

El Departamento de Estadística del ITAM

anuncia la siguiente sesión (No. 201) de

EL SEMINARIO ALEATORIO

que, con el título

**Determinación de la muestra óptima en minería de datos;
metodología y algoritmos**

impartirá

Ángel Kuri

**Departamento de Académico de Computación
ITAM**

Resumen

La minería de datos es una colección de técnicas cuyo propósito final es encontrar elementos de interés en grandes volúmenes de datos. En principio no se conocen las características de los datos que se desea analizar. Sin embargo, desde un punto de vista muy práctico, es razonable suponer que en ellos se encuentran ciertos patrones. Dichos patrones se buscan de manera tal que tipifiquen los datos bajo observación. Llamo "U" al total de datos a analizar.

Se presume que los datos están almacenados en la memoria de una computadora; típicamente en la memoria secundaria. Un problema básico es, obviamente, que la búsqueda a través de muchos millones de bytes se puede hacer lenta y complicada aunque se usen técnicas para hacer búsquedas eficientes y almacenamientos compactos. Una alternativa para eficientar el proceso de búsqueda y análisis consiste en identificar un subconjunto de datos (que llamo "M") tales que conserven todos los patrones de U. Evidentemente el detalle (la granularidad) de los patrones determina las cotas de la submuestra "óptima". Aquí la optimalidad consiste, precisamente, en identificar la muestra que satisfaga que: a) La información es la misma en M que en U y b) Las distribuciones de M sean suficientemente parecidas a las de U. El asunto importante es cómo logramos (a) y (b)

Fecha: Viernes 2 de Octubre

Hora: 13:00 hrs.

Salón: B-2

Lista de seminarios próximos. Visite:

<http://estadistica.itam.mx/seminarios.html>

El Seminario Aleatorio está destinado tanto a profesores como a estudiantes, por lo que el Departamento de Estadística agradece a los profesores que colaboren invitando a sus alumnos a estas sesiones.

María F. Rojano Agraz

ITAM

Depto. Académico de Estadística

Rio Hondo # 1,

Col. Tizapán San Angel

C.P. 01000 México, D. F.

Tel. 5628-4000 ext. 3803

Fax 5628-4086