

 <p>UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS</p>	<p>UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS</p> <p>FACULTAD DE INGENIERIA</p> <p>SYLLABUS</p> <p><i>Página 1 de 6</i></p>	<p>FACULTAD DE INGENIERÍA</p> <p>Maestría en Ciencias de la Información y las Comunicaciones</p>
---	--	--

Maestría en Ciencias de la Información y las Comunicaciones

-Énfasis:

ESPACIO ACADÉMICO (ASIGNATURA): BIG DATA					
Código del espacio académico:					
Obligatorio		Básico		Complementario	
Electivo	X	Intrínseco		Extrínseco	
Fecha última actualización	26/06/2021		Grupo: 1		
Número de créditos:	3				
TIPO DE CURSO					
Teórico		Práctico		Teórico-Práctico	X
				Virtual	
ALTERNATIVAS METODOLÓGICAS					
Clase magistral		Seminario		Seminario-Taller	X
Taller		Prácticas		Proyectos tutorados	X
Otro:					
HORARIO					
Día	Horas		Salón		
Lunes	2				
Martes	2				
JUSTIFICACIÓN ESPACIO ACADÉMICO					
<p>En un mundo inundado de datos (Big Data), podemos tomar mejor las decisiones de negocio, con base en el uso estratégico de los datos.</p> <ul style="list-style-type: none"> • ¿Como los datos me permiten mejorar los productos a mis clientes? • ¿Como los datos mejoran la operación y reducen los riesgos? <p>¿Qué tienen en común las organizaciones digitales más exitosas del mundo, como Amazon y Alphabet (Google)?</p> <ul style="list-style-type: none"> • La respuesta: un enfoque en los datos. Donna Burbank, Directora General de Global Data Strategy, Ltd., dice que la conexión entre la excelente utilización de datos y el éxito empresarial no es una coincidencia. El Foro Económico Mundial ha declarado que, en el entorno actual y futuro del mercado, los datos son más valiosos que los activos físicos. <ul style="list-style-type: none"> ○ Los datos son el nuevo crudo de la humanidad <p>Por todo lo anterior se hace fundamental desarrollar dos grandes aspectos académicos en nuestros profesionales: Arquitecturas de Big Data, Bases de Datos NO-SQL y el uso</p>					



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

FACULTAD DE INGENIERIA
SYLLABUS

Página 2 de 6

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

estratégico de los datos gracias a la Inteligencia Artificial, el Machine Learning y el Análisis de Datos.

CONOCIMIENTOS PREVIOS: Estadística Básica y Ciencias de la Computación

PROGRAMACIÓN DE CONTENIDO

OBJETIVO GENERAL

El objetivo de este curso es estudiar y aplicar de forma práctica todos los conceptos de Arquitectura de Big Data, Machine Learning y Análisis de Datos, con el fin de que el estudiante pueda crear proyectos de big data y análisis de datos para una empresa y lograr la transformación digital de la empresa.

OBJETIVOS ESPECÍFICOS

Al finalizar el curso el estudiante estará en capacidad de:

- Conocer las arquitecturas de Big Data y los diferentes modelos de Base de Datos NO-SQL
- Conocer que es el análisis de datos y su aplicación en los negocios
- Conocer las diferentes etapas del análisis de datos
- Conocer diferentes técnicas y herramientas de análisis de datos
- Conocer las técnicas de machine learning y deeplearning para el análisis predictivo.
- Conocer las técnicas de NLP (Natural Language Procesing) para el análisis de data no estructurada texto.
- Aplicar el análisis de datos en un dominio de negocio utilizando alguna herramienta de análisis de datos.

COMPETENCIAS DE FORMACIÓN

GENERICAS

- Identificar los problemas y oportunidades de la empresa y diseñar las preguntas de análisis para responder con datos
- Identificar las fuentes de datos y ser capaz de estructurar los conjuntos de datos necesarios para la construcción de los modelos
- Construir los modelos descriptivos, diagnósticos y predictivos utilizando herramientas matemáticas, estadísticas y de machine learning.
- Comunicar de forma clara los resultados del analisis.

ESPECIFICAS

Cognitivas (Saber):

- Capacidad de identificar conocimientos y herramientas tecnologías conducentes a la resolución de problemas particulares en el desarrollo análisis de datos.
- Representar soluciones de problemas aplicando el modelamiento de este mediante la abstracción de lógica del negocio.
- Identificar los conceptos fundamentales necesarios para el rol de científico de datos.

Procedimentales / Instrumentales (Saber hacer):

- Modelar y evaluar problemas cuya solución requiere el uso de los diferentes conceptos manejados en el análisis de datos.



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

FACULTAD DE INGENIERIA
SYLLABUS

Página 3 de 6

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

- Modelar arquitecturas de big data, usando tecnologías y buenas prácticas de la industria.
- Identificar la importancia de las matemáticas, la estadística y las ciencias de la computación, áreas que constituyen la base de la disciplina de la Ciencia de los Datos y la Big Data.
- Utilizar los sistemas computacionales como herramienta de posibles soluciones a problemas específicos.

