



NATIONAL AUTONOMOUS UNIVERSITY OF MEXICO
SCHOOL OF ENGINEERING



COURSE SYLLABUS

DATA MINING	2933	9,10	8
Course	Code	Semester	Credits
ELECTRICAL ENGINEERING	COMPUTER ENGINEERING	COMPUTER ENGINEERING	
Division	Department	Undergraduate Program	

Course:

Hours /week:

Hours / Semester:

Compulsory

Lecture

4.0

Lecture

64.0

Elective

Practical

0.0

Practical

0.0

Total

4.0

Total

64.0

Mode: Lecture-based

Prerequisite course: None

Subsequent course: None

Course Objective(s)

The student will evaluate the principles of the machine learning approach. to design and implement various data mining algorithms.

Course Topics

No.	TITLE	HOURS
1.	Introduction to Knowledge Discovery in Databases (KDD) and Data Mining	8.0
2.	Data Mining Architecture	12.0
3.	Data Mining Process	18.0
4.	Evaluation	14.0
5.	Data Mining Application	12.0
		64.0
Practical Activities		0.0
		64.0

1. Introduction to Knowledge Discovery in Databases (KDD) and Data Mining

Objective: The student will understand the importance of data mining within the process of knowledge discovery in data.

knowledge in data.

Content:

1.1 Architecture of the Knowledge Discovery Process in Database (KDD).

1.1.1 Understanding the Business Domain.

1.1.2 Identification of Relevant Data.

1.1.3 Data Cleaning.

1.1.4 Data Transformation.

1.1.5 Identification of Data Mining Tasks.

1.1.6 Implementation of Data Mining Algorithms.

1.1.7 Interpretation and Evaluation of Data.

1.2 The role of Data Mining within KDD.

2. Data Mining Architecture

Objective: The student will interpret the general architecture of the data mining process by analyzing its elements.

Content:

2.1 Repositories.

2.2 Data Servers.

2.3 Knowledge Database.

2.4 Data Mining Process.

2.5 Evaluation.

3. Data Mining Process

Objective: The student will select appropriate models, methods and algorithms for data mining.

Content:

3.1 Supervised and Unsupervised Learning.

3.2 Predictive Models.

3.3 Descriptive Models.

3.4 Data Mining Methods and Algorithms.

3.4.1 Classification.

3.4.2 Regression.

3.4.3 Clustering.

3.4.4 Summarization.

3.4.5 Dependency Models, Correlation.

3.4.6 Association Rules.

3.4.7 Detection of Changes and Deviations.

3.5 Data Mining Process.

- 3.5.1** Identify the Predictive or Descriptive Model.
- 3.5.2** Identify the Method.
- 3.5.3** Identify the Algorithm.
- 3.5.4** Generate the Model.
- 3.5.5** Validate the Model.
- 3.5.6** Improve the Model.

4. Evaluation

Objective: The student will identify which evaluation method to use in a particular problem and how it works in order to establish a comparison between them.

Content:

- 4.1** Training and Verification.
- 4.2** Performance Prediction.
- 4.3** Cross-validation.
- 4.4** Comparison of Data Mining Schemes.
- 4.5** Probability Prediction.
- 4.6** Cost Estimation.
- 4.7** Numerical Prediction Evaluation.
- 4.8** The Principle of Minimum Length Description (MLD).
- 4.9** Application of MLD to Clustering.

5. Application in Data Mining

Objective: The student will identify which evaluation method to use in a particular problem and how it works in order to establish a comparison between them.

Content:

- 5.1** Rudimentary Rule Inference.
- 5.2** Statistical Modeling.
- 5.3** Decision Tree Construction.
- 5.4** Coverage Algorithms: Construction Rules.
- 5.5** Association Rules.
- 5.6** Linear Models.
- 5.7** Occurrence-Based Learning.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA



PROGRAMA DE ESTUDIO

MINERÍA DE DATOS

2933

9,10

8

Asignatura

Clave

Semestre

Créditos

INGENIERÍA ELÉCTRICA

INGENIERÍA
EN COMPUTACIÓN

INGENIERÍA
EN COMPUTACIÓN

División

Departamento

Licenciatura

Asignatura:

Obligatoria ☐

Optativa ☒

Horas/semana:

Teóricas

Prácticas

Total

Horas/semestre:

Teóricas

Prácticas

Total

Modalidad: Curso teórico

Seriación obligatoria antecedente: Ninguna

Seriación obligatoria consecuente: Ninguna

Objetivo(s) del curso:

El alumno evaluará los principios del enfoque del aprendizaje en máquinas para diseñar e implementar diversos algoritmos de minería de datos.

Temario

NÚM.	NOMBRE	HORAS
1.	Introducción al descubrimiento del conocimiento en base de datos (KDD) y la minería de datos	8.0
2.	Arquitectura de minería de datos	12.0
3.	El proceso de minería de datos	18.0
4.	Evaluación	14.0
5.	Aplicación en la minería de datos	12.0
		64.0
	Actividades prácticas	0.0
	Total	64.0

1 Introducción al descubrimiento del conocimiento en base de datos (KDD) y la minería de datos

Objetivo: El alumno comprenderá la importancia de la minería de datos dentro del proceso de descubrimiento del conocimiento en datos.

Contenido:

1.1 Arquitectura del proceso de descubrimiento del conocimiento en base de datos (KDD).

- 1.1.1 Entendimiento del dominio del negocio.
- 1.1.2 Identificación de los datos relevantes.
- 1.1.3 Limpieza de datos.
- 1.1.4 Transformación de datos.
- 1.1.5 Identificación de tareas de Minería de datos.
- 1.1.6 Implementación de algoritmos de minería de datos.
- 1.1.7 Interpretación y evaluación de datos.

