

Evaluación de Clasificadores ROC - AUC

MARIA DE LOS ANGELES CONSTANTINO GONZALEZ



Tecnológico
de Monterrey



Métricas de Evaluación para Problemas de Clasificación

▪ Exactitud de clasificación (razón de aciertos) / error

- Exactitud de Clasificación es el porcentaje de predicciones correctas (entre más alto mejor)
- Error de Clasificación es el porcentaje de predicciones incorrectas (entre más pequeño mejor)

▪ Matriz de Confusión.

- Proporciona un mejor entendimiento de cómo se está desempeñando el clasificador.
- Permite calcular sensibilidad, especificidad, precision, valor f1 que podrían coincidir mejor con el objetivo del negocio que la medida de exactitud

• Curvas ROC y Área Bajo la Curva (Area Under the Curve AUC)

- Permite visualizar el desempeño del clasificador en todos los umbrales de clasificación posibles, lo que ayuda a elegir un umbral que equilibre adecuadamente la sensibilidad y la especificidad
- Sigue siendo útil cuando hay alto desbalance en las clases (a diferencia de la exactitud de clasificación / error)
- Más difícil de usar cuando hay más de dos clases de respuesta

Métricas a partir de la matriz de confusión

n=192		Predicted: 0	Predicted: 1	
Actual: 0		TN = 118	FP = 12	130
Actual: 1		FN = 47	TP = 15	62
		165	27	

Especificidad

Razón de Verdaderos Negativos

$$e = \text{TN} / (\text{TN} + \text{FP})$$

Sensibilidad

Razón de Verdaderos Positivos

$$r = \text{TP} / (\text{TP} + \text{FN})$$

Exactitud

$$a = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FN} + \text{FP})$$

Precisión:

Razón de Predicciones Positivas Correctas

$$p = \text{TP} / (\text{TP} + \text{FP})$$

Curva ROC

(Receive operating characteristics curve)

- Es una gráfica de la relación entre la sensibilidad, razón de verdaderos positivos (TPR) con respecto a la razón de falsos positivos (FPR).
- Sensibilidad: Razón de verdaderos positivos (True positive rate)

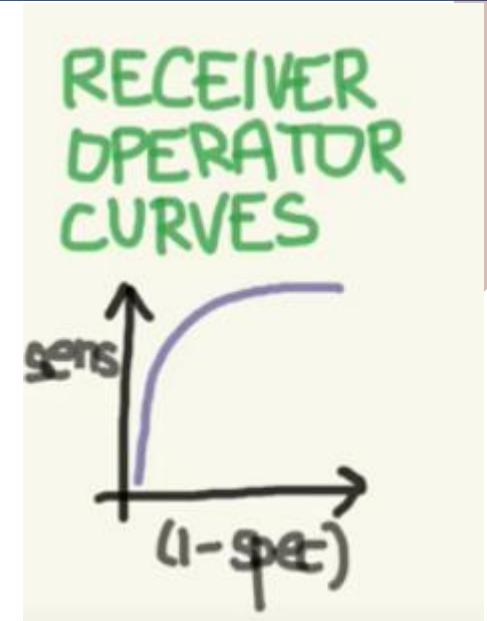
$$TPR = \frac{TP}{TP + FN}$$

- Razón de Falsos positivos (False positive rate) -"Falsa alarma"
(1 – Especificidad (Razón de Verdaderos Negativos (TNR)))

