

PROGRAMA DE CURSO

INGENIERÍA DE LA INFORMACIÓN

A. Antecedentes generales del curso:

Departamento	Ingeniería Industrial (DII)				
Nombre del curso	Ingeniería de la información	Código	IN4151	Créditos	6
Nombre del curso en inglés	<i>Information Engineering</i>				
Horas semanales	Docencia	3,0	Auxiliares	1,5	Trabajo personal 5,5
Carácter del curso	Obligatorio	X		Electivo	
Requisitos	CC1002: Introducción a la programación, IN3171: Modelamiento y optimización, IN3242: Estadística				

B. Propósito del curso:

El curso tiene como propósito que los y las estudiantes seleccionen y usen modelos de analítica de negocios provenientes de la ciencia de los datos y de la información, para mejorar la efectividad operacional y eficiencia de las organizaciones de los sectores públicos y privados, considerando sus necesidades de gestión de información. Para ello, adquieren habilidades para el diseño de modelos de datos utilizando los enfoques relacionales y NoSQL. Además identifican y aplican modelos supervisados y no supervisados de la ciencia de datos respondiendo a los problemas canónicos de gestión del desempeño organizacional, utilizando metodologías para la extracción de patrones de conocimiento, y su implementación en Python. Finalmente, identifican los procesos de negocios canónicos de las organizaciones para aplicar actividades de analítica predictiva, prescriptiva y descriptiva.

El curso tributa a las siguientes competencias específicas (CE):

CE1: Identificar, analizar y diagnosticar los diferentes elementos de los problemas complejos que surgen en las organizaciones, y que son claves para resolverlos.

CE2: Concebir y diseñar soluciones que crean valor para resolver problemas de las organizaciones, utilizando los conocimientos provenientes de la gestión de operaciones, tecnologías de información y comunicaciones, finanzas, economía y marketing.

CE3: Modelar, simular y evaluar problemas de gestión, para encontrar soluciones óptimas, a necesidades de la ingeniería industrial.

CE4: Emplear y aplicar los conocimientos de las distintas disciplinas constitutivas de la ingeniería industrial: gestión de operaciones, tecnologías de información y

comunicaciones, finanzas, economía y marketing, en las respectivas áreas funcionales de las organizaciones.

El curso tributa a las siguientes competencias genéricas (CG):

CG1: Comunicación académica y profesional:

Comunicar en español de forma estratégica, clara y eficaz, tanto en modalidad oral como escrita, puntos de vista, propuestas de proyectos y resultados de investigación fundamentados, en situaciones de comunicación compleja, en ambientes sociales, académicos y profesionales.

CG2: Comunicación en inglés:

Leer y escuchar de manera comprensiva en inglés una variedad de textos e informaciones sobre temas concretos o abstractos, comunicando experiencias y opiniones, adecuándose a diferentes contextos y a las características de la audiencia.

CG3: Compromiso ético:

Actuar de manera responsable y honesta, dando cuenta en forma crítica de sus propias acciones y sus consecuencias, en el marco del respeto hacia la dignidad de las personas y el cuidado del medio social, cultural y natural.

CG6: Innovación:

Concebir ideas viables y novedosas que generen valor para resolver necesidades latentes, materializadas en productos, servicios o en mejoras a procesos dentro de un sistema u organización, considerando el contexto sociocultural y económico y los beneficios para el usuario.

C. Resultados de aprendizaje:

Competencias específicas	Resultados de aprendizaje
CE1	RA1: Identifica necesidades de gestión de información en una organización pública o privada, contrastando las condiciones de la cadena de valor para mejorar el desempeño de la organización.
CE2, CE3	RA2: Crea modelos de datos acordes a las necesidades de información identificadas utilizando enfoques relacionales y NoSQL.
CE2, CE3, CE4	RA3: Selecciona y usa modelos de analítica predictiva para mejorar el desempeño de la gestión de la organización aplicando procesos de extracción de información y conocimiento (identificación de patrones).

CE2, CE3, CE4	RA4: Identifica los procesos de negocios canónicos de las organizaciones para aplicar actividades de analítica predictiva, prescriptiva y descriptiva.
Competencias genéricas	Resultados de aprendizaje
CG1	RA6: Elabora de forma clara y precisa informes técnicos de gestión, modelación y análisis de datos e información evidenciando en su escritura el desarrollo coherente de un problema, describiendo los métodos usados y sus resultados.
CG2	RA7: Lee diversos textos en Inglés (presentaciones, artículos científicos, reportes técnicos, etc.) para adquirir e incorporar conocimientos sobre conceptos, definiciones, herramientas y aplicaciones de la Ingeniería de la información.
CG3	RA8: Trabaja en sus tareas (informes, controles, quizzes, etc.) ajustándose con honestidad y responsabilidad a la normativa, respetando la propiedad intelectual de otros y ejerciendo un rol de creación individual.
CG3, CG6	RA9: Detecta necesidades considerando la situación actual del desempeño de la organización y dilemas éticos asociados, para proponer soluciones innovadoras a partir de las herramientas de la Ingeniería de la información.

