



# Big Data y Análisis de Datos

## Ingeniería de Software con Inteligencia Artificial

# Tecnologías de la Información

## Nivel Profesional Técnico



## CUADRO PROGRAMA

ESCUELA: TECNOLOGÍAS DE LA INFORMACIÓN  
CARRERA: INGENIERÍA DE SOFTWARE CON INTELIGENCIA ARTIFICIAL  
MÓDULO FORMATIVO: BIG DATA Y ANÁLISIS DE DATOS

### OPERACIONES

Define principales conceptos de Big Data  
Define las fuentes de generación de datos masivos  
Investiga el impacto del Big Data en la toma de decisiones  
Describe la arquitectura Big Data  
Define y estudia los sistemas de archivos distribuidos  
Crea bases de datos NoSQL  
Investiga los tipos de almacenamiento en la nube  
Describe los conceptos y principios de procesamiento de datos  
Investiga y describe las tendencias emergentes en Big Data  
Conoce las consideraciones éticas y legales en el manejo de datos

N°	Cod HT	TAREAS	Cod HO	HO-01	HO-02	HO-03	HO-04	HO-05	HO-06	HO-07	HO-08	HO-09	HO-10	HO-11
1	HT-01	Define los fundamentos e importancia del Big Data												
2	HT-02	Conoce el almacenamiento de datos masivos												
3	HT-03	Conoce herramientas de procesamiento de datos para Big Data												
4	HT-04	Revisa y conoce las aplicaciones en el mercado del Big Data												



Operación Nueva



Operación Repetida

## PROGRAMA DE FORMACIÓN PROFESIONAL

### HOJA DE PROGRAMACIÓN

Escuela: **TECNOLOGÍAS DE LA INFORMACIÓN**Módulo Formativo: **BIG DATA Y ANÁLISIS DE DATOS**Semestre: **VI**Carrera: **INGENIERÍA DE SOFTWARE CON INTELIGENCIA ARTIFICIAL**Módulo Ocupacional: **ESPECIALISTA EN SISTEMAS DE INFORMACIÓN Y CLOUD COMPUTING****Objetivo General:**

Al concluir este módulo, los participantes habrán adquirido la capacidad de comprender de manera sólida los fundamentos del Big Data y el análisis de datos, además de contar con las habilidades esenciales para abordar y gestionar conjuntos masivos de datos. Este curso también estimulará una comprensión crítica sobre la relevancia del Big Data en una variedad de sectores industriales.

SEM (SEMANA)	CONTENIDOS DE APRENDIZAJE			
	PROYECTOS TAREAS DE APRENDIZAJE	OPERACIONES	CONOCIMIENTOS TECNOLÓGICOS	CONOCIMIENTOS COMPLEMENTARIOS
1	Define los fundamentos e importancia del Big Data	<ul style="list-style-type: none"> <li>Define principales conceptos de Big Data.</li> <li>Define las fuentes de generación de datos masivos.</li> <li>Investiga el impacto del Big Data en la toma de decisiones empresariales.</li> </ul>	<ol style="list-style-type: none"> <li><b>Fundamentos del Big Data:</b> definición, características y evolución del concepto.</li> <li><b>Arquitectura general de soluciones Big Data:</b> capas de adquisición, almacenamiento, procesamiento y visualización.</li> <li><b>Tipos de datos:</b> estructurados, semi-estructurados y no estructurados.</li> <li><b>Fuentes de generación de datos masivos:</b> redes sociales, sensores IoT, transacciones, logs.</li> <li><b>Sistemas de información y su integración con Big Data:</b> ERP, CRM, SCM.</li> <li><b>Introducción al Cloud Computing:</b> definición, tipos de servicios (IaaS, PaaS, SaaS) y su relación con el procesamiento de Big Data.</li> </ol>	<ul style="list-style-type: none"> <li>Diferencias entre dato e información.</li> <li>Definición de Minería de datos.</li> </ul>
2	Conoce el almacenamiento de datos masivos	<ul style="list-style-type: none"> <li>Describe la arquitectura Big Data.</li> <li>Define y estudia los sistemas de archivos distribuidos GFS y HDFS.</li> <li>Crea bases de datos NoSQL (MongoDB, Apache Cassandra, etc.)</li> <li>Investiga los tipos de almacenamiento en la nube.</li> </ul>	<ol style="list-style-type: none"> <li><b>Arquitectura Big Data:</b> capas de ingesta, almacenamiento, procesamiento y visualización.</li> <li><b>Sistemas de archivos distribuidos:</b> definición, funcionamiento, ventajas y desventajas de HDFS (Hadoop Distributed File System) y GFS (Google File System).</li> </ol>	<ul style="list-style-type: none"> <li>Clúster de computadoras</li> <li>Procesamiento distribuido.</li> </ul>

## PROGRAMA DE FORMACIÓN PROFESIONAL

### HOJA DE PROGRAMACIÓN

Escuela: **TECNOLOGÍAS DE LA INFORMACIÓN**Módulo Formativo: **BIG DATA Y ANÁLISIS DE DATOS**Semestre: **VI**Carrera: **INGENIERÍA DE SOFTWARE CON INTELIGENCIA ARTIFICIAL**Módulo Ocupacional: **ESPECIALISTA EN SISTEMAS DE INFORMACIÓN Y CLOUD COMPUTING****Objetivo General:**

Al concluir este módulo, los participantes habrán adquirido la capacidad de comprender de manera sólida los fundamentos del Big Data y el análisis de datos, además de contar con las habilidades esenciales para abordar y gestionar conjuntos masivos de datos. Este curso también estimulará una comprensión crítica sobre la relevancia del Big Data en una variedad de sectores industriales.

