

PROGRAMA DE ESTUDIO DE ASIGNATURA - PEA
A. INFORMACIÓN GENERAL

Carrera:	TECNOLOGÍA SUPERIOR EN BIG DATA R		
Asignatura / Módulo:	MINERÍA DE DATOS I		
Código:	TSBD R-213		
Prerrequisitos:	INTRODUCCIÓN A LA BIG DATA,FUNDAMENTOS DE ESTADÍSTICA		
Correquisitos:	BASE DE DATOS I		
Campo de Formación:	ADAPTACIÓN E INNOVACIÓN TECNOLÓGICA		
Unidad de Organización Curricular:	UNIDAD PROFESIONAL		
Modalidad:	PRESENCIAL		
Itinerario:	1		
Número de Horas / Créditos:	192	Créditos:	4.00
Nivel:	SEGUNDO		
Área	Programación		
Organización de aprendizajes por modalidad, número de horas destinadas a cada componente.	Componente Docencia:	64	
	Componente de Prácticas de Aprendizaje:	64	
	Componente de Aprendizaje Autónomo:	64	
Horas de Tutoría (semanal):	1		

B. RESULTADOS DE APRENDIZAJE

b.1. Resultado de aprendizaje generales de la carrera (Perfil de egreso)	<ol style="list-style-type: none"> 1. Utiliza herramientas tecnológicas de avanzada, para adquirir y extraer datos necesarios para apoyar en la solución de problemas tanto a nivel empresarial como social. 2. Procesa datos estructurados y no estructurados para agruparlos e integrarlos con fin de obtener valor para el negocio. 3. Utiliza metodologías innovadoras basadas en TIC para capturar datos y extraer información que pueda entregar valor a las empresas. 4. Utiliza minería de datos para descubrir patrones en los datos que permitan solucionar problemas o mejorar servicios en las empresas. 5. Aplica su capacidad crítica e innovadora para proponer modelos de análisis de información basados en datos históricos, trabajando en conjunto con el Ingeniero de Datos, para llegar a inferencias estadísticas.
--	---

b.2. Resultados de aprendizaje específicos de la asignatura

1. Utiliza metodologías innovadoras basadas en TIC para capturar datos y extraer información que pueda entregar valor a las empresas.
 2. Adquiere y recopila datos de diversas fuentes, comprendiendo la importancia de la calidad y la relevancia de los datos para el análisis.
 3. Interpreta los resultados del análisis exploratorio, formulando hipótesis y planteando preguntas adicionales para investigar.
 4. Interpreta los resultados y las conclusiones de los modelos para respaldar la toma de decisiones basadas en datos en diferentes escenarios y aplicaciones.
- Eje transversal Interculturalidad: Promueve la sensibilización cultural y desarrolla habilidades interculturales entre los estudiantes. Se enfatiza la importancia de fomentar la colaboración y la experiencia internacional, lo que permitirá a los estudiantes trabajar de manera efectiva en equipos multidisciplinarios y diversas comunidades.
- Gestión ambiental: Aplican prácticas de desarrollo de software sostenibles en el contexto de la Minería de Datos.
- Equidad de género: Aprenden a trabajar para promover la igualdad de oportunidades y la inclusión de género en el desarrollo de software, garantizando una representación equitativa y justa de las perspectivas y contribuciones de todas las personas.
- Prevención del consumo de drogas Explorar Datos y extraer de Patrones utilizando técnicas de Minería de Datos, como la clasificación, agrupación (clustering) y asociación de patrones, para identificar tendencias y relaciones relevantes en los datos sobre el consumo de drogas.

C. OBJETIVOS DE LA ASIGNATURA

Objetivo General

Comprender los principios fundamentales, técnicas avanzadas y aplicaciones prácticas de la minería de datos, con el fin de adquirir las habilidades necesarias para analizar grandes conjuntos de datos, descubrir patrones significativos, tomar decisiones fundamentadas y generar valor en diversos contextos empresariales y científicos.

Objetivos Específicos

- 1. Comprender el concepto y la importancia de la minería de datos.
- 2. Aprender técnicas de recopilación y limpieza de datos.
- 3. Dominar herramientas y técnicas de análisis y visualización de datos.
- 4. Comprender los tipos de modelos y técnicas de evaluación.

