

SÍLABO DEL CURSO

BIG DATA MACHINE LEARNING

I. INFORMACIÓN GENERAL:

Facultad	Ingeniería	Carrera profesional	Ingeniería de Sistemas Computacionales	Ciclo	9°	Créditos			3
Código de curso	SIST1501	Requisitos	Taller de Base de Datos	Horas	HT	HP	HL	PC (**)	
					2	0	2	0	
Tipo de curso	Electivo	Modalidad del curso	Presencial	Periodo lectivo	2022-2				
El curso aporta a las competencias generales:		Responsabilidad Social y Ciudadanía: Actúa acorde a sus derechos y deberes ciudadanos con el fin de que sus acciones generen un impacto positivo en la sociedad y en la protección de los derechos humanos, propiciando mejoras en la realidad actual de su entorno a partir de la indagación del problema y/o el contexto.							
El curso aporta a las competencias específicas:		Competencia K: Capacidad de utilizar técnicas, habilidades y herramientas de la ingeniería moderna necesarias para la práctica de la ingeniería.							
El curso desarrolla el componente:		Investigación <input type="checkbox"/> Responsabilidad Social <input type="checkbox"/> Ciudadanía <input type="checkbox"/> Práctica Pre profesional <input checked="" type="checkbox"/>							
“En el presente semestre académico, por situación excepcional en el país, se podría reformular la secuencia y/o modalidad y/o estrategias didácticas de las actividades planificadas para el desarrollo de contenidos y/o evaluaciones en el curso, en concordancia con las disposiciones legales que puedan emitirse.”									

II. SUMILLA:

El curso es de naturaleza teórico-práctico, tiene como propósito contribuir a que el estudiante obtenga un claro entendimiento de los conceptos básicos de la importancia en la implementación de big data y su gestión, utilizando herramientas tecnológicas actuales para diseñar una implementación básica de machine learning.
Los temas principales son: fundamentos de BigData – MachineLearning: Manejo Spark, análisis de datos, uso de técnicas de ciencia de datos y la integración con Machine Learning.

III. LOGRO DEL CURSO:

Al finalizar el curso el estudiante presenta un proyecto de BigData integrado con Machine Learning basado en Spark, utilizando el lenguaje programación Python con las técnicas de ciencia de datos y las tecnologías Spark SQL, Streaming y GraphX.

IV. METODOLOGÍA DE ENSEÑANZA APRENDIZAJE

Para alcanzar el logro de aprendizaje del curso y de las unidades, el docente integra métodos activos, estrategias y técnicas de manera reflexiva y crítica, buscando motivar, estimular y guiar el aprendizaje del estudiante.
Se desarrolla el aprendizaje basado en resolución de casos en forma de ejercicios propuestos de manera individual y en equipo, para los cuales los estudiantes analizan problemas, gestionan requerimientos, diseñan y desarrollan soluciones haciendo uso herramientas y software para BigData y Machine Learning, que permitan dar solución a necesidades de diferente tipo. Elabora un trabajo de investigación donde representa los resultados obtenidos, haciendo uso adecuado del proceso de investigación.
El docente soporta su práctica pedagógica en un sistema multiplataforma y recursos multimedia que le permite el desarrollo de actividades sincrónicas y asincrónicas, así como la gestión de contenidos, videoconferencias y el uso de herramientas que permitan mejorar la experiencia formativa que promuevan el aprendizaje y el desarrollo de competencias generales y específicas en los estudiantes.

V. ORGANIZACIÓN DE UNIDADES DE APRENDIZAJE:

UNID	NOMBRE / LOGRO DE UNIDAD	SEM	SABERES ESENCIALES	ACTIVIDADES PC (**)
I	FUNDAMENTOS DE BIG DATA Al finalizar la unidad, el estudiante desarrolla un software aplicando los fundamentos y definiciones básicas de BIG DATA y sus diversos enfoques aplicado a un caso propuesto.	1	Presentación del sílabo del curso. Introducción al curso: Importancia de la Gestión de Datos. Definición de Machine Learning Enfoques de BigData.	NO APLICA
		2	Gestión de Datos con Spark	
		3	Spark en Cluster y Operaciones sobre RDDS (<i>Resilient Distributed DataSets</i>)	
		4	Evaluación T1.	

