



UNIMINUTO
Corporación Universitaria Minuto de Dios
Educación de calidad al alcance de todos
Vigilada MinEducación

Adquisición y adecuación de base de datos

Machine Learning

Taller. Adquisición y adecuación de base de datos

Descripción: El estudiante elaborará un taller sobre la adquisición y adecuación de base de datos, con el objetivo de preparar un conjunto de datos (dataset) para aplicarlo al modelo de machine learning a partir de las necesidades del problema. Esta actividad le permitirá desarrollar los resultados de aprendizaje de realizar el análisis descriptivo de los datos de acuerdo con el tipo de variable categórica o continua.

Ver: Tenga en cuenta las siguientes preguntas orientadoras para desarrollar la actividad:



¿Cuál es la metodología para adquirir datos que se aplicarán a los modelos de machine learning?

¿Qué procesamiento se debe realizar a los datos para su adecuación?

Juzgar: Consulte los siguientes recursos y con base en una lectura minuciosa, realice lo que se indica a continuación.

- Rothman, D. (2018). **Artificial Intelligence by Example: Develop Machine Intelligence from Scratch Using Real Artificial Intelligence use Cases**. pp. 30-45.
- UCI. Machine Learning Repository. (2021). **Base de datos 1**.
- UCI. Machine Learning Repository. (2021). **Base de datos 2**.
- Anexo 1.

Tome apuntes o subraye las ideas principales de los documentos consultados, con el fin de contar con los elementos conceptuales necesarios para resolver el taller de forma adecuada. Además, realice una investigación adicional en fuentes de información verídicas y confiables que le permita complementar la información.

Actuar: A partir de la información revisada y la consulta de los materiales de la semana, resuelva el taller.

1. Realice una lectura previa del documento y proceda a desarrollar el anexo que contiene las indicaciones para el desarrollo del taller, aplicando los conceptos abordados en la bibliografía y teniendo en cuenta el procedimiento que se menciona.
2. Redacte afirmaciones, descripciones o argumentos que soporten cada uno de los elementos e intégrelos a las categorías correspondientes:
 - Seleccione la base de datos a trabajar con el enlace indicado.
 - Descargue la data en el notebook (Google Colab); para acceder a Google colab ingrese al siguiente enlace: <https://colab.research.google.com/>
 - Posteriormente siga las instrucciones que describe el Anexo 1.
 - Preparare el conjunto de datos confiable a partir de la limpieza de datos, la identificación y eliminación de errores y datos duplicados.
 - Realice el análisis descriptivo de los datos de acuerdo con el tipo de variable categórica o continua.
3. Cuide la presentación, redacción y ortografía. Use correctamente las normas de redacción con el fin de que escriba un texto claro, preciso y coherente en el que se evidencie una estructura concreta y las ideas se relacionen de forma fluida y articulada.
4. Recuerde que sus referencias y citas deben apegarse a las normas APA.
5. Incluya un apartado en el que agregue el título del trabajo, su nombre completo, el nombre del curso y el número de la semana, el nombre del docente y la fecha de entrega.

Devolución creativa-metacognición: Al finalizar el taller, el estudiante tendrá las habilidades para preparar conjuntos de datos confiables, utilizando herramientas como los notebooks de Google Colab. El adecuado procesamiento de los datos garantiza un alto nivel de eficiencia de los modelos de machine learning.

Tiempos y recursos educativos para el desarrollo de la actividad

Tipo de actividad	Individual
Tiempos	Desarrollo de la actividad: 6 Consulta de materiales: 2
Recursos	<p>Recursos básicos:</p> <ul style="list-style-type: none"> • Rothman, D. (2018). <u>Artificial Intelligence by Example: Develop Machine Intelligence from Scratch Using Real Artificial Intelligence use Cases</u>. Pp. 30-45. • UCI. Machine Learning Repository. (2021). <u>Base de datos 1</u>. • UCI. Machine Learning Repository. (2021). <u>Base de datos 2</u>. • <u>Anexo 1</u>. <p>Recurso complementario:</p> <ul style="list-style-type: none"> • Bishop, C. (2006). <u>Pattern Recognition and Machine Learning</u> (67-76). Information Science and Statistics.