Forma de Evidenciarlo (Evaluación)

UNIDAD 1:Portafolio de ejercicios,Prueba Objetiva.
UNIDAD 2:Portafolio de ejercicios,Prueba Objetiva.
UNIDAD 3:Portafolio de ejercicios,Prueba Objetiva.
UNIDAD 4:Portafolio de ejercicios,Prueba Objetiva.

CONTENIDOS DE ENSEÑANZA – APRENDIZAJE

CONTENIDOS	CD	AA	APE
UNIDAD DIDÁCTICA: 1			
Introducción a Minería de Datos			
1.1. Definición y conceptos de Minería de Datos.			
1.2. Aplicaciones y casos de uso de Minería de Datos en diferentes industrias.			
1.3. Proceso de Minería de Datos:			
1.3.1. Metodología CRISP-DM (Cross Industry Standard Process for Data Mining).	16.00	16.00	16.00
1.3.2. SEMMA (Sample, Explore, Modify, Model, Assess).			
1.3.3. KDD (Knowledge Discovery in Databases).			
1.4. Tipos de datos y características			
1.4. Tipos de datos y características			
1.4.1. Datos estructurados			
1.4.2. Datos no estructurados.			
1.5. Fases del proyecto de Minería de Datos.			
UNIDAD DIDÁCTICA: 2			
Preparación de Datos			
2.1. Recopilación y adquisición de datos.			
2.2. Limpieza de datos			
2.2.1. Detección y manejo de valores atípicos	16.00	16.00	16.00
2.2.2. Datos faltantes y ruido.			
2.3. Transformación de datos:			
2.3.1. Normalización			
2.3.2. Discretización			
2.3.3. Codificación de variables categóricas.			
2.4. Selección de características relevantes.			

CONTENIDOS		CD	AA	APE
UNIDAD DIDÁCTICA: 3				
Análisis Exploratorio de Datos				
3.1. Análisis y visualización de datos 3.1.1. Pandas 3.1.2. Matplotlib 3.1.3. Seaborn. 3.2. Técnicas de visualización para comprender la distribución de datos y detectar patrones. 3.3. Análisis de correlación y relaciones entre variables. 3.4. Identificación de patrones y tendencias en los datos.		16.00	16.00	16.00
UNIDAD DIDÁCTICA: 4				
Modelado y Evaluación				
Modelado y Evaluación 4.1. Modelos de Minería de Datos: 4.1.1. Clasificación 4.1.2. Regresión 4.1.3. Clustering 4.1.4. Asociación. 4.2. Evaluación de modelos 4.2.1. Métricas de rendimiento 4.2.2. Técnicas de validación cruzada. 4.3. Ajuste de hiperparámetros y optimización de modelos. 4.4. Interpretación de resultados y toma de decisiones basada en los modelos.		16.00	16.00	16.00
Actividades Prácticas Programadas: (Por cada unidad que se tenga el componente práctico, realizar una breve descripción y justificación de la práctica).	Unidad 1.- Guía Práctica: Análisis de casos las Aplicaciones de la Minería de Datos en Diversas Industrias Unidad 2.- Guía Práctica: Limpieza, Transformación y Selección de Datos para Análisis Avanzado Unidad 3.- Guía Práctica: Exploración Visual y Análisis de Datos para Descubrir Patrones y Tendencias Unidad 4.- Guía Práctica: Construcción, Evaluación y Optimización de Modelos Predictivos en Minería de Datos			
ESTRATÉGIAS METODOLÓGICAS Y RECURSOS DIDÁCTICOS				
Estratégicas metodológicas:				
Evaluación auténtica Aprendizaje basado en problemas (ABP) Ejercicios, demostraciones y simulaciones situadas Aprendizaje basado en proyectos (ABP) Gamificación				
Recursos didácticos:				
Pizarra y tiza o marcadores Proyector y pizarra digital Mapas conceptuales y esquemas Internet y recursos en línea Laboratorios y talleres Software educativo				
BIBLIOGRAFÍA BASE Ubicación: Biblioteca ISTA				
María Pérz Marqués, (2015). Big Data: Técnicas, Herramientas Y Aplicaciones.. México: Alfaomega Grupo Editor.				
María Pèrez Marques, (2020). Big Data: Técnicas, Herramientas Y Aplicaciones.. Colombia: Alfaomega Grupo Editor.				

