

Universidad Nacional Abierta y a Distancia
Vicerrectoría Académica y de Investigación
Curso: Análisis de datos
Código: 202016908

Guía de actividades y rúbrica de evaluación – Tarea 5
 Proyecto Análisis de Datos

1. Descripción de la actividad

Tipo de actividad: Independiente	
Momento de la evaluación: Final	
Puntaje máximo de la actividad: 125 puntos	
La actividad inicia el: jueves, 16 de mayo de 2024	La actividad finaliza el: miércoles, 29 de mayo de 2024
Con esta actividad se espera conseguir los siguientes resultados de aprendizaje: Resultado de aprendizaje 2: Aplicar algoritmos de Machine Learning supervisado según el problema, empleando métodos de modelado predictivo como regresión y clasificación	
La actividad consiste en: La actividad consiste en el desarrollo de los siguientes ítems. <ul style="list-style-type: none"> El lenguaje a utilizar es Python, el cual se trabajará mediante Jupyter notebooks, utilizando Anaconda. Descargar el siguiente dataset <p>Dataset – Titanic dataset: https://www.kaggle.com/datasets/yasserh/titanic-dataset Set de datos sobre los pasajeros del Titanic. Permite practicar modelos de clasificación y regresión para predecir quién sobrevivió al naufragio.</p> <p>o descárguelo del entorno de aprendizaje junto a la guía de actividades con el nombre Anexo 6 - Dataset Titanic.zip</p>	

- Seleccionar **uno** de los modelos de Aprendizaje Supervisado vistos en la Tarea 3 (Regresión Lineal, Regresión Logística o Árboles de decisión) y realizar los siguientes pasos:
 1. Realizar un análisis exploratorio de los datos para identificar relaciones entre variables, valores atípicos, tendencias, etc.
 2. Preprocesar los datos limpiándolos, tratando valores faltantes y transformándolos según sea necesario.
 3. Seleccionar las características más relevantes para entrenar el modelo utilizando selección de características.
 4. Dividir el dataset en Train y Test para evaluar correctamente el modelo.
 5. Entrenar el modelo configurando los diferentes hiperparámetros.
 6. Evaluar el desempeño del modelo en el conjunto de Test con métricas como precisión, recall, F1-score, etc.
 7. Realizar las diferentes gráficas que permitan visualizar los resultados del modelo.
 8. Interpretar, analizar y documentar los resultados obtenidos
- Cargar en GitHub el código del modelo diseñado <https://github.com/>

Para el desarrollo de la actividad tenga en cuenta que:

En el entorno de Información inicial debe:

- Revisar la Agenda del Curso, donde se publican las fechas de inicio y cierre de cada actividad para que se realice su entrega en las fechas allí estipuladas.
- Revisar el foro Noticias del Curso y el foro General del Curso.

En el entorno de Aprendizaje debe:

- Revisar el Syllabus del curso
- Ingresar a la Unidad 2 – Aprendizaje Supervisado - Contenidos y referentes bibliográficos y realizar las lecturas recomendadas
- Participar en el Foro de discusión - Tarea 5 – Proyecto Análisis de Datos.

En el entorno de Evaluación debe:

- Entregar el desarrollo de la actividad en formato PDF en el espacio Tarea 5 - Proyecto Análisis de Datos - Rúbrica de evaluación y entrega de la actividad

Evidencias de trabajo independiente:

Las evidencias de trabajo independiente para entregar son:

En el Entorno de Evaluación **Tarea 5 - Proyecto Análisis de Datos**, subir un único archivo en formato PDF el cual debe nombrarse *Tarea5-Nombreestudiante*, ejemplo: *Tarea5-SandraPatiño* y debe contener:

- Portada.
- Introducción (indicando el contenido del trabajo a entregar).
- Objetivos
- Enlace de GitHub donde se pueda consultar el código del modelo diseñado.
- Interpretación de los resultados del modelo
- Conclusiones (Mínimo 3 conclusiones sobre el aprendizaje obtenido en el desarrollo del curso)
- Referentes bibliográficos con Normas APA.

Evidencias de trabajo grupal:

En esta actividad no se requieren evidencias de trabajo grupal.

