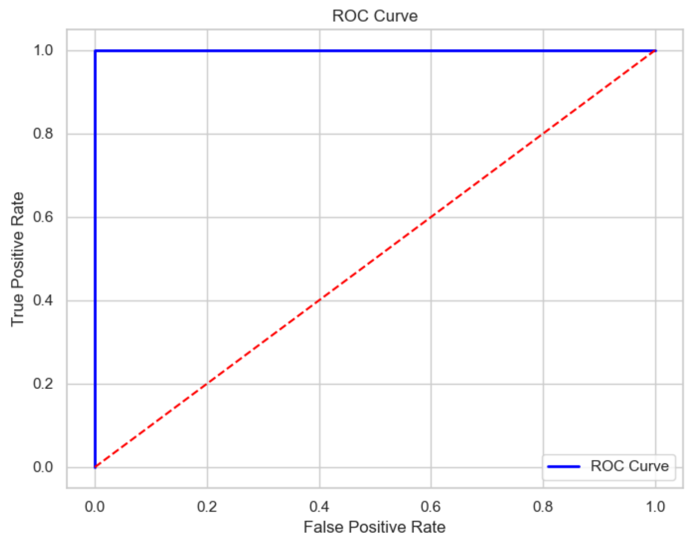
2. For predicting heart attacks, using a threshold metric like precision is suitable. Precision measures how accurate the model is in identifying true positive cases, which is crucial for predicting heart attacks accurately. High precision means fewer false alarms, ensuring that when the model predicts a heart attack, it's likely to be correct, which is essential for timely intervention and patient safety.

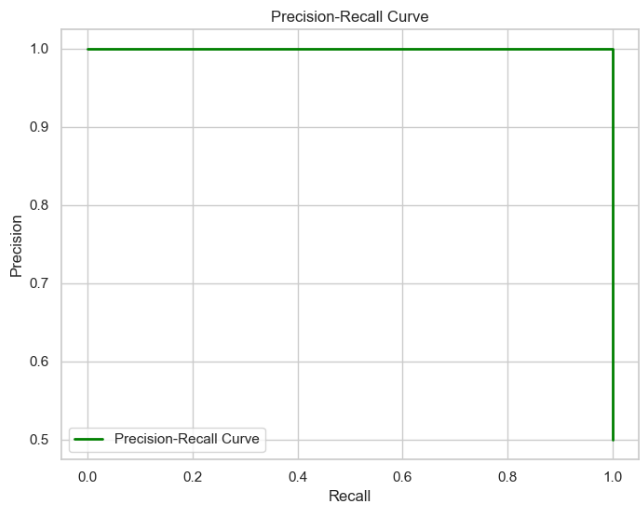
Curves differences (7 pts)

Create a new table with two columns: 'label' and 'predictions'. Fill the table with 10 rows of values representing binary labels and corresponding predicted probabilities. Ensure that the values chosen for the predictions are deliberately selected to make the ROC curve and precision-recall curve look different - which means that according to one curve the model is good, and according to the second one it's bad. After filling the table, draw both the ROC curve and precision-recall curve using the provided data. Explain the values that you chose and the reason for the change.

Label	Prediction	Prediction≥0.5	Prediction < 0.5
1	0.99	1	0
0	0.9	1	0
1	0.98	1	0
0	0.8	1	0
1	0.97	1	0
0	0.7	1	0
1	0.96	1	0
0	0.6	1	0
1	0.95	1	0
0	0.5	0	1

$$\underline{For\ the\ ROC\ curve:}$$
$$TPR = \frac{TP}{TP+FN} = \frac{5}{5} = 1, \quad FPR = \frac{FP}{FP+TN} = \frac{5}{5} = 1$$
$$\underline{For\ the\ Precision - Recall\ curve:}$$
$$Precision = \frac{TP}{TP+FP} = \frac{5}{10} = 0.5\,, \quad Recall = \frac{TP}{TP+FN} = \frac{5}{5} = 1$$





We chose these values to ensure that we achieve a perfect ROC curve with the highest AUC while keeping the precision-recall curve as a straight line, consistently at one for precision and increasing with recall.

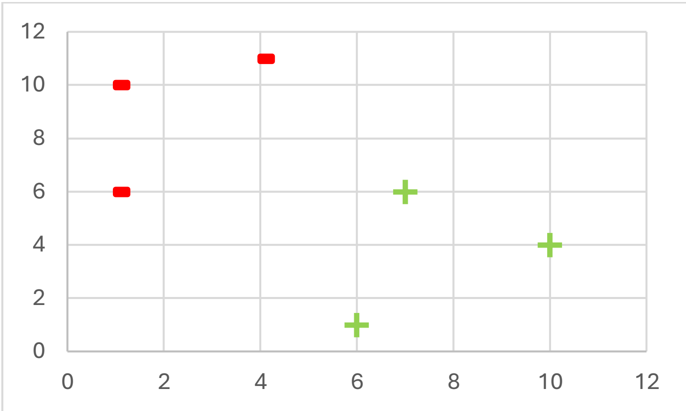
Consequently, we obtained much better (almost perfect) results on the ROC curve with a perfect AUC, but poorer results on the precision-recall curve.

4.

