



Rosario, 21 de diciembre de 2022.-

VISTO El expediente I.D. N° 8141899 presentado por el Consejo Departamental de Ingeniería en Sistemas de Información, relacionado con el programa analítico de la asignatura electiva "Minería de Datos", de la carrera Ingeniería en Sistemas de Información, y

**CONSIDERANDO**

Que los objetivos y contenidos del mismo se ajustan a la reglamentación vigente.

Que dicho programa cuenta con el aval del respectivo Consejo Departamental.

Que la Comisión de Enseñanza evaluó la presentación y aconsejó su aprobación.

Por ello y atento a las atribuciones otorgadas por el artículo 85° del Estatuto Universitario.

**EL CONSEJO DIRECTIVO DE LA FACULTAD REGIONAL ROSARIO  
DE LA UNIVERSIDAD TECNOLÓGICA NACIONAL**

**RESUELVE:**

ARTÍCULO 1°.- Aprobar el programa analítico de la asignatura electiva "Minería de Datos" para el Quinto Nivel de la carrera Ingeniería en Sistemas de Información – Plan 2008, que se agrega como Anexo I de la presente resolución. A partir del ciclo lectivo 2023.

ARTÍCULO 2°.- Establecer que la misma tendrá validez durante cuatro ciclos lectivos consecutivos, según la Ordenanza N° 1383 – Lineamientos para la implementación de asignaturas electivas para las carreras de grado en el ámbito de la Universidad.

ARTÍCULO 3°.- Regístrese. Comuníquese. Cumplido, archívese.

**RESOLUCIÓN N° 671**

UTN
FRRo
C.D.
S.R.

Ing. Rubén Fernando CICCARELLI  
Decano

Ing. Antonio Luis MUIÑOS  
Secretario Académico

# RESOLUCION N° 671 ANEXO N° 1



UNIVERSIDAD TECNOLÓGICA NACIONAL  
FACULTAD REGIONAL ROSARIO  
DEPARTAMENTO DE INGENIERÍA EN SISTEMAS DE INFORMACION

## Programa analítico de asignatura electiva Minería de Datos

<b>Carrera:</b>	Ingeniería en Sistemas de Información				
<b>Departamento:</b>	Ingeniería en Sistemas de Información				
<b>Titulación<sup>1</sup>:</b>	X	Ingeniería en Sistemas de Información		Analista universitario de Sistemas	
<b>Plan de Estudio:</b>	2008 - ordenanza 1150			<b>Área<sup>2</sup>:</b>	Gestión Ingeniería
<b>Dictado:</b>	Anual	X	Cuatrimstral	<b>Nivel:</b>	5
				<b>Electiva:</b>	Si
<b>Carga horaria Semanal:</b>	6 hs		<b>Carga horaria total de la asignatura:</b>		96 hs
<b>Fecha de Confección<sup>3</sup>:</b>	06/04/2021			<b>Versión<sup>4</sup></b>	9.99

<b>Fundamentación de la asignatura:<sup>5</sup></b>	La asignatura pertenece al Area Ingeniería cuyo objetivo es formar al alumno en la gestión y toma de decisiones para su aplicación al desarrollo de sistemas predictivos, de asignación, utilización y distribución de recursos
<b>Objetivos Generales<sup>6</sup>:</b>	<ul style="list-style-type: none"> <li>- Distinguir estrategias para gestionar un proyecto de ciencia de datos.</li> <li>- Aplicar estrategias de extracción, visualización y transformación de datos en forma de patrones útiles y aplicables en el desarrollo de sistemas inteligentes.</li> <li>- Crear modelos que resuelvan situaciones problemáticas particulares en el contexto del negocio.</li> <li>- Comprender la importancia de la Extracción de Conocimiento para la toma de decisiones en cualquier organización</li> <li>- Conocer las distintas metodologías que existen a la hora de desarrollar un proyecto de ciencia de datos</li> <li>- Formar al alumno en la aplicación de los distintos modelos comprendidos en la ciencia de datos justificando su aplicación</li> <li>- Analizar información proveniente del BIG DATA y distintas herramientas para su realización</li> </ul>

### Programa de contenido analítico

<b>Unidad temática N°: 1</b>
<b>Eje Conceptual:</b> Introducción a la ciencia de datos, metodologías de desarrollo, importancia para la toma de decisiones
<b>Objetivo:</b> Introducir al alumno en la nueva disciplina. Distinguir estrategias para gestionar un proyecto

<sup>1</sup> Indique los títulos de la carrera para los que se propone el programa analítico. Márquelos con una cruz.