$$FPR = \frac{FP}{TN + FP}$$

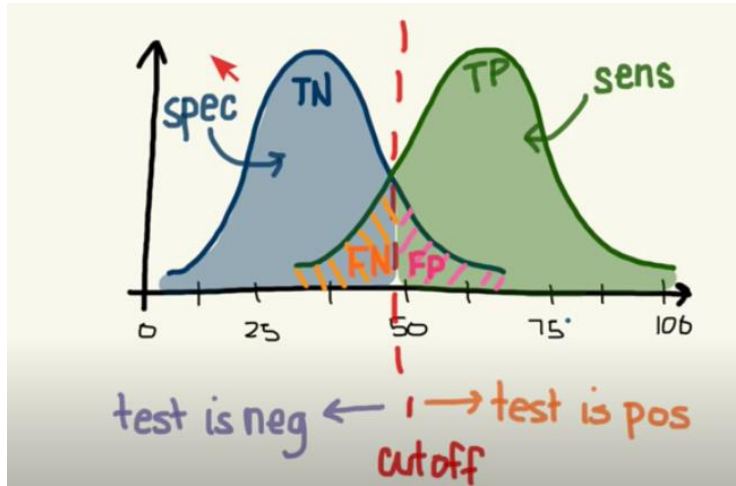
Especificidad

$$TNR = \frac{TN}{TN + FP}$$

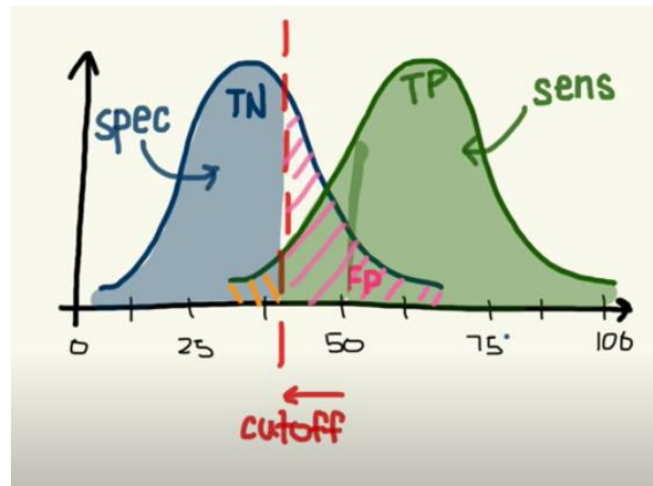


Sensibilidad VS Especificidad

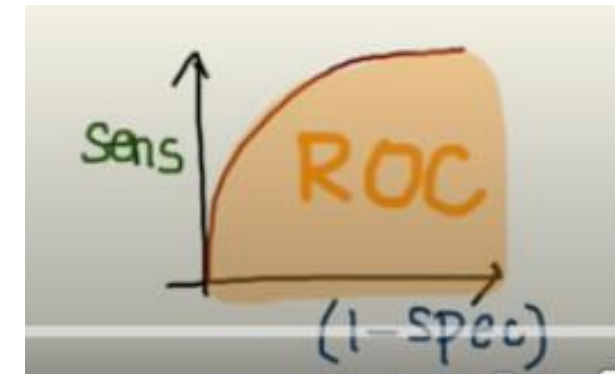
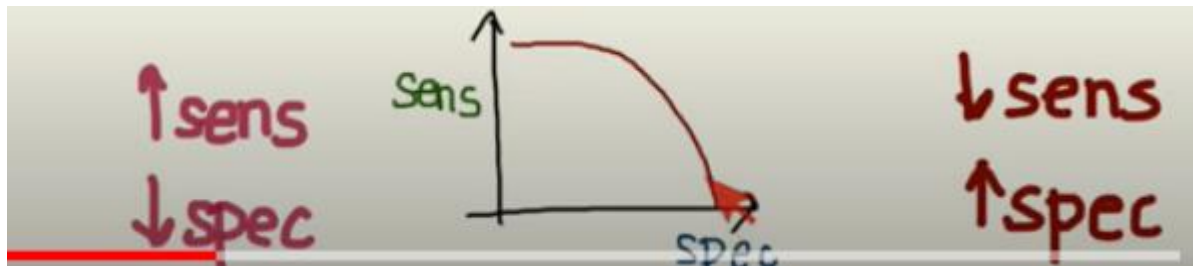
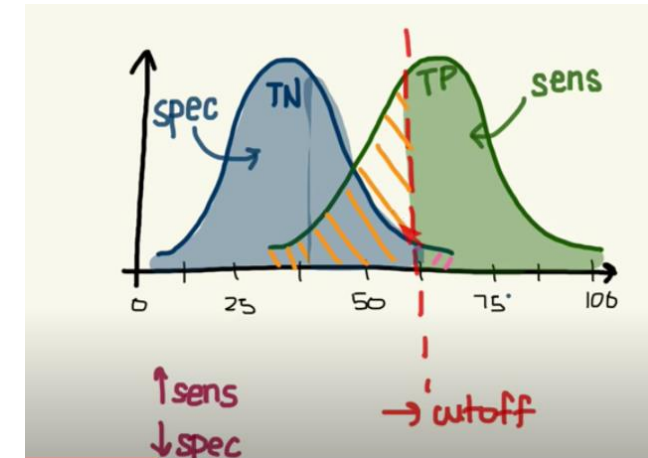
Ajuste del Umbral (Threshold) (cutoff)