D. Unidades temáticas:

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
1	RA1, RA7	Fundamentos de la Gestión de la Información y Analítica de Negocios	2 semanas
Contenidos		Indicador de logro	
1.1. Diferencias entre información y conocimiento en base a la noción de dato. 1.2. Función de los datos, información y conocimiento en la toma de decisiones y control de gestión del desempeño. 1.3. Desafíos organizacionales de gestión de datos, información y conocimiento en la cadena de valor/procesos de negocios. 1.4. Definición de Analítica de Negocios.		El/la estudiante: 1. Compara información y conocimiento en base a la noción de datos para establecer sus granularidades, niveles de abstracción y otras diferencias. 2. Contrasta condiciones de la cadena de valor de la organización para identificar desafíos de gestión de datos, información y conocimiento. 3. Explica integralmente qué se entiende por analítica de negocios. 4. Lee en Inglés textos explicativos identificando los principales conceptos y definiciones de la ingeniería de la información.	
Bibliografía de la unidad		Hernández et al., 2004; Kenett and Shmueli 2016; Shmueli et al., 2019; Sharda et al., 2020.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
2	RA2, RA6, RA7, RA8, RA9	Gestión de Datos e Información	4 semanas
Contenidos		Indicador de logro	
2.1. Bases de datos, conceptos y definiciones. 2.2. Bases de datos relacionales – MySQL 2.3. Extracción de información con SQL. 2.4. Bases de datos No-SQL (Not Only SQL) – Mongo DB. 2.5. Nociones de seguridad de datos. 2.6. Metodologías de extracción de información y conocimiento – KDD, Crisp-DM.		El/la estudiante: 1. Modela, normaliza e implementa modelos de datos relacionales en RDS o equivalentes. 2. Diseña e implementa consultas SQL sobre una base de datos relacional. 3. Modela e implementa modelos de bases de datos No-SQL en Mongo DB o equivalente. 4. Comprende las metodologías KDD y Crisp-DM para extraer patrones desde los datos.	
Bibliografía de la unidad		Elmasri and Navathe, 2007; Velásquez, 2013; Silberschatz et al., 2014., Shmueli et al., 2019; Mongo DB, 2021.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
3	RA3, RA6, RA7, RA8	Introducción a la Ciencia de Datos e Información para la Gestión	8 semanas
Contenidos		Indicador de logro	
3.1. Conceptos y definiciones de ciencia de datos e información. 3.2. Definiciones de aprendizaje de máquinas, sobreajuste, subajuste y generalización. 3.2. Preparación de datos – calidad, transformación, limpieza, análisis exploratorio. 3.3. Métodos de extracción y selección de características. 3.4. Métricas de evaluación – exactitud, precisión, recall,		El/la estudiante: 1. Utiliza las metodologías KDD y/o Crisp-DM para la extracción de patrones desde los datos. 2. Comprende los conceptos de aprendizaje de máquinas, sobreajuste, subajuste y generalización. 3. Evalúa la calidad de los datos, realizando análisis exploratorios, los que estandariza para la aplicación de modelos de ciencia de datos. 4. Usa técnicas para extraer y seleccionar características de los datos, y aplicar y evaluar modelos supervisados.	

<p>F1-score, curva ROC, etc.</p> <p>3.5. Análisis supervisado – Naive Bayes, Árboles de clasificación, SVM, otros modelos.</p> <p>3.6. Aplicaciones: fraude financiero, detección de anomalías, churn prediction, intención de compra, riesgo de pacientes, otros.</p> <p>3.7. Nociones de análisis no supervisado – (K-means, jerárquico)</p> <p>3.8. Aplicaciones: segmentación de clientes, market basket analysis.</p>	<p>5. Utiliza y evalúa modelos no supervisados sobre los datos.</p> <p>6. Identifica y compara aplicaciones canónicas de los modelos supervisados y no supervisados para la gestión de las organizaciones.</p>
<p>Bibliografía de la unidad</p>	<p>Hernández et al., 2004; Chapman and Feit, 2015; Hastie et al., 2016; Kuhn and Johnson, 2016; Shmueli et al., 2019; Tan, 2019; Kaggle, 2021.</p>