<sup>2</sup> Área a la que pertenece la asignatura

<sup>3</sup> refiere a la fecha en que se confecciona o desarrolla la versión

<sup>4</sup> Si el programa no es la primera vez que se entrega se produce un cambio en el número de versión cambio. Si el cambio es significativo cambia el entero sino los dígitos después del punto.

<sup>5</sup> Importancia para la formación profesional en función del perfil del egresado

<sup>6</sup> Objetivos generales que justifican la inclusión de la asignatura.

1/5



de ciencia de datos.

**Temas:**

Introducción, Definición. Tipos de datos y modelos. Extracción del conocimiento. Modelos, tareas y técnicas.

Ciencia de datos como sistema de soporte a la toma de decisiones.

Metodologías utilizadas en la Ciencia de Datos. Fases. Comparación de metodologías: KDD, CRISP-DM, SEMMA, P3TQ.

Fuentes de datos - Almacenes de datos. Procesamiento de grandes volúmenes de datos.

Análisis exploratorio de los datos. Visualización de datos. Extracción y Transformación de datos. ACP (Análisis de Componentes Principales). Análisis Factorial.

**Unidad temática N°: 2**

**Eje Conceptual:** Técnicas aplicadas en Ciencia de Datos

**Objetivo:** Crear modelos que resuelvan situaciones problemáticas particulares en el contexto del negocio. Estudiar las distintas técnicas aplicables. Entender la utilización de las mismas en problemas concretos. Ingeniería de Datos.

**Temas:**

Clusters. Fundamentos. Distancias. Algoritmos de agrupamiento y clasificación.

Reglas de asociación.

Regresión Logística

Árboles de decisión y sistemas de reglas

Análisis discriminante. Caso de dos poblaciones. Discriminador lineal. Poblaciones normales. Probabilidad de error

Combinación de modelos.

Ingeniería de Datos. BIG DATA Analítico.

Minería de Textos

Otras técnicas complementarias que contribuyan al entendimiento de la extracción de conocimiento.

**Unidad temática N°: 3**

**Eje Conceptual:** Proyectos de Ciencia de Datos: evaluación de modelos, análisis de factibilidad, implementación de la solución.



**Objetivo:** Aplicar estrategias de extracción, visualización y transformación de datos en forma de patrones útiles y aplicables en el desarrollo de sistemas inteligentes. Comprender la importancia de considerar a la ciencia de datos como un proceso y la implicancia de la misma en el proceso decisorio de la empresa. Entender las fases de la implementación de un programa de minería y la evaluación de modelos candidatos.

**Temas:**

Evaluación de modelos. Evaluación de clasificadores, evaluación de hipótesis basada en precisión, evaluación de hipótesis basada en costos. Comparación de técnicas de aprendizaje. Evaluación de modelos de agrupamiento. Evaluación de reglas de asociación. Otros criterios de evaluación

Interpretación, difusión y uso.

Implantación de la solución. Determinar necesidades y objetivos de negocio. Implantación progresiva. Integración con las herramientas y proyectos de la organización. Recursos necesarios.

Actualización y revisión de modelos.

## Bibliografía<sup>7</sup>

**Obligatoria o básica:**

Título	Autor/es	Editorial	Año de Edición
Ciencia de Datos	García – Molina – Berlanga – Patricio – Bustamante – Padilla.	Alfaomega	2018
Tratamiento de Datos con R, Statistics y SPSS	Gonzalez – Liste – Felpeto	Ediciones Díaz de Santos	2011
Introducción a la minería de datos	Hernández Orallo, Ramírez Quintana y Ferri Ramírez	Pearson Prentice Hall	2004
Análisis de datos multivariantes.	PEÑA, Daniel	Mc Graw Hill	2002

**Complementaria:**

Título	Autor/es	Editorial	Año de Edición
Probabilidad y estadística aplicadas a la ingeniería.	Montgomery, D C y Runger, G C.	Editorial Mc Graw Hill – México	1996
Técnicas de Análisis Multivariante de Datos. Aplicaciones con SPSS	Pérez, César	Pearson. Prentice Hall, España	2004.
Métodos multivariados aplicados al análisis de datos	Dallas Johnson	International Thomson Editores	2000

<sup>7</sup> Para textos: citar autor, título, ciudad, editorial, año. Para revistas: citar autor, título del artículo, nombre de la revista, n°, lugar, edición, año, páginas., Para sitios web: dirección de la página.



