

# **Curso del Data Mining al Big Data**

**Instructor: Dr. Luis Carlos Molina Félix**

## **Presentación.**

Las bases de datos y los sistemas de administración de datos han jugado un papel primordial en el crecimiento y éxito de las organizaciones corporativas, los cambios en la economía y la globalización de las empresas han aumentado su importancia. Esto ha dado origen a que los departamentos de tecnología de la información inicien la búsqueda y explotación de nuevas herramientas y paradigmas para el procesamiento de datos y de la información generada.

Con ello inicia el concepto de Inteligencia de Negocios permitiendo a las organizaciones realizar análisis sobre un conjunto de datos con el objetivo de obtener y mejorar el conocimiento de las operaciones de la empresa y ayudar en la toma de decisiones, incrementando o manteniendo la competitividad de la organización.

Existe un crecimiento en el uso de técnicas y enfoques tradicionales para hacer uso de bases de datos empresariales y herramientas de análisis de datos, por lo que en los diferentes mercados y tipos de organizaciones se ha requerido un mayor estudio y entendimiento sobre estas herramientas, modelos, procesos, algoritmos, etc., referentes al descubrimiento de conocimiento y que gracias a la integración de tecnologías como Minería de Datos con Inteligencia de Negocios es posible ofrecer de manera automatizada e inteligente información de suma importancia para ejecutar acciones para la solución de problemas organizacionales.

La Minería de Datos es un proceso fundamental en Inteligencia de Negocios, ya que permite obtener conocimiento a partir de los datos almacenados, mediante un proceso no trivial de extracción de información implícita, previamente desconocida y potencialmente útil. Descubrir conocimiento implica buscar patrones de comportamiento

aún no conocidos en los datos. El conocimiento se puede manifestar como: patrones, reglas de conocimiento, asociaciones, grupos, restricciones, tendencias, etc. Los dominios de aplicación prácticamente abarcan cualquier campo y tipo de empresas (comercialización, producción, finanzas, energía, gobierno, tecnología, etc.).

## Objetivos

Al final del curso los participantes estarán en capacidad de:

- Conocer de manera general las técnicas y enfoques de Minería de Datos y Big Data
- Exploración y uso de fuentes de datos para análisis y toma de decisiones
- Comprender el proceso de la Minería de Datos para extraer conocimiento desde base de datos y la aplicación de estas técnicas a la solución de problemas de reconocimiento de patrones, clasificación y pronóstico
- Diseñar, desarrollar y analizar programas de software para el análisis de datos, tales como reglas de conocimiento, asociaciones, grupos, restricciones, tendencias, patrones, etc.

## Curriculum Vitae

Luis Carlos Molina Félix, realizó sus estudios de maestría en la Universidad de Sao Paulo, Brasil y de doctorado en la Universidad Politécnica de Cataluña. Desde 1996 se ha dedicado de manera ininterrumpida al tema de Minería de Datos (Data Mining). Entre algunos proyectos que ha conducido destacan una gran cantidad de bancos, para temas de prevención de fraude en tarjetas de crédito y lavado de dinero, también tiendas de conveniencia para analizar los perfiles de clientes y análisis de canasta y empresas de telefonía celular para análisis de fuga de cliente y mejora de la rentabilidad. Su modelo de análisis de información para telefonía celular fue presentado en un Congreso Mundial de Telefonía Móvil en Singapur con una gran aceptación.

Ha sido autor, junto con Ramón Sangüesa de libro “Data Mining, una introducción”. También es autor del artículo “Data Mining: Torturando los datos hasta que confiesen”. Actualmente está por publicar el artículo: Torturando los datos hasta que confiesen 10 años después.

Desde el 2005 trabaja como Director en una empresa que se dedica a Data Mining, ha impartido más de 80 conferencias sobre el tema en varios países y se encuentra iniciando proyectos de Big Data.

## Temario

- Introducción.
  - Estado del arte de Data Mining en México.
  - Definición
  - Alcances
- Higienización de datos.
  - Orígenes
  - Clasificación
  - Estrategias de limpieza de datos
- Bases de datos.
  - Características
  - Tipos de variables
  - Cáncer en las bases de datos
- Modelo.
  - Cáncer del Modelo
- Evaluación del modelos
  - Matriz de confusión
- Metodología.
  - CRISP-DM
- Técnicas.

- Árboles de decisión
- Reglas de asociación
- Clustering
  
- Minería de Datos.
  - Detección de la clase
  - Clasificación de las variables
  - Ejercicios usando Weka
  - Proyecto de Minería de Datos.
  - Interpretación.
  - Explicación de un proyecto de inicio a fin
  
- Como vender un proyecto de Minería de Datos.
  - Selección del cliente
  - Discurso
  - Valor agregado
  -
  
- Otras tecnologías relacionadas a Data Mining:
  - Web Mining
  - Video Mining
  - Audio Mining
  
- Relación entre Data Mining y Big Data
  
- Ventajas del Big Data ante Data Warehouse

Costo:

Miembros del IEEE \$5,500.00 + IVA

No Miembros del IEEE \$6,500.00 + IVA

Incluye material del curso, diploma, café y comida.

Cupo mínimo de inscripción: el curso está supeditado a un número mínimo de personas inscritas. Si está interesado en el curso favor de preinscribirse.

Formas de pago: Depósito Bancario.

- No. de Cuenta: **9200-1013-463** de SANTANDER  
SUCURSAL #7978 PABELLON VISTA HERMOSA, RÍO  
MAYO No.1209, COL. VISTA HERMOSA C.P. 62290,  
CUERNAVACA, MORELOS , a nombre de GRUPO IITSE  
TECHNOLOGY SA de CV
- Clave Bancaria Estandarizada (CLABE) depósitos o  
transferencias nacionales: **0145-4092-0010-1346-31.**
- CÓDIGO "ABA" EQUIVALE AL BENEFICIARIO DEL BANCO  
SANTANDER S.A: **BMSXMXMM**
- Clave SWITF depósitos o transferencias  
internacionales: **021000021**

Para obtener cuota de miembro del IEEE deberá anexar copia de su credencial 2013. Una vez realizado el pago favor de enviar una copia del mismo y los siguientes datos con los cuales se realizará la factura.

- Nombre
- Domicilio
- RFC

Fecha y horario:

Lugar: Cuernavaca, Morelos, México. Fecha: Mayo 6 y 7 del 2013  
Horario: 09:00 a 18:00 horas.

Dirigido a:

Todos aquellos profesionales interesados en ampliar sus conocimientos en técnicas avanzadas de análisis de datos. Profesionales que se desempeñen en áreas de Business Intelligence, sistemas de información, Data Warehouse, riesgos y todas las áreas involucradas en el manejo de grandes volúmenes de información; y público en General.

Requisitos:

Conocimientos previos de estadística básica.

Informes:

M.C. Mario Guillén Rodríguez  
Gerencia de Análisis de Redes  
Instituto de Investigaciones Eléctricas  
Cuernavaca, Morelos. México.  
Tel. 01 777-3623811 ext. 7451  
Email: [mguillen@iie.org.mx](mailto:mguillen@iie.org.mx)