

**Profesor: Ph.D. David Díaz S.**  
**E-mail Profesor: [ddiaz@unegocios.cl](mailto:ddiaz@unegocios.cl)**  
**E-mail Tareas: [BI-DM@unegocios.cl](mailto:BI-DM@unegocios.cl)**

---

## **PRESENTACIÓN DEL CURSO**

El objetivo de éste curso es la revisión avanzada de las tecnologías digitales y sistemas de Inteligencia de Negocios y demostrar su importancia a través del uso de casos de estudio y tareas. El curso definirá qué se entiende por Business Intelligence (BI) y sus componentes tecnológicos y cómo pueden ser aplicados en las problemáticas de diferentes de industrias. En particular, se entenderá BI como la agrupación de grandes cantidades de datos y su análisis en orden de encontrar patrones interesantes que puedan ayudar a la toma de decisiones de negocios y a la realización de las funciones empresariales. BI considera una amplia gama de tecnologías que permiten a los usuarios el recolectar, almacenar, acceder y analizar datos de manera de mejorar el proceso de administración de información centrado en el cliente. BI consiste de tres componentes tecnológicos integrados: Data Warehousing, Data Mining y Reporte Analítico. El curso investigará principalmente los últimos dos componentes mencionados.

---

## **OBJETIVOS DEL CURSO**

Este curso está diseñado para que los alumnos puedan:

1. Entender los principios de Business Intelligence y sus implicancias para la innovación en los negocios.
  2. Entender los diferentes componentes tecnológicos de BI, a saber, data, text, web mining, data warehousing y reporte analítico.
  3. Evaluar el impacto de las diferentes tecnologías de BI para la innovación en los negocios.
  4. Entender cómo las tecnologías BI son utilizadas en aplicaciones de negocios: análisis de patrones, segmentación de clientes, market basket analysis, pronóstico, modelamiento predictivo y sentiment analysis, minería de textos, y análisis de redes sociales.
  5. Evaluar críticamente la evolución de dichas tecnologías y sus impactos.
  6. Demostrar la habilidad de usar tecnologías BI para resolver problemas de negocios.
-

## **METODOLOGÍA**

La metodología considera una combinación de clases teóricas, casos de estudio, y uso de software de BI. Luego de cada clase expositiva habrá un laboratorio aplicado. Los alumnos deberán seguir paso a paso las instrucciones entregadas en una guía preparada para tales fines, y desarrollar un ejercicio de data mining, además de entregar en la clase siguiente el desarrollo de ejercicios propuestos.

Por la naturaleza del curso, los tópicos serán tocados en el nivel de profundidad acorde al tiempo disponible y al nivel de preparación previa de los alumnos. Por esto, se considera de suma importancia, la participación del alumno para tener una adecuada retroalimentación, y también, la dedicación del estudiante en cuanto a investigar y profundizar en aquellos temas donde sienta mayor debilidad o aquellos que despierten mayor interés. Lecturas complementarias serán entregadas con la debida anticipación.

## **EVALUACIÓN**

La evaluación del curso considera los siguientes componentes:

### **Laboratorios y Tareas [L]**

Se llevarán a cabo al menos 5 laboratorios, incluyendo tópicos de preparación de datos, minería de datos, minería de texto, y análisis de redes sociales. Al final de cada Laboratorio será entregado una tarea o ejercicio propuesto, el que deberá ser entregado al mail de tareas una semana después. **Existirá una penalización de 0.5 décimas en la nota por cada media hora de atraso en la entrega.**

### **Prueba Solemne [S1]**

Se realizará una Prueba Solemne de carácter individual, en la que se controlen tanto los conceptos teóricos, como prácticos de las materias comprendidas en el curso.

### **Trabajo Grupal 1 [NT1]**

Se realizará un trabajo en grupos de 5-6 alumnos. El objetivo de este Proyecto es que los alumnos puedan encontrar grupos de clientes en una base de datos real, realizando un trabajo de segmentación con base en técnicas de data mining. La nota será calculada en base a la capacidad de toma de decisiones de negocios, calidad de las recomendaciones comerciales entregadas y a una presentación grupal que se llevará a cabo en fecha por definir. **Existirá una penalización de 0.5 décimas en la nota por cada media hora de atraso en la entrega.**