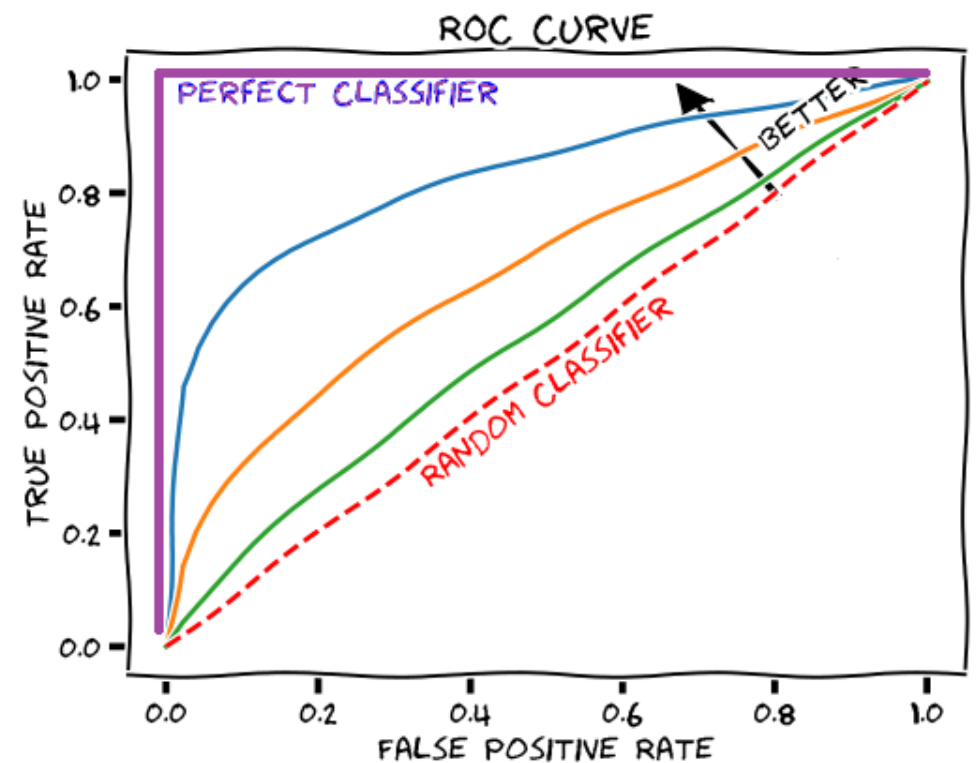
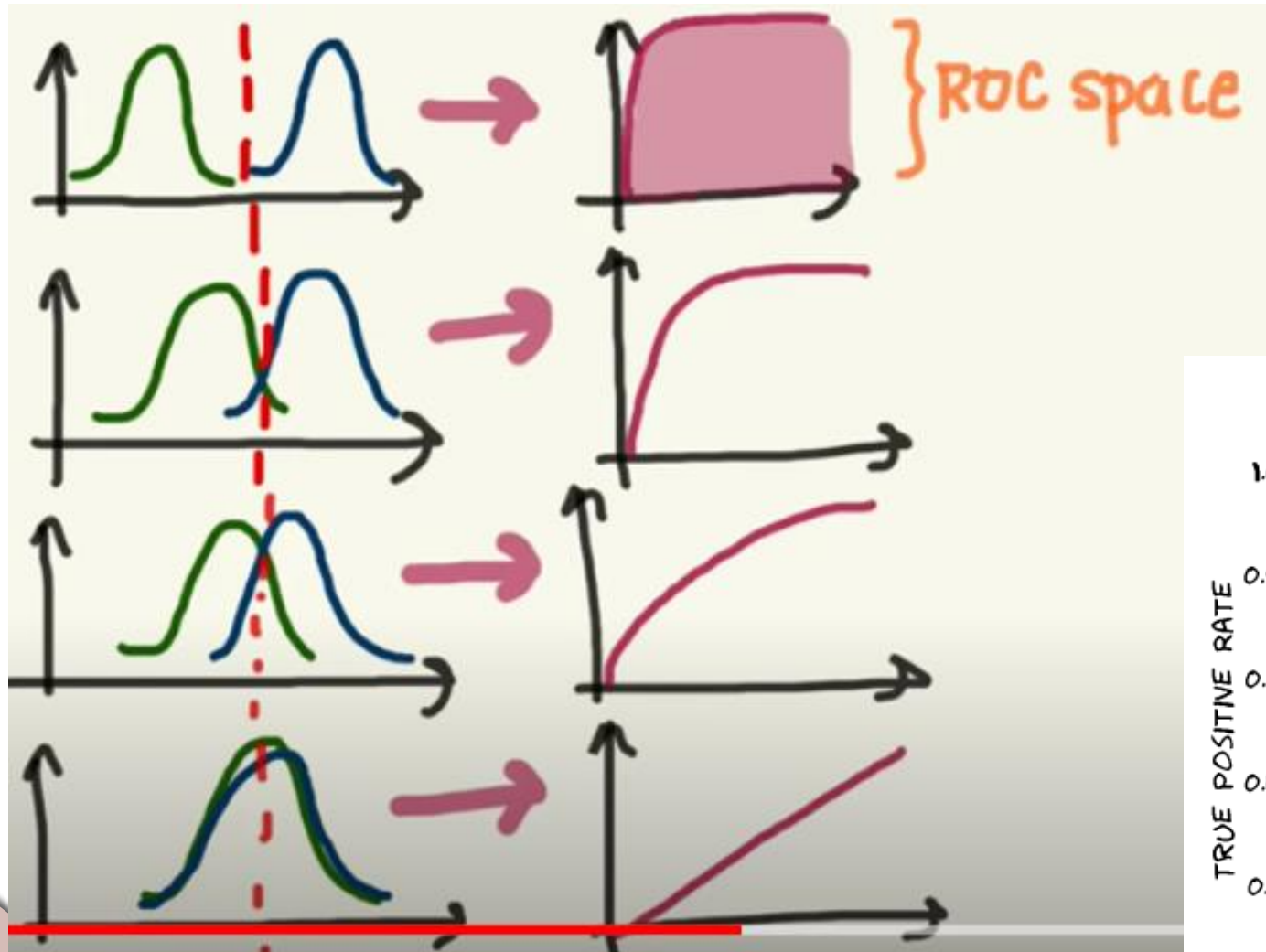
Sensibilidad aumenta,
Especificidad disminuye



