



**FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA**

DATOS GENERALES DE IDENTIFICACIÓN DEL CURSO				
<b>NOMBRE DEL CURSO</b>		<b>HRS. TOTALES:</b>	<b>TEORÍA</b>	<b>PRÁCTICA</b>
Big Data: principios y aplicaciones del análisis de datos masivos		20	8	12
<b>NIVEL:</b>	<b>MODALIDAD:</b>	<b>ÁREA:</b>	<b>UNIDAD QUE SOLICITA EL REGISTRO:</b>	
Medio Superior	Presencial	Físico Matemático	Centro de Educación Continua Unidad Morelos	
<b>PERIODO DE IMPARTICIÓN</b>			<b>SEDE</b>	
10 al 13 de diciembre de 2019			Centro de Educación Continua Unidad Morelos	
<b>DIRIGIDO A:</b>		<b>REQUISITOS ACADÉMICOS Y ADMINISTRATIVOS DE INGRESO DEL PARTICIPANTE:</b>		
Estudiantes y público en general de nivel medio superior en adelante con conocimientos básicos en computación.		Este programa educativo considera el siguiente perfil de ingreso: >> Académico: Computación básica, destreza en el manejo de archivos y carpetas, instalación de software en una PC y conocimiento del funcionamiento general de la Internet. >> Administrativo: Formato de Registro y ficha de depósito		
<b>RECURSOS DE OPERACIÓN DEL CURSO</b>				
<ul style="list-style-type: none"><li>- Aula equipada con 15 computadoras</li><li>- Material de apoyo</li><li>- Pizarrón, borrador, plumones</li><li>- Proyector</li></ul>				

<b>JEFA DE EDUCACIÓN CONTINUA</b>	<b>REVISÓ Y AVALÓ</b>	<b>DIRECTORA DEL CENTRO O UNIDAD</b>
Mtra. Karen Itzel Adán Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz

CENTRO DE EDUCACIÓN CONTINUA  
UNIDAD "MORELOS"



## FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

### FUNDAMENTACIÓN

Hoy en día, las tecnologías de la información (TI) ya no solo están enfocadas a la obtención de datos, a la clasificación, la gestión, la distribución y el procesamiento de la información. Además, según la consultora Gartner (2019), en la actualidad se generan más datos en dos días que en toda la historia contemporánea. Al respecto, según TICbeat (2017): "en 2015 se generaron nada menos que 145 Zettabytes, con la previsión de que las personas, máquinas y objetos conectados a Internet hagan lo propio con 600 Zettabytes de datos en 2020" (s. p.). No obstante, a pesar de la gran cantidad de datos, se calcula que aproximadamente el 10% de estos son útiles y solamente otro 10% de datos serán procesados y convertidos en valor de negocio mediante analítica de Big Data.

Su evolución está permitiendo que las tecnologías se transformen en conocimiento y sabiduría, orillando a las empresas y a las organizaciones en general a adoptar nuevas estrategias para poder aprovechar dicho conocimiento y a reconocer su íntima relación con otras tecnologías, como son la inteligencia artificial, la analítica de datos o el Internet de las cosas (como fuente de enormes volúmenes de datos).

Por esta razón, esta acción formativa ofrece herramientas para el manejo de ciencia y volumen de datos desde su obtención, análisis, visualización y conjunción de los mismos en la internet, hasta llegar a su procesamiento siendo moldeados y probados por la propia empresa de tal forma que se transformen en el alcance de resultados para su implementación tecnológica dentro de la organización y se identifiquen nuevas oportunidades llevando a un constante crecimiento. Al término, el participante estará habilitado para manejar grandes volúmenes de datos cuyo análisis desarrollen ideas que conduzcan a mejores decisiones y movimientos de negocios estratégicos.

