

## Ejercicio práctico nº 11 Resultados

		Equipo		8		
Instrucción	Categoría	Nivel de logro	Puntos	Ideal	Obtenido	Observaciones
	Datos	Obtienen una muestra de datos pertinentes para crear y evaluar modelos de regresión, cumpliendo todas las restricciones indicadas en el enunciado (semilla, tamaño, balanceo, etc.).  <b>IMC: peso (en kilogramos) dividida por estatura (en metros) al cuadrado.</b> <b>EN: sobrepeso (<math>IMC \geq 23,2</math>) y no sobrepeso (<math>IMC &lt; 23,2</math>).</b> <b>Pertinente: (1) debe incluir la variable de respuesta correctamente calculada y no incluir las variables evidentemente correlacionadas con ella, (2) las observaciones deben aparecer revueltas.</b> <b>Semilla: número de cinco dígitos.</b> <b>Tamaño: 100.</b> <b>Balanceo: 50% con sobrepeso.</b>	3	3	2	Usan una semilla de 4 dígitos (no de 5 dígitos)
		Obtienen una muestra de datos pertinente para poder crear y evaluar modelos de regresión, cumpliendo al menos dos de las restricciones indicadas en el enunciado.	2			
		Obtienen una muestra útil de datos para poder crear y evaluar modelos de regresión.	1			
		No responden, o el código está muy incompleto, o no se puede seguir, o no obtienen una muestra útil de los datos necesarios.	0			
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Selección de predictores	Escriben código en R correcto que utiliza funciones de los paquetes <code>leaps</code> y <code>caret</code> para explorar el conjunto válido de potenciales predictores, según las instrucciones del enunciado, y que escoge un subconjunto de ellos, respetando el tamaño indicado, de acuerdo al criterio de búsqueda solicitado y la técnica de muestreo especificada para evitar el sobreajuste (si corresponde).  <b>RLM 1: exhaustiva de subconjuntos (Leaps), BIC u otro, no aplica muestreo, peso ~ 2-8 predictores.</b> <b>RLM 2: RFE (caret), <math>R^2</math>, 5 repeticiones de validación cruzada de 5 pliegues (5x5CV), IMC ~ 10-20 predictores.</b> <b>RLogitM: RFE (caret), ROC-AUC, validación cruzada dejando uno fuera (LOOCV), EN ~ 2-6 predictores.</b>	4	12	12	
		Escriben código en R correcto que utiliza funciones de los paquetes <code>leaps</code> y <code>caret</code> para explorar el conjunto válido de potenciales predictores, según las instrucciones del enunciado, y que escoge un subconjunto de ellos, respetando dos de las condiciones solicitadas (tamaño, criterio de búsqueda, técnica de muestreo si corresponde).	3			
		Escriben código en R correcto que utiliza funciones de los paquetes <code>leaps</code> y <code>caret</code> para explorar un conjunto de potenciales predictores, y que escoge un subconjunto de ellos respetando el criterio de búsqueda o la técnica de muestreo solicitados.	2			
		Escriben código en R correcto que escoge un subconjunto de predictores útiles para construir el modelo de regresión solicitado.	1			
		No responden, o el código está muy incompleto, o no se puede seguir, o no se observa un proceso de selección de variables útiles para poder construir el modelo de regresión solicitado.	0			
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Modelo	Construyen correctamente un modelo de regresión pertinente para predecir la variable de respuesta indicada, usando como predictores las variables seleccionadas por el proceso de búsqueda y aplicando la técnica de muestreo para evitar el sobreajuste especificada, reproduciendo esta última las que mantiene cada vez que actualiza el modelo.  <b>RLM 1: bootstrapping.</b> <b>RLM 2: 5x5CV, la primera vez con RFE.</b> <b>RLogitM: LOOCV, la primera vez con RFE.</b>	3	9	9	
		Construyen correctamente un modelo de regresión pertinente para predecir la variable de respuesta indicada, usando como predictores las variables seleccionadas por el proceso de búsqueda y aplicando la técnica de muestreo para evitar el sobreajuste especificada, pero solo la primera vez.	2			
		Construyen correctamente un modelo de regresión pertinente para predecir la variable de respuesta indicada.	1			
		No responden, o el código está muy incompleto, o no se puede seguir, o no obtienen un modelo de regresión pertinente para predecir la variable de respuesta indicada.	0			
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Bondad de ajuste	Escriben comentarios y código en R correcto que verifica las condiciones que garantizan que el modelo obtenido tiene un buen nivel de ajuste, interpretando explícita y correctamente los resultados obtenidos en cada paso y tomando acciones correctivas apropiadas de ser necesarias o indicando los riesgos asociados a utilizarlos sin corregir estos problemas.  <b>(1) Disminución significativa de la varianza en RLM y de la devianza en RLogitM, (2) residuos independientes y (3) sin patrones, (4) relaciones lineales y (5) homocedásticas entre los predictores y la variable de respuesta en RLM, o (4) relaciones lineales entre los predictores y el logit de las probabilidades de la variable de respuesta.</b>	4	12	9	En el modelo de peso faltó comprobar la linealidad de las relaciones entre los predictores y la variable respuesta.  En el modelo de IMC faltó incluir las relaciones lineales entre los predictores seleccionados y la variable dependiente IMC.  Y en el modelo de EN faltó evaluar explícitamente la disminución significativa de la devianza.  <b>(3 + 3 + 3 = 9 puntos)</b>
		Escriben comentarios y código en R correcto que verifica al menos tres de las condiciones que garantizan que el modelo obtenido tiene un buen nivel de ajuste, tomando acciones correctivas de ser necesarias o comentando los riesgos asociados a utilizarlo sin corregir estos problemas.	3			
		Escriben código en R correcto que verifica al menos dos de las condiciones que garantizan que el modelo obtenido tiene un buen nivel de ajuste.	2			
		Escriben código en R correcto que verifica al menos una de las condiciones que garantizan que el modelo obtenido tiene un buen nivel de ajuste.	1			
		No responden, o el código está muy incompleto, o no verifica al menos dos de las condiciones que garantizan que el modelo obtenido tiene un buen nivel de ajuste.	0			

