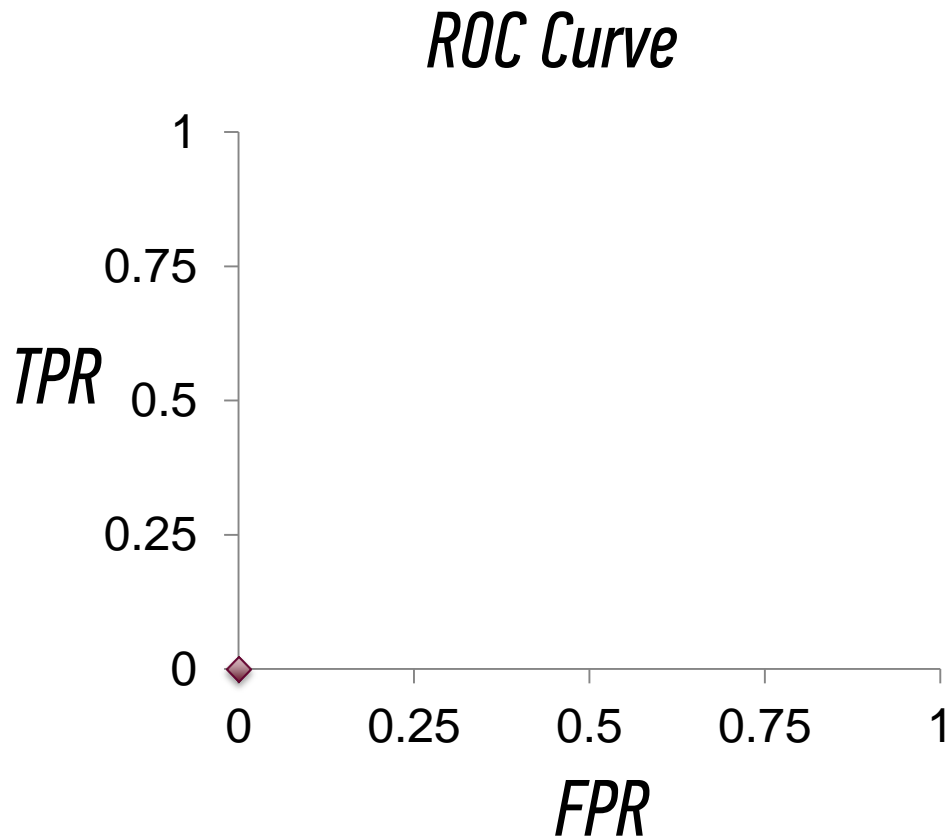


Email Number	Score	True Label
5	0.99	Spam
8	0.82	Spam
2	0.60	Spam
1	0.60	Ham
7	0.48	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham

*Every email is assigned a “spamminess” score by our classification algorithm. To actually make our predictions, we choose a numeric cutoff for classifying as spam.*

*An ROC Curve will help us to visualize how well our classifier is doing without having to choose a cutoff!*

Email Number	Score	True Label
5	0.99	Spam
8	0.82	Spam
2	0.60	Spam
1	0.60	Ham
7	0.48	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham



Email Number	Score	True Label
5	0.99	Spam
8	0.82	Spam
2	0.60	Spam
1	0.60	Ham
7	0.48	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham

TPR: When actual value is **spam**, how often is prediction **correct**?

FPR: When actual value is **ham**, how often is prediction **wrong**?

Cutoff	TPR (y)	FPR (x)	Cutoff	TPR (y)	FPR (x)
0			0.50		
0.05			0.65		
0.15			0.85		
0.25			1		

Email Number	Score	True Label
5	0.99	Spam
8	0.82	Spam
2	0.60	Spam
1	0.60	Ham
7	0.48	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham

TPR: When actual value is **spam**, how often is prediction **correct**?

FPR: When actual value is **ham**, how often is prediction **wrong**?

Cutoff	TPR (y)	FPR (x)	Cutoff	TPR (y)	FPR (x)
<b>0</b>	1	1	<b>0.50</b>	0.75	0.25
<b>0.05</b>	1	0.75	<b>0.65</b>	0.5	0
<b>0.15</b>	1	0.5	<b>0.85</b>	0.25	0
<b>0.25</b>	1	0.25	<b>1</b>	0	0

Email Number	Score	True Label
5	0.99	Spam
8	0.98	Spam
2	0.97	Spam
1	0.97	Ham
7	0.96	Spam
3	0.95	Ham
4	0.94	Ham
6	0.93	Ham

*Q: Would the ROC Curve (and AUC) change if the **scores** changed but the **ordering** remained the same?*

*A: Not at all! The ROC Curve is only sensitive to **rank ordering** and does not require **calibrated scores**.*