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
4	RA4, RA6, RA7, RA8, RA9	Procesos de Negocios y Analítica	1 semana
Contenidos		Indicador de logro	
<p>4.1. Procesos de negocios – conceptos y definiciones.</p> <p>4.2. Representación de actividades de analítica en procesos de negocios canónicos.</p>		<p>El/la estudiante:</p> <ol style="list-style-type: none"> Identifica los procesos de negocios canónicos de la cadena de valor y de apoyo, asociándolos a problemas de gestión del desempeño de la organización. Asocia procesos de negocios a actividades de analítica predictiva, prescriptiva y descriptiva. 	
Bibliografía de la unidad		Apuntes del curso.	

E. Estrategias de enseñanza - aprendizaje:

El curso considera las siguientes estrategias de enseñanza-aprendizaje:

- Clases expositivas.
- Aplicación de herramientas computacionales (laboratorios).
- Análisis de casos.

F. Estrategias de evaluación:

El curso tiene distintas instancias de evaluación que consideran:

- Tareas en grupo:
 - Tarea 1: Diseñar e implementar modelos de datos relacionales y NoSQL en base a un caso de negocio (privado o sector público).
 - Tarea 2: Analizar exploratoriamente y preprocesar datos. Desarrollar modelos de análisis supervisado en base a un caso de negocio (privado o sector público).
 - Tarea 3: Desarrollar modelos de análisis supervisado en base a un caso de negocio (privado o sector público).
 - Tarea 4: Desarrollar modelos de análisis no supervisado en base a un caso de negocio (privado o sector público).
- Tres controles individuales sobre el contenido de las cátedras y lecturas en Inglés asociadas a los temas que se están trabajando.

Al inicio de cada semestre, el cuerpo académico informará sobre la cantidad y tipo de evaluaciones, así como las ponderaciones correspondientes.

G. Recursos bibliográficos:

Bibliografía obligatoria:

- (1) Elmasri, R., and Navathe, S., (2007). Fundamentos de Sistemas de Bases de Datos, Addison-Wesley.
- (2) Kuhn, M., Johnson, K. (2016). Applied Predictive Modeling, Springer.
- (3) Mongo DB. La base de datos líder para aplicaciones modernas (2021). Disponible en línea: <https://www.mongodb.com/es> (accedido el 2021/08/05)
- (4) Sharda, R., Delen, D., Turban, E. (2020). Analytics, Data Science, & Artificial Intelligence: Systems for Decision Support, Pearson, Prentice Hall, 11th Edition.
- (5) Shmueli, G., Bruce, P. C., Gedeck, P., and Patel, N. R. (2019). Data Mining for Business Analytics: Concepts, Techniques, and Applications in Python, New

Jersey, John Wiley & Sons.

- (6) Tan, P., Steinbach, M., Kumar, V. (2019). Introduction to Data Mining, Pearson Higher Education.

Bibliografía complementaria:

- (1) Chapman, C., Feit, E. M. (2015). R for Marketing Research and Analytics, Springer.
- (2) Hastie, T., Tibshirani, R., Friedman, J. (2016). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2nd Edition.
- (3) Hernández, O., Ramírez-Quintana, M. J., Ferri-Ramírez, C. (2004). Introducción a la Minería de Datos, Pearson, Prentice Hall.
- (4) Kaggle. Google online community of data scientists and machine learners (2021). Disponible en línea: www.kaggle.com (accedido el 2021/08/05)
- (5) Kenett, R. S., and Shmueli, G. (2016). Information Quality: The Potential of Data and Analytics to Generate Knowledge, John Wiley & Sons.
- (6) Silberschatz, A., Korth, Sudarshan (2014). Fundamentos de Bases de datos, McGraw-Hill, 6a Edición.
- (7) Johnson, T. (2004). Exploratory Data Mining and Data Quality, Wiley.
- (8) Velásquez, J. (2013). Web mining and privacy concerns: Some important legal issues to be considered before applying any data and information extraction technique in web-based environments. *Expert Systems with Applications* 40 (13), 5228-5239.
- (9) Weske, M. (2012). Business Process Management: Concepts, Languages, Architectures, Springer.

H. Datos generales sobre elaboración y vigencia del programa de curso:

Vigencia desde:	Otoño, 2022
Elaborado por:	Ángel Jiménez
Validado por:	Validación académico par: Sebastián Ríos COMDOC, CTD de Industrial
Revisado por:	Área de Gestión Curricular