UNID	NOMBRE / LOGRO DE UNIDAD	SEM	SABERES ESENCIALES	ACTIVIDADES PC (**)
II	FUNDAMENTOS DEL MACHINE LEARNING Al finalizar la unidad, el estudiante implementa una estructura básica de machine learning aplicado a un caso propuesto.	5	Machine Learning	NO APLICA
		6	Deep Learning	
		7	Taller Machine Learning	
III	INTEGRACIÓN MACHINE LEARNING - BIGDATA Al finalizar la unidad el estudiante expone un modelo integrado de BigData con Machine Learning propuesto utilizando Spark y Ciencia de Datos.	8	Evaluación Parcial	NO APLICA
		9	Integración BigData – Machine Learning	
		10	Análisis de Resultados	
		11	RNA	
		12	Evaluación T2.	
IV	MODELOS DEL ANÁLISIS DE DATOS Al finalizar la unidad, el estudiante implementa aplicaciones de los modelos de análisis de Datos aplicado a diversas con enfoque de BigData.	13	Modelos de Regresión	NO APLICA
		14	Modelos de Clasificación	
		15	Analítica de Datos	
		16	EVALUACIÓN FINAL.	
		(-)	EVALUACIÓN SUSTITUTORIA.	

(**) PC= Práctica de Campo. / La práctica de campo solo aplica en los cursos que tienen horas PC declaradas en el Plan de Estudios/ Las actividades de práctica de campo se detallan en la ficha de práctica de campo que se encuentra como anexo al sílabo.

VI. SISTEMA DE EVALUACIÓN

El docente selecciona, según la naturaleza del curso, diferentes herramientas de evaluación soportados en la plataforma virtual de aprendizaje (LMS), para evidenciar el desempeño de los estudiantes. El docente puede emplear: foros virtuales, exposiciones, cuestionarios o formularios virtuales, y entregas de tareas como: portafolios de evidencias, ensayos, diseño de proyectos, entre otros que considere pertinentes.

EVALUACIÓN	PESOS	SEM	Descripción de la Evaluación (Evidencia que debe presentar el estudiante)
T1(a)	15%	4	Evaluación T1: Examen práctico/ Resolución de Casos/ Autoinstructivo
Evaluación Parcial(a)	30%	8	Examen Pparcial: Examen práctico/ Resolución de Casos/ Autoinstructivo
T2(a)	15%	12	Evaluación T2: Examen práctico/ Resolución de Casos/ Autoinstructivo
Evaluación Final(a)	40%	16	Evaluación Final: Examen práctico/ Proyecto final / Exposición de proyecto/ Paper / Autoinstructivo
Evaluación Sustitutoria(a)	-----	(-)	Evaluación Examen práctico

(a) Los calificativos deben ser publicados en el sistema de acuerdo al Calendario Académico establecido para el presente Semestre.

(-) Ver en el calendario académico los días en que se realizarán las evaluaciones sustitutorias

VII. VI. BIBLIOGRAFÍA BÁSICA:

Nº	AUTOR	TÍTULO	AÑO	ENLACE URL
1	Tomasz Drabas and Denny Lee	PySpark Cookbook : Over 60 Recipes for Implementing Big Data Processing and Analytics Using Apache Spark and Python	2018	https://ebookcentral.bibliotecaupn.elogim.com/lib/upnpe/detail.action?docID=5446040

VIII. INFORMACIÓN COMPLEMENTARIA:

REFERENCIA	ENLACE
The BigData Agenda: Data Ethics and Critical Data Studies Annika Richterich University of Westminster Press 2018	https://repositorio.upn.edu.pe/handle/11537/28756
Sistema de inteligencia de negocios en la administración de cultivos para la producción agrícola	https://laccei.org/LACCEI2021-VirtualEdition/meta/FP262.html