



I. INFORMACIÓN GENERAL

CURSO	:	Aplicaciones de Data Science
CÓDIGO	:	CC219
CICLO	:	202301
CUERPO ACADÉMICO	:	Montoya Cubas, Carlos Fernando
CRÉDITOS	:	4
SEMANAS	:	16
HORAS	:	4 H (Teoría) Semanal
ÁREA O CARRERA	:	Ciencias de la Computacion

II. MISIÓN Y VISIÓN DE LA UPC

Misión: Formar líderes íntegros e innovadores con visión global para que transformen el Perú.

Visión: Ser líder en la educación superior por su excelencia académica y su capacidad de innovación.

III. INTRODUCCIÓN

Los datos son valiosos, y saber gestionarlos genera ventajas profesionales, aprendes a predecir patrones de comportamiento ayudando a la toma de decisiones estratégicas, identificando nuevas oportunidades de negocio o prediciendo acciones futuras en diferentes campos organizacionales.

El potencial de la ciencia de datos es muy grande, diversas disciplinas siguen en constante innovación de sus algoritmos al servicio de la minería de datos y diversas tecnologías se siguen proponiendo para el manejo de grandes cantidades de datos semi-estructurados y no estructurados. Por otro lado, el internet de las cosas permite la generación de grandes cantidades de datos, imágenes o textos. La inteligencia de negocios continúa buscando nuevas estrategias para el modelado y mejora del negocio. Los requerimientos de tipo y velocidad de análisis, la cantidad y tipo de datos, así como el nivel de experiencia de negocio hacen de la Ciencia de datos un reto constante.

IV. LOGRO (S) DEL CURSO

Al finalizar el curso, el estudiante presenta un proyecto de aplicación de ciencia de datos evidenciando el uso de técnicas y herramientas básicas de ciencia de datos, así como su correcta explicabilidad, para generar acciones y resultados específicos que le permitan explotar, interpretar y mostrar insights de los datos de origen diverso.

V. UNIDADES DE APRENDIZAJE

UNIDAD N°: 1 Introducción a las aplicaciones en ciencia de datos

LOGRO

El estudiante aprenderá la importancia de la ciencia de datos, así como sus principales aplicaciones tanto industriales como científicas.

TEMARIO

Contenido 1:

- Importancia de la ciencia de datos.
- Conceptos básicos de ciencia de datos, aprendizaje supervisado y no supervisado.

Actividades de aprendizaje:

- Presentación y análisis de diapositivas.
- Resolución de ejercicios.

HORA(S) / SEMANA(S)

1

UNIDAD N°: 2 Procesamiento del Lenguaje Natural**LOGRO**

El estudiante podrá utilizar las habilidades que necesita para aplicar la ciencia de datos a las diferentes tareas de la minería de textos.

TEMARIO

Contenido 2:

- Introducción a Procesamiento de Lenguaje Natural
- Task - Tareas NLP
- Desafíos del NLP
- Comparación NLP y CL (Computacional Linguistics)
- Conceptos básicos de Lingüística
- NLP y Machine Learning (Aprendizaje Automático)

Contenido 3:

- Recuperación de Información
- Tokens y Tipos
- Extracción de Vocabulario
- Modelo de Espacio Vectorial
- Frecuencia de Documento Invertida
- Similitudes entre vectores

Contenido 4:

- Modelamiento de Lenguaje
- Problema del Modelamiento del Lenguaje
- Repaso de Conceptos Probabilísticos
- Trigramas
- Clasificador Bayesiano

Contenido 5:

- Estado del arte en NLP
- Sequence to Sequence Models, Attention, and the Transformer

Actividades de aprendizaje:

- Presentación y análisis de diapositivas.
- Resolución de ejercicios.

Evidencias de aprendizaje:

PC1: Práctica Calificada 1.

HORA(S) / SEMANA(S)

2,3,4

UNIDAD N°: 3 Sistemas de recomendación

LOGRO

El estudiante aprenderá a desarrollar sistemas de recomendación, así como su correcta clasificación y validación.

TEMARIO

Contenido 6:

- Introducción
- Framework para problemas de recomendación
- Clasificación vs Regresión
- Predicción de puntuación con regresión
- Filtrado colaborativo (CF)
- Enfoques de vecindario
- CF - Factorización de matrices
- Factorización de matrices probabilísticas
- CF con Deep Learning
- Codificadores automáticos para CF
- Estrategia de modelado

Contenido 7:

- Evaluación de sistemas de recomendación
- Estudio por usuario
- Estudio por lotes
- Prueba de producción
- Precisión, Recall, F-score

Contenido 8:

- Modelos Avanzados
- Comentarios explícitos
- Mínimos cuadrados alternos
- Factorización de matrices asimétricas
- Codificadores automáticos de eliminación de ruido

Actividades de aprendizaje:

- Presentación y análisis de diapositivas.
- Resolución de ejercicios.

Evidencias de aprendizaje:

TP: Trabajo parcial.

Evaluación Parcial (EP)

HORA(S) / SEMANA(S)

5,6,7,8

UNIDAD N°: 4 Visión Computacional

LOGRO

El estudiante podrá aplicar conocimientos de ciencia de datos para resolver las principales tareas de visión computacional, así como explicar la decisión de los algoritmos utilizados.