Especificidad aumenta,
Sensibilidad disminuye



Curva ROC



<https://www.youtube.com/watch?v=21lgj5Pr6u4>

Curva Roc – Módulo sklearn

```
y_pred_prob_pos = modelo.predict_proba(X_test)[:,-1]
```

```
fpr, tpr, thresholds = metrics.roc_curve(y_test, y_pred_prob_pos)
```

Primer argumento: Valores verdaderos,

Segundo argumento: Probabilidades de la clase positiva

Curva ROC

Recibe `y_test` y `y_pred_prob_pos`

Regresa 3 objetos `fpr` (false positive rate), `tpr` (true positive rate) y `thresholds`

```
plt.plot(fpr, tpr)
```

```
plt.xlim([0.0, 1.0])
```

```
plt.ylim([0.0, 1.0])
```

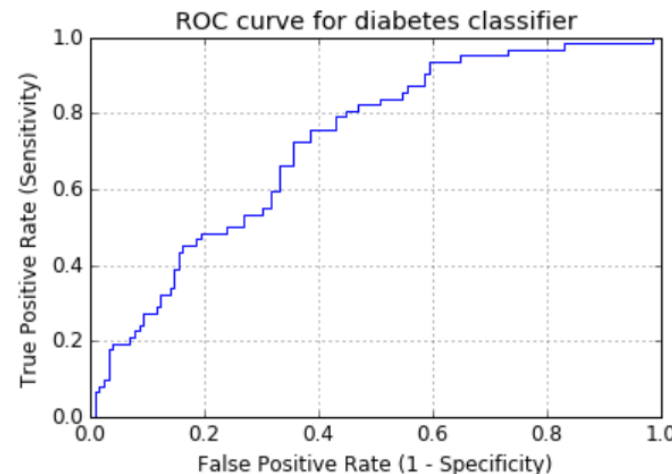
```
plt.rcParams['font.size'] = 12
```

```
plt.title('ROC curve for diabetes classifier')
```

```
plt.xlabel('False Positive Rate (1 - Specificity)')
```

```
plt.ylabel('True Positive Rate (Sensitivity)')
```

```
plt.grid(True)
```



La curva ROC puede ayudar a elegir un umbral que equilibre la sensibilidad y la especificidad de una manera que tenga sentido para el contexto particular.

Métrica: Área bajo la curva (AUC)

¿Qué clasificador es mejor C_1 o C_2 ?

- Depende de la región a la que se refiera.

¿Se puede calcular una sola medida que lo indique?

- Sí, calcular el área bajo la curva (AUC)

AUC es útil como una medida sumativa para en un sólo número evaluar el desempeño del clasificador

