

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
SYLLABUS
ANALÍTICA DE DATOS

FACULTAD: CIENCIAS Y EDUCACIÓN

PROYECTO CURRICULAR: MATEMÁTICAS

NOMBRE DEL DOCENTE:

ÁREA DE FORMACIÓN: ELECTIVA

ESPACIO ACADÉMICO: Asignatura (X), Grupo de Trabajo (),
Cátedra ()

Obligatorio () : Básico () Complementario ()

Electivo (X) :

CÓDIGO:

NUMERO DE ESTUDIANTES:

GRUPO:

NÚMERO DE CRÉDITOS: 2

TIPO DE CURSO: TEÓRICO PRÁCTICO TEO-PRAC: x

Alternativas metodológicas:

Clase Magistral (x), Seminario (), Seminario – Taller (), Taller (), Prácticas (), Proyectos tutoriados (), Otro: _____

HORARIO: Total Horas Semanales Lectivas: ____2____

DIA

HORA

SALON

I. JUSTIFICACIÓN DEL ESPACIO ACADÉMICO

El auge y crecimiento masificado del nuevo mundo de los datos, ha provocado que hayan surgido, en los últimos años, diversas herramientas, basadas en inteligencia artificial, para manipular estos datos y abstraer conocimiento a partir de allí, y a su vez ha crecido la demanda de capacidades para este tipo de especialidad.

Muy bien lo indican Azevedo y Santos (2008) en donde proponen metodologías ampliamente reconocidas y utilizadas como SEMMA y CRISP-DM. Multinacionales como Microsoft con su plataforma AZURE e IBM con Watson Analytics entre otras confirman este hecho.

Conocimientos previos: Se recomienda poseer conocimientos básicos en algún lenguaje de programación en particular (estructuras y funciones).

II. PROGRAMACIÓN DEL CONTENIDO

OBJETIVO GENERAL

Presentar una metodología clara para la introducción a la analítica de datos, así como las herramientas básicas, basada en estándares reconocidos mediante la aplicación práctica con el fin de llevar a cabo los diversos pasos que se deben tener en cuenta en esta labor.

OBJETIVOS ESPECÍFICOS

- Tener claros los pasos que se deben llevar a cabo en un proceso de analítica de datos.
- Manipular el lenguaje Python (en entorno Jupyter Notebook) y los paquetes más sobresalientes (Numpy, Matplotlib, Pandas, Sklearn, Scikit-Learn, Tensorflow, NLTK, Keras entre otros)
- Abarcar problemas de aprendizaje supervisado y no supervisado con algoritmos de Machine Learning.

PROPÓSITOS DE FORMACIÓN

Inducir al estudiante en el entorno de la analítica de datos, ampliamente aplicado en muchos campos y utilizado como base la Matemática y la Estadística. Además conocer terminología y variantes posibles con el fin de generar interés para su profundización posterior.

UNIDADES TEMÁTICAS Y/O PROBLEMÁTICAS

Introducción (Campos de aplicación, metodologías, herramientas)
Lenguaje Python (Entornos de desarrollo, preparación, instalación)
Acceso a fuentes de datos
Extracción, limpieza, preparación y visualización de datos
Creación de modelos
Prueba y validación
Casos prácticos

III. ESTRATEGIAS

Metodología Pedagógica y Didáctica:

Clases magistrales alternadas con sesiones de ejercicios y problemas en trabajo cooperativo. Cada tema se presenta de manera concisa con suficientes ejemplos ilustrativos. Es indispensable la generación y construcción de resultados fundamentales en la teoría y su demostración rigurosa.

Tipo de Curso	Horas			Horas profesor/semana	Horas Estudiante/semana	Total Horas Estudiante/semestre	Créditos
	TD	TC	TA	(TD + TC)	(TD + TC +TA)	X 16 semanas	
	2	0	2	2	4	64	2

Trabajo Presencial Directo (TD): trabajo de aula con plenaria de todos los estudiantes.

Trabajo Mediado _ cooperativo (TC): Trabajo de tutoría del docente a pequeños grupos o de forma individual a los estudiantes.

Trabajo Autónomo (TA): Trabajo del estudiante sin presencia del docente, que se puede realizar en distintas instancias: en grupos de trabajo o en forma individual, en casa o en biblioteca, laboratorio, etc.)

IV. RECURSOS

Medios y Ayudas: *El docente debe mostrar al estudiante los pormenores del trabajo matemático mediante sus exposiciones en el tablero. La utilización de recursos audiovisuales, recursos computacionales son convenientes como mediadores en el trabajo matemático.*

BIBLIOGRAFÍA

TEXTOS BÁSICOS

- González, Raúl. Python para todos
- Gopi Subramanian. Python Data Science Cookbook. Packt Publishing Ltd, 2015

TEXTOS COMPLEMENTARIOS

- Hands-On Machine Learning with Scikit-Learn and TensorFlow.
- Deep learning. Principios y fundamentos. Anna Bosch Rué, Jordi Casas Roma, Toni Lozano Bagén. Editorial UOC, 2019

REVISTAS

DIRECCIONES DE INTERNET

<https://www.python.org/>

<https://pypi.org/>

<https://jupyter.org/try>

<https://www.anaconda.com/products/individual>

<https://keras.io/examples/>

<https://matplotlib.org/>

V. ORGANIZACIÓN /TIEMPOS

Semana 1: Introducción (Campos de aplicación, metodologías, herramientas), Lenguaje Python (Entornos de desarrollo, preparación, instalación).

Semana 2,3,4: Uso de estructuras Python, sintaxis, variables. Lectura y escritura de archivos.

Semana 5,6: Manipulación de datos en Dataframe (Paquete Pandas), talleres prácticos

Semana 7,8: Acceso a fuentes de datos, extracción, limpieza, preparación y visualización de datos.

Semana 9,10,11,12: Aprendizaje supervisado y no supervisado con Machine Learning (Regresión, clasificación, clustering).

Semana 13, 14: Creación de modelos, prueba y validación

Semana 15, 16: Casos prácticos.

VI. EVALUACIÓN

La evaluación debe ser coherente con la metodología. Para incentivar el estudio permanente y cultivar la disciplina, se recomiendan pruebas escritas cortas y frecuentes.

Los estudiantes deben iniciarse en la aplicación de las formas de lenguaje, expresión y argumentación en matemáticas.

	TIPO DE EVALUACIÓN	FECHA	PORCENTAJE
PRIMER CORTE	Semana 8° de clases		35%
SEGUNDO CORTE	Semana 16° de clases		35%
EXAMEN FINAL	Luego de la semana 16°		30%