

Ejercicio Práctico en Orange: Clasificación de Clientes (Regresión Logística)

1. Muestren la tabla de métricas del Test and Score. ¿Qué modelo logró el mejor AUC y cuál el mejor F1-score?

- El modelo con mayor valor de AUC fue *Logistic Regression*.
- El valor más alto de F1-score también fue logrado por *Logistic Regression*.

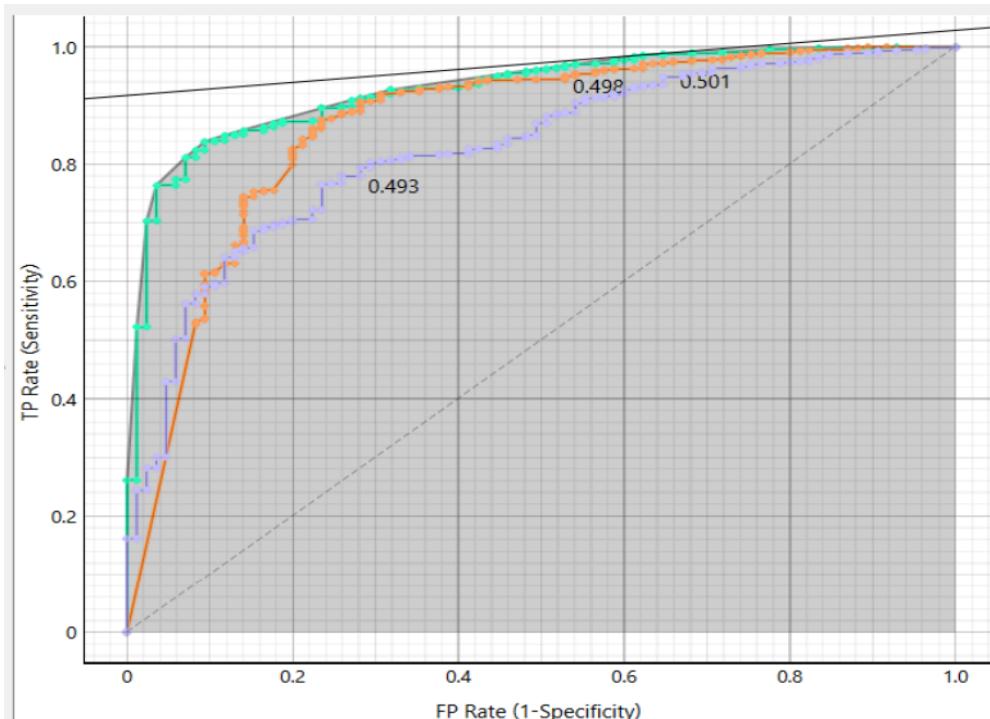
Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.925	0.915	0.910	0.908	0.915	0.498
Random Forest	0.867	0.908	0.896	0.892	0.908	0.398
Naive Bayes	0.825	0.784	0.820	0.891	0.784	0.353

2. A partir de la Matriz de Confusión de la Regresión Logística, identifiquen y reporten los valores de TP, TN, FP y FN.

- **Verdaderos Positivos (TP): 41**
- **Verdaderos Negativos (TN): 712**
- **Falsos Positivos (FP): 26**
- **Falsos Negativos (FN): 44**

		Predicted		
		no	yes	Σ
Actual	no	712	26	738
	yes	44	41	85
Σ		756	67	823

3. Analicen la curva ROC: ¿Cuál de los tres modelos exhibe el mejor trade-off entre Tasa de Verdaderos Positivos (TPR) y Tasa de Falsos Positivos (FPR)? Justifiquen su respuesta basándose en la gráfica.



A. Cálculo de Métricas

Incluyan fórmula y resultado con hasta 4 decimales:

1. Precisión (Precision)

$$\text{Precision} = \text{VP} / (\text{VP} + \text{FP}) = 350 / (350 + 100) = 0,7777\ 7$$

2. Exhaustividad (Recall o Sensibilidad)

$$\text{Recall} = \text{VP} / (\text{VP} + \text{FN}) = 350 / (350 + 150) = 0,7$$

3. Exactitud (Accuracy)

$$(VP + VN) / (VP + VN + FP + FN) = (350 + 100) / (350 + 100 + 150 + 4400) = 0,09$$

4. Puntaje F1 (F1-score)

$$F1=2 \times (\text{Precision} \cdot \text{Recall}) / (\text{Precision} + \text{Recall}) = 2 * (0.7777 * 0.7) / (0.7777 + 0.7) = 0.7368$$

B. Análisis Estratégico

1. Impacto de un FP (Falso Positivo):

Si el modelo pronostica que un cliente abandonará pero en realidad no lo hace, la empresa puede incurrir en gastos innecesarios al ofrecer incentivos a un cliente que no los necesita.

2. Impacto de un FN (Falso Negativo):

En este caso, se considera que el cliente seguirá pero termina abandonando. Esto puede afectar significativamente la retención y ocasionar costos por pérdida de ingresos y adquisición de nuevos usuarios.

3. Métrica de Prioridad:

Si el objetivo es evitar gastos superfluos en clientes que iban a quedarse, la métrica clave sería la **Precisión**, ya que minimiza los falsos positivos.