### Unidades Académicas de colaboración

NINGUNA

### Sector con el que se vincula el programa

PÚBLICO Y PRIVADO

### Introducción

Las organizaciones se enfrentan a un mercado tecnológico cambiante y con nuevas oportunidades de aprovechar herramientas y metodologías para así mejorar su competitividad e identificar las mejores prácticas en materia de manejo de datos. Debido a que la sociedad por sí sola es incapaz de procesar y memorizar todo ese mundo cambiante y torrente de datos (Big Data) que continuamente la está bombardeando desde los medios sociales, los teléfonos inteligentes y demás fuentes de datos, surge la necesidad de estudiar las tecnologías emergentes que permiten afrontar el crecimiento continuo de las mismas y sobre todo el manejo de los datos. Lo anterior, mediante programas educativos diseñados con prototipos que ilustren cómo abordar los cambios tecnológicos y cómo resolver los problemas.

JEFA DE EDUCACIÓN CONTINUA

Mtra. Karen Itzel Adán Delgado

REVISAR Y AVALAR

Ing. Brisa Arlett Aguilar Méndez

DIRECTORA DEL CENTRO O UNIDAD

Dra. Karla Edith Campos Díaz

INSTITUTO POLITÉCNICO NACIONAL  
CENTRO DE EDUCACIÓN CONTINUA  
UNIDAD "MORELOS"



### FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

Este programa educativo permitirá una mirada global a la tecnología así como las fuentes de datos masivos y áreas de aplicación, mediante el conocimiento de herramientas en el tema de minería de datos y Big Data, ofreciendo al participante una visión tecnológica dentro de las organizaciones para aprovechar sus datos y utilizarlos para identificar nuevas oportunidades, lo que conducirá a movimientos de negocios más inteligentes, operaciones más eficientes, mayores ganancias y clientes más felices beneficiándose en tema de reducción de costos, mejor toma de decisiones, productos y servicios.

Durante el desarrollo de 20 horas, el participante conocerá la estructura base sobre el conocimiento del uso, manejo y aprovechamiento de grandes cantidades de datos fomentando el desarrollo tecnológico dentro de las empresas u organizaciones mediante la instalación de software WEKA y la creación de 3 prototipos de aplicación en el campo de la medicina.

#### Evaluación y tipo de reconocimiento que se entregará

El proceso de formación se evaluará de manera continua, tomando en cuenta los diferentes momentos de la evaluación formativa, a saber:

- **Evaluación diagnóstica:** En esta etapa, el facilitador realizará actividades de integración grupal y con fines diagnósticos con la intención de reconocer aptitudes y actitudes de los integrantes del grupo, con el propósito de establecer estrategias que permitan potenciar el desarrollo de competencias en los participantes. Esta etapa no tendrá ninguna valoración.
- **Evaluación formativa:** En esta etapa se realizarán las siguientes acciones formativas:
  - ✓ *Actividades de aprendizaje:* Contempla el desarrollo de prototipos tecnológicos destinados al ámbito médico a través del trabajo en equipo, lo que permitirá a los participantes sistematizar sus ideas y formular sus proyectos.
  - ✓ *Participación colaborativa.* Los participantes se dividirán en grupos para trabajar dinámicas grupales o trabajo colaborativo que permita arribar a la conformación de prototipos.La evaluación formativa tendrá una valoración del 80% sobre la calificación final.
- **Evaluación sumativa:** Para la evaluación final, los participantes diseñarán un sistema de reparación de fallas mecánicas y eléctricas basado en la nube para todas las marcas de autos y modelos fabricados desde 1970 a la fecha. Su aportación a la calificación final será del 30%.

Para la acreditación del curso, el participante debe cumplir un mínimo de 80% de asistencia y calificación.

**Reconocimiento:** Constancia de participación emitida por la Dirección de Educación Continua del Instituto Politécnico Nacional.

JEFA DE EDUCACIÓN CONTINUA	REVISÓ Y AVALÓ	DIRECTORA DEL CENTRO O UNIDAD
Mtra. Karen Itzel Adán Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz
		INSTITUTO POLITÉCNICO NACIONAL CENTRO DE EDUCACIÓN CONTINUA UNIDAD "MORELOS"





### FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

#### COMPETENCIA GENERAL A DESARROLLAR:

Utiliza la herramienta de minería de datos WEKA para la resolución de problemas relacionados con grandes volúmenes de información, contemplando las tendencias en el manejo de datos para la conformación de un sistema de reparación de fallas mecánicas y eléctricas.