1.2 El papel de la minería de datos dentro del KDD.

2 Arquitectura de minería de datos

Objetivo: El alumno interpretará la arquitectura general del proceso de minería de datos mediante el análisis de sus elementos.

Contenido:

- 2.1 Repositorios.
- 2.2 Servidores de datos.
- 2.3 Base de datos de conocimientos.
- 2.4 Proceso de minería de datos.
- 2.5 Evaluación.

3 El proceso de minería de datos

Objetivo: El alumno seleccionará los modelos, métodos y algoritmos apropiados para el minado de datos.

Contenido:

- 3.1 Aprendizaje supervisado y no supervisado.
- 3.2 Modelos predictivos.
- 3.3 Modelos descriptivos.
- 3.4 Métodos y algoritmos de minería de datos.
 - 3.4.1 Clasificación.
 - 3.4.2 Regresión.
 - 3.4.3 Agrupación.
 - 3.4.4 Sumarización.
 - 3.4.5 Modelos de dependencia, correlación.
 - 3.4.6 Reglas de asociación.
 - 3.4.7 Detección de cambios y desviaciones.
- 3.5 Proceso de minería de datos.
 - 3.5.1 Identificar el modelo predictivo o descriptivo.
 - 3.5.2 Identificar el método.
 - 3.5.3 Identificar el algoritmo.
 - 3.5.4 Generar el modelo.
 - 3.5.5 Validar el modelo.
 - 3.5.6 Mejorar el modelo.

4 Evaluación

Objetivo: El alumno identificará qué método de evaluación utilizar en un problema en particular y su funcionamiento para poder establecer una comparación entre ellos.

Contenido:

- 4.1 Entrenamiento y verificación.
- 4.2 Predicción del rendimiento.
- 4.3 Validación cruzada.
- 4.4 Comparación de esquemas de minería de datos.
- 4.5 Predicción de probabilidades.
- 4.6 Conteo del costo.
- 4.7 Evaluación de predicción numérica.
- 4.8 El principio de la descripción de longitud mínima (LDM).
- 4.9 Aplicación de la LDM a agrupamientos.

5 Aplicación en la minería de datos

Objetivo: El alumno experimentará con los algoritmos de aprendizaje pertinentes para procesar los datos representados en un formato dado.

Contenido:

- 5.1 Inferencia de reglas rudimentarias.
- 5.2 Modelación estadística.
- 5.3 Construcción de árboles de decisión.
- 5.4 Algoritmos de cobertura: Reglas de construcción.
- 5.5 Reglas de asociación.
- 5.6 Modelos lineales.
- 5.7 Aprendizaje basado en ocurrencias.

Bibliografía básica

Temas para los que se recomienda:

BRITOS, Paola, HOSSIAN, Alejandro, et al.

Minería de datos basada en sistemas inteligentes

Nueva Librería, 2005

Todos

HERNÁNDEZ, Orallo, RAMÍREZ, José

Introducción a la minería de datos

México

Pearson/Prentice Hall, 2003

Todos

MITCHELL, Tom

Machine Learning

McGraw Hill, 1997

Todos

WITTEN, Ian, FRANK, Eibe

Data Mining

San Diego

Morgan Kaufmann Publishers, 2001

Todos

Bibliografía complementaria**Temas para los que se recomienda:**

AGGARWAL, Charu

Data Mining: The Textbook

New York

Springer, 2015

Todos

HAN, Jiawei, KAMBER, Micheline, PEI, Jian

Data Mining: Concepts and Techniques

3th Edition

Boston

Morgan Kaufmann, 2011

Todos

ZAKI, Mohammed J., MEIRA, Wagner Jr

Data Mining and Analysis: Fundamental Concepts and Algorithms USA

Cambridge University Press 2014

Todos

Sugerencias didácticas

Exposición oral	<input checked="" type="checkbox"/>
Exposición audiovisual	<input type="checkbox"/>
Ejercicios dentro de clase	<input checked="" type="checkbox"/>
Ejercicios fuera del aula	<input type="checkbox"/>
Seminarios	<input type="checkbox"/>
Uso de software especializado	<input type="checkbox"/>
Uso de plataformas educativas	<input type="checkbox"/>

Lecturas obligatorias	<input checked="" type="checkbox"/>
Trabajos de investigación	<input checked="" type="checkbox"/>
Prácticas de taller o laboratorio	<input checked="" type="checkbox"/>
Prácticas de campo	<input type="checkbox"/>
Búsqueda especializada en internet	<input type="checkbox"/>
Uso de redes sociales con fines académicos	<input type="checkbox"/>

Forma de evaluar

Exámenes parciales	<input checked="" type="checkbox"/>
Exámenes finales	<input checked="" type="checkbox"/>
Trabajos y tareas fuera del aula	<input checked="" type="checkbox"/>

Participación en clase	<input checked="" type="checkbox"/>
Asistencia a prácticas	<input checked="" type="checkbox"/>

Perfil profesiográfico de quienes pueden impartir la asignatura

Licenciatura en Ingeniería en Computación, Ciencias de Computación, Matemáticas Aplicadas o una carrera similar. Deseable haber realizado estudios de posgrado, contar con conocimientos y experiencia en el área de las bases de datos e ingeniería de software especialidad en minería de datos, contar con experiencia docente o haber participado en cursos o seminario de iniciación en la práctica docente.