FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS UNIVERSIDAD DE CHILE

PROGRAMA DE CURSO

Código	Nombre			
IN 5526	WEB INTELLIGENCE			
Nombre en Inglés				
WEB INTELLIGENCE				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas Trabajo Personal
6	10	3,0	1,5	5,5
Requisitos			Carácter del Curso	
IN3501 Tecnologías de Información y Comunicaciones para la Gestión IN3702 Investigación de Operaciones			Electivo de la carrera de Ingeniería Civil Industrial	

Competencias a las que tributa el curso

Competencias específicas

CE2: Concebir soluciones a los problemas que surgen en las organizaciones, utilizando los conocimientos provenientes de la gestión de operaciones, tecnologías de información y comunicaciones, finanzas, economía y marketing.

CE6: Desarrollar habilidades para liderar equipos de trabajo manejando las relaciones interpersonales.

Competencias genéricas:

CG1: Comunicar ideas y resultados de trabajos profesionales o de investigación, en forma escrita y oral, tanto en español como en inglés.

CG5: Gestionar su auto-aprendizaje en el desarrollo del conocimiento de su profesión, adaptándose a los cambios del entorno.

Propósito del curso

El curso IN 5526 tiene como propósito que el estudiante sea capaz de desarrollar aplicaciones de Web Intelligence que permitan extraer información y conocimiento desde los datos web, para luego ser usados en el desarrollo de nuevas tecnologías web para el futuro, como, por ejemplo, en sitios web adaptativos.

La estrategia metodológica a utilizar es activo-participativa permitiendo que el estudiante desarrolle destrezas y conocimientos en conceptos claves de Web Intelligence y su importancia en los negocios. Esta estrategia incluye resolución de casos simples, de forma individual y colectiva, y también el desarrollo de un proyecto de implementación dividido en tareas, que incliyen metodologías KDD para extracción de conocimiento desde datos y algoritmos y modelos para extracción de patrones, entre otros temas relevantes. En el desarrollo del curso, se busca fomentar el trabajo en equipo, gestionando el autoaprendizaje de los estudiantes, por medio de diversas actividades como resolución de ejercicios, método de proyecto, entre otras. El docente acompaña el proceso, resolviendo dudas, corrigiendo, proponiendo perspectivas de trabajo.

Resultados de Aprendizaje
CE2-CG5–**RA1**: **Analiza los conceptos claves de Web Intelligence, considerando la principal metodología KDD (Knowledge Discovery in Data Bases), y puede explicar su importancia en los negocios.** CE2–CE6-CG5–**RA2: Desarrolla un proyecto de levantamiento e implementación de un proyecto de Web Mining manejando la componente técnica y a la vez teniendo la vista de negocio** CE2–CE6–CG1–**RA3: Diseña con su grupo una solución informática a un desafío real y contingente en materia de obtención de conocimiento útil para un negocio y expone su propuesta de manera clara y fundamentada, utilizando un lenguaje técnico adecuado al contexto académico.**

Metodología Docente	Evaluación General
La metodología de trabajo será activo–participativa y su carácter es teórico–práctico. Se trabajará: - Clases expositivas con estructura de INICIO - DESARROLLO - CIERRE. - Trabajo por proyecto - Análisis de casos simples - Resolución de casos simples	La evaluación considera: - 2 Controles - 3 Tareas - Examen

Unidades Temáticas

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
1	RA1	Introducción a Web Intelligence	1
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
1.1. Conceptos Iniciales 1.2. El proceso KDD 1.3. Minería Web		El estudiante: 1. Maneja distinciones como World Wide Web y E-bussiness, entre otros conceptos. 2. Conoce las etapas del proceso KDD y como se aplica teóricamente sobre los datos para llegar al conocimiento 3. Identifica los principales problemas de minería de datos	5, 6, 7

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
2	RA1-RA2	Data Mining y Machine Learning	3
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
2.1 Introducción al aprendizaje de máquinas 2.2 Regresiones: Lineal y Logística 2.3 Redes Neuronales a. Perceptrón b. Perceptrón multicapa 2.4 K-means 2.5 Árboles de decisión 2.6 Redes Bayesianas 2.7 K-Nearest Neighbor (KNN) 2.8 Máquinas de soporte vectorial (SVM)		El estudiante: 1. Conoce el funcionamiento de las principales herramientas provistas por el aprendizaje de máquinas (algoritmos para clasificación, asociación, regresión, etc.) para solucionar los diversos problemas en Data Mining.	5, 6, 7

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
3	RA2–RA3	MINERIA DE CONTENIDO WEB (WEB CONTENT MINIG)	4
Contenidos		Indicador de logro	Referencias a la Bibliografía
3.1. Introducción 3.2. Minería de Texto 3.3. Clasificación del texto contenido en una página Web. 3.4. Segmentación de grupos de páginas con contenido similar 3.5. WEBSOM 3.6. Sumarización automática del texto de una página Web 3.7. Extracción de componentes clave 3.8. Procesamiento de Lenguaje Natural (NLP) 3.9. Aplicaciones		El estudiante: Aprende a procesar el contenido de un sitio Web, principalmente el texto. En primer lugar, se espera que aprenda a manejar el texto contenido en un sitio de manera vectorial utilizando el enfoque de text mining clásico, así como también considerando nuevas técnicas de procesamiento de lenguaje natural para, por ejemplo, hacer minería de opiniones.	1, 2, 3, 7

