

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				
<p>Competencias específicas</p> <p>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.</p> <p>CE6: Desarrollar habilidades para liderar equipos de trabajo manejando las relaciones interpersonales.</p> <p>Competencias genéricas:</p> <p>CG1: Comunicar ideas y resultados de trabajos profesionales o de investigación, en forma escrita y oral, tanto en español como en inglés.</p> <p>CG5: Gestionar su auto-aprendizaje en el desarrollo del conocimiento de su profesión, adaptándose a los cambios del entorno.</p>				
Propósito del curso				
<p>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.</p> <p>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 incluyen 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.</p>				

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 Intelligence aplicado, 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
<p>La metodología de trabajo será activo-participativa y su carácter es teórico-práctico. Se trabajará:</p> <ul style="list-style-type: none"> - Clases expositivas con estructura de INICIO - DESARROLLO - CIERRE. - Análisis de casos presentes en literatura reciente. - Trabajo individual y grupal a partir de tópicos principales. 	<p>La evaluación considera:</p> <ul style="list-style-type: none"> - 3 Tareas individuales - 1 Proyecto Grupal

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 conceptos iniciales como E-bussiness, E-commerce y Sitio Web adaptativo. 2. Conoce las etapas del proceso KDD y como se aplica teóricamente sobre los datos para llegar al conocimiento 3. Identifica la taxonomía básica de la Minería Web y la motivación detrás del área y sus divisiones.	4, 5, 6, 7, 8, 9, 15	

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
2	RA1-RA2-RA3	Data Mining y Machine Learning	3
Contenidos		Indicador de logro	Referencias a la Bibliografía
2.2 Minería de Datos y Aprendizaje de Máquinas. 2.4 Clustering 2.5 Árboles de Decisión 2.6 Support Vector Machines 2.7 Redes Neuronales Artificiales 2.8 Framework Bayesiano 2.9 K-Nearest Neighbours 2.10 Meta-algoritmos 2.11 Métricas de Desempeño	El estudiante: 1. Conoce el concepto de Minería de Datos y sus tareas comunes 2. Conoce el concepto de Aprendizaje de Máquinas, su idea básica y sus distintos enfoques. 3. Maneja la teoría básica detrás de algunos algoritmos de Aprendizaje de Máquina conocidos. 4. Conoce el concepto de Meta algoritmos y algunas alternativas existentes 5. Maneja las métricas de desempeño básicas para un algoritmo de aprendizaje supervisado.	5, 6, 7, 15, 16	

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
3	RA2-RA3	Minería de Contenido Web	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
3.1. Minería de Texto 3.1.1. Construcción de Representación Vectorial del texto 3.1.2. Aplicaciones 3.2. Herramientas de Procesamiento de Lenguaje Natural (NLP) 3.3. Clasificación de Contenido de Página Web		El estudiante: <ul style="list-style-type: none"> - Entiende la motivación detrás de representar el texto como un vector de palabras por página. - Conoce los pasos básicos para construir la representación vectorial y la métrica de similitud entre documentos. - Conoce algunas aplicaciones de negocio para la minería de texto. - Conoce de forma básica algunas herramientas de NLP. - Maneja los distintos enfoques para realizar clasificación de contenido identificado en texto. 	1, 2, 3, 7, 8, 9

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
4	RA2-RA3	Minería de Estructura Web	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
4.1. Minando la estructura de la Web 4.2. El algoritmo PageRank 4.3. El algoritmo HITS 4.4. Identificación de comunidades en la Web		El estudiante: <ul style="list-style-type: none"> - Entiende la motivación detrás de la representación de la estructura Web como un grafo y los enfoques para explotarla. - Conoce la teoría básica de los algoritmos PageRank y HITS, como medios para poder generar jerarquías entre sitios Web, junto con sus ventajas y desventajas. - Conoce algunos enfoques para identificar comunidades en la Web. 	1, 3, 8, 9

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
5	RA2-RA3	Minería de Uso Web	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
5.1. Minando el uso de la Web 5.2. El proceso de reconstrucción de sesiones 5.3. Estado del arte en Procesamiento de datos de Weblogs		El estudiante: <ul style="list-style-type: none"> - Entiende cómo puede representar el uso de la Web a partir de los Logs o registros del servidor. - Conoce el proceso de reconstrucción de sesiones, con su objetivo, sus distintas alternativas, problemas y soluciones. - Conoce un resumen del estado del arte asociado al procesamiento de Weblogs y la identificación de patrones en estos. 	1, 2, 3, 8, 9

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
6	RA2-RA3	Minería de Opiniones Web	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
6.1. Caracterizando una Opinión 6.2. Proceso de Minería de Opiniones 6.3. Tareas que involucran Minería de Opiniones 6.4. Procedimiento alternativo 6.5. Aplicaciones 6.6. Estado del arte 6.6.1. Extracción de aspectos 6.6.2. Detección de posturas 6.6.3. Aspectos productos en		El estudiante: <ul style="list-style-type: none"> - Entiende la motivación detrás de la minería de Opiniones. - Maneja un enfoque para caracterizar una Opinión. - Conoce teoría actual en relación a distintas propuestas para abordar el problema. - Identifica distintas tareas relacionadas con la Minería de Opiniones. - Conoce posibles aplicaciones prácticas de la Minería de Opiniones. - Conoce algunas nuevas tareas presentes en el estado del arte. 	12, 13, 14

Número	RA al que tributa	Nombre de la Unidad	Duración en Semanas
7	RA1	Entendiendo el Comportamiento del Usuario Web	2
Contenidos		Indicador de logro	Referencias a la Bibliografía
7.1. Comportamiento del usuario Web 7.2. El uso de Eye Tracking para caracterizar comportamiento 7.3. Identificación de Web Site Keyobjects 7.4. Clustering de Web Site Keywords 7.5. Aplicando neurociencia para modelar el comportamiento del usuario en la Web 7.6. Medición del proceso de toma de decisión del usuario Web 7.7. Search Engine Optimization 7.8. Sistemas de Recomendación		El estudiante: <ul style="list-style-type: none"> - Conoce algunas aplicaciones experimentales presentes en el estado del arte, para caracterizar el comportamiento del usuario en la Web. - Conoce el concepto de Search Engine Optimization, su valor y sus etapas. - Conoce el concepto de Sistema de Recomendación, sus usos prácticos y los enfoques para poder desarrollarlo. 	2, 11

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 analysis and data mining” 2nd Edition. Wiley Interscience 2014.
7. S. Chakrabarti, “Mining the Web, Discovering Knowledge from HyperText Data”. Morgan Kaufmann Publisher 2003.
8. B. Liu, “Web Data Mining”, Springer 2011.
9. A. Scime, “Web Mining: Application and techniques”. IDEA Group Publishing 2005.
10. 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.
11. 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.
12. Balazs, J. A., & Velásquez J. D. (2016). Opinion mining and information fusion: a survey. *Information Fusion*, 27, 95-110.
13. Hemmatian, F. & Sohrabi, M. K. (2019). A survey on Classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, 1-51.
14. Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
15. Han, J., Pei, J., & Kamber, M. (2014). *Data mining: concepts and techniques 3rd Edition*. Elsevier.
16. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Vigencia desde:	2021
Elaborado por:	Víctor Hernández M., Juan D. Velásquez
Validado por:	
Revisado por:	Comisión de Docencia del DII