TEMARIO

Contenido 9:

- Introducción a la visión computacional
- Clasificación de imágenes

Contenido 10:

- Detección de objetos
- Análisis facial

Contenido 11:

- Unet
- Detección de anomalías.

Contenido 12:

- Explicabilidad.

Actividades de aprendizaje:

- Presentación y análisis de diapositivas.
- Resolución de ejercicios.

Evidencias de aprendizaje:

PC2: Práctica Calificada 2.

HORA(S) / SEMANA(S)

9,10,11,12

UNIDAD N°: 5 eXplainable AI**LOGRO**

El estudiante podrá entender el porqué de la toma de decisiones de las principales técnicas de ciencia de datos.

TEMARIO

Contenido 13:

- Introducción a eXplainable IA
- Terminología Básica

Contenido 14:

- Modelo Interpretables
- Regresión Lineal
- Regresión logística

Contenido 15:

- Modelos Lineales Generalizados
- Interacción entre variables
- Modelos Aditivos Generalizados

Actividades de aprendizaje:

- Presentación y análisis de diapositivas.
- Resolución de ejercicios.

Evidencias de aprendizaje:

TF: Trabajo Final.

Examen Final (EB)

HORA(S) / SEMANA(S)

13,14,15,16

UNIDAD N°: 6 Outcome ABET 5 - Trabajo multidisciplinario

LOGRO

Actúa efectivamente como miembro o líder en un equipo, cuyos miembros juntos proporcionan liderazgo, crea un entorno de colaboración e inclusión, establece metas, planifica tareas y cumple objetivos.

TEMARIO

Contenido 15: ABET 5 - Trabajo multidisciplinario - Presentacion Trabajo Final (TF)

Contenido 16: Examen Final (EB)

HORA(S) / SEMANA(S)

15,16

VI. METODOLOGÍA

El estudiante cumple un rol activo en su aprendizaje, construyéndolo a partir de la reflexión crítica, análisis, discusión, evaluación, exposición e interacción con sus pares, y conectándolo con sus experiencias y conocimientos previos. Por ello, cada sesión está diseñada para ofrecer al estudiante diversas maneras de apropiarse y poner en práctica el nuevo conocimiento en contextos reales o simulados, reconociendo la importancia que esto tiene para su éxito profesional.

El curso se desarrolla haciendo uso de la Metodología Activa en cada una de sus dos sesiones semanales. La primera sesión es de 2 horas, en la cual se presentan los conocimientos teóricos y la otra sesión es de laboratorio de 2 horas en la cual se aplican los conocimientos adquiridos. Se hace uso extensivo de metáforas y representaciones gráficas y casos de estudio del tema desarrollado.

Cada semana se entregarán tareas académicas de acuerdo al tema realizado, El lenguaje de programación que se utilizará será Python.

Los entornos de desarrollo serán Anaconda, Google Colaboratory y kaggle.

VII. EVALUACIÓN

FÓRMULA

5% (TP1) + 10% (EA1) + 15% (TF1) + 15% (EB1) + 10% (TA1) + 10% (TA2) + 10% (TA3) + 10% (TA4) + 15% (DD1)

TIPO DE NOTA	PESO %
TA - TAREAS ACADÉMICAS	10
TA - TAREAS ACADÉMICAS	10
TP - TRABAJO PARCIAL	5
EA - EVALUACIÓN PARCIAL	10
TA - TAREAS ACADÉMICAS	10
TA - TAREAS ACADÉMICAS	10
DD - EVAL. DE DESEMPEÑO	15
TF - TRABAJO FINAL	15
EB - EVALUACIÓN FINAL	15

VIII. CRONOGRAMA

TIPO DE PRUEBA	DESCRIPCIÓN NOTA	NÚM. DE PRUEBA	FECHA	OBSERVACIÓN	RECUPERABLE
TA	TAREAS ACADÉMICAS	1	Semana 3		NO
TA	TAREAS ACADÉMICAS	2	Semana 5		NO
TP	TRABAJO PARCIAL	1	Semana 7		NO
EA	EVALUACIÓN PARCIAL	1	Semana 8		SÍ
TA	TAREAS ACADÉMICAS	3	Semana 11		NO
TA	TAREAS ACADÉMICAS	4	Semana 13		NO
DD	EVAL. DE DESEMPEÑO	1	Semana 15		NO
TF	TRABAJO FINAL	1	Semana 15		NO
EB	EVALUACIÓN FINAL	1	Semana 16		SÍ

IX. BIBLIOGRAFÍA DEL CURSO

https://upc.alma.exlibrisgroup.com/leganto/readinglist/lists/9509323320003391?institute=51UPC_INST&auth=LOCAL

ANEXO

En este anexo, se encuentran los reglamentos que todo alumno está obligado a leer y a cumplir en su rol de estudiante universitario en la UPC.

REGLAMENTO DE DISCIPLINA DE ALUMNOS :

<https://sica.upc.edu.pe/categoria/reglamentos-upc/sica-reg-26-reglamento-de-disciplina-de-alumnos>

REGLAMENTO PARA LA PREVENCIÓN E INTERVENCIÓN EN CASOS DE HOSTIGAMIENTO SEXUAL- UPC:

<https://sica.upc.edu.pe/categoria/normalizacion/sica-reg-31-reglamento-para-la-prevencion-e-intervencion-en-casos-de-hostiga>