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
4	RA1-RA2	MINERÍA DE ESTRUCTURA WEB (WEB STRUCTURE MINING)	1
Contenidos		Indicador de logro	Referencias a la Bibliografía
4.1. Introducción 4.2. El algoritmo HITS 4.3. El algoritmo Page Rank 4.4. Identificación de comunidades en la Web		El estudiante: Comprende la estructura de grafo de la Web y como identificar que páginas son relevantes dentro de un conjunto de páginas según su estructura de hipervínculos.	1, 3

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
5	RA1-RA2	MINERÍA DE USO WEB (WEB USAGE MINING)	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
5.1. Introducción 5.2. El proceso de reconstrucción de sesiones 5.3. Clusterización de sesiones 5.4. Clasificación del comportamiento del usuario en un sitio Web 5.5. Aplicación de reglas de asociación para descubrir patrones de navegación 5.6. Aplicación de patrones de secuencia para descubrir caminos comunes de acceso. 5.7. Una metodología para recomendaciones de navegación 5.8. Aplicaciones de WUM: Google Analytics		El estudiante: Aprende a procesar el contenido de un sitio Web, principalmente el texto. En primer lugar, se espera que aprenda a manejar el texto contenido en un sitio de manera vectorial utilizando el enfoque de text mining clásico, así como también considerando nuevas técnicas de procesamiento de lenguaje natural para, por ejemplo, hacer minería de opiniones.	1, 2, 3

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
6	RA1	ENTENDIENDO EL COMPORTAMIENTO DEL USUARIO WEB	1
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
6.1. Sistemas recomendadores en la web 6.2. Entendiendo las preferencias de los usuarios a través de Web Site Keyobjects y Eye Tracking 6.3. Aplicando neurociencia para modelar el comportamiento del usuario en la web 6.4. Web Intelligence y la Neurociencia		El estudiante: Combinando los datos y la información entregada por diversas técnicas de web mining, aprende a entender y modelar el comportamiento de los usuarios en la web, en vistas de generar aplicaciones que hagan la navegación más simple y rica. Se espera que el alumno consolide el aprendizaje adquirido y a construir y mantener sitios Web adaptativos.	2, 13

Número	RA	Nombre de la Unidad	Duración en Semanas
7	RA1	Tópicos Avanzados	1
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
7.1 Uso real de aplicaciones de Web Intelligence		El estudiante: Se entera de los usos prácticos de la aplicación de tecnologías de Web Intelligence en problemas actuales.	

Número	RA	Nombre de la Unidad	Duración en Semanas
8	RA1	DATOS PERSONALES, PRIVACIDAD Y MINERÍA WEB	1
Contenidos		**Indicador de logro**	**Referencias a la Bibliografía**
8.1 Datos personales 8.2 Protección legal de datos personales y privacidad 8.3 Desafíos		El estudiante: Se familiariza con el impacto que las técnicas de minería web tienen en la conservación de la privacidad de los usuarios y que conozca los conceptos básicos de protección de estos derechos de privacidad.	9,10

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS UNIVERSIDAD DE CHILE

Bibliografía General

1. J.D. Velásquez and L.C. Jain "Advanced Techniques in Web Intelligence part 1", Springer, 2010
2. J.D. Velásquez y L. Donoso "Tratamiento de Datos Personales en Internet", Thomson Reuters, 2013
3. J.D. Velásquez, V. Palade and L.C. Jain "Advanced Techniques in Web Intelligence part 2", Springer, 2013.
4. J.D. Velásquez and V. Palade "Adaptative Web Site". IOS Press, Netherland, 2007.
5. C.D. Manning, P. Raghavan, H. Schutze, "Introduction to Information Retrieval", Cambridge University Press 2008. http://www-csli.stanford.edu/~hinrich/information-retrieval-book.html
6. G. Myatt, "Making Sense of Data: A practical Guide to exploratory data analisis and data mining". Wiley Interscience 2007.
7. S. Chakrabarti, "Mining The Web, Discovering Knowlege From HyperText Data". MorganKaufmann Publisher 2003.
8. B. Liu, "Web Data Mining", Springer 2011.
9. A. Scime, "Web Mining: Application and techniques". IDEA Group Publishing 2005.
10. Tratamiento De Datos Personales En Internet, J Velásquez, L Donoso, Thomson, 2013
11. J.D. Velásquez. Web mining and privacy concerns: Some important legal issues to be consider before applying any data and information extraction technique in web-based environments, Expert Systems with Applications, 40(13):5228–5239, 2013
12. D. Shukla, S. B. Patel y A. K. Sen, «A literature review in health informatics using data mining techniques,» Int. J. Softw. Hardware Res. Eng. IJOURNALS, 2014.
13. Loyola, P., Martinez, G., Muñoz, K., Velásquez, J. D., Maldonado, P., & Couve, A. (2015). Combining eye tracking and pupillary dilation analysis to identify Website Key Objects. Neurocomputing, 168, 179-189.

Vigencia desde:	2017
Elaborado por:	Juan D. Velásquez, Andrés Córdova
Validado por:	
Revisado por:	Comisión de Docencia del DII