Sustento de Selección
Aborda las unidades 1,2,3 y 4
BIBLIOGRAFÍA DE CONSULTA: Ubicación: Biblioteca de ISTA

Sustento de Selección
Aborda las unidades 1,2,3 y 4

Periodo Académico: TSBD R OCTUBRE/2024 FEBRERO/2025

1.DATOS DE LA ASIGNATURA

CODIGO: TSBD R-213

MINERÍA DE DATOS I

Carrera:	TECNOLOGÍA SUPERIOR EN BIG DATA R
Ciclo/semestre:	SEGUNDO
Unidad de organización curricular:	UNIDAD PROFESIONAL
Campo de formación:	ADAPTACIÓN E INNOVACIÓN TECNOLÓGICA
Modalidad:	PRESENCIAL

Pre-requisitos		Co-requisitos	
Asignatura	Código	Asignatura	Código
FUNDAMENTOS DE ESTADÍSTICA	TSBD R-111	PROBABILIDAD Y ESTADÍSTICA	TSBD R-212
INTRODUCCIÓN A LA BIG DATA	TSBD R-115		

2. ACTIVIDADES DE APRENDIZAJE

Componentes	Horas semanales	Horas de ciclo
Horas docentes (Asistido por el profesor y colaborativo)	4.0	64
Horas Prácticas	4.0	64
Horas Autónomas	4.0	64
Total Horas	12.0	192

3. DOCENTE
3.1. Profesor Responsable

Mgtr. CHIMBO CORONEL VERONICA PAULINA

M2A

3.2 Horario de clases

Horas	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
9:00 a 10:00		M2A	M2A		M2A	
10:00 a 11:00		M2A			M2A	
11:00 a 12:00		M2A				

4. DESCRIPCIÓN DE LA ASIGNATURA

La materia de Minería de Datos es un curso interdisciplinario que combina conceptos de estadística, ciencias de la computación y negocios para explorar técnicas y metodologías utilizadas en el análisis y extracción de conocimiento a partir de grandes conjuntos de datos. Durante el curso, los estudiantes aprenderán sobre los fundamentos teóricos de la minería de datos, incluyendo los diferentes tipos de datos, métodos de preprocesamiento, algoritmos de modelado y técnicas de evaluación de modelos. Además, se enfocarán en aplicaciones prácticas de la minería de datos en diversas industrias, como finanzas, salud, marketing y tecnología. A través de actividades prácticas, proyectos y estudios de casos, los estudiantes desarrollarán habilidades para recopilar, limpiar, explorar, modelar y evaluar datos con el objetivo de tomar decisiones fundamentadas y generar valor en entornos empresariales y científicos. La materia puede incluir el uso de herramientas y lenguajes de programación como Python, R y SQL, así como la familiarización con software y plataformas de análisis de datos. Al final del curso, se espera que los estudiantes sean capaces de aplicar eficazmente técnicas de minería de datos para resolver problemas complejos y tomar decisiones informadas basadas en datos.

5. OBJETIVO DE LA ASIGNATURA

Comprender los principios fundamentales, técnicas avanzadas y aplicaciones prácticas de la minería de datos, con el fin de adquirir las habilidades necesarias para analizar grandes conjuntos de datos, descubrir patrones significativos, tomar decisiones fundamentadas y generar valor en diversos contextos empresariales y científicos.

5.1. OBJETIVOS ESPECÍFICOS

1. Comprender el concepto y la importancia de la minería de datos.
2. Aprender técnicas de recopilación y limpieza de datos.
3. Dominar herramientas y técnicas de análisis y visualización de datos.
4. Comprender los tipos de modelos y técnicas de evaluación.