#### PROGRAMA

TEMAS	COMPETENCIAS ESPECÍFICAS	ACTIVIDADES DE APRENDIZAJE	EVIDENCIAS DE APRENDIZAJE	HORAS
<b>Módulo 1. Una mirada global de la tecnología</b> 1.1 Qué es y para qué sirve el Big Data 1.2 Volumen, velocidad, variedad y veracidad de los datos masivos 1.3 Sacándole provecho a la tecnología 1.4 Administración, organización y análisis de los datos masivos 1.5 Almacenamiento de grandes volúmenes de datos 1.6 Procesamiento rápido de corrientes muy variopintas de datos 1.7 Minería de datos y aprendizaje de máquina	Utiliza componentes y aplicaciones de la herramienta de minería de datos como es WEKA con el propósito de pronosticar la posibilidad de sobrevivencia de un paciente con diabetes.	<b>Encuadre.</b> El instructor expone los objetivos, los contenidos, las formas de trabajo y de evaluación a lo largo del curso. Posteriormente, a través de dinámicas grupales y lluvia de ideas, establece el diagnóstico del grupo con el fin de adecuar las estrategias a las necesidades del grupo. Al finalizar, expone el tema para introducir al participante en dicha temática.  <b>Actividad 1.</b> Los participantes formarán equipos de trabajo y cada uno de ellos instalará la herramienta WEKA. Posteriormente, darán solución a la problemática presentada en el Prototipo de aplicación 1. Minería de datos clínicos como apoyo en el diagnóstico médico, donde el facilitador ilustrará mediante un ejemplo cuál es la mejor forma de diagnosticar a un paciente con problemas	Prototipo de aplicación e instalación del software WEKA. Valor 10%  Debate de los equipos para implementar el prototipo No. 1. Valor 5%	5 horas totales Teoría: 2 horas Práctica: 3 horas

JEFA DE EDUCACIÓN CONTINUA

Mtra. Karen Itzel Adán Delgado

REVISY AVALA

Ing. Brisa Arlett Aguilar Méndez

DIRECTORA DEL CENTRO O UNIDAD

Dra. Karla Edith Campos Díaz

INSTITUTO POLITÉCNICO NACIONAL  
CENTRO DE EDUCACIÓN CONTINUA  
UNIDAD "MORELOS"



FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

1.8 Segmentación de los datos y modelos predictivos 1.9 Prototipo de aplicación		de diabetes, así como el procedimiento más eficaz. Al término, el instructor fomenta la discusión mediante un debate con la finalidad de arribar a acuerdos sobre la utilidad del prototipo 1 y la implementación de mejoras.  El facilitador proporciona un cuestionario a los grupos de trabajo para evaluar los temas del primer módulo.	Cuestionario resuelto sobre los temas estudiados. Valor 5%	
<b>Módulo 2. Fuentes de datos masivos y áreas de aplicación del Big Data</b> 2.1 Fuentes de datos masivos 2.2 Comunicación entre usuarios (medios sociales) 2.3 Comunicación ser humano-máquina (acceso a la web) 2.4 Comunicación entre máquinas (RFIDs y sensores) 2.5 Aplicaciones del Big Data 2.6 Fidelización de clientes	Desarrolla un prototipo de terapias inteligentes con medicina de precisión (personalizada), con la capacidad de personalizar recetas médicas, empleando la técnica del aprendizaje de máquina no supervisado y basada en las aplicaciones para la comunicación entre máquinas y seres humanos.	<b>Actividad 2.</b> Los equipos de trabajo conformados en la actividad pasada, se reúnen para dar continuidad al desarrollo del prototipo propuesto en el módulo 1 en el que evaluarán, a través de un instrumento <i>ex profeso</i> , lo siguiente: <ul style="list-style-type: none"><li>Eficacia en la captura de datos del prototipo 1</li><li>Obtención de al menos un modelo para implementar el prototipo</li><li>Optimización del modelo.</li><li>Nuevos descubrimientos sobre el modelo optimizado o con el anterior</li></ul> De acuerdo con los resultados obtenidos, cada equipo preparará un reporte y una exposición de dicha información con el fin de intercambiar su experiencia con sus pares.	Validación del prototipo 1. Valor 5%  Reporte de avances del prototipo 1 y exposición por equipos. Valor 5%	5 horas totales: Teoría: 2 horas Práctica: 3 horas