2. Lineamientos generales para la elaboración de las evidencias de aprendizaje a entregar.

Para evidencias elaboradas **independientemente**, tenga en cuenta las siguientes orientaciones

1. Realizar un reconocimiento del contenido publicado en el Entorno de Aprendizaje (Unidad 2 – Aprendizaje Supervisado - Contenidos y referentes bibliográficos).
2. Intervenga en el foro de discusión aplicando las normas de Netiqueta Virtual, evidenciando siempre respeto por las ideas de sus compañeros y del cuerpo docente.
3. Antes de entregar el producto solicitado revise que cumpla con todos los requerimientos que se señalaron en esta guía de actividades, rúbrica de evaluación y por parte del tutor en el foro de discusión.
4. No cometa fraudes, ni plagios ni actos que atenten contra el normal desarrollo académico de las actividades.

Tenga en cuenta que todos los productos escritos individuales o grupales deben cumplir con las normas de ortografía y con las condiciones de presentación que se hayan definido.

En cuanto al uso de referencias considere que el producto de esta actividad debe cumplir con las normas **APA**

En cualquier caso, cumpla con las normas de referenciación y evite el plagio académico, para ello puede apoyarse revisando sus productos escritos mediante la herramienta Turnitin que encuentra en el campus virtual.

Considere que en el acuerdo 029 del 13 de diciembre de 2013, artículo 99, se considera como faltas que atentan contra el orden académico, entre otras, las siguientes: literal e) “El plagiar, es decir, presentar como de su propia autoría la totalidad o parte de una obra, trabajo, documento o invención realizado por otra persona. Implica también el uso de citas o referencias faltas, o proponer citad donde no haya coincidencia entre ella y la referencia” y liberal f) “El reproducir, o copiar

con fines de lucro, materiales educativos o resultados de productos de investigación, que cuentan con derechos intelectuales reservados para la Universidad”

Las sanciones académicas a las que se enfrentará el estudiante son las siguientes:

- a) En los casos de fraude académico demostrado en el trabajo académico o evaluación respectiva, la calificación que se impondrá será de cero puntos sin perjuicio de la sanción disciplinaria correspondiente.
- b) En los casos relacionados con plagio demostrado en el trabajo académico cualquiera sea su naturaleza, la calificación que se impondrá será de cero puntos, sin perjuicio de la sanción disciplinaria correspondiente.

3. Formato de Rúbrica de evaluación

Tipo de actividad: Independiente	
Momento de la evaluación: Final	
La máxima puntuación posible es de 125 puntos	
Primer criterio de evaluación: Implementa el modelo de aprendizaje supervisado Este criterio representa 60 puntos del total de 125 puntos de la actividad.	Nivel alto: El estudiante implementa de forma adecuada el modelo de Aprendizaje supervisado de su elección. Si su trabajo se encuentra en este nivel puede obtener entre 31 puntos y 60 puntos Nivel Medio: El estudiante implementa parcialmente el modelo de Aprendizaje supervisado de su elección. Si su trabajo se encuentra en este nivel puede obtener entre 2 puntos y 30 puntos Nivel bajo: El estudiante no implementa el modelo de Aprendizaje supervisado. Si su trabajo se encuentra en este nivel puede obtener entre 0 puntos y 1 punto
Segundo criterio de evaluación: Interpreta los resultados Este criterio representa 40 puntos del total de 125 puntos de la actividad	Nivel alto: El estudiante interpreta de forma adecuada los resultados obtenidos por los modelos. Si su trabajo se encuentra en este nivel puede obtener entre 21 puntos y 40 puntos Nivel Medio: El estudiante interpreta parcialmente los resultados obtenidos por los modelos. Si su trabajo se encuentra en este nivel puede obtener entre 2 puntos y 20 puntos Nivel bajo: El estudiante no interpreta los resultados obtenidos por los modelos. Si su trabajo se encuentra en este nivel puede obtener entre 0 puntos y 1 punto
Tercer criterio de evaluación: Presenta el documento con la	Nivel alto: El documento cuenta con las especificaciones solicitadas en la guía de actividades para su entrega. Si su trabajo se encuentra en este nivel puede obtener entre 13 puntos y 25 puntos Nivel Medio: El documento cuenta parcialmente con las especificaciones solicitadas en la guía de actividades para su entrega.

estructura solicitada y uso correcto de las normas APA.

Este criterio representa 25 puntos del total de 125 puntos de la actividad

Si su trabajo se encuentra en este nivel puede obtener entre 2 puntos y 12 puntos

Nivel bajo: El documento no cuenta con las especificaciones solicitadas en la guía de actividades para su entrega.

Si su trabajo se encuentra en este nivel puede obtener entre 0 puntos y 1 punto