6.RELACIÓN OBJETIVOS, RESULTADOS DE APRENDIZAJE, CONTENIDOS Y ESTRATEGIAS DE ENSEÑANZA –

OBJETIVO ESPECÍFICO	RESULTADOS DE APRENDIZAJE	FECHA INICIO	FECHA FIN	CONTENIDOS (UNIDADES Y TEMAS)	N° HORAS*			ESTRATEGIAS DE ENSEÑANZA
					CD	APE	AA	
1. Comprender el concepto y la importancia de la minería de datos. 2. Identificar aplicaciones prácticas en diversas industrias. 3. Aplicar metodologías estándar y tipos de datos. 4. Entender las fases generales de un proyecto de minería de datos.	1. Utiliza metodologías innovadoras basadas en TIC para capturar datos y extraer información que pueda entregar valor a las empresas.	16/10/2024	13/11/2024	UNIDAD DIDÁCTICA: 1 Introducción a Minería de Datos 1.1. Definición y conceptos de Minería de Datos. 1.2. Aplicaciones y casos de uso de Minería de Datos en diferentes industrias. 1.3. Proceso de Minería de Datos: 1.3.1. Metodología CRISP-DM (Cross Industry Standard Process for Data Mining). 1.3.2. SEMMA (Sample, Explore, Modify, Model, Assess). 1.3.3. KDD (Knowledge Discovery in Databases). 1.4. Tipos de datos y características 1.4.1. Datos estructurados 1.4.2. Datos no estructurados. 1.5. Fases del proyecto de Minería de Datos.	16.00	16.00	16.00	Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo Aprendizaje inverso (flipped classroom)
1. Aprender técnicas de recopilación y limpieza de datos. 2. Dominar la transformación de datos y selección de características. 3. Entender la importancia de la preparación de datos para la minería. 4. Aplicar técnicas para mejorar la calidad y utilidad de los datos.	2. Adquiere y recopila datos de diversas fuentes, comprendiendo la importancia de la calidad y la relevancia de los datos para el análisis	18/11/2024	11/12/2024	UNIDAD DIDÁCTICA: 2 Preparación de Datos 2.1. Recopilación y adquisición de datos. 2.2. Limpieza de datos 2.2.1. Detección y manejo de valores atípicos 2.2.2. Datos faltantes y ruido. 2.3. Transformación de datos: 2.3.1. Normalización 2.3.2. Discretización 2.3.3. Codificación de	16.00	16.00	16.00	Aprendizaje activo Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo

OBJETIVO ESPECÍFICO	RESULTADOS DE APRENDIZAJE	FECHA INICIO	FECHA FIN	CONTENIDOS (UNIDADES Y TEMAS)	N° HORAS*			ESTRATEGIAS DE ENSEÑANZA
					CD	APE	AA	
1. Aprender técnicas de recopilación y limpieza de datos. 2. Dominar la transformación de datos y selección de características. 3. Entender la importancia de la preparación de datos para la minería. 4. Aplicar técnicas para mejorar la calidad y utilidad de los datos.	2. Adquiere y recopila datos de diversas fuentes, comprendiendo la importancia de la calidad y la relevancia de los datos para el análisis.	18/11/2024	11/12/2024	variables categóricas. 2.4. Selección de características relevantes.	16.00	16.00	16.00	Aprendizaje activo Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo
1. Dominar herramientas y técnicas de análisis y visualización de datos. 2. Comprender la distribución y patrones de los datos. 3. Identificar correlaciones y relaciones entre variables. 4. Descubrir patrones y tendencias a través del análisis exploratorio.	3. Interpreta los resultados del análisis exploratorio, formulando hipótesis y planteando preguntas adicionales para investigar.	16/12/2024	15/01/2025	UNIDAD DIDÁCTICA: 3 Análisis Exploratorio de Datos 3.1. Análisis y visualización de datos 3.1.1. Pandas 3.1.2. Matplotlib 3.1.3. Seaborn. 3.2. Técnicas de visualización para comprender la distribución de datos y detectar patrones. 3.3. Análisis de correlación y relaciones entre variables. 3.4. Identificación de patrones y tendencias en los datos.	16.00	16.00	16.00	Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo

OBJETIVO ESPECÍFICO	RESULTADOS DE APRENDIZAJE	FECHA INICIO	FECHA FIN	CONTENIDOS	N° HORAS*			ESTRATEGIAS DE ENSEÑANZA
				(UNIDADES Y TEMAS)	CD	APE	AA	
1. Entender los tipos de modelos y técnicas de evaluación. 2. Aprender a evaluar y ajustar modelos de minería de datos. 3. Dominar estrategias de optimización y validación de modelos. 4. Interpretar resultados y tomar decisiones basadas en ellos.	4. Interpreta los resultados y las conclusiones de los modelos para respaldar la toma de decisiones basadas en datos en diferentes escenarios y aplicaciones. -Eje transversal Interculturalidad: Promueve la sensibilización cultural y desarrolla habilidades interculturales entre los estudiantes. Se enfatiza la importancia de fomentar la colaboración y la experiencia internacional, lo que permitirá a los estudiantes trabajar de manera efectiva en equipos multidisciplinarios y diversas comunidades. - Gestión ambiental: Aplican prácticas de desarrollo de software sostenibles en el contexto de la Minería de Datos. - Equidad de género: Aprenden a trabajar para promover la igualdad de oportunidades y la inclusión de género en el desarrollo de software, garantizando una representación equitativa y justa de	20/01/2025	12/02/2025	UNIDAD DIDÁCTICA: 4 Modelado y Evaluación Modelado y Evaluación 4.1. Modelos de Minería de Datos: 4.1.1. Clasificación 4.1.2. Regresión 4.1.3. Clustering 4.1.4. Asociación. 4.2. Evaluación de modelos 4.2.1. Métricas de rendimiento 4.2.2. Técnicas de validación cruzada. 4.3. Ajuste de hiperparámetros y optimización de modelos. 4.4. Interpretación de resultados y toma de decisiones basada en los modelos.	16.00	16.00	16.00	Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo Aprendizaje inverso (flipped classroom)

OBJETIVO ESPECÍFICO	RESULTADOS DE APRENDIZAJE	FECHA INICIO	FECHA FIN	CONTENIDOS	N° HORAS*			ESTRATEGIAS DE ENSEÑANZA
				(UNIDADES Y TEMAS)	CD	APE	AA	
1. Entender los tipos de modelos y técnicas de evaluación. 2. Aprender a evaluar y ajustar modelos de minería de datos. 3. Dominar estrategias de optimización y validación de modelos. 4. Interpretar resultados y tomar decisiones basadas en ellos.	las perspectivas y contribuciones de todas las personas. - Prevención del consumo de drogas Explorar Datos y extraer de Patrones utilizando técnicas de Minería de Datos, como la clasificación, agrupación (clustering) y asociación de patrones, para identificar tendencias y relaciones relevantes en los datos sobre el consumo de drogas.	20/01/2025	12/02/2025	UNIDAD DIDÁCTICA: 4 Modelado y Evaluación Modelado y Evaluación 4.1. Modelos de Minería de Datos: 4.1.1. Clasificación 4.1.2. Regresión 4.1.3. Clustering 4.1.4. Asociación. 4.2. Evaluación de modelos 4.2.1. Métricas de rendimiento 4.2.2. Técnicas de validación cruzada. 4.3. Ajuste de hiperparámetros y optimización de modelos. 4.4. Interpretación de resultados y toma de decisiones basada en los modelos.	16.00	16.00	16.00	Aprendizaje basado en problemas (ABP) Aprendizaje basado en proyectos (ABP): Aprendizaje colaborativo Aprendizaje inverso (flipped classroom)
TOTAL HORAS CICLO					64	64	64	

Componente Docencia: CD Aprendizaje Práctico Experimental: APE Aprendizaje Autónomo: AA

7.CRITERIOS PARA LA ACREDITACIÓN DE LA ASIGNATURA

1. Todas las carreras de la institución utilizarán un mismo proceso evaluación excepto para la carrera de Tecnología Superior en Administración de Infraestructura y Plataformas Tecnológicas.