JEFA DE EDUCACIÓN CONTINUA  
  
Mtra. Karlem Itzel Adan Delgado

REVISÓ Y AVALÓ  
  
Ing. Brisa Arlett Aguilar Méndez

DIRECTORA DEL CENTRO O UNIDAD  
  
Dra. Karla Edith Campos Díaz



FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

2.7 Desarrollo de nuevos productos		<b>Actividad 3.</b> Con base en la valoración e intercambio de opiniones y sugerencias de la actividad previa, cada equipo resolverá los problemas detectados en el prototipo y dará continuidad para resolver los problemas detectados y conformar el Prototipo 2. Terapias inteligentes con medicina de precisión (personalizada). Al concluir, cada equipo intercambiará su experiencia a través de una exposición.	Propuesta de prototipo 2 y de exposición resultados. Valor: 10%	
<b>Módulo 3. Arquitectura del Big Data</b> 3.1 Arquitectura del cuestionamiento en Google 3.2 Arquitectura estándar del Big Data 3.3 Ejemplos de arquitecturas: IBM Watson, Netflix, Ebay, Ticket Master, LinkedIn y Paypal 3.4 Procesamiento distribuido con Hadoop 3.5 La plataforma de Hadoop	Utiliza aplicaciones para la construcción del prototipo de inteligencia artificial como auxiliar en el tratamiento de enfermedades idiopáticas, basadas en el procesamiento en paralelo y distribución de la memoria tanto en arquitecturas estándar como comerciales, enfatizando en el sistema de archivos distribuidos de la plataforma Hadoop de Apache (HDFS).	<b>Actividad 4.</b> Cada equipo retomará el prototipo que construyeron en el módulo 2 y, de acuerdo con el resultado, elaborarán un reporte que dé respuesta a lo siguiente: <ul style="list-style-type: none"><li>• Determinar formas de implementación del aprendizaje no supervisado en WEKA.</li><li>• Número de experimentos (iteraciones) que tuvieron que realizar para encontrar la receta solicitada.</li><li>• Número de grupos en lo que tuvieron que dividir los datos para encontrar la receta solicitada.</li><li>• Grupo en el que la máquina clasificó al paciente.</li></ul>	Reporte de avances del prototipo 2 y de exposición grupal. Valor 10%	5 horas totales: Teoría: 2 horas Práctica: 3 horas

JEFE DE EDUCACIÓN CONTINUA	REVISY AVALA	DIRECTORA DEL CENTRO O UNIDAD
Mtra. Karen Itzel Adán Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz



**FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA**

3.6 Objetivos del diseño HDFS 3.7 Arquitectura maestro-esclavo 3.8 Sistema de almacenamiento en bloques 3.9 Escritura y lectura de archivos locales y oleadas de datos en HDFS 3.10 Archivos secuenciados 3.11 YARN (Yet Another Resource Negotiator)		<ul style="list-style-type: none"><li>Determinar las dosis que más beneficiarán a un paciente en particular, para cada uno de los nutrientes sugeridos por la máquina.</li></ul> <p>Al término, cada equipo realizará una exposición de resultados.</p> <p><b>Actividad 5.</b> En equipos, los participantes realizarán la construcción del prototipo 3, Inteligencia artificial como auxiliar en el tratamiento de enfermedades idiopáticas y expondrán los avances del prototipo propuesto en el módulo 2.</p>	Propuesta de prototipo 3. Valor: 15%	
<b>Módulo 4: Otras tecnologías para la implementación del Big Data</b> 4.1 Qué es y cómo trabaja el MapReduce de Google 4.2 La oleada de datos (streaming) de Hadoop 4.3 Los lenguajes Hive y Pig de Apache 4.4 Bases de datos NoSQL	Utiliza principios de las bases de datos no relacionales (NoSQL), así como lenguajes Pig y Hive para acceder con mayor eficiencia a caudales de datos (streaming) al complementarse con los servicios que ofrece la nube (cloud computing), para la conformación de un sistema de reparación de fallas mecánicas y eléctricas.	<p><b>Actividad 6.</b> Cada equipo de trabajo expone el prototipo propuesto en el módulo 3. Inteligencia artificial como auxiliar en el tratamiento de enfermedades idiopáticas</p> <p><b>Actividad integradora.</b> En equipo, los participantes desarrollan una propuesta de un sistema de reparación de fallas mecánicas y eléctricas basado en la nube para todas las marcas de autos y modelos fabricados desde 1970 a la fecha.</p>	Exposición por grupos Valor total: 10% <p>Propuesta del sistema de reparación de fallas mecánicas y eléctricas Valor 30%</p>	5 horas totales: Teoría: 2 horas Práctica: 3 horas

JEFA DE EDUCACIÓN CONTINUA	REVISÓ Y AVALÓ	DIRECTORA DEL CENTRO O UNIDAD
Mtra. Karen Ivel Adán Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz





**FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA**

4.5 Procesamiento de oleadas de datos con Spark				
4.6 Sistemas de ingestión de datos				
4.7 Sistemas de mensajería				
4.8 Kafka				
4.9 Computación en la nube (cloud computing) para Big Data				

**BIBLIOGRAFÍA:**

1. Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
2. Data, B. (2014). Principles and best practices of scalable realtime data systems. *N. Marz J. Warren. Henning*.
3. Maass, W., Parsons, J., Puro, S., Rosales, A., Storey, V. C., & Woo, C. C. (2017). Big data and theory. *Encyclopedia of big data*, 1-5.
4. Medrano Parado, S. Z. (2017). Modelo de minería de datos usando machine learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016.
5. Sánchez, M. J. H., & Mirón, C. C. Aplicaciones de la inteligencia a la medicina.

JEFE DE EDUCACIÓN CONTINUA	REVISAR Y AVALAR	DIRECTORA DEL CENTRO O UNIDAD
Mtra. Karen Itzel Adán Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz





## FORMATO DE REGISTRO DE CURSOS DE FORMACIÓN A LO LARGO DE LA VIDA

### DATOS DEL INSTRUCTOR

MÓDULO/TEMA	NOMBRE (S) DE (LOS) INSTRUCTOR(ES)
Módulo 1: Una mirada global de la tecnología Módulo 2: Fuentes de datos masivos y áreas de aplicación del Big Data Módulo 3: Arquitectura del Big Data Módulo 4: Otras tecnologías para la Implementación del Big Data	Mtro. Sergio López González

\*Adjuntar formato de Currículum Vitae de cada Instructor

### DATOS DE COORDINADOR ACADÉMICO

Nombre:	Mtra. Kareem Itzel Adan Delgado
Cargo:	Jefa del Departamento de Educación Continua
Número Telefónico y extensión:	55. 5729 6000 Ext. 81704 / 81702
Correo Electrónico:	<a href="mailto:Kadand@ipn.mx">Kadand@ipn.mx</a> / <a href="mailto:econtinua.cecmorelos@gmail.com">econtinua.cecmorelos@gmail.com</a>

JEFA DE EDUCACIÓN CONTINUA	REVISÓ Y AVALÓ	DIRECTORA DEL CENTRO O UNIDAD
Mtra. Kareem Itzel Adan Delgado	Ing. Brisa Arlett Aguilar Méndez	Dra. Karla Edith Campos Díaz