### **Trabajo Grupal 2 [NT2]**

Se realizará un trabajo en grupos de 5-6 alumnos. Éste consistirá en desarrollar un modelo de pronóstico o predictivo. Específicamente, se trabajará en grupos utilizando una

metodología colaborativa, donde cada uno de los alumnos cumplirá un rol en el Proyecto. La nota será calculada en base a la capacidad de toma de decisiones de negocios, calidad de las recomendaciones comerciales entregadas, al poder predictivo final alcanzado, a la calidad del reporte entregado y a una presentación grupal que se llevará a cabo en fecha por definir. **Existirá una penalización de 0.5 décimas en la nota por cada media hora de atraso en la entrega.**

## **CONDICIONES DE APROBACIÓN**

Las notas finales por actividad se calcularán de la siguiente manera:

<b>NL</b>	<b>= Nota de Laboratorios y Tareas (promedio simple)</b>
<b>S1</b>	<b>= Prueba Solemne</b>
<b>NT1</b>	<b>= Nota Trabajo 1. Segmentación de Clientes.</b>
<b>NT2</b>	<b>= Nota Trabajo Final</b>

Al final del semestre, dependiendo de la cantidad de evaluaciones realizadas, se determinará la eliminación de un porcentaje de peores notas para cada alumno. Esto será de criterio de la coordinación cátedra.

$$\text{Nota Final} = 30\% \text{NL} + 20\% \text{S1} + 25\% \text{NT1} + 25\% \text{NT2}$$

## **OBSERVACIONES IMPORTANTES**

En el caso de comportamientos académicos indebidos, en particular, copia o fraude en una tarea, control o proyecto, el alumno será calificado con nota 1,0 en la actividad respectiva. Estas notas no podrán ser eliminadas. No obstante lo anterior, si la falta lo amerita el alumno podrá ser sancionado con la reprobación del curso, previa consulta a los canales pertinentes.

El medio de información oficial del curso es el sitio web, por lo que los estudiantes están obligados a consultarlo frecuentemente (al menos 3 veces por semana). Las informaciones más importantes se publicarán en el apartado de Noticias/Avisos.

En las tareas grupales se espera una participación de todos los integrantes, ya que en caso de una exposición oral, la evaluación de un integrante del grupo podría reflejar la nota de todo el grupo.

Las tareas que no cumplan con las pautas de entrega establecidas, se calificarán con 1,0, al igual que aquellas que sean iguales o similares a otras, pues serán consideradas copia.

**Inasistencias a Actividad calificada con nota:**

En caso de inasistencia a cualquier actividad evaluativa, el alumno será calificado con nota mínima (1.0) en dicha evaluación y deberá justificar su inasistencia siguiendo los procedimientos ya establecidos por la Facultad.

---

## **MATERIAL DEL CURSO: TEXTOS Y ARTÍCULOS**

Como textos guías se utilizarán los siguientes libros:

“Data Mining Techniques for Marketing, Sales and Customer Relationship Management”, de Michael J.A. Berry and Gordon S. Linoff, Editorial John Wiley, 2004.

Como lecturas complementarias se recomiendan:

“Data Mining: Practical Machine Learning Tools and Techniques” de Ian H. Witten, Eibe Frank, Segunda Edición.

“Data Mining: introductory and Advanced Topics” de Margaret H. Durham, Prentice Hall, 2002.

“Database Processing”, de Kroenke, D., Editorial Pearson - Prentice Hall. Edición Internacional, 2006

“The Data Warehouse Lifecycle Toolkit” de Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker. Editorial John Wiley and Sons, 2007

Este curso está inspirado en el curso de Andrew NG dictado en Standford University. Más info en <http://cs229.stanford.edu/>

## **CURRÍCULUM DEL PROFESOR**

### **David Díaz Solís**

David Díaz, Ingeniero Comercial, Universidad de Chile. Magíster en Finanzas, Universidad de Chile. Doctor of Philosophy (Ph.D.) in Business Administration (Business Intelligence), The University of Manchester. Díaz es académico jornada completa del Departamento de Administración de la Universidad de Chile, desarrollando actividades de docencia en pregrado, MBA y Magíster, investigación académica y extensión. El Dr. David Díaz cuenta con una destacada trayectoria de publicaciones académicas relacionadas con la utilización de técnicas de modelamiento de datos financieros, siendo su principal línea de investigación los Sistemas de Información y la aplicación de la filosofía Business Intelligence en problemas financieros y de negocios. Asimismo, ha desarrollado asesorías a gobiernos y empresas nacionales y extranjeras.

