



PROCEDIMIENTO DE EVALUACIÓN Y ACREDITACIÓN DE LAS COMPETENCIAS PROFESIONALES

CUESTIONARIO DE AUTOEVALUACIÓN PARA LAS TRABAJADORAS Y TRABAJADORES

UNIDAD DE COMPETENCIA

“UC2493_3: Entrenar modelos en sistemas de inteligencia artificial basados en aprendizaje automático”

LEA ATENTAMENTE LAS INSTRUCCIONES

Conteste a este cuestionario de **FORMA SINCERA**. La información recogida en él tiene **CARÁCTER RESERVADO**, al estar protegida por lo dispuesto en la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Su resultado servirá solamente para ayudarle, **ORIENTÁNDOLE** en qué medida posee la competencia profesional de la “UC2493_3: Entrenar modelos en sistemas de Inteligencia Artificial basados en aprendizaje automático”.

No se preocupe, con independencia del resultado de esta autoevaluación, Ud. **TIENE DERECHO A PARTICIPAR EN EL PROCEDIMIENTO DE EVALUACIÓN**, siempre que cumpla los requisitos de la convocatoria.

Nombre y apellidos del trabajador/a: NIF:	Firma:
Nombre y apellidos del asesor/a: NIF:	Firma:

INSTRUCCIONES CUMPLIMENTACIÓN DEL CUESTIONARIO:

Las actividades profesionales aparecen ordenadas en bloques desde el número 1 en adelante. Cada uno de los bloques agrupa una serie de actividades más simples (subactividades) numeradas con 1.1., 1.2.,..., en adelante.

Lea atentamente la actividad profesional con que comienza cada bloque y a continuación las subactividades que agrupa. Marque con una cruz, en los cuadrados disponibles, el indicador de autoevaluación que considere más ajustado a su grado de dominio de cada una de ellas. Dichos indicadores son los siguientes:

1. No sé hacerlo.
2. Lo puedo hacer con ayuda.
3. Lo puedo hacer sin necesitar ayuda.
4. Lo puedo hacer sin necesitar ayuda, e incluso podría formar a otro trabajador o trabajadora.