- Valor AUC más alto indica Mejor clasificador

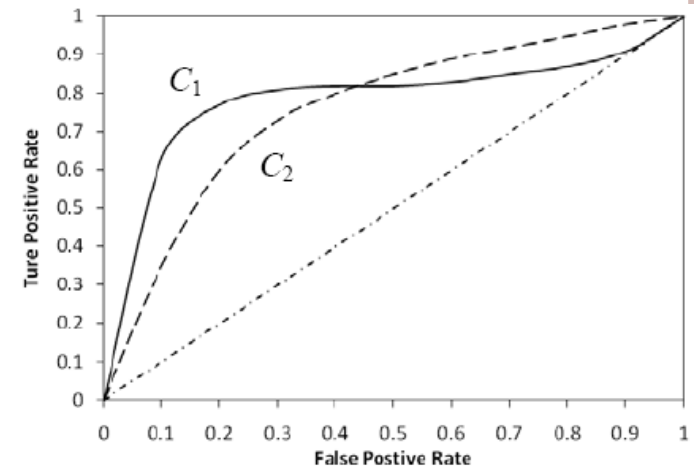


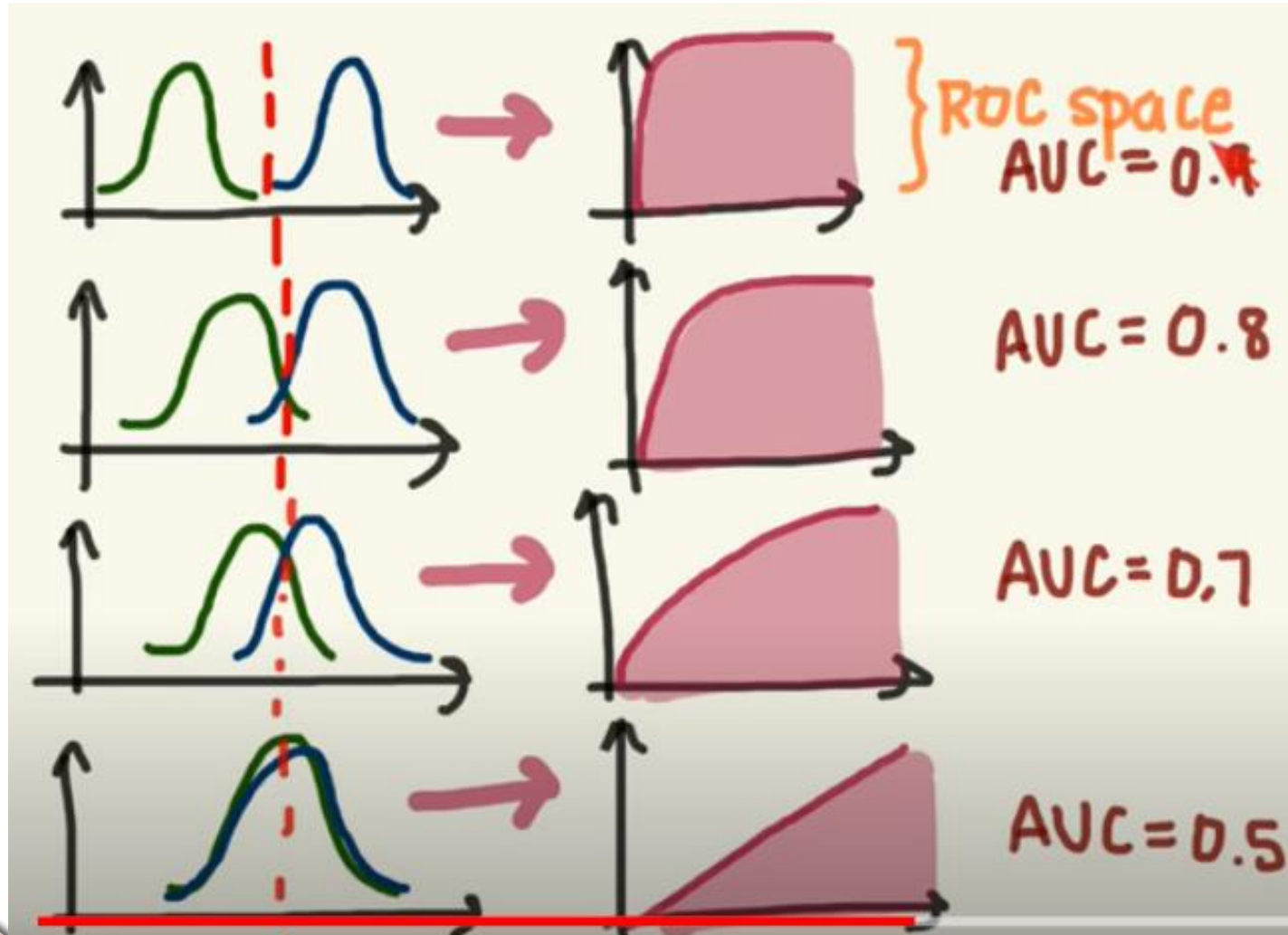
Fig. 3.8. ROC curves for two classifiers (C_1 and C_2) on the same data

Métrica:

Área bajo la curva (AUC)