SEM (SEMANA)	CONTENIDOS DE APRENDIZAJE			
	PROYECTOS TAREAS DE APRENDIZAJE	OPERACIONES	CONOCIMIENTOS TECNOLÓGICOS	CONOCIMIENTOS COMPLEMENTARIOS
			<b>3) Bases de datos NoSQL:</b> definición, características (escalabilidad horizontal, ausencia de esquema fijo, etc.), comparación con bases de datos relacionales. <b>4) Tipos de bases de datos NoSQL:</b> documentales (MongoDB), de columnas (Cassandra), clave-valor (Redis), grafos (Neo4j). <b>5) Estructuras de datos JSON:</b> uso en bases NoSQL, ventajas para intercambio y almacenamiento. <b>6) Sistemas de almacenamiento en la nube:</b> almacenamiento en bloques, archivos y objetos. Servicios como AWS S3, Google Cloud Storage, Azure Blob Storage.	
3	Conoce herramientas de procesamiento de datos para Big Data	<ul style="list-style-type: none"> <li>Estudia los Frameworks de procesamiento de datos Hadoop Spark y Flink.</li> <li>Describe los conceptos y principios de MapReduce.</li> </ul>	<b>1) Frameworks de procesamiento de datos:</b> conceptos, estructura y funcionamiento de Hadoop, Spark y Flink. <b>2) MapReduce:</b> modelo de programación distribuida basado en funciones de mapeo y reducción. <b>3) Comparativa entre Spark y Hadoop MapReduce:</b> velocidad de	<ul style="list-style-type: none"> <li>Almacenamiento de datos por RAID de discos.</li> </ul>



## PROGRAMA DE FORMACIÓN PROFESIONAL

### HOJA DE PROGRAMACIÓN

**Escuela:** TECNOLOGÍAS DE LA INFORMACIÓN

**Módulo Formativo:** BIG DATA Y ANÁLISIS DE DATOS

**Semestre:** VI

**Carrera:** INGENIERÍA DE SOFTWARE CON INTELIGENCIA ARTIFICIAL

**Módulo Ocupacional:** ESPECIALISTA EN SISTEMAS DE INFORMACIÓN Y CLOUD COMPUTING

#### Objetivo General:

Al concluir este módulo, los participantes habrán adquirido la capacidad de comprender de manera sólida los fundamentos del Big Data y el análisis de datos, además de contar con las habilidades esenciales para abordar y gestionar conjuntos masivos de datos. Este curso también estimulará una comprensión crítica sobre la relevancia del Big Data en una variedad de sectores industriales.

SEM (SEMANA)	CONTENIDOS DE APRENDIZAJE			
	PROYECTOS TAREAS DE APRENDIZAJE	OPERACIONES	CONOCIMIENTOS TECNOLÓGICOS	CONOCIMIENTOS COMPLEMENTARIOS
			<p>procesamiento en memoria, tolerancia a fallos, y flexibilidad.</p> <p>4) <b>Componentes de procesamiento distribuido:</b> nodos maestros y trabajadores, clúster managers, drivers, ejecutores.</p> <p>5) <b>Contenedores y virtualización ligera:</b> uso de Docker para implementar entornos Big Data modulares.</p> <p>6) <b>Pipeline de procesamiento de datos:</b> etapas desde la ingesta hasta la agregación de resultados.</p>	
4	Revisa y conoce las aplicaciones en el mercado del Big Data	<ul style="list-style-type: none"> <li>Investiga y describe las tendencias emergentes en Big Data.</li> <li>Conoce las consideraciones éticas y legales en el manejo de datos masivos.</li> </ul>	<p>1) <b>Panorama actual del Big Data:</b> tendencias en tiempo real, analítica predictiva, analítica prescriptiva y visualización avanzada.</p> <p>2) <b>Big Data e IoT:</b> arquitectura de sensores, generación de datos de telemetría, y flujos de datos.</p> <p>3) <b>Servicios de nube para análisis de datos:</b> características y beneficios de soluciones como Amazon Web Services, Google Cloud Platform y Azure para Big Data.</p> <p>4) <b>Clasificación de datos según sensibilidad:</b> datos personales, confidenciales, sensibles y públicos.</p>	<ul style="list-style-type: none"> <li>Business Intelligence (BI).</li> <li>Ética informática.</li> </ul>



## PROGRAMA DE FORMACIÓN PROFESIONAL

### HOJA DE PROGRAMACIÓN

**Escuela:** TECNOLOGÍAS DE LA INFORMACIÓN

**Módulo Formativo:** BIG DATA Y ANÁLISIS DE DATOS

**Semestre:** VI

**Carrera:** INGENIERÍA DE SOFTWARE CON INTELIGENCIA ARTIFICIAL

**Módulo Ocupacional:** ESPECIALISTA EN SISTEMAS DE INFORMACIÓN Y CLOUD COMPUTING

**Objetivo General:**

Al concluir este módulo, los participantes habrán adquirido la capacidad de comprender de manera sólida los fundamentos del Big Data y el análisis de datos, además de contar con las habilidades esenciales para abordar y gestionar conjuntos masivos de datos. Este curso también estimulará una comprensión crítica sobre la relevancia del Big Data en una variedad de sectores industriales.

SEM (SEMANA)	CONTENIDOS DE APRENDIZAJE			
	PROYECTOS TAREAS DE APRENDIZAJE	OPERACIONES	CONOCIMIENTOS TECNOLÓGICOS	CONOCIMIENTOS COMPLEMENTARIOS
			5) <b>Conceptos legales aplicados a Big Data:</b> introducción a GDPR, Ley de Protección de Datos Personales, y principios de privacidad por diseño.	



**SENATI**