1: Buscar correlaciones entre las variables utilizando herramientas que incorporan técnicas de la estadística y el aprendizaje automático, con anterioridad al diseño y entrenamiento de modelos y siguiendo las indicaciones proporcionadas por el científico de datos responsable.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
1.1: Buscar las correlaciones entre variables, tanto lineales como no lineales, aplicando técnicas de la estadística y del aprendizaje automático ('machine learning'- ML-).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.2: Aplicar los procedimientos de categorización y codificación de las variables tales como mapas de bits ('one-hot vectors') u otros mediante software o herramientas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.3: Aplicar las transformaciones matemáticas tales como logaritmos, exponenciales, funciones trigonométricas u otras antes del entrenamiento de los modelos, teniendo en cuenta la distribución de sus valores dentro del conjunto de datos.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.4: Seleccionar las variables de entrada más explicativas para tenerlas en cuenta durante el diseño de los modelos, considerando aquellas que tienen mayor correlación con las variables de salida o variables objetivo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.5: Redactar el documento que describe las variables, incluyendo todos los detalles analizados en relación a las variables tanto de entrada como de salida y a las correlaciones encontradas, con el fin de servir para tomar decisiones durante el diseño de los modelos y para ser integrado en un informe final.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2: Reducir la dimensión de las muestras de los conjuntos de datos, mediante programación o herramientas software, para obtener una representación de los mismos mediante variables latentes siguiendo instrucciones de la persona responsable.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
2.1: Utilizar las técnicas de reducción de la dimensión, tales como análisis de componentes principales -PCA-, t-SNE, autocodificadores ('autoencoders') basados en redes neuronales y otras, disponibles como funcionalidades en herramientas software existentes, eligiendo aquellas que realizan una mayor reducción de la dimensión mientras conservan la mayor variabilidad (capacidad explicativa) de las muestras.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.2: Programar los métodos y procedimientos, usando el lenguaje definido para el proyecto, realizando la reducción de la dimensión de las muestras con objeto de crear una copia del conjunto de datos con las transformaciones ya aplicadas, o bien aplicando las transformaciones durante el entrenamiento, antes de que las muestras se utilicen como entrada a los modelos para su entrenamiento y evaluación.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.3: Documentar el trabajo realizado, indicando: - Las técnicas de reducción de la dimensión aplicadas al conjunto de datos con el que se esté trabajando, - Los valores concretos de los parámetros de configuración de cada técnica, - Los resultados obtenidos en cuanto a reducción de la dimensión y pérdida de explicación de la variabilidad, enfatizando el balance entre la proporción de reducción y la pérdida de variabilidad en función de uno o más parámetros.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3: Visualizar los datos, representándolos gráficamente, con la finalidad de corroborar las correlaciones encontradas, verificando la reducción de la dimensión aplicada sobre el conjunto de datos con los que se esté trabajando.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
3.1: Generar las gráficas, escribiendo programas al efecto, utilizando varias representaciones, con el fin de verificar tanto las correlaciones encontradas como las transformaciones aplicadas para reducir la dimensión de las muestras, bajo la supervisión de la persona responsable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.2: Elaborar la información recogida mediante gráficas, usando medios didácticos tales como diapositivas o infografías, organizándola para presentarla al equipo de trabajo, con el fin de sacar conclusiones de manera conjunta y así decidir colectivamente qué subconjunto de todas las transformaciones	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3: Visualizar los datos, representándolos gráficamente, con la finalidad de corroborar las correlaciones encontradas, verificando la reducción de la dimensión aplicada sobre el conjunto de datos con los que se esté trabajando.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
probadas se utilizarán para probar su efecto con los distintos modelos a entrenar y evaluar.				
3.3: Recoger las gráficas en un informe, presentándolas en formatos tales como histogramas, mapas de dispersión y otras técnicas de representación, explicando cómo se han obtenido y aportando una valoración de cada gráfica, destacando aspectos cualitativos como la correlación entre las variables latentes fruto de las transformaciones y/o reducciones y las variables salida.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4: Diseñar modelos basados en aprendizaje automático ('machine learning' -ML-) para aplicarlos sobre el conjunto de datos con el objeto de abordar el problema planteado según su tipo -de regresión o de clasificación-, siguiendo las indicaciones proporcionadas por el científico de datos responsable.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
4.1: Elegir las técnicas de aprendizaje automático tales como redes neuronales, máquinas de soporte vectorial, modelos de mixturas de gaussianas, árboles de decisión u otras para diseñar los modelos en base al análisis exploratorio de los datos y el estudio visual realizados previamente.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.2: Ensayar los modelos diseñados, programando código para entrenarlos, realizando experimentos que permitan ver qué combinaciones de parámetros de configuración conducen a mejores resultados, tales como: - El coeficiente de aprendizaje para las redes neuronales ('learning rate'), el número de capas, el número de neuronas en cada capa, los tipos de activación a utilizar en las capas ocultas, - El número de grupos ('clusters') para el algoritmo 'k means', y - El número de árboles a generar en el caso de 'random forests'.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.3: Seleccionar las métricas a utilizar para evaluar modelos tales como: - La desviación porcentual entre valores predichos y reales en problemas de regresión como la predicción del volumen de ventas semanal por producto, - La ratio de falsos positivos y la de falsos negativos para un problema de clasificación, - Otras métricas relacionadas con las ratios anteriores como 'accuracy', 'precision', 'recall' o 'sensitivity', 'specificity', y área bajo la curva ROC ('Receiver Operating Characteristic'), teniendo en cuenta las técnicas a aplicar y el problema a resolver.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4: Diseñar modelos basados en aprendizaje automático ('machine learning' -ML-) para aplicarlos sobre el conjunto de datos con el objeto de abordar el problema planteado según su tipo -de regresión o de clasificación-, siguiendo las indicaciones proporcionadas por el científico de datos responsable.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
4.4: Evaluar los modelos diseñados en base a las métricas seleccionadas, con el fin de decidir, bien la creación de nuevos modelos desde cero, bien el rediseño de modelos anteriores, en función de que los resultados obtenidos se ajusten a los esperados, siguiendo indicaciones de la persona responsable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.5: Incluir los comentarios en el código escrito, detallando el papel que juega una variable, describiendo las funciones y algoritmos con el fin de ser reutilizado en el desarrollo de soluciones a otros problemas similares y siguiendo las normas establecidas por la organización.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.6: Generar la documentación, detallando los modelos diseñados con todas sus variantes, con el fin de que sirva como punto de partida para el informe que incluirá los resultados de evaluar el rendimiento de cada uno de los modelos.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5: Evaluar los modelos diseñados, definiendo un subconjunto de test, cuando el conjunto de datos no lo tuviera previamente definido, y creando nuevas particiones de validación y entrenamiento una vez separado el subconjunto de test.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
5.1: Dividir el conjunto de datos en dos subconjuntos, entrenamiento y test, si no estaban previamente definidos, usando las muestras del subconjunto de test solamente para evaluar el modelo y respetando que el subconjunto de test represente un porcentaje del conjunto de datos completo entre el 10% y el 30%.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.2: Volver a dividir el subconjunto de entrenamiento en dos nuevos subconjuntos, entrenamiento propiamente dicho y validación, escogiendo las muestras de manera aleatoria, pero conservando la misma distribución de clases.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.3: Llevar a cabo la validación cruzada de los modelos, generando distintas particiones de entrenamiento y validación, donde el número de distintas particiones lo indicará la persona responsable, repitiendo el proceso completo de entrenamiento y evaluación del modelo por cada una de las particiones, decidiendo en qué iteración del entrenamiento se guarda la versión del modelo para ser evaluado con el subconjunto de test.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5: Evaluar los modelos diseñados, definiendo un subconjunto de test, cuando el conjunto de datos no lo tuviera previamente definido, y creando nuevas particiones de validación y entrenamiento una vez separado el subconjunto de test.	INDICADORES DE AUTOEVALUACIÓN			
	1	2	3	4
5.4: Evaluar cada modelo entrenado y validado con cada partición con el subconjunto de test, mostrando la media aritmética, la varianza u otra medida de las métricas elegidas para evaluar los modelos a partir de las particiones creadas para la validación cruzada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.5: Crear el informe de la evaluación, tomando como punto de partida el informe con todos los detalles de los modelos diseñados e incluyendo gráficas y tablas que muestren las métricas elegidas, para evaluar los modelos y su evolución durante el entrenamiento con respecto a los subconjuntos de validación y que muestren los resultados de los distintos modelos diseñados y sus variantes, siempre con respecto al mismo subconjunto de test con el fin de poder comparar rápidamente el rendimiento y la robustez de cada modelo y sus variantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>