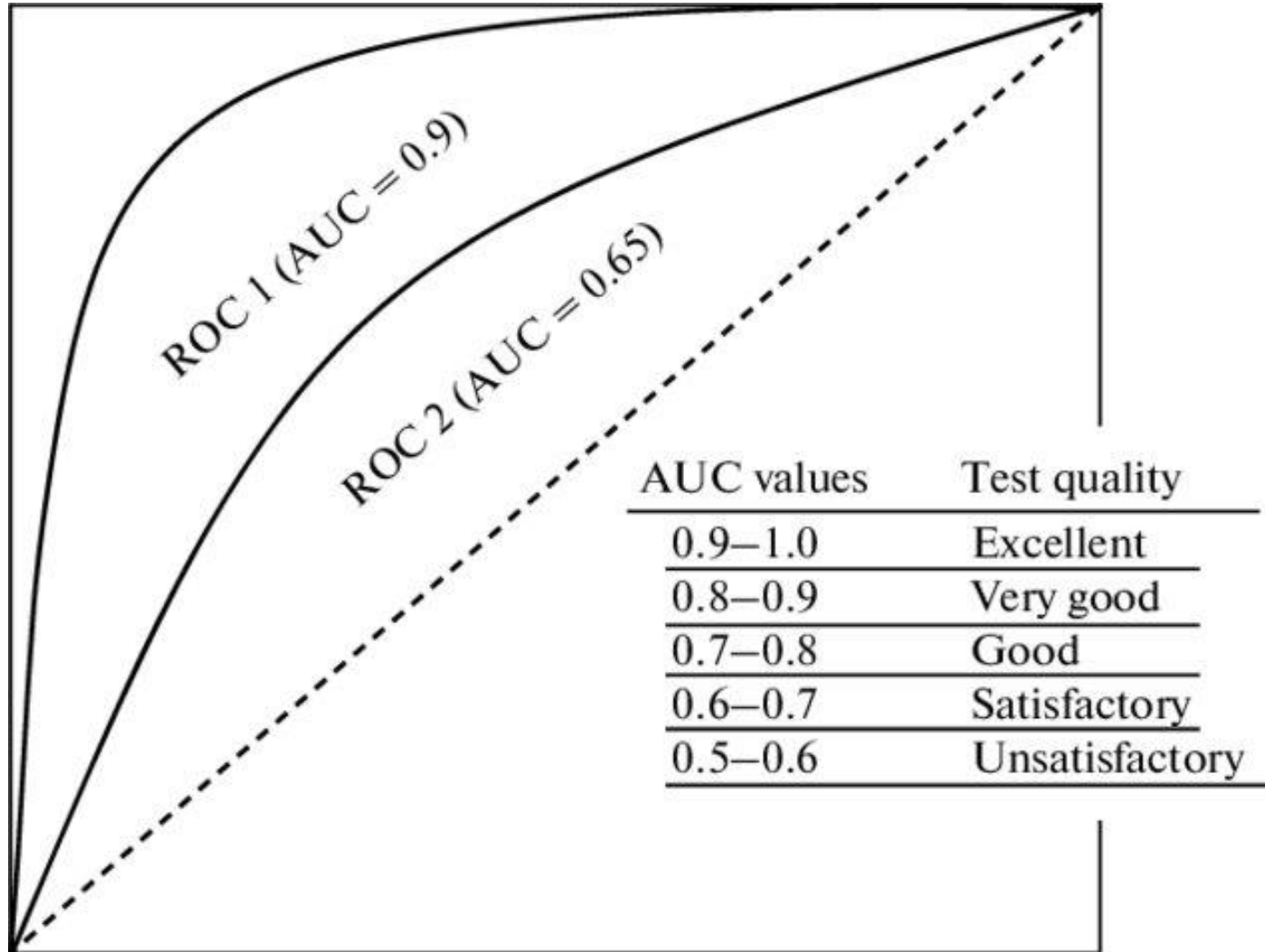
- AUC es el **porcentaje** que está **debajo de la curva ROC**
- Si AUC para C_i es mayor que la de C_j , se dice que C_i es mejor que C_j
 - Si un clasificador es perfecto, su valor AUC es 1
 - Si un clasificador hace todas las predicciones aleatorias, su valor AUC es 0.5.
- **AUC sklearn metric**
IMPORTANTE: el primer argumento son los valores verdaderos, el segundo argumento son las probabilidades generadas.
- `print(metrics.roc_auc_score(y_test, y_pred_prob))`
- 0.724565756824