2. La estructura de Evaluación constará con :

- Parcial 1 – Evaluado de 0 a 10 puntos
- Parcial 2 – Evaluado de 0 a 10 puntos
- Examen de Recuperación - Evaluado de 0 a 10 puntos

FORMA DE EVALUACIÓN	PUNTAJE	OBSERVACIÓN
PARCIAL 1	0-10	Mínimo 3 actividades calificadas y promediadas sobre 10
PARCIAL 2	0-10	Mínimo 3 actividades calificadas y promediadas sobre 10
PROMEDIO	0-10	Si el promedio es mayor o igual a 3 puntos y menor a 7 puntos el estudiante tiene derecho a rendir un Examen de Recuperación.
EXAMEN DE RECUPERACIÓN	0-10	El examen de recuperación debe ser igual o mayor a 7 puntos.

3. Componentes de los elementos de la estructura de evaluación:

a. Cada Parcial dentro de la estructura de evaluación debe constar mínimo con 3 actividades calificadas y promediadas sobre 10 puntos para obtener el promedio del Parcial. Las actividades de cada Parcial deben estar alineadas a la forma de evidenciar los resultados de aprendizaje plasmadas en el PEA de cada asignatura.

Las actividades a ser evaluadas quedan a libertad de cátedra de cada docente, generando las actividades según la naturaleza del contenido del PEA y la experticia del docente en la ejecución del PEA para lograr los resultados de aprendizaje planteados. Para evaluar las actividades los docentes deben plasmar como se evaluará (Técnica de Evaluación) para conocimiento de los estudiantes y como respaldo del docente en caso de solicitud de recalificación. b. Las calificaciones Parcial 1 y Parcial 2 se promediarán y dará como resultado el componente de Promedio.

c. Si el promedio es mayor o igual a 7 puntos la asignatura se APRUEBA

d. Si el promedio es menor a 3 puntos la asignatura se REPRUEBA

e. Si el promedio es mayor o igual a 3 puntos y menor a 7 puntos el estudiante tiene derecho a rendir un Examen de Recuperación.

f. El examen de recuperación debe ser igual o mayor a 7 puntos.

i. El examen de recuperación debe cumplir con el formato establecido en el instructivo de documentos académicos y debe ser entregado a las direcciones de carrera (examen con respuestas y sin respuestas)

g. El promedio final es igual al Promedio + el 50% de la nota obtenida en el Examen de Recuperación, si esta suma es mayor a 8 puntos se asigna al PROMEDIO FINAL solo 8 puntos como máximo en este caso se APRUEBA la asignatura, si el promedio final es menor a 7 puntos la asignatura se REPRUEBA.

ESCALA CUANTITATIVA	9,50 a 10	8,50 a 9,49	7,50 a 8,49	7,00 a 7,49	Menor a 7,00
EQUIVALENCIA	Excelente	Muy bueno	Bueno	Regular	Deficiente

FORMA DE EVIDENCIARLO

UNIDAD 1 Portafolio de ejercicios, Prueba Objetiva.
UNIDAD 2 Portafolio de ejercicios, Prueba Objetiva.
UNIDAD 3: Portafolio de ejercicios, Prueba Objetiva.
UNIDAD 4 Portafolio de ejercicios, Prueba Objetiva.

8. REFERENCIAS REQUERIDAS PARA EL APRENDIZAJE DE LA ASIGNATURA

8.1 Bibliografía Base

Código	Bibliografía
2057	Maria Pérez Marqués, (2015). Big Data: Técnicas, Herramientas Y Aplicaciones.. México: Alfaomega Grupo Editor.
2100	Maria Pérez Marques, (2020). Big Data: Técnicas, Herramientas Y Aplicaciones.. Colombia: Alfaomega Grupo Editor.

8.2 Bibliografía Complementaria

Bibliografía
Han, J., Pei, J., & Kamber, M. (2011). Data Mining: Concepts and Techniques. Morgan Kaufmann.

8.3 Linkografía

URL
UCI Machine Learning Repository. (s.f.). UCI Machine Learning Repository. Recuperado de https://archive.ics.uci.edu/ml/

9. APROBACIÓN

ELABORADO		APROBADO POR	
DOCENTES		DIRECTOR/A DE LA CARRERA	
 _____ Mgtr. CHIMBO CORONEL VERONICA PAULINA		 _____ Mgtr. BERNAL PRISCILA	
Fecha:	14-oct-2024	Fecha:	14-oct-2024