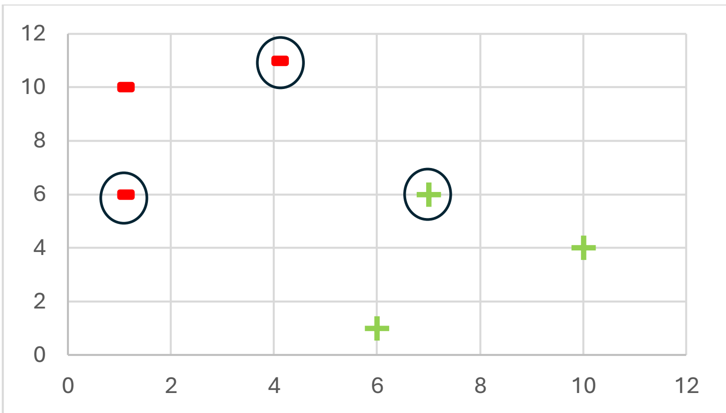
x1	x2	class
1	6	-
1	10	-
4	11	-
6	1	+
7	6	+
10	4	+

Based on the hyperplane, calculate the margin size.

Starting with this graph:



We choose our support vectors to be those 3 Points:



So our support vectors are: $\begin{pmatrix}1\\6\end{pmatrix}, \begin{pmatrix}4\\11\end{pmatrix}, \begin{pmatrix}7\\6\end{pmatrix}$. , we add 1 for the bias $\begin{matrix}1\\6\\1\end{matrix}, \begin{matrix}4\\11\\1\end{matrix}, \begin{matrix}7\\6\\1\end{matrix}$

Calculations:

$$\begin{cases} \alpha_1 \begin{pmatrix}1\\6\\1\end{pmatrix}\begin{pmatrix}1\\6\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}4\\11\\1\end{pmatrix}\begin{pmatrix}1\\6\\1\end{pmatrix} + \alpha_3 \begin{pmatrix}7\\6\\1\end{pmatrix}\begin{pmatrix}1\\6\\1\end{pmatrix} = -1 \\ \alpha_1 \begin{pmatrix}1\\6\\1\end{pmatrix}\begin{pmatrix}4\\11\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}4\\11\\1\end{pmatrix}\begin{pmatrix}4\\11\\1\end{pmatrix} + \alpha_3 \begin{pmatrix}7\\6\\1\end{pmatrix}\begin{pmatrix}4\\11\\1\end{pmatrix} = -1 \\ \alpha_1 \begin{pmatrix}1\\6\\1\end{pmatrix}\begin{pmatrix}7\\6\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}4\\11\\1\end{pmatrix}\begin{pmatrix}7\\6\\1\end{pmatrix} + \alpha_3 \begin{pmatrix}7\\6\\1\end{pmatrix}\begin{pmatrix}7\\6\\1\end{pmatrix} = 1 \end{cases}$$

$$38\alpha_1 + 71\alpha_2 + 44\alpha_3 = -1$$

$$71\alpha_1 + 138\alpha_2 + 95\alpha_3 = -1$$

$$44\alpha_1 + 95\alpha_2 + 86\alpha_3 = 1$$

$$\alpha = \begin{pmatrix}-\frac{61}{225}\\ \frac{3}{25}\\ \frac{4}{225}\end{pmatrix} \longrightarrow -\frac{61}{225}\begin{pmatrix}1\\6\\1\end{pmatrix} + \frac{3}{25}\begin{pmatrix}4\\11\\1\end{pmatrix} + \frac{4}{225}\begin{pmatrix}7\\6\\1\end{pmatrix} =$$

$$\begin{pmatrix}-\frac{61}{225}\\ -\frac{366}{225}\\ -\frac{61}{225}\end{pmatrix} + \begin{pmatrix}\frac{12}{25}\\ \frac{33}{25}\\ \frac{3}{25}\end{pmatrix} + \begin{pmatrix}\frac{28}{225}\\ \frac{24}{225}\\ \frac{4}{225}\end{pmatrix} = \begin{pmatrix}1/3\\ -1/5\\ -2/15\end{pmatrix}$$

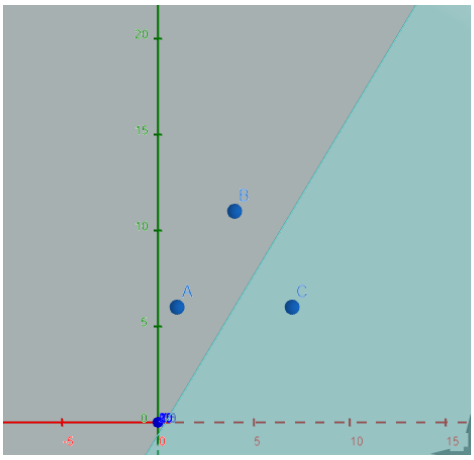
The hyperplane is $\begin{pmatrix}1/3\\ -1/5\end{pmatrix}$ with offset $-2/15$

When $z=0$, $y(x) = \frac{5}{3}x - \frac{10}{15}$

$\text{Distance}\left(y(x), \begin{pmatrix}1\\6\\0\end{pmatrix}\right) = 2.57$

So margin size overall (either side x2) $= 5.14$

3D Plotting:



We can see our “by hands” calculations are equal to “by program” calculations:

$f: z = \frac{1}{3}x - \frac{1}{5}y - \frac{2}{15}$

$A = (1, 6, 0)$

$B = (4, 11, 0)$

$C = (7, 6, 1)$

$g: y = \frac{5}{3}x - \frac{10}{15}$

$a = \text{Distance}(A, g)$

$= 2.57$

$b = \text{Distance}(B, g)$

$= 2.57$

+ $\text{Distance}(C, g)$

$= 2.57$