ROC: AUC

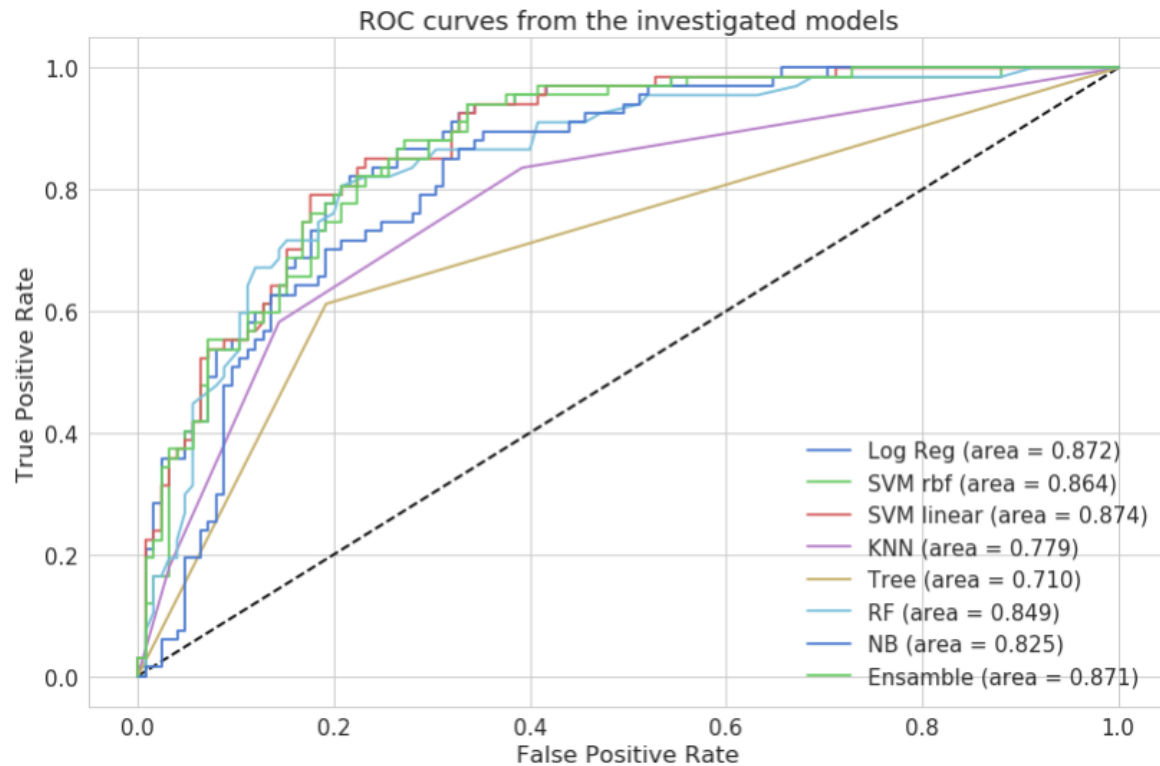


<u>AUC</u>	<u>Quality Of Test</u>
0.9-1	Excellent
0.8-0.9	Good
0.7-0.8	Fair
0.6-0.7	Poor
0.5-0.6	Fail

ROC - AUC



Varias curvas ROC en una gráfica usando sklearn



```
# Logistic regression
modelLR = LogisticRegression()
modelLR.fit(train_X, train_Y)
y_pred_prob_lr = modelLR.predict_proba(test_X)[:,-1]
fpr_lr, tpr_lr, thresholds_lr = roc_curve(test_Y, y_pred_prob_lr)
roc_auc_lr = auc(fpr_lr, tpr_lr)
roc_auc_score
```

```
plt.plot(fpr_lr, tpr_lr, label='Log Reg (area = %0.3f)' % roc_auc_lr)
```

Scikit Learn - functions

Scoring	Function
Classification	
'accuracy'	<u>metrics.accuracy_score</u>
'f1'	<u>metrics.f1_score</u>
'precision'	<u>metrics.precision_score</u>
'recall'	<u>metrics.recall_score</u>
'roc_auc'	<u>metrics.roc_auc_score</u>

References

- Ramesh Sharda; Dursun Delen; Efraim Turban. . Chapter 4. Predictive Analysis. In Business, Analytics and Data Science, 4^a. Ed. Pearson, 2017
- <https://www.ritchieng.com/machine-learning-evaluate-classification-model/>
- Videos: [Intuitive sensitivity and specificity](https://www.youtube.com/watch?v=U4_3fditnWg) https://www.youtube.com/watch?v=U4_3fditnWg
- Video: [Tradeoff between Sensitivity and Specificity](#)
- Video: [ROC Curves](#)
- https://scikit-learn.org/stable/modules/model_evaluation.html
- <https://www.ibm.com/garage/method/practices/reason/evaluate-and-select-machine-learning-algorithm/>
- <https://glassboxmedicine.com/2019/02/23/measuring-performance-auc-auroc/>
<https://www.kaggle.com/code/mnpathak1/model-comparison-with-roc-curves-and-more/notebook>