Rol del tutor: A través de los canales de comunicación del curso, acompañará y guiará la elaboración de la actividad, orientando el proceso de análisis. Asimismo, aclarará inquietudes, ampliará conceptos, brindará retroalimentación y, finalmente, evaluará la actividad de acuerdo con los criterios establecidos en la rúbrica.

Forma de entrega:

- Guarde el documento en el formato correspondiente .pptx., si su elección fue una herramienta digital, obtengan el enlace proporcionado por la herramienta que utilizaron y ubíquelo en una hoja Word con sus datos.
- Nombre el archivo de la siguiente manera:
primerapellido_primernombre_tipodeactividad. Por ejemplo:
romero_luis_presentacióndediapositivas.
- Dé clic en **Agregar entrega** para subir el trabajo a la plataforma.
- Arrastre y suelte el documento en **Archivos enviados**.
- Dé clic en el botón **Guardar cambios** para finalizar la entrega. De lo contrario, el archivo no se guardará y no se habrá entregado.

Rúbrica. Taller

Criterios		No evaluable (0)	Insuficiente (1.9)	Regular (2.9)	Bueno (3.9)	Excelente (5)	Puntos
Forma	Ortografía, gramática, cohesión y coherencia	<ul style="list-style-type: none"> Se detecta plagio total, se envió otra actividad o la actividad no cumple con los criterios mínimos solicitados para ser evaluada, la actividad está incompleta, no entregó. 	<ul style="list-style-type: none"> Se observan ocho o más errores gramaticales y de ortografía. El taller es confuso, ya que sus ideas no se relacionan adecuadamente. <p>NO SE CONSIDERAN LOS ERRORES REPETIDOS.</p>	<ul style="list-style-type: none"> Se observan de cinco a siete errores gramaticales y de ortografía. El taller es ambiguo y sus ideas no se expresan de forma fluida ni articulada. <p>NO SE CONSIDERAN LOS ERRORES REPETIDOS.</p>	<ul style="list-style-type: none"> Se observan de uno a cuatro errores gramaticales y de ortografía. El taller es claro y coherente, sus ideas se expresan de forma fluida y articulada, aunque con mínimas imprecisiones. <p>NO SE CONSIDERAN LOS ERRORES REPETIDOS.</p>	<ul style="list-style-type: none"> No se observan errores gramaticales ni faltas de ortografía. El taller es claro y coherente y sus ideas se expresan de forma fluida y articulada. 	
	Puntaje	0	0.63	0.70	0.90	1	
Fondo	Aplicación de conceptos	<ul style="list-style-type: none"> Se detecta plagio total, se envió otra actividad o la actividad no cumple con los criterios mínimos solicitados para ser evaluada, la actividad está incompleta, no entregó. 	<ul style="list-style-type: none"> No se aplican los conceptos necesarios para desarrollar el taller, por lo tanto, no se evidencian comprensión ni apropiación de la temática. 	<ul style="list-style-type: none"> Se aplican, con bastantes imprecisiones y de forma incompleta, los conceptos necesarios para desarrollar el taller, evidenciando escasa comprensión y apropiación de la temática. 	<ul style="list-style-type: none"> Se aplican, con algunas imprecisiones, los conceptos necesarios para desarrollar el taller, evidenciando comprensión y apropiación básicas de la temática. 	<ul style="list-style-type: none"> Se aplican de forma precisa todos los conceptos necesarios para desarrollar el taller, evidenciando comprensión y apropiación de la temática. 	
	Puntaje	0	0.63	1.10	1.5	2	
	Referencias bibliográficas	<ul style="list-style-type: none"> Se detecta plagio total, se envió otra actividad o la actividad no cumple con los criterios mínimos solicitados para ser evaluada, la actividad está incompleta, no entregó. 	<ul style="list-style-type: none"> No se colocan citas ni referencias en todo el documento. 	<ul style="list-style-type: none"> No se utiliza el formato de citas y referencias APA. 	<ul style="list-style-type: none"> El formato para el tipo de citas y referencias APA tiene algunos errores. 	<ul style="list-style-type: none"> El formato para el tipo de citas APA y número de referencias corresponden con los criterios solicitados. 	
	Puntaje	0	0.64	1.10	1.5	2	
						Calificación total	