Actitudinales (Ser):

- Actuar estratégicamente dentro de un grupo de trabajo, cuyas soluciones constituyan análisis de datos.
- Actuar éticamente en el uso del análisis de datos y la big data como apoyo en las diferentes actividades humanas.
- Comunicarse estratégicamente haciendo uso de las tecnologías apoyadas en software.
- Actuar en contextos académicos y profesionales con un enfoque culto, ético y humanístico.
- Interpretar la realidad y proponer enfoques innovadores en contextos sociales.
- Presentar los trabajos de forma estética y usable, conforme al contexto al cual se destina

PROGRAMA (UNIDADES TEMÁTICAS Y CONTENIDO DETALLADO)

- Introducción a la Big Data y la importancia en el mundo hoy
- Big Data
 - Definiciones
 - Arquitecturas
 - Map-Reduce
 - Hadoop
 - Spark
 - Bases de Datos NO-SQL
- Análisis y ciencia de los datos
 - Los datos y la toma de decisiones en tu vida y en las organizaciones
 - Etapas en el análisis de datos
 - Tipos de Análisis
 - Descriptivo
 - Predictivo
 - Prescriptivo
 - Machine Learning
- Aplicación del Proceso de Análisis de Datos
 - Definición del objeto de análisis y elaboración de las preguntas de análisis
 - Definición de las variables y las fuentes de datos
 - Depuración de los datos
 - Análisis Descriptivo y Estadística Básica
 - Análisis de una variable



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

**UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS**

**FACULTAD DE INGENIERIA
SYLLABUS**

Página 4 de 6

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

- Análisis de 2 variables
- Análisis y presentación grafica de los datos
- Análisis Predictivo
 - Linear regression modeling
 - Classification modeling and Text Mining
 - Clustering and Similarity: Retrieving Documents
 - Recommender systems
 - Otras Técnicas de Machine Learning
 - ❖ Trees
 - ❖ Etc.
 - ❖ Deep Learning.
- NLP Natural Language Processing

ESTRATEGIAS

Metodología Pedagógica y Didáctica:

- Exposiciones magistrales por parte del docente de todos los conceptos teóricos del curso
- Desarrollo de casos prácticos individuales semana a semana
- La actividad colaborativa es primordial. La interacción dentro del espacio académico define un escenario de mediación cultural, en el cual la contrastación teórica, la verificación práctica y el diálogo colectivo son estrategias para la mejor asimilación de conceptos, la detección de fallos en las implementaciones y la generación de puntos de innovación.

La presentación de contenidos se realiza a partir de núcleos problémicos y la aplicación de los mismos en su resolución a partir de ejercicios prácticos y un proyecto de clase. La clase magistral está centrada en la presentación de prácticas de ejemplo, de las cuales se apropiarán conceptos previos tomados de lecturas, videos y podcast realizados como objetos de aprendizaje.

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	
	4	2	3	6	9	144	3

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.

RECURSOS



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

FACULTAD DE INGENIERIA
SYLLABUS

Página 5 de 6

FACULTAD DE INGENIERÍA
Maestría en Ciencias de la
Información y las Comunicaciones

- Clases magistrales para introducir los conceptos básicos utilizando diferentes recursos: presentaciones, videos, consulta de enlaces de interés.
- Talleres individuales de aplicación de los conceptos básicos
- Revisión bibliográfica, lecturas y exposiciones por parte de los estudiantes.
- Desarrollo, en grupo, de un proyecto de curso para implementar realizar un caso de análisis de datos y big data aplicado en una empresa pública o privada
- Correo Institucional: El docente cuenta con un correo institucional mediante el cual los estudiantes podrán contactarlo en caso de requerirlo.

ORGANIZACIÓN/TIEMPOS

Espacios, Tiempos, Agrupamientos:

Se recomienda trabajar una unidad cada cuatro semanas, trabajar en pequeños grupos de estudiantes, utilizar Internet para comunicarse con los estudiantes para revisiones de avances y solución de preguntas (esto considerarlo entre las horas de trabajo cooperativo).

Semana/ unidad temática	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Introducción	X															
2. Arquitectura Big Data		X	X													
3. Bases de datos grafos y semánticas				X												
4. Proceso de Análisis de Datos					X	X										
5. Curación y depuración de datos							X									
6. Análisis descriptivo								X	X							
7. Análisis diagnóstico										X						
8. Análisis predictivo y Machine learning											X	X	X			
9. NLP														X	X	
10. Proyecto final																X

EVALUACIÓN

Es importante tener en cuenta las diferencias entre evaluar y calificar. El primero es un proceso cualitativo y el segundo un estado terminal cuantitativo que se obtiene producto de

 <p>UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS</p>	<p>UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS</p> <p>FACULTAD DE INGENIERIA</p> <p>SYLLABUS</p> <p><i>Página 6 de 6</i></p>	<p>FACULTAD DE INGENIERÍA</p> <p>Maestría en Ciencias de la Información y las Comunicaciones</p>
---	--	--

la evaluación. Para la obtención de la información necesaria para los procesos de evaluación se requiere diseñar distintos formatos específicos de autoevaluación, coevaluación y heteroevaluación. Una formación en competencias requiere: 1. Evaluación de los aprendizajes de los estudiantes en sus dimensiones: individual/grupo, teoría/práctica, oral escrita. 2. Autoevaluación: la evaluación del desempeño del estudiante realizado por el mismo. 3. Coevaluación del desempeño de los estudiantes entre estudiantes y docente. 4. Evaluación del desempeño docente.

La evaluación se realizará teniendo en cuenta:

	TIPO DE EVALUACIÓN	FECHA	PORCENTAJE
PRIMERA NOTA	Evaluación Teórica de los conceptos de Arquitecturas de Big Data y Bases de Datos NOSQL		20%
SEGUNDA NOTA	Evaluación teórica de análisis de datos y machine learning		20%
TERCERA NOTA	Presentación de proyectos aplicados cada semana		30%
CUARTA NOTA	Proyecto final		30%

ASPECTOS A EVALUAR DEL CURSO

Evaluación de los aprendizajes de los estudiantes en sus dimensiones: individual/grupo, teórica/práctica, oral/escrita. Se puede considerar la autoevaluación y la coevaluación del curso: de forma oral entre estudiantes y docente.