## Ejercicio práctico nº 11 Resultados

			Equipo		8		
Instrucción	Categoría	Nivel de logro	Puntos	Ideal	Obtenido	Observaciones	
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Generalidad	Escriben comentarios y código en R correcto que verifica las condiciones que garantizan que el modelo obtenido es generalizable, interpretando explícita y correctamente los resultados obtenidos en cada paso y tomando acciones correctivas apropiadas de ser necesarias o comentando los riesgos asociados a utilizarlos sin corregir estos problemas.	4	12	9	En el modelo de peso los diagnósticos de multicolinealidad y análisis de influencia no están explícitamente detallados.	
		<b>(1) No hay multicolinealidad problemática, (2) no hay presencia de casos sobreinfluyentes, (3) no hay indicios de sobreajuste.</b>					
		Escriben comentarios y código en R correcto que verifica las condiciones que garantizan que el modelo obtenido es generalizable, tomando acciones correctivas apropiadas de ser necesarias o comentando los riesgos asociados a utilizarlos sin corregir estos problemas.	3			En el modelo de IMC están los diagnósticos básicos de generalidad presentes, pero sin exhaustividad.	
		Escriben comentarios y código en R correcto que verifica al menos dos de las condiciones que garantizan que el modelo obtenido es generalizable, comentando las posibles acciones correctivas o los riesgos asociados a utilizarlos sin corregir estos problemas.	2			Y el modelo de EN incluye validación adecuada pero faltan diagnósticos adicionales como la multicolinealidad problemática.	
		Escriben comentarios y código en R correcto que verifica al menos una de las condiciones que garantizan que el modelo obtenido es generalizable	1			<b>(3 + 3 + 3 = 9 puntos)</b>	
No responden, o el código está muy incompleto, o no verifica al menos una de las condiciones que garantizan que el modelo obtenido es generalizable.	0						
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Calidad predictiva	Escriben código R correcto que evalúa la calidad predictiva del modelo obtenido en datos no usados para construirlo, usando una o más métricas pertinentes y caracterizando su distribución si corresponde.	4	12	10	En el modelo de Peso el RMSE es calculado con bootstrapping, pero falta análisis detallado de su distribución.	
		<b>Pertinente:</b> <b>RLM: una medida del error (MSE, RMSE, MAD, etc.) que comete el modelo.</b> <b>RLogitM: una métrica apropiada para clasificación (exactitud, sensibilidad, especificidad, área bajo la curva ROC).</b> <b>Distribución de la métrica: corresponde con bootstrapping y 5x5CV, no para LOOCV.</b>					
		Escriben código R correcto que evalúa la calidad predictiva del modelo obtenido en datos no vistos para construirlo, usando una métrica pertinente.				3	En el modelo de IMC el RMSE es calculado con validación 5x5CV, sin análisis de distribución.
		Escriben código R correcto que evalúa la calidad predictiva de ambos modelos obtenidos, usando una métrica pertinente en datos no utilizados para su construcción.				2	<b>(3 + 3 + 4 = 10 puntos)</b>
		Escriben código R correcto que evalúa la calidad predictiva de alguno de los modelos solicitados.				1	
El código está muy incompleto, no se puede seguir o no permite evaluar la calidad predictiva de alguno de los modelos obtenidos.	0						
Aplicar puntaje por cada modelo según nivel de logro en cada uno.	Conclusión	Entregan conclusiones correctas sobre el ajuste, generalidad y calidad predictiva del modelo obtenido, basadas en las evaluaciones y arreglos realizados.	3	9	7	En el modelo de EN, las conclusiones están centradas en calidad predictiva, pero falta concluir sobre generalidad y ajuste.	
		Entregan conclusiones correctas sobre dos de los aspectos solicitados del modelo obtenido, basadas en las evaluaciones y arreglos realizados.	2			<b>(3 + 3 + 1 = 7 puntos)</b>	
		Entregan conclusiones correctas sobre uno de los aspectos solicitados del modelo obtenido, basadas en las evaluaciones y arreglos realizados.	1				
		No hay conclusiones o estas no se entienden o no se basan en lo realizado.	0				
	Código fuente	El script está completo, ordenado y bien indentado, es fácil de seguir y no requiere cambios para que funcione.	5	5	5		
		El script está bastante completo (> 80%) y ordenado, es fácil de seguir y funciona cambiando un detalle (como olvidar cargar algún paquete).	3				
		El script está relativamente completo (>50 %), se puede seguir con relativa facilidad y funciona cambiando varios detalles.	1				
		No responden, o el script está muy incompleto como para evaluar buenas prácticas o cuesta seguirlo o requiere cambios importantes para poder ejecutarse.	0				
	Ortografía y redacción	El script está comentado paso a paso, con claridad (basta una lectura para entender) y con buena redacción y ortografía (<= 5 errores), usando vocabulario propio de la disciplina y el contexto del problema	4	4	4		
		El script está bien comentado, usando vocabulario propio de la disciplina, con claridad y ortografía y redacción aceptables (<=10 errores)	2				
		El script contiene comentarios en los pasos importantes y con menos de 15 errores de ortografía y redacción	1				
		El script no contiene muchos comentarios o estos contienen muchos errores de ortografía y redacción (> 15)	0				
TOTAL			78	67			