Con respecto a consultorías, ha desarrollado múltiples proyectos de valorización de empresas, predicción de impagos, evaluación de proyectos, modelamiento de tasas de interés, diseño e implementación de sistemas de información de negocios, entre otras. Sus investigaciones han sido publicadas en prestigiosas revistas académicas de América Latina, Estados Unidos y Europa en las áreas temáticas de Modelación de Precios Accionarios, Comportamiento de Tasas de Interés, y

Modelos de Quiebra de Bancos y Empresas Financieras, Sistemas Expertos, Detección de Fraude en Bolsas de Valores, Utilización de Inteligencia Artificial y Data Mining, entre otros.

**PROGRAMACIÓN Y CONTENIDOS**

Fecha	Contenidos	Aplicaciones	Ejemplo o Lab
<b>Módulo 1</b>	<b>Introducción, Aplicaciones Generales en Data Mining y repaso</b>		
<b>Módulo 2</b>	<b>Metodología CRISP-DM</b>	<b>Metodología de Exploración de Datos</b>	
<b>Módulo 3</b>	<b>Entendimiento del Negocio y preparación de los Datos.</b>	<b>Operaciones de Bases de datos Financieros, Unión, Intersección, filtro, calidad de datos, limpieza, etc.</b>	<b>Ayudantía 1 Lab 1 importación y pre-procesamiento de datos y Lab 2 pre-procesamiento de datos y visualización</b>
<b>Módulo 4</b>	<b>Modelamiento: Reglas de Asociación</b>	<b>Detección de patrones simples y en series de tiempo.</b>	<b>Lab Reglas de Asociación</b>
<b>Módulo 5</b>	<b>Modelamiento: Análisis de Conglomerados</b>	<b>Detección de similitudes, grupos y clústers en datos. Detección de outliers. Uso en segmentación de mercados, customer profiling.</b>	<b>Lab Clustering</b>
<b>Módulo 6</b>	<b>Intro problemas de clasificación</b>		
<b>Módulo 7</b>	<b>Modelamiento: Clasificación. Árboles de Decisión</b>	<b>Predicción de Churn Rate, Impago</b>	<b>Lab Análisis de Clasificación 1</b>
<b>Módulo 8</b>	<b>Modelamiento con muestras no Balanceadas</b>	<b>Over sampling, sub-sampling, precision, recall. Cost sensitive learning. Validación cruzada, bootstrapping, gráficos de ROC y LIFT</b>	<b>Lab Análisis de Clasificación 2</b>
<b>Módulo 9 (opcional)</b>	<b>Analytics en Contexto E-commerce</b>	<b>Impacto del BI en modelos de negocios y Economía</b>	
<b>Módulo 10 (opcional)</b>	<b>Analisis de Series de Tiempo 1</b>	<b>Modelos de regresión con y sin Estacionalidad</b>	<b>Proyección Millas Aereas, Nro Pasajeros de Tren</b>

<b>Módulo 11</b> <b>(opcional)</b>	<b>Análisis de Series de Tiempo 2</b>	<b>Modelos Autoregresivos con y sin estacionalidad</b>	<b>Proyección Ventas de Shampoo</b>
<b>Módulo 12</b> <b>(opcional)</b>	<b>In-Bound Marketing</b>	<b>Caso Hubspot</b>	
<b>Módulo 13</b> <b>(opcional)</b>	<b>Minería de Texto. Parte 1.</b>	<b>Introducción. Técnicas estadísticas. Lectura de Noticias Financieras automatizada. Análisis de Reclamos.</b>	
<b>Módulo 14</b> <b>(opcional)</b>	<b>Minería de Texto. Parte 2.</b>	<b>Introducción. Técnicas Lingüísticas. Caso Estudio. Clasificación de Noticias Financieras automatizada. Análisis de Reclamos.</b>	
<b>Módulo 15</b> <b>(opcional)</b>	<b>Social Network Analysis</b>	<b>Definición y Cálculo de Métricas. Caso aplicado a Directorios Financieros.</b>	
<b>Módulo 16</b> <b>(opcional)</b>	<b>Webmining</b>	<b>Lectura y clasificación noticias financieras, twitts</b>	
<b>Presentaciones Finales Copa</b>			
<b>Módulo 17</b>	<b>Repaso y Cierre</b>		