## Propuesta Pedagógica

Estrategias utilizadas en el dictado de la asignatura:

- Clases teórico-prácticas siguiendo los contenidos y secuencia especificados en el programa.
- Vinculación permanente de los contenidos teóricos con casos prácticos, para reforzar la comprensión y el proceso de aprendizaje.
- Práctica en laboratorio con el uso de herramientas de software libres y propietarias:
  - Pentaho: software libre usado para procesos de Extracción Transformación y Carga.
  - IBM SPSS Statistics: software para el análisis estadístico
  - R / Rapid Miner: software libre para ciencia de datos
  - IBM SPSS Modeler: software para minería de IBM
  - Otras herramientas de software libre disponibles en el mercado que cubran los temas dictados en la materia.
- Contenido audiovisual para mejorar la comprensión de los temas.
- La cátedra cuenta con un sitio web, en el cual se pone a disposición del alumno todo el material didáctico. En la misma se publican también las novedades relacionadas a trabajos prácticos y exámenes.
- Al finalizar cada tema del curso, se proponen trabajos prácticos cuyo objetivo es fijar y aplicar los contenidos vistos en clase.
- Todos los trabajos prácticos se realizan en grupos, para promover capacidades de trabajo en equipo.

En el transcurso del cuatrimestre el alumno deberá realizar un trabajo práctico integrador, cuyo objetivo es vincular e integrar todos los temas de la materia. Los docentes de la cátedra monitorean y corrigen periódicamente los avances del mismo, efectuándose la entrega completa al finalizar el cuatrimestre.

Anualmente, la cátedra organiza "Jornadas de Ciencia de Datos", donde empresas referentes en la disciplina exponen casos reales de aplicación reforzando los temas tratados en el curso y la vinculación Universidad-Empresa.

**Evaluación:**

Trabajos prácticos de Ciencia de Datos con defensa  
Exposición de trabajos de investigación en clase abierta  
Parciales sobre temas dictados.  
Final: Examen final teórico

El/los trabajos prácticos de regularización pueden recuperarse en el primer llamado de las mesas de julio/agosto, noviembre/diciembre y febrero/marzo inmediatamente posteriores al cursado.

### Aprobación No DIRECTA

*Regularidad:*

*Trabajos Prácticos específicos cubriendo los temas dictados durante el cursado.*

*Entregas de trabajos prácticos exigidos en tiempo y forma.*

*Exposición de trabajos de investigación en clase abierta.*

*Examen final:*

*Defensa teórica del Trabajo Práctico Integrador y examen oral de temas dictados en la materia.*



### Aprobación DIRECTA

*Condiciones de aprobación directa:*

*Entregas de trabajos prácticos exigidos en tiempo y forma.*

*Defensa de los trabajos prácticos.*

*Exposiciones orales de temas investigados.*

*Parcial de los contenidos de la materia.*

### **Asignaturas Correlativas del plan<sup>8</sup>**

<b>Asignaturas regulares para el cursado:</b>	Simulación
<b>Asignaturas aprobadas para el cursado:</b>	Probabilidad y Estadística Gestión de Datos
<b>Asignaturas aprobadas para rendir:</b>	

### **Justificación de correlatividades**

Se necesitan conocimientos de base de datos y acceso a bases de datos mediante lenguaje SQL para los procesos de extracción y transformación de datos.

Se necesitan conocimientos de modelos estadísticos usados para la toma de decisiones.

### **Asignaturas Equivalentes respecto del plan anterior<sup>9</sup>**

<b>Asignatura/s equivalente respecto del plan anterior:</b>	
---	--

<sup>8</sup> Todos los cuadros deben estar completados.

<sup>9</sup> Consignar asignaturas que se pueden otorgar como equivalentes para las posibles solicitudes de cambio de plan.